
N

Nano (Evenescent-Wave)-Particle Image Velocimetry

► [Micro/Nano Flow Characterization Techniques](#)

Nano Manipulation

► [3D Micro/Nanomanipulation with Force Spectroscopy](#)

Nano(Evanescent-Wave)-Particle Velocimetry

Minami Yoda
G. W. Woodruff School of Mechanical Engineering,
Georgia Institute of Technology, Atlanta, GA, USA

Synonyms

[Total internal reflection \(fluorescence\) velocimetry](#)

Definition

Evanescent-wave particle velocimetry is a technique for measuring interfacial flow velocities, typically within 500 nm of the wall, by illuminating fluorescent particles, or beads, up to about 1 μm in diameter by evanescent waves, and tracking the motion of these beads over time as they are convected by the flow.

Since the illumination intensity decreases farther away from the wall, the three-dimensional position of each particle can be determined from a single image of the particles along a plane parallel to the wall from both the brightness and the location of the particle image.

Introduction

The interfacial transport of various aqueous solutions in the region less than 1 μm from the wall is of interest in many nanofluidic and microfluidic applications. The most important transport quantity in the typically steady, laminar, and incompressible flows in these applications is the velocity. The energy required to drive an aqueous solution through nanochannels and microchannels, for example, is determined by the velocity gradient at the channel wall, usually calculated by differentiating the velocity field. Measuring this velocity field requires techniques with a spatial resolution at least an order of magnitude less than the overall dimension of the nanochannel or microchannel.

The leading microfluidic velocimetry technique, micro-particle image velocimetry (μPIV), measures fluid velocity distributions by following the motion of suspended fluorescent polystyrene (PS) particles less than 3 μm in diameter [1, 2]. Because PS has a density very close to that of most aqueous solutions, these suspended particles should follow the flow. Micro-PIV can measure the two in-plane components of the velocity without disturbing the flow by illuminating the entire flow and imaging a single two-dimensional plane of the flow. And, because it can use the same

optical path to illuminate and image the flow, μ PIV requires only the addition of a single optical window to the microfluidic device.

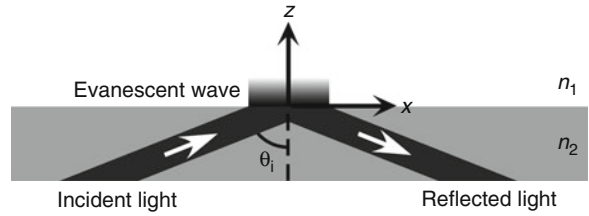
In the near-wall region, the velocity is mainly parallel to the wall because the fluid cannot penetrate the wall. Imaging the velocity field within $1\ \mu\text{m}$ of the wall in a plane parallel to the wall with μ PIV is difficult, however, because the thickness of the plane imaged by the technique is at least $2\ \mu\text{m}$, and optical interference (e.g., glare, reflections) greatly reduces the quality of the image. Instead of illuminating the entire flow, however, evanescent waves, reviewed later, can be used to illuminate only the flow within $0.5\ \mu\text{m}$ of the wall.

In general, particle velocimetry techniques determine the flow velocity field from measurements of particle velocities assuming that the particles follow the flow. The tracers used here will not follow the flow because of Brownian diffusion, which is significant for these small particles. And near the wall, Brownian fluctuations will not necessarily have a mean of zero because the diffusion varies with distance from the wall. Furthermore, the particles will not uniformly sample the velocities near the wall. The particle concentration will vary along the direction normal to the wall due to particle-wall electric double layer and van der Waals interactions. These phenomena (briefly discussed in the next section), which play a significant role for suspended particles within $1\ \mu\text{m}$ of the wall, should be considered when determining the flow velocity field from these particle velocity data.

Key Concepts

Evanescent Waves

From basic optics, it is known that when light of wavelength λ is incident upon the planar interface between two homogeneous dielectric media with refractive indices n_1 and n_2 (where $n_2 > n_1$) at an angle of incidence θ_i exceeding the critical angle $\theta_c \equiv \sin^{-1}(n_1/n_2)$, total internal reflection (TIR) occurs, where all of the incident light is reflected, remaining in the more dense medium of refractive index n_2 (Fig. 1). The Fresnel conditions for TIR, however, give a nonzero amplitude transmission coefficient, and so there is a wave transmitted into the lower refractive-index medium that cannot, on average, carry any energy across the interface [3].



Nano(Evanescent-Wave)-Particle Velocimetry, Fig. 1 Total internal reflection of light at the planar interface between two dielectric media. The reflected beam is slightly offset along the x -direction by the Goos-Hänchen shift

This transmitted *evanescent wave*, which propagates parallel to the interface along the x -axis, has a purely imaginary wave number, and hence an amplitude that decays exponentially as it propagates into the less dense medium of refractive index n_1 . The intensity of the evanescent wave is then:

$$I = I_o \exp\left\{-\frac{z}{z_p}\right\} \quad (1)$$

where I_o is the intensity of the evanescent wave at the interface, z is the distance normal to the interface measured from the interface, and z_p , the intensity-based penetration depth, is given by:

$$z_p = \frac{\lambda}{4\pi} [n_2^2 \sin^2 \theta_i - n_1^2]^{-1/2} \quad (2)$$

The intensity of the evanescent wave is maximum at the interface, and this maximum value can be several (more than 4) times the intensity of the light incident upon the interface [4]. Based on Eq. 1, the evanescent wave only illuminates the region within a few penetration depths of the wall. Evanescent waves, because they are confined to a region less than $1\ \mu\text{m}$ from the wall, have been used in total internal reflection fluorescence (TIRF) microscopy to visualize cell-substrate contact and measure the binding kinetics of proteins to cell surface receptors in biophysics [4] and measure colloidal forces in surface science [5].

EDL and Van der Waals Interactions

When a solid surface is exposed to an aqueous solution, the surface will become charged by dissociation of surface groups or adsorption of ions, for example. This charged surface will then in turn attract mobile counterions from the solution to form a charged

screening layer, known as the electric double layer (EDL), whose thickness is characterized by the Debye length scale κ^{-1} , which is less than 50 nm for all but very dilute aqueous solutions [6]. Both the glass microchannel wall and PS particle surfaces will become negatively charged for aqueous solutions at moderate pH values, and the particle-wall surfaces, which both have a net positive charge, will then have a repulsive interaction between their EDLs over distances of order κ^{-1} that varies with the ionic strength and temperature of the solution.

The particle and wall surfaces will also experience attractive intermolecular van der Waals interactions. The combination of EDL repulsion mediated by van der Waals attraction, described by the Derjaguin-Landau-Verwey-Overbeek (DLVO) theory [6], leads to a strongly nonuniform particle distribution along wall-normal, with a particle depletion zone immediately next to the wall.

Hindered Brownian Diffusion

Colloidal particles with radii $a \leq 0.5 \mu\text{m}$ suspended in an unbounded fluid are subject to significant Brownian diffusion due to Stokes drag and thermal effects. The stochastic fluctuations due to Brownian diffusion are described by the Stokes-Einstein diffusion coefficient:

$$D_{\infty} = \frac{kT}{6\pi\mu a} \quad (3)$$

where k is the Boltzmann constant, T is the absolute temperature of the fluid, and μ is the viscosity of the fluid. In the near-wall region, however, the fluctuations of the particle position due to Brownian diffusion are reduced by hydrodynamic interactions with the wall. Moreover, the presence of the wall hinders diffusion normal and parallel to the wall differently. For a particle whose center is at a wall-normal distance $z = O(a)$ from the wall, the correction factor for diffusion parallel to the wall is given by the Faxén relation [7]:

$$\beta_{\parallel} \equiv \frac{D_{\parallel}}{D_{\infty}} = 1 - \frac{9}{16} \frac{a}{z} + \frac{1}{8} \left(\frac{a}{z}\right)^3 - \frac{45}{256} \left(\frac{a}{z}\right)^4 - \frac{1}{16} \left(\frac{a}{z}\right)^5 \quad (4)$$

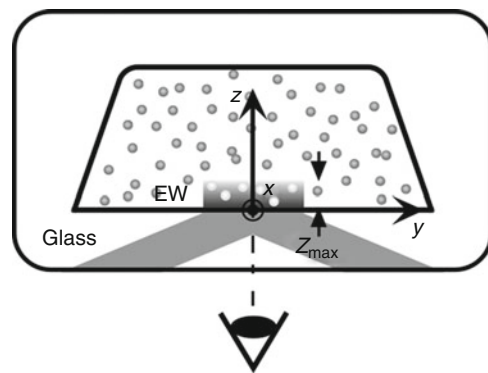
where D_{\parallel} is the coefficient for hindered diffusion parallel to the wall. The correction factor for diffusion normal to the wall, which is an infinite series in z/a [8], can be approximated by regression of the series [9] as:

$$\beta_{\perp} \equiv \frac{D_{\perp}}{D_{\infty}} = \frac{6h^2 + 2ah}{6h^2 + 9ah + 2a^2} \quad (5)$$

where D_{\perp} is the coefficient for hindered diffusion normal to the wall and $h = z - a$ is the particle edge-wall separation. Near the wall, the diffusion coefficients become functions of the particle wall-normal position. Given that this position will change over time due to diffusion normal to the wall, the near-wall fluctuations of particles due to Brownian diffusion will not have a mean of zero, unlike the fluctuations described by Eq. 3.

Evanescent-Wave Particle Velocimetry

In evanescent-wave particle velocimetry, evanescent waves generated by the TIR of a laser beam at the interface between the glass wall of the channel and the flowing aqueous solution are used to illuminate fluorescent PS particles of radius $a = 50\text{--}500 \text{ nm}$ suspended in the flow (Fig. 2). The evanescent waves can be generated either by coupling the laser beam into the channel using a prism, or by using a high numerical



Nano(Evanescent-Wave)-Particle Velocimetry, Fig. 2

Sketch of the cross-section of flow through a glass microchannel containing suspended fluorescent particles. The particles are illuminated by evanescent waves (EW) from the TIR of a laser beam at the glass wall-fluid interface, and imaged through the same (lower) wall of the channel. The flow is along the x -direction

aperture TIRF microscope objective. For a beam of blue light at $\lambda = 488$ where $n_1 = 1.34$ (assuming that the refractive index of the aqueous solution is that of water) and $n_2 = 1.46$, $\theta_c = 66.6^\circ$ and $z_p \approx 100\text{--}200$ nm by Eq. 2 for angles of incidence a few degrees greater than the critical angle. The fluorescence from the particles in the x - y plane parallel to the bottom wall is imaged from below through a microscope objective by an intensified or electron multiplying CCD (ICCD and EMCCD, respectively) camera.

The maximum z -distance from the wall Z_{\max} where tracer particles are visible in the images is usually determined by the noise characteristics of the camera. For a 16-bit EMCCD camera image, for example, the background noise can be as great as 8% of the maximum grayscale value. In this case, $Z_{\max} = 2.5z_p$ based on Eq. 1, since the images of particles beyond that distance cannot be distinguished from the background noise in the images. The displacement of the particles along the x - and y -axes in the image plane can then be determined from a pair of images separated by a time interval Δt , and the two components of the velocity of the particle in the plane are simply that displacement divided by Δt .

Nanoparticle Image Velocimetry

In nanoparticle image velocimetry (nPIV), the average velocity of a group of particles is calculated by cross-correlating the part of the first image in the pair which contains the images of these particles (the interrogation window) with the second image in the pair. The displacement of the group of particles is then taken to be the shift between the two images required to obtain the maximum cross-correlation [1]. The spatial resolution of the velocity data along x and y is determined by the corresponding dimensions of the interrogation window, while the spatial resolution of the velocity data along z is Z_{\max} .

The images of particles closer to the wall will be brighter than those farther from the wall because of the exponential decay of the illumination. So, although nPIV data represent the average velocity of the particles within a distance Z_{\max} of the wall, the average is biased toward the velocities of particles closer to the wall, because their brighter images will be more heavily weighted in the cross-correlation. This bias can be eliminated by using particle-tracking approaches where the displacements of individual

particles are determined from a pair of images by finding the (x, y) positions of each particle center in each image.

Particle Distributions

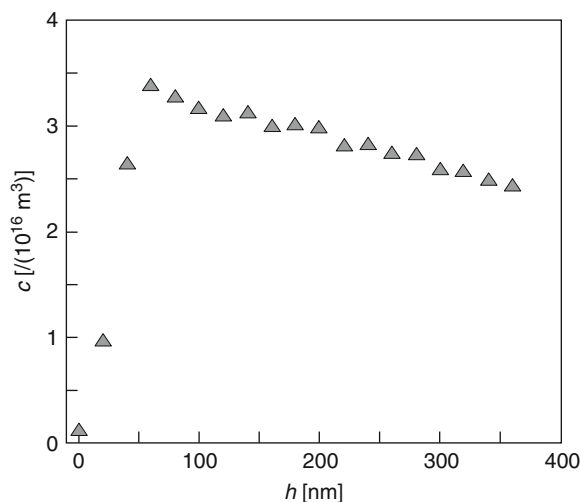
The nonuniform nature of the illumination, however, also makes it possible to determine the z -position of each particle center. If the intensity I_p of the image of a fluorescent particle excited by evanescent waves follows Eq. 1, the particle edge-wall separation is given by:

$$h = z_p \ln \left\{ \frac{I_p^0}{I_p} \right\} \quad (6)$$

where I_p^0 is the intensity of the image of a particle touching the wall (i.e., at $h = 0$), and the wall-normal position of the particle center $z = h + a$. In practice, particle-to-particle variations in I_p^0 measured in calibration experiments are quite large, with standard deviations as great as 9%, due to variations in the particle properties. For example, the radius of commercially available fluorescent PS particles typically has a standard deviation as great as 10%, based on dynamic light scattering measurements.

This close to the wall, the flow velocity is mainly parallel to the wall. The uncertainty in the z -position of the particles is also quite large. The wall-normal positions of the particles estimated using Eq. 6 are therefore used to determine the distribution of the particles (instead of the velocity component) along the z -direction. The particle edge-wall separations of a large ensemble of particles should then give a good estimate of the steady-state distribution of the particles along the wall-normal direction.

Directly measuring this distribution turns out to be important in determining the wall-normal position of near-wall velocities estimated from particle displacements. Figure 3 shows a typical particle distribution in terms of the particle number density c estimated from more than 7×10^4 images of fluorescent PS particles with a mean radius of 110 nm. The particles were suspended at a bulk number density $c_0 \approx 3 \times 10^{16} \text{ m}^{-3}$ in a 1 mM aqueous sodium tetraborate solution and convected by a weak Poiseuille flow driven by a pressure gradient of a few kPa/m. Here, I_p was taken to be the intensity averaged over the area of the particle image. As expected, the region immediately next to the



Nano(Evanescent-Wave)-Particle Velocimetry, Fig. 3 The number density profile of fluorescent PS particles measured by MnPTV with a mean radius of 110 nm

wall at $h = 0$ is depleted of particles; as h increases, the particle concentration increases from a value near zero at $h = 0$ to a maximum value at $h \approx 60$ nm. The concentration then slowly decreases to a value near c_0 . The shape of $c(h)$ is typical of repulsive EDL interactions mediated by attractive van der Waals forces.

Multilayer Nanoparticle Tracking Velocimetry

In multilayer nanoparticle tracking velocimetry (MnPTV), a single pair of particle images illuminated by evanescent waves is processed to obtain a velocity profile, typically consisting of three data points, all still at $z \leq Z_{\max}$. First, the particle displacements parallel to the wall between the first and second image of the pair are determined using particle-tracking, and the displacements normal to the wall are determined from the brightness of the particle image using Eq. 6. Each particle displacement is then assigned to its average wall-normal position over the two images in the pair. Next, the particle velocity components parallel to the wall (i.e., along x and y) are divided, based upon the average wall-normal position, into a few (usually three) layers, and averaged over all the particles in each layer. Finally, the average velocity parallel to the wall in each layer is assigned to the average distance from the wall sampled by the particles calculated from $c(h)$.

Although the nonuniform particle distribution shown in Fig. 3 only affects the position of the velocity, and not the velocity itself, the region of depleted particles next to the wall will shift the MnPTV result nearest the wall away from the wall. This shift can significantly affect the accuracy of near-wall velocity data, as will be shown in the next section. By ensemble averaging over the particles in each layer, MnPTV reduces the uncertainty in the resulting velocity and in the wall-normal position of that velocity.

As mentioned previously, the particles may not follow the flow because of hindered Brownian diffusion. The effects of Brownian diffusion have been modeled using Fokker-Planck and Langevin approaches [10, 11], and it has been shown that the discrepancy between the particle and fluid velocities due to Brownian diffusion increases with a dimensionless group $\Pi \equiv D_{\infty}(\Delta t)/Z_{\max}^2$ that compares the displacements due to Brownian diffusion with the wall-normal dimension of the region illuminated by the evanescent wave for $\Pi > 1$. The effects of Brownian diffusion can, therefore, be reduced by minimizing the time interval between the two images in the pair Δt . By reducing Z_{\max} to the thickness of a single layer, MnPTV should also have less error due to Brownian effects than other types of evanescent-wave particle velocimetry.

The particles will also lag the flow if there is a large velocity gradient, or shear, which causes the particles to rotate. The velocity of a sphere U near a solid planar wall convected by a simple shear flow with velocity of $\dot{\gamma}z$ (where $\dot{\gamma}$ is the shear rate) is already reduced by 3% if the particle edge-wall separation $h/a = 2$, and by 8% at $h/a = 0.5$ [12]. Given that there are very few particles so close to the wall due to the repulsive particle-wall EDL interactions, however, the effects of fluid shear are unlikely to have a significant effect on ensemble-averaged velocities.

Examples of Application

Over the past decade, evanescent-wave particle velocimetry has been used to study a variety of phenomena. Studies of the hindered Brownian diffusion of $a = 100$ – 500 nm fluorescent PS particles [13] have reported good agreement with theory for diffusion parallel to the wall, but poor agreement for diffusion normal to the wall. Investigations using larger

$a = 1.5 \mu\text{m}$ particles [14], however, reported good agreement with theory for diffusion both parallel and normal to the wall.

Evanescent-wave particle velocimetry has also been used to study electroosmotic flows [15] and flow inside the EDL [16], as well as slip of Newtonian liquids in Poiseuille flow. Slip lengths as great as 96 nm were observed in hydrophobic channels at a shear rate of $1,800 \text{ s}^{-1}$ [17], while MnPTV studies have reported slip lengths as great as 23 nm in hydrophobic channels at a shear rate of $2,250 \text{ s}^{-1}$ [18]. The MnPTV studies noted that accounting for the nonuniform near-wall distribution tracers was important in improving the accuracy of the estimated slip lengths, since assuming uniformly distributed tracers would have shifted the velocity estimate in the layer nearest the wall by 24 nm closer to the wall, increasing the slip length by as much as 51 nm, or more than three times the reported value.

Finally, a few studies have used other types of tracers. Near-wall flow velocities have been estimated from the displacements of both fluorescently labeled λ DNA molecules [19] and ZnS-overcoated CdSe quantum dots [20] illuminated by evanescent waves.

Cross-References

- ▶ [Micro/Nano Flow Characterization Techniques](#)
- ▶ [Nanoparticle Tracking Analysis](#)

References

1. Adrian, R.J., Westerweel, J.: Particle Image Velocimetry. Cambridge University Press, Cambridge (2011)
2. Wereley, S.T., Meinhart, C.D.: Recent advances in micro-particle image velocimetry. *Annu. Rev. Fluid Mech.* **42**, 557–576 (2010)
3. Hecht, E.: Optics, 3rd edn. Addison-Wesley, Reading (1998)
4. Axelrod, D., Burghardt, T.P., Thompson, N.L.: Total internal reflection fluorescence. *Annu. Rev. Biophys. Bioeng.* **13**, 247–268 (1984)
5. Prieve, D.C.: Measurement of colloidal forces with TIRM. *Adv. Colloid Interface Sci.* **82**, 93–125 (1999)
6. Israelachvili, J.: Intermolecular and Surface Forces, 2nd edn. Academic, London (1992)
7. Faxén, H.: Der Widerstand gegen die Bewegung einer starren Kugel in einer zähen Flüssigkeit, die zwischen zwei parallelen Ebenen Wänden eingeschlossen ist. *Ann. Phys.* **4**, 89–119 (1922)
8. Brenner, H.: The slow motion of a sphere through a viscous fluid towards a plane surface. *Chem. Eng. Sci.* **16**, 242–251 (1961)
9. Bevan, M.A., Prieve, D.C.: Hindered diffusion of colloidal particles very near to a wall: revisited. *J. Chem. Phys.* **113**, 1228–1236 (2000)
10. Sadr, R., Hohenegger, C., Li, H., Mucha, P.J., Yoda, M.: Diffusion-induced bias in near-wall velocimetry. *J. Fluid Mech.* **577**, 443–456 (2007)
11. Huang, P., Guasto, J.S., Breuer, K.S.: The effects of hindered mobility and particle depletion of particles in near-wall shear flows and the implications for nanovelocimetry. *J. Fluid Mech.* **637**, 241–265 (2009)
12. Goldman, A.J., Cox, R.G., Brenner, H.: Slow viscous motion of a sphere parallel to a plane wall. Part II. Couette flow. *Chem. Eng. Sci.* **22**, 653–660 (1967)
13. Choi, C.K., Margraves, C.H., Kihm, K.D.: Examination of near-wall hindered Brownian diffusion of nanoparticles: experimental comparison to theories by Brenner (1961) and Goldman et al. (1967). *Phys. Fluids* **19**, 103305 (2007)
14. Huang, P., Breuer, K.S.: Direct measurement of anisotropic near-wall hindered diffusion using total internal reflection velocimetry. *Phys. Rev. E* **76**, 046307 (2007)
15. Sadr, R., Yoda, M., Zheng, Z., Conlisk, A.T.: An experimental study of electro-osmotic flow in rectangular microchannels. *J. Fluid Mech.* **506**, 357–367 (2004)
16. Sadr, R., Yoda, M., Gnanaprakasam, P., Conlisk, A.T.: Velocity measurements inside the diffuse electric double layer in electroosmotic flow. *Appl. Phys. Lett.* **89**, 044103 (2006)
17. Huang, P., Guasto, J.S., Breuer, K.S.: Direct measurement of slip velocities using three-dimensional total internal reflection velocimetry. *J. Fluid Mech.* **566**, 447–464 (2006)
18. Li, H.F., Yoda, M.: An experimental study of slip considering the effects of nonuniform colloidal tracer distributions. *J. Fluid Mech.* **662**, 269–287 (2010)
19. Gai, H., Li, Y., Silber-Li, Z., Ma, Y., Lin, B.: Simultaneous measurements of the flow velocities in a microchannels by wide/evanescent field illuminations with particle/single molecules. *Lab Chip* **5**, 443–449 (2005)
20. Pouya, S., Koochesfahani, M., Snee, P., Bawendi, M., Nocera, D.: Single quantum-dot (QD) imaging of fluid flow near surfaces. *Exp. Fluids* **39**, 784–786 (2005)

Nanoassembly

- ▶ [Dielectrophoretic Nanoassembly of Nanotubes onto Nanoelectrodes](#)
- ▶ [Ligand-Directed Gold-Phage Nanosystems](#)

Nanobelts

- ▶ [Chemical Vapor Deposition \(CVD\)](#)

Nanobiosensors

- ▶ [Biosensors](#)

Nanobonding

- ▶ [Nanorobotic Spot Welding](#)

Nanocalorimeter

- ▶ [Nanocalorimetry](#)

Nanocalorimetric Sensors

- ▶ [Nanocalorimetry](#)

Nanocalorimetry

Jean-Luc Garden and Olivier Bourgeois
Institut Néel CNRS-UJF, Grenoble, France

Synonyms

[ac-Calorimetry](#); [Attocalorimetry](#); [Biocalorimetry](#); [Differential scanning calorimetry](#); [Femtocalorimetry](#); [Heat capacity](#); [Isothermal calorimetry](#); [Microcalorimetry](#); [Nanocalorimeter](#); [Nanocalorimetric sensors](#); [Nanoscale thermal analysis](#); [Nanotechnology](#); [Relaxation calorimetry](#); [Thermodynamics of small systems](#); [Thermometry](#)

Definition

Calorimetry is the part of thermodynamics which aims to measure any quantity of heat (enthalpy, specific heat, heat release) stored, released or brought into play in any state of matter, in a reaction, or in phase transitions [27].

More precisely, the terminology of “Nanocalorimetry” may cover different concepts depending on the area of science where it is used. It concerns any calorimetric method in which either the samples to be studied have a size in the range of the nanometer scale or the measured energies involved are of the order of the nanojoule or below.

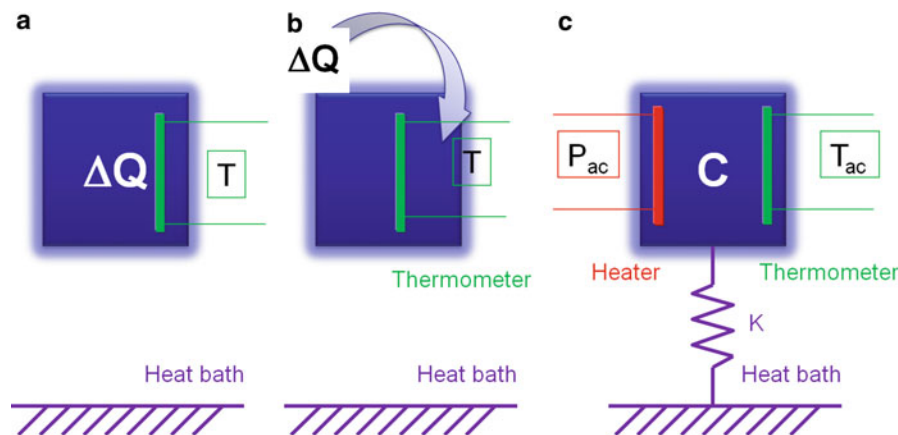
Although these two types of definitions are very different, they both have something in common. They are indeed concerned with the development of miniaturized sensors and more generally with ultrasensitive experiments imposed when small systems are involved. Indeed, most of the modern sensors are built to carry out measurements at the ultimate limit of what can be detected by calorimetry on thin films or very small volume samples.

In this entry, the key issues governing an experiment of nanocalorimetry are discussed. In particular, for each point, recent experiments conducted in this area over the world are presented. This will provide a non-exhaustive overview of what is currently done in nanocalorimetry. Particular attention will be given to micro- and nanofabrication technologies as well as highly sensitive thermal technique necessary to achieve an experiment of nanocalorimetry. Here, bolometry is not discussed; despite its thematic proximity; the measurement of radiation and possible refrigeration at the mesoscopic scale has been described extensively in other works [19].

Nanocalorimetry

Introduction

In calorimetry, thermal isolation is the major issue. The heat capacity of a sample under study, or the measurement of exchanges of energy between the sample and its environment, is properly measured only if the sample is correctly isolated from its environment (adiabatic conditions). In practice, the thermal



Nanocalorimetry, Fig. 1 Thermal schemes for calorimetric experiments (a) a quantity of heat ΔQ is released in the sample, an increase of temperature will appear. This system is assumed to be infinitely isolated from the heat bath. (b) An external heat is supplied to the sample also infinitely isolated from the heat bath.

By measuring the increase of temperature we will get access to the heat capacity of the sample (adiabatic calorimetry). (c) The sample is linked to the heat bath through a thermal conductance K ; a thermometer and a heater allow the measurement of the thermal properties (C and K)

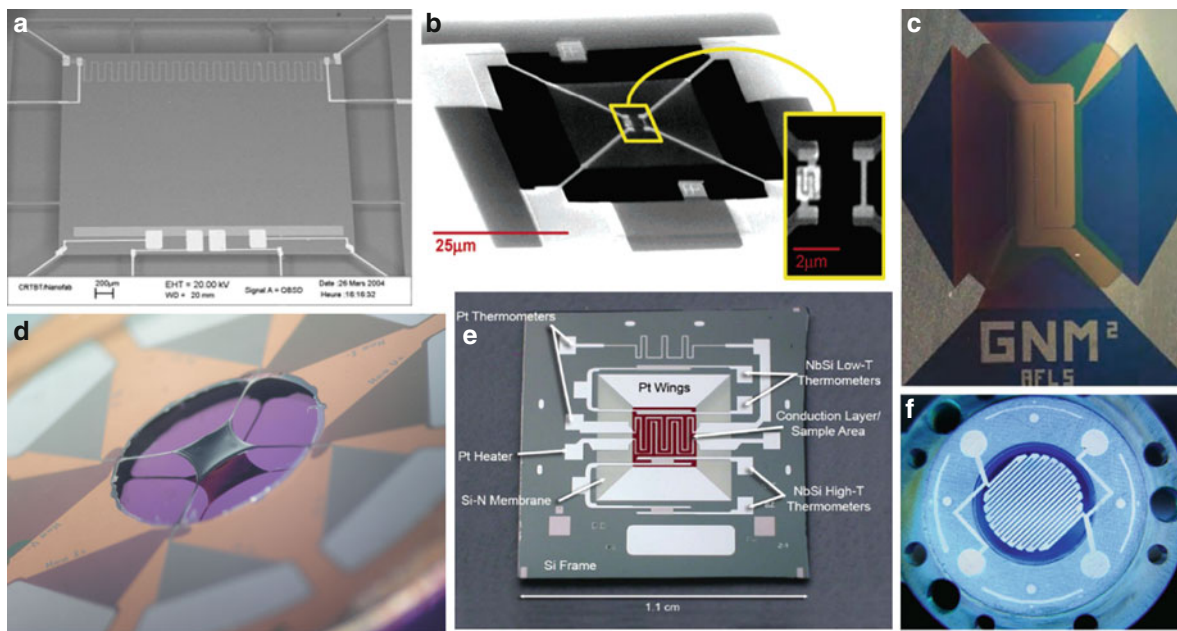
isolation is not perfect, and one overcomes this problem by calibrating correctly the thermal link between the sample and its surroundings. As discussed later in this entry, the thermal relaxation rate, τ , which determines the condition of adiabaticity, is a crucial parameter for the selection of measurement methods adapted to the physics to be studied (see the section “[Experimental Techniques](#)”). This problem of thermal isolation is more problematic in nanocalorimetry because the samples are very small in size and therefore have small masses and thus very small heat capacities. Therefore, to fulfill the requirement of adiabaticity, the experimentalist makes thin suspended membranes micrometer thick that will support small objects to be studied (see [Fig. 1](#)).

One of the advantage of working with membrane sensor is the reduction of addenda (the sample holder and the sensitive measuring elements are called “addenda”); this is especially important when working with very small systems. The small thickness of the membrane reduces also the thermal coupling to the outside providing then thermal isolation. Some membranes are structured to further limit the exchange of heat between the measurement area and the thermal bath. In this case, they look like micro-trampolines suspended by arms that allow the passage of current leads (see [Fig. 2](#)).

Another point is that, since in the case of nanocalorimetry the amount of energy to be measured

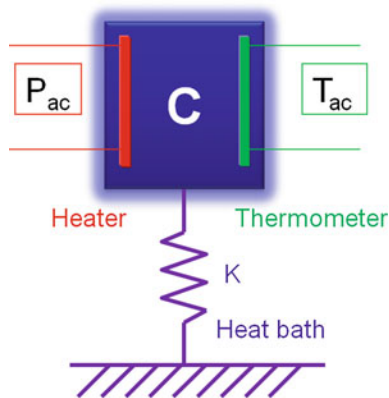
is small, the resolution of the calorimetric measurement must be sufficiently high to access the expected thermal properties. A first means to increase the signal-to-noise ratio (SNR) of the measurement is to reduce the noise of the detected signals. In other words, one must develop a low noise electronic chain adapted to the chosen experimental technique and also adapted to the sensor converting temperature changes in measurable signals (usually voltage). This aspect, which concerns the sensitive measurements and thermometry, is discussed later on (see the section “[Thermometry](#)”). A second way is to reduce significantly the heat capacity of sample holder and measuring devices (thermometers and heaters).

In this case, the use of nanotechnology and microfabrication is indispensable for the realization of sample holder for which the heat capacity is at least equivalent (or smaller) than the sample heat capacity being studied. The use of these technologies is crucial for measuring heat capacities of objects of very small size or for the detection of low energy in relevant thermodynamic transformations. Therefore, micro- and nanotechnologies are essential keys to nanocalorimetry. This justifies the use of thin microfabricated membranes for thermal isolation (see [Fig. 3](#)). Each membrane contains the sensitive elements. These elements are micro-machined in thin layers of sub-micrometer thicknesses deposited by vacuum evaporation technique or by magnetron



Nanocalorimetry, Fig. 2 Various sensors made in silicon, silicon nitride, or polymer. (a) Silicon membrane for very low temperature heat capacity measurement by ac calorimetry [3]. (b) Silicon nitride sensor made by e-beam lithography used for relaxation calorimetry [10]. (c) Silicon nitride sensor used for fast scanning calorimetry [31]. (d) Polyparaxylylene membrane

for phase transition detection in thin magnetic films around 100 K [32]. (e) Silicon nitride sensor used for relaxation calorimetry over a wide range of temperature (Cooke 2010). (f) Polyimide membrane for heat capacity measurement of thin polymer films at ambient temperature [18]



Nanocalorimetry, Fig. 3 Scheme of the principle of a nanocalorimetric measurement on membrane

sputtering. Microphotolithography techniques are used to shape the geometry of the sensitive elements and define their impedances, and therefore define the sensitivity of the calorimetric measurement.

Another major issue in calorimetry is the temperature homogeneity of the sample and the addenda. Indeed, the low thermal diffusivity of some samples

limits the dynamic of temperature variations during the measurement. The relaxation time of thermal diffusion in the sample and addenda, τ_{diff} , also depends on the geometry of nanocalorimeters. In general, a thin layer of gold, or highly diffusive material, is deposited by vacuum evaporation technique or sputtering on the measurement area so that the temperature is as uniform as possible in the sample. Finally, the thin thickness of the membranes reduces the thermal coupling to the outside providing then thermal isolation, while the isothermal layer ensures thermal homogeneity of the sensitive area. The experimental timescale Δt_{exp} , the timescale over which the thermal measurement occurs has to be slower than the diffusion time (to ensure a homogeneous temperature) and faster than the thermalization time of the sensing part to the heat bath. This last point can be mathematically summarized by the two following inequalities:

$$\tau_{\text{diff}} \ll \Delta t_{\text{exp}} \ll \tau$$

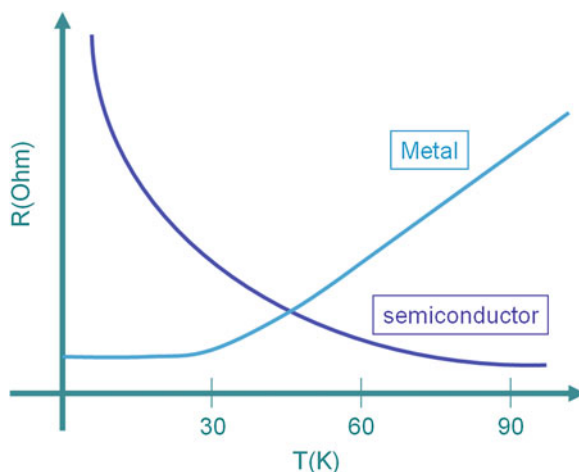
where τ_{diff} is the thermal diffusion time and τ is the thermal relaxation time.

Thermometry

The basis of the measurement of heat is essentially based on the ability of the experimentalist to accurately measure temperature. To obtain a sensitive measurement of temperature, it requires access to a physical quantity X , which varies greatly with temperature. This large variation with temperature is characterized by what is called the temperature coefficient $\frac{1}{X} \frac{dX}{dT}$. A highly sensitive measurement of temperature variation will allow a direct access to very small energy, which is the case generally when nanosystems are involved. The physical quantity X will be anything that can be accurately measured: a volume, a pressure, or an electrical quantity (capacitance, resistance, etc. . .).

Some of the most used technique to measure temperature in nanocalorimetric experiments are detailed below.

Resistive thermometry: One of the most commonly used physical quantities to measure a temperature is the resistance. The resistivity of pure metals is changing a lot with temperature from very high temperature down to 30 K. Then, these metals will be perfect thermometers in this temperature range. The most common thermometer is platinum; this metal when elaborated in thin films geometry has a positive temperature coefficient of resistance of about $\frac{1}{R} \frac{dR}{dT} = 3 \times 10^{-3} \text{K}^{-1}$; the resistance is decreasing when the temperature is decreased (Fig. 4). Other materials have also a strong change of resistance with temperature like semiconductor or like Anderson or Mott insulator. These materials have a huge increase of resistance as the temperature is lowered because the electron transport is more and more limited as the thermal activation is diminished. These materials (germanium, carbon, niobium silicon, niobium nitride) are widely used as thermometer at low temperature. They are much more efficient than metals below 40 K with temperature coefficient that can be above 1K^{-1} ! Temperatures as low as few millikelvin, close to the absolute zero, may be measured using this type of thermometry. The limitation comes from the applied current necessary to measure the resistance; the Joule heating creates parasitic power. This dissipated power may induce temperature gradients and then induces error when estimating the real temperature of a nanosystem.



Nanocalorimetry, Fig. 4 Variation of resistance versus temperature in the cases of metals and semiconductors

Noise thermometry: Any electrons in a regular resistance are subject to Brownian motion. This Brownian motion takes its origin in the temperature activation. This movement of electrons gives birth to a varying voltage across the resistor which is called the Johnson–Nyquist noise. Due to its origin this noise is a function of the temperature through $V = \sqrt{4k_B TR \Delta f}$, where k_B is the Boltzmann constant, T the temperature, R the resistance, and Δf the frequency window where this noise is measured. By measuring the voltage versus time, knowing the resistance, one can deduce the temperature of the resistor R . This thermometry is not very sensitive but it has the major advantage of not dissipating any spurious power because its measurement does not require any electrical current.

Thermocouple thermometry: A thermocouple is composed of a junction between two different materials. A temperature difference between a “hot” and a “cold” junction will be converted into a voltage (Seebeck effect). The measurement of this voltage is a widespread way of measuring a temperature. However, the amplitude in volt per Kelvin produced by a usual thermocouple is small (of the order of magnitude of several hundreds of microvolts per degree), it needs at least a few hundred couples in a system called thermopile to obtain sensitivities equivalent to those of resistive thermometers.

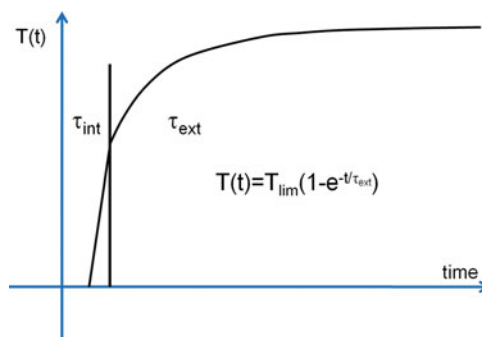
A thermopile is a device composed of plenty of thermocouples connected electrically in series and thermally in parallel. Each couple participates in the total impedance of the sensor and therefore increases the thermal noise of electrons in conductors (Johnson–Nyquist noise). The fact that it is not necessary to polarize and thus generate thermal power in the sample to obtain a measuring signal is a major advantage of this technique.

When working close to room temperature, most calorimeters and differential nanocalorimeters use thermopile to measure directly the temperature difference between the sample cell and a reference neutral cell (see DSC). In this case, the “hot” and “cold” junctions are perfectly thermally coupled to the sample and reference, respectively (or vice versa). For zero signal detection devices, two thermopiles connecting sample/thermal bath and reference/thermal bath, respectively, are mounted in opposition in order that same temperature elevations of the sample and references with respect to thermal bath gives approximately zero signal. Knowing the thermal link (thermal exchange coefficient) between the sample and the reference with respect to the surrounding, the voltage from the thermopile is proportional to the difference of heat flows (W) exchanged between the sample and the reference with respect to thermal bath, respectively. One prefers to present the converted signals collected by thermopiles in watts rather than in Kelvin because their direct integration along time gives the heat absorbed or released by the sample during the experiment of calorimetry. In conclusion, the thermopile is the ideal element to detect differential temperature or differential heat flow between two objects without direct generation of power in one or (and) the other of these objects.

Experimental Techniques

Principle of Measurement

Calorimetry is the measurement of heat exchanges between a system for which the thermal properties have to be studied and its environment (thermal bath). There are two types of calorimetric measurements. The first is the measurement of the specific heat.



Nanocalorimetry, Fig. 5 Illustration of the two significant timescales in a nanocalorimetric experiment

In this case, the experimentalist provides a given heat flux to the sample and measures the resulting temperature rise (Fig. 1c). The second concerns the measures of energy released or absorbed by a sample during any transformations or physicochemical interaction at constant temperature; one speaks in this case about isothermal calorimetry. In all cases, the experiment of calorimetry consists in measuring a change in temperature. Thermometry and the measurement electronics are therefore two essential elements for nanocalorimetry. The temperature sensor is then chosen as a function of the particular experimental methods used; the latter being in fact adapted to the physical phenomena that the experimentalist wishes to study.

Let us introduce two characteristic times which will be very useful in the rest of this entry. First the internal thermalization time noted τ_{int} is defined; this time is related to the diffusion of heat inside the sample to be studied. Secondly, the external thermalization time noted τ_{ext} is defined. This time is given by the ratio C/K where C is the heat capacity of the system and K the thermal conductance of the link to the bath (see Fig. 1). If a heat power is supplied to the sample at a rate faster than the internal thermal time, then the temperature of the sample is not homogeneous (see Fig. 5). On the other hand, if a power is supplied to a sample over a very long time, then the external thermal time will be dominant. A gradient of temperature will be established following an exponential law between the sample and the heat bath as shown in the Fig. 5.

Experimental Methods

Adiabatic Calorimetry

Regarding the specific heat, various experimental methods allow its measurement. The most traditional is the adiabatic calorimetry in which the sample and its addenda are isolated as much as possible from their environment. A thermal power of known value is sent at a timescale much smaller than the thermal relaxation time τ_{ext} . The measured temperature rise is inversely proportional to the heat capacity of the sample (plus addenda)

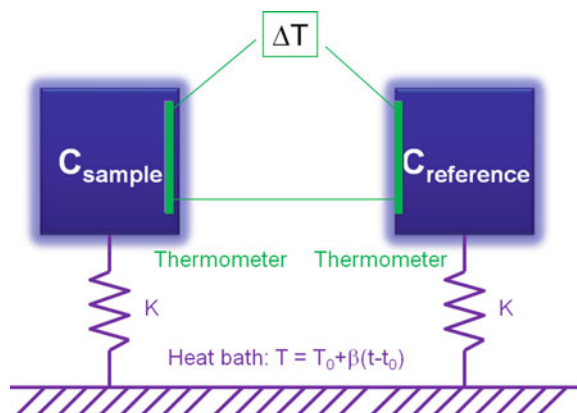
$$\Delta T = \frac{\int_0^{\Delta t} P dt}{C}$$

Then, it remains simply to change the temperature of the thermal bath and carry out a new measurement in order to have a final $C(T)$ curve. In practice, adiabaticity is not perfect, and the heat exchange coefficient has to be taken into account, and therefore energy losses between the sample and the thermal bath. This is especially true in nanocalorimetry because objects to be studied are small and therefore their heat capacity is small to. It is therefore very difficult to thermally isolate them, and the adiabatic nanocalorimetry method cannot be implemented. Scientists have developed calorimetric techniques better suited to the measurement of small samples to overcome the problem of heat loss. They are presented in the following sections.

Differential Scanning Calorimetry

A second widespread calorimetric experimental technique is called differential scanning calorimetry (DSC) [11, 22]. It is a continuous method in which the sample temperature follows a temperature ramp imposed by the experimentalist (Fig. 6). The fact that it is differential indicates that what is measured is directly the temperature difference between the sample and a neutral reference. However, one must keep in mind that any calorimetric method can be conceived in differential mode. In the case of DSC, differential heat capacity ΔC and differential heat flux ΔP (between sample and reference) measured during the ramp temperature obey the following equation:

$$\Delta C = \frac{\Delta P + \tau \frac{d(\Delta P)}{dt}}{\beta}$$



Nanocalorimetry, Fig. 6 Thermal scheme for DSC experiment. One cell contains the sample and the other cell contains a neutral reference. During the scanning temperature rate the temperature difference (or heat flux difference) is recorded as a function of time

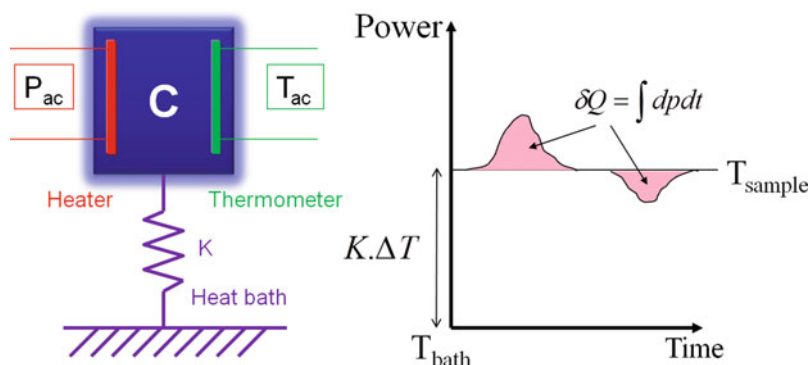
where β is the scanning temperature rate and τ the thermal relaxation time. This differential heat flux (or thermal power) is related to the differential temperature by means of the heat exchange coefficient K :

$$\Delta P = K\Delta T$$

Compared with the adiabatic calorimetry, here the heat exchange coefficient K also plays a role at the same level than the heat capacity C . The former is measured by means of a prior calibration in quasi-static mode where the above equation is valid. In DSC, the scanning rate $\beta = dT/dt$, is an important factor since it determines the signal intensity. Indeed, the first equation above shows that for a given differential heat capacity, the differential signal ΔP is for a part proportional to this scanning rate. However, in this case, the power of resolution of the measurement is reduced due to the presence of the second term revealing the dynamic of the DSC. Again, the thermal time constant plays a key role at high scanning rates. It should be noted that the above equations are only valid for a large thermal diffusivity in the sample and the sample holder, so that the time constant of heat diffusion (τ_{int}) does not become the limiting factor. They are also valid under the assumption of perfect thermal symmetry of the calorimetric head (second-order terms are not shown) [11, 22]. Classical DSC based micro- and nanotechnologies are rare. Let us mention SiN membrane DSC for heat capacity

Nanocalorimetry,

Fig. 7 Thermal scheme for isothermal calorimetry experiment. The power necessary to maintain the temperature of the sample at a constant value is represented versus time. The integration of this power provides the energy absorbed or dissipated by the sample at such temperature



detection of nl-range liquid droplets [47] and MEMS-DSC device where protein folding processes are analyzed via the measurement of differential heat capacity [44].

Fast speed DSC: In nanocalorimetry, the samples to be studied are so small that the internal thermalization time is not limiting. In this case, scientists have circumvent the problem of isolation of small samples in using scanning rates with values going from about 10^4 to 10^6 K/s. This new experimental technique, called “Fast Speed DSC,” has grown significantly in recent years. They use microsensors, typically SiN thin membranes, produced through micro- and nanotechnology [1, 33, 42]). These very high speeds yield to high sensitivities and allow measurements of very thin films from 100 to 1,000 K [15, 31]. A last feature of these rapid calorimetric measurements is that, at such high temperature rates, kinetics of studied thermal event can be observed. This is also true for high-frequency nanocalorimetric methods as seen in the following.

Isothermal Calorimetry

A third widely used measurement method is called Isothermal Calorimetry [26]. It consists of measuring absorption (or removals) of energy in a sample over time at constant temperature (Fig. 7). Titration nanocalorimetry is an experimental method issued from Isothermal Calorimetry in which a compound A in solution (titrant) interacts with compound B in solution (titrate), producing or absorbing energy at constant temperature. If an aliquot (few percent in volume) of the titrant solution containing the molecule A interacts with the solution containing the molecule B, then the experiment can be carried out N times to achieve complete disappearance of B. In this way,

through this type of experience, not only the enthalpy of reaction between molecules A and B can be obtained (during the firsts interactions), but also, thanks to the saturation curve $\Delta H = f(N)$, the constant of reaction between A and B is obtainable, like for acid-base chemical titrations. The constant of reaction allows access to the Gibbs free energy at a certain temperature. With calorimetric titration, the thermodynamic parameters ΔH , ΔG , and ΔS could be accessed for the reaction A/B which provide complete thermodynamic data at determined temperatures. This technique is mainly used in chemistry, biochemistry, and biology [37]. Since in this field, firstly biochemical and biophysical reactions generally have relatively low energy values, and secondly available volumes of biological samples are generally reduced (simply because they are expensive due to the cost of the synthesis) the use of nanocalorimetry is a natural choice. In the past decades, the development of isothermal titration nanocalorimeters has been increasing a lot [20, 45]. Another approach to induce interaction of A with B is to mix the two solutions in the sensitive area via two differentiated liquid inputs. The mixture produces or absorbs energy in the sensitive area which is detected by the sensor and then flows through a third flow output path. One speaks about flow nanocalorimeters [24, 28, 30, 35, 49]. An original way of binding has been obtained by means of electrostatic mixing [41].

A major technical difficulty in the development of isothermal titration or flow nanocalorimeters is the need for a coupling of microfluidic techniques with that of the microfabrication of nanosensors. Indeed, one needs to bring the various liquid reactants on the measurement area while maintaining the highest thermal insulation of the same area. Certain isothermal nanocalorimeter measures only the power or heat

released or absorbed by a small biological object already positioned (e.g., living cells) along time [8, 23, 43]. In the field of biocalorimetry, volumes of analyzed samples comprise between few nanoliter and few microliter, and minimum detectable powers are between few nanoWatt and few hundredth of nanoWatt depending on the nanocalorimeters designs. Integration of power measured versus time yields to minimum recorded energies from few tenth of nanoJoule to few microJoule (see tables showing performances of various isothermal calorimeters in Refs. [6] and [28]).

Low Temperature Calorimetry

Traditionally, at very low temperatures, the most used calorimetric method is the adiabatic calorimetry. However, as the adiabatic nanocalorimetry is difficult to carry out at low dimension, scientists have developed new techniques known as dynamic calorimetry techniques, which overcome the problem of heat loss. There are two main different methods: the relaxation calorimetry and ac-calorimetry [5]. Relaxation calorimetry overcomes poor insulation in directly measuring the exponential decay of the temperature of the sample compared to the thermal bath after a pulse of power. Measuring the thermal conductance K is necessary to have the heat capacity. On contrary, ac-calorimetry depends only on the frequency of temperature oscillations to be placed in adiabatic condition or quasi-adiabatic conditions. The sample temperature oscillates at a frequency such that no heat loss holds over one period of oscillation. In this case, it is not necessary to measure K . These different experimental techniques suitable for nanocalorimetry are described below.

Relaxation calorimetry: Relaxation calorimetry is to apply a heat pulse on the sensitive area of a sensor. An increase in the temperature of the sample appears; this increase will be faster than heat leak because the internal diffusion time is very short. The technique involves recording the decrease of the temperature sensor over time:

$$T = T_{bath} + \frac{P_0}{K} \exp\left(-\frac{t}{\tau}\right)$$

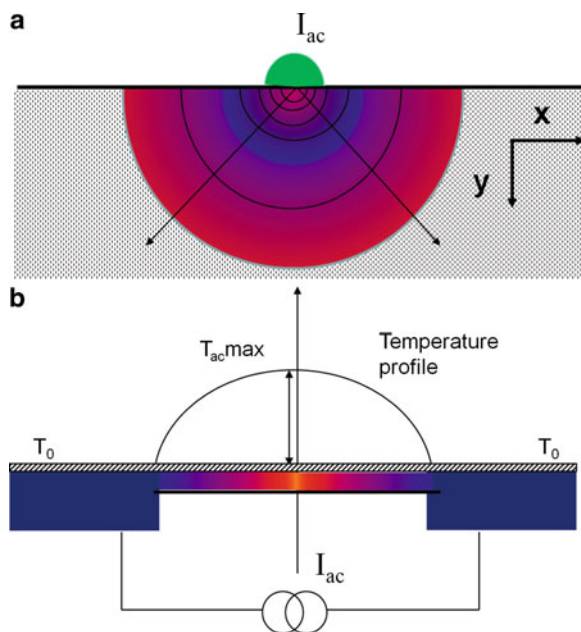
This decay is exponential and depends on the ratio between the heat capacity of the sensor with the sample

and the thermal conductance to the bath. The thermal conductance can be estimated by other means, hence from the exponential decay the heat capacity can be deduced. This measurement is usually performed on a membrane where the thermometer and heater are lithographed. It is a sensitive method which may cover a large temperature range [12, 50]. One weak point of this technique is that the measure is not based on an oscillating method, and therefore the SNR is degraded by the presence of thermal drift. This technique has been applied in some recent cases of specific heat measurement on very small sensors at low temperature [10, 39].

AC-calorimetry: In ac calorimetry, an input thermal power $P(t) = P_{dc} + P_{ac}$ constituted by a dc and a ac term is supplied to a system connected to a thermal bath by means of a thermal conductance of known value [13, 25, 40, 51]. In focusing on the ac term only, the corresponding oscillating temperature T_{ac} is measured. At low frequencies the measurement is not adiabatic or quasi-adiabatic, and at high frequencies the temperature of the sample is not homogeneous anymore. This means that the thermal relaxation time and the thermal diffusion time have to be carefully taken into account. However, in choosing the appropriate frequency range, it is possible to circumvent the problem of thermal insulation of small objects to be measured by nanocalorimetry. It remains to design a calorimetric device with an appropriate working frequency range. The frequency of the temperature oscillation has to be faster than the relaxation time to the heat bath τ_{ext} (adiabatic condition) and slower than the diffusion time in the sensor τ_{int} . Under these circumstances, the heat capacity of the sample is simply obtained by the ratio of the ac power on the ac temperature:

$$C = \frac{P_{ac}}{i\omega T_{ac}}$$

The measured heat capacity is actually a complex number with real and imaginary components of different physical meaning. These are obtained by means of the measure of either the amplitude or the phase of the oscillating temperature. This method is particularly sensitive because locking-amplifiers with narrow bandwidth filter can be used for oscillating temperature recording but also because dc or low frequency thermal



Nanocalorimetry, Fig. 8 Schemes of the 3ω method with the thermal wave in the plane (perpendicular to the transducer) (a) or along the sample (parallel to the transducer) (b)

drifts are less important. Heat capacity resolutions $\Delta C/C$ of the order of 10^{-5} to 10^{-4} are currently obtained while in relaxation calorimetry or other methods resolutions lay between 10^{-3} and 10^{-2} . Thus, ac-calorimetry is a technique particularly adapted to nanocalorimetry because, on one side it allows measures on small objects, and on another side it allows measurement in the nanoJoule/Kelvin or nanoJoule/picoJoule ranges [16, 32]. Although ac calorimetry has been developed in the field of low temperature physics [34]), for few decades it has been adapted to different temperature ranges in several domains of research (magnetic films, biological objects, polymers, etc. . .) where the detection of fine phase transitions or phase transformations were researched [17, 18, 46]. More precisely, the record in sensitivity was obtained in attojoule calorimetry (10^{-18} J) which allowed access to complete new physics [3]. The principal inconvenient of ac-calorimetry is the useable frequency range limited to no more than one or two decades when measurement of the dynamic of the studied system via $C(\omega)$ is wanted. This disadvantage is avoided in another dynamic calorimetric method called 3ω method (Fig. 8).

Other Methods

3ω -method. In the 3ω method, the thermal power that generates temperature oscillations in the sample is provided by the thermometer itself. A current of frequency f passes through the thermometer-like heater resistance (called the transducer), which produces a thermal power oscillating at frequency $2f$ due to Joule effect. As the transducer temperature varies at that frequency, this results in a $3f$ term which contains all the thermal information of the sample [2, 14]:

$$\delta T_{ac} = \frac{P_0}{K\sqrt{1 + (\omega\tau)^2}} \text{ or } \delta T_{ac} = \frac{P_0}{\sqrt{2\omega CK}}$$

The two above equations apply depending on the geometrical design of the sensor. ω is the pulsation, related to the frequency through $\omega = 2\pi f$. At very low frequency, the amplitude of the temperature oscillation is only related to the thermal conductance. If the measurement is made as a function of the frequency, the heat capacity can be extracted through the estimation of the thermalization time $\tau = \frac{c}{K}$.

The purely electrical term oscillating at frequency f , which is much larger than the $3f$ component, has generally to be removed using a Wheatstone bridge circuit, otherwise the quality of the measurement will suffer. This method, non-adiabatic by principle, was used to measure the thermal properties of material of very small thermal diffusivity for which a frequency dependence of the specific heat was expected. It has, for example, been successfully used to study the frequency-dependent specific heat at the glass transition of some glass formers over several decades of frequency [2]. Depending on the geometry of the couple transducer with sample, the thermal effusivity $c \times k$ (product of specific heat by thermal conductivity) is measured or, upon particular geometrical conditions, the thermal conductivity k is directly measured at low frequency (see equations above) [52]. In nanocalorimetry, this method has been successfully applied to the measurement of the thermal conductivity of crystalline silicon wires of nanometric sizes for which phonon-blocking effects were observed at very low temperatures [4, 21]. Recent experiments use the 3ω method for thermal properties measurement of liquid-like samples [9, 36].

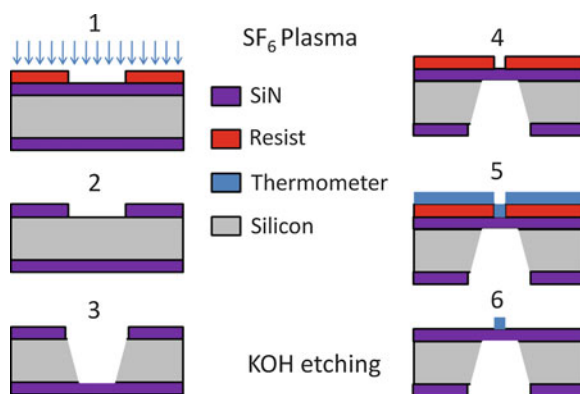
Nanocalorimetry, Table 1 Thermal parameters for various materials from high temperature to very low temperature. The thermal conductivity is given in W/cm K and the specific heat in J/g K

	Si		SiN		Pure copper		Polymer (PTFE)		Glass	
	c	k	c	k	c	k	c	k	c	k
1,000 K	8.2	0.3	4	0.2	0.5	4	NA	NA	3	2.10^{-2}
300 K	6.7	1.5	2	0.1	0.3	4	0.9	$3 \cdot 10^{-3}$	1	10^{-2}
100 K	2.5	8	1	4.10^{-2}	0.25	5	0.3	$2 \cdot 10^{-3}$	0.6	8.10^{-3}
10 K	2.10^{-3}	0.3	10^{-3}	6.10^{-3}	8.10^{-4}	160	0.18	10^{-3}	6.10^{-3}	10^{-3}
1 K	5.10^{-7}	$\sim 10^{-3}$	$\sim 10^{-6}$	5.10^{-4}	$\sim 10^{-5}$	40	$\sim 10^{-3}$	7.10^{-5}	7.10^{-6}	5.10^{-4}
0.1 K	$\sim 10^{-11}$	NA	$\sim 10^{-9}$	NA	$\sim 10^{-6}$	0.4	NA	NA	8.10^{-8}	10^{-5}

Unconventional methods: measurement of caloric curve. In this technique developed for the in-flight measurement of clusters (hundred of atom), one uses the photofragmentation of the clusters by a laser beam to estimate the internal energy U as a function of temperature. The clusters ions are selected using a mass spectrometer and thermalized in helium gas. Then they are irradiated by a laser, the fragments produced are analyzed by a second mass spectrometer. Their number depends directly on the inner energy of each clusters, hence its measurement gives the internal energy of the cluster. Its derivative ($C = \frac{\partial U}{\partial T}$) will give the heat capacity of the clusters ions (see Refs. [7, 38]). Melting point and heat capacity of cluster of 140 atoms or less could be measured using this technique, a performance never equaled with other techniques. Others unconventional methods of nano- and micro-calorimetry exist but are not described in this brief entry.

Micro- and Nanosensors

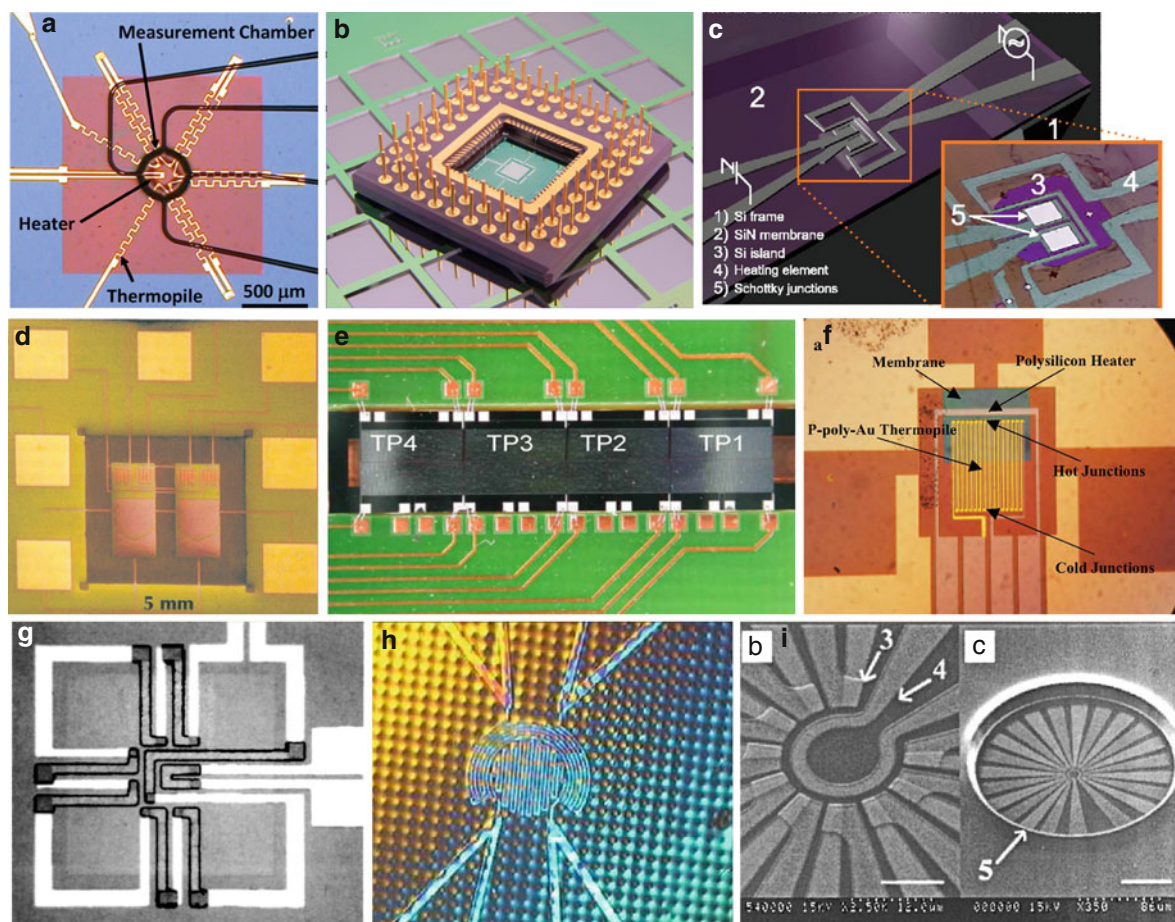
It is the need to do thermodynamic measurements on specific physical or chemical nanosystem that will determine what type of sensor that must be developed for nanocalorimetry. Secondly, the choice of the experimental technique (as detailed in the previous section) will also influence the conception of the sensing part. Several examples of nanocalorimeter based on micro- or nanofabrication are shown below where the best results have been obtained in terms of sensitivity and/or resolution. Great care must be taken in the choice of materials along with the geometry of the system (thermometry and heater); it will fully determine the thermal performance of the sensor. Two main parameters should orient the choice of materials for the sample



Nanocalorimetry, Fig. 9 Example of microfabrication procedure of a suspended membrane of SiN. 1 Spinning of resist and photolithography, 2 remove the resist, 3 chemical etching of silicon with KOH, 4 photolithography of the transducer, 5 deposition of the transducer, 6 lift-off

holder: small heat capacity and a complete absence of phase change in the working temperature area to avoid spurious signal coming from the holder. The geometry will be chosen to reduce the leak to the heat bath or to control it through a dedicated thermal link. Thermal data of the main materials used for manufacturing sensors are summarized in Table 1. This table will be fruitfully used during the design of a new sensor depending of its future working temperature range or its particular specifications (materials compatibility).

The first attempt to measure the heat capacity of thin films was done by G.D. Zally et al. [48]. He built a calorimeter based on a very thin pyrex membrane where thermometer and heater were attached. Since this attempt, numerous different sensors have been built in various materials from silicon, silicon nitride, glass to polymer. Some of them are detailed below.



Nanocalorimetry, Fig. 10 Various sensors for isothermal or biological liquid-sample nanocalorimetric experiments. (a) Parylene membrane for isothermal chemical and biological interactions [28], (b) silicon sensor for isothermal chemical and biological interactions (Photograph courtesy www.sensor.nl, [42]), (c) silicon nitride sensor for small-volume-liquid calorimetric detection [20], (d) polyimide membrane for electrostatic-liquid-mixing calorimetric detection [41], (e) silicon

microfluidic chamber for flow-through isothermal calorimetry [29], (f) glass microfluidic reaction chamber for flow-through calorimetry [49], (g) silicon nitride membrane for small-volume-liquid calorimetric detection [8], (h) micro-posts sustained polyimide membrane for small-volume-liquid calorimetric detection [18], (i) silicon nitride suspended sensor for isothermal detection of living cells [23]

Silicon, Silicon Nitride, or Diamond Membranes

As mentioned above, the choice of the materials is given by the specification of the experiment to be performed. The best material for application in a wide temperature range (from 40 K to 1,000 K) is, for sure, silicon nitride (SiN). This amorphous material is totally inert in that temperature window and can be easily manipulated. Fabrication of membrane out of SiN is mastered in all cleanrooms in the world using chemical etching of silicon by KOH (see the sketch in Fig. 9). Thermometer and heater can be lithographed on top, which permits the measurement of thermal

properties. The design of the heat link and the choice of materials for the transducers will set the performance of the calorimeter.

If one wants to apply nanocalorimetry technique at low or very low temperature, SiN is not the best choice. As it can be seen in the Table 1, the heat capacity is high as compared to single crystal silicon. Then the fabrication of silicon membrane is more appropriate and should be preferred. Moreover, the structuring of the membrane will be absolutely necessary to create a well-defined isotherm. Indeed, the thermal conductance of silicon (or SiN) is still high below 10 K, and

then a proper thermal isolation lies in suspending a membrane through isolated arms (see Fig. 2a or b). In Fig. 10, examples (c), (g), and (i) are silicon membrane for isothermal flow-through or small-volume-liquid nanocalorimetric measurements.

Polymer Membranes

In nanocalorimetry, thin polymer membranes are also used to insure good thermal insulation of the measurement area. This type of membrane is mostly used for medium to high temperatures because of the low thermal conductivity of polymers (see Table 1). At very low temperatures, they are less attractive because their heat capacity is very high (see Table 1). To ensure uniformity of temperature in these sensors, the low thermal diffusivity of these materials is generally eliminated through the deposition of a thin layer of metal diffusing heat over the entire surface of the sensitive area (e.g., gold, aluminum, silver). Figures 2d, f and 10a, d, h are examples of polymer membranes for experiments of ac-nanocalorimetry or isothermal nanocalorimetry on thin magnetic films, thin polymer films, and low volumes chemical or biological objects or reactions. The thickness of these membranes can range from a few hundreds of nanometers to several tens of micrometers. The use of micro- and nanotechnologies is obviously necessary to deposit and lithographically machined sensing elements on the membrane, as well as a possible step of membrane design through dry etching techniques (e.g., reactive ion etching or oxygen plasma).

Conclusion

In that not complete review, various experimental methods named “nanocalorimetry” have been presented. It has been shown that it concerns all calorimetric method for which the dimensions of the objects studied are in the order of magnitude of nanometer, or measured energy or power are in the range of nanojoule or nanowatt. The review has been based on numerous examples of nanocalorimeters existing in the literature from low temperature to room temperature with application in condensed matter, chemistry, biophysics, or biology in general. The aspect of micro- and nanofabrication, essential for the achievement of nanocalorimetric devices, has been emphasized as well as the importance of sensitive instrumentation,

electronic conditioning, and the choice of materials for the thermal detectors.

Calorimetry is by essence a universal method of measurement, covering by this way a wide spectrum of different researches. The emergence of nanocalorimetry in recent decades has allowed the exploration of new and original fields of research. With the evolution of modern micro- and nanofabrication technologies, many new developments will appear especially toward more sensitive sensor especially in biophysics. Significant room for improvement still exists in many areas of nanoscience, which makes nanocalorimetry quite an open subject of research.

Cross-References

- ▶ [Biosensors](#)
- ▶ [Nanotechnology](#)
- ▶ [Thermal Actuators](#)
- ▶ [Thermal Conductivity and Phonon Transport](#)
- ▶ [Thermoelectric Heat Convertors](#)

References

1. Anahory, Y., Guihard, M., Smeets, D., Karmouch, R., Schiettekatte, F., Vasseur, P., Desjardins, P., Hu, L.A., Allen, L.H., Leon-Gutierrez, E., Rodriguez-Viejo, J.: Fabrication, characterization and modeling of single-crystal thin film calorimeter sensors. *Thermochim. Acta* **510**, 126–136 (2010)
2. Birge, N.O., Dixon, P.K., Menon, N.: Specific heat spectroscopy: Origins, status and applications of the 3 omega method. *Thermochim. Acta* **304**(305), 51–66 (1997)
3. Bourgeois, O., Skipetrov, S., Ong, F., Chaussy, J.: Attojoule calorimetry of mesoscopic superconducting loops. *Phys. Rev. Lett.* **94**, 057007 (2005)
4. Bourgeois, O., Fournier, Th, Chaussy, J.: Measurement of the thermal conductance of silicon nanowires at low temperature. *J. Appl. Phys.* **101**, 016104 (2007)
5. Bourgeois, O.: Thermal nanosystems and nanomaterials. Book series: Topics in applied physics, pp. 537–567. Springer, Berlin (2009)
6. Braissant, O., Wirz, D., Göpfert, B., Daniels, A.U.: Biomedical use of isothermal microcalorimeters. *Sensors* **10**, 9369–9383 (2010)
7. Breaux, G.A., Benirschke, R.C., Sugai, T., Kinnear, B.S., Jarrold, M.F.: Hot and solid gallium clusters: Too small to melt. *Phys. Rev. Lett.* **91**, 215508 (2003)
8. Chancellor, E.B., Wiksw, J.P., Baudenbacher, F., Radparvar, M., Osterman, D.: Heat conduction calorimeter for massively parallel high throughput measurements with

- picoliter sample volumes. *Appl. Phys. Lett.* **85**, 2408–2410 (2004)
9. Choi, S.R., Kim, D.: Real-time thermal characterization of 12 nl fluid samples in a microchannel. *Rev. Sci. Instrum.* **79**, 064901 (2008)
 10. Chung Fon, W., Schwab, K.C., Worlock, J.M., Roukes, M.L.: Nanoscale, phonon-coupled calorimetry with sub-attojoule/Kelvin resolution. *Nano Lett.* **5**, 1968 (2005)
 11. Claudy, P.: *Analyse Calorimétrique Différentielle*. Lavoisier, Paris (2005)
 12. Cooke, D.W., Michel, K.J., Hellman, F.: Thermodynamic measurements of submilligram bulk samples using a membrane-based “calorimeter on a chip”. *Rev. Sci. Instrum.* **79**, 053902 (2008)
 13. Corbino, O.M.: Thermal oscillations in lamps of thin fibers with alternating current flowing through them and the resulting effect on the rectifier as a result of the presence of even-numbered harmonics. *Phys. Z.* **11**, 413–417 (1910)
 14. Corbino, O.M.: Periodic resistance changes of fine metal threads, which are brought together by alternating streams as well as deduction of their thermo characteristics at high temperatures. *Phys. Z.* **12**, 292–295 (1911)
 15. Efremov, M.Y., Olson, E.A., Zhang, M., Schiettekatte, F., Zhang, Z.S., Allen, L.H.: Ultrasensitive, fast, thin-film differential scanning calorimeter. *Rev. Sci. Instrum.* **75**, 179–191 (2004)
 16. Fominaya, F., Fournier, T., Gandit, P., Chaussy, J.: Nanocalorimeter for high resolution measurements of low temperature heat capacities of thin films and single crystals. *Rev. Sci. Instrum.* **68**, 4191–4195 (1997)
 17. Garden, J.-L., Château, E., Chaussy, J.: Highly sensitive ac nanocalorimeter for microliter-scale liquids or biological samples. *Appl. Phys. Lett.* **84**, 3597–3599 (2004)
 18. Garden, J.-L., Guillou, H., Lopeandia, A.F., Richard, J., Heron, J.-S., Souche, G.M., Ong, F.R., Vianay, B., Bourgeois, O.: Thermodynamics of small systems by nanocalorimetry: From physical to biological nano-objects. *Thermochim. Acta* **492**, 16–28 (2009)
 19. Giazotto, F., Heikkilä, T.T., Luukanen, A., Savin, A.M., Pekola, J.P.: Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications. *Rev. Mod. Phys.* **78**, 217 (2006)
 20. Hakala, T.K., Toppari, J.J., Torma, P.: A hybrid method for calorimetry with subnanoliter samples using Schottky junctions. *J. Appl. Phys.* **101**, 034512 (2007)
 21. Heron, J.-S., Bera, C., Fournier, T., Mingo, N., Bourgeois, O.: Blocking phonons via nanoscale geometrical design. *Phys. Rev. B* **82**, 155458 (2010)
 22. Höhne, G.W.H., Hemminger, W.F., Flammersheim, H.-J.: *Differential scanning calorimetry*, 2nd edn. Springer, Berlin/Heidelberg (2010)
 23. Johannessen, E.A., Weaver, J.M.R., Bourova, L., Svoboda, P., Cobbold, P.H., Cooper, J.M.: Micromachined nanocalorimetric sensor for ultra-low-volume cell-based assays. *Anal. Chem.* **74**, 2190–2197 (2002)
 24. Köhler, J.M., Zieren, M.: Micro flow calorimeter for thermoelectrical detection of heat of reaction in small volumes. *Fresenius J. Anal. Chem.* **358**, 683–686 (1997)
 25. Kraftmakher, Y.: Modulation calorimetry and related techniques. *Phys. Rep.* **356**, 1–117 (2002)
 26. Ladbury, J.E., Chowdhry, B.Z. (eds.): *Biocalorimetry, applications of calorimetry in the biological sciences*. Wiley, Chichester (1998)
 27. Lavoisier, A., de Laplace, P.S.: *Mémoire sur la chaleur*, p. 355. Académie des Sciences, Paris (1780)
 28. Lee, W., Fon, W., Axelrod, B.W., Roukes, M.L.: High-sensitivity microfluidic calorimeters for biological and chemical applications. *Proc. Nat. Acad. Sci.* **106**, 15225–15230 (2009)
 29. Lerchner, J., Wolf, A., Wolf, G., Baier, V., Kessler, E., Nietzsche, M., Krügel, M.: A new micro-fluid chip calorimeter for biochemical applications. *Thermochim. Acta* **445**, 144–150 (2006)
 30. Lerchner, J., Wolf, A., Schneider, H.-J., Mertens, F., Kessler, E., Baier, V., Funfak, A., Nietzsche, M., Krügel, M.: Nano-calorimetry of small-sized biological samples. *Thermochim. Acta* **477**, 48–53 (2008)
 31. Lopeandia, A.F., Valenzuela, J., Rodriguez-Viejo, J.: Power compensated thin film calorimetry at fast heating rates. *Sensor. Actuat. A* **143**, 256–264 (2008)
 32. Lopeandia, A.F., Andre, E., Garden, J.-L., Givord, D., Bourgeois, O.: Highly sensitive parylene membrane-based ac-calorimeter for small mass magnetic samples. *Rev. Sci. Instrum.* **81**, 053901 (2010)
 33. Minakov, A.A., Adamovsky, S.A., Schick, C.: Non-adiabatic thin-film (chip) nanocalorimetry. *Thermochim. Acta* **432**, 177–185 (2005)
 34. Minakov, A.A., Roy, S.B., Bugoslavsky, Y.V., Cohen, L.F.: Thin-film alternating current nanocalorimeter for low temperatures and high magnetic fields. *Rev. Sci. Instrum.* **76**, 043906 (2005)
 35. Nam, S.K., Kim, J.K., Cho, S.C., Lee, S.K.: Design and characterization of a high resolution microfluidic heat flux sensor with thermal modulation. *Sensors* **10**, 6594–6611 (2010)
 36. Park, B.K., Park, J., Kim, D.: Three-omega method to measure thermal properties of subnanoliter liquid samples. *Rev. Sci. Instrum.* **81**, 066104 (2010)
 37. Russell, D.J., Hansen, L.D.: Calorimeters for biotechnology. *Thermochim. Acta* **445**, 151–159 (2006)
 38. Schmidt, M., Kusche, R., Kronmüller, W., von Issendorf, B., Haberland, H.: Experimental determination of the melting point and heat capacity for a free cluster of 139 sodium atoms. *Phys. Rev. Lett.* **79**, 99 (1997)
 39. Suh, K.S., Kim, J.W., Kim, H.J., Park, Y.D., Kim, K.H.: Enhanced accuracy in a silicon-nitride-membrane-based microcalorimeter with variation of lateral layout. *Thermochim. Acta* **490**, 1–7 (2009)
 40. Sullivan, P.F., Seidel, G.: steady-state ac-temperature calorimetry. *Phys. Rev.* **173**, 679–685 (1968)
 41. Torres, F.E., De Bruyker, D., Bell, A.G., Wolkin, M.V., Peeters, E., Williamson, J.R., Anderson, G.B., Schmitz, G.P., Recht, M.I., Schweizer, S., Scott, L.G., Ho, J.H., Elrod, S.A., Schultz, P.G., Lemer, R.A., Bruce, R.H.: Enthalpy arrays. *Proc. Nat. Acad. Sci.* **101**, 9517–9522 (2004)
 42. van Herwaarden, A.W.: Overview of calorimeter chips for various applications. *Thermochim. Acta* **432**, 192–201 (2005)
 43. Verhaegen, K., Baert, K., Simaels, J., Van Driessche, W.: A high-throughput silicon microphysiometer. *Sensor. Actuat.* **82**, 186–190 (2000)

44. Wang, L., Wang, B., Lin, Q.: Demonstration of MEMS-based differential scanning calorimetry for determining thermodynamic properties of biomolecules. *Sensor. Actuat. B* **134**, 953–958 (2008)
45. Xu, J., Reiserer, R., Tellinghuisen, J., Wikswo, J.P., Baudenbacher, F.J.: A microfabricated nanocalorimeter: Design, characterization, and chemical calibration. *Anal. Chem.* **80**, 2728–2733 (2008)
46. Yao, H., Ema, K., Fukada, H., Takahashi, K., Hatta, I.: ac nanocalorimeter for measuring heat capacity of biological macromolecules in solution. *Rev. Sci. Instrum.* **74**, 4164–4168 (2003)
47. Youssef, S., Podlecki, J., Al Asmar, R., Sorli, B., Cyril, O., Foucaran, A.: MEMS scanning calorimeter with serpentine-shaped platinum resistors for characterizations of microsamples. *J. Microelectromech. Syst.* **18**, 414–423 (2009)
48. Zally, G.D., Mochel, J.M.: Fluctuation heat capacity in superconducting thin films of amorphous BiSb. *Phys. Rev. Lett.* **27**, 1710 (1971)
49. Zhang, Y.Y., Tadigadapa, S.: Calorimetric biosensors with integrated microfluidic channels. *Biosens. Bioelectron.* **19**, 1733–1743 (2004)
50. Queen, D.R., Hellman, F.: Thin film nanocalorimeter for heat capacity measurements of 30 nm films. *Rev. Sci. Instrum.* **80**, 063901 (2009)
51. Tagliati, S., Rydh, A.: Absolute accuracy in membrane-based ac nanocalorimetry. *Thermochim. Acta* **522**, 66–71 (2011)
52. Cahill, D.G.: Thermal-conductivity measurement from 30-K to 750-K - the 3-omega method. *Rev. Sci. Instrum.* **61**, 802–808 (1990)

Nanocarriers

► [Acoustic Nanoparticle Synthesis for Applications in Nanomedicine](#)

Nanochannels for Nanofluidics: Fabrication Aspects

Aeraj ul Haque¹ and Alope Kumar²

¹Biodetection Technologies Section, Energy Systems Division, Argonne National Laboratory, Lemont, IL, USA

²Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Synonyms

[Nanofluidic Channels](#); [Nanopores](#); [Nanoslits](#)

Definition

Nanochannels or nanoslits are defined as channels with at least one physical dimension equal to or less than 100 nm. The term has, however, been loosely applied to cover channels with a critical dimension as large as 500 nm. Nanopore is a term sometimes used interchangeably with nanochannels. Here, nanochannels refer to channels where fluid flow is on the surface of the substrate, while nanopores are referred to as channels where flow is through the substrate. Another important concept in the field of nanochannels is 1D and 2D nanochannels. A 1D nanochannel can be classified as a channel with both physical dimensions in the nanometer regime, while a 2D nanochannel is defined as one with only one dimension in the nanometer range. These terms, however, have been used interchangeably by some researchers leading to confusion. A better parameter is aspect ratio, that is, the ratio between the height and width or between the diameter and length (e.g., in case of nanotubes). A ratio of more than 100 would represent a high aspect ratio. Nanofluidics refers to study and manipulation of fluid flow in nanochannels. This is the area where nanochannels are finding most promising applications and the following discussion will only focus on applications of nanochannels to nanofluidic technologies.

Overview

Nanochannels are defined as channels that have at least one physical dimension that is less than or equal to 100 nm. The area of research that deals with the behavior of fluids, particularly ionic fluids in nanochannels is called nanofluidics. While nanofluidic encompasses analysis of liquid, gas, or super critical fluid in nanochannels, it is the study of ionic liquids and charged particles in nanochannels that has generated the most interest. This is because of the unique nanoscale electrostatic interactions, entropy, electrostatics, and steric phenomenon observed in nanochannels that are not observed in microchannels [1]. In nanochannels, the Debye screening length that defines the extent to which surface charge of the nanochannel walls interacts with the bulk fluid is now comparable to the critical dimension of the nanochannel [2]. This proximity allows the wall surface charge to interact with charged particles such as

ion, proteins, or DNA in fluid inside the nanochannel, and surface charge effects can no longer be ignored as is the case in microfluidics. The dominance of ionic interactions leads to unique transport phenomenon caused by complex electrical double layer-electrical double layer (EDL-EDL) interactions, dominant streaming potentials, and particle/biomolecule-EDL interactions [1].

This behavior of fluids in nanochannels is now being exploited for applications targeting biomolecule separation and detection. Some of the techniques used include *co-ion exclusion*. Here the concentration of co-ions is highly depleted in the nanochannel compared to the bulk solution. Electrokinetic flow causes other interesting phenomenon like *concentration polarization* where ions of opposing charges are concentrated at the anode and cathode [3]. Using this approach DNA was shown to be depleted at one end and concentrated at the other end of a nanochannel. These features of nanofluidics have led to extensive fundamental studies and they are now being exploited for biological applications [1]. Biosensing using nanofluidics is beneficial because the small length scales and electrostatic interactions can be leveraged for facile separation of biomolecules, increasing the speed and sensitivity of bioassays. Many applications have targeted DNA separation, concentration, and detection.

Nanochannels are also relevant to cellular organisms. Every living cell is host to a broad collection of nanochannels that traverse the cell membrane and take part in active transport of ions and key nutrients. The most important ones are ion channels; they are central to balancing the intra- and extracellular concentrations of the physiological ions Na^+ , K^+ , Ca^{2+} , and Cl^- [4] maintaining the cell membrane potential. Also important are aquaporins that nonselectively transport water, ions, and metabolites through the cell. These naturally occurring macromolecular protein-based nanochannels play other significant roles including cell signaling and are the target of many therapeutic drugs. It is no coincidence that the first attempts at nanochannel synthesis involved the naturally occurring protein nanopore – α -hemolysin, a 10-nm long and 1.4-nm internal diameter protein that was involved in lipid bilayer membrane. The setup was used to study DNA translocation through the nanopore in the presence of an electric field, and was promoted as an extremely sensitive single molecule detector. Since then many efforts have focused on fabricating nanopores as small as 2 nm

that can be employed for unique applications in single molecule detection and DNA sequencing.

Nanofabrication is a key enabler of nanofluidic research and like microfluidics, the area of nanochannel fabrication has greatly benefitted from the rich knowledge in silicon processing available from the microelectronics industry. Silicon and silicon oxide are, therefore, the most commonly used substrates for nanochannel fabrication. Recently, other cheaper materials such as polydimethylsiloxane (PDMS) and polymethylmethacrylate (PMMA) are becoming popular alternatives for creating nanochannels.

Nanochannels can be fabricated using either top-down approach or bottom-up approach. The top-down approach involves selective deposition and removal of material from substrates generally using classical nanofabrication methods such as photolithography, wet and dry etching, and substrate bonding. This is currently the dominant approach and complex features can be fabricated with relative ease. The equipment and protocols used for top-down fabrication are however expensive, which can drive up the cost of nanochannel devices substantially. In the bottom-up approach, nanochannels are either created through self-assembly or through electrochemical treatment of a suitable substrate such as alumina. This approach can be more affordable and a large array of nanochannels can be created very easily, though spatial control is difficult. Therefore, it is not easy to fabricate complicated geometries and interfacing with the macro-world might also be a challenge.

Historical Perspective

In 1939, bacterial gramicidin, a naturally occurring antibiotic and a nanochannel were isolated by Rene Dubos. The nAChR ion channel was the first naturally occurring ion channel to be isolated. In 1996, α -hemolysin was incorporated into a lipid bilayer and in a first, the possibility of using nanochannels for performing a practical application – DNA sequencing was demonstrated [5]. Fabrication of tunable nanopores in SiO_2 and SiN with minimum dimensions ~ 1.8 nm using ion beam sculpting was first reported by Golovchenko and colleagues at Harvard in 2001. Dekker's group at Delft demonstrated the patterning of 20-nm diameter pores in Si, SiO_2 , and SiN with electron beam lithography (EBL) and then shrinkage of these pores under tunneling electron microscope

(TEM). Nanoimprint lithography was developed in 1996 by Chou and colleagues [6] and in 1997, Stern et al. [2] demonstrated the fabrication of 50-nm high nanochannels by etching a buried interfacial amorphous silicon sacrificial layer.

Basic Methodology and Key Findings

Fabricating Nanochannels with the Top-Down Approach

Generally the top-down techniques used for fabricating nanochannels employ fabrication protocols borrowed from the microelectronics industry. These include photolithography, physical deposition, and wet and dry etching, to name a few.

Photolithography

The simplest and most convenient method of fabricating nanochannels is through conventional UV photolithography. A UV-sensitive photopolymer or photoresist is spun coat on silicon, SiO₂, glass, or other appropriate substrate and then exposed to a UV light source through a glass mask. This mask has the nanochannel pattern defined in it. The pattern can also be exposed serially using a UV laser writer. The photopolymer is developed in a chemical developer that removes the unexposed or exposed resist (depending on the resist tone) and transfers the pattern in the photoresist. This is followed with a wet or dry etching step that etches the channel in the substrate. The resolution of conventional UV photolithography used in academia is generally in the micron range (1–3 μm). With advanced photolithography techniques using deep UV, extreme UV light source, or X-rays and expensive diamond masks, resolution below 100 nm can be obtained. These instruments and techniques are, however, very expensive and not readily available in the academia. Therefore, fabrication of nanochannels by photolithography is generally achieved by controlling the etch depth. Nanochannels fabricated using conventional photolithography are, therefore, 2D nanochannels. The channels are closed by bonding silicon, glass, or PDMS cover on the substrate. High aspect ratio nanochannels with depth as small as 2 nm have been fabricated with this method [7].

Electron Beam Lithography (EBL)

In EBL, a high-powered electron beam is used to transfer the pattern into an appropriate e-beam-

sensitive resist. This is a maskless process as the pattern is directly written into the resist with the e-beam. The most commonly used e-beam resist is polymethylmethacrylate (PMMA), but now more advanced resists such as ZEP and Hydrogen Silsequioxane (HSQ) that provide better resolution are becoming increasingly popular. Beam diameter smaller than 10 nm are possible; so theoretically line widths of that order can be directly patterned onto the e-beam resist. Practically however, this is a tough task as inelastic scattering often causes over exposure of the resist. Also at very small dimensions over development of the resist is always a concern. This can be overcome using cold development processes at sub-zero temperatures [8]. The patterns can then be transferred in the underlying substrate (Si, glass etc.) with a wet or dry etching process. EBL combined with RIE has also been used to fabricate nanopores that have diameters in the tens of nanometers in Si, SiO₂, and Si₂N₃ membranes. In an interesting discovery made by Dekker's group [9], exposure of nanopores formed in silicon membranes to high-energy electron beams was found to fluidize the silicon and cause shrinkage of the nanopores. Switching the beam off quenches the shrinkage process and the new reduced diameter pore retains its shape. Using this method single-nanometer size pores were fabricated successfully. The central advantage of EBL is the higher degree of control and the possibility of obtaining channels with varying widths in the nanometer regime. The key drawback of EBL is that due to its serial patterning nature, it is a slow technique and thus throughput is very low. It also requires highly trained staff and is relatively expensive.

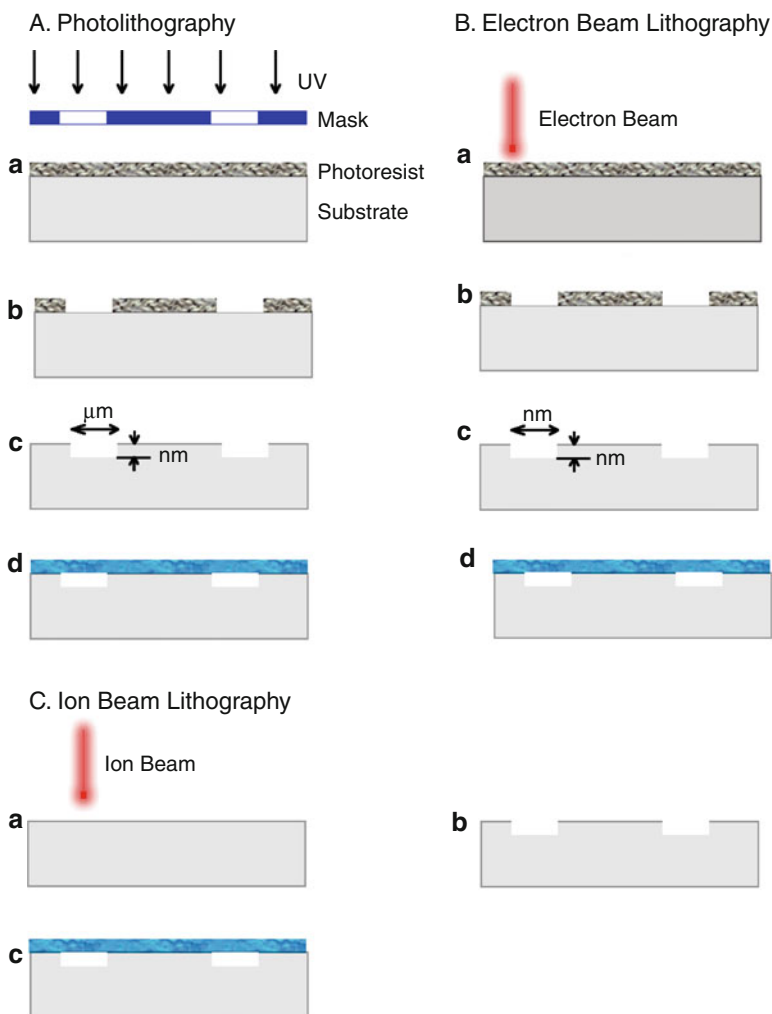
Ion Beam Lithography (IBL)

Similar to EBL, Ion Beam Lithography is a serial patterning technique that uses a focused beam of ions such as protons, or heavy ions such as gallium, to directly write the pattern. The heavier ions scatter less and therefore, IBL is assumed to yield better resolution compared to EBL. It can be used for directly etching and fabricating nanochannels in substrates like silicon and glass without the need of a photoresist, a major advantage over EBL. Nanochannels down to 20 nm were directly milled with IBL in silicon and glass, and proton beams were also used to fabricate nanochannels in polymers such as PMMA. Nanopores in silicon oxide and silicon nitride membranes as small as 5 nm have been reported with IBL [10]. This technique,

Nanochannels for Nanofluidics: Fabrication Aspects, Fig. 1

(A) Photolithography

Nanochannel patterning using photolithography; (a) Photoresist spun coat on the nanochannel substrate and exposed to UV light through a photomask that has the nanochannel pattern defined in it. After development of the photoresist (b) the substrate is etched with RIE to the desired nm depth. A second substrate is bonded on this substrate to close the nanochannel. (B) Nanochannel definition with EBL; (a) An electron beam is used to directly write the nanochannel pattern on an e-beam-sensitive resist spun coat on the substrate. After photoresist development (b) RIE is used to etch nanochannels in the substrate. (C) With IBL an ion beam is used to directly etch the pattern in an appropriate substrate



however, suffers from the same drawbacks as EBL – it has very low throughput, requires highly skilled researchers for operation, and is expensive. The doping of the substrate with the heavy metal ions is also a concern (Fig. 1).

Sacrificial Technique

In sacrificial techniques, an intermediate sacrificial layer is generally used between the two substrates that will be bonded to form the nanochannels. A less than 100 nm thick layer of a sacrificial material in the pattern of the nanochannel is first deposited on the nanochannel substrate in the desired thickness. This thickness later defines the depth of the nanochannel. For this reason it is important that a sacrificial material is chosen that can be deposited at a thickness below 100 nm with high uniformity. Commonly used

sacrificial materials include amorphous silicon, polysilicon, and even metals such as chromium and gold. The deposited sacrificial layer is first patterned generally using photolithography and dry or wet etching into the desired nanochannel shape. The second nanochannel substrate is then bonded with the substrate that was patterned with the sacrificial layer. The two substrates bond only in places where they come in direct contact with each other and do not bond with the sacrificial layer. In another variation, a structural layer can also be deposited on the sacrificial layer. The sacrificial layer is now wet etched in an appropriate etchant that selectively etches the sacrificial layer, but does not or minimally affects the substrates, leaving behind nanochannels. Using this approach it is possible to make high aspect ratio channels, and nanochannels that were 20-nm deep,

0.5–200- μm wide and up to 4-mm long were reported [2]. Since wet etching of the channels is solely based on diffusion inside the channel, this is a slow approach. It can take a couple of days to etch through the entire sacrificial layer. Also because selectivity of etchant toward sacrificial layer versus the substrate is not always 100%, tapering is often observed in nanochannels fabricated using this technique.

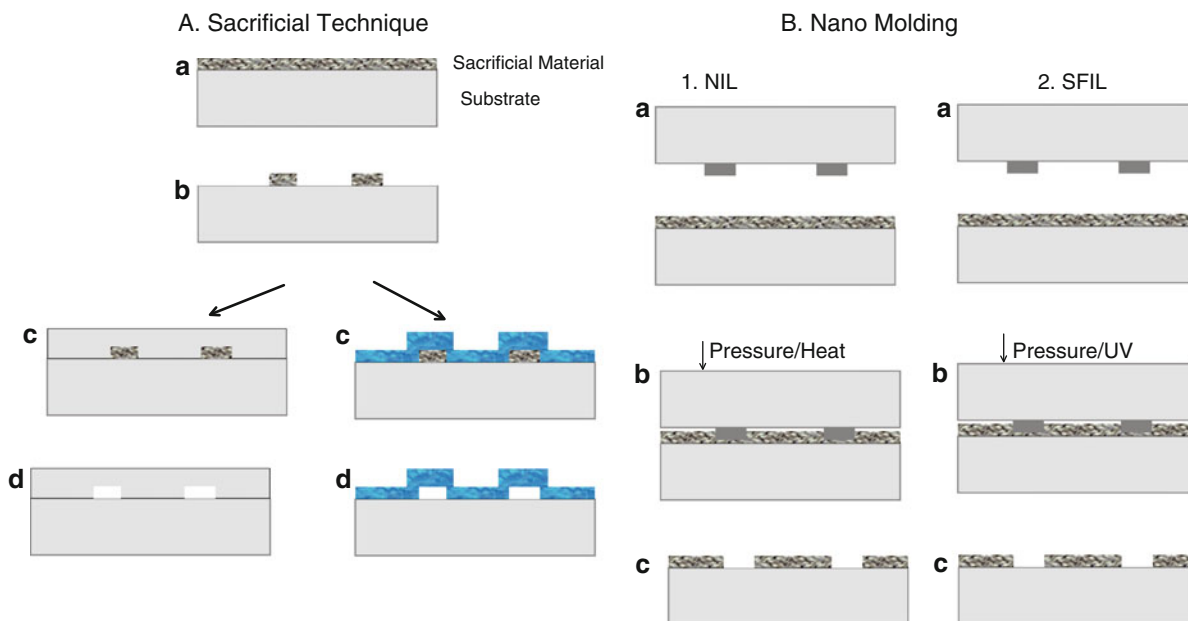
Nanomolding Techniques

Micromolding of polymer elastomeric materials provides a cheap and easy way for rapid prototyping of devices and therefore occupies widespread popularity in the microfluidics arena. Polydimethylsiloxane (PDMS) is by far the most popular material used for micromolding of fluidic channels because it can retain the features of the mold with up to 99% accuracy. Nanochannel fabrication in elastomeric materials is much more difficult mainly because at nanometer dimensions, they tend to deform and collapse [6]. This inhibits direct molding of nanochannels in PDMS using molds of the nanochannels. If designed properly, however, this deformation and collapse can actually be used for fabricating nanochannels. One method is crack-induced generation of nanochannels in PDMS. When a slab of PDMS is surface oxidized, for example, in an oxygen plasma, the surface hardens creating an elastic modulus mismatch between the surface and the bulk PDMS. Application of a uniaxial tensile force generates cracks in the slab forming triangular nanochannels. The nanochannels can be sealed by oxygen plasma bonding with another PDMS slab, glass, silicon or by using this slab as a mold for transferring the nanochannel pattern into another PDMS layer. Nanochannels with heights of 80 nm have been fabricated with this technique. In an alternative approach, thinner and thicker PDMS slabs that have both been surface oxidized in oxygen plasma are bonded together and a uniaxial tensile force is applied. This causes crack generation at the weak oxidized bonding interface producing diamond-shaped nanochannels. In both these cases, tuning of the tensile force controls the depth of the nanochannels. In a similar approach called wrinkle nanochannels (WNC), a stretched slab of PDMS is exposed to an oxygen plasma or UV light source and then relaxed. This creates wrinkle patterns along the PDMS surface. The wavelength of this sinusoidal wave pattern is controlled by the difference in elastic moduli of the

PDMS slab and the substrate on which it is bonded, while the amplitude depends on the applied strain. With this technique nanoslits with widths and heights in the tens of nanometers have been fabricated.

Other researchers have used the inherent elasticity of PDMS that allows it to deform around structures to fabricate nanochannels. In the first method, shallow beams were fabricated on a silicon substrate using reactive ion etching (RIE). A thin film of oxygen plasma-treated PDMS is bonded on this silicon substrate. The PDMS film deforms in the crevices of the substrate in a wave like pattern and bonds to the top of the beams and part of the base of the crevice. Nanosized gaps at the edge of the beams and the point where PDMS bonds at the base of the crevice result in the formation of nanochannels. In a similar technique called roof collapse technique, precision steps with strictly controlled height-to-width ratio of 0.2 were fabricated [6]. These are then used to mold microchannels in PDMS. When this PDMS slab is bonded to a substrate, the roof of the microchannel collapses. While the bulk of the channel is sealed nanoslits still remain open at the edges of the microchannels. With this technique nanochannels with heights as small as 60 nm have been fabricated. This technique is particularly applicable for creating variously shaped nanochannels.

Unlike elastomers, thermoplastics such as poly (methyl methacrylate) (PMMA), cyclic olefin co (polymer) (COC), and poly(ethylene terephthalate) (PET) deform negligibly and can retain their shape on direct imprinting with nanoscale features. This is the principle behind nano imprint lithography (NIL) pioneered by Chou et al. [11, 17]. A nanomold or template is first created generally using EBL followed by RIE in silicon or glass, or by direct writing using focused ion beam (FIB). A thermoplastic resist is spin coated on another substrate such as silicon and heated above its glass transition temperature (T_g). The template is now pressed on the spun coat thermoplastic under vacuum and the thermoplastic is allowed to cure. The template is released and a short RIE is performed to remove the residual thermoplastic leaving the nanochannel patterns embossed in it. The thermoplastic can itself be used as the material for the nanochannels or the thermoplastic can be used as a mask for etching the underlying substrate with RIE. With this technique 20-nm deep nanochannels that are centimeters in length were fabricated for nanofluidic applications. In a similar



Nanochannels for Nanofluidics: Fabrication Aspects, Fig. 2 (A) Nanochannel patterning with sacrificial technique; (a) An appropriate sacrificial material is deposited on a substrate and the nanochannel pattern defined with photolithography (b). Then either a second substrate is directly bonded on top of this substrate or another structural material is deposited (c). The

technique called step and flash imprint lithography (SFIL), a UV curable polymer is spun coat on an appropriate substrate. A template of the nanochannel pattern is fabricated in a transparent substrate as discussed before and pressed on the photoresist. The photoresist is exposed to UV light through the transparent template curing it and embossing the pattern in it. The template is removed and a short RIE step is performed to remove the residual photoresist layer. Another embossing technique is called reversal imprint. Here, a liquid polymer or prepolymer is spin coated directly on the template that is previously surface treated with fluorosilanes or other adhesion inhibitors. The template with liquid polymer is now transferred on another substrate and hardened. The template is removed such that the hardened polymer is left on the substrate. In a novel technique, silica nanowires were used as the nanochannel template. Silica nanowires, produced by tapering single-mode optical fibers in an alcohol flame [6], were positioned on a glass substrate using a scanning tunneling microscope (STM). The nanowire was hot embossed in a polycarbonate substrate and then etched away with hydrofluoric acid yielding the nanochannel in

sacrificial material is then etched leaving the nanochannels. (B) A thermoplastic (1a) or a photosensitive (2a) polymer is spun coat on a substrate and embossed with a mold of the nanochannels. The polymer is then applied to heat and pressure (1b) or UV light (2b) curing the polymer. The mold is removed (1c, 2c) leaving the nanochannels defined in the polymer

polycarbonate. The polycarbonate nanochannel was closed with PDMS. The need for STM position makes this technique very slow. Also this nanowire molding is confined to straight nanochannels.

The greatest advantage of nanomolding techniques is that they can yield high throughput and are relatively inexpensive. Particularly, NIL is very popular since equipment and materials for nanomolding using this technique are now available commercially through NanoNex[®]. The real expensive part in nanomolding techniques is the template that still has to be fabricated with EBL or FIB, but once the mold is machined it can be reused to imprint multiple devices (Fig. 2).

Miscellaneous Top-Down Techniques

Edge Lithography takes the advantage of asymmetry at the edge of a structure to fabricate nanochannels. In one example, polysilicon was deposited at the edge of a step on SiO₂ film. This was followed with anisotropic etching of polysilicon and the SiO₂ film such that the thickness of the polysilicon is converted into the width of the nanochannel. In another example, a 10-nm thick SiO₂ film was deposited on a Si₂N₃ substrate and then further coated with another Si₂N₃ masking layer. With

controlled under etching 20 nm high and 200 nm wide nanochannels were fabricated. *Phase shift lithography* (PSL) is also considered as an edge lithography technique [2] that uses phase-shift photomasks to beat the diffraction limit of UV light. These are transparent masks that include trenches with depth $d = \lambda [2(n-1)]$ where λ is the wavelength of light and n is an integer. The phase shift in light caused by the trenches result in destructive interference, and interference patterns that are below 100 nm resolution down to 30 nm have been reported with this technique. *Interference lithography (IL)* is another interference pattern-based lithography technique. Here, the photoresist is exposed to the interference pattern produced by two coherent laser beams whose fringe period is given by $T = (\lambda/2) \sin(\theta/2)$ where θ is the angle between the two light waves. IL is difficult and the setup to produce the interference patterns is complex but it can yield a dense array of nanochannels at a high rate without a mask. There is, however, limited flexibility in the number of shapes that can be patterned with this technique.

A novel technique was demonstrated by Han's group using *junction gap breakdown* to fabricate nanogap nanochannels in PDMS [12]. Two chevron-shaped microchannels were fabricated on either side of a horizontal microchannel such that there was a 40 μm gap between the tip of the chevron channels and the central horizontal channel. The channels were filled with an ionic buffer and a high voltage was applied with the horizontal channel as the anode while the chevron channels as cathode. This caused a junction gap breakdown at the tip of the chevron channels resulting in nanogap nanochannels. Pulling optical fibers and capillaries that are heated above their glass transition temperature can also be used to fabricate nanochannels. Co-axial fibers that were a mixture of silica sol-gel and motor oil were extruded to yield silica nanochannels with an inner diameter of 20 nm after annealing and removal of the oil. Only vertical nanochannels shaped like venturics can be fabricated with this technique (Fig. 3).

Atomic Force Microscopy (AFM) Nanolithography has also been used to create nanoscale features and depressions on substrates. Using high aspect ratio Si tips 80-nm wide and 30-nm deep channels were fabricated in PMMA [6]. AFM nanolithography like EBL and FIB is a serial technique and is in fact even more tedious and slower than the beam-based techniques.

While throughput can be increased using multiple probe tips such as on Dip Pen Nanolithography systems available from NanoInk[®], limitations in scan area and the collateral damage around the indentation site reduce the practicality of this technique.

An attractive method of fabricating nanochannels is the fabrication of microchannels in a substrate followed by anisotropic deposition of another material. Controlled deposition converts the microchannel to a nanochannel. Using this approach parylene was deposited in microchannels etched in silicon. The parylene deposited preferentially at the top of the channel compared to the base. This caused the channel to close at top forming parylene nanotubes. As parylene is biocompatible these nanotubes can find biomedical applications such as in implants or as synthetic arteries or veins. A layer-by-layer polyelectrolyte multilayer deposition approach and non-conformal deposition was also used to convert microchannels to nanochannels [13]. Microchannels 1 μm wide and 15 μm deep were fabricated in silicon with photolithography and deep RIE. Plasma Enhanced Chemical Vapor Deposition (PECVD) was then used to deposit silica on the device. The silica deposited preferentially on the top of the channels reducing its width at that location. Next conformal Low Pressure Chemical Vapor Deposition (LPCVD) of polysilicon was deposited that closed the microchannel at the top. After thermal oxidation to reduce the channel width to the desired level, layer-by-layer deposition of polyelectrolyte films was performed to reduce the channel width to 200 nm.

Bottom-Up Nanochannel Fabrication Techniques

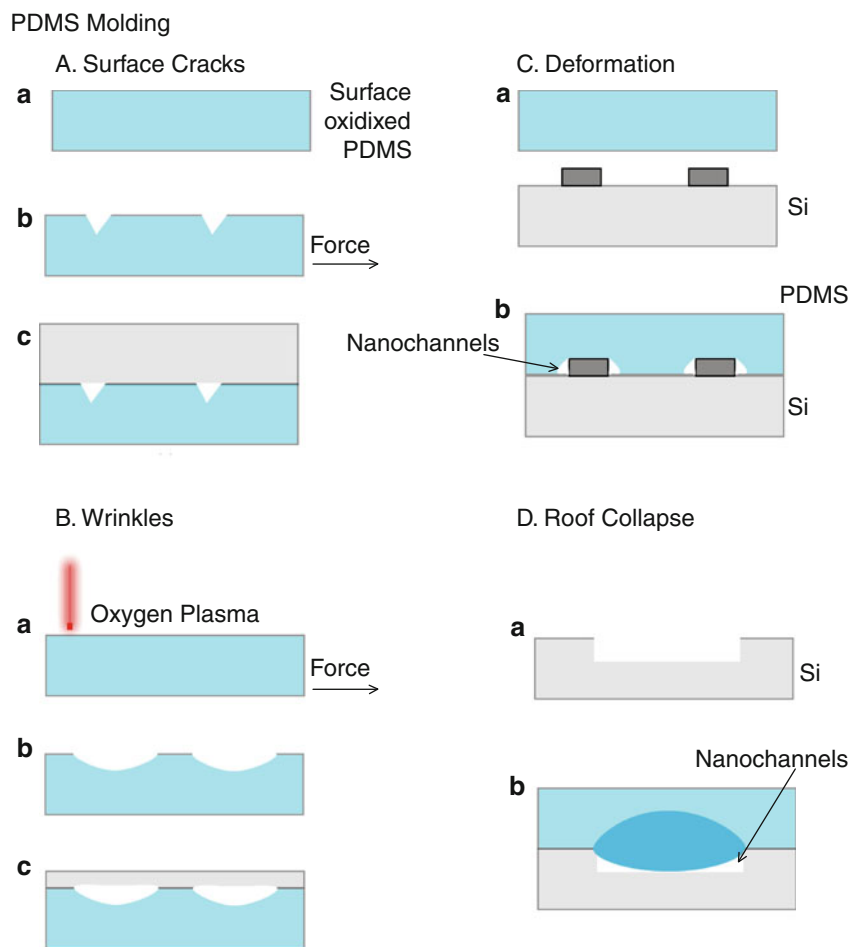
Bottom-up fabrication of nanochannels involves processes where nanochannels are created electrochemically, through controlled atomic synthesis and assembly of synthetic nanochannels/nanopores and through self-assembly of polymers or naturally occurring nanochannels.

Electrochemical Synthesis

Thin sheets of aluminum can be anodized in acid solutions such as oxalic acid, sulfuric acid, or phosphoric acid with application of voltage/voltage cycles to produce aluminum oxide nanochannel arrays. These alumina nanochannel arrays called anodic aluminum

Nanochannels for Nanofluidics: Fabrication Aspects, Fig. 3

Fabrication of nanochannels in PDMS using (A) production of nanochannels through formation of cracks and (B) formation of wrinkles. Both techniques use surface oxidation of PDMS and application of a tensile force to create these nanochannels. Bonding with a second PDMS, silicon or glass layer closes the nanochannels. Nanochannels can also be formed in PDMS using deformation around pillars patterned in a substrate (C) or collapsing of the roof that yields nanochannels at the bottom corners of the pillar (D).



oxide (AAO) or porous anodic alumina (PAA) have been around for a long time and are primarily used as nanofilters. Recently interest in using AAO as nanotemplates for ordered carbon nanotube (CNT) or nanowire growth and in nanofluidic applications has grown considerably. The anodization process involved the growth of an initial oxide layer on the surface of the aluminum film. Oxygen molecules then penetrate and oxidize the aluminum only at sites that provide low electrical resistance. The oxide dissolves in the acid resulting in pore nucleation and growth at these sites. The applied electrical field causes the pores to penetrate perpendicular to the aluminum surface until they traverse the entire alumina film. The pore radius, lengths, and densities can easily be tuned by controlling the number of the anodization voltage cycles and the applied voltage [14]. NIL was also used to produce dimples on the aluminum film that act as anodization nucleation sites [2].

Nanotubes: Carbon Nanotubes and Nanowires

With advances in physical vapor deposition techniques and catalysis it is now possible to synthesize nanotubes and nanowires with excellent control. Among these, carbon nanotubes (CNTs) have enjoyed great attention. CNTs are rolls of graphene, an atomic layer of carbon arranged in classic honey comb pattern. They can be single-walled (SWCNTs) and multi-walled (MWCNTs) and can have diameters from a couple of nanometers to hundreds of nanometers. They have excellent mechanical and electrical properties (tensile strength in the GPa range) and can be conducting or semi-conducting. They can easily be functionalized with biomolecules, leading to their vast use in bio-applications. CNTs can be synthesized using arc discharge, laser ablation, and chemical vapor deposition (CVD and MOCVD) methods. In the arc-discharge process, carbon atoms are vaporized in helium plasma ignited by an electric arc produced by applying high

currents between a carbon anode and cathode. This process produces MWCNTs. To produce SWCNTs a metal catalyst such as iron or cobalt is added to the system. Laser ablation is similar to arc discharge, that is, it involves evaporation of a carbon target but now high-intensity laser pulses are employed for ablation. The process is generally performed in a tube furnace heated to 1,200°C. The furnace is flushed with an inert gas that blows the synthesized CNTs downstream where they are collected. The CVD process involves synthesis of CNTs by the decomposition of a hydrocarbon such as acetylene in the presence of an appropriate catalyst. After nanotube synthesis the reactor is cooled down and the CNTs collected. By controlling the temperature and pressure inside the reactor, careful selection of precursor hydrocarbon and catalyst SWCNTs and MWCNTs can be produced in the same reactor. The CVD process is particularly advantageous for growing ordered arrays of CNTs on substrates that have been appropriately seeded with a catalyst. An MOCVD process was used to fabricate CNT array in a PAA template that has found applications in CNT transistors and as a glucose biosensor [14]. While CNTs are readymade nanopores, their application is tricky mainly because of integration with the macro-world. CNT nanopores were fabricated in by trapping CNTs with dielectrophoresis and then spin coating SU-8 on top of it that is patterned to produce fluidic reservoirs.

Nanowires can also be used as templates for the fabrication of nanochannels. Nanowires can be synthesized using various techniques such as electroplating (for metallic nanowires), chemical vapor deposition, vapor–liquid–solid method, etc. Bottom-up fabrication of SiO₂ nanochannels was demonstrated with Si nanowires. The Si nanowires were surface oxidized via a timed oxidation process such that the core was still Si. Isotropic etching of the Si core was performed with XeF₂ gas that yielded SiO₂ nanochannels with hollow cores.

Block Copolymers

Block copolymers are soft materials that are composed of two or more immiscible macromolecules. These macromolecules can self-assemble into nanoscopic arrangements with each macromolecule comprising the block polymer maintaining its original properties [15]. Block copolymers can be used to form nanoporous membranes that have ordered nanopores.

Polystyrene-block-PDMS (PS-PDMS) was spin coated on a silicon substrate. The block copolymer self-assembled to form ~16 nm diameter vertical cylinders of PS arranged uniformly at a pitch of ~23 nm in a PDMS block. Exposure to oxygen plasma etched the PS leaving nanopores in the unetched PDMS. The PDMS was later used as a mask for etching the underlying substrate. More recently efforts are being made to arrange long channel like arrays of block copolymers. One of the block copolymer phases can then be selectively etched to form long nanochannels. Misalignment over long distances limits the maximum feature length.

Self-assembly of Naturally Occurring Nanochannels

Naturally occurring protein nanochannels are present in the cell phospholipid bilayer and the intracellular membranes of all eukaryotic and prokaryotic cells. They perform critical cellular functions from maintaining the ion balance, transport of metabolites to cell signaling, communication, and balancing osmotic pressure. Particularly important are ion channels that are also the target of therapeutics. Efforts have been directed toward isolating ion channels and using them as nanochannels for other applications. Porin MspA, a 3.1 nm internal diameter *Mycobacterium smegmatis* channel protein selective for hydrophobic nutrients, was self-assembled inside highly ordered pyrolytic graphite using hydrophobic interactions [16]. In another approach 1.4–4.6 nm nanopores were fabricated by self-assembly of α -hemolysin monomers in a lipid bilayer. Thus far demonstrations of self-assembled isolated naturally occurring nanochannels or their monomers are limited and practical applications are nonexistent. This is mainly because of difficulty of integrating such nanochannels with complex nanofluidic devices. Nonetheless, this represents a very interesting area of research and applications are expected to grow as better ways of integrating the top-down and bottom-up approaches are developed.

Future Directions

Nanofluidics is a frontier area of science today, and ease and affordability of nanofabrication can be key attributes to the broader embrace of the technology. Nanochannel fabrication on substrates like silicon and glass has its advantages such as high fidelity, but

fabrication in such cases can suffer from high fabrication costs and low throughput. Advances in polymer-based nanofabrication can offset some of these disadvantages. However, polymer-based fabrication techniques can suffer from the inherent lack of mechanical rigidity of the material, leading to deformation issues. As advances in nanofabrication are made, the quest for customized surface modification of channel walls and integration of various components into nanochannels will also increase.

Cross-References

- ▶ [Applications of Nanofluidics](#)
- ▶ [Electron Beam Lithography \(EBL\)](#)
- ▶ [Nanoimprint Lithography](#)
- ▶ [Surface-Modified Microfluidics and Nanofluidics](#)

References

1. Napoli, M., Eijkel, J.C.T., Pennathur, S.: Nanofluidic technology for biomolecule applications. *Lab Chip* **10**, 957–985 (2010)
2. Abgrall, P., Nguyen, N.T.: Nanofluidics devices and their applications. *Anal. Chem.* **80**, 2326–2341 (2008)
3. Kovarik, M.L., Jacobson, S.C.: Nanofluidics in lab-on-a-chip devices. *Anal. Chem.* **81**, 7133–7140 (2009)
4. Hille, B.: *Ion Channels of Excitable Membranes*, 3rd edn. Sinauer, Sunderland, MA (2001)
5. Dekker, C.: Solid-state nanopores. *Nat. Nanotechnol.* **2**, 209–215 (2007)
6. Chantiwas, R., Park, S., Soper, S.A., Kim, B.C., Takayama, S., Sunkara, V., Hwang, H., Cho, Y.K.: Flexible fabrication and applications of polymeric nanochannels and nanoslits. *Chem. Soc. Rev.* **40**, 3677–3702 (2011)
7. Duan, C., Majumdar, A.: Anomalous ion transport in 2-nm hydrophilic nanochannels. *Nat. Nanotechnol.* **5**(12), 848–852 (2010)
8. Ocola, L.E., Stein, A.: Effect of cold development on improvement in electron-beam nanopatterning resolution and line roughness. *J. Vac. Sci. Technol. B* **24**, 3061–3065 (2006)
9. Storm, A.J., Chen, J.H., Ling, X.S., Zandbergen, H.W., Dekker, C.: Fabrication of solid-state nanopores with single-nanometre precision. *Nat. Mater.* **2**, 537–540 (2003)
10. Lo, C.J., Aref, T., Bezryadin, A.: Fabrication of symmetric sub-5 nm nanopores using focused ion and electron beams. *Nanotechnology* **17**, 3264–3267 (2006)
11. Chou, S.Y., Krauss, P.R., Renstrom, P.J.: Imprint lithography with 25-nanometer resolution. *Science* **272**, 85–87 (1996)
12. Lee, J.H., Chug, S., Kim, S.J., Han, J.: Poly(dimethylsiloxane)-based protein preconcentration using a nanogap generated by junction gap breakdown. *Anal. Chem.* **79**, 6868–6873 (2007)
13. DeRocher, J.P., Mao, P., Han, J., Rubner, M.F., Cohen, R.E.: Layer-by-layer assembly of polyelectrolytes in nanofluidic devices. *Macromolecules* **43**, 2430–2437 (2010)
14. Claussen, J.C., Franklin, A.D., Ul-Haque, A., Porterfield, D.M., Fisher, T.S.: Electrochemical biosensor of nanocube-augmented carbon nanotube networks. *ACS Nano* **3**, 37–44 (2009)
15. Jackson, E.A., Hillmyer, M.A.: Nanoporous membranes derived from block co-polymers: From drug delivery to water filtration. *ACS Nano* **4**, 3548–3553 (2010)
16. Mijatovic, D., Eijkel, J.C.T., Van den Berg, A.: Technologies for nanofluidic systems: top-down vs bottom-up – a review. *Lab Chip* **5**, 492–500 (2005)
17. Perry, J.L., Kandlikar, S.G.: Review of fabrication of nanochannels for single phase liquid flow. *Microfluid. Nanofluid.* **2**, 185–193 (2006)

Nanocluster

- ▶ [Synthesis of Subnanometric Metal Nanoparticles](#)

Nanocoil

- ▶ [Nanorobotics for NEMS Using Helical Nanostructures](#)

Nanocombs

- ▶ [Chemical Vapor Deposition \(CVD\)](#)

Nano-Concrete

- ▶ [Nano-engineered Concrete](#)

Nanocrystalline Materials

- ▶ [Computational Study of Nanomaterials: From Large-Scale Atomistic Simulations to Mesoscopic Modeling](#)
- ▶ [Nano-sized Nanocrystalline and Nano-twinned Metals](#)

Nanodentistry

Hans Deyhle, Simone Hieber and Bert Müller
Biomaterials Science Center (BMC), University of
Basel, Basel, Switzerland

Synonyms

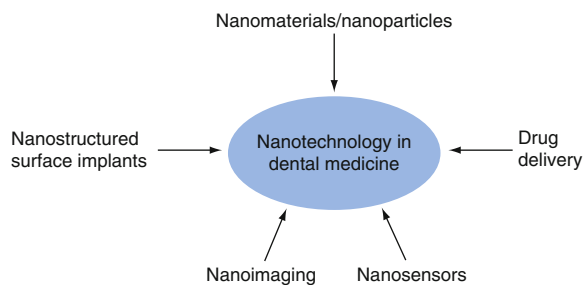
[Nanotechnology in dental medicine](#)

Definition

Nanodentistry is defined as the science and technology of diagnosing, treating, and preventing oral and dental diseases, relieving pain, and of preserving and improving dental health, applying materials structured on the nanometer scale [1]. The nanotechnology considers nanostructures between 1 and 100 nm exhibiting properties and functionalities that fundamentally differ from that of other length scales, as the surface of the nanostructures dominates the material properties usually given by the bulk. The nanomaterials are not only promised to improve the properties and functionalities of dental products but also to lead to the development of innovative, novel products for the benefit of patients. In particular, nanomaterials have been part of engineered products, such as implants, dental and surgical instruments, dental care wares, as well as diagnostics tools.

Overview

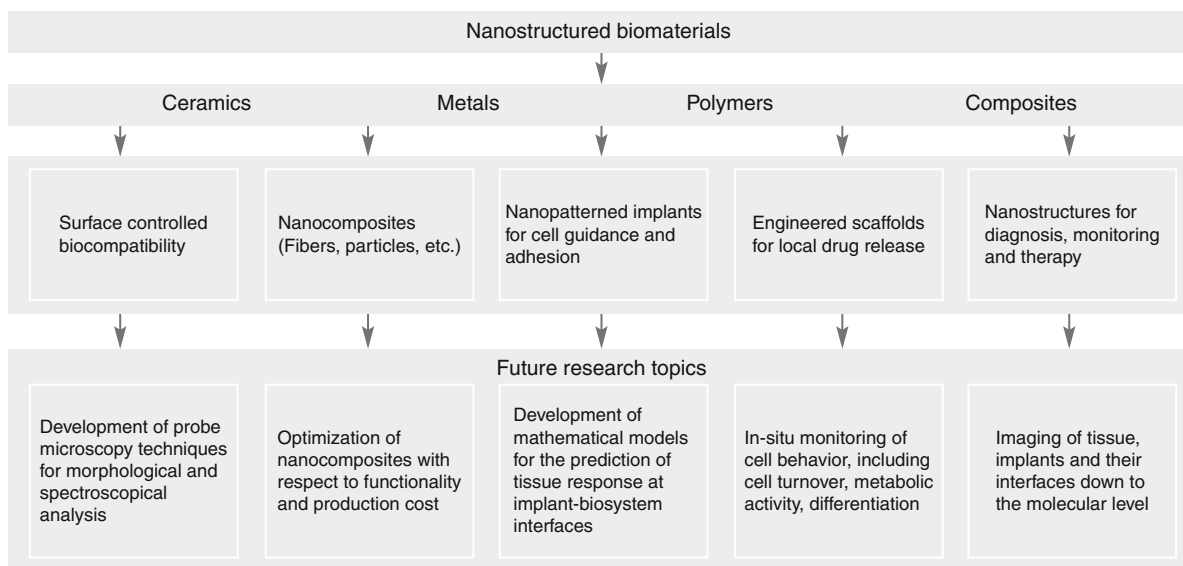
The cover story of R. A. Freitas Jr. in the *Journal of the American Dental Association* more than a decade ago [2] introduced the term *nanodentistry* to a larger community. He developed visions effectively using dental nanorobots for orthodontic realignments, for dentition regeneration procedures and, finally, oral health maintenance. He already elucidated the increasing influence of nanomaterials and the importance of tissue engineering. It has been pointed out that properly configured dentifrobots will recognize and subsequently destroy pathogenic bacteria residing in the plaque and elsewhere.



Nanodentistry, Fig. 1 Key issues currently related to nanotechnology in dental medicine

Although most of Freitas' ideas are still science fiction, today many applications of nanometer-scale components are known in the field of dentistry, which include nanoparticles in sensitive toothpastes, nanostructured surfaces of dental implants to improve osseointegration, and to reduce inflammatory reactions of different kinds of drugs (cp. Fig. 1). Further applications still on the research stage are classified into nanoimaging to reveal the nanometer-sized features in human tissues and man-made implants or nanosensors to quantify physical, chemical, or biological parameters within the oral cavity.

Dental materials are ceramics, metals, and polymers or any combination (cp. Fig. 2). Applying the knowledge and experience of nanotechnology to solve problems in the field of dental medicine, nanometer-sized patterns have been created on surfaces of implants and surgical instruments and incorporated into the bulk of the dental materials to accomplish the dedicated functionality. Nanostructures can improve the biocompatibility of dental implants or carry drugs toward the target within the oral cavity. Reactive nanostructures are going to emerge in diagnosis, monitoring, and therapy to enhance the patient treatments. Diverse activities augment the basic research of bio-nano-materials. Scanning probe microscopy techniques, for instance, enable us to morphologically and spectroscopically analyze hard and soft tissues to understand the structure–function relationship on the nanometer scale. Sophisticated mathematical simulations allow tailoring the material properties by the combination of nanometer-sized components. The outcome also predicts the tissue response and the self-organization at the implant–biosystem interfaces down to the molecular scale. Highly advanced in vivo measurements support cell biologists to determine the cell behavior with



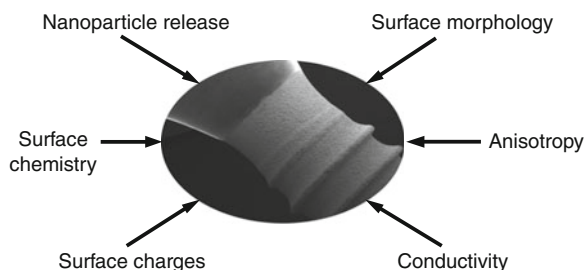
Nanodentistry, Fig. 2 Materials science for nanodentistry

respect to proliferation, differentiation, and metabolism, especially at the interface between the dental material and the surrounding tissues. For this target, high-resolution imaging facilities have become more and more vital to quantify the different kinds of nanostructures with the desired precision.

Biocompatibility

Biocompatibility cannot be simply described as a specific property of the selected biomaterial, as well known from mechanical quantities such as Young's modulus. It is a peculiar term and defined as "... the ability of a material to perform with an appropriate host response in a specific application." [3]. This means it not only depends on the material itself but also from the location within the human body where the material is to be placed. Consequently, biocompatibility is not only related to intrinsic material properties. Its classification relies on standards (mainly ISO 10993 [4]). Therefore, the regular updates of these standards may change the status of a material concerning biocompatibility. Natural scientists and engineers are often not familiar with such a practice and, hence, usually terminate their work on such research areas.

The encouraging argument, however, correlates to the experimental observations that biocompatibility of



Nanodentistry, Fig. 3 Several parameters allow tailoring the biocompatibility of dental implants

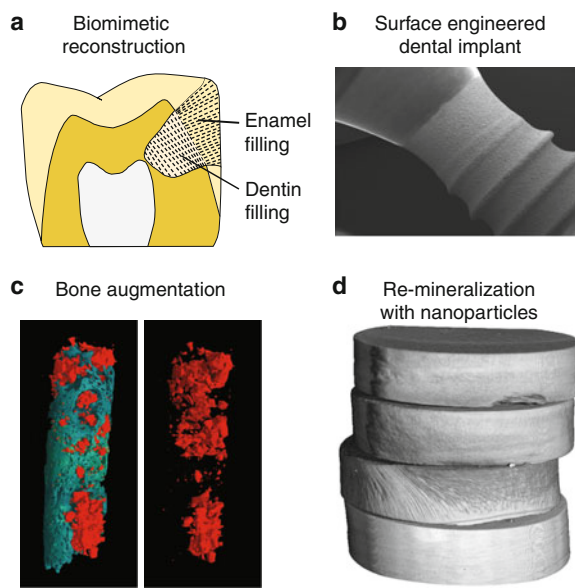
implants can be intentionally manipulated to optimize the clinical outcome. Figure 3 illustrates how some selected phenomena can influence the biocompatibility. For a dental titanium implant, for example, the surface is made rough down to the molecular level by sandblasting and etching procedures to ensure osseointegration and to reduce the inflammatory reactions [5]. Surface morphology can offer certain angles for dedicated protein absorption and activity. In addition, the surface chemistry is commonly tailored concerning oxide thickness, stoichiometry, and normally functionalized to obtain a hydrophilic surface. Further relevant factors are particle and ion release, surface charges, electrical conductivity, and anisotropies as present on patterned surface micro- and nanostructures. The measurement of the surface roughness on real implants is more than challenging,

since there is no single parameter and no single experimental setup, which allow simultaneously determining roughness on the micro- and nanometer scales.

Key Research Findings

Figure 4 shows representative applications of nanotechnology in dental medicine. The fillers of reconstruction materials, the surfaces of dental implants, the different kinds of bone-augmentation ceramics and glass ceramics, as well as the re-mineralization of hard tissues by means of nanoparticles exploit nanotechnology by now and possess enormous growing potential. Nature-analogue, anisotropic restorations will replace today's dental filling and inlay materials as well as artificial teeth [6]. Researchers are still searching how the anisotropy can be introduced to obtain nanostructures similar to the ones present in healthy human teeth.

It belongs to the key impacts to realize that the ceramic components in dentin are orthogonally oriented to the nanostructures of the same size in enamel (see Fig. 5). This finding well observable in scanning small-angle X-ray scattering could explain why the dentin-enamel junction acts as an effective crack barrier. Small-angle X-ray scattering in scanning mode, as performed at the cSAXS beamline (Swiss Light Source, Paul Scherrer Institut, Switzerland) allows uncovering the mean orientation, abundance, and anisotropy of nanometer-sized components in human tissues over macroscopic areas [7]. It is in general difficult to isolate the contribution of specific components to the scattering signal, see chapter "Imaging the human body down to the molecular level." The collagen in human dentin, however, yields a characteristic signal related to its 67 nm-wide periodicity. Due to the sensitivity of scattering data to periodic structures, it is possible to extract a signal unequivocally related to collagen, thus allowing gathering component-specific information. Several 400 μm -thin slices from a human third molar were examined. Figure 5 shows the processed scattering signal from the features with sizes between 60 and 70 nm. The mean orientation of the scattering signal is coded according to the color wheel. The brightness relates to the abundance of nanostructures, while color saturation matches the degree of the anisotropy. The orientation of the scattering signal in both enamel and dentin is

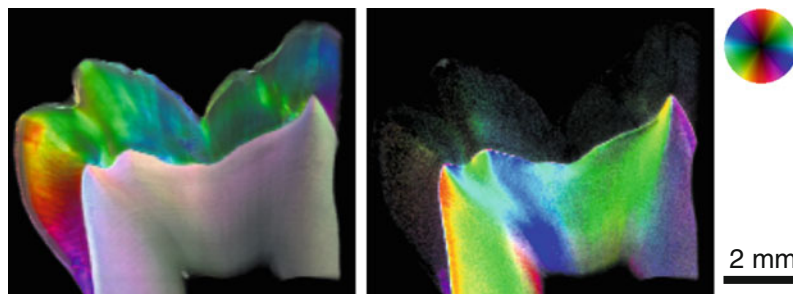


Nanodentistry, Fig. 4 Applications of nanotechnology in dentistry with growing potential

perpendicular to the crystallite orientation. On the left-hand side, the information from the total scattering signal is shown, which is mainly produced by the ceramic components of the tooth. Enamel and dentin can clearly be distinguished due to their different scattering potential. The strong color in the enamel is related to the high degree of anisotropy, that is, the high degree of orientation of the calcium phosphate crystallites, and their strong scattering potential in this range. In the dentin, the anisotropy is less distinct, and decreases with increasing distance from the dentin-enamel junction (DEJ). Moreover, a perpendicular orientation of the scattering signal, and thus of the nanostructures, between enamel and dentin is observed. While the crystallites are mainly oriented perpendicular to the DEJ in enamel, parallel to the main direction of mechanical load, the dentin structures appear parallel to the DEJ, perpendicular to the mechanical loading. It is well known that the DEJ acts as an effective crack barrier, impeding the propagation of cracks from the enamel into the dentin. The image on the right in Fig. 5 shows only the signal related to the collagen matrix. Due to the distinctive periodicity of 67 nm along the collagen fibrils, the scattering signal is oriented parallel to the fibrils. No significant signal is found in the enamel. The mean scattering orientation of the collagen is found to be mainly perpendicular to

Nanodentistry,

Fig. 5 Processed small-angle X-ray scattering data in the range between 60 and 70 nm. The mean orientation of the scattering signal is according to the color wheel. The brightness relates to the nanoparticle abundance, while color saturation codes their degree of anisotropy



that of the total signal on the left-hand image, revealing a parallel arrangement of collagen fibrils with the calcium phosphate crystallites. The collagen abundance is maximal near the DEJ and decreases toward the pulp. Abrupt changes in collagen fibril orientation are found along lines connecting the tooth cusps and the pulp. In this manner, the structure of the human teeth along the entire nanometer range is evaluated.

As pointed out above, the part of dental implants to be fixed in the bone is made rough. It contains nanostructures which resemble the apatite crystallites in human bone and therefore offer nanostructures with angles for suitable protein absorption. The suppliers have their secrets for the surface engineering processing, which includes sandblasting, etching, and electrochemical functionalization. Surface-engineered dental implants are very well established and do have a high success rate. During the next decade, we can expect a significant reduction in price so that a broader community of our aging society can benefit from such surface-engineered implants with dedicated nanostructures.

The calcium phosphate phases for bone augmentation gain more and more importance in our aging population. The resorbable calcium phosphate ceramics or bio-glasses do actively support the bone formation. These bone substitutes are combined with autologous bony tissues to repair larger and larger defects arising from cancer surgery, trauma, or just tooth loss.

There are numerous reasons for the de-mineralization of teeth. The mineral loss, especially owing to attacks of caries bacteria, has a massive effect on lifetime and function of teeth. Therefore, strategies have been developed to at least partially re-mineralize the enamel and dentin as the robust alternative to mechanical removal and conventional filling the defect. Optimized dental hygiene has reached

restricted success in small caries lesions. Researchers all over the world use slices of dentin to explore the re-mineralization capacity of a variety of nanotechnology-based ceramics particles. Using a set of slices, a direct comparison of the different materials and procedures has become possible.

Future Research

The dental nano-biomaterials have to be investigated to judge the toxicity and the biocompatibility and to pursue the personalized medicine with respect to more and more frequent material incompatibilities. The development of nanocontainers for a variety of drugs aims for targeted delivery preventing undesired side effects. For larger caries regions, which grow fast into the dentin, the collagen network might be still available, but the mineral phase is already dissolved. The return of the calcium phosphates and traces of other minerals into the affected region for restoration purposes is challenging but not impossible, as the collagen network could provide nucleation sites for the re-mineralization.

Bio-inspired dental fillings and inlays have to exhibit the anisotropy and orientation of organic and inorganic nanometer-sized components as found in enamel and dentin. Today, the dentists already use different fillings for dentin and enamel but without oriented anisotropic ceramics. Therefore, scientists and engineers have to discover opportunities to incorporate anisotropic nanometer-sized ceramics into the fillers and to carry them into the desired direction. Hence nature-analogue mechanical properties should be obtained. Based on these oriented nanostructures, one should search for self-repairing materials. Here again the human body provides concepts: the remodeling of bone.

Cross-References

- ▶ [AFM in Liquids](#)
- ▶ [Atomic Force Microscopy](#)
- ▶ [Basic MEMS Actuators](#)
- ▶ [Bioadhesion](#)
- ▶ [Bioinspired Synthesis of Nanomaterials](#)
- ▶ [Biomimetics](#)
- ▶ [Biosensors](#)
- ▶ [Chitosan Nanoparticles](#)
- ▶ [Confocal Laser Scanning Microscopy](#)
- ▶ [Electron Microscopy of Interactions Between Engineered Nanomaterials and Cells](#)
- ▶ [Imaging Human Body Down to Molecular Level](#)
- ▶ [In Vivo Toxicity of Titanium Dioxide and Gold Nanoparticles](#)
- ▶ [Microfabricated Probe Technology](#)
- ▶ [Nanomechanical Properties of Nanostructures](#)
- ▶ [Nanomedicine](#)
- ▶ [Nanoparticles](#)
- ▶ [Nanostructured Functionalized Surfaces](#)
- ▶ [Scanning Electron Microscopy](#)
- ▶ [Selected Synchrotron Radiation Techniques](#)
- ▶ [Self-repairing Materials](#)

References

1. Dosch, H., Van de Voorde, M.H. (eds.): Genesys White Paper: A new European partnership between nanomaterials science & nanotechnology and synchrotron radiation and neutron facilities. Max-Planck-Institut für Metallforschung, Stuttgart (2009)
2. Freitas, R.A.: Nanodentistry. *J. Am. Dent. Assoc.* **131**, 1559–1565 (2000)
3. Williams, D.F.: *Williams Dictionary of Biomaterials*. Liverpool University Press, Liverpool (1999)
4. ISO 10993: Biological evaluation of medical devices, Institute ANS. www.iso.org (2009)
5. Müller, B.: Natural formation of nanostructures: from fundamentals in metal heteroepitaxy to applications in optics and biomaterials sciences. *Surf. Rev. Lett.* **8**, 169–228 (2001)
6. Bar-Cohen, Y. (ed.): *Biomimetics: Nature-Based Innovation*. CRC Press/Taylor & Francis Group, Abingdon (2011)
7. Müller, B., Deyhle, H., Bradley, D., Farquharson, M., Schulz, G., Müller-Gerbl, M., et al.: Scanning x-ray scattering: Evaluating the nanostructure of human tissues. *Eur. J. Clin. Nanomed* **3**, 30–33 (2010)

Nanodevices

- ▶ [Nanostructures for Energy](#)

Nano-ecotoxicology

- ▶ [Ecotoxicity of Inorganic Nanoparticles: From Unicellular Organisms to Invertebrates](#)

Nano-effects

- ▶ [Physicochemical Properties of Nanoparticles in Relation with Toxicity](#)

Nanoelectrodes

- ▶ [Dielectrophoretic Nanoassembly of Nanotubes onto Nanoelectrodes](#)

Nanoencapsulation

Mary Cano-Sarabia and Daniel Maspoch
CIN2 (ICN-CSIC), Catalan Institute of
Nanotechnology, Bellaterra, Barcelona, Spain

Synonyms

[Molecular encapsulation](#)

Definition

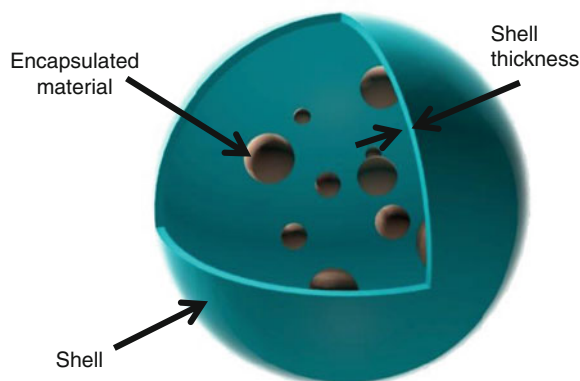
Nanoencapsulation is defined as the entrapping of active ingredients in nanometer-sized capsules.

Overview

Nanoencapsulation is defined as the technology of packaging nanoparticles of solid, liquid, or gas, also known as the core or active, within a secondary material, named as the matrix or shell, to form nanocapsules (see Fig. 1) [1]. The core contains the active ingredient (e.g., drugs, perfumes, biocides, vitamins, etc., see Table 1), while the shell isolates and protects the core from the surrounding environment. This protection can be permanent or temporal, in which case the core is generally released by diffusion or in response to a trigger, such as shear, pH, or enzyme action, thus enabling their controlled and timed delivery to a targeted site [2, 3].

Nanocapsules may range from 1 to 1,000 nm in size and they have a multitude of different shapes, depending on the materials and methods used to prepare them [2]. The structure of encapsulated ingredients (see Table 1), which largely depends on the selected shell material and nanoencapsulation method, can be classified into two main categories (see Fig. 2): capsules with (a) a core that is surrounded by a shell of the matrix material or (b) a core that is entrapped within a continuous network of the matrix material. Variations of these morphologies include capsules with multiple cores or multilayered capsules [1]. However, the most significant feature of nanocapsules is their nanoscopic size that provides a large surface area. The total surface area is inversely proportional to the capsule diameter. This large surface area is appropriate for incorporating recognition species (functionalization with peptides, antibodies, organic polymers, etc.), sites of adsorption and desorption, chemical reactions, and light scattering, among others.

Many different materials can be used as encapsulating matrices, which must be selected depending on the critical properties needed for each intended application (see Table 1). The majority of these carriers are proteins (gelatin and albumin), polysaccharides (dextrin, starch, gums), fats, liposomes [4], biopolymers, co-polymers (poly(lactic-co-glycolic acid)), micelles, organogels, dendrimers, solid nanoparticles (SLN), polymeric nanoparticles, emulsion-based systems, and metal-organic particles [5].



Nanoencapsulation, Fig. 1 Scheme of a nanocapsule

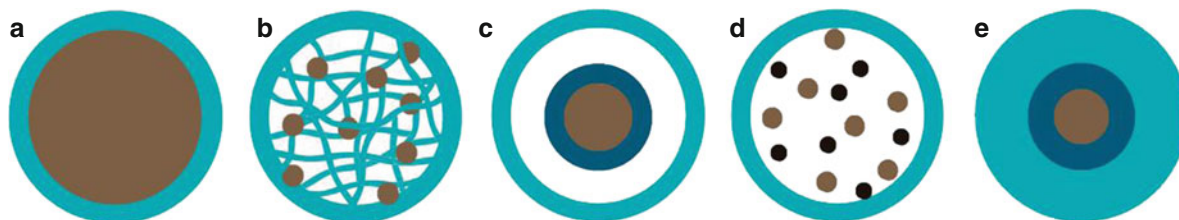
The main reason for nanoencapsulating species is to ensure that the encapsulated material reaches the area of action without being adversely affected by the external environment through which it passes. Thus, the principal reasons for encapsulation are: (a) separation of incompatible materials, (b) conversion of liquids to free-flowing solids [6], (c) increased stability (protection of the encapsulated against oxidation or deactivation due to reaction in the environment), (d) masking organoleptic properties like color, taste, and odor of substances, (e) safe handling of toxic materials, and (f) controlled and targeted release of encapsulated active compounds. In addition, nanoencapsulation allows to modify the release of encapsulated active materials, providing sustained release (maintaining the right concentration), long lasting release (and therefore improving effects), targeted release (improving adhesion, penetration, or recognition of tissues and cells), or triggered release (mainly by environmental changes in pH, temperature). Encapsulation and release modification (sustained, triggered, or targeted release) also reduces doses, and therefore, potential toxicity of drugs. This is of the main interest in medical applications. Moreover, since the encapsulated molecule is not in direct contact with the environment, irritation at the site of administration is commonly reduced.

Because of this wide range of advantages and properties, nanoencapsulation technologies have applications in many industrial sectors, including the medical, pharmaceutical, cosmetics, chemical, textile, construction, agricultural, and food industries. In fact,

Nanoencapsulation, Table 1 Core and shell materials commonly used in different nanoencapsulation applications

Application	Materials
Food	C: acidulants (acetic, sorbic, lactic), flavoring agents (citrus/mint oils, oleoresin, oregan, menthol), sweeteners (sugar, aspartame), colorants (β -carotene), lipids (fish oil, linoleic oil, palmitic acid), vitamins and minerals (vitamin A, D, E and K, folic acid), salts, preservatives and enzymes and microorganisms (lipase, invertase, penicillium roqueforti) S: carbohydrates (sucrose, malto- and cyclo-dextrines, chitosan), gums (agar, arabic, gum acacia, sodium alginate), lipids (oils, paraffin, stearic acid, fats, beeswax, phospholipids, Stealth [®] liposomes), celluloses, proteins (albumin, casein, gelatin, gluten, peptides), synthetic elastomers (polyacrylamide, polyacrylate, polyethylene, polyvinyl alcohol, polyvinyl acetate), synthetic polymers (acrylonitrile, polybutadiene), silicon dioxide
Pharmaceutics	C: hormones (insulin), vaccines, oncologic (doxorubicin, cisplatin) and hematology actives (anthracyclines), proteins (bovine serum albumin, globulin, myoglobin), enzymes (glucose oxidase), antibiotics (amoxicillin), drugs (felodipine, nicardipine, heparin), bacterias (lactobacillus acidophilus), and hepatocytes S: carbohydrates (chitosan, alginate, dextran), proteins (polylysine, polyglutamic, albumin, peptides), synthetic elastomers (poly(alkyl acrylate), poly(ethylene glycol)), celluloses (ethyl cellulose), synthetic polymers (poly(lactide-co-glycolide), polyethylene glycol), lipids (phospholipids), carbon nanotubes, erythrocytes, gums (alginate), silicon dioxide, phyllosilicate
Agriculture	C: Insecticides (endosulfan), herbicides (fenamiphos), pheromones, fertilizers, pesticides, microbicides S: Polyurea, polyurethane, gelatin, gum arabic, chitin, chitosan
Textile	C: PCM, emollients (aloe vera, vitamin E, lanolin), dyes, fabrics softeners, flame retardants S: Poly(styrene), urea-formaldehyde, ethyl cellulose
Paintings	C: Biocide (OIT), pigments, anti-foulant agents, anticorrosion agents S: Melamine formaldehyde, gelatine – acacia, silica, zeolite
Personal care	C: natural ingredients, essential oils (menthol, lime, lavender), preservatives, vitamins (A, E, C), retinol, and triclosan S: carbohydrates (sucrose, malto- and cyclo-dextrines, chitosan), cellulose, silicon dioxide

C core, S shell



Nanoencapsulation, Fig. 2 Morphologies of nanocapsules (from left to right): (a) single-core capsule, (b) dispersed core in polymer gel, (c) multilayer capsule, (d) dual-core capsule, and (e) single-core-multi-shell capsule

this technique is already commonplace within a range of industries, but it is accepted that only around 10% of potential applications are currently being exploited. For example, in the case of medicine delivery, the core material (e.g., drugs and vaccines) can be targeted and released in a controlled manner and at a specific location (see Fig. 3). In agrochemistry, the nanocapsules were proposed for controlling the release of insecticides or pesticides, and improving their efficiency and decreasing the quantity of active substances for a given application; or, in the textile industry,

lasting textile fragrances and skin softeners in the form of nanocapsules have been developed [6].

Basic Nanoencapsulation Techniques

Before considering the fabrication of nanoencapsulated products, one should first determine which are all the characteristics, properties, and capabilities that such products must satisfy for each intended application (Table 2). For this, the following questions must be

Nanoencapsulation,

Fig. 3 (a) Nanoliposomes loaded with the chemotherapeutic agent doxorubicin in commercial product Doxil[®] (Ortho Biotech) and (b) SEM image of doxorubicin liposomal in Nuclisomes[®] formulation (Image taken from reference [20])

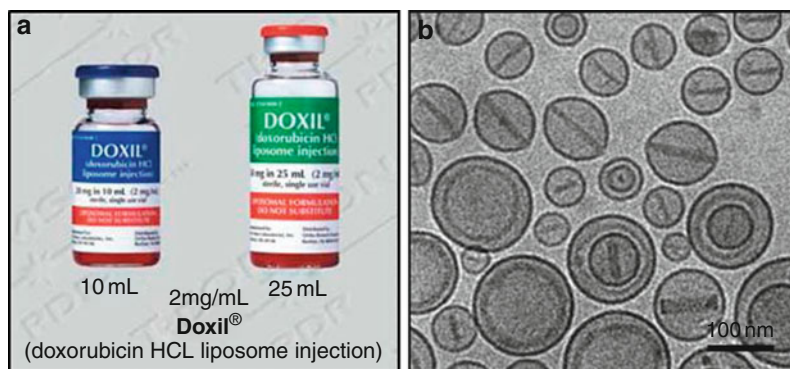
**Nanoencapsulation,**

Table 2 Several nanoencapsulation techniques and the steps involved in each process

Nanoencapsulation technique	Process involved
Sol-gel	a. Solution of core and polymer b. Formation of sol phase c. Gelation d. Solidification
In situ polymerization	a. Preparation of core solution b. Addition of droplets of monomer
Coacervation	a. Formation of a three-immiscible chemical phase b. Deposition of the coating c. Solidification of the coating
Rapid expansion of supercritical solution	a. Preparation of solution of core and shell materials in CO ₂ b. Depressurization through a nozzle
Liposome entrapment	a. Microfluidization b. Ultrasonication c. Reverse-phase evaporation
Inclusion complexes	Preparation of complexes by mixing or grinding
Spray drying	a. Preparation of a dispersion b. Homogenization of the dispersion c. Atomization of the dispersion
Solvent evaporation	a. Preparation of solution of polymer and core b. Solvent evaporation by heating
Electrocoextrusion	a. Preparation of core solution and wall solution b. Simultaneous spraying of two solutions from two coaxial capillaries

taken into consideration before choosing which nanoencapsulation methodology is the more appropriate [2]:

1. What functionality should the encapsulated ingredients (core) provide to the final product?
2. What kind of coating material (matrix) should be selected?
3. What processing conditions must the encapsulated ingredients survive before releasing its content?
4. What is the optimal concentration of the active ingredient in the nanocapsule?

5. By what mechanism should the ingredient be released from the nanocapsules?
6. What are the particle size, density, and stability requirements for the encapsulated ingredient?
7. What are the cost constraints of the encapsulated ingredient?

Similarly, an understanding of the physicochemical properties of the core and factors that control the interfacial and aggregation behavior of the matrix materials is crucial to choose suitable processes for nanoencapsulation [1]. Today, a multitude of techniques are used

for nanoencapsulating active species. The more popular techniques can be classified in three main families: chemical, physicochemical, and physicomechanical encapsulation processes [3]. The chemical and physicochemical processes include those methods in which chemical reactions are involved in the nanocapsule formation. In contrast, in the physicomechanical methods, no chemical reactions are involved in the nanocapsule formation, and only shape fabrication usually takes place. Some of the techniques that are commonly used are briefly discussed below.

Chemical Processes

The chemical processes are building-up techniques where the nanocapsules are synthesized through the nucleation and growth processes involving elementary building blocks, such as ions, atoms, and molecules. The most important chemical methods include the emulsions, suspensions, precipitations, sol-gels, and polymerizations. The advantages of these chemical methods are that they provide nanocapsules with high purity, high uniformity, small particle size, narrow size distribution, dispersibility, good chemical homogeneity, and more reactivity.

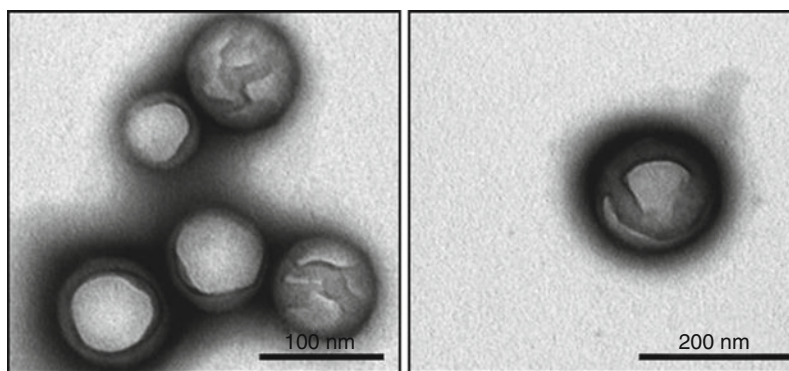
The majority of chemical methods are based on the initial formation of an emulsion. An emulsion is a suspension of small droplets (diameters below 1,000 nm), usually stabilized by an emulsifier, of one liquid in a second liquid in which the first will not mix. These emulsions are conventional emulsions, multi-layered emulsions, microemulsions, nanoemulsions, solid lipid emulsions, or double emulsions. Generally, the “emulsification” methods consist on first creating an emulsion of a liquid core in a continuous phase using shaking, stirring, homogenizing, or spray processes. In a second step, a shell is formed via organic polymerization, or the emulsion homogenized by microfluidizers, or the suspension atomized for solvent evaporation. Control over the properties of emulsion-based nanoencapsulation systems can be achieved by tailoring the dispersed phase as well as the structure of emulsions [1]. In general, the so-called microemulsions are defined as thermodynamically stable transparent or translucent isotropic dispersions of nanodroplets with a size ranging between 5 and 1,000 nm, whereas nanoemulsions are metastable dispersions of nanoscale droplets (<100 nm) and optically transparent. This last property is of interest, for example, to the food and beverage industry as it

enables the delivery of flavors and bioactive ingredients in clear emulsions. For example, kinetic stable nanoemulsions of triglyceride oils may be achieved using high-pressure homogenizers as the final step [7].

Emulsion encapsulation methods include the “in situ polymerization” and the “interfacial polymerization” methodologies, among others. In the “in situ polymerization,” the monomer is added dropwise to an emulsion, which is formed by the aqueous polymerization solution, the active ingredient, and an emulsifier, under continuous stirring. Initially, a low molecular weight polymer is formed in the continuous phase, and as the polymerization proceeds, the polymer gradually grows and simultaneously entraps the core material to form the final nanocapsules. For example, insulin-loaded poly(alkyl cyanoacrylate) nanocapsules are prepared using this technique [8]. On the other hand, in the “interfacial polymerization” technique, the capsule shell is formed at/on the droplet surface by polymerization of the reactive monomers. First, the core material is dispersed into the continuous phase where the monomer is dissolved, forming an emulsion. A co-reactant is then added to the mixture, provoking the polymerization at the core interface, thus generating the capsule shell. As an illustrative example, this method is used for the encapsulation of polyaminoacids in PLGA-polaxamer nanocapsules (see Fig. 4) [9].

The “sol-gel” method is a process based on the preparation of a solution, sol and gel, followed by the solidification and heat treatment of the organic and inorganic compounds. “Sol-gel” encapsulation involves the evolution of inorganic networks through the formation of a colloidal suspension (sol) and gelation of the sol to form a network in a continuous liquid phase (gel). In this method, the capsules can range between submicron sizes to a few tens of microns [10]. The most popular use of this process is the production of immobilized biomaterials, such as enzymes and microorganisms. Its main advantages are the very mild operating conditions as well as the small particle size achievable and the affordable raw materials. From an industrial point of view, this process is easily scaled up even considering the tight cost-in-use constraint characteristics of the food industry [11]. For example, this method can be used for the immobilization of lipase enzymes within a phyllosilicate and vinyltriethoxysilane sol-gel matrices and beta-glucosidase in alginate-silicate sol-gel matrices.

Nanoencapsulation,
Fig. 4 TEM image of
a polyaminoacid loaded in
PLGA: poloxamer
nanocapsules (Image taken
from reference [9])



In addition to the chemical nanoencapsulation techniques described above, nanoencapsulation can also be carried out by “suspension cross-linking.” This is the method of choice for encapsulating proteins and polysaccharides. This method involves the dispersion of a polymer aqueous solution containing the core material in an immiscible organic solvent to form an emulsion. Droplets are hardened by covalent cross-linking and directly converted to capsules. The cross-linking process is accomplished either thermally or by the use of cross-linking agents. Albumin nanocapsules containing doxorubicin and magnetite particles are prepared by this technique. Salmon calcitonin-loaded chitosan-PEG nanocapsules are prepared with this methodology as well.

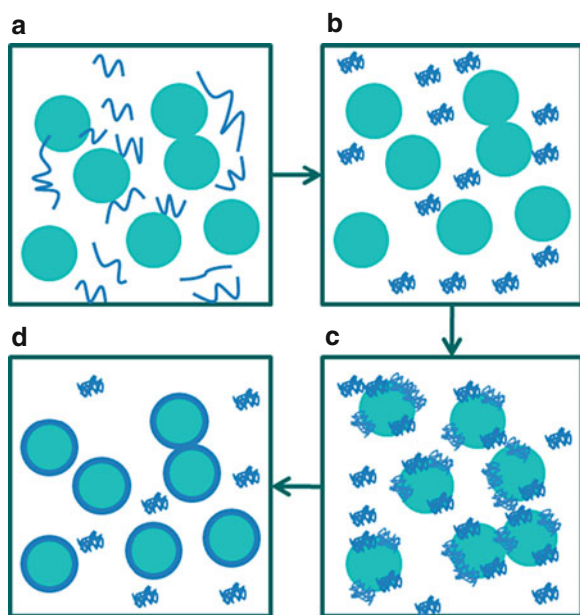
Finally, nanoencapsulation can also be carried out via “hydrothermal methods” where chemical reactions in water at high temperatures and high pressures are involved for the nanocapsule formation, and via “oxidation processes” that oxidize and deoxidize directly the raw materials in the liquid phase or quasi-liquid phase state to prepare the nanocapsules.

Physicochemical Processes

Physicochemical methods are the processes that combine both the chemical and physical methods during the nanoencapsulation. For example, “phase inversion nanoencapsulation” (PIN) is a method for highly efficient encapsulation of drugs, proteins, or plasmid DNA into biodegradable polymer nanospheres. Capsules are produced with a limited size distribution, ranging from 100 nm to 10 μm in diameter, and labile drugs are efficiently encapsulated without denaturation. In this method, the active compound is added to a dilute polymer solution (normally in methylene chloride), which is then poured rapidly into an unstirred bath of non-

solvent (petroleum) at a solvent to non-solvent ratio of 1:100, causing nano- and microspheres to form spontaneously. The drug dicumarol, proteins such as bovine serum albumin, hormones such as insulin and plasmid DNA have been successfully encapsulated using the PIN system in combination with biologically erodible, biologically adhesive polymers like fumaric acid, sebacic acid, poly-lactide-co-glycolide (PLGA), polylactic (PLA) acid, and polystyrene.

“Coacervation and phase separation” is another physicochemical method based on the gradual desolvation of a homogeneous polymer solution into a polymer-rich phase (coacervate) and the poor polymer phase (coacervation medium). Phase separation processes are divided into simple and complex coacervation. The mechanism of capsule formation is identical in both processes, except in the way in which the phase separation is carried out (see Fig. 5). In a simple coacervation, a desolvation agent, such as gelatin or ethyl cellulose, is added to the aqueous or organic media for phase separation, whereas complex coacervation involves complexation between two oppositely charged polymers, both soluble in an aqueous media. Nanoencapsulation by coacervation is carried out by preparing an aqueous polymer solution (1–10%) at 40–50°C where the hydrophobic core material is also dispersed/dissolved. Stabilizers are usually added to the mixture to keep the individuality of the final particles. Then, a suitable desolvating agent is gradually added to the mixture leading to the formation of partially desolvated polymer molecules, and hence their precipitation on the surface of the core material. Finally, the prepared capsules are stabilized by cross-linking, desolvation, or thermal treatment (e.g., cooling at 5°C). An illustrative example is that of the gelatin nanocapsules (840 nm) loaded with the



Nanoencapsulation, Fig. 5 Schematic illustration of a typical coacervation process, involving (a) dispersion of the core material in a solution of the shell polymer, (b) initial separation of the coacervate from the solution, (c) coating of the core material by microdroplets of the coacervate, and (d) coalescence of the coacervate to form continuous shell by reticulation at the interface

protein/peptide drug bovine serum albumin (BSA), which have been produced by a modified coacervation method free of desolvating agents.

Another common physicochemical approach is the use of molecular systems that self-assemble into capsules at given conditions, such as the aggregation of lipid molecules into spherically closed bilayer structures at low concentrations, the so-called vesicles or liposomes, or molecular inclusion complexes, such as cyclodextrins and dendrimers or hyperbranched polymers. “Liposomes” are the smallest artificial vesicles of spherical shape that can be produced from natural nontoxic phospholipids and cholesterol. When phospholipids are dispersed in an aqueous phase, the liposomes are formed spontaneously, thus encapsulating either aqueous or lipid-soluble materials. In general, they are considered very versatile tools in biology, biochemistry, and nanomedicine because they are currently used for delivery of vaccines, hormones, enzymes, and vitamins (see Fig. 3) [4]. The great advantage of liposomes over other nanoencapsulation techniques is the stability that

liposomes impart to water-soluble material in high water activity applications [2]. Recently, it has been published as a very interesting nanoencapsulation technology that produced capsules coined “colloidosomes” because of their resemblance to liposomes, while using colloid particles as shell material. Also, the ultimate achievement in enzyme protection and delivery is represented by polymeric vesicles, which are called polymersomes.

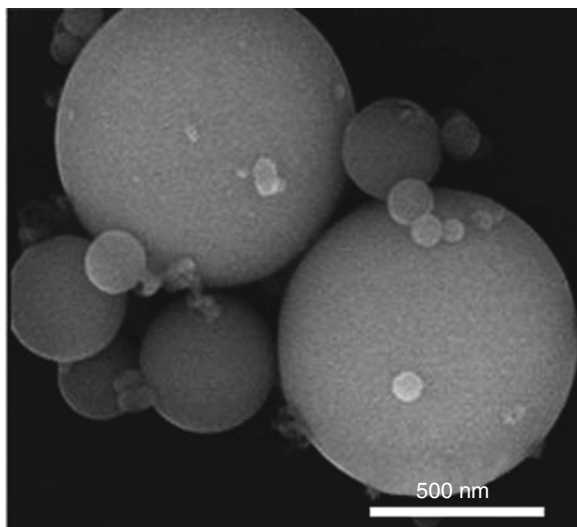
“Inclusion complexes” generally refers to the supramolecular association of the encapsulated substance into a cavity-bearing complex or dendrimer (shell). Small organic molecules can displace the water from the inner cavity and form thermodynamically stable complexes. The encapsulated unit is kept within the cavity by hydrogen bonding, van der Waals forces, or by the entropy-driven hydrophobic effect. Examples include the encapsulation of hydrophobic vitamins (A, E, or K) in cyclic oligosaccharides such as β -cyclodextrins [11].

Other physicochemical processes for nanoencapsulation are based on *supercritical fluids*. In these methods, the core ingredient is dispersed in a matrix material solubilized in a supercritical fluid (usually carbon dioxide). Removal of the carbon dioxide results in the core being encapsulated within the matrix material. For example, in *rapid expansion of supercritical solutions* (RESS), the dispersion is sprayed through a nozzle and a particulate material containing the core is formed [11]. Titanium dioxide and silica nanoparticles coated with poly(lactide-co-glycolide) are prepared using the RESS process [12] (Fig. 6).

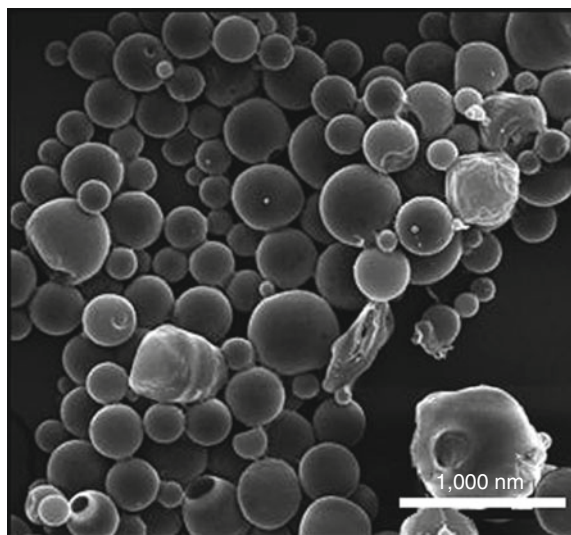
In addition to the nanoencapsulation techniques described above, other physicochemical processes are “solid lipid nanoparticles” (SLN), “layer-by-layer deposition” (e.g., encapsulation by polyelectrolyte multilayer), and “controlled precipitation.”

Physicomechanical Processes

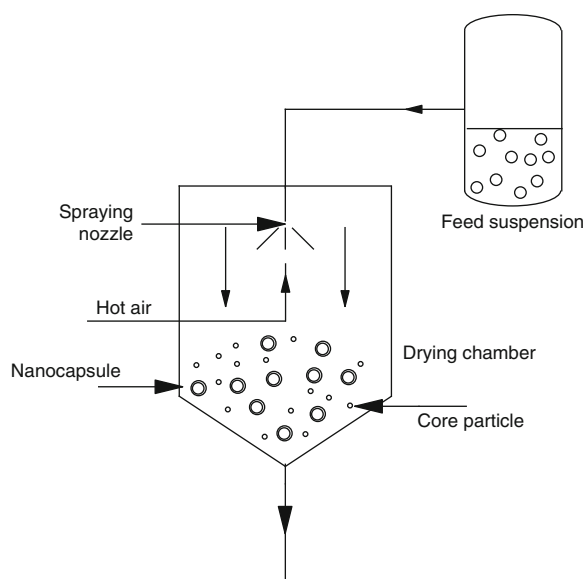
The most important physicomechanical methodologies for encapsulating desired species at the nanometer scale are: (1) spray drying, (2) solvent evaporation/solvent extraction, and (3) electroencapsulation-based techniques. Among them, “spray drying” is a process widely used in the encapsulation of fragrances, oils, and flavors. In a “spray drying process,” core particles are dispersed in a polymer solution of wall material and atomized into a hot chamber (see Fig. 7). As the solvent evaporates, the shell material solidifies onto



Nanoencapsulation, Fig. 6 SEM image of PLGA-encapsulated SiO₂ powder prepared by rapid expansion of supercritical suspension process (Image taken from reference [12])



Nanoencapsulation, Fig. 8 SEM image of oxaprozin/methylated- β -cyclodextrin-loaded PLGA nanocapsules (Image taken from reference [15])

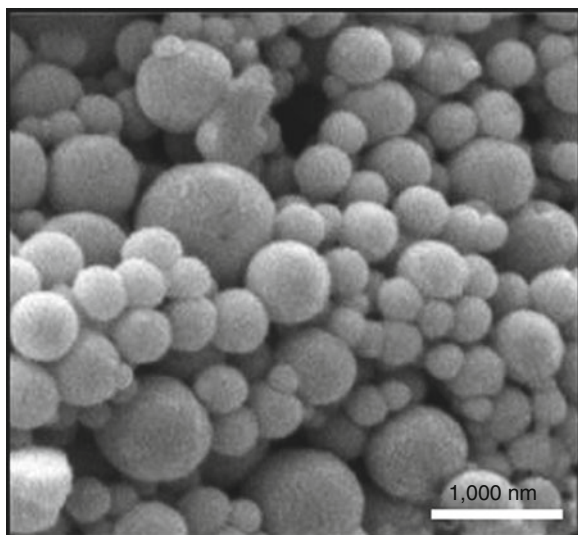


Nanoencapsulation, Fig. 7 Schematic representation of the spray drying process

the core particles, obtaining polycore or matrix-type capsules [13]. The one-droplet-to-one-particle mechanism during the spray drying sets the lower limit of the particle size to 100 nm. For example, this technique is used for preparing 260 nm in diameter amifostine/PLGA nanocapsules.

“Solvent evaporation/solvent extraction” is a method based on the dissolution of a polymer in a water-immiscible organic solvent, such as dichloromethane or chloroform, where the core material is dissolved or dispersed as well. The mixture is then added dropwise to a stirring aqueous solution to form small polymer droplets containing the encapsulated material, and then, the solvent is eliminated. Here, the coat material shrinks around the core material and encapsulates it. The solvent elimination can be done by simple solvent evaporation (heat or reduced pressure) or by solvent extraction with a third liquid, which is precipitant for the polymer and miscible with both water and the solvent [14]. This method is suitable for the preparation of fluorescein-loaded nanocapsules based on PLGA using droplet-based microfluidic platforms. PLGA nanocapsules loaded with naproxen and oxaprozin-cyclodextrin are also prepared using this methodology (see Figs. 8 and 9) [15]. In some cases, the “solvent evaporation/solvent extraction” method is also combined with the “freeze drying” process, for example, in the encapsulation of hypericin in polylactic acid.

Recently, pharmaceuticals and cosmetic or food industries seek for new, more efficient technologies for nanocapsule production, which do not degrade the chemical or biological properties of the material.



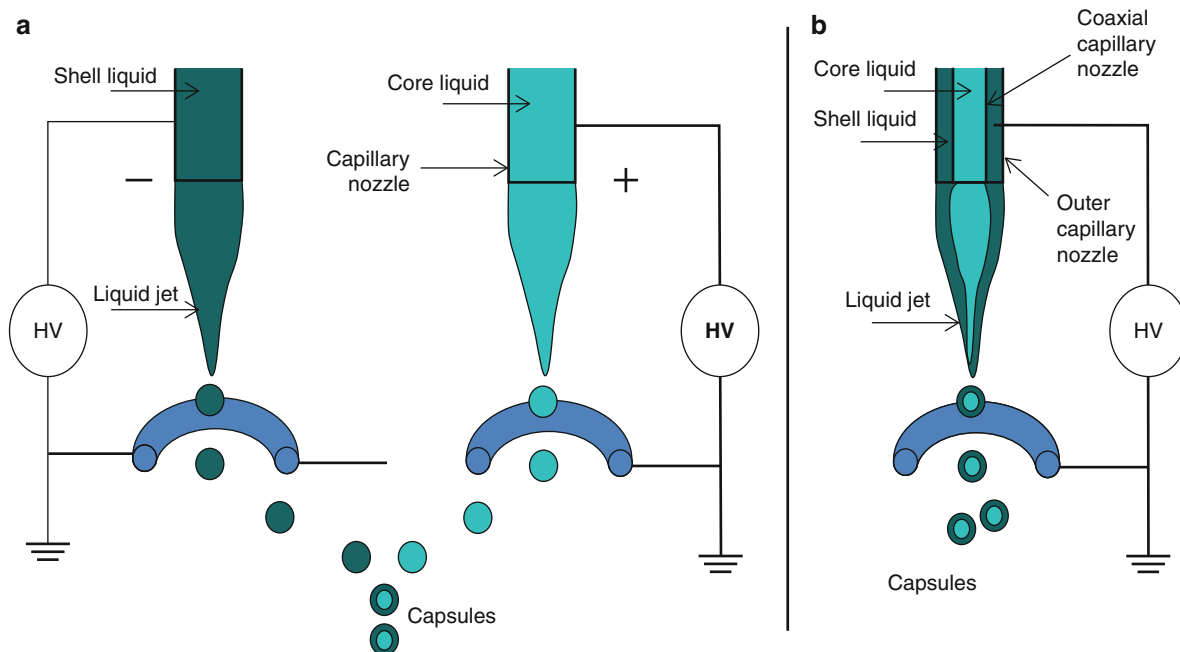
Nanoencapsulation, Fig. 9 SEM image of naproxen-PLAG nanocapsules prepared by solvent evaporation/extraction technique (Image taken from reference [14])

Electroencapsulation-based techniques are excellent candidates to be some of these technologies. “Electrospraying” is a straightforward bottom-up method to manufacture capsules of submicrometer diameters. It is a method of liquid atomization by means of electrical forces. In “electrospraying,” the liquid flowing out of a capillary nozzle is subjected to an electrical shear stress by maintaining the nozzle at high electric potential, which forces the liquid to be dispersed into fine droplets. The nanocapsules are produced by previously dissolving the polymer and the active compound into the solvent(s). With this method, the droplet sizes can range from hundreds of micrometers down to several tens of nanometers, and they can be controlled by electrical means, that is, by adjusting the flow rate and voltage applied to the nozzle [16]. Recently, research in electronanoencapsulation is aimed at developing new drug-delivery systems, medicine production, and ingredient dosages in the cosmetic and food industries, since the utilization of electrical forces can increase the effectiveness of nanoencapsulation. As a consequence, some “electrospraying” techniques have recently been described. For example, “impacting of two oppositely charged droplets” is a method where both the droplets’ streams are emitted from two separated capillary nozzles maintained at opposite potentials, one of the capillaries at the positive and the other at the negative

(see Fig. 10a). The droplets collide due to Coulomb attraction, forming a capsule via submerging the droplet of higher surface tension within that of the smaller surface tension. Another method is the “electrospraying/evaporation of colloidal suspension,” in which a suspension is initially electrosprayed followed by the shell solidification by solvent evaporation. Similarly, in the “electrospraying/gelatinization of colloidal suspension” method, a suspension of a core material is electrosprayed into a bath containing a gelatinizing or polymerizing agent. This agent forms a hard envelope on the core material. Among other actives, electrospraying has been tested for the encapsulation of the drug paclitaxel in PLGA and polycaprolactone. Recently, nanocapsules composed of natural polymers (e.g., chitosan) for ampicillin and elastin-like polypeptides for doxorubicin have also been fabricated using the “electrospraying” methods [13].

The “electrocoextrusion” process is another electronanoencapsulation technique based on the simultaneous spraying of two different liquids from two coaxial capillaries that are maintained at the same potential (see Fig. 10b). In this method, the core liquid flows from the central capillary, and the envelope liquid flows through the annular nozzle between the capillaries. It is important to note that the core liquid must have a high resistivity and the envelope sufficiently high conductivity in order to be used in this methodology [16]. At the present time, the smallest capsules reported using this methodology have a diameter of 150 nm, and they were produced using a photopolymer (Somos 6120-DuPont) and ethylene glycol.

Besides the formation of spherical nanocapsules, electronanoencapsulation technologies can also allow the encapsulation into nanofibers. Indeed, “electrospinning” is a method to produce nanofibers where a suspension of the core material in a polymer solution is supplied from a spinneret and forms a droplet at the spinneret exit. In this methodology, an electric field is applied to the suspension by immersing an electrode in it and placing the counter-electrode some distance from the spinneret, and then jetting sets in. This field provokes an electrically induced bending instability in the solution jet, which causes stretching of the bent sections of the jet. The solvent eventually evaporates, the jet dries and solidifies, and the spun nanofibers are deposited on the counter-electrode.



Nanoencapsulation, Fig. 10 Schematic representation of two electroencapsulation processes: (a) electrospraying of two oppositely charged droplets and (b) electrocoextrusion

The nanofibers are collected onto the edge of a grounded collector disk. This method has been used in biotechnology, that is, for the capturing of living cells (virus and bacteria) in poly(vinyl alcohol) nanofibers.

In addition to the techniques described above, other physicomachanical methods that can be used for encapsulating desired species at the nanoscale are the “melt-solidification,” “centrifugal suspension separation,” “vibrating and coaxial nozzle,” and “pan-coating” processes.

Release Mechanisms

Different release mechanisms of encapsulated materials provide controlled, sustained, and targeted release of core materials. There are three main mechanisms by which encapsulated material is delivered from a capsule: (1) mechanical rupture of the capsule wall, (2) dissolution (impact/shear/pressure) or melting of the wall, and (3) diffusion through the wall. The desired mechanism depends on the nature of the application. For example, carbonless copy paper, and scratch and sniff perfumes rely on mechanical rupture

of shell to release the core contents. In these applications, the rupture may be caused by pressure or due to propagation of cracks. In contrast, the detergent industry uses dissolution of shell wall of powder detergents for the release of encapsulated protease enzyme in order to remove bloodstains from clothing. Similarly, in the pharmaceutical industry, nanoencapsulated products are designed for sustained and controlled release by either degradation or diffusion through the shell. For example, doxorubicin is encapsulated in liposomes to control its release into the body reducing its cardiotoxicity [17].

Applications

Nanoencapsulation is believed by many scientists and technological fields to be one of the most important scientific topics of the next decade with a huge impact on medicine, pharmacology, construction, cosmetics, chemistry, textile, agriculture, and food industry, among others. Nowadays, nanoencapsulation is being used in a diverse range of markets, including, but not limited to agrochemicals, aromatherapy, paints, coatings, colorants, pharmaceuticals, personal care,

Nanoencapsulation, Table 3 Examples of nanoencapsulated materials in the market

Product	Company	Industry
Caelyx [®]	Sequus Pharma	Pharmaceutical
Myocet [®]	Cephalon Limited	
LBL-Intra [®]	Capsulation	
PharmFilm [®]	MonoSol Rx	
Doxil [®]	Alza Corp.	
Lipozyme IM-60	NNBN America Inc.	
AmBisome [®]	Gilead	
MicroMatrix [™]	ABN	Food
VitaDHA [™]	Blue California	
Pharmasmooth [™]	Balchem Corp.	
Cavamax [®] , Omegadry [®]	Wacker Chemie AG	
Plenitude [®]	L'Oréal	Personal care
Primasys [®]	Cognis	
Tinoderm [®]	BASF	
Cutanova [®]	Dr.Rimpler GmbH	

adhesives, biocides, nutraceuticals, oil and gas, textiles, paper systems, electronics, and imaging systems (see Table 3). Depending on the application, the nanoencapsulated product can be delivered as a dry, free-flowing powder, as slurry, or in the form of a wet filter cake. A brief explanation of some applications in some of the most powerful sectors is given below.

Agriculture

Nanoencapsulation appears to be the most promising technique for short-term application in the field. One of the most important applications of nanoencapsulated products is in the area of crop protection in agriculture. In the case of controlling parasitic plants, encapsulation can be used to solve problems regarding phytotoxicity on the crop of the herbicides used against the parasite. For example, glyphosate, imidazolinones, and sulfonylureas can be encapsulated within a polymeric shell. Germination stimulants can be also delivered into the soil inside nanocapsules protecting them from degradation.

Pharmaceutics and Medicine

Nanoencapsulation currently has the highest potential in the pharmaceutical industry because it allows the possibility to deliver a drug at the right moment, in the

right place, and at an adequate concentration. One of the simplest examples is the protection of orally applied drugs from the attack of acids in the stomach before they can be adsorbed in the intestine. This problem is solved in many cases by the encapsulation of the active ingredient in chitosan and PLGA (poly (lactic-co-glycolic acid) nanocapsules. In addition, both polymers are widely used in many aspects of biomedical research ranging from tissue engineering to drug delivery because they are highly biocompatible, easily available, and provide excellent delivery capabilities. For example, ascorbyl palmitate can be encapsulated in a chitosan derivative with an efficiency of 84% at 56% drug loading. The nanocapsules display no short-term cytotoxicity against the human skin melanoma. On the other hand, nanoencapsulation is also used for cancer therapy. Indeed, two nanoencapsulated products named as Doxil[®] and Abraxane[®] have recently been approved by the European Medicines Agency (EMA). Doxil[®] is a liposomal system for doxorubicin delivery and treatment of ovarian carcinoma (see Fig. 3) and Abraxane[®] is an albumin nanoparticle Taxol conjugate for the treatment of metastatic breast cancer.

Food Industry

Today's food industry makes use of nanoencapsulation to develop "functional foods" with ingredients that improve the nutritional content of food without affecting the taste, aroma, or texture of food, and enhance the shelf life and stability of the ingredient and the finished food product [1]. Nanoencapsulation is used to protect sensitive food ingredients (i.e., flavors, oils, vitamins) as well as to provide viable texture blending, appealing aroma release, and taste, odor, and color masking. This technology also enables food companies to incorporate minerals, vitamins, flavors, and essential oils [7]. In addition, nanoencapsulation can simplify the food manufacturing process by converting liquids to solid powder, reducing costs allowing low-cost handling equipment. Moreover, nanocapsules also help fragile and sensitive materials (oils, vitamins, enzymes) to survive processing and packaging conditions, and stabilize the shelf life of the active compound [18]. For example, there are several commercial nanoencapsulated products in the market that offer

a specific supplement of Omega-3 DHA and of a natural form of folate for pregnant and lactating women, such as VitaDHA™ Materna.

Personal Care

Nowadays nanoencapsulation technologies play an important and growing role in the health and beauty industry. They offer an ideal and unique carrier system for active ingredients aside from an effective protection against photodegradation. Their striking effects can enhance the tactile and visual appearance of a variety of cosmetic and personal care products, including bath gels, lotions, tanning, makeup, aromatherapy, etc. Furthermore, nanoencapsulation can make significant savings for formulators as it can reduce the amount of active ingredients needed. For example, the retinyl palmitate oil, which is the most stable form of vitamin A and plays an important role in cellular differentiation and carcinogenesis prevention, has been recently prepared as a novel nanocapsule antiaging formulation. Flexible polymeric nanocapsules with retinyl palmitate core and PLA shell as a carrier of drugs or simply as a reservoir of vitamin A were prepared. The main advantage of these capsules is that they can easily pass through biological barriers due to deformability characteristics, reaching deeper layers of the skin and allowing a uniform permeation in the skin [19]. Another example are liposomes as sources of magnesium ascorbyl phosphate, a stable derivative of vitamin C, which allow higher effects of vitamin, such as whitening and antioxidants thanks to the targeted activity provided by the liposome.

Other Applications

Besides many classical applications, nanoencapsulation technologies provide a large number of other applications. One of them is the preparation of high surface area carbon nanocapsules (100 nm) with hollow cores and porous shells. Such mesoporous carbon nanocapsules are of extensive interest for catalysis, separation, gas storage, and fuel cells, among others applications. Another novel application of nanoencapsulation is the preparation of carbon nanocapsules that serve as heat transfer fluid for heat dissipation systems.

Future Directions for Research

The nanoencapsulation of active compounds for controlled release applications is a promising alternative to solve some of the major problems of active ingredients faced by industries. Today, a wide range of encapsulated products have been developed, manufactured, and successfully marketed in the pharmaceutical, medical, and cosmetic industries. However, nanoencapsulation has still found a small market in many other industries, such as food, construction, etc. In these potential fields, nanoencapsulation is still far from being fully developed since the development time is rather long and requires multidisciplinary cooperation. For this reason, great advances in nanoencapsulation are highly expected in the near future. These advances will include the development of novel materials to be used as matrices, more stable, efficient, and biocompatible nanocapsules and advanced nanoencapsulation technologies, as well as their transferring to industrial-scale mass production. Such improvements and novel capabilities will expand the scope of application for nanoencapsulation technologies in many industries, where such innovations will certainly open new opportunities to develop a novel generation of products.

Cross-References

- ▶ [Liposomes](#)
- ▶ [Nanomedicine](#)
- ▶ [Nanoparticles](#)

References

1. Augustin, M.A., Hemar, Y.: Nano- and micro-structured assemblies for encapsulation of food ingredients. *Chem. Soc. Rev.* **38**, 902–912 (2009)
2. Desai, K.G.H., Park, H.J.: Recent developments in microencapsulation of food ingredients. *Dry Technol.* **23**, 1361–1394 (2005)
3. Jyothi, N.V.N., Prasanna, P.M., Sakarkar, S.N., Prabha, K.S., Ramaiah, P.S., Srawan, G.Y.: Microencapsulation techniques, factors influencing encapsulation efficiency. *J. Microencapsul.* **27**, 187–197 (2010)
4. Ciobanu, M., Heurtault, B., Schultz, P., Ruhlmann, C., Muller, C.D., Frisch, B.: Layersome: development and optimization of stable liposomes as drug delivery system. *Int. J. Pharm.* **344**, 154–157 (2007)

5. Imaz, I., Rubio-Martinez, M., Garcia-Fernandez, L., Garcia, F., Ruiz-Molina, D., Hernando, J., Puentes, V., Maspoch, D.: Coordination polymer particles as potential drug delivery systems. *Chem. Commun.* **46**, 4737–4739 (2010)
6. Jaworek, A.: Electrostatic micro- and nanoencapsulation and electroemulsification: A brief review. *J. Microencapsul.* **25**, 443–468 (2008)
7. Wooster, T.J., Golding, M., Sanguansri, P.: Impact of oil type on nanoemulsion formation and Ostwald Ripening stability. *Langmuir* **24**, 12758–12765 (2008)
8. Dange, C., Vranckx, H., Balschmidt, P., Couvreur, P.: Poly (alkyl cyanoacrylate) nanospheres for oral administration of insulin. *J. Pharm. Sci.* **86**, 1403–1409 (1997)
9. d'Angelo, I., Parajó, Y., Horváth, A., Kéri, G., La Rotonda, M.I., Alonso, M.J.: Improved delivery of angiogenesis inhibitors from PLGA: poloxamer blend micro- and nanoparticles. *J. Microencapsul.* **27**, 57–66 (2010)
10. Ciriminna, R., Sciortino, M., Alonzo, G., Schrijver, A.D., Pagliaro, M.: From molecules to systems: Sol – Gel microencapsulation in silica-based materials. *Chem. Rev.* **111**, 765–789 (2010)
11. Gouin, S.: Microencapsulation: industrial appraisal of existing technologies and trends. *Trends Food Sci. Technol.* **15**, 330–347 (2004)
12. Kongsombut, B., Tsutsumi, A., Suankaew, N., Charinpanitkul, T.: Encapsulation of SiO₂ and TiO₂ fine powders with poly(DL-lactic-co-glycolic acid) by rapid expansion of supercritical CO₂ incorporated with ethanol cosolvent. *Ind. Eng. Chem. Res.* **48**, 11230–11235 (2009)
13. Peltonen, L., Valo, H., Kolakovic, R., Laaksonen, T., Hirvonen, J.: Electro spraying, spray drying and related techniques for production and formulation of drug nanoparticles. *Expert Opin. Drug Deliv.* **7**, 705–719 (2010)
14. Javadzadeh, Y., Ahadi, F., Davaran, S., Mohammadi, G., Sabzevari, A., Adibkia, K.: Preparation and physicochemical characterization of naproxen-PLGA nanoparticles. *Colloids Surf B Biointerfaces* **81**, 498–502 (2010)
15. Mura, P., Maestrelli, F., Cecchi, M., Bragagni, M., Almeida, A.: Development of a new delivery system consisting in 'drug-in cyclodextrin-in PLGA nanoparticles'. *J. Microencapsul.* **27**, 479–486 (2010)
16. Jaworek, A., Sobczyk, A.T.: Electro spraying route to nanotechnology: an overview. *J. Electrostat.* **66**, 197–219 (2008)
17. Bonacucina, G., Cespi, M., Misici-Falzi, M., Palmieri, G.F.: Colloidal soft matter as drug delivery system. *J. Pharm. Sci.* **98**, 1–42 (2009)
18. Chiu, Y.T., Chiu, C.P., Chien, J.T., Ho, G.H., Yang, J., Chen, B.H.: Encapsulation of lycopene extract from tomato pulp waste with gelatin and poly(γ -glutamic acid) as carrier. *J. Agric. Food Chem.* **55**, 5123–5130 (2007)
19. Teixeira, Z., Zanchetta, B., Melo, B.A.G., Oliveira, L.L., Santana, M.H.A., Paredes-Gamero, E.J., Justo, G.Z., Nader, H.B., Guterres, S.S., Durán, N.: Retinyl palmitate flexible polymeric nanocapsules: characterization and permeation studies. *Colloids Surf. B* **81**, 374–380 (2010)
20. Fondell, A., Edwards, K., Ickenstein, L., Sjöberg, S., Carlsson, J., Gedda, L.: Nuclisome: a novel concept for radionuclide therapy using targeting liposomes. *Eur. J. Nucl. Med. Mol. Imaging* **37**, 114–123 (2010)

Nano-engineered Concrete

Konstantin Sobolev¹ and Florence Sanchez²

¹Department of Civil Engineering and Mechanics, University of Wisconsin-Milwaukee, Milwaukee, WI, USA

²Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, USA

Synonyms

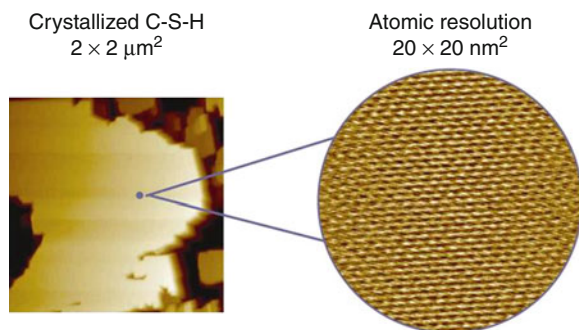
Nano-concrete

Definitions

What Is Nano-concrete?

Portland cement concrete is a complex nanostructured, multiphase, multiscale composite material that evolves over time [1, 2]. The elementary block calcium-silicate-hydrate (C-S-H), which holds the concrete composite together, is also a nanostructured material (Fig. 1) [3]. The properties of concrete exist in multiple length scales (nano to micro to macro). Processes occurring at the nanoscale ultimately affect the engineering properties and long-term performance of concrete [2].

The nanotechnology of concrete can be represented by two directions: nano-science and nano-engineering (or nanomodification) [1, 2, 4]. As defined by Sanchez and Sobolev [2] nano-science “deals with the measurement and characterization of the nano- and microscale structure of cement-based materials to better understand how this structure affects macroscale properties and performance through the use of advanced characterization techniques and atomistic or molecular level modeling.” Nano-engineering includes “the techniques of manipulation of the structure at the nanometer scale to develop a new generation of tailored, multifunctional, cementitious composites with superior mechanical performance and durability potentially having a range of novel properties such as: low electrical resistivity, self-sensing capabilities, self-cleaning, self-healing, high ductility, and self-control of cracks [2].” Concrete can be nano-engineered by the incorporation of nano-sized (less than 100 nm) building blocks or objects (e.g., nanoparticles and



Nano-engineered Concrete, Fig. 1 Nanoscale structure of C–S–H crystallized on calcite substrate and revealed by AFM (Ca/Si = 0.9). Reprinted with permission from the American Ceramic Society Bulletin, 2005, vol. 84, N 11

nanotubes) to control material behavior and add novel properties, or by the grafting of molecules onto cement particles, cement phases, aggregates, and additives (including nano-sized additives) to provide surface functionality, which can be adjusted to promote specific interfacial interactions.

Introduction

Though nanotechnology is a new emergent technology, nano-sized objects have existed for millennia, from ancient times when humans began to use nano-sized materials in glass [5], to modern times where “classic” photography employs silver nanoparticles sensitive to light. Natural materials have been shown to exhibit exceptional mechanical properties that are in some cases unmatched by man-made materials. For example, the exceptional performance of biomaterials such as mollusk shells has been attributed to the presence of calcium-based nanocrystals [6]. One of the most promising directions in nanotechnology is a better understanding of the relationship between microstructure and strengthening mechanisms employed by nature with the goal to ultimately mimic nature’s design principle and self-organization (bottom-up approach) [5].

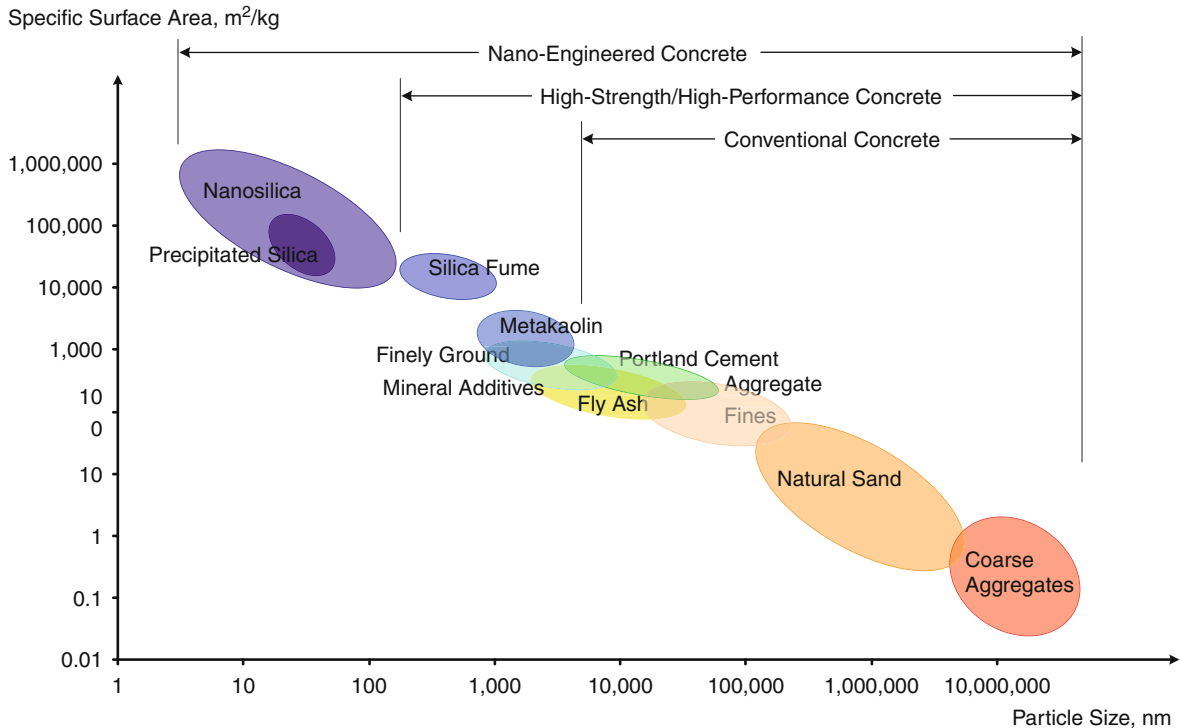
The nanoparticle is the elementary building block in nanotechnology and is comprised of up to thousands of atoms combined into a cluster of 1–100 nm. The conditions of the bottom-up production of nanoparticles allow for control of the size and shape of the particle. A reduction in size provides an exceptional surface area-to-volume ratio and changes in the surface

energy, surface chemistry, and surface morphology of the particle, altering its basic properties and reactivity [1, 2, 5, 6]. Nanoparticles have been shown to significantly enhance the mechanical performance of a variety of materials, including metals, polymers, ceramic, and concrete composites [1, 5–7].

By far, the most significant development has been the synthesis of fullerenes (C_{60}) and fullerene-like structures such as carbon nanotubes [5]. Since their discovery, carbon nanotubes have been used in numerous applications varying from superconductors, optical and electronic devices, sorbents, drug delivery systems, and nanostructured materials. The exceptional tensile strength of carbon nanotubes (as much as 20 times higher than that of steel) makes them ideal reinforcing components for structural applications, such as bridges and high-rise buildings. Polymer-based nanofibers and nanotubes [5, 8, 9] are promising nanomaterials that could open the door for new developments and innovative applications in building and construction industry.

Understanding the Structure at the Nano-level and “Bottom-Up” Design

Most research on nanotechnology in concrete has focused to date on the investigation of the structure and mechanical properties of concrete at the nanoscale [1–4, 10, 11]. Recent advances in instrumentations have made it possible to characterize the structure of concrete at the nanoscale and to measure the local mechanical properties of its micro- and nanoscopic phases [10]. Significant progress in understanding nanoscale processes in cementitious materials has been achieved thanks to the use of nanoscale characterization techniques [1, 2, 10, 11]. These advanced techniques include nuclear magnetic resonance, atomic force microscopy, micro- and nano-indentation, neutron scattering, ultrasonic force microscopy, and focus-ion beam (FIB) nanotomography. For example, the use of atomic force microscopy (AFM) has revealed that, contrary to general thought, nanoscale C–S–H has in fact a highly ordered structure (Fig. 1) [3]. A better understanding of the structure of concrete at the nano-level will allow for a better control of concrete performance and even the tailoring of desired properties and is expected therefore to affect the method of production and use of concrete.



Nano-engineered Concrete, Fig. 2 Particle size and specific surface area related to concrete materials (Adapted from [1])

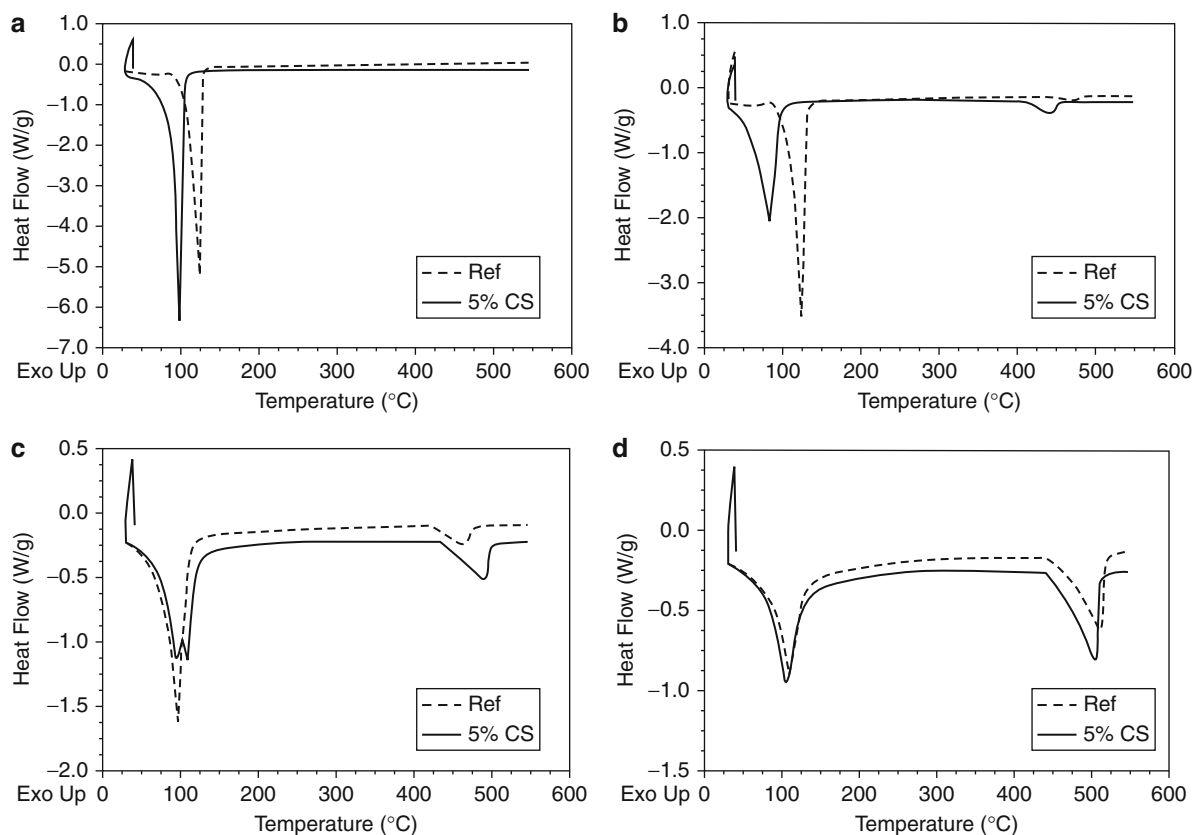
Another application of nanotechnology in concrete has come from the “bottom-up” possibilities of nano-chemistry with the development of new products such as novel superplasticizers and new coating materials [12, 13]. One such superplasticizer was recently developed by BASF and was specifically designed for controlling the slump retention characteristic of the concrete mixture [13]. Another example is the development of coating materials with new self-cleaning properties, discoloration resistance, anti-graffiti protection, and high scratch-and-wear resistance. These self-cleaning materials are based on photocatalyst technology [12]. Titanium dioxide (TiO_2) is used as a photocatalyst for the decomposition of organic compounds. TiO_2 is active under exposure to UV light, exhibiting self cleaning and disinfecting properties. Another aspect of self-cleaning is provided by the hydrophilicity of the surface, which helps to prevent dust and dirt from attaching to it.

In the past, major developments in concrete technology have been achieved through the use of superfine particles such as fly ash and silica fume. Recent advances in nano-chemistry and the development of new methods for synthesis of nanoparticles are now

expected to offer a new range of possibilities for improvement of concrete performance [14, 15] (Fig. 2). Incorporation of nanoparticles into conventional construction materials can provide the materials with advanced or smart properties that are of specific interest for high-rise, long-span, or intelligent infrastructure systems [1, 2, 10, 11, 15].

Performance of Concrete with Nanoparticles

The addition of nanoparticles has been shown to improve various performance of concrete [1, 2, 15–20]. For example, nanosilica (silicon dioxide nanoparticles, nano- SiO_2) has been shown to improve workability and strength in high-performance and self-compacting concrete [1, 2, 14, 15]. Improvement in concrete performance has been attributed to several effects of the nanoparticles. Well-dispersed nanoparticles can act as crystallization centers for cement hydrates, thereby accelerating the hydration reactions and can act as filler, filling the voids between the cement grains, thus reducing the material porosity [15–17]. Well-dispersed nanoparticles have also been shown to promote the



Nano-engineered Concrete, Fig. 3 Differential scanning calorimetry of reference (Ref.) and samples containing 5% of CS at different hydration times: (a) 1 h hydration; (b) 4 h hydration; (c) 8 h hydration; and (d) 24 h hydration [20]

formation of smaller-sized crystals (e.g., $\text{Ca}[\text{OH}]_2$ and AFm), thus densifying the microstructure, and to promote crack arrest and interlocking, improving the material mechanical properties.

Most of the research on the addition of nanoparticles to cementitious systems has been performed using nano- SiO_2 [1, 2, 14–20]. Nano- SiO_2 has been found to accelerate the hydration reactions of C_3S , improve concrete workability and early age strength, be more efficient in enhancing strength than silica fume, and improve the bond between the aggregates and the cement paste [14–20]. The following provides a few examples of recent studies in this area highlighting some of the effects of nano- SiO_2 addition.

The accelerating effect of nano- SiO_2 (colloidal silica, CS) on the hydration of C_3S (alite, Ca_3SiO_5) has been reported by Björnström et al. [17]. Differential scanning calorimetry (DSC) for samples containing up to 5% of colloidal silica (Fig. 3) showed the initial dependence of C–S–H formation on the

addition of colloidal silica and suggested that colloidal silica mostly affects the initial silica polymerization rates rather than the amount of product ultimately formed [17].

Porro et al. [16] reported that the compressive strength of cement pastes increased with the reduction in particle size of nano- SiO_2 . The strength improvement was attributed to the formation of larger silicate chains of the C–S–H gel in cement pastes containing nano- SiO_2 .

In contrast to silica fume (SF), the addition of nano- SiO_2 at dosages of 1–5% has been reported by Qing et al. [18] to increase the thickness and viscosity of superplasticized cement pastes and shorten the initial and final setting times. Nano- SiO_2 significantly improved the compressive strength in the early stages of hardening. After 3 days of curing, the cement pastes with 5% nano- SiO_2 were 41% stronger than the reference pastes. At 28 days, the strength was improved by 25%. Without exception, the strength of the cement pastes

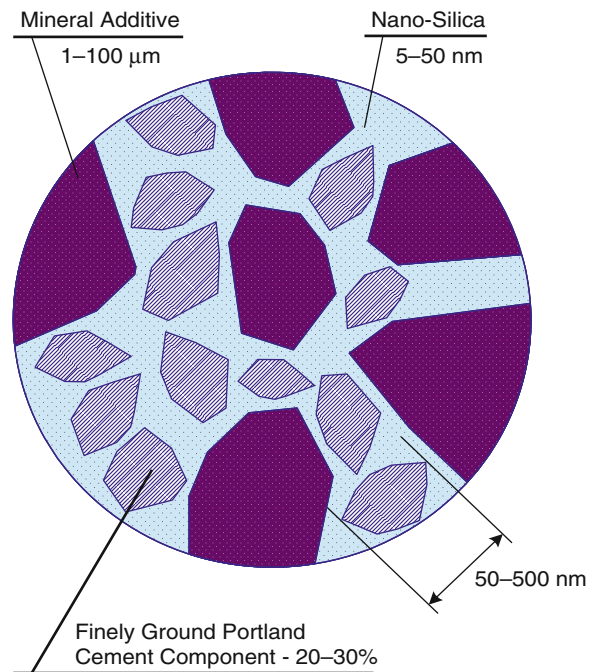
containing nano-SiO₂ improved with increasing dosage of the nano-additive.

The addition of nano-SiO₂ has been shown to significantly improve the pozzolanic activity of fly ash. In a study performed by Li [19], the addition of 4% nano-SiO₂ to concretes prepared using 50% type F fly ash, a water-to-cement ratio of 0.28, and a superplasticizer dosage of 2% (by weight of binder) helped to achieve reaction levels typically obtained at 6 months after only 24 days. The pozzolanic reaction of nano-SiO₂ was shown to be very quick, reaching 70% of its ultimate value (achieved by “common” pozzolans at the age of 2 years) within 3 days and ~95% in a 2-week period. In contrast, plain fly ash reacted with lime very slowly, reaching a significant hydration degree only at the age of 2 years.

The addition of nano-SiO₂ at a dosage of 1–2% has been shown by Collepardi et al. [14] to improve the performance of low-heat self-compacting concrete (SCC). The addition of nano-SiO₂ to blended blast furnace slag concretes made the mixtures more cohesive and reduced bleeding and segregation but had little effect on the slump loss measured within the first 30 min. The best performance was obtained for the blended blast furnace slag concrete prepared using ground fly ash (GFA) as mineral additive, 2% nanosilica, and 1.5% superplasticizer with a compressive strength of 55 MPa at the age of 28 days, low bleeding, higher slump flow, and very little slump loss.

The addition of nano-SiO₂ has also been reported by Hosseini et al. [20] to improve the interfacial transition zone (ITZ) by combining with calcium hydroxide crystals to produce C–S–H leading to a denser, more uniform ITZ. This resulted in a significant improvement of the compressive strengths of recycled aggregate concretes.

As an alternative to nanoparticle addition, the idea of nano-binder was recently proposed to improve concrete performance [1]. In nano-binder, the mineral additives act as the primary components and nano-sized cement is used to fill the gaps between the particles (Fig. 4). The nano-sized cement thus acts as a glue to bind the less reactive mineral additives together. The nano-sized cement can be obtained either by colloidal milling of conventional (optionally high C₂S) portland cement clinker (top-down approach) or by self-assembly (bottom-up approach) through mechanochemically induced topochemical reactions [1].

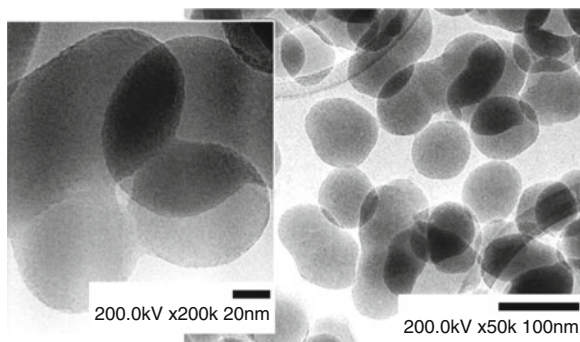


Nano-engineered Concrete, Fig. 4 The concept of nano-binder (nano-engineered cement) (Adapted from [1])

Tailoring of Nanoparticles for Optimal Performance

While small quantities of nanoparticles (less than 1% by weight of cement) are generally sufficient to improve the performance of composites [2, 14–20], practical application in concrete requires manufacturing a considerable amount of these particles [5–7]. Being able to produce nanoparticles in large quantities at a low cost is, therefore, a paramount. Several technologies have been used for the production of nanoparticles [5–7]. The sol-gel process is the most commonly used for the production of nanosilica (Fig. 5). The chemical procedure consists of the formation of a colloidal suspension (sol) and gelation of the sol to form a network in a continuous liquid phase (gel). The procedure can be summarized as follows [7, 15]:

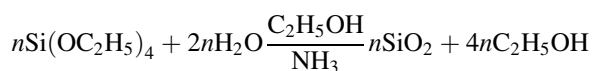
1. Hydrolysis of the precursor
2. Condensation and polymerization of monomers to form the particles
3. Growth of the particles
4. Agglomeration of the particles, followed by the formation of a network and gel structure



Nano-engineered Concrete, Fig. 5 The morphology and particle size distribution of nano-SiO₂ (TEM) synthesized by the sol-gel method [15]

5. Drying (optional) to remove the solvents and thermal treatment (optional) to remove the surface functional groups and obtain the desired crystal structure

Trimethylethoxysilane or tetraethoxysilane (TMOS/TEOS) is typically applied as the precursor in the production of nanosilica and the following reaction takes place [7, 15]:



pH, temperature, concentration of reagents, H₂O/Si molar ratio, and type of catalyst are important parameters that can affect the production and characteristics of the nanoparticles [7, 15].

Flores et al. [15] used the sol-gel process to synthesize nano-SiO₂ with a size range of 5–100 nm (Fig. 5).

Different experimental conditions were examined as summarized in Table 1. X-ray diffraction study showed that the obtained nano-SiO₂ was a highly amorphous material with predominant crystallite size of 1–2.5 nm [15]. The addition of 0.25% (by weight of cementitious material) of the synthesized nano-SiO₂ improved the early strength of portland cement mortars with a 17% strength increase at 1 day. The 28-day compressive strength was improved by as much as 10% and the flexural strength by as much as 25% [15]. Proper dispersion of the nano-SiO₂ within the cement paste is critical for improved performance [1, 15]. The use of polycarboxylate acrylic ester (PAE) superplasticizer, ultrasonification, and/or high-speed mixing has proven to be efficient in the dispersion of nano-SiO₂ [15]. The addition of 0.1% superplasticizer

and 0.25% nano-SiO₂ synthesized using the sol-gel process has shown a 20% increase in the compressive strength of cement mortars (Fig. 6).

Developing New Functionalities

Photocatalytic TiO₂ has been used in white cement-based concrete, providing the material with self-cleaning and air-purification capabilities [12, 21]. It is important for architectural concrete to maintain its aesthetic and decorative characteristics such as color over its entire service life, even in highly polluted urban environments. Photocatalytic materials provide a smart solution to effectively minimize contaminant buildup at the surface of architectural concrete through photocatalytic oxidation. One of the most popular photocatalytic materials used in cement is TiO₂. The anatase polymorph of TiO₂ (especially, nano-sized) is the most effective TiO₂-based photocatalytic material, outperforming rutile in cement-based materials [12, 21]. Photocatalytic nano-TiO₂ embedded into cement matrix has also been reported to be very effective for NO_x abatement [21].

Application of Carbon Nanotubes and Carbon Nanofibers

Due to extraordinary strength and tensile modulus, carbon nanotubes and carbon nanofibers (CNTs/CNFs) are potential candidates for use as nano-reinforcement in cement-based materials [1, 2, 5, 6, 22–26]. Vapor-grown CNTs are concentric cylinders made from rolled-up planes of hexagonally bound carbon atoms. Vapor-grown CNFs (Fig. 7) exhibit a “herringbone” or cup-stacked structure of nested graphene layers with a hollow core. The production of CNTs on an industrial scale remains expensive. In contrast, CNFs have a lower production cost and can be made at a higher production capacity [27]. The use of CNTs/CNFs in cement remains limited [2]. The proper dispersion of CNTs/CNFs within the cement paste is one of the main challenges [2, 22–26] due to their high hydrophobicity and strong self-attraction. A number of methods have been used to improve the dispersion of CNTs/CNFs in cement, including the use of surfactants, surface treatment with acids, high shear mixing, and sonication [22–26].

Nano-engineered Concrete, Table 1 Design and properties of nano-SiO₂ [15]

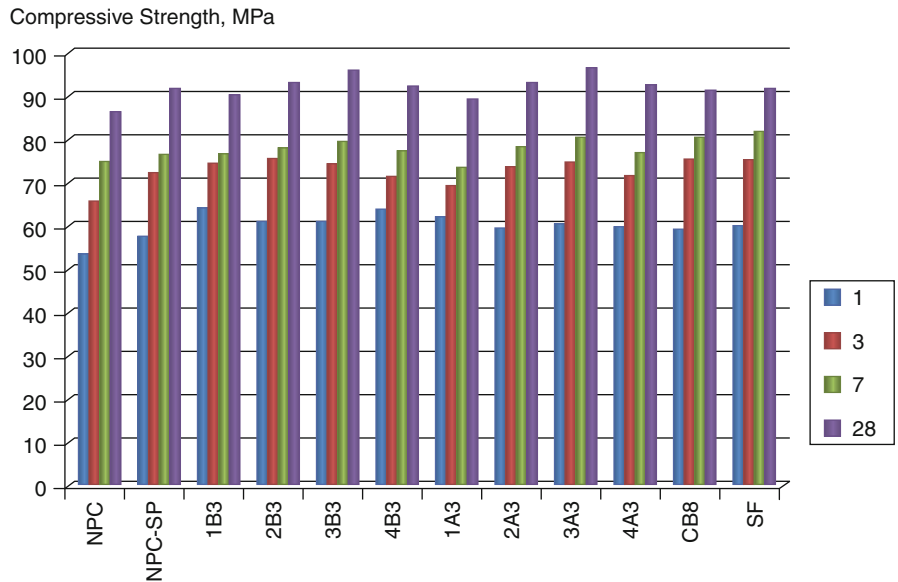
Specimen type ^a	Molar ratio TEOS/Etanol/H ₂ O	Reaction time, hours	Particle size (TEM), nm	Surface area (BET), m ² /kg
1B3	1/24/6	3	15–65	116,000
2B3	1/6/6	3	30, 60–70	145,000
3B3	1/6/24	3	15–20	133,000
4B3	1/24/24	3	5	163,000
1A3	1/24/6	3	5	510,100
2A3	1/6/6	3	<10	263,500
3A3	1/6/24	3	<10, 17	337,100
4A3	1/24/24	3	5	382,200

^a Sample coding

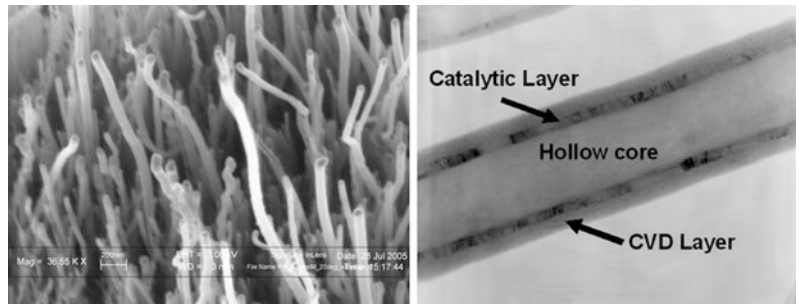
- First number denotes molar ratio combination as per as experimental matrix
- ABC — Last number corresponds to the reaction time in hours
- Letter – denotes reaction media: A-for acid (pH=2) and B-for base (pH=9)

Nano-engineered Concrete,

Fig. 6 Compressive strength of superplasticized mortars with nano-SiO₂ at different ages (NPC and NPC-SP are reference portland cement and superplasticized mortars; nanoparticle specification as per as Table 1, Fig. 5; CB-8 and SF are cembinder-8 and silica fume) [15]



Nano-engineered Concrete, Fig. 7 SEM (left) and TEM (right) image of carbon nanofibers [27]



To date mixed performance results have been achieved ranging from significant improvements for small addition of CNTs/CNFs to minimal or no improvement. Among the promising results, Makar et al. [23] demonstrated that CNTs can affect early age hydration and indicated the potential of CNTs for crack bridging and enhanced stress transfer. Shah et al. [24] showed that, after dispersion in water using surfactant and ultrasonic energy, small amounts of CNTs (0.048% and 0.08%) can result in a significant (~50%) increase in the Young's modulus of cement pastes.

Smart Stress-Sensing Composites

The incorporation of conductive nanomaterials (CNTs/CNFs, nano-Fe₂O₃, etc.) into cement-based composites has been reported to provide the materials with self-sensing capability [28, 29]. It has been shown that carbon fiber-reinforced cement may function as a piezoresistive strain sensor [28]. Due to the high conductivity of the carbon fibers compared to the cement matrix, the introduction of the fibers into cement decreases the electrical resistivity of the material. Strain sensitivity (i.e., fractional change in electrical resistance per unit strain) of as much as 700 has been obtained with incorporation of 0.2–0.5% (by volume) carbon microfibers [28]. Recently, a self-sensing cement material with multi-walled carbon nanotubes (MWCNT) has been proposed for traffic monitoring [29]. The piezoresistive properties of the MWCNT-cement composite enabled the detection of mechanical stresses induced by traffic flow.

Future Developments

Concrete science is expected to benefit significantly from the rapidly growing field of nanotechnology. Advances in the characterization of the nanoscale structure of concrete are providing new insights about concrete. New developments are underway in the nano-engineering of concrete. Some of the most promising include the development of [1]:

- Highly efficient superplasticizers for concrete and high-strength fibers with exceptional energy-absorbing capacity
- Concrete with self-cleaning properties, discoloration resistance, and anti-graffiti protection, based on photocatalytic technology
- Reinforced binders with nanoparticles or nanotubes
- Concrete with smart properties such as temperature-, moisture-, or stress-sensing capability, deformative properties, and non-shrinking and low thermal expansion
- Eco-binders modified with nanoparticles and produced with significantly reduced amount of portland cement
- Eco-binders with nanoparticles based on alternative systems (MgO, phosphate, geopolymers, gypsum)
- Self-healing concrete with controlled supply of internal moisture or solution of repair/bonding agent

Conclusions

Nanotechnology in concrete is still in its exploratory phase and full-scale applications in the field of construction are to date very limited. Nevertheless, the potential for nanotechnology to help improve the performance and durability of one of the oldest known construction materials is most promising.

Acknowledgments KS is grateful to Michelle M. Schoenecker, Ismael Flores, and Zhibin Lin for their help with the reported experiments and manuscript preparation. KS gratefully acknowledges the financial support from the National Science Foundation under NSF CMMI 0937652 and the NSF travel grant to attend the NICOM 3 conference, as well as support of CONACYT, PROMEP, and PAICYT (Mexico). F.S. gratefully acknowledges the financial support from the National Science Foundation under NSF CAREER CMMI 0547024 and the NSF travel grant to attend the NICOM3 conference.

Cross-References

- ▶ [Atomic Force Microscopy](#)
- ▶ [Carbon Nanotubes](#)
- ▶ [Nanoparticles](#)
- ▶ [Sol-Gel Method](#)
- ▶ [Spider Silk](#)

References

1. Sobolev, K., Ferrada-Gutiérrez, M.: How nanotechnology can change the concrete world: part 2. *Am. Ceram. Soc. Bull.* **11**, 16–19 (2005)
2. Sanchez, F., Sobolev, K.: Nanotechnology in concrete – a review. *Constr. Build. Mater.* **24**(11), 2060–2071 (2010)
3. Plassard, C., Lesniewska, E., Pochard, I., Nonat, A.: Investigation of the surface structure and elastic properties of calcium silicate hydrates at the nanoscale. *Ultramicroscopy* **100**(3–4), 331–338 (2004)
4. Scrivener, K.L.: Nanotechnology and cementitious materials. In: Bittnar, Z., Bartos, P.J.M., Nemecek, J., Smilauer, V., Zeman, J., (eds.) *Nanotechnology in construction: proceedings of the NICOM3, 3 rd International symposium on nanotechnology in construction*. Prague, Czech Republic, pp. 37–42. (2009)
5. Bhushan, B. (ed.): *Handbook of Nanotechnology*. Springer, Berlin (2004)
6. Atkinson, W.I.: *Nanocosm – Nanotechnology and the Big Changes Coming from the Inconceivably Small*, pp. 36–39. AMACOM, New York (2003)
7. Wilson, M., Smith, K.K.G., Simmons, M., Raguse, B.: *Nanotechnology – Basic Science and Emerging Technologies*. Chapman and Hall/CRC, Boca Raton/London (2000)
8. Mart, J., Mijangos, C.: Tailored polymer-based nanofibers and nanotubes by means of different infiltration methods into alumina nanopores. *Langmuir* **25**(2), 1181–1187 (2009)
9. Dalton, A.B., Collins, S., Razal, S.J., Munoz, E., Ebron, V.H., Kim, B.G., Coleman, J.N., Ferraris, J.P., Baughman, R.H.: Continuous carbon nanotube composite fibers: properties, potential applications, and problems. *J. Mater. Chem.* **14**, 1–3 (2004)
10. Trtik, P., Bartos, P.J.M.: Nanotechnology and concrete: what can we utilise from the upcoming technologies? In: *Proceeding of the 2nd Anna maria workshop: cement & concrete : trends & challenges*, Holmes Beach, FL, USA pp. 109–120 (2001)
11. Beaudoin, J.J.: Why engineers need materials science. *Concrete Int.* **21**(8), 86–89 (1999)
12. Watanabe, T., Kojima, E., Norimoto, K., Kimura, T., Machida, M., Hayakawa, M., Kitamura, A., Chikuni, M., Saeki, Y., Kuga, T., Nakashima, Y.: Multi-functional material with photocatalytic functions and method of manufacturing same. US patent US6294247 (2001)
13. Corradi, M., Khurana, R., Magarotto, R.: Controlling performance in ready mixed concrete. *Concrete Int.* **26**(8), 123–126 (2004)
14. Collepardi, M., Ogoumah-Olagot, J.J., Skarp, U., Troli, R.: Influence of amorphous colloidal silica on the properties of self-compacting concretes. In: *Proceedings of the international conference, Challenges in concrete construction – innovations and developments in concrete materials and construction*, Dundee, UK, pp. 473–483 (2002)
15. Flores I., Sobolev K., Torres L.M., Valdez P.L., E. Zarazua, Cuellar E.L.: Performance of cement systems with Nano- SiO₂ particles produced using sol-gel method. In: *TRB first international conference in North America on nanotechnology in cement and concrete*, Irvine, California, USA, 5–7 May 2010
16. Porro, A., Dolado, J.S., Campillo, I., Erkizia, E., de Miguel, Y., Sáez de Ibarra, Y., Ayuela, A.: Effects of nanosilica additions on cement pastes. In: Ravindra, R.K., Dhir, K., Newlands, M.D., Csetenyi, L.J. (eds.) *Applications of Nanotechnology in Concrete Design*. Thomas Telford, London (2005)
17. Bjornstrom, J., Martinelli, A., Matic, A., Borjesson, L., Panas, I.: Accelerating effects of colloidal nano-silica for beneficial calcium-silicate-hydrate formation in cement. *Chem. Phys. Lett.* **392**(1–3), 242–248 (2004)
18. Qing, Y., Zenan, Z., Deyu, K., Rongshen, C.: Influence of nano-SiO₂ addition on properties of hardened cement paste as compared with silica fume. *Constr. Build. Mater.* **21**(3), 539–545 (2007)
19. Li, G.: Properties of high-volume fly ash concrete incorporating nano-SiO₂. *Cement Concrete Res.* **34**, 1043–1049 (2004)
20. Hosseini, P.B., Booshehrian, A., Delkash, M., Ghavami, S., Zanjani, M.K.: Use of Nano-SiO₂ to improve microstructure and compressive strength of recycled aggregate concretes. In: *NICOM3, 3 rd International symposium on nanotechnology in construction*, Prague, Czech Republic (2009)
21. Cassar, L., Pepe, C., Tognon, G., Guerrini, G.L., Amadelli, R.: White cement for architectural concrete possessing photocatalytic properties. In: *Proceedings of the 11th international congress on the chemistry of cement (ICCC)*, Durban, South Africa, p 2012 (2003)
22. Sanchez, F.: Carbon nanofiber/cement composites: challenges and promises as structural materials. *Int. J. Mater. Struct. Integr.* **3**(2–3), 217–226 (2009). Special issue on nanotechnology for structural materials
23. Makar, J.M., Margeson, J., Luh, J.: Carbon nanotube/cement composites – early results and potential applications. In: *Proceedings of 3 rd international conference on construction materials: performance, innovations and structural implications*. Vancouver, 22–24 Aug 2005, p. 1–10
24. Shah S.P., Konsta-Gdoutos M.S., Metaxa Z.S., Mondal P.: Nanoscale modification of cementitious materials. In: Bittnar, Z., Bartos, P.J.M., Nemecek, J., Smilauer, V., Zeman, J. (eds.) *Nanotechnology in construction: proceedings of the NICOM3, 3 rd International Symposium on Nanotechnology in Construction*, Prague, Czech Republic, p. 125–130 (2009)
25. Gay, C., Sanchez, F.: Performance of carbon nanofibers/cementitious composites with a high-range water-reducer. *J. Transport. Res. Board* **2142**, 109–113 (2010). *Nanotechnology in Cement and Concrete, Volume 2*
26. Sanchez, F., Ince, C.: Microstructure and macroscopic properties of hybrid carbon nanofibers/silica fume cement composites. *Compos. Sci. Technol.* **69**(7–8), 1310–1318 (2009)
27. Pyrograf-III carbon nanofiber http://pyrografproducts.com/Merchant5/merchant.mvc?Screen=cp_nanofiber, Accessed 25 May 2011
28. Chung, D.: Piezoresistive cement-based materials for strain sensing. *J. Intell. Mater. Syst. Struct.* **13**(9), 599–609 (2002)
29. Li, H., Xiao, H.G., Ou, J.P.: A study on mechanical and pressure-sensitive properties of cement mortar with nanophase materials. *Cement Concrete Res.* **34**, 435–438 (2004)

Nanoengineered Hydrogels for Cell Engineering

Esmail Jabbari

Biomimetic Materials and Tissue Engineering Laboratory, Department of Chemical Engineering, Swearingen Engineering Center, Rm 2C11, University of South Carolina, Columbia, SC, USA

Synonyms

[Nanostructured hydrogels](#); [Patterned hydrogels](#)

Definition

Nanoengineered hydrogels are soft, high water content, hydrophilic networks with nanoscale topographical or biomolecular patterns for mimicking the extracellular matrix of tissues in order to control cell fate and function.

Introduction

Hydrogels are three-dimensional (3D) cross-linked polymeric networks that retain a significant fraction of water in their structure without dissolving when placed in physiological solution. Due to their high water content, biomolecules like peptides and proteins and cells immobilized in hydrogels are more likely to retain their biological activity. Poly(ethylene glycol) (PEG)-based hydrogels, due to minimal interaction of the ethylene oxide repeating unit with biomolecules, do not elicit an immune response. Therefore, PEG hydrogels are used extensively for cell immobilization and encapsulation. Furthermore, PEG networks can be conjugated with bioactive molecules to produce cell-responsive hydrogels for applications in tissue engineering and regenerative medicine [1]. The cell fate including shape, morphology, phenotype, and function is determined by interaction with soluble and insoluble factors in the extracellular matrix (ECM) [2]. These interactions are between the functional motifs of the ECM proteins and binding sites of the glycoproteins embedded in the cell's plasma membrane. The functional motifs as well as geometrical and morphological

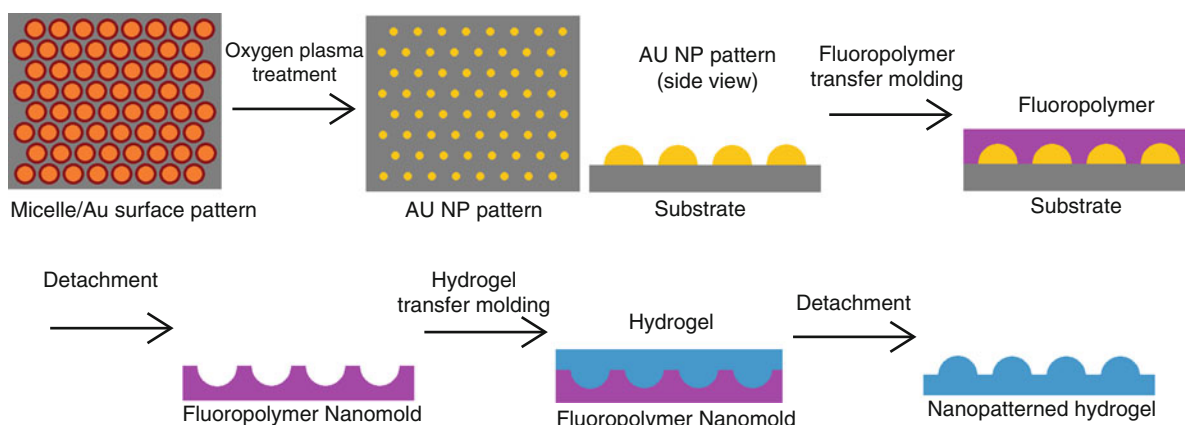
features of the ECM are ordered at the nanoscale. As a result, synthetic hydrogels patterned with biologically active molecules or topographical nanostructures are important in understanding the role of ECM on cell fate, and the engineering of cell environment at the nanoscale. Patterning of hydrogels at the nanoscale includes surface techniques (2D patterning) and bulk/matrix techniques (3D patterning). The cell-nanostructure interaction is two dimensional for surface-patterned hydrogels while the interaction is three dimensional for matrix-patterned hydrogels. Surface patterning techniques include electron beam patterning, micelle nanolithography, UV-assisted nanomolding, and nanostamping. Matrix techniques include polymer phase separation and multiphoton chemical patterning.

Electron Beam Patterning

Hydrogels are surface-patterned at the nanoscale using a focused electron beam. In this method, a thin layer of hydrogel precursor solution, spin-coated on silicon wafers, is exposed to a focused electron beam using a modified field-emission scanning electron microscope [3]. The extent of cross-linking and density of the pattern can be controlled with this technique by varying the pixel dwell time to irradiation dose and interpixel spacing. After irradiation, the pattern is exposed by dissolving unexposed areas with organic solvents and rinsing in water. Topographical patterns with sizes ranging 50–1,000 nm can be produced with this technique. As an example, poly(sodium 4-styrenesulfonate-co-poly(ethylene glycol) methacrylate) (pSS-co-pPEGMA) with heparin-mimicking thiol groups was patterned with electron beam and the patterned gel is used for localized attachment of vasculogenic proteins bFGF and VEGF [4]. As another example, Boc-protected aminoxy tetraethylene glycol methacrylate was patterned with electron beam, the Boc-protecting group was removed to expose the aminoxy groups for localized grafting of proteins to the patterned hydrogel via oxime ligation [5].

Micelle Nanolithography

Self-assembled array of metal-impregnated diblock copolymers is used for nanoscale surface patterning



Nanoengineered Hydrogels for Cell Engineering, Fig. 1 Schematic diagram for fabrication of a surface nanopatterned hydrogel by micelle nanolithography (Refs. [7, 11])



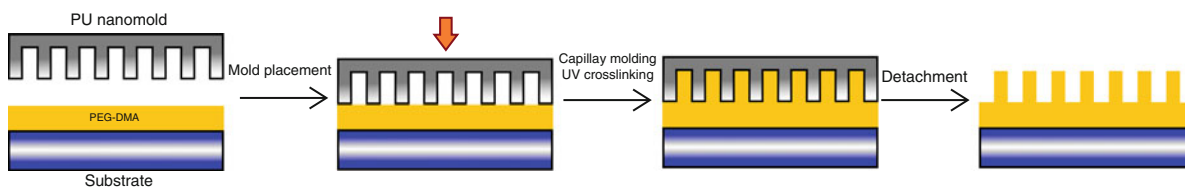
Nanoengineered Hydrogels for Cell Engineering, Fig. 2 Schematic diagram for fabrication of PDMS nanostamp (Ref. [8])

of hydrogels. In this technique, a monolayer of polystyrene-block-poly(2-vinylpyridine) (PS-*b*-P2VP) and HAuCl₄ micelles is formed on a substrate (glass or silicon wafers) by dip-coating [6]. Next, the micellar array is treated with oxygen plasma to remove the PS-*b*-P2VP copolymer and exposing the ordered array of gold nanoparticles (NPs) on the substrate, as shown in Fig. 1. The pattern of gold NPs is moved to a nonadhesive fluorosiloxane layer by transfer-molding, followed by transfer to a hydrogel surface as shown in Fig. 1. In a variation of this technique, polystyrene homopolymer, as an ordering interference reagent, is mixed with PS-*b*-P2VP and HAuCl₄ micelles to produce a disordered array of gold NPs on the substrate [7], leading to a disordered nanopattern on the hydrogel surface. In another variation, PtCl₄ is used in the place of HAuCl₄ to form a nanoscale pattern of platinum NPs on the substrate [8]. The platinum pattern is moved to a PDMS substrate by transfer-molding, as shown in Fig. 2, to produce a catalytic nanostamp for chemical patterning of hydrogel surfaces. In another interesting approach to generate micro- nanostructured hydrogels, an array of gold NPs on a glass substrate was coated with a photoresist layer [9]. After photolithography and removal of the uncovered gold NPs and the protecting resist layer, the microstructured nanoarray of gold NPs was covered with acrylated

PEG hydrogel precursor solution and cross-linked by ultraviolet (UV) radiation. Upon detaching from the glass substrate, a hydrogel surface with microstructured array of gold NPs was generated which could potentially be used to study the modulation of nano- and microscale patterns of biomolecules on cell function.

Nanomolding and Nanostamping

The block-copolymer-templated micelle lithography is used to fabricate stamps and molds in sub-100 nanoscale for patterning [8]. In this approach, a silicon substrate is patterned with platinum (Pt) NPs by micelle lithography and the bare silicon surface is covalently covered with a fluorosiloxane monolayer. Next, a PDMS layer is spin-coated on the patterned substrate followed by curing. Upon peeling the PDMS layer, a Pt nanopatterned catalytic stamp and a silicon nanopatterned mold are produced. The catalytic stamp can be used to print nanoscale biomolecular patterns on hydrogels. The pattern on silicon/glass substrate can be moved to a polyurethane (PU), PDMS, or other soft hydrophobic substrates [10] for subsequent patterning of hydrogel surfaces [11]. As an example, in UV-assisted capillary lithography [12], a



Nanoengineered Hydrogels for Cell Engineering, Fig. 3 Schematic diagram for UV-assisted capillary nanomolding (Ref. [12])

nanopatterned PU mold is placed over a thin layer of methacrylated PEG precursor solution to make conformal contact by capillary forces, as shown in Fig. 3. After cross-linking of the PEG layer by UV-assisted photopolymerization, the PU mold is peeled off to produce a PEG hydrogel with a surface pattern ranging 50–500 nm in size. Direct molding of a hydrogel precursor solution with nanotemplated quartz followed by cross-linking can also be used to form hydrogels with nanopatterned surface topography [13].

Copolymer Phase Separation

Phase separation of amphiphilic di- and triblock copolymers at the nanoscale is used to form hydrogels with less than 50 nm matrix nanopatterns. For example, amphiphilic triblock copolymers of poly(ϵ -caprolactone) and PEG were allowed to self-assemble in aqueous solution to form 30–40 nm ordered cylindrical structures [14]. Following electron beam cross-linking of the microphase separated mixture, the PCL phase was removed by hydrolysis to form a porous hydrogel, which can potentially be patterned in 3D with cell-responsive functional groups. In a similar approach, microphase separation of a blend of polystyrene-block-poly(ethylene oxide) (PS-*b*-PEO) and arginine-glycine-aspartic acid (RGD)-conjugated PS-*b*-PEO was used to form nanoscale domains of RGD integrin-binding peptide in the hydrogel matrix [15]. By manipulating phase separation of the diblock copolymer, domain sizes of 8–14 nm, lateral spacing of 62–44 nm, and island densities of 450–900 per μm^2 were formed. In a related approach, a blend of methacrylated PEG-*b*-PLA-*b*-PEG (PLA denotes poly(lactic acid)), methacrylated biotin-PEG, and a self-assembling lyotropic liquid crystalline mesophase as the polymerization template was used to generate biotinylated hydrogels with lamellar matrix morphology [16]. The biotin-avidin chemistry was used to

attach integrin-binding RGD peptide to the nanostructured matrix. In a similar approach, a mixture of blank alginate chains and integrin-binding RGD-coupled alginate chains was cross-linked in aqueous solution to form nanopatterned RGD islands in an alginate matrix [17]. Similarly, a mixture of thiol-derivatized hyaluronic acid (HA) and acrylated nanogels was cross-linked to form a uniformly distributed nanostructured HA hydrogel [18].

Cell Engineering with Nanopatterned Gels

When cardiomyocytes were seeded on PEG hydrogels patterned with 50–150 nm nanopillars, cell adhesion was significantly enhanced and formed 3D aggregates, which retained their conductive and contractile properties [12]. Neonatal rat ventricular myocytes (NRVMs) seeded on anisotropically nanofabricated hydrogel substrates, with ridges/grooves in the 50–800 nm range, had perfect alignment with the ridge/groove direction, the direction of cell contraction was closely aligned with the pattern direction, and the action potential direction was anisotropic, when compared to unpatterned hydrogel substrates [19]. The spreading of NIH-3T3 fibroblasts increased from 1,100 to 1,700 μm^2 as 3D lateral spacing of RGD integrin-binding domains in a PS-*b*-PEO hydrogel matrix was decreased from 62 to 44 nm [15]. The phosphorylation of Y397 focal adhesion kinase, cell spreading, and osteogenic differentiation of MC3T3 preosteoblasts depended on the spacing of RGD islands in an alginate matrix. Closer spacing of RGD islands favored phosphorylation of Y397 and cell spreading while wider spacing favored differentiation, and proliferation was a function of RGD density in the matrix [17]. Block copolymer micelle lithography followed by ligand conjugation was used to form a gradient in the spacing of cyclic-RGD peptide, specifically recognized by $\alpha_v\beta_3$ integrins, on a substrate

[20]. The gradient in cyclic-RGD spacing induced cell orientation toward shorter spacing. In another study, integrin clustering on MC3T3-E1 osteoblasts depended on the local order of RGD ligand arrangement on the substrate for average ligand spacings >70 nm [7]. Cell adhesion was lower on substrates with ordered RGD pattern and higher on disordered patterns. Human embryonic stem cells seeded on polyurethane gels with 350-nm grooved arrays, fabricated by UV capillary force lithography, differentiated directly into the neuronal lineage in the absence of differentiation-inducing agents [21]. Submicron pattern of cell-repulsive PEG hydrogel on an adhesive substrate can discriminate between different cell types. At the highest PEG density with overlapping patterns, neither neurite cells nor astrocytes adhered to the substrate [22]. As the spacing between PEG patterns was increased to 2 μm , neurons were able to grow and form long axons while astrocytes were unable to grow. At higher PEG spacing of 10 μm , both neurons and astrocytes were able to grow on the patterned substrates. In an interesting study, poly(vinyl alcohol) (PVA) nanogels were used to form nanoarrays of tethered lipid bilayer rafts (tLBRMs) on a solid substrate to study the effect of cell membrane's lipid rafts on signal transduction [23]. In this work, PVA hydrogel was nanopatterned on a gold substrate using UV nanolithography to form selective nanowells. After covering the gold nanowells with a monolayer of vesicle rupturing promoter, tLBRMs were selectively immobilized on a gold substrate between the PVA nanogels by monolayer-assisted vesicle fusion. These studies clearly demonstrate that engineered hydrogels can play a major role in understanding the effect of topography and ECM ligands, patterned at the nanoscale, on cell shape, morphology, phenotype, and fate.

Future Direction

A number of techniques such as electron beam patterning and micelle lithography have been developed for nanoscale patterning of hydrogel surfaces. However, there is a need to develop techniques for 3D patterning of hydrogels at the nanoscale, as the interaction of cells with ligands in the ECM is three dimensional. Related to that, the effect of surface pattern of adhesive ligands on cell shape and morphology has been studied

extensively but there is a need to conduct similar studies in 3D within a nanopatterned hydrogel matrix. Those studies should be extended to the effect of 3D pattern of ECM ligands on cell differentiation, lineage selection, and fate.

Cross-References

- ▶ [Bioadhesion](#)
- ▶ [Bioinspired Synthesis of Nanomaterials](#)
- ▶ [Biomimetic Synthesis of Nanomaterials](#)
- ▶ [Biomimetics](#)
- ▶ [BioPatterning](#)
- ▶ [Dip-Pen Nanolithography](#)
- ▶ [Electron Beam Lithography \(EBL\)](#)
- ▶ [Injectable Hydrogel](#)
- ▶ [MEMS on Flexible Substrates](#)
- ▶ [Nanoimprint Lithography](#)
- ▶ [Nanoimprinting](#)
- ▶ [Nanopatterned Surfaces for Exploring and Regulating Cell Behavior](#)
- ▶ [Nanoscale Printing](#)
- ▶ [Nanotechnology](#)
- ▶ [Nanotechnology in Cardiovascular Diseases](#)
- ▶ [Self-assembly](#)
- ▶ [Self-assembly of Nanostructures](#)
- ▶ [Smart Hydrogels](#)
- ▶ [Stereolithography](#)

References

1. He, X.Z., Ma, J.Y., Jabbari, E.: Effect of grafting RGD and BMP-2 protein-derived peptides to a hydrogel substrate on osteogenic differentiation of marrow stromal cells. *Langmuir* **24**, 12508–12516 (2008)
2. Fisher, O.Z., Khademhosseini, A., Langer, R., Peppas, N.A.: Bioinspired materials for controlling stem cell fate. *Acc. Chem. Res.* **43**, 419–428 (2009)
3. Hong, Y., Krsko, P., Libera, M.: Protein surface patterning using nanoscale PEG hydrogels. *Langmuir* **20**, 11123–11126 (2004)
4. Christman, K.L., Vazquez-Dorbatt, V., Schopf, E., Kolodziej, C.M., Li, R.C., Broyer, R.M., Chen, Y., Maynard, H.D.: Nanoscale growth factor patterns by immobilization on a heparin-mimicking polymer. *J. Am. Chem. Soc.* **130**, 16585–16591 (2008)
5. Christman, K.L., Broyer, R.M., Schopf, E., Kolodziej, C.M., Chen, Y., Maynard, H.D.: Protein nanopatterns by oxime bond formation. *Langmuir* **27**, 1415–1418 (2010)
6. Blummel, J., Perschmann, N., Aydin, D., Drinjakovic, J., Surrey, T., Lopez-Garcia, M., Kessler, H., Spatz, J.P.:

- Protein repellent properties of covalently attached PEG coatings on nanostructured SiO₂-based interfaces. *Biomaterials* **28**, 4739–4747 (2007)
7. Huang, J., Grater, S.V., Corbellini, F., Rinck, S., Bock, E., Kemkemer, R., Kessler, H., Ding, J., Spatz, J.P.: Impact of order and disorder in RGD nanopatterns on cell adhesion. *Nano Lett.* **9**, 1111–1116 (2009)
 8. Mizuno, H., Buriak, J.M.: Nanoscale patterning of organic monolayers by catalytic stamp lithography: scope and limitations. *ACS Appl. Mater. Interfaces* **1**, 2711–2720 (2009)
 9. Aydin, D., Louban, I., Perschmann, N., Blümmel, J., Lohmüller, T., Cavalcanti-Adam, E.A., Haas, T.L., Walczak, H., Kessler, H., Fiammengo, R., Spatz, J.P.: Polymeric substrates with tunable elasticity and nanoscopically controlled biomolecule presentation. *Langmuir* **26**, 15472–15480 (2010)
 10. Graeter, S.V., Huang, J., Perschmann, N., López-García, M., Kessler, H., Ding, J., Spatz, J.P.: Mimicking cellular environments by nanostructured soft interfaces. *Nano Lett.* **7**, 1413–1418 (2007)
 11. Diez, M., Mela, P., Seshan, V., Möller, M., Lensen, M.C.: Nanomolding of PEG-based hydrogels with Sub-10-nm resolution. *Small* **5**, 2756–2760 (2009)
 12. Kim, D.-H., Kim, P., Song, I., Cha, J.M., Lee, S.H., Kim, B., Suh, K.Y.: Guided three-dimensional growth of functional cardiomyocytes on polyethylene glycol nanostructures. *Langmuir* **22**, 5419–5426 (2006)
 13. Gaston, A., Khokhar, A.Z., Bilbao, L., Sáez-Martínez, V., Corres, A., Obieta, I., Gadegaard, N., Nanopatterned, U.V.: UV curable hydrogels for biomedical applications. *Microelectron. Eng.* **87**, 1057–1061 (2009)
 14. Kang, J., Beers, K.J.: Synthesis and characterization of PCL-b-PEO-b-PCL-based nanostructured and porous hydrogels. *Biomacromolecules* **7**, 453–458 (2006)
 15. George, P.A., Doran, M.R., Croll, T.I., Munro, T.P., Cooper-White, J.J.: Nanoscale presentation of cell adhesive molecules via block copolymer self-assembly. *Biomaterials* **30**, 4732–4737 (2009)
 16. Clapper, J.D., Pearce, M.E., Guymon, C.A., Salem, A.K.: Biotinylated biodegradable nanotemplated hydrogel networks for cell interactive applications. *Biomacromolecules* **9**, 1188–1194 (2008)
 17. Comisar, W.A., Kazmers, N.H., Mooney, D.J., Linderman, J.J., Engineering, R.G.D.: Nanopatterned hydrogels to control preosteoblast behavior: a combined computational and experimental approach. *Biomaterials* **28**, 4409–4417 (2007)
 18. Bencherif, S.A., Washburn, N.R., Matyjaszewski, K.: Synthesis by AGET ATRP of degradable nanogel precursors for in situ formation of nanostructured hyaluronic acid hydrogel. *Biomacromolecules* **10**, 2499–2507 (2009)
 19. Kim, D.-H., Lipke, E.A., Kim, P., Cheong, R., Thompson, S., Delannoy, M., Suh, K.-Y., Tung, L., Levchenko, A.: Nanoscale cues regulate the structure and function of macroscopic cardiac tissue constructs. *Proc. Natl. Acad. Sci. USA* **107**, 565–570 (2010)
 20. Hirschfeld-Warneken, V.C., Arnold, M., Cavalcanti-Adam, A., López-García, M., Kessler, H., Spatz, J.P.: Cell adhesion and polarisation on molecularly defined spacing gradient surfaces of cyclic RGDfK peptide patches. *Eur. J. Cell Biol.* **87**, 743–750 (2008)
 21. Lee, M.R., Kwon, K.W., Jung, H., Kim, H.N., Suh, K.Y., Kim, K., Kim, K.-S.: Direct differentiation of human embryonic stem cells into selective neurons on nanoscale ridge/groove pattern arrays. *Biomaterials* **31**, 4360–4366 (2010)
 22. Krsko, P., McCann, T.E., Thach, T.-T., Laabs, T.L., Geller, H.M., Libera, M.R.: Length-scale mediated adhesion and directed growth of neural cells by surface-patterned poly(ethylene glycol) hydrogels. *Biomaterials* **30**, 721–729 (2009)
 23. Lee, B.K., Lee, H.Y., Kim, P., Suh, K.Y., Kawai, T.: Nanoarrays of tethered lipid bilayer rafts on poly(vinyl alcohol) hydrogels. *Lab Chip* **9**, 132–139 (2009)
-
- ## Nanofabrication
- ▶ [Integration of Nanostructures within Microfluidic Devices](#)
-
- ## Nanofatigue Tester
- ▶ [Nanoindentation](#)
-
- ## Nano-FET
- ▶ [Nanostructure Field Effect Transistor Biosensors](#)
-
- ## Nanofibrous Materials and Composites
- ▶ [Computational Study of Nanomaterials: From Large-Scale Atomistic Simulations to Mesoscopic Modeling](#)
-
- ## Nanofluidic Channels
- ▶ [Nanochannels for Nanofluidics: Fabrication Aspects](#)
-
- ## Nanofluidics
- ▶ [Electrokinetic Fluid Flow in Nanostructures](#)
 - ▶ [Integration of Nanostructures within Microfluidic Devices](#)

Nanofocusing of Light

► [Light Localization for Nano-optical Devices](#)

Nanogap Biosensors

J. Tanner Nevill and Daniele Malleo
Fluxion Biosciences, South San Francisco, CA, USA

Definition

A nanogap biosensor is an arrangement of two electrodes separated by no more than 300 nm that is used to electrically detect biologically relevant materials, reactions, or interactions in solution. The limit of 300 nm is imposed because it represents the practical upper limit of the characteristic thickness of the electrical double layer, which forms at all charged surfaces in aqueous solutions.

Overview

Nanogap biosensors are made using various micro- and nano-fabrication techniques to create a nanometer-scale detection region. This region is interrogated using electrical techniques to sense and measure the presence or activity of a biomolecule. All-electrical measurements offer the promise of low cost, rapid, and label-free detection. Many arguments are made for using nanogap biosensors, including higher sensitivity, reduction of parasitic effects, overlapping double layers, and the use of molecular-sized cavities. Liang and Chou use a nanogap electrode arrangement inside of a nanofluidic channel like a pore for detecting (and eventually sequencing) single strands of DNA [1]. Im et al. exploit the fact that a molecule can occupy a significant portion of the nanogap, thereby making it very sensitive to low numbers of molecules: By functionalizing the surface of the nanogap, analyte binding significantly changes the electrical characteristic of the sensing electrodes [2].

The electrical double layer, a cloud of counterions that balances out the charge present on any surface in solution, is often seen as a parasitic component of

electrical measurements: This layer shields the bulk solution from the induced electric field thus decreasing the sensitivity of alternating current detection techniques. Because nanogaps have dimensions similar or smaller than the double layer thickness, the electric potential in a nanogap distributes differently than in a macroscale gap. In fact, nanogap devices typically create environs where the volume of interest consists exclusively of double layer regions, actually eliminating the bulk solution from the measurement. Elimination of the bulk response is desirable in many situations, where measurement and quantification of very few or even single molecule events is of interest, making the concept of overlapping double layers in nanogaps very compelling.

The sensors developed thus far can be classified by the geometry of the detection region. According to the number of dimensions along which the sensing cavity exists, there are three categories as depicted in Fig. 1: 1D nanogaps refer to point gap junctions, 2D devices consist of coplanar electrodes or “fractured planes,” and 3D nanogaps are “nanocavities,” where the electrode surfaces forming the cavity have a significant area relative to the gap length.

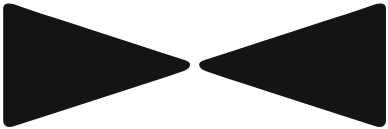





As form follows function, the measurement techniques and applications can be correlated with the geometry. 1D and 2D nanogaps are typically interrogated with DC voltages to produce a resistive measurement, while 3D nanogaps are mostly interrogated using AC measurement techniques, to measure a complex impedance response. Consequently, the discussion of nanogap biosensors is divided below into two sections.

One- and Two-Dimensional Nanogap Biosensors

Advantages and Disadvantages

The key advantage of one- and two-dimensional nanogaps as biosensors is the opportunity for surface-bound molecules to occupy a significant portion of the interelectrode spacing, thereby increasing sensitivity, even enabling single molecule detection [3].

The main drawback of these sensors is the difficulty in ensuring specificity in the measurements. For this reason, the electrodes in the device need to have been functionalized appropriately before sample introduction, to enable the determination of presence

Geometry	Representative diagram	
1-D [Point junction]	<p style="text-align: center;">Top view</p> 	<p style="text-align: center;">Cross-section</p> 
2-D [Co-planar]	<p style="text-align: center;">Top view</p> 	<p style="text-align: center;">Cross-section</p> 
3-D [Contains volume]	<p style="text-align: center;">Top view</p> 	<p style="text-align: center;">Cross-section</p> 

Nanogap Biosensors, Fig. 1 Definition of the three categories of nanogap geometry. Form follows function and each geometry is typically tuned for a different function

or absence of specific molecules quickly and with off-the-shelf electrical meters: IV (current–voltage) meters, electrochemical measurement stations, or patch clamp amplifiers.

From a fabrication point of view, while it is in principle very easy to fabricate one-dimensional nanogap electrodes by electromigration, repeatability of this process can be an issue; conversely, fabrication via electron beam lithography allows for extreme control and repeatability but has slower throughput and higher cost.

Methods of Fabrication

One-dimensional nanogap structures are most commonly realized using electromigration, electron-beam lithography, or a combination of the two techniques. In some cases to shrink gaps to nanometer dimensions, methods such as surface-catalyzed metal deposition or dielectrophoretic transport of metal nanoparticles are used. The latter methods are advantageous as they enable fabrication of many nanogap electrodes at once. On the other hand, electron beam lithography offers the elegant advantage of directly writing

nanoscale features in photoresist, which, incorporated with ion milling, lift-off, or unique metal deposition techniques, allows for the creation of nanogap electrodes. Electromigration is an intriguing technique, because it was an originally undesirable effect often encountered in the electronics industry. As metal lines got smaller and smaller, the relatively higher current densities would cause breaks to occur. These breaks were a result of a momentum transfer between the electrons comprising current and the atoms in the metal lattice [4]. Breaks tended to happen at grain boundaries or along defects in the lattice, and were highly dependent on temperature. Researchers interested in moletronics (a field of study whereby individual molecules are used to perform electronic functions) recognized that this technique could be used to create nanogap electrode junctions. This method offers the ability to get very small point gaps in an inexpensive manner, but there is a notable lack of control. Numerous groups have made strides in creating feedback control for electromigrated nanogap devices in order to reliably control the gap sizes in addition to offering parallel fabrication methods [5].

Another method is the chemical deposition of metal or electroplating which allows for the creation of nanogap devices using traditional fabrication methods that would not normally be capable of nanoscale dimensions. Lastly, a few groups have used either mechanical manipulation of electrodes to create nanogap devices or have mechanically broken wires to create nanoscale gaps [6].

Two-dimensional nanogap devices are similarly realized using metal deposition over a sacrificial linear feature to create a linear nanogap. Oblique metal deposition is a simple way to create a metal overhang in order to form a nanogap [7]. Similar to one-dimensional nanogap sensors, electron beam lithography is a suitable alternative for 2D nanogap fabrication as well as the use of sacrificial layers [8].

Measurement Methods

Direct Current Measurements

The most elementary way of acquiring data from a nanogap biosensor is to apply a DC (direct current) potential, while measuring current flow through the sensor electrodes: When the gap is “open,” no current flows through except for a small tunneling current. When a biomolecule, or other sample, is attached to both the electrodes or otherwise closes the circuit, a larger direct current starts to flow. Specificity in the measurement can be attained by appropriate functionalization of the electrodes’ surface.

In the case of measurements carried out in solution a background current is continuously measured due to ion transport in the solution. Sample attachment to the electrodes is not necessary, and just proximity of charged species (such as DNA molecules) can be detected by measuring increases in current.

Ohm’s law, $I = \frac{V}{R}$, is in most cases sufficient to interpret the data acquired.

The tunneling current by itself can be used to characterize the device and estimate the distance between two electrodes, if the gap is too small to be detected by optical or electron microscopy.

The equation describing the tunneling current between two electrodes is:

$$I = \frac{K_1 A}{s^2} [X^2 \exp(-k_2 s X) - Y^2 \exp(-k_2 s Y)] \quad (1)$$

where $X = \sqrt{\phi - \frac{V}{2}}$, $Y = \sqrt{\phi + \frac{V}{2}}$, $k_1 = 6.32 \times 10^{10} \text{Vs}^{-1}$, $k_2 = 1.025 \text{J}^{-\frac{1}{2}}$, A is the emission area in cm^2 , s is the

electrode separation or distance in \AA , ϕ is the barrier height, and V is the voltage in Volts [9].

Current–Voltage (IV) Measurements

To determine the behavior and characteristics of the molecule under study, or the sensor itself, DC potential can be swept in incremental steps from a negative to a positive potential and the resultant current traces drawn: Asymmetries and/or nonlinearities in the measured IV traces can be used to extrapolate characteristics of the sample under study.

Three-Dimensional Nanogap Biosensors

Advantages and Disadvantages

Three-dimensional nanogap devices are significantly different from one- and two-dimensional nanogaps in fabrication methods, measurement techniques, and in many cases, applications. 3D nanogaps contain a significant volume relative to the nanogap width. This nanocavity allows for both surface bound and free molecules to be monitored. Alternating current measurement techniques are typically used (e.g., dielectric or impedance spectroscopy), which provide data from which meaningful biological parameters can be extrapolated (e.g., binding efficiency of biomolecules, hydrodynamic radius of proteins).

The key advantages of nanogap biosensors using impedance detection techniques (e.g., impedance spectroscopy) are a homogeneous interrogation electric field and the opportunity for surface-bound molecules to occupy a significant portion of the detection volume, thereby increasing sensitivity.

The disadvantages of these devices include complex fabrication, difficulty in coupling to external equipment without introducing parasitic components, and the need for on-sensor reference electrodes and controls.

Methods of Fabrication

Three-dimensional nanogap devices are almost exclusively fabricated by provisionally incorporating a sacrificial layer of nanoscale dimensions, which is used to separate two conducting (metal or semiconductor) electrodes. When this sacrificial layer is removed or etched away, in the final steps of the fabrication protocol, a nanocavity is created.

A common technique is to use a thermally grown oxide layer, which creates a precise thickness by tightly controlling growth conditions. This layer can then later be fully or partially removed with a wet etch.

The materials used to form the electrodes vary, but the significant majority of nanogap devices use either gold or semiconductor electrodes. Gold has numerous advantages over semiconductors for many of these devices, including biocompatibility, ease of surface modification (through thiol-based chemistries), stability (resistance to oxidation), and off-the-shelf availability. One disadvantage of gold is the lack of compatibility with some fabrication processes. In such cases semiconductor electrodes have been used, despite their inherent shortcomings: mediocre conductivity and susceptibility to spontaneous surface oxidation [10].

Overlapping Electrical Double Layers

In order to understand why 3D nanogap devices are so interesting as biosensors, the topic of overlapping electrical double layers must be covered. The electrical double layer at a single electrode surface is discussed in depth by Lyklema [11]. At a charged surface, electromigration of charged species balances with diffusion to create a diffuse layer of counterions that shields the bulk solution from the electrode's potential. Inside a nanogap, the situation is much different: The double layers can occupy the entire volume. The presence of the bulk solution, which can be measured in terms of parasitic capacitance and resistance, is virtually nonexistent. Depending on the solution's properties, the double layers can actually "overlap." This term has been used to describe the situation where the Debye length is greater than the distance between the electrodes. The Debye length (λ_D) is the characteristic length scale of the double layer thickness, and is defined for a symmetric electrolyte:

$$\lambda_D = \sqrt{\frac{\varepsilon_0 \varepsilon R T}{2 z^2 F^2 c_0}} \quad (2)$$

where ε_0 is the permittivity of free space, ε is the relative permittivity of the solution, R is the gas constant, T is the temperature in Kelvin, z is the chemical valence of the electrolyte, F is the Faraday constant, and c_0 is the molar concentration of the electrolyte.

Nanogap Biosensors, Table 1 Debye lengths (λ_D) as a function of electrolyte concentration expressed as molarity (M)

Debye lengths	
c_0 (M)	λ_D (nm)
10^0	0.30
10^{-1}	0.96
10^{-2}	3.04
10^{-3}	9.62
10^{-4}	30.4
10^{-5}	96.2
10^{-6}	304

Table 1 shows theoretical Debye lengths for a symmetric monovalent electrolyte at room temperature. These lengths can range from less than a nanometer to hundreds of nanometers, so for many of the nanogap devices in use and under development, overlapping double layers are a commonality.

To investigate the electric potential that results from the distribution of ions or charges inside of a nanogap, Poisson's equation is used:

$$\nabla^2 \phi = \frac{-\rho_E}{\varepsilon} \quad (3)$$

where ϕ is the electric potential and ρ_E is the charge density:

$$\rho_E = F \sum_i z_i c_i \quad (4)$$

where i represents the ion species. If the nanogap is to be considered a closed system, mass must be conserved, so the following equation is also required:

$$\frac{\partial c_i}{\partial t} = -\nabla \cdot J_i + R_i \quad (5)$$

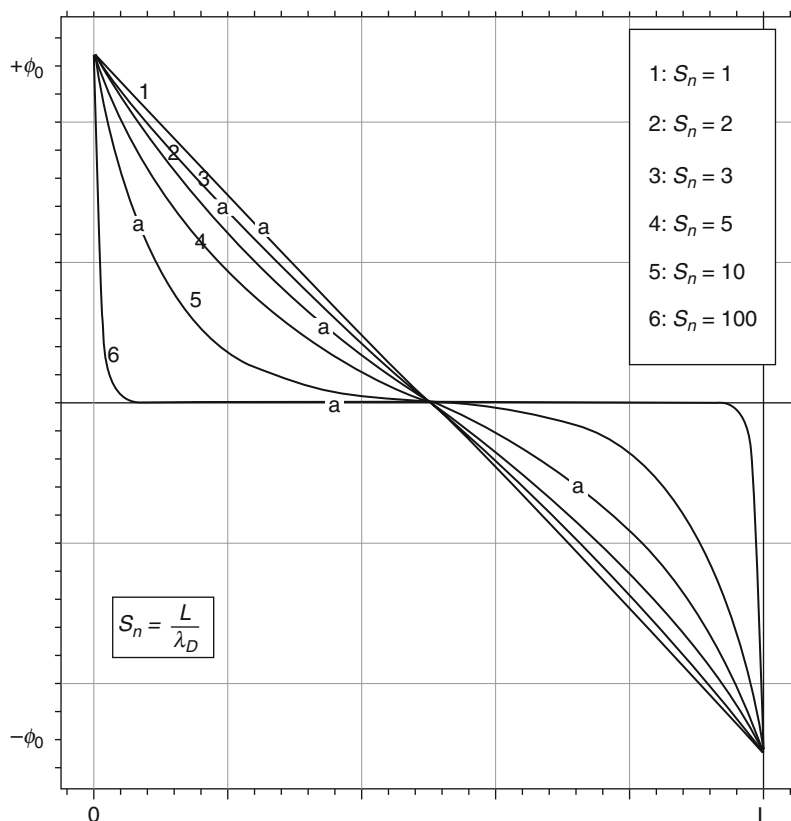
where J_i is the flux density in mol/cm²/s and R_i is the production of species i , which is assumed to be zero at the voltages common for impedance spectroscopy. The flux density can be shown as:

$$J_i = -D_i \nabla c_i - z_i F c_i u_i \nabla \phi \quad (6)$$

where D_i is the diffusion coefficient and u_i is the ion mobility. This represents diffusive flux and electric flux, as represented by the two terms in this equation, respectively. Flux across the boundaries is assumed

Nanogap Biosensors,

Fig. 2 Numerical solution of overlapping double layers. The electrode surfaces are represented by the *left* and *right* plot borders. The *y*-axis is the potential, and the *x*-axis is the distance between the electrodes. The potential distribution is plotted as the dimensionless number, S_n



to be zero, and the potential on the two electrode surfaces are:

$$\phi(x=0) = -\phi(x=L) = \phi_0 \quad (7)$$

The distance between the electrodes, L , is varied. A dimensionless ratio of the distance between the electrodes and the Debye length is defined to be S_n :

$$S_n = \frac{L}{\lambda_D} \quad (8)$$

The results from this solution are shown in Fig. 2. Although this analysis is based on steady-state assumptions, it is a good approximation for most impedance spectroscopy systems, since the relaxation frequency of the electrical double layer is typically faster than the frequency of the current injected by the analyzer.

It is clear from the numerical solution shown in Fig. 2 that when the double layers overlap, the potential drop across the electrodes approaches

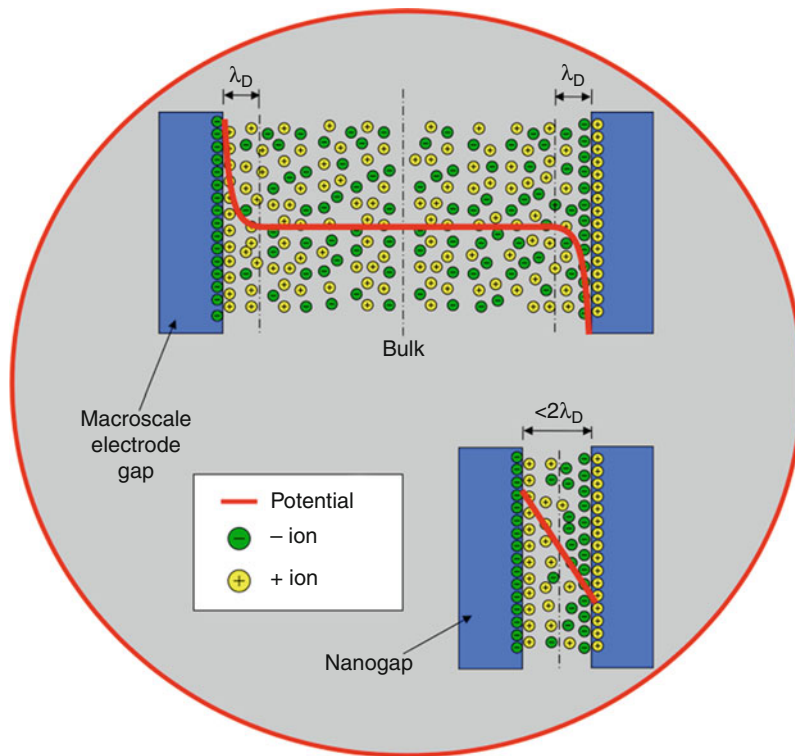
a linear regime. This is significantly different from a macroscopic gap, where virtually all of the potential drops across the double layers and the bulk solution in a macro-gap typically exhibits a null electric field. This is an important point for biosensors, since the entire solution should ideally be subjected to the same conditions. For macroscale electrode sensors, molecules in the middle of the gap will behave much differently from molecules found on or near the electrode surfaces. The uniform field inside a nanogap is a significant advantage for these devices, since all molecules inside of the gap will see the same electrical conditions. The bulk solution is essentially removed, thereby providing, at least in theory, a more homogeneous response. Figure 3 depicts a simple illustration of the electrical double layers in a macroscale gap and a nanogap with overlapping double layers.

Measurement Methods

The detection method is typically achieved by measuring the complex impedance response to either a single excitation frequency or a full range of frequencies, or

Nanogap Biosensors,

Fig. 3 Electrical double layers and overlapping electrical double layers. Schematic of ion distribution inside macroscale gaps and nanogap devices, and resultant potential distribution. λ_D represents the Debye length



spectrum: The impedance spectrum of a system is measured by applying a frequency-dependent excitation voltage and measuring the resulting current. From the impedance spectrum and the geometrical parameters of the system, the complex permittivity or dielectric spectrum is derived.

Kaatze and Feldman have comprehensively reviewed the basic principles of dielectric spectroscopy and the methods in use to measure the dielectric properties of liquid samples over the frequency range from about 10^{-6} – 10^{12} Hz [12].

Conventionally, a small AC (alternating current) potential, $V^*(j\omega)$, swept over a range of frequencies is applied. The electrical current response, $I^*(j\omega)$, is measured and the complex impedance $Z^*(j\omega)$, of the system is:

$$|Z^*(j\omega)| = \frac{V^*(j\omega)}{I^*(j\omega)} = Z_{RE}(j\omega) + jZ_{IM}(j\omega) \quad (9)$$

Here $Z^*_{RE}(j\omega)$ and $Z^*_{IM}(j\omega)$ are the real and imaginary parts of the complex impedance, respectively.

The magnitude and phase angle of the complex impedance are

$$|Z^*(j\omega)| = \sqrt{[Z_{RE}(j\omega)]^2 + [Z_{IM}(j\omega)]^2} \quad (10)$$

and

$$\angle Z^*(j\omega) = \arctan \frac{Z_{RE}(j\omega)}{Z_{IM}(j\omega)} \quad (11)$$

respectively.

Knowing the geometrical and physical parameters of the biosensor, it is possible to model the physical mechanisms of conductivity and polarizability with equivalent circuit analysis. Using networks of resistors and capacitors, the response of the system under controlled conditions can be predicted, and consequently the detection of biological events that cause deviations from the predicted response.

Determination of an appropriate equivalent model requires some a priori knowledge of the chemistry and physics of the system. Because multiple models can be

found to fit any acquired set of impedance data, the choice of the “most appropriate” model is often the result of a complex balance between complexity, in terms of raw number of circuit elements employed, and efficiency, in terms of which chemical and physical phenomena of the system need to be characterized and with what level of sensitivity.

In general, it is of the utmost importance to design a nanogap sensor system where the number of parasitic circuit elements (which tend to increase the complexity of the resultant equivalent circuit model) is minimized, while the impedance response of the system is maximally affected by specific biological events taking place in the nanogap. This is usually achieved by a combination of clever design, appropriate choice of materials, and a reasoned choice of range of frequencies to be measured.

To maximize sensitivity, it is advantageous to have large active areas and small interelectrode distances, maximizing device capacitance. In the particular case where biomolecules are sensed as they bind to the surface and modify the double layer, it is also convenient to minimize the contribution of the bulk solution to the overall impedance, which leads to a sensitivity response that is inversely proportional to the interelectrode distance. Defining ξ as the *quality factor*, the two dependencies can be formulated as follows:

For a generic parallel-plate capacitive sensor, the measured quantity is capacitance, while the sensed parameter is the permittivity ϵ . The two are related by the well-known expression:

$$C = \epsilon \frac{A}{d} \quad (12)$$

where A is the area of the electrodes, d the distance separating them, and ϵ the permittivity of the dielectric between the two plates. To maximize the measured capacitance given a fixed permittivity it is necessary to maximize the area-to-interelectrode distance ratio. Therefore, the first quality factor can be defined as

$$\xi_1 = \frac{C}{\epsilon} = \frac{A}{d} \quad (13)$$

For the specific case of sensors sensitive to surface binding events where bulk contributions (due to the total volume enclosed by two parallel electrodes) are to

be minimized, a second quality factor can be defined as the area-to-volume ratio, i.e.,

$$\xi_2 = \frac{A}{V} = \frac{A}{A \cdot d} \quad (14)$$

The combination of these two factors yields:

$$\xi = \xi_1 \xi_2 = \frac{A}{d^2} \quad (15)$$

This dimensionless parameter can be used to quantitatively compare the efficiency of different nanogap sensors, as well as to optimize the design of new sensors.

Although alternating current methods are the most common methods used with 3D nanogap biosensors, changes in direct currents (and similarly, shifts in IV curves) have been used [13].

Key Research Findings

Some key pieces of literature are briefly listed below. This list is not meant to be exhaustive, but rather illustrative of the most significant research findings achieved using nanogap sensors, and a starting point for further exploration.

One- and Two-Dimensional Nanogap Biosensors

One- and two-dimensional nanogap devices have been used mostly in proof-of-concept studies, to determine the ability to detect and to a limited extent, characterize nucleic acids (DNA, RNA) binding events, and presence of other biologically relevant biomolecules such as proteins and antibodies. Chen et al. detected specific target DNA concentrations of as low as 10 picomolar identifying single-base-pair mismatches [14]. Liang et al. observed electrical signals caused by 1.1 kilobase-pair (kbp) double-stranded (ds)-DNA passing through the gap in the nanogap detectors with a gap equal to or less than 13 nm [15]. More recently, Cingolani et al. demonstrated that hybridization events are detected by the discrete increase in conductivity when the target AuNC–DNA conjugates bridge the electrode gap [16].

The electrical detection of other biomolecular interactions has also been demonstrated. For example, Haguët et al. have detected the capture of antibodies as

well as the specific interaction between the biotin and the streptavidin molecules [17]. Subsequently, the same group has shown that the same method can be extended to the detection of immunoglobulin G (IgGs) in serum using protein probes [18].

Three-Dimensional Nanogap Biosensors

Three-dimensional nanogap resistive sensors with gaps of 5, 10, and 15 nm have been used to detect biotin–streptavidin binding events by measuring peak increases in direct current [13].

Single-frequency AC impedance measurements have been used successfully with three-dimensional nanogap sensors adopting metal electrodes and integrated reference sensing on-chip, for the detection of thrombin, a blood clotting factor, and to investigate its binding properties to a specific antibody and RNA-aptamer. A similar device was used to study the interaction of the Rev peptide with a corresponding RNA anti-Rev aptamer in concentration spanning the range of 100 nM–2 μ M. The parasitic effect of the double layer was significantly reduced by using measurement frequencies of 800–1,280 MHz [19].

AC impedance spectroscopy has also been used to specifically detect the presence of thrombin in solution by looking at impedance spectra recorded using an 20mV AC signal swept over a frequency range of 10 Hz–100 kHz [20].

Similarly, hybridization of poly-A nucleotides to immobilized poly-T between nanogap electrodes has been shown to result in a measurable change in capacitance over a frequency range of 100 Hz–10 kHz [21].

Future Directions

At this point in time, the development of nanogap biosensors is still in the research phase, i.e., there are no commercially available nanogap biosensors. However, the promise of sensitive, label-free devices with small footprints is alluring. Great strides have been made in terms of nanogap manufacturing and detection. Future work is needed in scale up of manufacturing, as few groups have achieved repeatable large-scale fabrication. This will be essential in developing sensors in quantities at low cost.

For applications relating to nucleic acid determination and sequencing, increases in sensitivity and

specificity are necessary for these biosensors to compete with existing industry standard sensors which offer fluorescence-based readouts.

Specific label-free, all-electrical nucleic acid detection can still be valuable for diagnostic point of care devices, although additional work is needed to enable multiplexed detection which will allow the analysis of a panel of biomarker sequences, rather than a single short target sequence.

Similarly useful will be nanogap devices that enable the detection of specific proteins in serum or blood, forfeiting the cost and complexity associated with traditional antigen–antibody assays such as ELISAs. Devices aimed at point-of-care testing are particularly exciting because this application capitalizes on the inherent advantages of a nanogap: small volumes, small footprint, low power consumption, and label-free, all-electrical detection. The future will likely see many point-of-care biosensors using electrical-based detection, and perhaps nanogap technology can play an important role.

Cross-References

- ▶ [Biosensors](#)
- ▶ [Dry Etching](#)
- ▶ [SU-8 Photoresist](#)
- ▶ [Wet Etching](#)

References

1. Liang, X., Chou, S.Y.: Nanogap detector inside nanofluidic channel for fast real-time label-free DNA analysis. *Nano Lett.* **8**, 1472–1476 (2008)
2. Im, H., et al.: A dielectric-modulated field-effect transistor for biosensing. *Nat. Nanotech* **2**(7), 430–434 (2007)
3. Gu, B., et al.: Nanogap field-effect transistor biosensors for electrical detection of avian influenza. *Small* **5**(21), 2407–2412 (2009)
4. Chiras, S., Clarke, D.R.: Dielectric cracking produced by electromigration in microelectronic interconnects. *J. Appl. Phys.* **88**, 6302–6312 (2000)
5. Fernandez-Martinez, I., Gonzalez, Y., Briones, F.: Parallel nanogap fabrication with nanometer size control using III-V semiconductor epitaxial technology. *Nanotechnology* **19**(27), 275302 (2008)
6. Higuchi, Y., et al.: Application of simple mechanical polishing to fabrication of nanogap flat electrodes. *Jpn. J. Appl. Phys.* **45**, L145–L147 (2006)
7. Li, T., Hu, W., Zhu, D.: Nanogap electrodes. *Adv. Mater.* **22**(2), 286–300 (2009)

8. Lazzarino, M., et al.: Twin cantilevers with a nanogap for single molecule experimentation. *Microelectron. Eng.* **83** (4–9), 1309–1311 (2006)
9. Stroschio, J.A., Kaiser, W.J.: Scanning tunneling microscopy. In: Celotta, R., Lucatorto, T. (eds.) *Methods of experimental physics*, vol. 27. Academic, San Diego (1993)
10. Ionescu-Zanetti, C., et al.: Nanogap capacitors: sensitivity to sample permittivity changes. *J. Appl. Phys.* **99**(2), 024305 (2006)
11. Lyklema, J.: *Fundamentals of Interface and Colloid Science*, p. 768. Academic, London (1995)
12. Kaatz, U., Feldman, Y.: Broadband dielectric spectrometry of liquids and biosystems. *Meas. Sci. Technol.* **17**(2), R17 (2006)
13. Jang, D.Y., et al.: Sublithographic vertical gold nanogap for label-free electrical detection of protein-ligand binding. *J. Vac. Sci. Technol. B* **25**(2), 443–447 (2007)
14. Chen, F., et al.: Electrochemical approach for fabricating nanogap electrodes with well controllable separation. *Appl. Phys. Lett.* **86**(12), 123105 (2005)
15. Liang, X., Chou, S.Y.: Nanogap detector inside nanofluidic channel for fast real-time label-free DNA analysis. *Nano. Lett.* **8**(5), 1472–1476 (2008)
16. Maruccio, G., et al.: A nanobiosensor to detect single hybridization events. *Analyst* **134**(12), 2458–2461 (2009)
17. Haguët, V., et al.: Combined nanogap nanoparticles nanosensor for electrical detection of biomolecular interactions between polypeptides. *Appl. Phys. Lett.* **84**(7), 1213 (2004)
18. Marcon, L., Stievenard, D., Melnyk, O.: Characterization of nanogap chemical reactivity using peptide-capped gold nanoparticles and electrical detection. *Bioconjug. Chem.* **19**(4), 802–805 (2008)
19. Schlecht, U., et al.: Detection of Rev peptides with impedance-sensors – Comparison of device-geometries. *Biosens. Bioelectron.* **22**(9–10), 2337–2340 (2007)
20. Mannoor, M.S., et al.: Nanogap dielectric spectroscopy for aptamer-based protein detection. *Biophys J.* **98**(4), 724–732 (2010)
21. Choi, Y.K., et al.: Sublithographic nanofabrication technology for nanocatalysts and DNA chips. *J. Vac. Sci. Technol. B* **21**(6), 2951–2955 (2003)

Nanogrippers

Peter Bøggild
DTU Nanotech – Department of Micro- and Nanotechnology, Technical University of Denmark, Lyngby, Denmark

Synonyms

[Nanotweezers](#)

Definition

Nanogrippers are micro- or nanoactuators equipped with high-precision dual end-effectors, capable of applying controllable, opposing forces on nanoscale objects, and thereby allowing these to be lifted free from a surface, held in a well-defined position, moved to another position, and placed in a controlled way.

Overview

Nanogrippers allow structures with at least one dimension being of order 100 nm or smaller to be manipulated in a predictable manner. Such grippers consist of one or more actuators, which determine the gap between two or more end-effectors. The actuators are operated through variations of a physical quantity, which is typically an electrical current or voltage, but might also be the temperature, a magnetic field, or even light. While there is a wide selection of actuators available for conventional microelectromechanical systems (MEMS), nearly all nanogrippers which have been demonstrated in practical manipulation of nanostructures so far have either used electrostatic or electrothermal actuation. These actuation principles combine simplicity and compactness of design with predictable actuation behavior, and are well suited for downscaling. The nanogripper itself could in principle have any size, as long as this does not prevent it from accessing and manipulating nanostructures in a controlled manner. There are, however, several reasons why grippers shown to be effective in manipulation of nanostructures tend to have an overall size significantly smaller than microgrippers; large grippers tend to block the line of sight of the microscope which is necessary to align the gripper to the object. By using conventional contact photolithography, the lateral resolution with which the end-effectors – the parts of the gripper touching the sample – can be shaped, is limited to around 1 μm . While high-precision fabrication techniques such as electron beam lithography and focused ion beam lithography can be used to achieve more well-defined end-effectors, nanogrippers have been realized with multiwalled carbon nanotubes acting as both electrostatic actuators and end-effectors. Nanogrippers can be used to manipulate objects which are difficult or time-consuming to approach, touch and move, or to place them in special positions. While

manipulation is inherently slow compared to conventional parallel fabrication techniques in terms of mass-fabrication, gripper-based nanoassembly has shown to be a flexible, rapid prototyping technique with a short turn-around time for small scale production of nanodevices. Robotic manipulation using microgrippers has for instance been used to break off individual carbon nanotubes from their initial positions, and mount these onto cantilever probes for high aspect ratio atomic force microscopy. Other applications for nanogrippers include using the dual tips as a two-probe for fast electrical characterization, for assembly of nanostructures using dielectrophoresis or to arrange nanoscale objects for subsequent structural analysis with a Transmission Electron Microscope.

Design Considerations

The choice of actuator principle has considerable influence on the size, shape, actuation range, accuracy, generated force and thereby the applicability of a gripper tool, and for nanoscale gripping the options are in practice much fewer as compared to microgrippers. Piezoelectric actuators made from ceramic materials can generate very precise motion with a high force, but are notoriously difficult to integrate in a small tool. A large number of actuation principles have been proposed for nanogrippers and nanotweezers, however, nearly all attempts so far resulted in working devices have been made with either electrostatic or electrothermal actuation. Both of these principles provide a reasonable compromise between ease of fabrication and device integration, high reliability, high resolution, a large generated force, and a small overall footprint. The discussion will mainly focus on these two, as a treatment of the numerous other potential actuation principles and their role in hypothetical nanogrippers is beyond the scope of this chapter.

With a very few exceptions all nanogrippers perform their motion in a single plane. Conventional bulk micro- and nanofabrication is highly suited for making in-plane actuators for electrostatic and electrothermal actuation, which is very convenient for miniaturization. Where electrostatic actuation uses Coulomb attraction between oppositely charged capacitor plates to generate a force, electrothermal actuation arises from exploiting the thermal expansion of materials, when heated up by an electrical current. In a review

article, Sahu [1] discuss the most relevant actuator options for micro and nanomanipulation, and lists the key design characteristics, which are adopted for nanogrippers in the following:

Size – small size to allow interaction with nanoscale components. An important trade-off for all types of grippers is to make the actuators, or at least the end-effectors, small and precise enough for delicate manipulation with a good line-of-sight between the microscope objective lens and the manipulated object, while at the same time providing a sufficiently solid structure to overcome surface forces between the object and the substrate or in some cases to break off the object from its initially fixed position. Many designs fail to be ever used in practice because they either make it extremely difficult to observe the manipulation process due to their bulkiness, or because the force they provide is too small to detach the objects of interest from a surface in a controlled way. Finally, larger actuators often suffer from out-of-plane bending of up to tens of micrometers, requiring constant realignment, and can lead to damage of the samples. This is not necessarily a problem for manipulation of microscale objects, but can make nanomanipulation extremely difficult and cumbersome.

Range of motion and applied force – depending on the size of the objects, as well as how much force is required to release them from their initial position, a proper balance between range of motion (actuation range) and gripping force must be chosen – generally, this is a trade-off given by the choice of actuation principle and design.

Actuation resolution – the deflection per input voltage/current should be small enough to accurately administer the force, for example, in the range 0.1 nm/mV for a voltage source. Since the deflection is basically quadratic in drive signal, both for electrothermal and electrostatic nanogrippers, the resolution depends on the deflection.

Contact and force-sensing – to avoid damage to samples, and to allow detection of the nanogripper, contact and force-sensing are very attractive features. This is best done with either capacitive or piezoresistive readout, which unfortunately both add to the complexity of the devices, and therefore make design of nanogrippers even more challenging. In practice, analysis of microscope images often serves as a convenient way to both detect

contact and to infer the applied forces. Knowing the spring constant of a gripper, the gripping force can be estimated from the observed deviation from its known mechanical equilibrium position. It is still very challenging to detect forces in a range that is relevant for manipulation of nanoscale objects [22], in particular if such sensitive force-sensing components have to be integrated with a minute nanogripper device. For biological nano- and microscale objects, the relevant force range is pN to nN. NanoNewton force resolution for microgrippers can be achieved by capacitive read-out [2]. Manipulation of hard, rigid objects like silicon nanowires or multiwalled carbon nanotubes may involve forces of order nN, while for picking up (detachment or breaking) of such structures forces in the μN range are relevant [3].

Power – a low power dissipation is beneficial for avoiding damage to the samples, and in the case of electrothermal grippers also to protect the grippers themselves. High peak temperatures in such grippers may cause deformation or failure not just in polymer but also metallic or silicon actuator structures. Careful topological design of the layout [4], or proper choice or combinations of materials [5] can significantly reduce the peak temperature at the end-effector and the overall heat dissipation to the environment.

Fabrication and material selection – Both for electrostatic and electrothermal actuators, the choice of material and fabrication strategy can have a great impact. While silicon is often chosen as a default material due to the availability of proven fabrication methods and its excellent mechanical properties, there are sometimes more suited alternatives. For instance, the much larger thermal expansion coefficient of the polymer SU-8 in combination with a metallic heating layer can provide large actuation stroke at far lower temperatures than pure metallic or silicon actuators [5]. With photolithography, it is possible to fabricate end-effectors capable of manipulating nanoscale objects with dimensions down to around 100 nm [4, 6, 7]. For entering the sub-100 nm range, electron beam lithography (EBL) is an obvious choice due to its order-of-magnitude improvement of the resolution compared to photolithography. Alternatively, focused ion beam milling can be used to fine-tune end-effectors [8].

Micro- and nanoscale surface interaction – while van der Waals interactions comprise the most consistent and ubiquitous type of surface force, several other kinds of interactions interfere with nanogripper operation depending on the environment and the involved surfaces. Such effects are mostly but not always unwanted. In humid conditions, capillary forces are the most prominent among a wide range of surface/contact interactive effects [9], while electrostatic forces due to charging by the electron beam can lead to strong stiction and unpredictable motion and behavior of nanostructures for instance inside a scanning electron or ion microscopes where the constant bombardment with charged particles can lead to strong charge accumulation. In both focused electron and ion microscopy, the intense beam can also directly lead to “gluing” or fusing of objects, by deposition of carbonaceous material.

Scaling, Surface Forces, and Pick and Place

The concept of nanogrippers is based on the presumption that manipulation using opposing forces is somehow useful analogous to human hands, tweezers, and grippers in interacting with the macroscale environment. While robotic grippers comprise one cornerstone in the industrial civilization with more than a million industrial robots in operation worldwide, many of which are using gripper-like end-effectors, it is not at all obvious that manipulation can ever be as important on a scale that is six orders of magnitude smaller. In the macroscale world, mechanical forces and gravity play an important role, simply due to the volumetric size and mass of the objects compared to the surface area. A stone is picked up using a hand, while sand or dust is far easier to wipe up with a wet cloth, due to capillary forces. The dominant role of surface forces in nanomanipulation can be illustrated by considering a sphere of radius R in close contact with a flat surface, assuming these to be separated by a small distance D . Both the van der Waals force [9]:

$$F_{vdW} = \frac{A_H R}{6D^2} \quad (1)$$

and the capillary force

$$F_{cap} = 4\pi\gamma_L R \cos \theta \quad (2)$$

for this system are linearly dependent on R . The Hamaker constant $A_H \approx 10^{-19} J$ represents the combined van der Waals interactions acting between the atoms in the sphere and the surface, $\gamma_L = 0.072 \text{ J/m}^2$ is the surface tension of water, and θ is the contact angle. Comparing the magnitude of these forces to the gravitational force, $F_g = 4\pi R^3 \rho g/3$, where ρ is the density and g is the acceleration of gravity, the surface forces exceed the gravitational force for objects smaller than roughly 10^{-3} m . While for macroscopical objects, surface roughness generally decreases the effective contact area and thus the surface forces by orders of magnitude, surfaces and nanostructures encountered in nanomanipulation are ever so often in much more intimate contact – either because the nanostructures are small compared to the scale of the surface features, or because the surfaces involved are atomically smooth, such as the surface of a carbon nanotube. Even in the extreme case of the effective contact area being reduced by three orders of magnitude due to surface roughness effects, the surface forces still dominate over the gravitational forces for objects in the submicron scale.

If the object can be picked up by adhesion forces using just a single end-effector, it appears that there is not much use of dual end-effectors; it is frequently observed in practice that the object remains attached to one of the end-effectors after the gap has been opened. The macroscale analogy is picking up sand or salt using a pair of tweezers as compared to using a moist cloth. In this case, the capillary forces are clearly more effective than a mechanical force. However, since the surface forces can keep nanoscale objects strongly fixed to a surface, the possibility of applying a *larger* mechanical force may facilitate picking up the objects [10]. For nanowires or carbon nanotubes grown directly on a substrate, the dual end-effector is highly suited; epitaxial nanowires and carbon nanotubes typically have a maximal strength of order 10 GPa, with defect-free single-walled carbon nanotubes closer to 60 GPa, significantly larger than for instance hardened steel (1–2 GPa). A dual end-effector can both apply the mechanical force necessary to break off the object, maintain it in a fixed position

between the jaws [10], and finally deliver it with the correct orientation at the target device or position. Dual end-effectors are also convenient for electrical measurements of an object held between the jaws [11].

Environment and Microscopy

As seen from the above discussion, the performance characteristics of a nanogripper depend on a complex interplay between the material properties, actuation principle, size, and design of the gripper, but also the application scenario in which the gripper are to be used. A key aspect of a nanogrippers usability is indeed the manipulation environment. This is determined by the environmental tolerances of the samples to be manipulated, the requirements, and limitations of the microscopy technique needed for monitoring the process, and finally the physical space required to host the nanogripper, samples as well as the often quite bulky positioning manipulators. Issues related to environment and microscopy are discussed in the following.

Nanomanipulation may be performed in vacuum, ambient, or liquid conditions. A liquid environment is often chosen for manipulation of biological samples, and in this case the liquid is typically a saline solution, which is electrically conducting. The biological objects are typically damaged at elevated temperatures, which nearly rules out electrothermal actuators as these invariably dissipate heat into the surrounding liquid when immersed. Similarly, most electrostatic actuators reported on today require application of voltages larger than a few volts, which leads to hydrolysis and consequently bubble formation in the liquid, which prevents any form of nanomanipulation. Finally, the liquid environment makes optical microscopy the only viable option, which even at the very best conditions makes alignment or even visualization of immersed sub-100 nm structures extremely difficult. The three-dimensional nature of nanogrippers excludes ultrahigh resolution monitoring possible with scanning probe microscopy, as such techniques can only image and manipulate 2D and quasi-2D surfaces. Although atomic force microscope probes with integrated carbon nanotube tweezers have been fabricated by Akita and coworkers [12], a combination of scanning electron microscopy and a separate scanning probe was used in a later demonstration [13]. There has

thus been little progress in gripper-based manipulation of nanostructures in a liquid environment so far; here, optical tweezers is a more powerful method alongside other far-field manipulation strategies such as dielectrophoresis [22].

Ambient or atmospheric conditions allow both electrostatic and electrothermal grippers to operate in a useful way. Just as for a liquid environment, optical microscopy is the only practical option, providing imaging of wire-like structure (such as carbon nanotubes) with sizes down to roughly 100 nm and in some cases tens of nanometers. In ambient conditions, and in particular in very humid conditions [9], water films will reside on most surfaces and lead to strong capillary forces as discussed above. These surface forces can make picking and placing of objects difficult.

Vacuum conditions allow electron microscopes to be used, and this is a strong advantage. The scanning electron microscope allows imaging of nanostructures with lateral dimensions down to 1 nm (such as single-walled carbon nanotubes), while providing sufficient space inside the specimen chamber for the manipulation stages required to move the nanogripper. Capillary forces are largely eliminated, and the absence of air and water makes the behavior of the gripper more predictable. However, the constant and concentrated irradiation with electrons may damage the samples, just as particularly biological samples cannot stay alive in vacuum. Scanning electron microscopes also tend to deposit thin layers of carbonaceous contamination, which can influence the sample, change the morphology, or even bind the samples in unwanted positions. A consequence of this is sometimes that the manipulation has to be conducted with a strict control of the imaging, to avoid excessive contamination. This electron beam induced contamination (EBiD) can be used to an advantage, providing a handy “glue,” which can be applied simply by focusing the electron beam at a junction until a sufficient amount of carbonaceous material has mechanically connected the two objects [4, 6, 14].

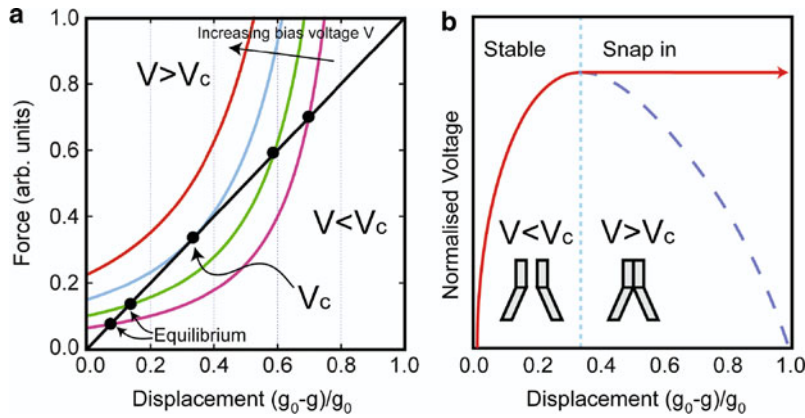
Other types of electron microscopes are relevant to mention. The transmission electron microscope provides atomic resolution in ultrahigh vacuum conditions, but allows far less than a cubic centimeter of sample space. It is possible to manipulate a single sharp tip inside a TEM using highly specialized nanomanipulators, but so far, nanogrippers as such have not been used in situ TEM.

The environmental scanning electron microscope (ESEM) allows up to about 10 mbar gas to be inside the sample chamber, with a moderate degradation of the imaging resolution. The ESEM is a very flexible system for nanogripper operation, providing both high resolution monitoring, a certain degree of control of the environment and sufficient space to operate one or more nanogrippers in situ. The dual beam FIB/SEM combines a scanning electron microscope with the focused ion beam system, which is highly convenient for combination of nanogripper manipulation with in situ shaping of nanostructures and nanotools, deposition of metal electrodes, or fastening glue by EBiD. To date, most gripper-based manipulation of nanoscale objects has been done inside some type of scanning electron microscope.

Finally, the nanomanipulation or nanorobotic system used to perform the manipulation is a key factor in the success of a nanogripper. A deep discussion of nanomanipulation systems is beyond of the scope of this text, however, a few, highly successful examples of platforms for complex manipulative operations include for instance the semiautomatic nanomanipulation system by Eichhorn, Fatikow, and coworkers [14] with integrated depth detection using advanced three-dimensional scanning electron microscopy and image recognition technology, and commercial, high-resolution multi-manipulator workstations from the US company Zyvex Corporation and German companies Klocke Nanotechnik GmbH, Kleindieck Nanotechnik GmbH, and Smaract GmbH.

Electrostatic Nanogrippers

The electrostatic actuator is very widespread in MEMS, due to its convenient simplicity and ease of integration with standard silicon fabrication processes. It has a fast response time, provides nanometer-precision motion, consumes very little power, and thus dissipates virtually no heat to the surroundings. High-precision capacitive position sensors can be made with the same fabrication process and a similar capacitor layout, which means that integration of capacitive position sensors and actuators is natural and straightforward using bulk microfabrication. Nevertheless, some of the simplest, smallest, most well-known electrostatic nanogrippers are constructed through manually nanoassembly, rather than using bulk microfabrication.



Nanogrippers, Fig. 1 (a) Force against displacement showing the mechanical equilibrium points where the linear elastic restoration force is balanced by the electrostatic force. At the critical voltage V_c , there is a singular equilibrium point, beyond

which there are no stable solutions; the actuator thus experience “pull-in” or “snap-in”. (b) the normalized bias voltage versus displacement illustrates the sudden snap-in event at V_c

Electrostatic Actuation

Electrostatic actuators are variable capacitors with air, liquid, or vacuum as a dielectric. A voltage difference V applied to two plates leads to a net electrostatic force acting between these. Depending on the design, the force F_{el} is translated into motion of two or more end-effectors which can close around an object and hold it, thus constituting an electrostatic gripper. The electrostatic gripper comes in mainly two varieties, the plate-type or the comb-type actuator. In the generic plate-type actuator the capacitor gap is variable, while in the generic comb-type actuator the effective (overlap) area is varied.

For a plate capacitor with area A and gap size g , and absolute permittivity ϵ the capacitance is $C = \epsilon A/g$. The capacitor plates are assumed to be kept in position by a spring with the spring constant k . With an applied voltage V across the capacitor plates the change in potential energy ΔU for the spring-capacitor-voltage source system [15] is

$$\Delta U = \frac{1}{2}k(g - g_0)^2 - \frac{1}{2}CV^2, \tag{3}$$

where g_0 is the initial gap size. In equilibrium the total potential energy is minimum, $\nabla(\Delta U) = 0$, and thus for actuation with one degree of freedom actuation the balance equation becomes:

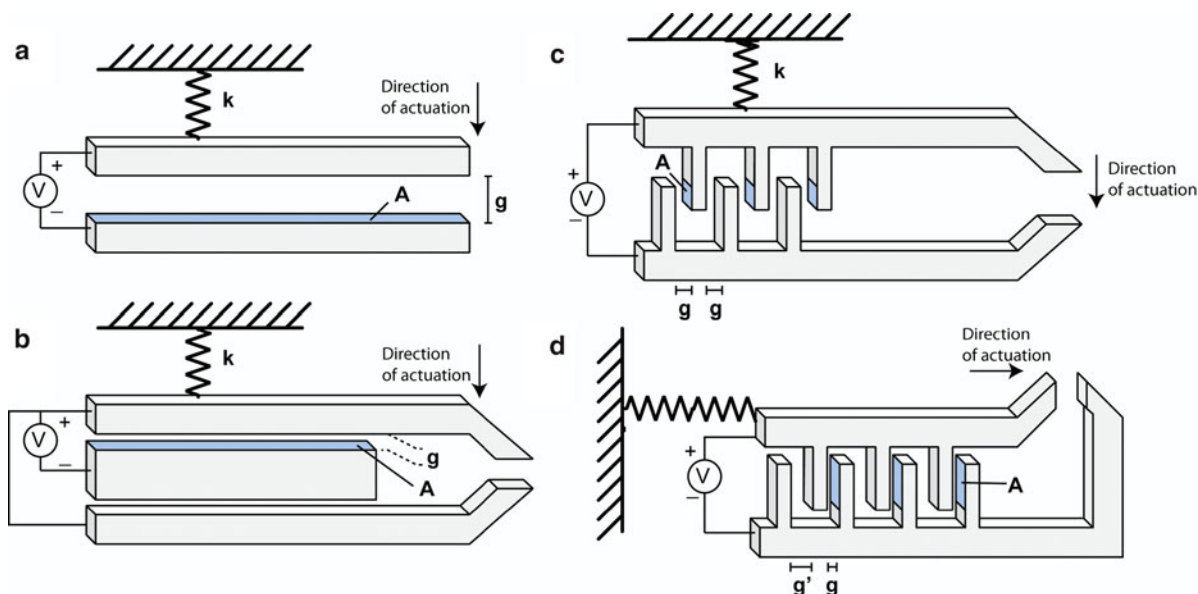
$$g - g_0 = \frac{1}{2} \frac{V^2}{K} \frac{\partial C}{\partial g} = - \frac{\epsilon AV^2}{2kg^2}. \tag{4}$$

This corresponds to a spring force $F_m = -k(g - g_0)$ balancing the electrostatic force $F_e = \epsilon AV^2/2g^2$, which is quadratic in V . As the gap is decreased the magnitude of the capacitance gradient increases. At a critical voltage V_c corresponding to the gap size $g = 2g_0/3$, the spring force can no longer balance the electrostatic force, and the system becomes unstable, which is illustrated in Fig. 1. The actuator will experience a “pull-in,” where the gap size is reduced to zero, possibly leaving the system in this state, even after the voltage has been reduced to below V_c .

The instability can be avoided with the comb-type design, where the capacitors are equipped with interleaved, sliding pins. With this type of actuator, the capacitance is changed through the area rather than the gap size, as shown in Fig. 2. The force generated by comb-type actuators scales with the number of gaps, so higher forces can be achieved by increasing the area of the actuators. The displacement is thus [1]:

$$\Delta x = N \frac{\epsilon h V^2}{kg^2}, \tag{5}$$

where N is the number of gaps, h is their height, and the area between each capacitor pair is $A = h\Delta x$, see Fig. 2. Comb-drives are generally larger and more bulky than plate capacitor drives, and have only been used in a few instances for nanoscale manipulation. The operating voltage for electrostatic nano- and microgrippers ranges from a few [12] to several



Nanogrippers, Fig. 2 Different types of electrostatic actuators for nanogrippers. (a) Basic plate capacitor, where the gap size is changed upon increasing the bias voltage V , while the area A is constant. The movable part is thought to be connected to the device via an elastic support, here represented by a spring with a spring constant k . (b) A fixed center electrode the two end-effectors to remain at the same potential, thus preventing current from running through the gripped object. (c) A comb-drive used

to close two end-effectors. In this configuration, the overlap area A between the two arrays of capacitors is varied, rather than the gap size g , which avoids the pull-in instability. (d) Comb-like drive where the direction of motion of the movable part is lateral rather than vertical, with capacitor gaps g' and g . This configuration suffers from the pull-in instability, since it is essentially of the parallel plate type

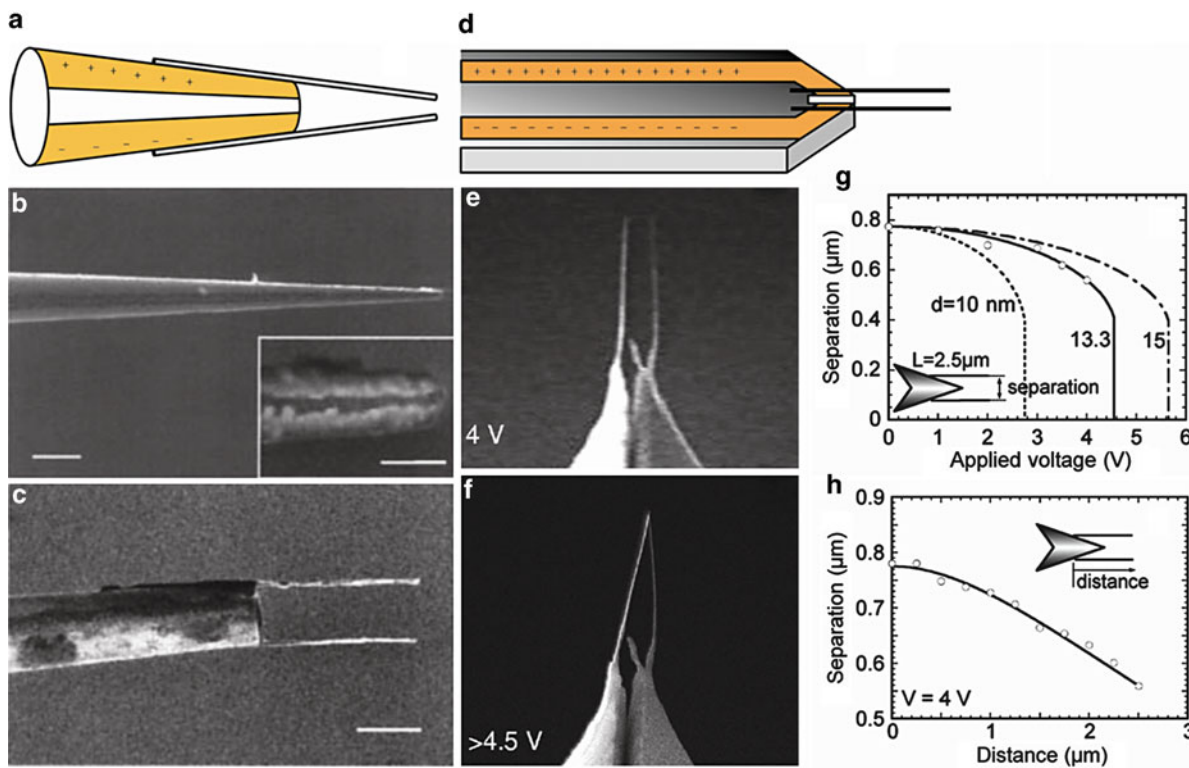
hundred volts, depending on the design, the size, and the requirements, with a resolution of order 1 nm.

Nanotube Nanotweezers

Although microgrippers of different types were developed already in the 1990, the first and most spectacular type of gripper tool capable of manipulating nanostructures is the nanotube nanotweezers [11]. This gripper is comprised of multiwalled carbon nanotubes glued to the sides of a glass microcapillary under an optical microscope. On each side of the glass capillary, a metal electrode is evaporated to provide electrical contact to the carbon nanotubes. By applying a moderate voltage the nanotubes were attracted toward each other, until pull-in at 8.5 V where the gap g was roughly equal to $2g_0/3$. The authors demonstrated picking up of 500 nm diameter fluorescent nanobeads. As an added feature, the good electrical properties of the carbon nanotubes allowed electrical characterization of the held object. Following this approach, Akita et al. [12] integrated carbon nanotubes with a atomic force microscope probe, which was split in two electrically separated parts using focused ion beam milling, see Fig. 3.

Batch-Fabricated Grippers

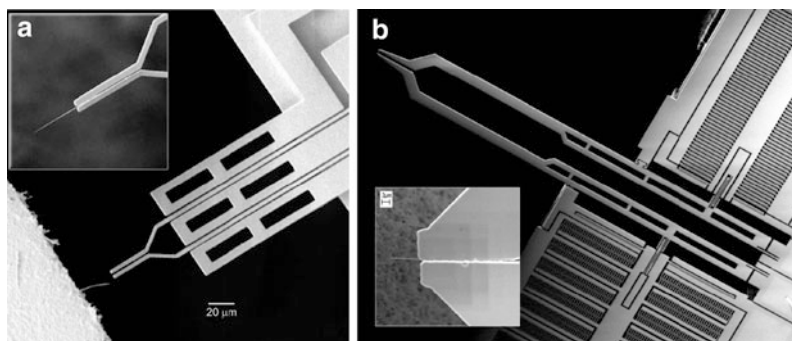
While an impressive range of electrostatic microgrippers has been fabricated using bulk microfabrication, the majority are not designed to engage in manipulation of nanoscale objects, as this require very precisely manufactured end-effectors as well as high resolution and precision of motion. Nelsons and coworkers [16] developed a range of advanced, highly integrated, and very well-characterized electrostatic grippers, which has since been commercialized by the Swiss company FemtoTools, see Fig. 4b. These grippers have a well-functioning force feedback sensing, have been tested thoroughly in air, liquid, and SEM, and are fully integrated with measurement and control electronics and computer interfaces. Although primarily designed for microscale manipulation, these are capable of nanoscale manipulation, as shown in the inset in Fig. 4b. Yamahata and coworkers reported a precision-manufactured electrostatic nanogripper with an integrated displacement sensor intended for biomolecules, and demonstrated trapping of bundles of DNA molecules using dielectrophoresis rather than mechanical



Nanogrippers, Fig. 3 (a) Kim and Liebers nanotube nanotweezers [11] with multiwalled carbon nanotubes glued to the sides of a glass capillary, where an electrode layer has been deposited on each side. (b) Magnified view of the capillary after metal deposition. (c) The nanotweezers after mounting of the nanotubes. (d) Nanotweezers made by focused ion beam milling a metallized cantilever to separate the electrode into two [12].

(e) Two nanotubes attached to the tips at 4 V. (f) At a bias voltage of 4.5 V, the nanotubes pulls in. (g) The separation versus applied voltage, showing pull-in occurring at different voltages depending on the initial separation. (h) Variation of gap separation for different initial length of the nanotubes, all at a bias voltage of 4 V

N

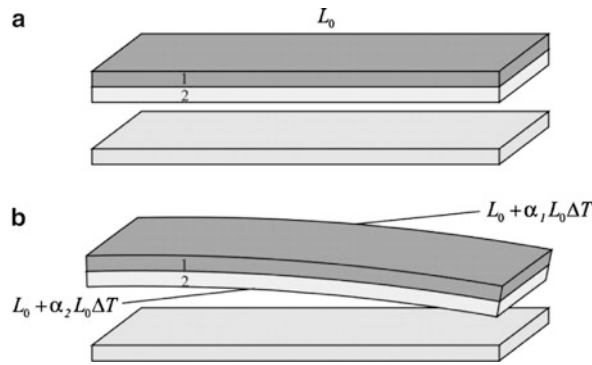


Nanogrippers, Fig. 4 Two examples of electrostatic actuators; (a) a plate capacitor driven microgripper with both open and close capability [10]. The inset shows the gripper holding a silicon nanowire with a diameter of 200 nm and (b) Gripper

with lateral comb-drives and integrated capacitive position read-out. The inset shows a similar gripper manipulating a silicon nanowire with a length of approximately 6 μm (Courtesy of V. Eichhorn and Femtotools)

gripping [22]. Mølhave [10] reported manipulation of silicon nanowires inside a scanning electron microscope using a plate capacitor gripper with five legs, of which

two served as thin flexible actuator arms, and three as rigid electrodes, allowing the gripper to both open and close compared to its neutral position, see Fig. 4.



Nanogrippers, Fig. 5 Bimorph actuator used as a gripper. Two layered beams with different thermal expansion coefficients α_1 and α_2 and initial length L_0 will deflect sideways

upon a uniform temperature change ΔT . This could be used to form a compact nanogripper

Thermal Grippers

The volume of most solids will increase with temperature, and this effect can be exploited in several ways to provide accurate, yet forceful motion. The thermal expansion of for instance a rod with length L as the temperature is changed from T_s to T_h can be written

$$L(T_h) = L(T_s)(1 + \alpha\Delta T), \Delta T = T_h - T_s \quad (6)$$

where α can often be treated as a constant, nearly independent of temperature. The thermal expansion coefficient varies between about 1–5 for Si, Ge, and C, 5–30 ppm/K for cleanroom relevant metals to 52 ppm/K for a polymer material such as SU-8. A temperature increase of 100 K thus only lead to a 0.5% expansion of silicon, hence, a key design feature is amplification of this expansion to a useful actuation range. A basic strategy is to obtain a different thermal expansion for two objects, which are connected mechanically in a way that amplifies the deformation.

In a bimorph actuator, this can be done by coupling two layers of material with different thermal expansion coefficient α_1 and α_2 to each other in a single beam as illustrated in Fig. 5. A homogeneous heating of the structure will then cause the beam to deflect towards the side with the smallest expansion. The bimorph actuation principle is well suited for downscaling, as demonstrated in the work by Sul and coworkers [17], which was capable of delivering 1 μ N at a temperature of 430 K. The actuator was realized by depositing

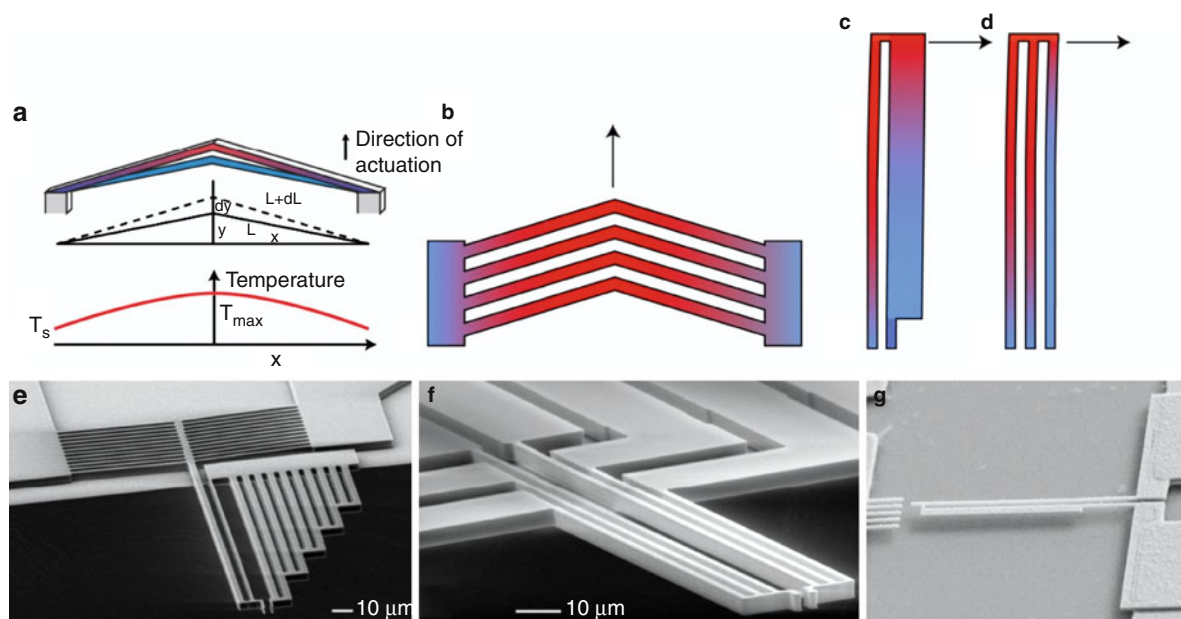
aluminum on multiwalled carbon nanotubes. Yet, there are no examples in literature of nanomanipulation carried out with a bimorph gripper.

Electrothermal Actuation

Correspondingly, a current I passing through two rods with different cross section, will be lead to different temperature increase through Joule heating, as the dissipated power in a volume element of length dx and cross-section A is given by $RI^2 = \rho(dx/A)A^2J^2 = \rho AJ^2 dx$, where the current density is $J = I/A$ and the $\rho = RA/dx$ is electrical resistivity. The temperature distribution can be found by solving the heat conduction continuity equation, which in its simplest form reads

$$\frac{d^2T}{dx^2} = -\frac{\rho J^2}{\kappa}, \quad (7)$$

where κ is the thermal conductivity. The heat exchange with the surrounding air/liquid is ignored, which corresponds to an electrothermal device in vacuum. Moreover, energy loss by radiation is omitted. While this can be solved analytically for temperature distribution and actuation response with many basic actuator designs, finite element method is often used for its greater applicability to nontrivial designs [4]. Two often used designs are the Guckel actuator (also called thin-beam-thick-beam) and the rib-cage actuator, also called the Chevron actuator [1]. The Guckel actuator is a loop where the two conducting arms have different cross section, and the thick beam has a hinge to allow flexible sideways motion. The rib-cage actuator



Nanogrippers, Fig. 6 Electrothermal actuators. (a) A V-actuator is pushed in the upward direction, as the opposing beams expand upon heating. By passing a current through the structure it will attain a parabolic temperature distribution with the peak temperature T_{max} in the center. (b) Rib-cage or “Chevron” actuator consisting of many V-actuators in parallel. (c) Thin-beam-thick-beam or “Guckel” actuator. As current is passed through the loop, the larger heat generated in the

narrowest beam will force deflection sideways towards the cooler, wide beam. (d) The three-beam actuator operates by passing a current through either the left or right two legs, consequently allowing for actuation in two directions, or opening as well as closing capability of the gripper. (e) Silicon rib-cage actuator [6]. (f) Silicon three-beam actuator [3]. (g) Thin-beam-thick-beam actuator with submicron dimensions [18]

consists of opposing rods, which are angled slightly, so that the longitudinal expansion of the rods forces the actuator to move perpendicular to the beam direction, as indicated in Fig. 6, which also amplifies the motion.

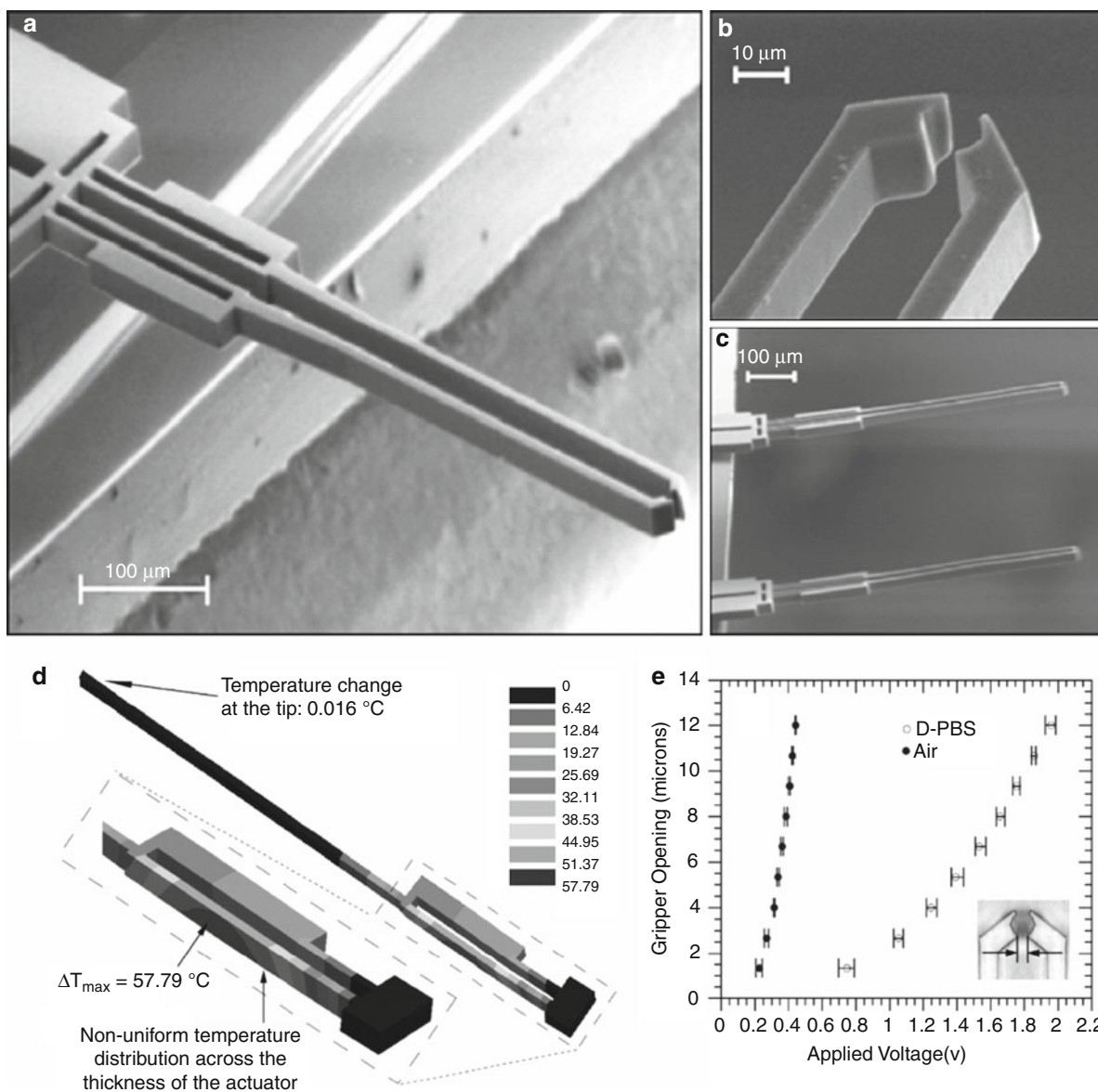
Carlson and Eichhorn [6] used an advanced nanorobotic manipulation system equipped with an electrothermal rib-cage gripper to break off a multiwalled carbon nanotube from a substrate, mount it as a tip of an atomic force microscope probe, and finally fasten the nanotube to the probe by Electron Beam Induced Deposition. The US company Zyvex manufactures a line of advanced electrothermal microgrippers. The conversion from electrothermal actuation to a mechanical motion of the end-effectors is done through a highly optimized design where a capacitive comb-structure provides force readout. While these grippers were clearly designed for micron-size or larger objects the precision and overall size of the end-effectors should allow manipulation of submicron objects.

Chronis [5] employed a hybrid material approach to achieve a large actuation range at a much lower

temperature than what can be achieved in bulk silicon or metal (Fig. 7). Thin metallic structures for conducting the electrical current was combined with the high thermal expansion coefficient material SU-8 (52 ppm/K). The gripper has submicron features, can open almost $10 \mu\text{m}$ at a bias voltage of around 1 V, and most notably impose a very small heat load onto the surroundings. Finite element calculations showed a temperature increase at the end-effector of a fraction of a degree, while the overall temperature increase was around 60 K for a gripper opening of $12 \mu\text{m}$. When using such a hybrid device with different expansion coefficients, there is a risk of out-of-plane bending due to the bimorph effect. In this case, however, the thickness of the SU-8 layer was so much larger than the metal film, that out-of-plane bending could be avoided.

Topology Optimization

Compared to electrostatic nanotweezers where the basic capacitor structures have proven their worth in millions of commercial MEMS devices, there is more room for improvement for electrothermal actuators. Electrostatic

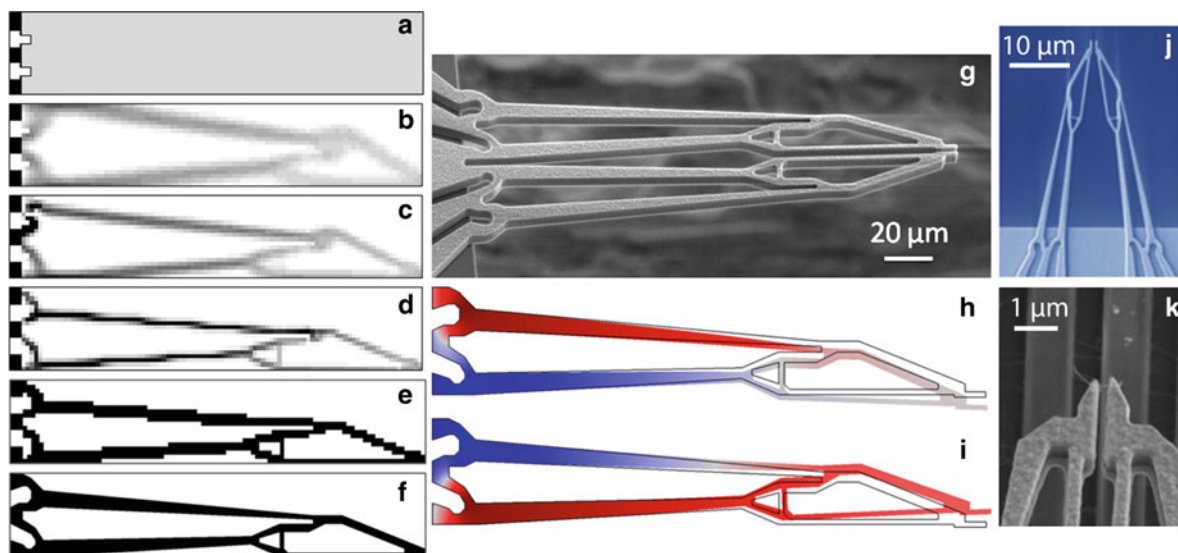


Nanogrippers, Fig. 7 (a) Micro/nanogripper with electrothermal actuators defined in SU-8 with thin chromium-gold electrodes heating the Guckel-type actuators at the base. (b) Zoomed in and (c) overview micrographs of the gripper. (d) Finite element calculations showing negligible temperature

increase at the tip. (e) Actuation curves for the gripper in both liquid (D-PBS) and in air. Due to different thermal resistance to liquid as compared to air, opening the gripper in liquid requires far higher applied voltage [5]

actuation is governed by large thermal gradients, and depends on the detailed shape of the actuator as well as the material properties in the mechanical, electric, and thermal domain. While the formalism outlined above (7) is capable of predicting the behavior of any design, electrothermal actuators are often based on rather simple variations of either the thin-beam-thick-beam actuator or the rib-cage/V-shaped actuator.

For nanogrippers, however, there are good reasons to push beyond the standard solutions. When downsizing the actuators to make overall smaller and more precise grippers, the available force is typically also decreased. To avoid strong out-of-plane bending, a high aspect ratio of the gripper arms is necessary, making it difficult at the same time to reduce the size and maintain a high lateral spring constant. Just as for



Nanogrippers, Fig. 8 Topology-optimized three-beam nanogrippers. (a) Illustration of the initial domain, where the black rectangles indicate the fixed support structures and electrode connections. (b–d) During the optimization process, the material is redistributed toward a more optimal shape. (e, f) Optimized shape converted to black&white and smoothed. (g) Scanning electron micrograph of topology-optimized gripper

fabricated in silicon. (h, i) Close and open modes of the actuator (red/blue indicate hot/cold regions), showing that the peak temperature is moved toward the anchors rather than the end-effectors as in the original device [4]. (j) 30 μm long topology optimized nanogripper, with a gap size of 300 nm. (k) Nanogripper lifting out a multiwalled carbon nanotube from a narrow trench [8]

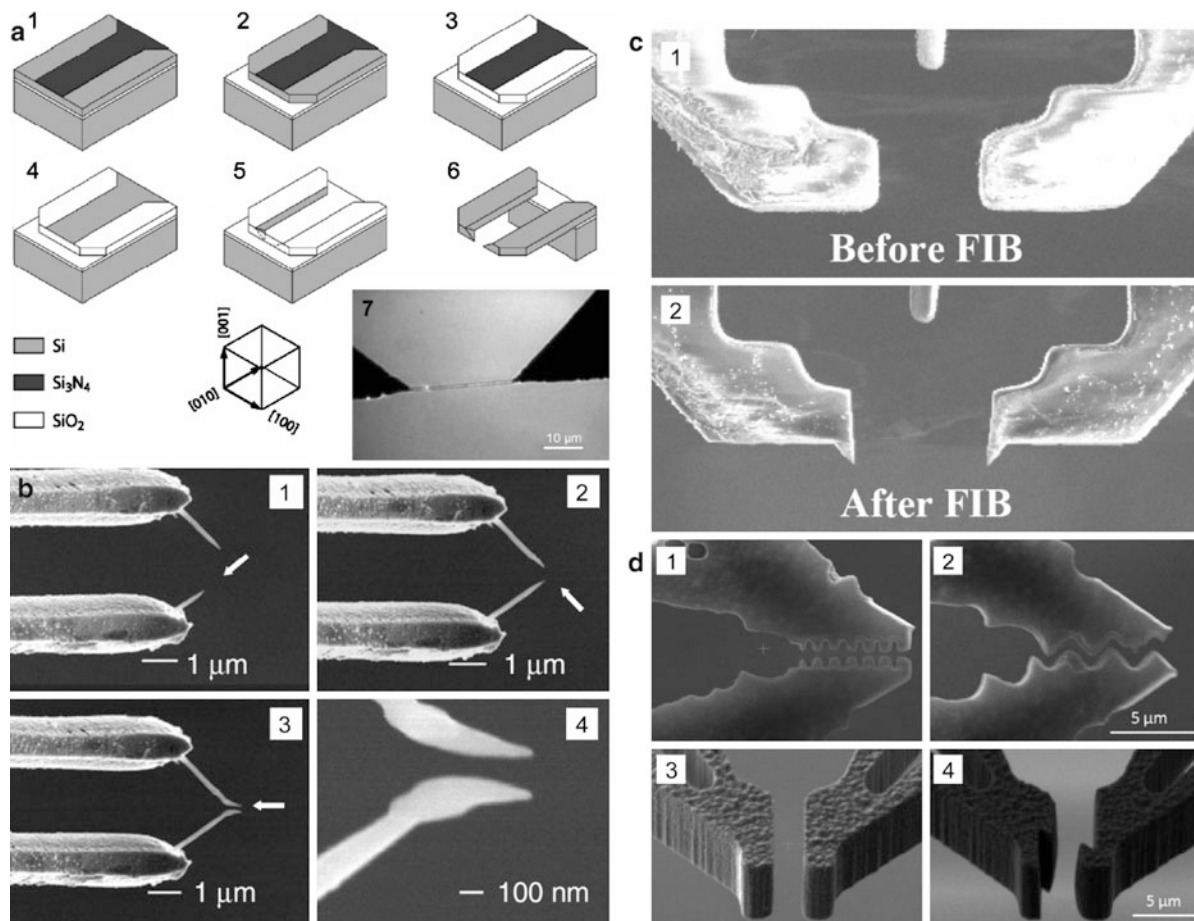
macroscopical tweezers, the optimal shape of a gripper in terms of accessibility and actuation stroke is long and slender, while for strength and force it should be short and wide. Sardan et al. [4] replaced the three-beam actuator with a topology-optimized actuator, and achieved two orders of magnitude higher spring constant at the same actuation stroke (1 μm), while the peak temperatures were conveniently moved toward the base away from the end-effectors (see Fig. 8). This strategy leads to compact grippers capable of successful, repetitive mounting of carbon nanotubes onto AFM probes, TEM grids, and fixed electrodes [4].

Cagliani and coworkers [8] used electron beam lithography to batch-fabricate topology-optimized nanogrippers with a length of 35–50, a gap size of 300 nm, reducing the overall size by a factor of 2–6 times compared to Ref. [4]. The reduced size and thickness of the end-effectors in combination with the enhanced rigidity given by the topology-optimized layout, made it possible to break off as-grown Gallium Phosphide nanowires from the substrate inside a scanning electron microscope. By coating the grippers with a thin layer of gold, bundles of single-walled carbon nanotubes could be lifted up and mounted

inside a Transmission Electron Microscope for subsequent electrical and structural characterization. This device is so far the smallest batch-fabricated nanogripper in operation (Fig. 8 j-k).

Nanoengineering of End-Effectors

Just as for macroscale grippers, the size and shape of the end-effectors – the critical interfacial part of the gripper where the contact is made to the nanostructures – is vitally important for successful application of the tool. While conventional nanofabrication techniques such as electron beam lithography are sometimes used [8], this critical part of the gripper is often engineered using creative and unusual solutions. Elegant use of photolithography combined with an anisotropic potassium hydroxide wet etch allowed Yamahata [22] to create sharp opposing tips on the end-effectors in a batch process, see Fig. 9a. Boggild et al. [19] used angled Electron Beam induced Deposition (EBiD) to define needle-shaped end-effectors with a gap size of 100 nm inside a scanning electron microscope, as shown in (Fig. 9b). A gap size of just 25 nm could be achieved by gradually allowing the deposited material on the two tips to grow until the



Nanogrippers, Fig. 9 Customization of end-effectors. (a) Anisotropic wet etch with potassium hydroxide (KOH) etch can be used to create sharp edges and tips in silicon. A narrow gap with ultrasharp opposing tips was created in a batch process using KOH etch by Yamahata [22]. (b) Electron beam induced deposition performed in the direction of the *white arrow* shows gradual buildup of dual supertips on silicon dioxide end-effectors with a gap of 100 nm. The method was used to create gaps down to 25 nm. [19]. (c) Focused ion beam milling used to

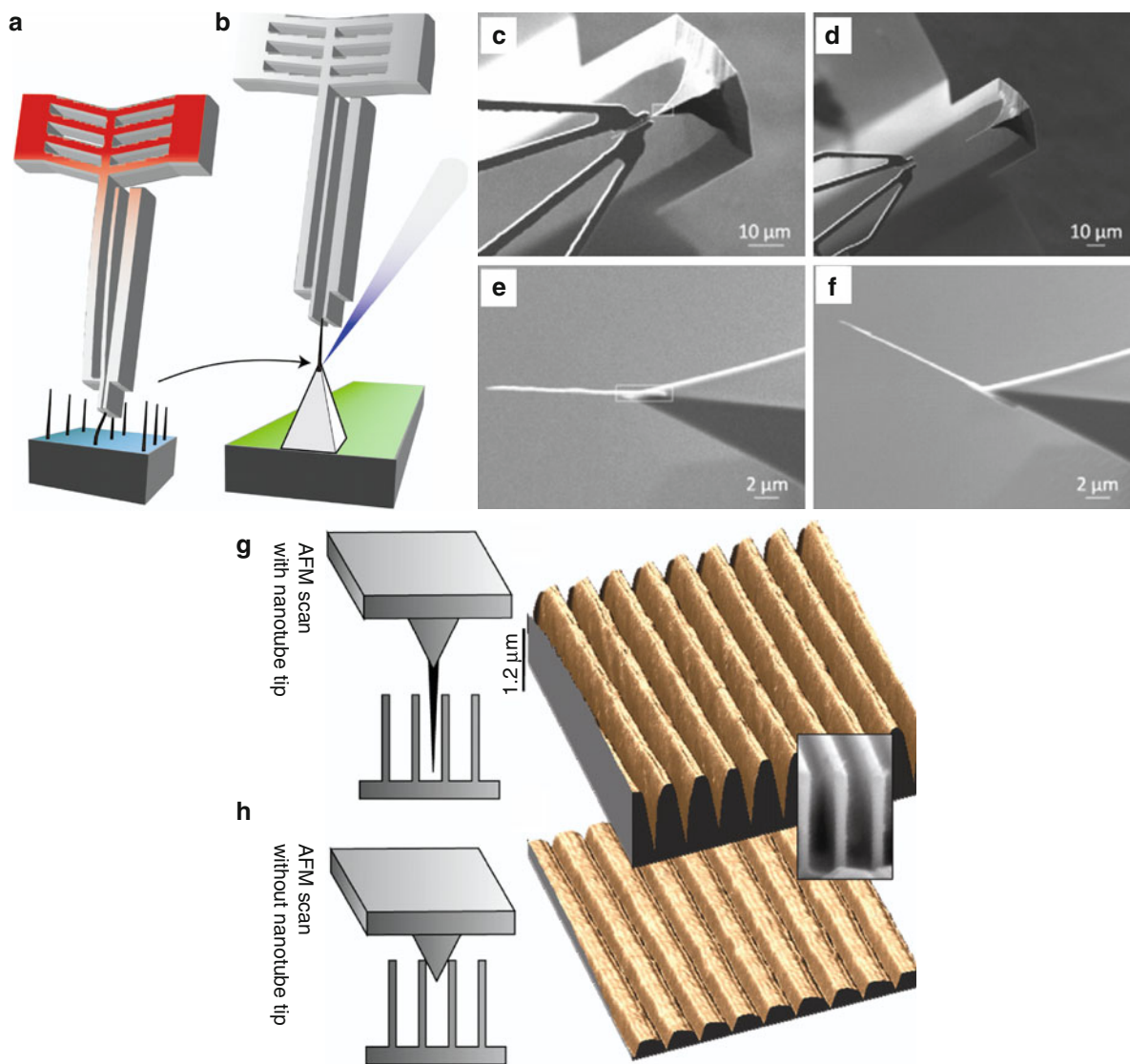
reshape end-effectors [20]. (d.1) Silicon microgripper fabricated with an initial closed gap, which is subsequently opened and further shaped by the focused ion beam milling for reduced contact area. (d.2) FIB milling is done so that the gripper will apply a finite force even in the neutral position, that is, with no applied voltage. (d.3, 4) show end-effectors equipped with a self-locking mechanism, so the gripper when pushed together will lock in this position similar to some macroscopic tweezers (Courtesy of Özlem Sardan and Mikkel Klarskov)

desired gap size was achieved. Focused ion beam milling is highly convenient for in situ reshaping of nanodevices, since it can both perform milling of material as well as deposition using precursor gasses. FIB milling has been used to define end-effectors with a more well-defined or smooth contact plane (Fig. 9c), to achieve a smaller contact area for reduced surface force interactions (Fig. 9d.1) or to engineer the end-effector so that it applies a force in its *neutral* position, and has to be opened rather than closed during the actuation phase (Fig. 9d.2). Finally, FIB can be used

for more exotic mechanism such as self-locking grippers which can keep a nanostructure mechanically fixed even when “released” (Fig. 9d.4).

Applications

In terms of nanomanipulation, the objects of interest may be small in either one, two, or three directions, that is, plates, wires, or particles, and this makes quite a difference in terms of how they can be handled and



Nanogrippers, Fig. 10 (a, b) Illustration of pick and place of vertically aligned multiwalled carbon nanotubes from their as-grown position to the apex of a scanning probe tip, where the *blue triangle* illustrates the electron beam induced deposition used to fasten the nanotube to the probe. (c, d) The nanogripper during and after mounting of the nanotube tip. (e, f) Examples of

different configurations for the nanotubes on the AFM probe. (g, h) AFM scanning of deep trenches with and without the supertip, showing a clear improvement of the nanotube-enhanced probe compared to a commercial super-sharp AFM probe

with which tools. In the following, the feasibility and optimal strategy of manipulating such objects will be discussed in the light of possible applications.

Quasi 0D Structures: Particles

There are a few examples of nanotweezers used for picking up of submicron particles in demonstration of

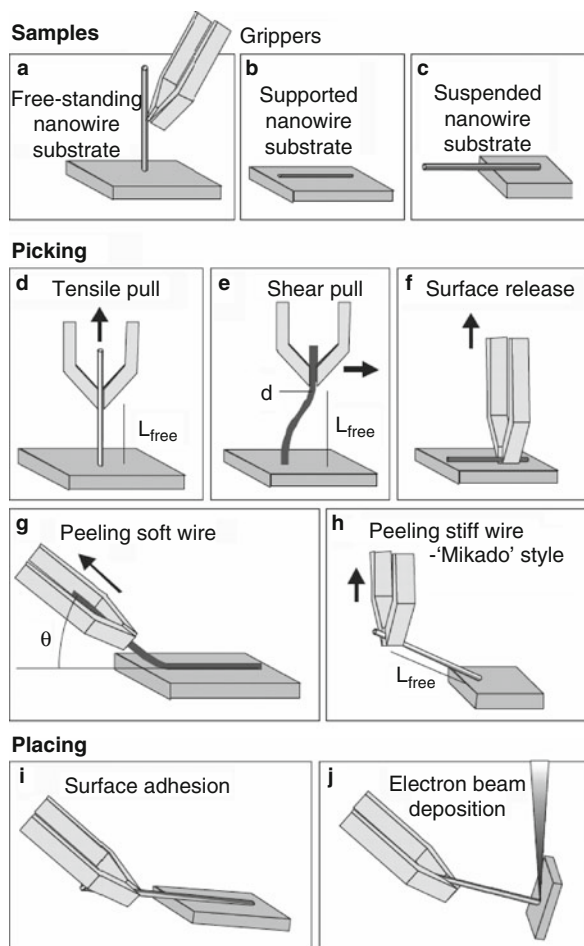
the capability of a nanogripper [11, 13]. Such manipulation can also be combined with electrical characterization of an object if the end-effectors are conducting, similar to for instance carbon nanotubes [11]. For the manipulation process itself, a simple tip will in most cases be effective for particles as numerous pick-and-place experiments using scanning probe tips have

shown. Submicron and nanoscale particles will in the majority of practical scenarios be notoriously difficult to pick up using a gripper, either because the gripper force is too weak to overcome the surface forces, or too bulky to access the particle in a predictable way. This situation is comparable to lifting up grains of sand using fingers, which is possible but hardly convenient. While proof-of-principle demonstrations are possible, it is doubtful that mechanical nanogrippers will ever play an important role in real applications involving nanoparticles.

Quasi 1D Structures: Fibers, Wires, and Tubes

Just as for the macroscale, the most obvious and convenient object to handle with a gripper or a pair of tweezers is a quasi-one-dimensional object, such as a carbon nanotube. The oblong shape of the object makes access far easier than for nanoparticles, with excellent visibility as the nanowires/nanotube does not block the vision. Moreover, the fact that the free end of the nanowires or nanotube can easily be positioned and even fixed by Electron Beam induced Deposition (EBiD) of Focused Ion Beam deposition provides a straightforward and reliable pick-and-place process, which can be automatized [14]. Manipulation of nanowires/nanotubes fall in two categories:

1. Investigation of the mechanical, electrical, or structural properties of individual nanowires/nanotubes. The mechanical properties of nanostructures can be measured by inferring the gripping force from the visible deformation of the nanogripper upon contact, given that the spring constant is known, or by using sensors integrated in the gripping device [2]. The electrical properties can be measured if the end-effectors are conducting [8, 11] or if the nanostructure is placed onto electrodes [4]. The structural properties can be measured by transporting the nanostructure to a Transmission Electron Microscope [8] or onto a TEMgrid [4].
2. Direct application of the nanostructure to perform a function in a device. Microfabricated electrothermal grippers were used in a semiautomated assembly procedure [4] to first break off multiwalled carbon nanotubes from their as-grown position on a substrate, transport them, and align them to an atomic force probe, and finally hold them in place while electron beam induced deposition was used to fasten the nanotube to the probe.

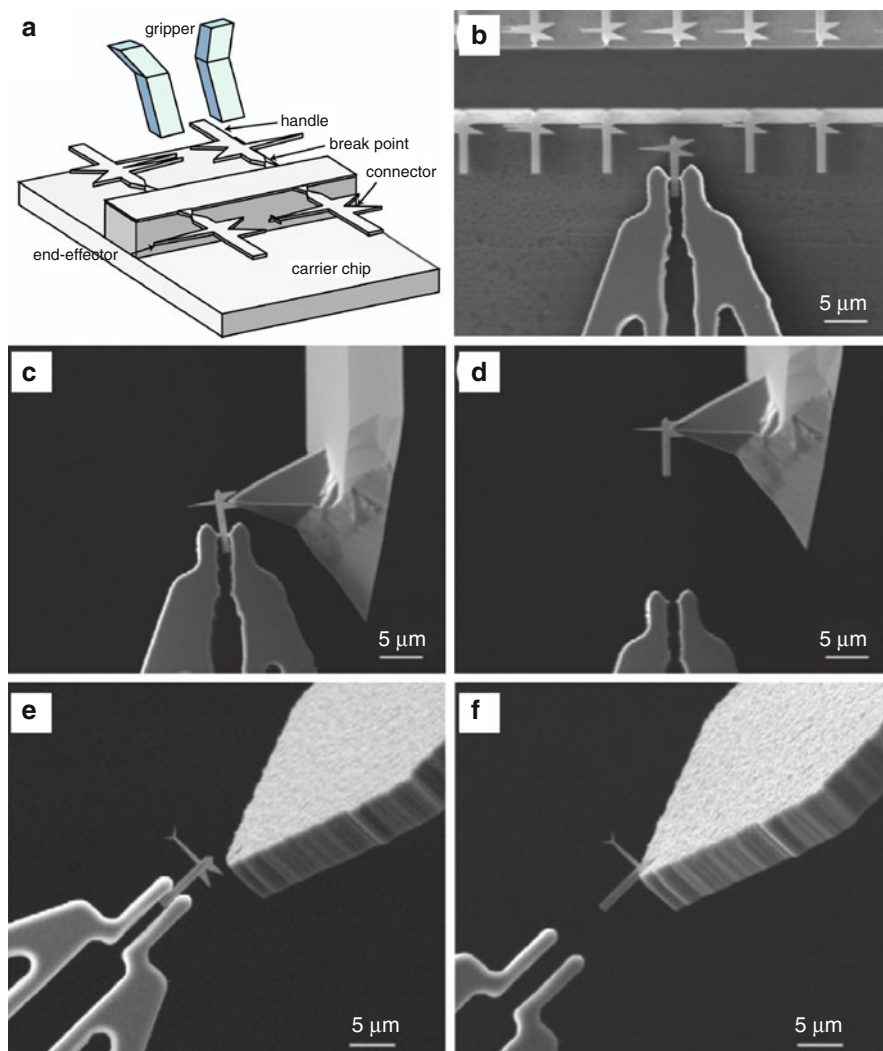


Nanogrippers, Fig. 11 (a–c) Initial configurations of rod-like nanostructures. Picking strategies: (d) Tensile pull. (e) Shear pull and (f) Surface release. (g) Peeling of soft wire and (h) peeling of rigid wire. Strategies for placing using either (i) surface adhesion or (j) Electron Beam induced Deposition (EBiD) [10]

Figure 10a, b illustrates the basic procedure, and the device during (Fig. 10c) and after (Fig. 10d) mounting of carbon nanotube is shown as scanning electron images. Figure 10e and 10f shows two configurations with different angle of the carbon nanotube tip, and Fig. 10g and 10h shows AFM scans of 2.5 mm deep microfabricated trenches with and without the carbon nanotube supertip attached to the AFM probe. The trenches are well resolved by the supertip-enhanced AFM probe.

Detaching nanorods from a fixed position using a nanogripper is nontrivial problem. Mølhave and

Nanogrippers, Fig. 12 (a) Illustration of the “nanobits” concept. (b) Arrays of premade scanning probe tips fabricated with electron beam lithography, with a single “nanobit” being detached using a nanogripper. (c) Mounting of nanobits onto AFM scanning probe tip. (d) The nanobit-enhanced scanning probe ready to scan. (e) A sidetip-equipped nanobit tip mounted in a focused ion beam slit in an AFM scanning probe (f) [21]



coworkers [10] divided the possible arrangements of nanorods on a surface into freestanding, supported, and (partly) suspended initial states, and compared different pick-and-place strategies for such structures. Cartridges of carbon nanotubes suspended from a razor-sharp knife edge were used by Akita and Nakayama [13] to assemble nanotubes on scanning probe tips to be used as nanotweezers.

Freestanding nanowires and nanotubes are very convenient objects for manipulation: they can be catalytically grown with great uniformity and in large numbers using physical or chemical vapor deposition techniques, with positions and diameter determined by pre-positioned catalytic particles. Thus, it is straightforward to produce large arrays of accessible

nanostructures, with well-controlled dimensions, position, and physical properties. The deterministic nature of such a periodic arrangement directly facilitates automation. Detachment of as-grown structures, however, may require a considerable force as the often very high maximal strength must be overcome. Mølhave et al. considered mainly two approaches to detach freestanding 1D nanostructures; tensile pull along the longitudinal axis of the nanorod, and shear pull in the perpendicular direction with respect to the substrate (sideways). In practical experiments as well as finite element calculations by Andersen et al. [3], it turned out that shear pulling is far the most effective strategy for multiwalled carbon nanotubes (Fig. 11).

Quasi 2D Structures: Flakes, Membranes, and Slices

Microgrippers can be used to manipulate thin slices of material, which have been cut out by focused ion beam milling, onto a transmission electron grid for inspection of the cross-sectional structure. Although the thickness of such slices is often just of order of hundred nanometers to allow transmission of electrons through the slice, the grippers do not have to be particularly small for this task; in fact, a considerable force may be required to detach the TEM samples from their support. Nanogrippers might be envisioned for manipulation of very small flakes of material onto TEM grids, or for other plate or flake-like nanostructures. Kumar and coworkers [21] demonstrated predictable pick and place of nanoscale flake-like scanning probe supertips as shown in Fig. 12. These detachable, 130 nm thin silicon nitride membrane tips, termed “nanobits,” were defined by electron beam lithography and arranged in large reservoirs for convenient access. The most successful strategy for picking up these objects turned out to be grasping at the edges of the flakes with the gripper and breaking the tips off with a sideways motion (see Fig. 12), similar to the strategy used to detach as-grown carbon nanotubes [6]. In these experiments, several nanobits were mounted on Atomic Force Microscope tips and used for AFM scanning of deep trenches and sidewalls, showing considerable improvement over conventional AFM probes.

Non-manipulative Applications

Even when not actuated mechanically, nanogrippers provide two individual fingers with a narrow spacing, and which can easily be made conductive. This has shown to be convenient for a number of tasks that are not themselves mechanical manipulation including “grasping” nanostructures through dielectrophoresis [22], electrical characterization ([4, 8, 11, 22]), transporting nanostructures into a TEM for structural analysis, and in situ experiments [8] as well as alternatively placing nanostructures on a TEM grid for analysis. Nanostructures with diameters in the 10 nm range are exceedingly difficult to manipulate through mechanically actuation of dual end-effectors.

Conclusion

Despite recent progress in gripper technology as well as robotic manipulation systems, reliable operation of nanogrippers still remains a challenge. The necessity of real-time observation of the mechanical interactions on the nanoscale, surface forces, as well as the difficulty in creating gripper devices small enough to deal with sub-100 nm structures is limiting the transition from literature’s numerous proposed nanogrippers concepts to devices that support practical applications. All in all, state-of-the-art nanogrippers are highly useful tools for basic nanomaterial research, in addition to providing a shortened turnaround time for single device prototyping and perhaps even small scale production of specialized nanodevices, which involves nanostructures in the 10–1,000 nm range.

Cross-References

- ▶ [Basic MEMS Actuators](#)
- ▶ [Carbon Nanotubes](#)
- ▶ [DNA Manipulation Based on Nanotweezers](#)
- ▶ [Electron-Beam-Induced Deposition](#)
- ▶ [Focused-Ion-Beam Chemical-Vapor-Deposition \(FIB-CVD\)](#)
- ▶ [Manipulating](#)
- ▶ [Nanorobotic Assembly](#)
- ▶ [Nanorobotic Spot Welding](#)
- ▶ [Nanorobotics](#)
- ▶ [Nanorobotics for NEMS using Helical Nanostructures](#)
- ▶ [Thermal Actuators](#)

References

1. Sahu, B., Taylor, C.R., et al.: Emerging challenges of microactuators for nanoscale positioning, assembly, and manipulation. *J. Manuf. Sci. Eng. Trans. ASME* **132**(3), 0309171–03091716 (2010)
2. Kim, K., Liu, X.Y., et al.: Nanonewton force-controlled manipulation of biological cells using a monolithic MEMS microgripper with two-axis force feedback. *J. Micromech. Microeng.* **18**(5), 3100–3105 (2008)

3. Andersen, K.N., Petersen, D.H., et al.: Multimodal electrothermal silicon microgrippers for nanotube manipulation. *IEEE Trans. Nanotechnol.* **8**(1), 76–85 (2009)
4. Sardan, O., Eichhorn, V., et al.: Rapid prototyping of nanotube-based devices using topology-optimized microgrippers. *Nanotechnology* **19**(49), 495503–1–495503-9 (2008)
5. Chronis, N., Lee, L.P.: Electrothermally activated SU-8 microgripper for single cell manipulation in solution. *J. Microelectromech. Syst.* **14**(4), 857–863 (2005)
6. Carlson, K., Andersen, K.N., et al.: A carbon nanofibre scanning probe assembled using an electrothermal microgripper. *Nanotechnology* **18**(34), 345501-1–345501-9 (2007)
7. Molhave, K., Hansen, T.M., et al.: Towards pick-and-place assembly of nanostructures. *J. Nanosci. Nanotechnol.* **4**(3), 279–282 (2004)
8. Cagliani, A., Wierzbicki, R., et al.: Manipulation and in situ transmission electron microscope characterization of sub-100 nm nanostructures using a microfabricated nanogripper. *J. Micromech. Microeng.* **20**(3), 035009-1–035009-7 (2010)
9. Bhushan, B. (ed.): *Handbook of Nanotechnology*. Springer, Berlin (2004)
10. Molhave, K., Wich, T., et al.: Pick-and-place nanomanipulation using microfabricated grippers. *Nanotechnology* **17**(10), 2434–2441 (2006)
11. Kim, P., Lieber, C.M.: Nanotube nanotweezers. *Science* **286**(5447), 2148–2150 (1999)
12. Akita, S., Nakayama, Y., et al.: Nanotweezers consisting of carbon nanotubes operating in an atomic force microscope. *Appl. Phys. Lett.* **79**(11), 1691–1693 (2001)
13. Akita, S., Nakayama, Y.: Manipulation of nanomaterial by carbon nanotube nanotweezers in scanning probe microscope. *Jpn. J. Appl. Phys. Pt. 1 Regul. Pap. Short Notes Rev. Pap.* **41**(6B), 4242–4245 (2002)
14. Eichhorn, V., Fatikow, S., et al.: Nano lab: a nanorobotic system for automated pick-and-place handling and characterization of CNTs. In: *ICRA: 2009 IEEE International Conference on Robotics and Automation, Kobe*, vols. 1–7, pp. 1643–1648 (2009)
15. Senturia, S.D.: *Microsystem Design*. Academic, Boston (2001)
16. Beyeler, F., Neild, A., et al.: Monolithically fabricated microgripper with integrated force sensor for manipulating microobjects and biological cells aligned in an ultrasonic field. *J. Microelectromech. Syst.* **16**(1), 7–15 (2007)
17. Sul, O., Yang, E.H.: A multi-walled carbon nanotube-aluminum bimorph nanoactuator. *Nanotechnology* **20**(9), 095502-1–095502-5 (2009)
18. Lee, J.S., Park, D.S.W., et al.: Sub-micron metallic electrothermal actuators. *J. Micromech. Microeng.* **15**(2), 322–327 (2005)
19. Boggild, P., Hansen, T.M., et al.: Fabrication and actuation of customized nanotweezers with a 25 nm gap. *Nanotechnology* **12**(3), 331–335 (2001)
20. Chen, B.K., Zhang, Y., et al.: From microgripping to nanogripping. In: *MEMS 2010: 23 rd IEEE International Conference on Micro Electro Mechanical Systems, Hongkong. Technical Digest*, pp. 296–299 (2010)
21. Kumar, R.T.R., Hassan, S.U., et al.: Nanobits: customizable scanning probe tips. *Nanotechnology* **20**(39), 395793-1–395793-6 (2009)
22. Yamahata, C., Collard, D., et al.: Silicon nanotweezers with subnanometer resolution for the micromanipulation of biomolecules. *J. Microelectromech. Syst.* **17**(3), 623–631 (2008)

Nanohardness

- ▶ [Nanomechanical Properties of Nanostructures](#)

Nanohardness Tester

- ▶ [Nanoindentation](#)

Nanohelix

- ▶ [Nanorobotics for NEMS Using Helical Nanostructures](#)

Nanoimprint Lithography

- ▶ [Nanoscale Printing](#)
- ▶ [Nanoimprinting](#)

Nanoimprinting

Wei Chen and Chunlei Wang
 Department of Mechanical and Materials Engineering,
 Florida International University, Miami, FL, USA

Synonyms

[Nanoimprint lithography](#)

Definition

Nanoimprinting is a type of lithographic technology that provides a simple and efficient approach of patterning various nanostructures with sub-10-nm resolution.

Overview

Nanoimprinting can promise state-of-the-art smaller electronics. Although nanoimprint lithography (NIL) is not yet an industrial process with well-established standard, extensive interests in NIL processes come from a large community of sensor, biochip, and nano-optics manufactures and institutes [1]. Nanoimprinting has received increased research motivation due to its ability to create smaller functional device structures such as nanoscale transistors, single electron memory, nanofluidic channels, etc., via an accurate and efficient nanolithography process [2].

NIL is very similar to conventional hot embossing, which is based on stamping of a pattern into a polymer softened by raising the temperature of the polymer just above its glass transition temperature. The hot embossing stamp can be made in a variety of ways including micromachining from silicon, LIGA (a German acronym for Lithographie (Lithography), Galvanoformung (Electroplating), Abformung (Molding)), and machining using a CNC (computer numerical control) tool (for making large features). A wide variety of polymers have been successfully hot embossed with submicron size features, including PC (polycarbonate) and PMMA (polymethylmethacrylate).

In the current encyclopedia, a general introduction about NIL process, the involved materials such as resist and mold (template), as well as NIL manufacturers are presented. The detailed and intensive discussions about NIL are lucubrated in the related review papers [3, 4].

Key Concepts

Nanoimprinting Process

NIL, i.e., thermal NIL process, was first reported by Chou S. et al. from University of Minnesota in 1995 [5, 6]. Subsequently, Haisma J. et al. from Philips research laboratories developed photo-NIL,

i.e., UV-NIL process [7]. The processes of NIL mainly involve the following steps: spin-coating or dispensing resist (polymer or monomer) on the substrate, pressing resist layer with patterned mold, solidifying patterned resist by heating or UV exposure, separating mold (demolding) and postprocessing. The schematic drawing of the NIL fabrication process is shown in Fig. 1. Briefly, a resist with thermoplastic or photosensitive characteristic is first spin-coated or dispensed on a wafer substrate. As for photo-NIL process, there is a predeposited planar layer on the substrate. Then, a hard mold containing nanometer-scale surface-relief features is pressed into the above thin resist cast. Note that a transparent mold is requested in the UV-NIL. As shown in Fig. 1 (Step 2 and 3), the mold protrusions are squeezed into the soft resist layer. After the treatment of heating or UV exposure, the patterns of mold are solidified inside the resist. The mold is then removed from the imprinted resist layer, i.e., demolding process. The desired nanostructured patterns in the residual layer are obtained after the appropriate pattern definition process using ion etching post treatment.

The NIL process is suitable for fabricating various nanodevices, whose basic components can be prepared through traditional UV-NIL process. Figure 2 shows examples of patterns transferred into silicon [8]. The residual nanostructure layer with an aspect ratio greater than six is obtained after reactive ion etching (as depicted in Fig. 2a). Maintaining the fidelity of the resist mask by controlling the residual layer thickness allows transfer features down to 13 nm as shown in Fig. 2b. These various nanostructures offer basic units to fabricate photonics or other functional nanodevices at low cost and high throughput.

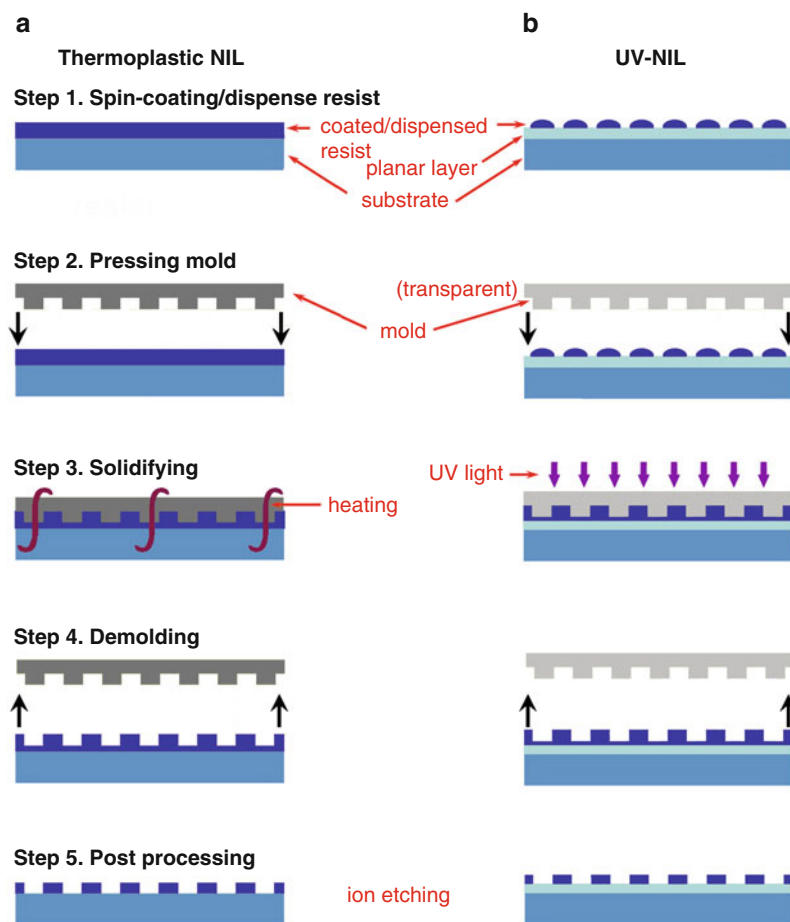
Nanoimprinting Resist

The resist materials used in imprinting should be able to deform easily under an applied pressure and have sufficient mechanical strength. The resist should also have good mold-releasing properties to maintain their structural integrity during the demolding process since imprint lithography makes a conformal replica of the surface-relief patterns by mechanical embossing.

Thermoplastic Resist Materials

The thermoplastic polymers such as PMMA and PS (polystyrene) can be easily deposited on the substrate with desired thicknesses by spin-coating process. These polymers are usually used as imprint resists in

Nanoimprinting,
Fig. 1 Schematic of typical
 (a) thermal NIL and (b) UV-
 NIL process

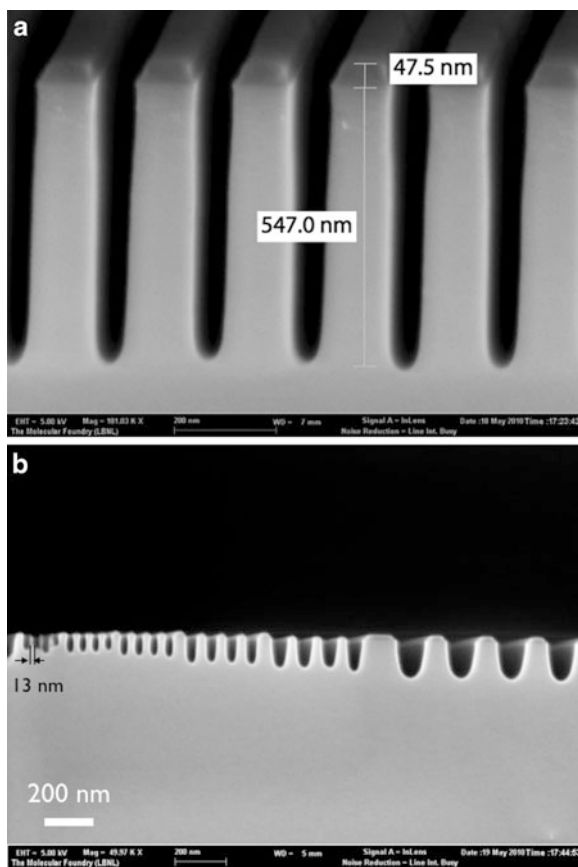


thermal NIL. The glass transition temperature (T_g) of the polymer is a key parameter for the determination of the solidifying treatment temperature in thermal molding processes. Typically a suitable imprint temperature is chosen 70–90°C above the resist’s T_g , so that the polymer materials reach a viscous flow state. In order to preserve the imprinted pattern, the mold and the imprinted resist must be cooled down to below T_g . In addition, there is a trade-off between the imprinting temperature and the thermal stability, which indicates that although a low- T_g material can be used in NIL for the sake of reducing the processing temperature, the imprinted patterns are also unstable and tend to deform at temperatures close to the imprinting temperature.

In a NIL process, the polymers also need to have a sufficiently low viscosity for efficient squeeze to complete the imprinting process within a practical time frame [9]. The viscosity of a polymer material

not only depends on the temperature, but also depends strongly on the polymer’s molecular weight (M_w) relative to the so-called critical molecular weight (M_c) of a given polymer. The M_c can be interpreted as the molecular weight at which a temporary network of entanglements spans over macroscopic dimensions. In practice, low-molecular-weight polymers with $M_w < M_c$ can be imprinted at lower temperatures, lower pressures, or within shorter times. Therefore, the choices of T_g and M_w are both important in maintaining the structural stability of the imprinted patterns.

Various alternatives have been exploited to lower the viscosity of thermoplastic materials. These include dissolving the polymer in its monomer [10] or another solvent and using a PDMS (polydimethylsiloxane) stamp for solvent evaporation [11]. It is inferred that thermally curable or thermosetting polymers are



Nanoimprinting, Fig. 2 SEM pictures of imprinted gratings after transfer into silicon wafers. (a) 90-nm-linewidth gratings with 547-nm depth into silicon and 47.5-nm-thick UV-NIL resist remaining after the etching steps. (b) Grating with variable pitch/linewidth. The minimum feature size is 13 nm with an etching depth of around 50 nm [8]

excellent resist for NIL because of the possibility of low-pressure imprinting and good mechanical integrity after cross-linking by thermal treatment.

UV-NIL Resist Materials

Different from thermoplastic materials requiring high temperature and pressures, photo-NIL (UV-NIL) materials can be handled at ambient temperatures. UV-curable NIL materials are composed of a mixture of monomers or prepolymers and a suitable photoinitiator, and often chemicals are added which decrease the effect of radical scavengers on photopolymerization [12]. The monomer fluid with low Young modulus and low viscosity is less sensitive to the

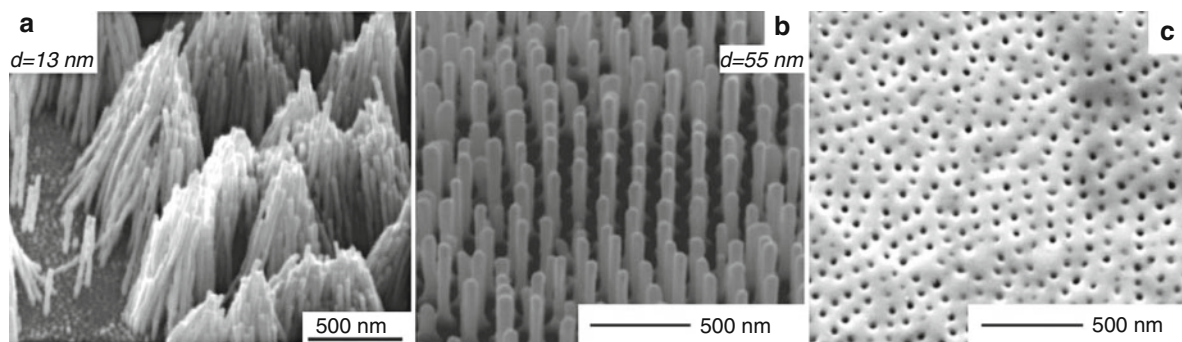
effects of pattern density in imprinting process, which allows the use of a small-area mold and the patterning of large-area substrate by a step-and-repeat process. Immediately during contact of the stamp with the liquid mixture, filling of the mold starts by capillary forces, which pulls the stamp toward the substrate. Therefore, the general strategy is as follows: Low viscosities are needed both for rapid dispensing and filling of mold cavities. Thin resin layers on top of a thicker transfer layer are used to achieve a homogeneous film thickness. Cross-linking and photopolymer conversion are adapted to achieve high curing speed and high etch resistance in the following breakthrough plasma etching process. Shrinkage and etching rate need to be controlled for optimum pattern transfer. Often trade-offs are needed to achieve a good balance between good physicochemical properties wetting and curing kinetics and the suitability for the NIL process [13]. For specific applications such as the dual damascene process for the structuring of interconnects on microchips, resists with adapted dielectric properties have been developed [14].

It should be recognized that NIL can not only be used to form patterns in a polymer resist, but can also be extended to create desired structures in many other polymer systems, especially those that have special functionalities, or can be used to form functional polymer device structures directly.

Nanoimprinting Mold

The mold in the nanoimprint technique plays the same role as the photomask in photolithography. The mold used in NIL can essentially be any type of solid material that has a high strength and durability, which can be made of metals, dielectrics, or semiconductors. Various hard materials such as Si, SiO₂, SiC, silicon nitride and sapphire are suitable for imprinting molds [15]. Diamond is also investigated as a potential mold material for NIL [16]. These studies showed that Si and SiO₂ have sufficient hardness and durability for the nanoimprint application, which make themselves a very good pair for the NIL process. In addition, the UV transparent properties are essential for mold materials of UV-NIL, which render UV light pass through mold itself and irradiate the resist layer.

After spin-coated with a resist, the mold materials are patterned with various features using UV lithography, interference lithography, and electron-beam



Nanoimprinting, Fig. 3 SEM images of (a) Pt-BMG mold with 13-nm-diameter rods fabricated by embossing on porous alumina; (b) Pt-BMG mold with 55-nm-diameter rods; (c) The Pt-BMG mold, after crystallization, imprints holes into the same Pt-BMG [17]

lithography depending on the size requirements. A minimum lateral feature size can be obtained using electron-beam lithography; interference lithography is fitted for large-area periodic features while UV lithography for microscale and larger features. Then, a hard masking layer, such as a metal, can be deposited over the patterned resist template, followed by a lift-off process that removes the resist template and the material on top, leaving a patterned mask layer on the substrate. Lastly reactive ion etching process is used to selectively etch away the mold materials in the unmasked area and produce surface-relief features required for NIL.

Silicon-based molds are expensive to make and are not very durable. Schroers J. et al. at Yale University developed the nanopatterning of amorphous metals which can be used as inexpensive templates for nanoimprinting [17]. These bulk metallic glasses (BMGs) could be ideal materials for small-scale applications due to their superior mechanical and homogeneous properties. These molds with features as small as 13 nm can be used to imprint onto polymers and other materials. Figure 3 shows SEM images of Pt-BMG nanorods with 13-nm (Fig. 3a) and 55-nm (Fig. 3b) diameters fabricated by embossing on porous alumina. The Pt-BMG substrate with nanorods was pressed into a PMMA sheet at 160°C. Figure 3c shows an SEM image of 55-nm-diameter holes imprinted in Pt-BMG using the crystallized Pt-BMG mold. The results demonstrate the ability of metallic glasses to precisely replicate mold features ranging from 250 to 13 nm. The two attractive features of metallic glass molding – the ability to replicate features from tens of micrometers down to 13 nm and the possibility of using BMGs

as nanotemplates – stem directly from the unique softening behavior and homogeneous structure of metallic glasses.

Variation of Nanoimprinting

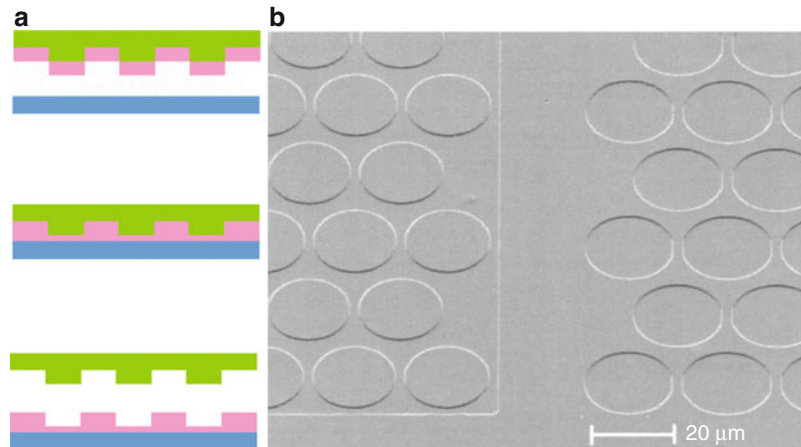
Some innovative processes such as roller-type NIL, combined thermal/UV-NIL, and reverse tone NIL are developed to meet various requirements, which make conventional NIL more flexible and practicable.

Conventional NIL process cannot significantly improve the throughput in the patterning of large-area product with low cost because it is not a continuous process. To overcome this problem, a continuous roller-type NIL is proposed and developed. In roller-type NIL process, only the adjacency area of contact line is pressed for a given time, significantly reducing the printing force and providing better uniformity [18].

Combined thermal/UV-NIL can address the issue of replicating patterns with various sizes and complexity by the separate lithography steps [19]. Sole NIL process experiences difficulties in replicating larger features or resulting in incomplete pattern.

In order to obtain nanosized patterns on the flexible substrate, such as polymer membrane, a reversal imprinting technique was developed [20]. Different from conventional NIL process, a resist layer is first spin-coated on a patterned hard mold and then transferred to a substrate under an elevated temperature and pressure (as shown in Fig. 4a). The reverse NIL process can be accomplished under different pattern transfer modes by controlling imprinting temperature and degree of surface planarization of the spin-coated mold. Figure 4b shows PMMA patterns created by reversal imprinting at 175°C after

Nanoimprinting, Fig. 4 (a) Schematic illustrations of the patter transfer processes in reversal nanoimprinting at temperatures well above the glass transition temperature; (b) Patterns in PMMA created by reversal imprinting at 175°C on a 50-mm-thick Kapton film: 190-nm-deep micrometer-sized mold coated with a 7% solution [20]



spin-coating the mold with 190-nm-deep micrometer-sized features.

Main Manufacturers for Nanoimprinting

The NIL machines are very important for the development and applications of NIL technology. In this entry, the most famous manufactures in the world are listed. The general information about these enterprises is excerpted from their homepages and the corresponding websites, which are attached in the last sentence of every paragraph.

AMO

AMO, i.e., Gesellschaft für Angewandte Mikro- und Optoelektronik mbH, is an important enterprise in the area of nanofabrication. AMO provides fabrication and development services for UV nanoimprint with quartz templates and soft-UV nanoimprint with elastomeric templates. AMO offers the integration of imprint material, templates, and tools with customer-oriented process development. In-house fabrication for high-resolution quartz, elastomeric templates, and UV-curable imprint resists can be supplied. Based on their profound equipment configuration, they can offer up to full 6" or 8" wafer step and repeat. For More details information, readers can visit the website of AMO: <http://www.amo.de>.

EV

EV Group (EVG), a leading supplier of wafer bonding and lithography equipment for the MEMS,

nanotechnology, and semiconductor markets, today unveiled a new technology capability that enables ultrahigh-resolution patterning of features down to 12.5 nm. EV Group provides a complete product line for UV-based nanoimprint lithography including single step UV-imprinting systems and step and repeat large-area UV-nanoimprinting systems. The official website of EV is <http://www.evgroup.com>.

HP's NLS

HP (Hewlett-Packard Co.) has licensed its NIL technology to startup Nanolithosolutions (NLS) Inc.. Nanolithosolution business is to empower researchers and technology developers with the most easy-to-use nanolithography tool to develop next-generation nanotechnology applications. NLS provides complete turnkey solution for nanolithography applications. The product offering includes nanofabrication equipments, all the necessary chemicals and materials, imprinting mold preparation process and imprint fabrication process, and custom process development services. The website of NLS Inc. is <http://www.nanolithosolution.com>.

MII

Molecular Imprints, Inc. (MII) is the market and technology leader for high-resolution, low cost-of-ownership NIL systems and solutions in the hard disk drive and semiconductor industries. Molecular Imprints systems feature its innovative Jet and Flash™ Imprint Lithography (J-FIL™) technology to help create the extremely small features required in today's state-of-the-art storage disk drives and

memory devices. J-FIL has further application in emerging light emitting diodes (LED), solar energy, and biotechnology markets. <http://www.molecular-imprints.com> is the website of MII.

Nanonex

Nanonex was founded in 1999 by Stephen Y. Chou, a Professor of Princeton University and a renowned pioneer in nanotechnology. Built on the team, expertise, and IPs of 8 years worth of multi-million-dollars research of Professor Chou's university research group, Nanonex began its own development in 2000 and quickly delivered products to the market. Today Nanonex products are used in both research and manufacturing by industry and academia. In addition to its own IPs, Nanonex has received exclusive licenses for the related technologies developed by Prof. Chou's research group. Nanonex's current team consists of seasoned business and technical members from industry as well as the experts in nanoimprint technology and applications. The mission of Nanonex is to deliver, today, user-friendly nanoimprint lithography (NIL) tools and solutions to a broad spectrum of markets and to both experts and non-experts of microfabrication. The Nanonex solution makes a turnkey nanoimprint operation today a reality. The website of Nanonex is <http://www.nanonex.com>.

OAI

OAI is a Silicon Valley-based manufacturer of advanced precision equipment for the Semiconductor, MEMS, Microfluidics, Display, and Photovoltaic and Solar Industries. The company offers a broad portfolio of field-proven products that include mask aligners, UV exposure systems, UV light sources, Nano Imprint Systems, and PV Solar Simulators and I-V Testers. The website of OAI is <http://www.amo.de>.

Obducat

Obducat is the world-leading supplier of lithography solutions for manufacturing and replication of advanced micro- and nanoscale structures. Obducat provides viable and cost-effective lithography solutions that will give a competitive edge to the customers, enabling them to deliver breakthrough applications and achieve improved profitability and success. Obducat is the first company to commercialize nanoimprint lithography (NIL) and electron-beam recorder (EBR). Today it is the market leader

with the largest installed base worldwide of NIL. The details about Obducat are shown in the website: <http://www.obducat.com>.

SUSS MicroTec

SUSS MicroTec is a supplier of equipment and process solutions for microstructuring applications with more than 60 years of engineering experience. The solution portfolio covers all performance relevant steps for wafer processing ranging from coating, baking, developing, aligning to wafer bonding and is complemented by specialized add-ons such as photomasks, nanoimprint lithography tools, and optical lenses. The website of SUSS MicroTec is <http://www.suss.com>.

Future Directions for Nanoimprinting

Nanoimprinting technique offers overwhelming advantages for nanoscale device fabrications, including low cost, resolution scalability, and pattern repeatability. The simplicity of NIL has made it appealing to researchers in various fields, such as next-generation HDD storage and future semiconductor devices. However, standard processes or criteria about NIL are not established, which affects the NIL industrial application and development. In addition, the fabrications of high-quality templates with high-resolution periodic features are still key issues limiting the popularization of NIL. Despite all these difficulties, a large community of researchers and enterprises, such as Molecular Imprints, Inc. (MII) and AMO, are trying to establish standard NIL processes and develop significantly advanced materials and molds. Toshiba Corporation has validated the use of its NIL technology in developing 22-nm CMOS devices in 2007. Furthermore, Toshiba fabricated narrow trench features at dimensions down to 18 nm using MII's Imprio 250 system. IBM also focused on understanding the basic physics and chemistry of the nanoimprint process, from initial resist dispense to removal of the template after cure, and using that knowledge to design optimum resist materials and interfacial surface treatments. In the near future, NIL will become a preferred and universal process in nanofabrication technology.

Cross-References

- ▶ [DUV Photolithography and Materials](#)
- ▶ [Nanoscale Printing](#)

References

1. Ofir, Y., Moran, I.W., Subramani, C., Carter, K.R., Rotello, V.M.: Nanoimprint lithography for functional three-dimensional patterns. *Adv. Mater.* **22**, 3608–3614 (2010)
2. Glinsner, T., Kreindl, G., Kast, M.: Nanoimprint lithography. *Optik Photonik* **2**, 42–45 (2010)
3. Guo, L.J.: Nanoimprint lithography: methods and material requirements. *Adv. Mater.* **19**, 495–513 (2007)
4. Schiff, H.: Nanoimprint lithography: an old story in modern times? A review. *J. Vac. Sci. Technol. B* **26**, 458–480 (2008)
5. Chou, S.Y., Krauss, P.R., Renstrom, P.J.: Imprint of sub-25 nm vias and trenches in polymers. *Appl. Phys. Lett.* **67**, 3114–3116 (1995)
6. Chou, S.Y., Krauss, P.R., Renstrom, P.J.: Imprint lithography with 25-nanometer resolution. *Science* **272**, 85–87 (1996)
7. Haisma, J., Verheijen, M., Heuvel, K.V.D.: Mold-assisted nanolithography: a process for reliable pattern replication. *J. Vac. Sci. Technol. B* **14**, 4124–4128 (1996)
8. Peroz, C., Dhuey, S., Volger, M., Wu, Y., Olynick, D., Cabrini, S.: Step and repeat UV nanoimprint lithography on pre-spin coated resist film: a promising route for fabricating nanodevices. *Nanotechnology* **21**, 445301 (2010)
9. Hirai, Y.: Polymer science in nanoimprint lithography. *J. Photopolym. Sci. Technol.* **18**, 551–558 (2005)
10. Jung, G.Y., Ganapathiappan, S., Li, X., Ohlberg, D.A.A., Olynick, D.L., Chen, Y., Tong, W.M., Williams, R.S.: Fabrication of molecular-electronic circuits by nanoimprint lithography at low temperatures and pressures. *Appl. Phys. A* **78**, 1169–1173 (2004)
11. Khang, D.Y., Lee, H.H.: Room-temperature imprint lithography by solvent vapor treatment. *Appl. Phys. Lett.* **76**, 870–872 (2000)
12. Long, B.K., Keitz, B.K., Willson, C.G.: Materials for step and flash imprint lithography (S-FIL[®]). *J. Mater. Chem.* **17**, 3575–3580 (2007)
13. Colburn, M., Choi, B.J., Sreenivasan, S.V., Bonnecaze, R.T., Willson, C.G.: Ramifications of lubrication theory on imprint lithography. *Microelectron. Eng.* **75**, 321–329 (2004)
14. Schmid, G.M., Stewart, M.D., Wetzal, J., Palmieri, F., Hao, J.J., Nishimura, Y., Jen, K., Kim, E.K., Resnick, D.J., Liddle, J.A., Willson, C.G.: Implementation of an imprint damascene process for interconnect fabrication. *J. Vac. Sci. Technol. B* **24**, 1283–1291 (2006)
15. Guo, L.J.: Recent progress in nanoimprint technology and its applications. *J. Phys. D* **37**, R123–R141 (2004)
16. Taniguchi, J., Tokano, Y., Miyamoto, I., Komuro, M., Hiroshima, H.: Diamond nanoimprint lithography. *Nanotechnology* **13**, 592–596 (2002)
17. Kumar, G., Tang, H.X., Schroers, J.: Nanomoulding with amorphous metals. *Nature* **457**, 868–872 (2009)
18. Tan, H., Gilbertson, A., Chou, S.Y.: Roller nanoimprint lithography. *J. Vac. Sci. Technol. B* **16**, 3926–3928 (1998)
19. Pfeiffer, K., Fink, A., Gruetzner, G., Bleidiessel, G., Schulz, H., Scheer, H.: Multistep profiles by mix and match of nanoimprint and UV lithography. *Microelectron. Eng.* **57–58**, 381–387 (2001)
20. Huang, X.D., Bao, L.R., Cheng, X., Guo, L.J., Pang, S.W., Yee, A.F.: Reversal imprinting by transferring polymer from mold to substrate. *J. Vac. Sci. Technol. B* **20**, 2872–2876 (2002)

Nanoindentation

Bharat Bhushan and Manuel L. B. Palacio
 Nanoprobe Laboratory for Bio- & Nanotechnology
 and Biomimetics, The Ohio State University,
 Columbus, OH, USA

Synonyms

[Elastic modulus tester](#); [Nanofatigue tester](#);
[Nanohardness tester](#); [Nanoscratch tester](#)

Definition

Nanoindentation refers to the experiment conducted using a depth-sensing indenter at the nanometer or sub-micrometer scale. The nanoindentation apparatus continuously monitors the load and the position of the indenter relative to the surface of the specimen (depth of an indent) during the indentation process. Sharp indenters can be used to measure hardness, elastic modulus, continuous stiffness, scratch resistance, film-substrate adhesion, residual stresses, time-dependent creep and relaxation properties, fracture toughness, and fatigue.

Overview

Mechanical properties of solid surfaces and thin films are of interest as the mechanical properties affect the tribological performance of surfaces [1–5]. Among the mechanical properties of interest, one or more of which

can be obtained using commercial and specialized indentation testers, are elastic–plastic deformation behavior, hardness, Young’s modulus of elasticity, scratch resistance, film–substrate adhesion, residual stresses, time-dependent creep and relaxation properties, fracture toughness, and fatigue. Indentation measurements can assess structural heterogeneities on and underneath the surface, such as diffusion gradients, precipitates, presence of buried layers, grain boundaries, and modification of surface composition. Physical contacts at sliding interfaces in magnetic storage devices and micro/nanoelectromechanical systems (MEMS/NEMS) occur at very low loads; thus, friction and wear of sliding surfaces is primarily controlled by the physical and chemical properties of a few surface atomic layers. In this entry, the applications of nanoindentation to nanotechnology will be illustrated for the digital micromirror device (DMD), probe-based ferroelectric information storage, and biological MEMS sensor devices.

Nanoindentation Apparatus

In the depth-sensing indentation instruments, the load-indentation depth is continuously monitored during the loading and unloading processes. Commercially available nanoindenters are made by various manufacturers, some of which include Agilent, Oak Ridge, TN; Hysitron, Minneapolis, MN; CSIRO, Lindfield, Australia; CSM Instruments SA, Neuchatel, Switzerland, and Micro Materials, Wrexham, UK.

An example of a common nanoindentation apparatus is the Nanoindenter XP, manufactured by Agilent. The apparatus continuously monitors the load and the position of the indenter relative to the surface of the specimen (depth of an indent) during the indentation process. The area of the indent is then calculated from knowledge of the geometry of the tip of the diamond indenter. Mechanical properties measurements can be made at a minimum penetration depth of about 20 nm (or a plastic depth of about 15 nm) [6]. Specifications and commonly used operating parameters for commercial nanoindenters are given in Table 1.

The NanoIndenter XP consists of three major components: the indenter head, an optical/atomic force microscope, and an *x-y-z* motorized precision table for positioning and transporting the sample between the optical microscope and indenter, Fig. 1a. The load on the indenter is generated using a voice coil in permanent magnet assembly, attached to the top of

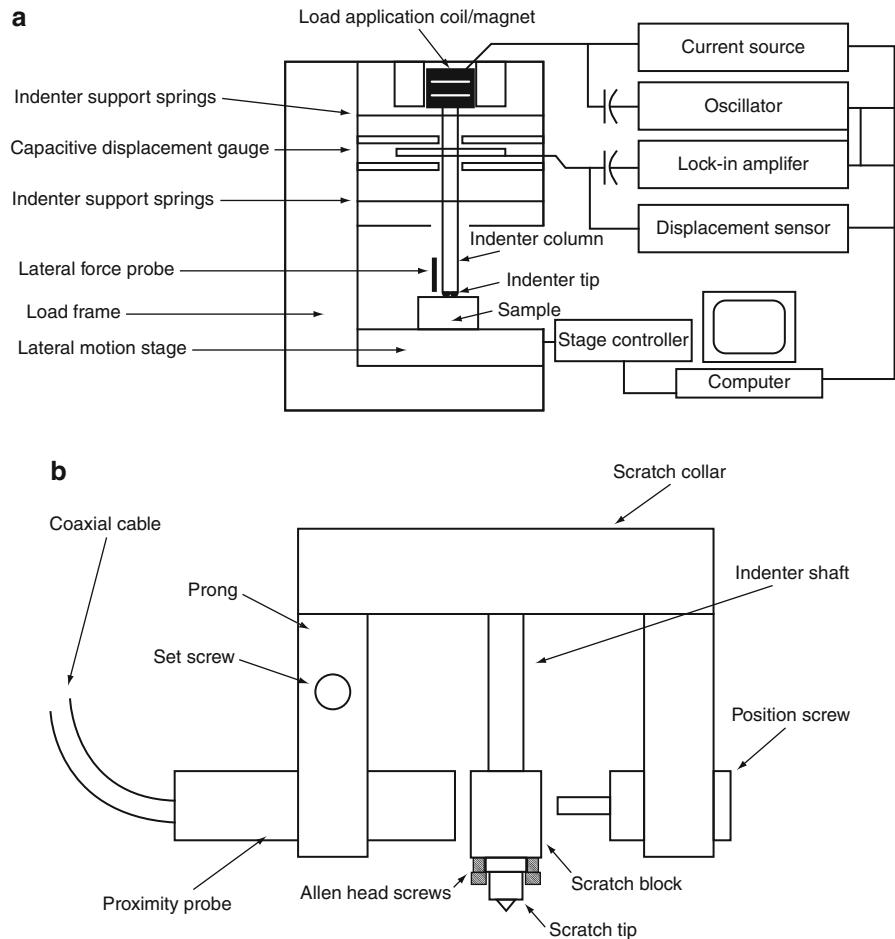
Nanoindentation, Table 1 Specification and commonly used operating parameters for a commercial NanoIndenter

Load range	
Standard heads	0–500 mN
High load head	0–10 N
Load resolution	
Standard head	±50 nN
High load head	±50 nN
Vertical displacement range	0–500 mm
Vertical displacement resolution	±0.05 nm
Typical approach rate	10 nm/s
Typical indentation load rate	10% of peak load/s
Typical indentation displacement rate	10% peak displacement/s
Optical microscope magnification	Up to 1,500 <i>x</i>
AFM objective magnification	Up to one million <i>x</i>
Spatial resolution of the X-Y-Z table	±400 nm in the X and Y directions
Area examined in a single series of indentations	150 × 150 mm
Minimum penetration depth	~20 nm
Continuous stiffness option	
Frequency range	10–150 Hz
Time constant	0.33 s
Smallest measurable distance	0.1 nm
Scratch and tangential force option	
Scratch velocity	Max. 100 mm/s with 20 points/mm
Tangential displacement range	2 mm
Tangential displacement resolution	100 nm
Tangential load resolution	10 mN
Minimum measurable tangential load	10 mN

the indenter (loading) column. The generated load is the vector product of the current through the coil and the magnetic field strength of the permanent magnet. (Other methods of load application are found in other instruments, e.g., electrostatic.) Two interchangeable indenter heads are available with a load range of 50–840 mN. The displacement of the indenter is measured using a three-plate capacitive displacement sensor. The indenter column is attached to the moving plate. Two outer plates are maintained at equal and opposite drive voltages. The output voltage of the center-pickup plate is uniquely related to the position of that plate in the capacitive gap. The displacement system is calibrated using a laser-interferometric

Nanoindentation,

Fig. 1 Schematics of the nanoindenter XP (a) showing the major components – the indenter head, a X-Y-Z motorized precision table, and tangential force option, and (b) detailed schematic of the tangential force option hardware (not to scale and the front and rear prongs not shown)



calibration system. This plate-and-indenter assembly is supported by two leaf springs cut in such a fashion to have very low stiffness. The motion is damped by airflow around the central plate of the capacitor, which is attached to the loading column.

A diamond tip is attached at the bottom of the indenter rod. The indenter head assembly is rigidly attached to the “U” beam below which the x - y - z table rides, Fig. 1a. The optical microscope is also attached to the beam. The position of an indent on a specimen is selected using the optical microscope. The specimens are held on an x - y - z table whose position relative to the microscope or the indenter is controlled with a joystick. The spatial resolution of the position of the table in the x - y plane is ± 250 nm and its position is observed on the computer screen. The nanoindenter is placed on a vibration-isolation table, and the entire apparatus is enclosed in a heavy wooden cabinet to ensure the thermal stability of the samples.

The Indenters

The main requirements for the indenter are high elastic modulus, no plastic deformation, low friction, smooth surface, and a well-defined geometry that is capable of making a well-defined indentation impression. The first four requirements are satisfied by choosing the diamond material for the tip. A well-defined perfect tip shape is difficult to achieve. Berkovich is a three-sided pyramid and provides a sharply pointed tip compared with the Vickers or Knoop indenters, which are four-sided pyramids and have a slight offset (0.5–1 mm) [6]. Because any three nonparallel planes intersect at a single point, it is relatively easy to grind a sharp tip on an indenter if Berkovich geometry is used. However, an indenter with a sharp tip suffers from a finite but an exceptionally difficult-to-measure tip bluntness. In addition, pointed indenters produce a virtually constant plastic strain impression and there is the additional problem of assessing the elastic

modulus from the continuously varying unloading slope. Spherical indentation overcomes many of the problems associated with pointed indenters. With a spherical indenter, one is able to follow the transition from elastic to plastic behavior and thereby define the yield stress [6]. However, a sharper tip is desirable, especially for extremely thin films requiring shallow indentation. Therefore, the Berkovich indenter is most commonly used for measurements of nanomechanical properties.

The Berkovich indenter, directly brazed to a 304 stainless steel holder, is a three-sided (triangular-based) pyramidal diamond, with a nominal angle of 65.3° between the (side) face and the normal to the base at apex, an angle of 76.9° between edge and normal, and with a radius of the tip less than $0.1 \mu\text{m}$.

Another three-sided pyramidal indenter, the cube corner indenter, can displace more than three times the volume of the Berkovich indenter at the same load, thereby producing much higher stresses and strains in the vicinity of the contact and reducing the cracking threshold. This makes this indenter ideal for the estimation of fracture toughness at relatively small scales. The spherical indenter initiates elastic contact and then causes elastic–plastic contact at higher loads, and is well suited for the examination of yielding and work hardening.

Nanoscratch and Lateral Force Measurements

The nanoscratch and lateral force option which allows making of the scratches of various lengths at programmable loads is described here. Lateral (friction) forces can also be measured simultaneously. The additional hardware for the lateral force option includes a set of proximity (capacitance) probes for measurement of lateral displacement or force in the two lateral directions along x and y , and a special “scratch collar” which mounts around the indenter shaft with hardness indenter, Fig. 1b. A scratch block is mounted on the end of the indenter shaft, in line with the proximity probes and the positioning screws. The scratch tip is attached to the scratch block with two Allen head screws. The scratch tip can be a Berkovich indenter or a conventional conical diamond tip with a tip radius of about 1–5 mm and an included angle of 60 – 90° (typically $1 \mu\text{m}$ of tip radius with 60° of included angle). A larger included angle of 90° may be desirable for a more durable tip. The tip radius should not be very small because it will get blunt readily [7].

During scratching, a load is applied up to a specified indentation load or up to a specified indentation depth, and the lateral motion of the sample is measured. In addition, of course, load and indentation depth are monitored. Scratches can be made either at the constant load or at ramp-up load. Measurement of lateral force allows the calculations of the coefficient of friction during scratching. As an example, for the Nanoindenter XP, the resolution of the proximity probe provides resolution of lateral force of about 2 mN; therefore, a minimum load of about 20 mN can be measured. Consequently, a minimum normal load of about 0.2 mN should be used for a sample with coefficient of friction of about 0.1 using the Nanoindenter XP. Microscopy of the scratch produced at ramp-up load allows the measurement of critical load required to break up of the film (if any) and scratch width and general observations of scratch morphology. Typically, draw acceleration (mm/s^2), and draw velocity (mm/s) are 10 mm/s^2 and 5 mm/s , respectively.

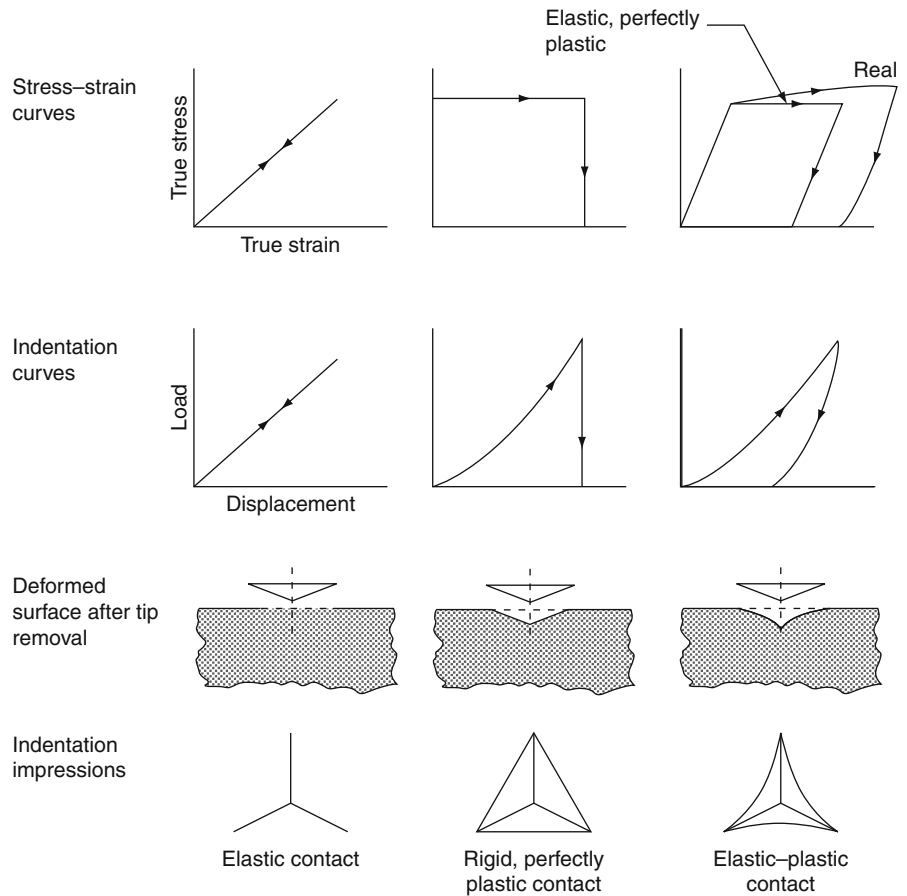
Analysis of Indentation Data

An indentation curve is the relationship between the load W and the displacement (or indentation depth or penetration depth) h , which is continuously monitored and recorded during indentation. Stress–strain curves, typical indentation curves, the deformed surfaces after tip removal, and residual impressions of indentation for ideal elastic, rigid-perfectly plastic and elastic-perfectly plastic, and real elastic–plastic solids are shown in Fig. 2. For an elastic solid, the sample deforms elastically according to Young’s modulus, and the deformation is recovered during unloading. As a result, there is no impression of the indentation after unloading. For a rigid-perfectly plastic solid, no deformation occurs until yield stress is reached, when plastic flow takes place. There is no recovery during unloading and the impression remains unchanged. In the case of elastic–plastic solid, it deforms elastically according to Young’s modulus and then it deforms plastically. The elastic deformation is recovered during unloading. In the case of an elastic-perfectly plastic solid, there is no work hardening.

All engineering surfaces follow real elastic–plastic deformation behavior with work hardening [1, 3, 5, 8, 9]. The deformation pattern of a real elastic–plastic sample during and after indentation is shown schematically in Fig. 3a [10]. The contact depth (h_c) is defined

Nanoindentation,

Fig. 2 Schematics of stress–strain curves, typical indentation curves, deformed surfaces after tip removal, and residual impressions of indentation, for ideal elastic, rigid-perfectly plastic, elastic-perfectly plastic (ideal), and real elastic–plastic solids



as the depth of indenter in contact with the sample under load. The depth measured during the indentation (h) includes the depression of the sample around the indentation in addition to the contact depth. The depression of the sample around the indentation ($h_s = h - h_c$) is caused by elastic displacements and must be subtracted from the data to obtain the actual depth of indentation or actual hardness. At peak load, the load and displacement are W_{max} and h_{max} , respectively, and the radius of the contact circle is a . Upon unloading, the elastic displacements in the contact region are recovered and when the indenter is fully withdrawn, the final depth of the residual hardness impression is h_f . A load–displacement curve schematic is shown in Fig. 3b.

Oliver and Pharr [10] developed an expression for h_c at the maximum load (required for hardness calculation) from h_{max} ,

$$h_c = h_{max} - \varepsilon W_{max}/S_{max} \tag{1}$$

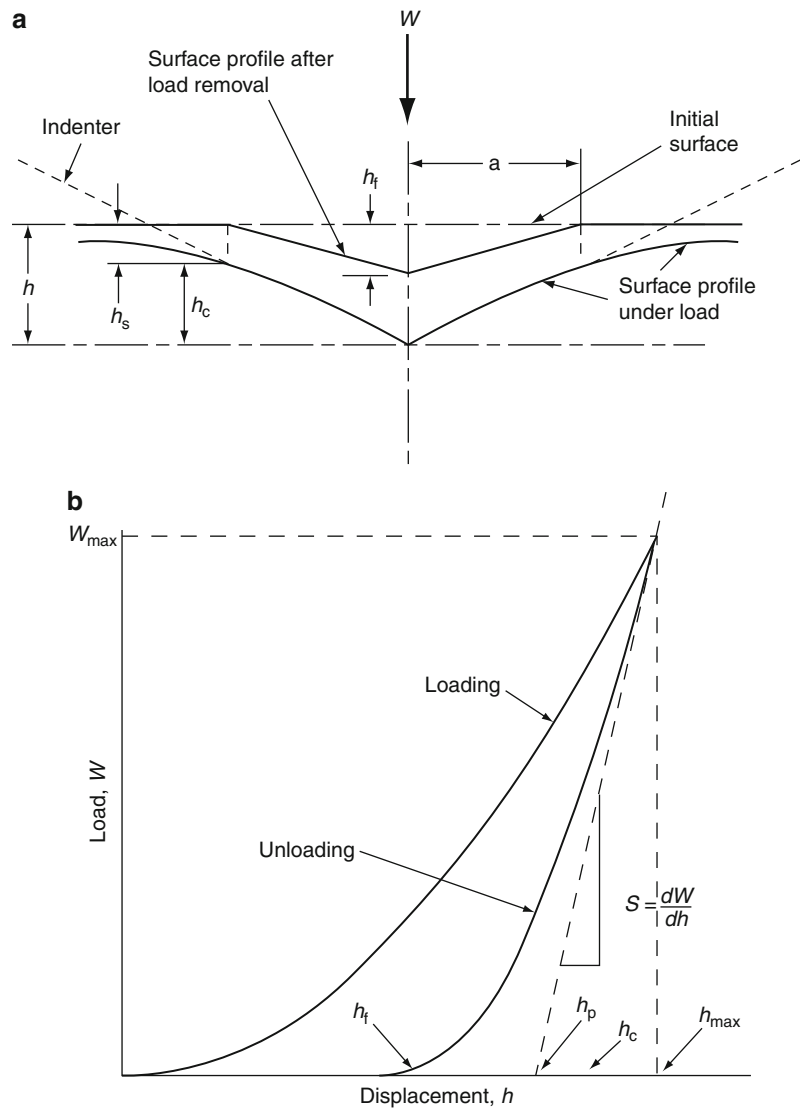
where $\varepsilon = 0.72$ for the conical indenter, $\varepsilon = 0.75$ for the paraboloid of revolution, and $\varepsilon = 1$ for the flat punch; and S_{max} is the stiffness ($=1/compliance$) equal to the slope of unloading curve (dW/dh) at the maximum load. Oliver and Pharr [10] assumed that behavior of the Berkovich indenter is similar to that of the conical indenter, since cross-sectional areas of both types of indenters varies as the square of the contact depth and their geometries are singular at the tip. Therefore, for Berkovich indenter, $\varepsilon \sim 0.72$. Thus, h_c is slightly larger than plastic indentation depth (h_p), given by

$$h_p = h_{max} - W_{max}/S_{max} \tag{2}$$

For a Vickers indenter with ideal pyramidal geometry (ideally sharp tip), projected contact area-to-depth relationship is given as [6, 11]

$$A = 24.5h_c^2 \tag{3}$$

Nanoindentation,
Fig. 3 (a) Schematic representation of the indentation process illustrating the depression of the sample around the indentation and the decrease in indentation depth upon unloading [10], and (b) schematic of load–displacement curve



Since the area-to-depth relationship is equivalent for both typical Berkovich and Vickers pyramids, Eq. 3 holds for the ideal Berkovich indenter as well. However, the indenter tip is generally rounded so that ideal geometry is not maintained near the tip. The actual indentation depth, h_c , produces a larger contact area than would be expected for an indenter with an ideal shape. For the real indenter used in actual measurements, the nominal shape is characterized by an area function $F(h_c)$, which relates projected contact area of the indenter to the contact depth

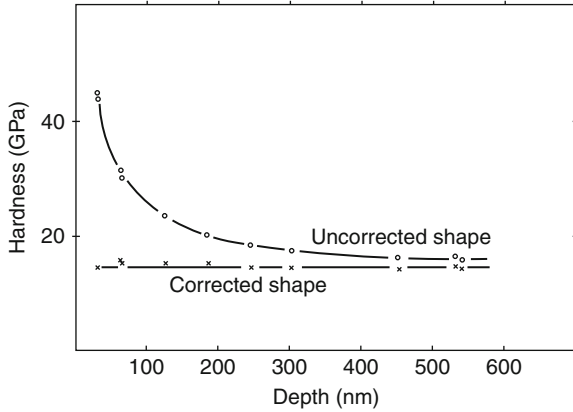
$$A^{1/2} = F(h_c) \tag{4}$$

The functional form must be established experimentally prior to the analysis.

Hardness Berkovich hardness HB (or H_B) is defined as the load divided by the projected contact area of the indentation. It is the mean pressure that a material will support under load. From the indentation curve, the hardness at maximum load can be obtained as

$$HB = W_{max}/A \tag{5}$$

where W_{max} is the maximum indentation load and A is the projected contact area at the peak load. The contact area at the peak load is determined by the geometry of



Nanoindentation, Fig. 4 Hardness as a function of indentation depth for polished single-crystal silicon (111) calculated from the area function with and without tip shape calibration [11]

the indenter and the corresponding contact depth h_c using Eq. 1 and 4. A plot of hardness as a function of indentation depth for polished single-crystal silicon (111), with and without tip shape calibration, is shown in Fig. 4. It should be noted that for this example, tip shape calibration is necessary and the hardness is independent of the corrected depth.

It should be pointed out that hardness measured using this definition may be different from that obtained from the more conventional definition in which the area is determined by direct measurement of the size of the residual hardness impression. The reason for the difference is that, in some materials, a small portion of the contact area under load is not plastically deformed, and, as a result, the contact area measured by observation of the residual hardness impression may be less than that at peak load. However, for most materials, measurements using two techniques give similar results.

Stresses during the indentation process can induce phase transformation in some ceramic materials, such as yttria-stabilized ZrO_2 . Stress-induced-phase transformation can cause a change in volume of indented material that will in turn affect the load–displacement data. A relationship between indentation pressure and phase transformation and resulting change in volume of indental material during the indentation would need to be studied.

Modulus of Elasticity Even though during loading a sample undergoes elastic–plastic deformation, the initial unloading is an elastic event. Therefore, the

Young’s modulus of elasticity or, simply, the elastic modulus of the specimen can be inferred from the initial slope of the unloading curve (dW/dh) called stiffness ($1/\text{compliance}$) (at the maximum load). It should be noted that the contact stiffness is measured only at the maximum load, and no restrictions are placed on the unloading data being linear during any portion of the unloading.

It has been found that the compliance for the indenter is approximately *independent of the shape* (with a variation of at most 3%) if the projected area is fixed [6]. The relationship for the compliance C (inverse of stiffness S) for an (Vickers, Knoop, and Berkovich) indenter is given as

$$C = \frac{1}{S} = \frac{dh}{dW} \approx \frac{1}{2E_r} \left(\frac{\pi}{A} \right)^{1/2} \quad (6)$$

where $\frac{1}{E_r} = \frac{1-\nu_s^2}{E_s} + \frac{1-\nu_i^2}{E_i}$, dW/dh is the slope of the unloading curve at the maximum load, E_r , E_s , and E_i are the reduced modulus and elastic moduli of the specimen and the indenter, and ν_s and ν_i are the Poisson’s ratios of the specimen and indenter. C (or S) is the experimentally measured compliance (or stiffness) at the maximum load during unloading, and A is the projected contact area at the maximum load.

The contact depth h_c is related to the projected area of the indentation A for a real indenter by Eq. 4. A plot of the measured compliance (dh/dW) the reciprocal of the corrected indentation depth obtained from various indentation curves (one data point at maximum load for each curve) should yield a straight line with slope proportional to $1/E_r$ [11]. E_s can then be calculated, provided Poisson’s ratio with great precision is known to obtain a good value of the modulus. For a diamond indenter, $E_i = 1,140$ GPa and $\nu_i = 0.07$ are taken.

The most frequently used method to measure initial unloading stiffness (S) is by Doerner and Nix [11], who measured S by fitting a straight line to the upper about one third portion of the unloading curve. The problem with this is that for nonlinear loading data, the measured stiffness depends on how much of the data is used in the fit. Oliver and Pharr [10] proposed a new procedure. They found that the entire unloading data are well-described by a simple power law relation:

$$W = B(h - h_f)^m, \quad (7)$$

where the constants B and m are determined by a least-square fit. The initial unloading slope is then found analytically, differentiating this expression and evaluating the derivative at the maximum load and maximum depth.

This analysis is based on an elastic solution that only accounts for sink-in (the indented material around the indenter below its original surface). However, in the more realistic case of elastic–plastic contact, sink-in or pile-up (the indented material around the indenter above its original surface) can occur depending on the specific mechanical properties of the material. For pile-up situations, the just-described method would underestimate the true contact area by as much as 50%. This in turn leads to overestimation of the hardness and elastic modulus. Based on some modeling, pile-up is significant only when $h_f/h_{max} > 0.7$ and the material does not appreciably work harden. Note that h_f/h_{max} equal to zero corresponds to fully elastic deformation and a value of 1 corresponds to rigid-plastic behavior. Compressive residual stresses result in pile-up whereas tensile stresses result in sink-in. Although some correction procedures have been proposed [12], the real contact area measurement requires imaging of indentation impressions.

Determination of Load Frame Compliance and Indenter Area Function The measured displacements are the sum of the indentation depths in the specimen and the displacements of suspending springs and the displacements associated with the measuring instruments, referred to as load frame compliance. Therefore, to determine accurately the specimen depth, load frame compliance must be known. This is especially important for large indentations made with high modulus for which the load frame displacement can be a significant fraction of the total displacement. The exact shape of the diamond indenter tip needs to be measured because hardness and elastic modulus depend on the contact areas derived from measured depths. The tip gets blunt and its shape significantly affects the prediction of mechanical properties (Fig. 4).

Oliver and Pharr [10] proposed a method for determining area functions. Their method is based only on one assumption that Young’s modulus is independent of indentation depth. They modeled the load frame and the specimen as two springs in series; thus,

$$C = C_s + C_f \tag{8}$$

where C , C_s , and C_f are the total measured compliance, specimen compliance, and load frame compliance, respectively. Combining Eqs. 6 and 8 yields

$$C = C_f + \frac{1}{2E_r} \left(\frac{\pi}{A}\right)^{1/2} \tag{9}$$

From Eq. 9, it should be noted that if the modulus of elasticity is constant, a plot of C as a function of $A^{1/2}$ is linear and the vertical intercept gives C_f . Accurate values of C_f are obtained when the specimen compliance is small, i.e., for large indentations.

To determine the area function and the load frame compliance, Oliver and Pharr made relatively large indentations in aluminum because of its low hardness. In addition, for the larger aluminum indentations (typically 700–4,000 nm deep), the area function for a perfect Berkovich indenter (Eq. 3) can be used to provide a first estimate of the contact.

Using the measured C_f value, they calculated contact areas for indentations made at shallow depths on the aluminum with measured E_r and/or on a harder fused silica surface with published values of E_r , by rewriting Eq. 9 as

$$A = \frac{\pi}{4} \frac{1}{E_r^2} \frac{1}{(C - C_f)^2} \tag{10}$$

from which an initial guess at the area function was made by fitting A as a function h_c data to an eighth-order polynomial

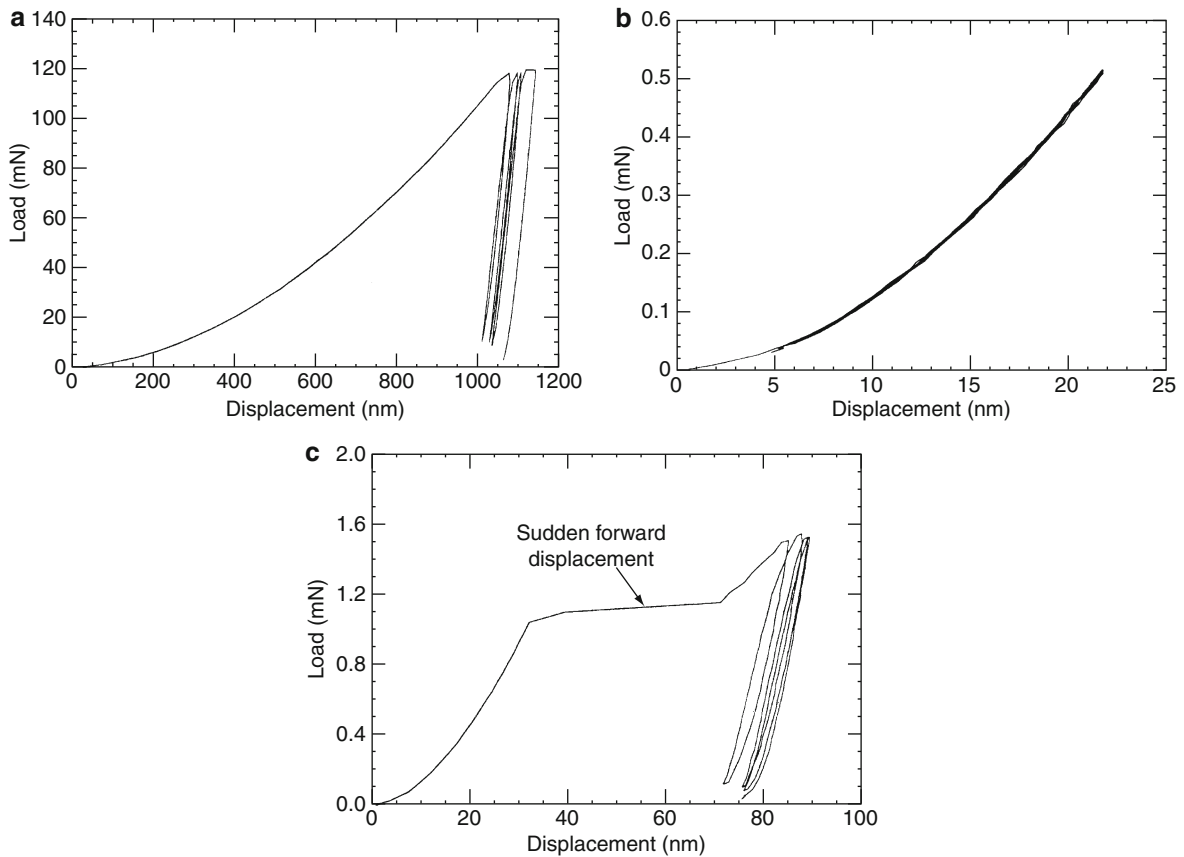
$$A = 24.5h_c^2 + C_1h_c + C_2h_c^{1/2} + C_3h_c^{1/4} + \dots + C_8h_c^{1/128} \tag{11}$$

where C_1 through C_8 are constants. The first term describes the perfect shape of the indenter; the others describe deviations from the Berkovich geometry due to blunting of the tip.

Key Research Findings

Examples of Measured Mechanical Properties

To illustrate the usage of nanoindentation techniques, typical data obtained on various bulk materials, coatings, and microstructures is presented below.



Nanoindentation, Fig. 5 Load versus displacement plot for an electro-polished single-crystal tungsten (a) at a peak load of 120 mN, (b) at a peak load of 0.5 mN (elastic contact), and (c) at a peak load of 1.5 mN showing the yield point [10]

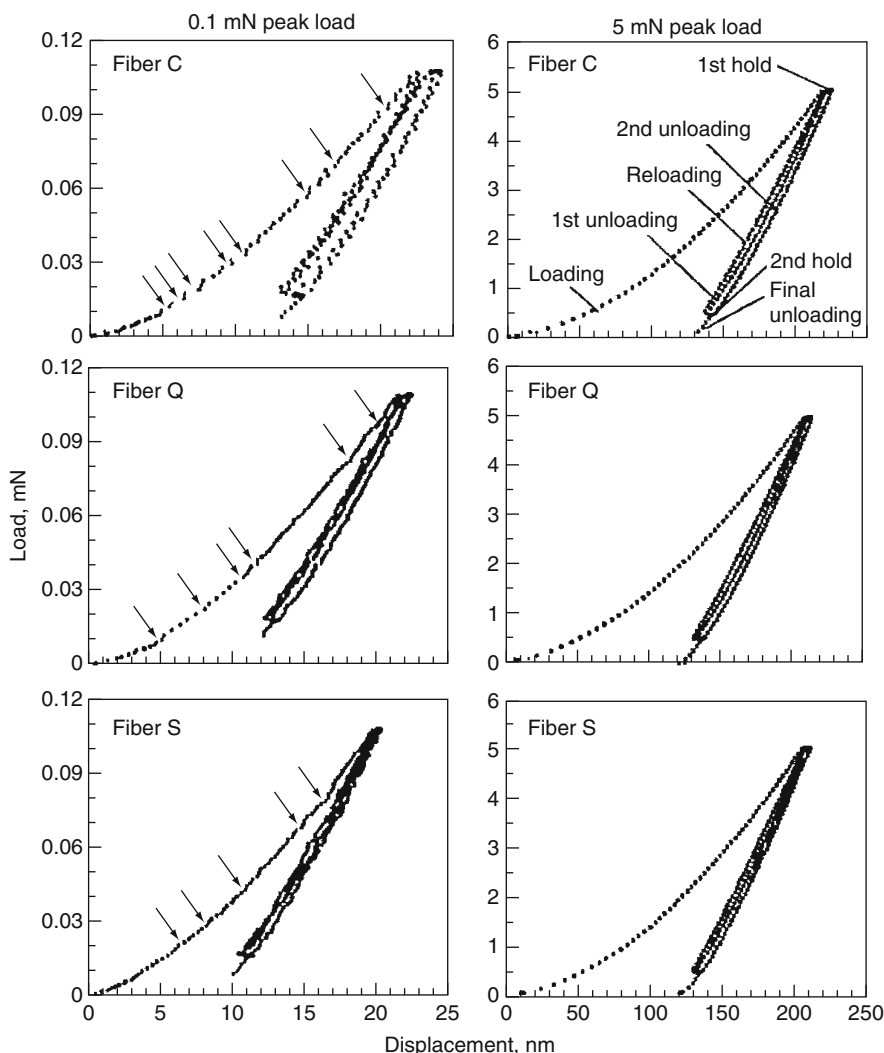
Load–Displacement Curves A variety of mechanical phenomena, such as the transition from elastic to plastic deformation, creep deformation, formation of subsurface cracks, and crystallographic phase transition, can be studied by the load–displacement curves obtained at different loading conditions [6]. The load–time sequence used by Oliver and Pharr [10] for study of various materials included three loading–unloading cycles, hold for 100 s at 10% of the peak load, reload, hold for 100 s, and unload. Load–displacement curves for electropolished single-crystal tungsten are shown in Fig. 5. The tungsten data are typical of materials in which the hardness is relatively small compared to the modulus, as is observed in most metals; most of the indenter displacement in these metals is accommodated plastically and only a small portion is recovered on unloading. Figure 5 shows that the peak load displacements shift to higher values in successive loading/unloading cycles. In addition, the relatively large

displacement just prior to final unloading is due to creep during the 100-s hold period at peak load. Indentation at a very low load of 0.5 mN caused only elastic displacements (Fig. 5b). At higher peak loads, the indentation is not just elastic (Fig. 5c). When a threshold load of about 1 mN is reached, a sudden jump in displacement corresponding to the onset of plasticity is observed, and a permanent hardness impression is formed. Tungsten at low loads (Fig. 5c) exhibit distinct hysteresis loops, as might be expected if there were a small amount of reverse plasticity upon loading. However, the looping degenerates with cycling after three or four cycles, the load–displacement behavior is largely elastic.

Another set of load–displacement curves is shown in Fig. 6 for glass fibers, which are used as reinforcements in numerous materials including plastic, rubber, cement as well as heat insulation materials [13]. The indents were made on polished cross sections of the

Nanoindentation,

Fig. 6 Representative load versus displacement plots of indentations made at 0.1 and 5 mN peak indentation loads on various glass fibers [13]



fibers. Fibers C, Q, and S are all 16 μm diameter fibers, each with different formulations. The load–displacement plots of indentations made at 0.1 mN peak loads on all fibers exhibit several displacement discontinuities of pop-in marks during the first and second loading steps of the indentation, as denoted by the arrows in Fig. 6. Pop-in marks in the loading curves result from sudden penetrations of the tip into the sample. Fiber S exhibits the smallest pop-in marks, followed by fibers Q and C, which suggests that Fiber S has the highest resistance to plastic deformation as well as crack formation and crack propagation.

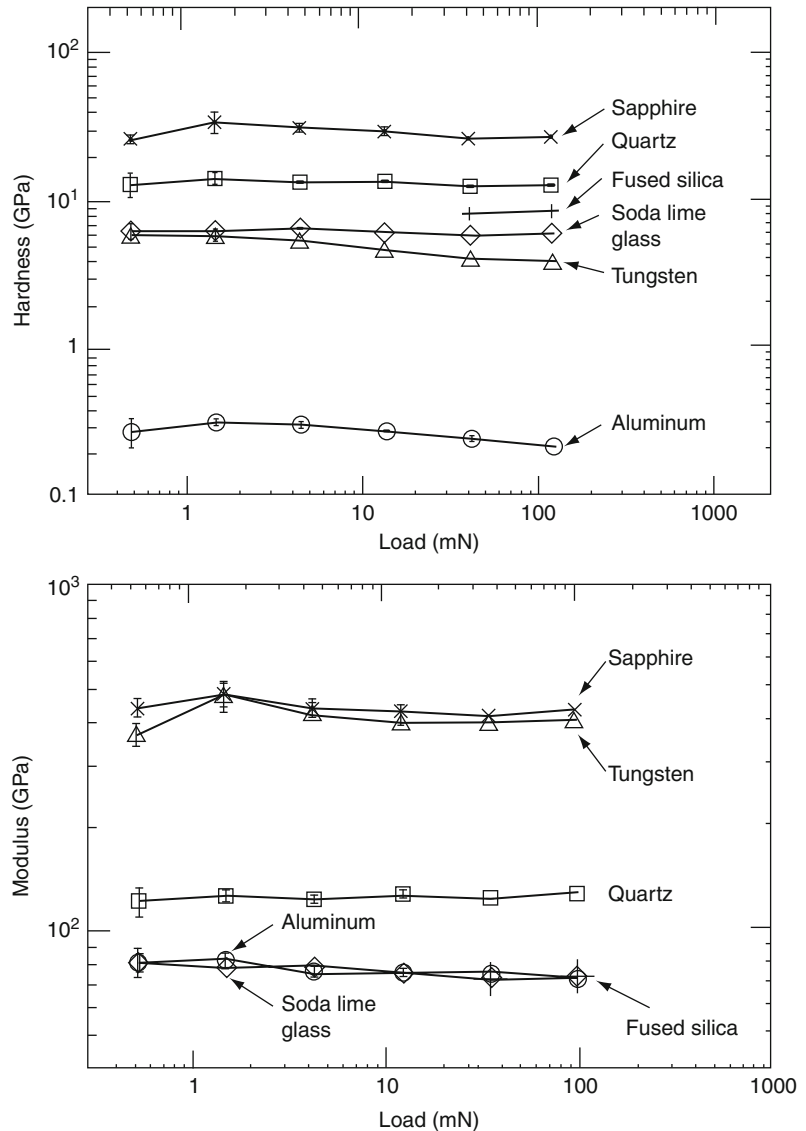
Hardness and Elastic Modulus Measurements The nanoindenter is commonly used to measure surface mechanical properties of bulk materials [6, 10].

Hardness and elastic moduli for six bulk materials and three materials used as a substrate for the construction of magnetic rigid disks and single-crystal silicon are presented in Fig. 7. The data show that there is a very little indentation size effect in several materials on the hardness values. In the case of aluminum and tungsten in Fig. 7, there is a modest increase in hardness at low loads, which could be due to surface-localized cold work resulting from polishing [10]. The modulus data also show that there is very little evidence for an indentation size effect; i.e., the moduli remain more or less constant over the entire range of load.

The nanoindenter is ideal for the measurement of mechanical properties of thin films and composite

Nanoindentation,

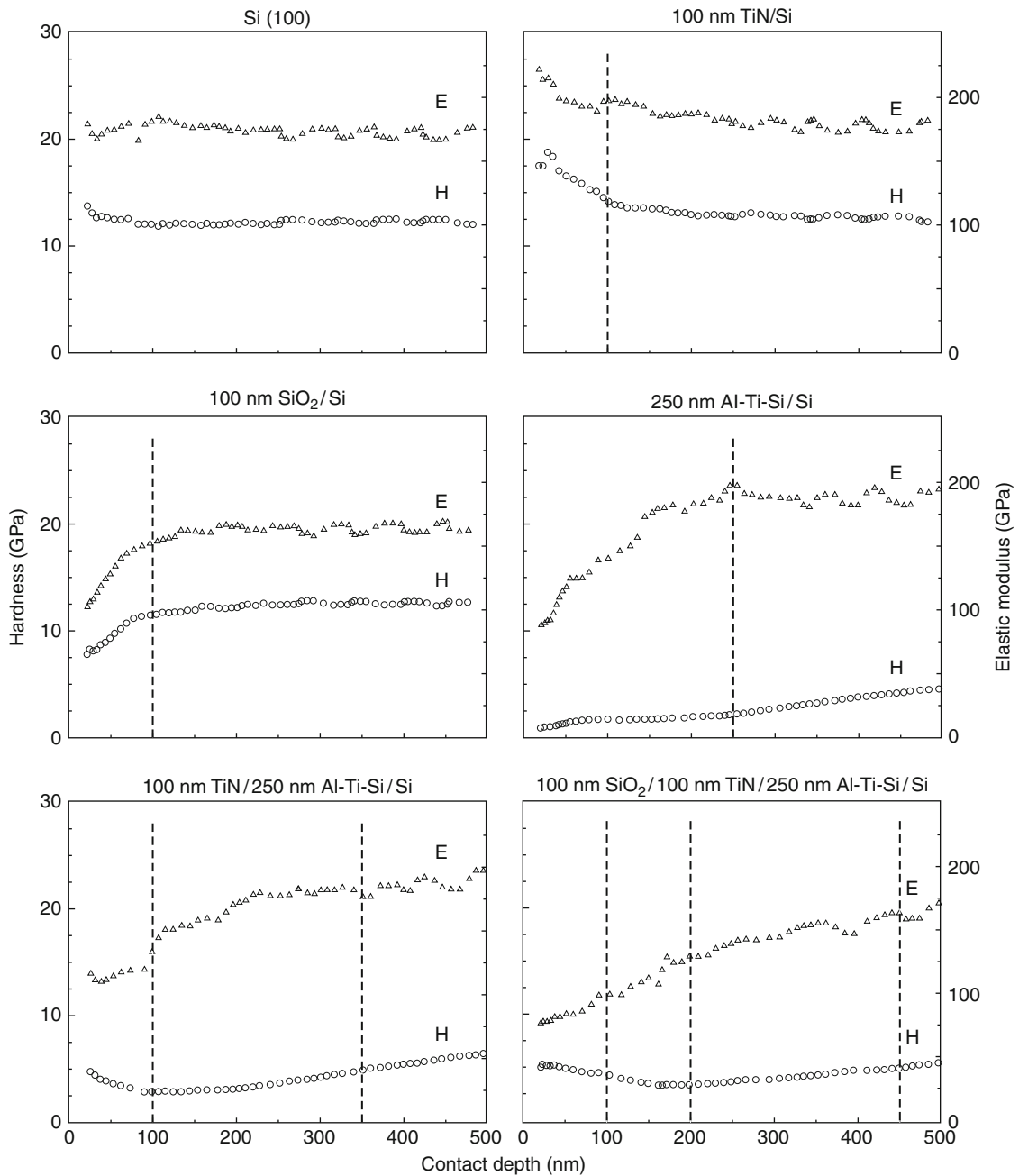
Fig. 7 Hardness and elastic modulus as a function of load for six bulk materials – mechanically – polished single-crystal aluminum, electro-polished single-crystal tungsten, soda-lime glass, fused silica, (001) single-crystal sapphire, and (001) single-crystal quartz [10]



structures. A number of investigators have reported hardness of composite structures with thin films on a substrate. It is widely accepted that to measure true hardness of the films, the indentation depth should not exceed 10% of the film thickness [14]. However, based on a finite element analysis of the indentation of thin films of various thicknesses it was found that the true hardness of the films could be obtained if the indentation depth does not exceed about 30% of the film thickness. At higher indentation depths, the composite hardness changes with the indentation depth [6].

The significance of the composite hardness and modulus on multilayered thin films is illustrated for

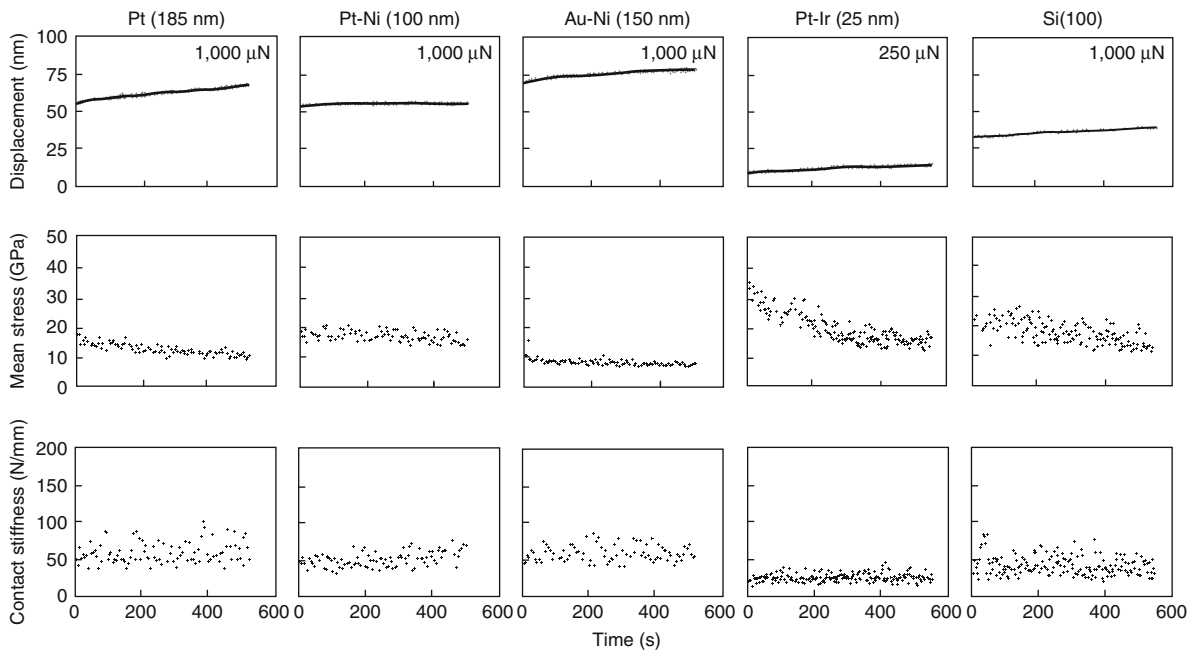
the components of the digital micromirror device (DMD). The DMD is the main component of commercial DLP™ projectors, invented at Texas Instruments (TI) in 1987. The DMD chip is composed of an array of micromirrors, along with a yoke and hinge array underneath. The contact between the tips at the end of the yoke and the underlying landing site is required for operation. Since the DMD chip is composed of various thin film structures, the mechanical properties of the component materials are important to the DMD performance. Figure 8 shows the hardness and elastic modulus as a function of contact depth for the DMD materials measured using the continuous stiffness



Nanoindentation, Fig. 8 Hardness and elastic modulus as a function of contact depth for the thin films used in the digital micromirror device (DMD) [15]

method, with data for Si (1 0 0) included for reference. The samples belong to three categories: bulk material, single-layer, and multilayered materials. For bulk material, Si (1 0 0), the hardness and elastic modulus does not change with contact depth, as expected. For a hard film on soft substrate, TiN/Si, the hardness and

elastic modulus values decrease with contact depth. For soft films on hard substrates, SiO₂/Si and Al alloy/Si, their hardness and elastic modulus values increase with contact depth. The hardness and elastic modulus of TiN/Si and SiO₂/Si gradually approach the hardness and elastic modulus values of Si (1 0 0) after



Nanoindentation, Fig. 9 Creep displacement, mean stress and contact stiffness as a function of time of the various metal-coated AFM probes used in ferroelectric data recording and Si(100) as reference [16]

the contact depths exceed 100 nm, due to the substrate effect [15].

For multilayered materials, TiN/Al alloy/Si and SiO₂/TiN/Al alloy/Si, there are two common observations. For both samples, the lowest hardness values were obtained near the TiN/Al alloy interfaces, which are approached at the contact depth of about 100 and 200 nm, respectively. It is observed that both samples show much lower hardness values (~ 5 GPa) at the contact depth of about 30 nm as compared to the corresponding single-layer materials. The substrate effect contributions are expected to be small at this contact depth.

It is known that for a metallic thin film, the measured hardness is sensitive to grain size, film texture, film thickness, level of adhesion of the film to the substrate, and presence of residual stress. When TiN was reactively sputtered onto Al alloy, the stress state and microstructure (grain size, texture) of TiN might be changed as compared to TiN/Si, since the film growth substrate is changed. According to the Hall–Petch relationship, in metal and alloy films, if the grain size of the TiN/Al alloy/Si was greater than that of the TiN/Si, then the hardness would decrease. It is believed that films grown by physical and chemical

vapor deposition techniques usually display a preferred orientation, which is dependent upon deposition process variables. Hardness reductions by as much as two orders have been observed in cubic coatings (e.g., TiC, TiN, ZrC) as a function of the preferred orientation plane [e.g., (1 1 1), (1 0 0) and (1 1 0)]. For SiO₂/TiN/Al alloy/Si, since SiO₂ is usually amorphous when deposited as films, the grain size and texture may not be the parameters responsible for the hardness decrease of SiO₂/TiN/Al alloy/Si [15].

Creep Resistance Most materials, including ceramics (and even diamond) are found to creep at temperatures well below half their melting points, even at room temperature. Indentation creep and indentation load relaxation (ILR) tests are commonly used to measure the time-dependent flow of materials. These techniques offer an advantage of being able to probe the deformation properties of a thin film as a function of indentation depth and location [6].

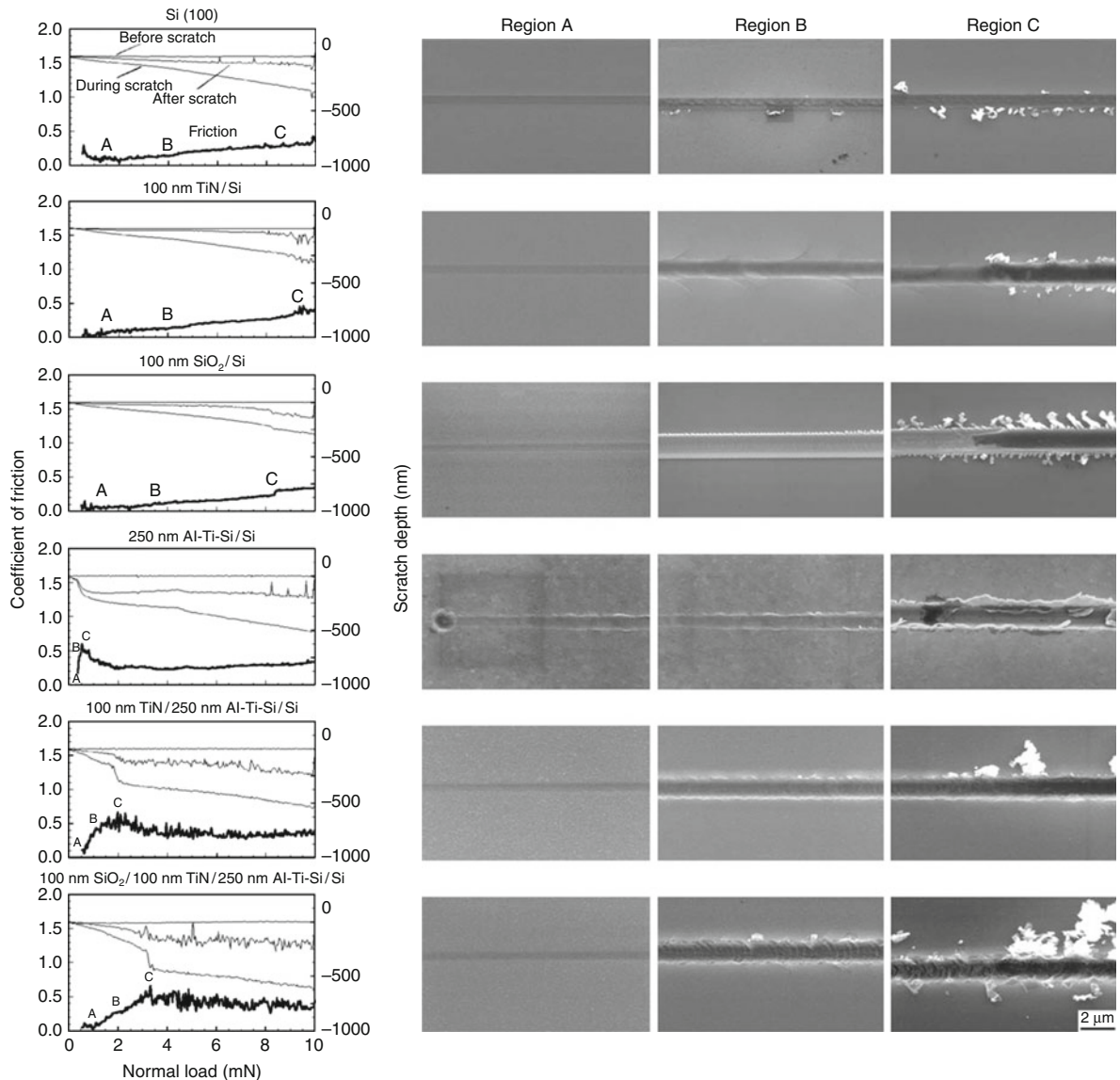
An example of nanoscale creep is shown in Fig. 9 for the noble metal thin films used for probe-based ferroelectric data recording. In this technology, a conductive atomic force microscope (AFM) probe with a noble metal coating is placed in contact with lead zirconate titanate (PZT) film, which serves as the

ferroelectric material. The long-term device performance will be compromised if the metal films exhibit excessive creep [16]. Figure 9 shows creep data for the various metal-coated probes and Si, where the change in displacement, mean stress, and contact stiffness was monitored while holding the tip at the maximum imposed load (P_{\max}). A lower P_{\max} was applied on Pt–Ir as it is a very thin coating; the selected load of 250 mN ensures that the tip is measuring the creep resistance of the coating and did not penetrate the substrate during the creep experiment. From the displacement plots, it is observed that the three alloyed coatings (Pt–Ni, Pt–Ir, and Au–Ni) exhibited less creep compared to Pt. The Si surface also shows significant creep, which is consistent with previous reports. This creep on silicon is a combination of two factors: initial creep can occur as the tip penetrates the native oxide surface, which is then followed by a dislocation glide plasticity mechanism [16, 17].

Microscratch Resistance of Various Materials Using Micro/Nanoscratch Technique The microscratch technique is commonly used for screening of bulk materials for wear resistance. As stated earlier, normal load applied to the scratch tip is gradually increased during scratching until the material is damaged. Friction force can be measured during the scratch test [7]. After the scratch test, the morphology of the scratch region including debris is observed in an SEM. Based on the combination of the changes in the friction force as a function of normal load and SEM observations, the critical load is determined and the deformation mode is identified. Any damage to the material surface as a result of scratching at a critical ramp-up load results in an abrupt or gradual increase in friction. The material may deform either by plastic deformation or fracture. Ductile materials (all metals) deform primarily by plastic deformation, resulting in significant plowing during scratching. Tracks are produced whose width and depth increase with an increase in the normal load. Plowing results in a continuous increase in the coefficient of friction with an increase in the normal load during scratching. Debris is generally ribbon-like or curly, whereas brittle materials deform primarily by brittle fracture with some plastic deformation. In the brittle fracture mode, the coefficient of friction increases very little until a critical load is reached at which the material fails catastrophically and produces fine debris, which is rounded, and the coefficient of friction increases rapidly above the critical load.

In the example for the digital micromirror device (DMD) mentioned above, the repeated contact between the yoke and landing site components eventually leads to wear over time. The wear resistance of these materials can be evaluated through a nanoscratch experiment. For the components of the DMD, Fig. 10 shows the coefficient of friction and scratch depth profiles as a function of increasing normal load, along with SEM images of three regions over scratches [15]. SEM images were taken at the beginning of the scratch (indicated by “A” on the friction profile), at the point of initiation of damage at which the coefficient of friction increases by a factor of two (indicated by “B” on the friction profile), and at the point of the initiation of the delamination of the films (for Si substrate, this region represents the position where severe damage began) (indicated by “C” on the friction profile). The normal load corresponding to region B is defined as the “critical load.” Before the critical load (from regions A to B), the coefficient of friction of the Si (1 0 0), TiN and SiO₂ films increased at a slower rate, and at the critical load (region B), these three samples exhibit a small increase in the coefficient of friction. However, for the remaining samples, i.e., Al alloy, TiN/Al alloy, and SiO₂/TiN/Al alloy, the coefficient of friction increased sharply from regions A to B. The slopes of the increase of the coefficient of friction in these three samples from regions A to B are inversely proportional to the hardness of the samples. For instance, the Al alloy shows the lowest hardness, but highest slope of the increase of the coefficient of friction. The increase of the coefficient of friction during scratching was attributed to the increasing plowing of the sample by the tip with increasing normal load, and the softer material led to deeper plowing of the tip and faster increase of the coefficient of friction. At region C, all films were delaminated and experienced an abrupt increase in the coefficient of friction, which was associated with catastrophic failure as well as significant plowing of the tip into the sample.

The SEM images show that all the samples experienced plastic deformation at the beginning of the scratch (region A). For Al alloy, because it is very soft, region A is very close to region B. The indent of the conical tip, the plastic deformation associated with the pile-up was observed at the beginning of the scratch for Al alloy. At region C, for Si (1 0 0) substrate, only some debris particles were found without any cracks on the side of the scratch, which is

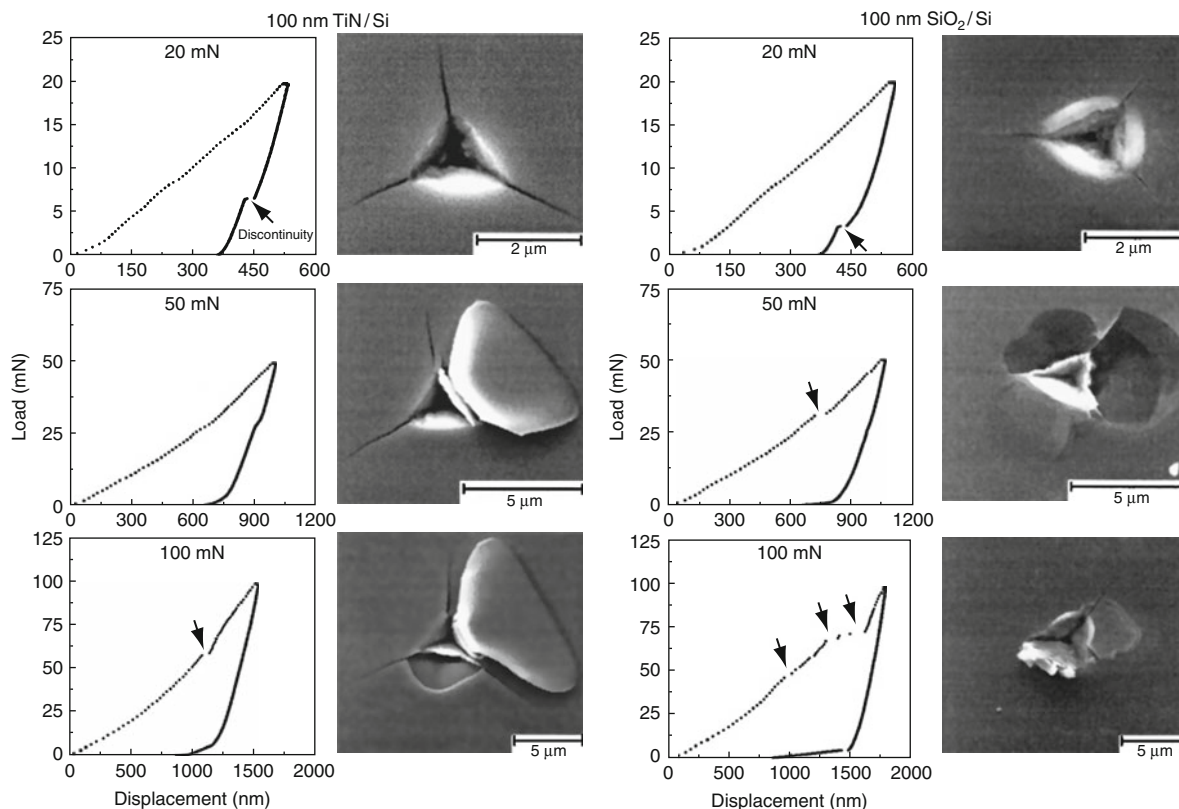


Nanoindentation, Fig. 10 Coefficient of friction and scratch depth profiles as a function of normal load for DMD thin films and SEM images of three regions over scratches: at the beginning of the scratch (indicated by “A” on the friction profile), at

the point of damage initiation (indicated by “B” on the friction profile), and at the point of film delamination (indicated by “C” on the friction profile). For Si substrate, region C is where severe damage began) [15]

responsible for the small increase in the coefficient of friction. For TiN, SiO₂ and Al alloy films, it is clear that at region C, the scratch tip reached the Si substrate and delaminated the film. The irregular cracks found on the side of the scratch in TiN indicate that TiN is brittle. However, it is interesting to observe that the SiO₂ film shows high ductility according to the SEM images at region C, where regular film pieces were found to attach to the sample surface even after the

film was delaminated. As expected, severe plowing and high pile-up were found for Al alloy. For TiN/Al alloy and SiO₂/TiN/Al alloy, a lot of debris were found at region C, indicating that more materials were torn away from the films. By investigating the corresponding scratch depths, it can be seen that at region C, the in situ scratch depth of TiN/Al alloy and SiO₂/TiN/Al alloy are about 350 and 450 nm, respectively, indicating that for both multilayer films,



Nanoindentation, Fig. 11 Load–displacement curves of indentations made at 20, 50, and 100 mN peak loads and corresponding SEM micrographs of the indents on two of the films used in the DMD. Solid arrows indicate discontinuity

during the unloading portion of the curve and the hollow arrows indicate a step during the loading portion of the load–displacement curve [19]

the tip reached the Si substrates. This may explain why more material was torn away from the multilayered films at region C. For the thin film materials used in DMD, the damage initiated at the “critical load” should be prevented in order to make the DMD operate properly. Therefore, the film topography and mechanical property changes due to scratch in the region from A to B, especially the film failure mechanism at region B, are of great interest for DMD [15].

Nanofracture Toughness Nanoindentation can be used to determine fracture toughness [6]. The indentation cracking method is especially useful for measurement of fracture toughness of thin films or small volumes. For fracture toughness measurement of ultrathin films ranging from 100 nm to few micrometers, traditional indentation or four-point flexure techniques cannot be used. Because of shallow indentation depths required in the indentation technique, it is difficult to measure a radial crack length under even SEM.

Li et al. [18] developed a novel technique based on nanoindentation in which through-thickness cracking in the coating is detected from a discontinuity in the load–displacement curve and energy released during cracking is obtained from the curve. Based on the energy released, fracture mechanics analysis is used to calculate fracture toughness. A cube corner is preferred because the through-thickness cracking of hard films can be accomplished at lower loads.

An example of the fracture toughness determination for the thin films used in DMD is shown in Fig. 11, which contains load–displacement curves of indentations made at 20, 50, and 100 mN peak indentation loads together with the scanning electron microscopy (SEM) micrographs of indentations on the TiN/Si and SiO₂/Si. Steps are found in the loading curves obtained at 100 mN for TiN/Si, and at 50 mN (and 100 mN) for SiO₂ as shown by arrows in Fig. 11. The steps in the loading curves of indentations result from the ring-like

through-thickness cracking as shown in the SEM images. These ring-like through-thickness cracking were used to calculate the fracture toughness together with the strain energy release obtained from the steps in the loading curves, which will be described later. At 20 mN, the radial cracks were found in both TiN/Si and SiO₂/Si. The discontinuity in the displacement observed during unloading is believed to be due to the formation of lateral cracks at the base of the median crack which results in the surface of the specimen being thrust upward [19].

Li et al. [18] suggest that the steps in the loading curve on the films result from the film cracking, and that the fracture process progresses in three stages: (1) first ring-like through-thickness cracks form around the indenter by high stresses in the contact area; (2) delamination and buckling occur around the contact area at the film/substrate interface by high lateral pressure; and (3) second ring-like through-thickness cracks and spalling are generated by high bending stresses at the edges of the buckled film. The fracture toughness was calculated from the second ring-like through-thickness cracking generated in the third stage. In this stage, the stress concentration at the end of the interfacial crack cannot be relaxed by the propagation of the interfacial crack. With an increase in indentation depth, the height of the bulged film increases. When the height reaches a critical value, the bending stresses caused by the bulged film around the indenter will result in the second ring-like through-thickness crack formation and spalling at the edge of the buckled film, which leads to a step in the loading curve. The step in the loading curve is totally from the film cracking and not from the interfacial cracking or the substrate cracking [19].

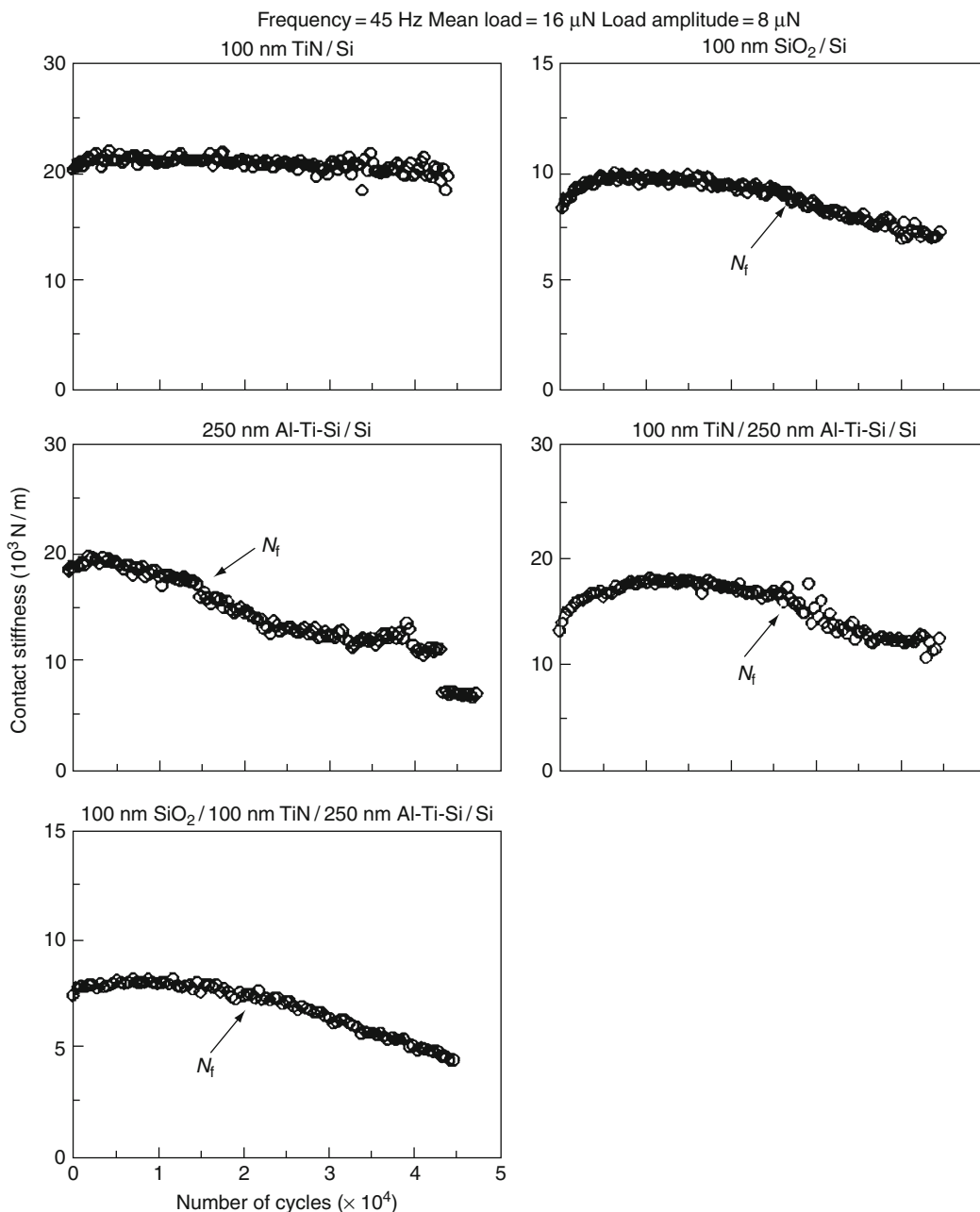
Nanofatigue Fatigue fracturing progresses through a material via changes within the material at the tip of a crack, where there is high stress intensity. There are several situations: cyclic fatigue, stress corrosion, and static fatigue. In this discussion, the focus is on cyclic fatigue, which results from the cyclic loading of machine components; e.g., the stresses cycle from tension and compression occurs in a loaded rotating shaft.

To compare the fatigue lives of the thin film structures used in the DMD, the contact stiffness data as a function of the number of cycles for TiN/Si, SiO₂/Si, Al alloy/Si, TiN/Al alloy/Si, and SiO₂/TiN/Al alloy/Si cyclically deformed at an oscillating load amplitude of

8 mN with a mean load of 16 mN at a frequency of 45 Hz are shown in Fig. 12. For all the samples, the contact stiffness increased to varying degrees at the beginning of the fatigue test as a result of strain hardening. TiN/Si and Al alloy/Si films have the longest and shortest N_f , respectively. By comparing the N_f of multilayered thin film structures with the corresponding single top layer films (~TiN and SiO₂), it can be seen that the multilayered thin film structures exhibited shorter N_f . Ranking of the nanoscale fatigue data is similar to results from hardness studies, i.e., the hardness of multilayered thin film structures is lower than that of the corresponding single top layer films [19].

The relationship between fatigue life and hardness is as follows. At the beginning of the fatigue test, the tip penetrates into the sample a certain displacement, depending on the hardness of the sample. Then the tip is held and a cyclic load is applied. As the number of cycles increases, the tip penetrates deeper. It is known that during the indentation fatigue test, even though there may be no cracks, the tip of the sharp indenter can sink deeper into the material and extend the plastic zone. At this stage, the compressive stress and the deformation volume may not be high enough to cause cracks. In this case, the contact stiffness would not decrease abruptly. It is known that voids and impurities act as original cracks. As the number of cycles reaches a critical value N_f , the original cracks beneath and around the tip would propagate due to the accumulated deformation, as well as new cracks that may develop at high deformation volume. This would cause the contact stiffness to drop suddenly. Since TiN/Si is the hardest sample, it resisted the tip to sink and consequently results in lower accumulated deformation volume. So it took a longer time for TiN/Si to fail due to the fatigue-induced crack. For softer materials, SiO₂/Si and Al alloy/Si, the tip sank easily and caused larger deformation volume and earlier fatigue failure. This explanation can also be used to understand the data that multilayered thin film structures with soft interlayer showed shorter fatigue life than the corresponding top hard films. The hardness of multilayered structures is lower than the top hard films responsible for shorter fatigue life [19].

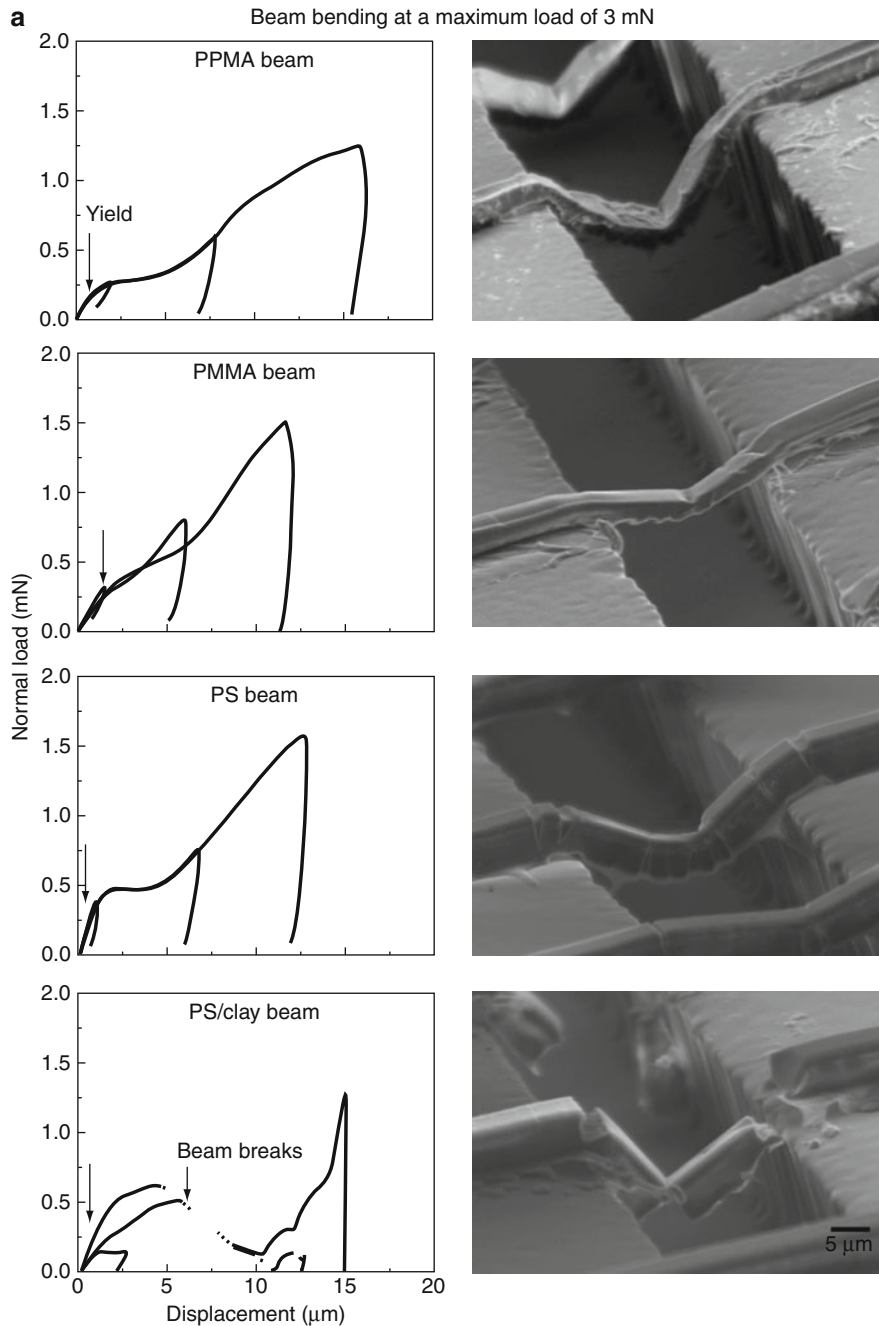
Deformation of Microbeams Microscale structures, such as beams, are common in MEMS/NEMS devices. In this example, the focus is on polymer microbeams designed for the BioMEMS device known as a cell force sensor. In the work by Palacio et al. [20], the



Nanoindentation, Fig. 12 Contact stiffness as a function of the number of cycles for various thin films used in the DMD [19]

materials considered were poly(propyl methacrylate), poly(methyl methacrylate), poly(styrene), and a poly(styrene)-nanoclay composite (abbreviated as PPMA, PMMA, PS and PS/clay, respectively). For this application, nanoindentation can be used for conducting bending experiments in the normal and lateral

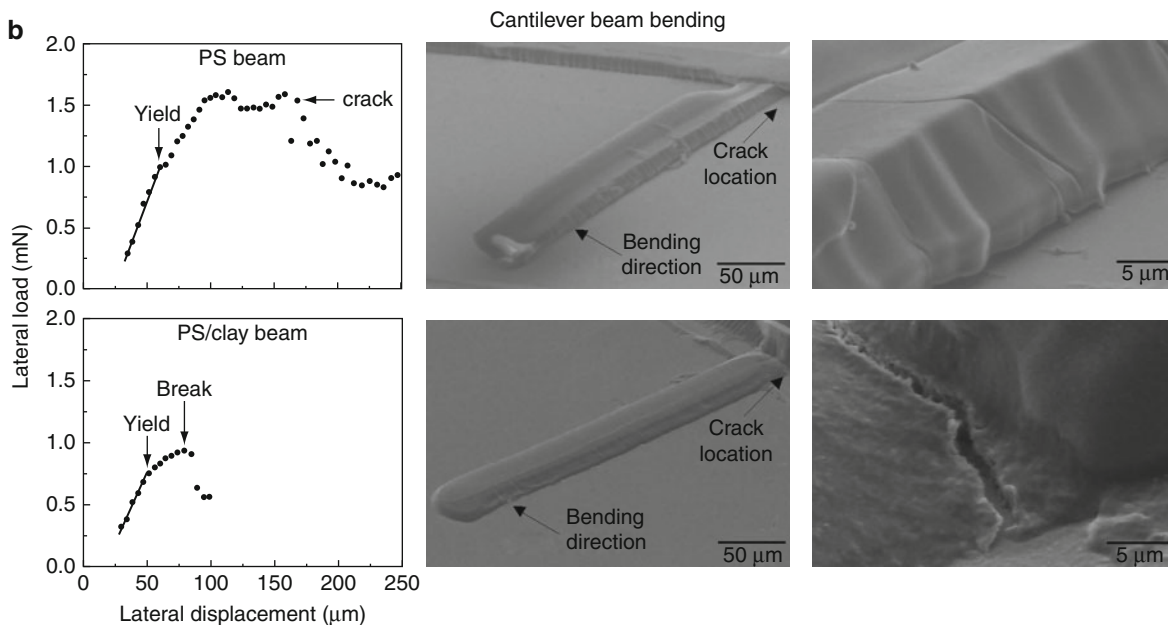
directions. These experiments provide information on the elastic modulus, yield strength, and breaking strength of these materials. This example illustrates how nanoindentation can be used beyond the more traditional experiments carried out on bulk materials and thin films.



Nanoindentation, Fig. 13 (continued)

Bending experiments were performed by Palacio et al. [20] with applied loads in the millinewton range in order to investigate the behavior of polymer beams beyond the elastic deformation regime. [Figure 13a](#)

shows normal load–displacement profiles at maximum imposed loads of 0.5, 1.5 and 3.0 mN applied on three different beams in one wafer sample. For PPMA, PMMA and PS, the loading portion overlaps,



Nanoindentation, Fig. 13 (a) Load–displacement curves for the normal beam bending of the PPMA, PMMA, PS and PS/clay polymer beams for use in the cell force BioMEMS sensor device. SEM images correspond to the beams subjected to 3 mN load. (b) Lateral bending of PS and PS/clay cantilever beams. The

lateral force is plotted as a function of the lateral displacement, and the arrows indicate the onset of yield for both PS and PS/clay, cracking for PS and breaking for PS/clay. SEM images indicate the bending direction and the crack locations [20]

indicating uniform beam quality. Slight variability in the loading profile is observed in the PS/clay composite, which is attributed to variable interfacial adhesion between the filler and the matrix and inhomogeneities in the dispersion of the clay particles.

The first inflection in the load–displacement profiles implies the onset of yield and its location is indicated by arrows. The three unfilled polymers (PPMA, PMMA and PS) exhibit ductility, as seen in the SEM images for the beams tested with an applied load of 3 mN. The beams stretch at the ends and underneath the center of the beam. These correspond to regions subjected to maximum tensile stress. The beams deformed symmetrically, indicating that the load was applied equidistant from the clamping points. The PS/clay nanocomposite beam exhibited slight yielding, followed by breaking, as indicated by an arrow in the load–displacement profile. After breaking, the load continued to increase because a predefined load is imposed by the nanoindenter. The corresponding SEM image for PS/clay shows a flat fracture surface, indicating that the crack propagated perpendicular to the direction of the applied stress [20].

From the load–displacement curves, the yield strength can be evaluated by using [20, 21]:

$$\sigma = \frac{3 Fl}{4 bh^2} \tag{12}$$

where F is the load corresponding to yield of the beam (as indicated by arrows in the load–displacement profile), l , b , and h are the beam’s length, width and height, respectively. It is assumed that the beam ends are clamped.

Figure 13b presents the results of the lateral bending experiments on PS and PS/clay cantilever beams. On the left column are the plots for the lateral force recorded as a function of the lateral displacement. From the linear elastic regime (as indicated by the straight line drawn over the data points), the modulus can be evaluated. For a cantilever beam with one end clamped, it is expressed as [20, 21]:

$$E = \frac{l^3}{3I} m \tag{13}$$

where l is the beam length, I is the area moment of inertia for the beam cross section, and m is the slope of the linear region of the force–displacement curve.

From the load–displacement curves, the yield strength can be evaluated by using [20, 21]:

$$\sigma = \frac{6Fl}{bh^2} \quad (14)$$

where F , l , b , and h are defined similarly as in Eq. 12. The force used for calculating the yield strength is taken at the point where a change in slope was observed (as indicated by the arrows on the left in Fig. 13b).

For PS, a steady decrease in the lateral load was observed at 160 mm, while for the PS/clay beam, the drop at 80 mm was abrupt (as indicated by the arrows on the right). SEM images reveal that for PS, this corresponds to crack formation at the base of the beam that did not run throughout the entire thickness of the beam. However for PS/clay, the beam broke, which implies that the addition of the nanoclay filler induces embrittlement of the beam.

Cross-References

- ▶ [Atomic Force Microscopy](#)
- ▶ [Friction Force Microscopy](#)

References

1. Bhushan, B.: Principles and Applications of Tribology. Wiley, New York (1999)
2. Bhushan, B.: Modern Tribology Handbook, vols. 1 and 2. CRC Press, Boca Raton (2001)
3. Bhushan, B.: Introduction to Tribology. Wiley, New York (2002)
4. Bhushan, B.: Springer Handbook of Nanotechnology, 3rd edn. Springer, Heidelberg (2010)
5. Bhushan, B.: Nanotribology and Nanomechanics I – Measurement Techniques and Nanomechanics, II – Nanotribology, Biomimetics, and Industrial Applications, 3rd edn. Springer, Heidelberg (2011)
6. Bhushan, B., Li, X.: Nanomechanical characterization of solid surfaces and thin films. *Int. Mater. Rev.* **48**, 125–164 (2003)
7. Bhushan, B., Gupta, B.K., Azarian, M.H.: Nanoindentation, microscratch, friction and wear studies of coatings for contact recording applications. *Wear* **181–183**, 743–758 (1995)
8. Gerberich, W.W., Gao, H., Sundgren, J.-E., Baker, S.P. (eds.): Thin Films: Stresses and Mechanical Properties VI. Symposium proceedings, materials research society, vol. 436. Materials Research Society, Pittsburgh (1996)
9. Johnson, K.L.: Contact Mechanics. Cambridge University Press, Cambridge (1985)
10. Oliver, W.C., Pharr, G.M.: An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *J. Mater. Res.* **7**, 1564–1583 (1992)
11. Doerner, M.F., Nix, W.D.: A method for interpreting the data from depth-sensing indentation instruments. *J. Mater. Res.* **1**, 601–609 (1986)
12. Tsui, T.Y., Pharr, G.M.: Substrate effects on nanoindentation mechanical property measurement of soft films on hard substrates. *J. Mater. Res.* **14**, 292–301 (1999)
13. Li, X., Bhushan, B., McGinnis, P.: Nanoscale mechanical characterization of glass fibers. *Mater. Lett.* **29**, 215–220 (1996)
14. Tabor, D.: The Hardness of Metals. Clarendon, Oxford (1951)
15. Wei, G., Bhushan, B., Jacobs, S.J.: Nanomechanical characterization of multilayered thin film structures for digital micromirror devices. *Ultramicroscopy* **100**, 375–389 (2004)
16. Palacio, M., Bhushan, B.: Nanotribological and nanomechanical properties of lubricated PZT thin films for ferroelectric data storage applications. *J. Vac. Sci. Technol. A* **26**, 768–776 (2008)
17. Bhushan, B., Kulkarni, A., Bonin, W., Wyrobek, J.T.: Nanoindentation and picondentation measurements using a capacitive transducer system in atomic force microscopy. *Philos. Mag. A* **74**, 1117–1128 (1996)
18. Li, X., Diao, D., Bhushan, B.: Fracture mechanisms of thin amorphous carbon films in nanoindentation. *Acta Mater.* **45**, 4453–4461 (1997)
19. Wei, G., Bhushan, B., Jacobs, S.J.: Nanoscale fatigue and fracture toughness measurements of multilayered thin film structures for digital micromirror devices. *J. Vac. Sci. Technol. A* **26**, 768–776 (2004)
20. Palacio, M., Ferrell, N., Hansford, D., Bhushan, B.: Nanomechanical characterization of polymer beam structures for BioMEMS applications. *Sens. Actuator A* **135**, 637–650 (2007)
21. Young, W.C., Budynas, R.G.: Roark’s Formulas for Stress and Strain. McGraw-Hill, New York (2002)

Nanointerconnection

- ▶ [Nanorobotic Spot Welding](#)

Nanojoining

- ▶ [Nanorobotic Spot Welding](#)

Nanomachining

- ▶ [Ultraprecision Machining \(UPM\)](#)

Nanomagnetism

- ▶ [Magnetic Nanostructures and Spintronics](#)

Nanomaterials

- ▶ [Computational Study of Nanomaterials: From Large-Scale Atomistic Simulations to Mesoscopic Modeling](#)
- ▶ [Integration of Nanostructures within Microfluidic Devices](#)
- ▶ [Nanostructures for Energy](#)

Nanomaterials and Nanoproducts

- ▶ [Physicochemical Properties of Nanoparticles in Relation with Toxicity](#)

Nanomaterials as Field Emitters

- ▶ [Field Electron Emission from Nanomaterials](#)

Nanomaterials for Electrical Energy Storage Devices

Shrikant C. Nagpure and Bharat Bhushan
Nanoprobe Laboratory for Bio- & Nanotechnology
and Biomimetics, The Ohio State University,
Columbus, OH, USA

Synonyms

[Accumulator](#); [Batteries](#); [Electric double layer capacitor](#); [Fuel cell](#); [Nanostructured materials](#)

Definition

An electrical energy storage device stores the input energy and delivers it in the form of electrical energy when needed.

A material that is engineered with morphological features on the order of less than 100 nm and exhibits different characteristics compared to its bulk counterpart is defined as a nanomaterial.

Overview

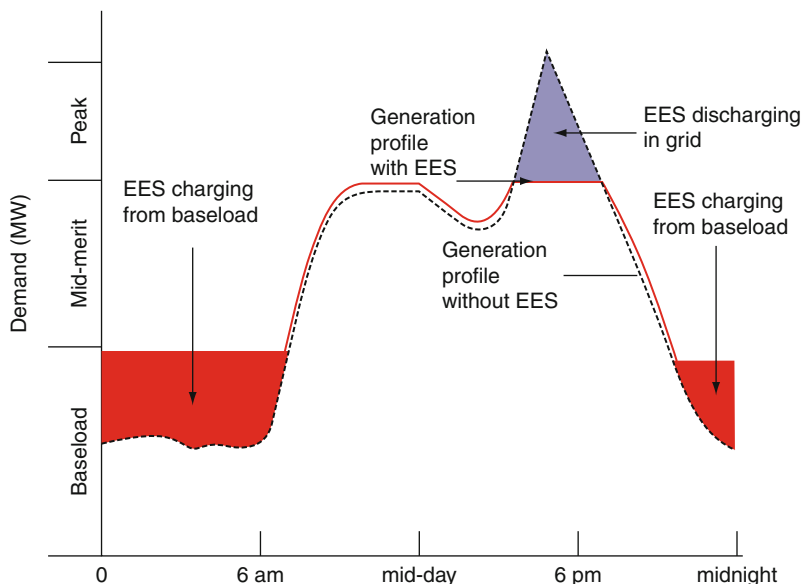
Need for Energy Storage Devices

Storage of electrical energy is one of the major research focuses of this century. Energy storage devices have already helped revolutionize the electronic gadget industry, but apart from this, energy storage devices of higher capacity and power rating can prove to be very beneficial in other stationary applications such as load-leveling in existing power plants and nonstationary applications such as electric vehicles.

The demand for electric power from an electric grid varies based on the season, day of the week, hour, and also due to transients. [Figure 1](#) shows a typical electrical power supply and demand curve over a 24-h period [53]. The power plants are always designed to meet the peak power demand. However, in a 24-h cycle, there are instances when the demand for the power supply is less than the peak power available. Thus, at times, power plants are forced to operate at lower output levels which prove to be uneconomical as the power plants are most efficient at their rated capacity. Energy storage devices can help balance this demand by shifting the load from peak to off-peak hours. During the off-peak hours the electrical energy can be stored in an energy storage device, and then this stored energy would be used during the peak hours. This load-leveling would help in designing power plants with lower rated capacity, thus reducing the cost of installation. The operating cost of a power plant can be reduced by operating the plants at the rated capacity for most of the time during its life. Another advantage is that the utility company can manage the demand with the existing power plants and defer the installation of the higher capacity power plants and thus reduce the overheads.

Nanomaterials for Electrical Energy Storage Devices, Fig. 1

A typical electrical power demand profile for a 24-h period. During early morning hours the demand is below the capacity of the power plant while the demand is at peak in the evening around 6 PM. During off-peak hours an electrical energy storage (EES) device would store the energy and during the peak hours the EES will be discharged. Thus EES would level the load and the actual demand cycle will be as shown by the solid line (Adapted from [53])



Electrical energy storage devices also play a vital role in the design and development of power plants based on other unconventional renewable energy sources, such as solar or wind energy [53]. The energy harvested in such power plants varies largely due to operating conditions such as weather, wind speed, and temperature. In this case, electrical energy storage devices can be used to store the energy during the smooth operation of the plant and overcome the intermittent nature of it. Thus, harvesting energy from unconventional sources would be more feasible.

Apart from these stationary applications energy storage devices are also necessary for certain nonstationary applications. Since about 2000, batteries have been seriously considered in the development of electric vehicles. The automotive industry has developed hybrid electric vehicles, plug-in hybrid electric vehicles, and electric vehicles, which would reduce dependence on oil and reduce greenhouse gas emissions.

Need for Nanomaterials in Energy Storage

Nanomaterials, due to their unique characteristics, are very instrumental in developing energy storage devices with high energy and power density. Energy conversion in energy storage devices takes place with a chemical reaction at the surface, charge transfer, etc. These processes occur at the nanoscale. Nanomaterials provide a large surface area to volume ratio for such processes to occur. A high surface area

permits a high contact area with the electrolyte and a larger number of reaction sites. In certain cases they enable electrode reactions to occur that cannot take place for materials composed of micrometer-sized particles.

The kinetics of the energy storage device are governed by the kinetics of the diffusion process within the device elements. Nanomaterials help in reducing the diffusion length and hence improve the kinetics of the operation of an energy storage device. Thus, nanomaterials can improve the charge-discharge rates. The electron transport within the electrode is also improved by nanometer-sized particles.

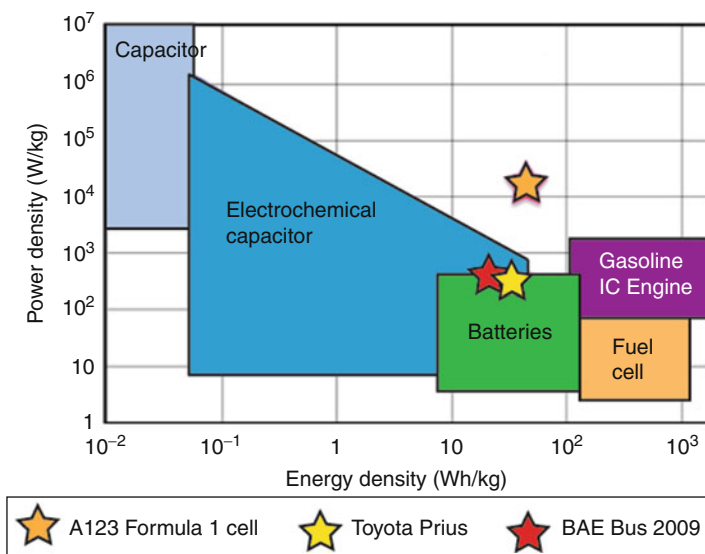
Since the charge transfer reactions occur mostly at the surface of the nanometer-sized particles, the induced stress and crystal lattice breakdown is less during the repeated charge-discharge cycles. This helps in improving the cycle life of the device.

Recent Trends in Electrical Energy Storage Devices

A Ragone plot, shown in Fig. 2, is used to compare the energy density and power density of various electrical energy storage (EES) devices [53, 54]. Energy is defined as the capacity to do work, so energy density represents the capacity of the EES, i.e., the amount of energy stored in the device. Power is defined as the rate of doing work, so power density represents the rate at which the EES is charged or discharged.

Nanomaterials for Electrical Energy Storage Devices, Fig. 2

Ragone chart showing the comparison of the power density and energy density for batteries, capacitors, and fuel cells and internal combustion engine. Energy density represents the capacity while power density represents the rate of the charging/discharging of EES device (Adapted from [54])



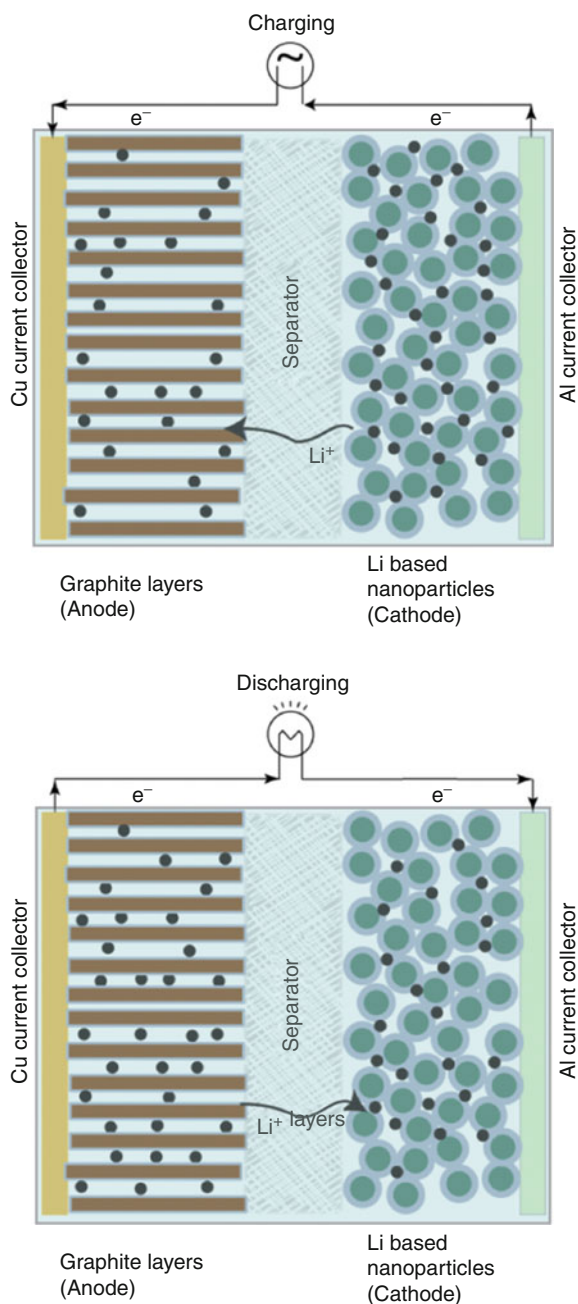
As seen in the Ragone plot (Fig. 2), a fuel cell has the highest energy density but poor power density, so fuel cells are suitable for applications where high energy storage is necessary. They also tend to operate at optimum efficiency only at a constant output. They have a certain warm-up time necessary to bring them to operating conditions, and as such they are often coupled with other EES. As opposed to fuel cells, an electrochemical capacitor has high power density but low energy density, so they are suitable for applications where rapid charge/discharge is necessary. However the output of the electrochemical capacitor depends largely on its state of charge, and thus it fails to deliver energy at constant voltage [53]. Batteries lie in between the fuel cell and the electrochemical capacitor on the Ragone plot. They have higher power density than fuel cells and do not require any warm-up time. They also have higher energy density than an electrochemical capacitor and have a flat voltage curve during discharge. Thus, batteries have been actively pursued as the EES for immediate future applications.

Li-Ion Batteries

In 1912, G. N. Lewis started developing lithium batteries. Lithium (atomic number 3, group 1, period 2) metal was used as an anode in these batteries. Lithium is an alkali metal, and being the lightest of the metals with the greatest electrochemical potential, it has the

largest energy density for its weight. However, lithium is very unstable and highly reactive, as it has only one valence electron, which makes it unsafe as the electrode material. As such, research shifted from a lithium battery to a lithium-ion battery, which is a much safer option. In 1991, Sony commercialized Li-ion batteries and started using them in its electronic products. Since then the importance and market for Li-ion batteries has been ever increasing.

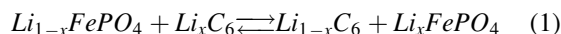
Figure 3 shows a schematic of a typical lithium-ion battery. In a typical cylindrical lithium-ion battery separator, anode, cathode and separator are stacked alternately then rolled and packed in a cylindrical can filled with electrolyte. The anode is the electropositive electrode which holds the lithium in the charged state. The cathode is the electronegative electrode which serves as the source for the lithium [3]. The electrodes are immersed in an electrolyte solution which is mostly a lithium hexafluorophosphate (LiPF_6) salt dissolved in organic solvent. A separator is used in between the anode and the cathode to avoid short-circuiting between the electrodes. The electrochemical reaction in the case of the lithium-ion battery is based on the intercalation of Li^+ ions in the electrodes. During charging the Li^+ ions are transferred from the cathode and intercalated into the anode. During discharging, a deintercalation reaction occurs, and Li^+ ions are transferred from the anode to the cathode while the electrons flow through the external circuit. The cell voltage is determined by the energy of the



Nanomaterials for Electrical Energy Storage Devices,

Fig. 3 Schematic showing operation of a Li-ion battery. In a cylindrical battery separator, anode, cathode and separator are stacked alternately then rolled and packed in a cylindrical can filled with electrolyte solvent containing Li salt. During charging Li^+ ions are inserted in the anode structure by virtue of intercalation and during discharging, these Li^+ ions are removed from the anode structure by virtue of deintercalation and transferred to the cathode

chemical reaction, such as given by equation 1, occurring in the cell [37].

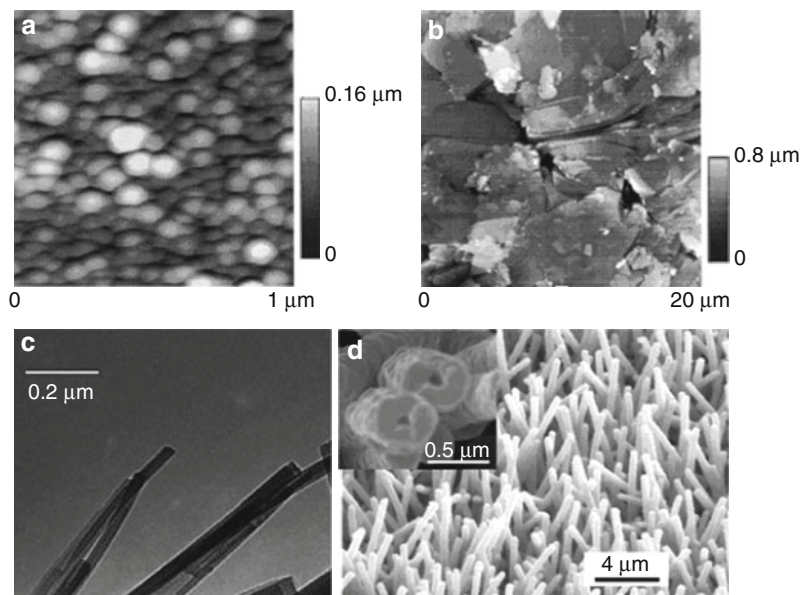


The anode and cathode electrodes used in Li-ion batteries are complex composites with active material bonded to the current collector with polymeric binders. The anode material is bonded to the copper current collector and the cathode material is bonded onto an aluminum current collector. The cathode material is also coated with carbon black to improve the electronic conductivity.

Anode Lithium is an ideal anode material for Li-ion batteries, but because of the plating out of lithium during the charge/discharge cycle and subsequent formation of dendrites, there is a risk of the cell being short-circuited. Instead, carbonaceous materials are preferred for anodes in most commercially available Li-ion batteries [22]. Studies have been conducted on graphite [6], C-C composite, mesocarbon microbeads (MCMB) [18], carbon nanotubes [13, 52], and carbon films. Evaporated carbon film and a natural graphite anode is shown in Fig. 4a, b respectively [22]. These anodes have a very good rate of lithium insertion/removal and thus improve the charge/discharge rate (power rating) of the battery, but in any Li-ion cell with a carbonaceous anode a solid electrolyte interphase (SEI) is formed during the cycling of the cell [21, 22, 43]. The SEI is formed from the electrolyte decomposition products [50]. This SEI layer prevents the graphite surface from further exfoliation and also prevents further reduction of the electrolyte and consumption of active lithium. In the case of nanoparticulate graphite, the consumption of active lithium would be still higher, and the excessive charge developed between the graphite surface and the SEI would result in a loss of overall cell voltage. Also, it is important to note that the lithium is intercalated into graphite at potentials less than 100 mV versus Li^+/Li . There is always a risk of lithium depositing on the graphite surface resulting in dendrite formation and fatal short-circuiting of the cell [6]. Carbon nanotubes exhibit twice the lithium storage compared with graphite but have similar problems of surface-layer formation and safety [6]. To overcome some disadvantages of graphite, the research has been focused on the titanium oxide.

Nanomaterials for Electrical Energy Storage Devices, Fig. 4

Different types of anode materials used in the Li-ion battery. (a) AFM image of a carbon film on glass prepared by electron-beam evaporation technique [22]; (b) AFM image of a natural graphite powder pressed together in a hydraulic press [22]; (c) TEM image of TiO_2 -B nanowires [6]; and (d) SEM image of mesoporous Co_3O_4 nanowire arrays grown on Ti foil. The inset shows the open tips of the nanowire [28]



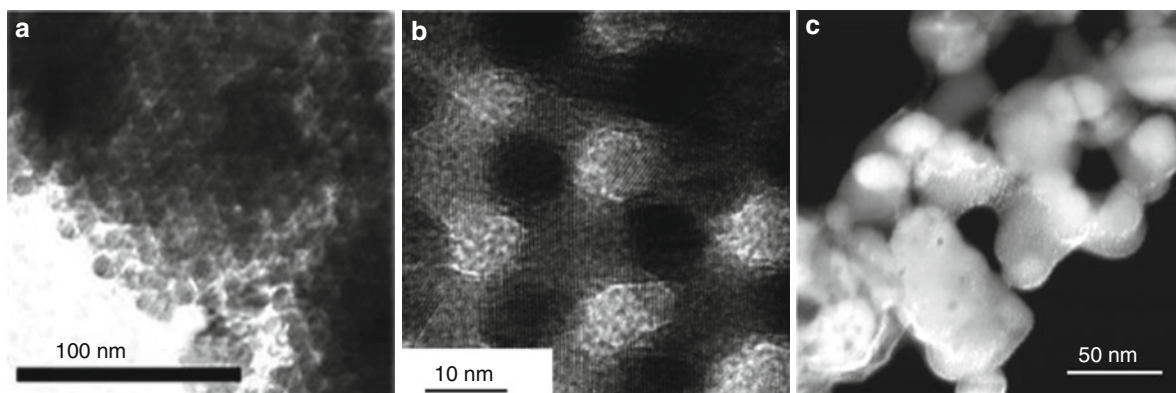
The nanoparticles of non-toxic $\text{Li}_4\text{Ti}_5\text{O}_{12}$ ($\text{Li}[\text{Li}_{1/3}\text{Ti}_{5/3}]\text{O}_4$) with spinel structure have high diffusion rates due to short diffusion distance. In this spinel structured oxide, lithium can be intercalated to give a composition range of $\text{Li}_{4+x}\text{Ti}_5\text{O}_{12}$, $0 < x < 3$ [14, 20], but the capacity of the $\text{Li}_4\text{Ti}_5\text{O}_{12}$ is only 150 mAh/g compared to 300 mAh/g capacity of graphite. Instead nanotubes/nanowires composed of TiO_2 -B, as shown in Fig. 4c, are actively studied [6]. This fifth polymorph of titanium dioxide has all the advantages of $\text{Li}_4\text{Ti}_5\text{O}_{12}$ but also has a capacity of 300 mAh/g. TiO_2 -B is also safer as it eliminates the risk of lithium plating on the electrode and can be cycled with higher rates as in $\text{Li}_4\text{Ti}_5\text{O}_{12}$. The TiO_2 -B nanowires exhibit better performance than TiO_2 -B nanoparticles as they have higher reversibility of intercalation (>99.9% per cycle, after the first cycle). Other similar oxide nanowire electrode materials of interest include Sn, Co, and V oxides [6]. Lithium metal alloys such as $\text{Li}_{4.4}\text{Sn}$ (993 mAh/g) and $\text{Li}_{4.4}\text{Si}$ (4,200 mAh/g) with high specific capacity are also being considered as anode materials in Li-ion batteries. The major drawback of these materials is the large volume change during the cycling of the cell to accommodate the large amounts of lithium [6].

Lithium's ability to react reversibly with the binary transition metal oxides is exploited in developing anode materials such as mesoporous Co_3O_4 , shown in Fig. 4d [28]. This reaction is not limited by the one

or two lithium atoms per host as in intercalation, and so such materials can have a higher capacity than conventional anode materials. The poor kinetics of these types of materials can be improved by using nanowires or nanoparticulates as the anode material.

Cathode The potential area of greatest opportunity for nanomaterials is the cathode of the Li-ion battery [53]. The most common cathode materials are layered oxides LiMO_2 , the spinels $\text{Li}[\text{M}_2]\text{O}_4$ and olivines LiMPO_4 where M is a transition metal atom [16]. The nanoparticulates of these materials reduce the diffusion length, provide a higher electrolyte/electrode contact area, and thus improve the charge/discharge kinetics of the Li-ion batteries. The nanoparticles of these lithium compounds are made by grinding, synthesis from solution, or by sol-gel approaches [6].

Initially a LiCoO_2 , as shown in Fig. 5a, was developed as the cathode material for Li-ion batteries [6, 16], but due to expensive and toxic cobalt they were later replaced by oxides such as $\text{LiNi}_{0.8}\text{Co}_{0.15}\text{Al}_{0.05}\text{O}_2$, $\text{LiNi}_{0.8}\text{Co}_{0.2}\text{O}_2$, $\text{Li}_{1-x}\text{Ni}_{1-y}\text{Co}_y\text{O}_2$, and $\text{LiMn}_{0.5}\text{Ni}_{0.5}\text{O}_2$ in commercial batteries [16, 23, 24]. These oxides have a high operating voltage, in the range of 2.75–4.3 V [16]. They have very high energy density and power density but they lack the necessary structural stability for deep discharge cycles, during which the host oxide structure collapses upon removal of more than 50% of the Li. Li spinel compounds, such as LiMn_2O_4 shown in Fig. 5b, has good structural stability [6]. They also have



Nanomaterials for Electrical Energy Storage Devices, Fig. 5 Different types of cathode materials used in the Li-ion battery. (a) TEM image of mesoporous LiCoO₂ transition metal

oxide [6], (b) TEM image of nanostructured LiMn₂O₄ spinel [6], and (c) TEM image of olivine structured LiFePO₄ nanoparticles embedded in carbon matrix [38]

higher operating voltage and high charge/discharge rate but low energy density. The most recent among these lithium compounds has been the olivine structured LiFePO₄ shown in Fig. 5c [38]. It has a lower operating voltage of ~3.3 V but demonstrates higher power and energy density along with a good structural stability.

Electrochemical Capacitors

Capacitors are commonly used in power and consumer electronic circuits. However they have very low capacitance in the range of few millifarads which make them unsuitable for large storage devices. In 1957, Becker filed the first patent describing the concept of electrochemical capacitors (EC) which have capacitance in the range of a few hundred farads [4, 46]. It was then commercialized by Nippon Electric Company (NEC), Japan, under a license from Standard Oil of Ohio (SOHIO) [46]. EC are also called supercapacitors or ultracapacitors. The two different types of EC are the electric double layer capacitor (EDLC) and the pseudocapacitor.

A schematic showing the operation of an EDLC is shown in Fig. 6. EDLC consists of two electrodes with a thin coating of active material having very high surface area. These electrodes are immersed in an electrolyte, mostly sulfuric acid or acetonitrile, containing both positive and negative ions. When the capacitor is connected to the external power source, the electric field attracts the opposite charged ions toward the surface of the electrode. The charge separation occurs near the surface of the electrode, and energy is stored in the EDLC. During discharging, as the

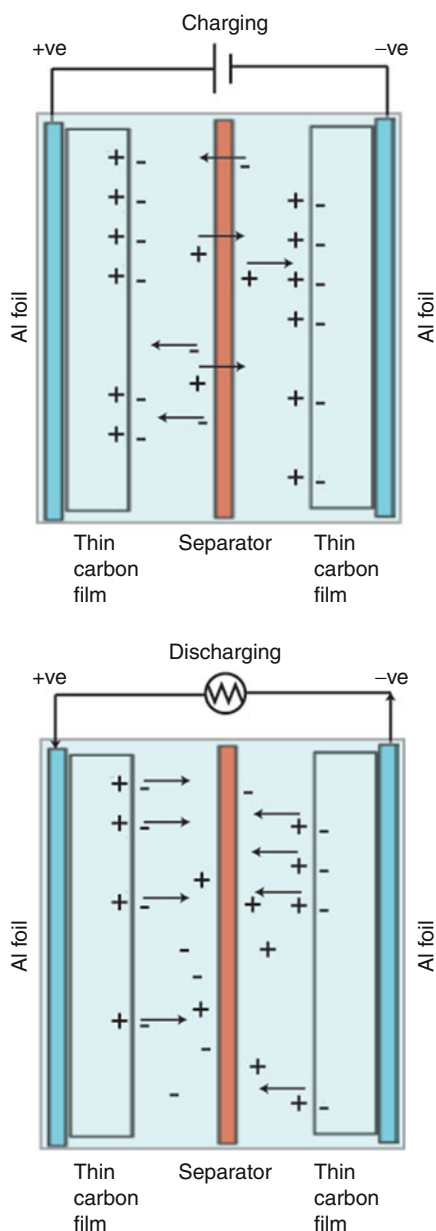
electrons flow through the load the ions disperse in the electrolyte solution and the electrode-electrolyte interface fades [25, 35]. A separator is used to prevent the tunneling of the electrons between the electrodes through the electrolyte. The other type of EC capacitor, the pseudocapacitor, has essentially the properties of the EDLC as well as a battery. Thus a charge is stored by the surface as well as subsurface activities. In EDLC there is no physical or chemical change in the electrode, and the ion adsorption process is highly reversible. Thus, theoretically EDLC can be charged/discharged without any damage to its components [34].

The charge separation at the electrode-electrolyte interface is given by Helmholtz's double layer capacitance C :

$$C = \frac{\epsilon_r \epsilon_o A}{d} \quad (2)$$

where ϵ_r is the electrolyte dielectric constant, ϵ_o is the dielectric constant of the vacuum, d is the effective thickness of the double layer (charge separation distance), and A is the electrode surface area [46].

Many materials used as an anode in Li-ion batteries are generally considered suitable for EC. In general, to achieve high capacitance between the charged double layer it is necessary to have a high specific surface area (SSA) of the double layer material and electrodes with high electronic conductivity. Carbonaceous material has high SSA, high electronic conductivity, electrochemical stability, and structural stability and is suitable for the EC. The various carbon-based materials



Nanomaterials for Electrical Energy Storage Devices,

Fig. 6 Schematic showing operation of an electrochemical capacitor (EC). During charging the ions in the electrolyte are attracted to the nanoporous electrodes with opposite charge. This creates the charge separation at the electrode-electrolyte interface and the energy is stored in EC. During discharging, as the electrons flow through the load the ions disperse in the electrolyte solution and the electrode-electrolyte interface fades

such as carbon fabrics, carbide-derived carbons, carbon nanotubes, onion-like carbon, and nanohorns are studied as potential EC electrode materials. Disordered microporous carbon (SiC-derived carbon),

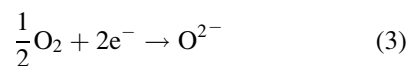
shown in Fig. 7a, and carbon nanotubes, shown in Fig. 7c, are among the various carbonaceous ideal EC electrode materials due to their high SSA [46].

Several other materials, such as V_2O_5 nanowires shown in Fig. 7b, and ruthenium oxides are actively being studied as EC electrode materials due to their broad range of morphologies [12, 49]. Hydrated ruthenium oxide has the highest capacity, 200 Wh/kg [53]. The other promising EC material is nickel oxide due to its high specific capacitance value compared to carbon materials [1, 10, 27, 33]. Thin films of nickel oxide can be easily synthesized through electrochemical route or sol-gel method. The nanostructured NiO thin films shown in Fig. 7d can be produced with specific microstructure, high surface area, and reduced crystallite size [39, 51].

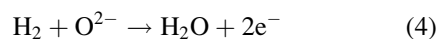
Fuel Cell

A fuel cell converts the energy stored in the source fuel to usable electrical energy. It is different from other EES in that the reactants flow through the cell rather than being sealed within the cell. There are several types of fuel cells, but the most common are the solid oxide fuel cell (SOFC) and the proton exchange membrane fuel cell (PEMFC) [7]. A schematic of SOFC is shown in Fig. 8a and schematic of PEMFC is shown in Fig. 8b. SOFC is the most fuel flexible among all fuel cells. Other than pure hydrogen, it can run on hydrocarbon, methanol, or even gasoline, but in the case of PEMFC pure hydrogen is the only choice of fuel. Essentially both fuel cells have similar principles of operation except that in SOFC, O^{2-} ions are conducted through the solid oxide membrane (SOM) while in PEMFC H^+ ions are conducted through the proton exchange membrane (PEM) [17, 19].

In SOFC, the oxygen is oxidized to O^{2-} as per the following half-cell reaction at the cathode:

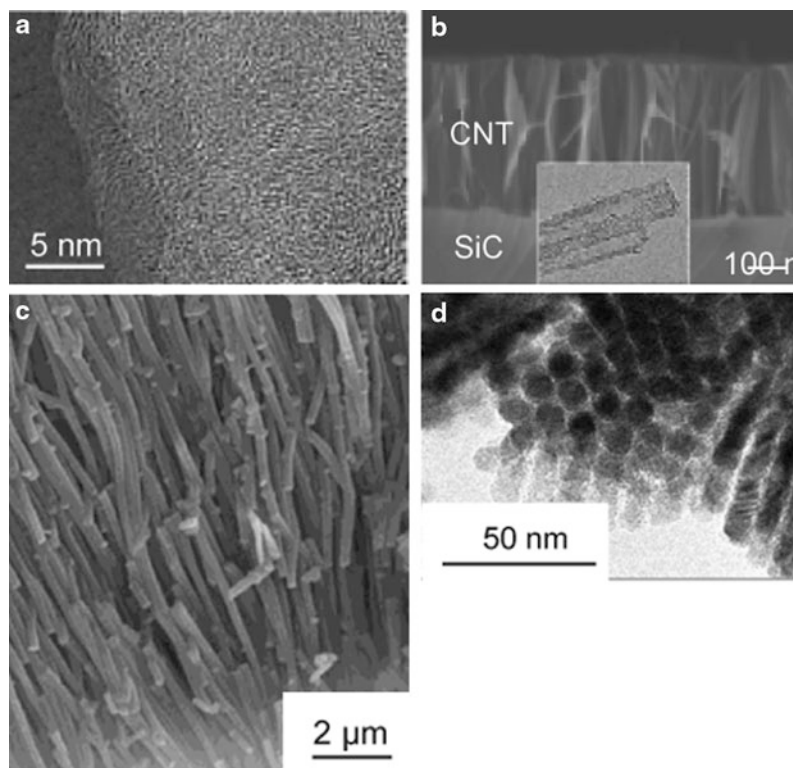


These O^{2-} ions are conducted through the solid oxide membrane to the anode where they are reduced by hydrogen with water as a by-product. The anode half-cell reaction is:

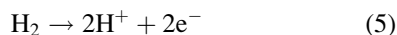


Nanomaterials for Electrical Energy Storage Devices, Fig. 7

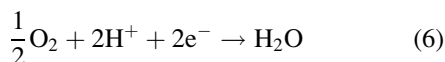
Different types of electrode materials used in EC. (a) TEM image of a disordered microporous carbon (SiC-derived carbon) [46], (b) SEM image of an array of carbon nanotubes (CNT) on Si-C (inset shows the TEM image of the same CNT) [46], (c) TEM image of V_2O_5 nanowires [49], and (d) TEM image of ordered mesoporous NiO [51]



In the case of PEMFC, hydrogen is reduced at the anode, and the H^+ ions are formed as per the following half-cell reaction:



These H^+ ions are conducted through the proton exchange membrane, and they are combined with oxygen at the cathode to form water as the by-product of the reactions. The half-cell reaction at the cathode is:

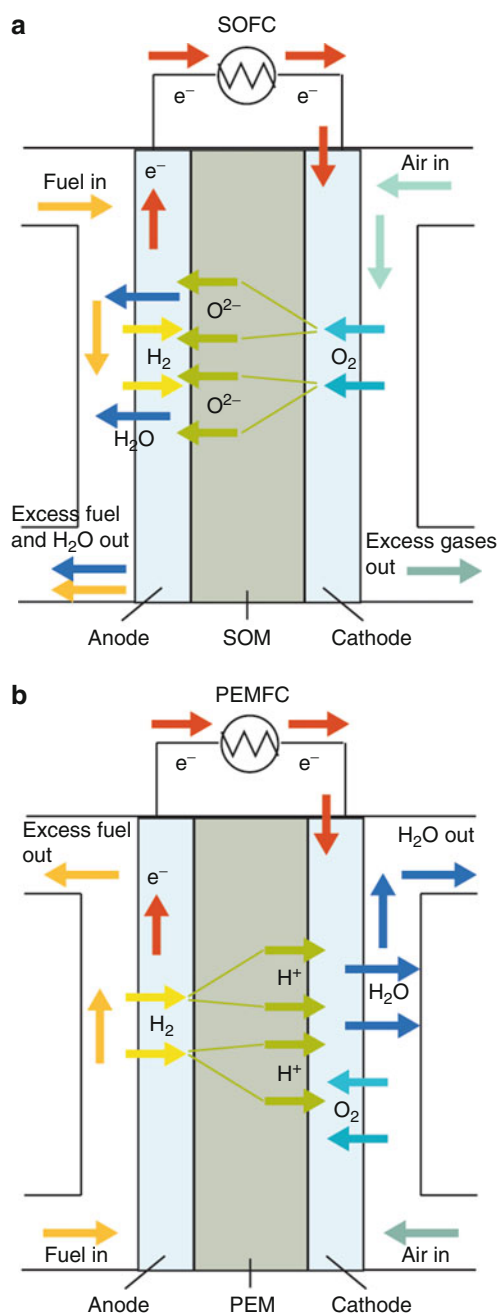


Due to the high operating temperatures, in the range of 500–1,000°C, SOFC is more cost effective as it does not need an expensive catalyst such as Pt, used in PEMFC, but the disadvantage of high operating temperatures is that the components of SOFC should demonstrate higher thermal and thermomechanical stability. In the case of the PEMFC the operating temperature is lower (70–100°C), but the proton exchange membrane needs to be hydrated to optimum

levels to maintain the proton conductivity [17]. Too much wetting of the membrane causes flooding of the cell. If the membrane is dry then it can also lead to cracking of the membrane and overheating of the cell [17].

A good electrolyte membrane for a fuel cell should have high ionic conductivity, low electronic conductivity, should be stable in both oxidizing and reducing environments, good mechanical properties, and good thermal stability. The electrodes should have high electronic conductivity and high (electro)catalytic activity. The anode catalyzes the oxidation of the fuel components by the ions transported through the electrolyte membrane and the cathode catalyzes the reduction of oxygen [19].

SOFC An extensive review of SOFC has been presented by Minh [36], Singhal [47], and Ormerod [41]. The most widely accepted electrolyte for SOFC is yttria-stabilized zirconia (YSZ) [39]. Figure 9a shows a complex structure of the composite cathode supported by a 240 μm thick dense YSZ electrolyte. The cathode fabricated by combustion chemical vapor deposition (CVD) consists of three porous



Nanomaterials for Electrical Energy Storage Devices,

Fig. 8 (a) Schematic showing operation of a solid oxide fuel cell (SOFC). Oxygen is reduced at cathode and the resulting O^{2-} ions are conducted through the solid oxide electrolyte membrane (SOM) to the anode where they are oxidized with the hydrogen to form water. (b) Schematic showing operation of a proton exchange membrane fuel cell (PEMFC). Hydrogen is oxidized at the anode and the resulting H^+ ions are conducted through the proton exchange membrane (PEM) to the cathode where they are reduced with O to form water

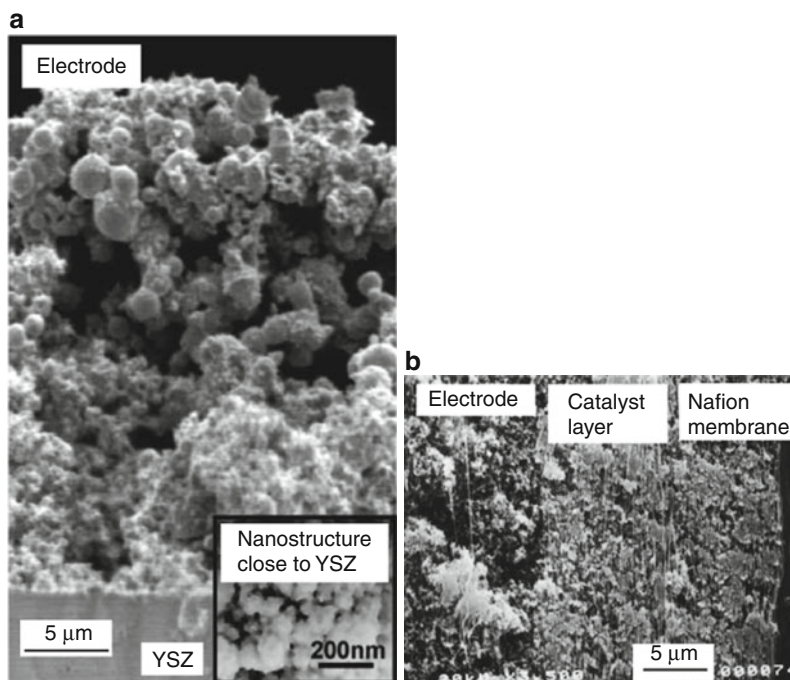
layer structures, with about 5 μm thick 60 wt.% lanthanum strontium manganite (LSM)-40 wt.% gadolinia doped ceria (GDC) fine agglomerates (0.5 μm diameter) at the bottom (close to YSZ electrolyte), followed by 5 μm thick 30 wt.% LSM-30 wt.% lanthanum strontium cobaltite (LSC)-40 wt.% GDC fine agglomerates (0.5 μm diameter), and 15 μm thick 60 wt.% LSC- 40 wt.% GDC coarse agglomerates (2–3 μm diameter) on the top (air side) [30]. The inset shows the nanostructure present in the bottom two layers. This nanostructure offers extremely high surface area for oxygen reduction. The anode is mostly a ceramic metal composite (cermet) comprised of Ni + YSZ [17]. The other electrolyte systems investigated for SOFC are strontium, magnesium-doped lanthanum gallate (LSGM), and gadolinium- or samarium-doped ceria (CGO or CSO).

PEMFC Operation of PEMFC and its materials has been covered extensively in Costamagna and Srinivas [8, 9]. In the case of PEMFC the most commonly used electrolyte is Nafion[®] (sulfonated polymer), manufactured by DuPont. Nafion is comprised of perfluorinated backbones, which provide chemical stability, and of sulfonated side-groups which aggregate and facilitate hydration, which is necessary for proton conductivity. Figure 9b shows a typical membrane-electrode assembly of PEMFC [5]. In this assembly, the Nafion[®] membrane is hot-pressed onto the electrode impregnated with Nafion[®] 5% solution by brushing technique. Other sulfonated polymers have been discussed by Kreuer [26]. For hydrogen/air fuel cells, Pt nanoparticles supported on carbon are utilized for both the anode and cathode.

Hydrogen storage Hydrogen is the optimal fuel for all types of fuel cells. Storage of hydrogen on-board a transportation vehicle continues to be one of the most challenging topics for the commercialization of hydrogen-fueled vehicles. Hydrogen has very high energy density by mass but very poor volumetric energy density. The higher volumetric energy density can be achieved for smaller but not lighter containers by compressing pure hydrogen to higher pressures, but compressing the pure hydrogen requires a high-powered compressor. Similarly, higher volumetric density can be achieved by liquid hydrogen, but it requires cryogenic storage systems as liquid hydrogen boils at 20.268°K, increasing the cost of insulation. To overcome these difficulties, storage of hydrogen by adsorption (physisorption) on materials with high

Nanomaterials for Electrical Energy Storage Devices, Fig. 9

Different types of electrode-electrolyte materials used in SOFC and PEMFC. (a) SEM image of strontium doped lanthanum manganite (LSM)–strontium doped lanthanum cobaltite (LSC)–gadolinia doped ceria (GDC) composite cathode electrode fabricated on yttria-stabilized zirconia (YSZ) electrolyte membrane used in SOFC. The inset shows the high resolution imaged of two nanostructured bottom layers close to the YSZ (adapted from [30]). (b) PEM membrane-electrode assembly used in PEMFC. SEM image of the cross-section of the Nafion[®] 117 membrane and electrode impregnated with the Nafion[®] 5% solution [5]

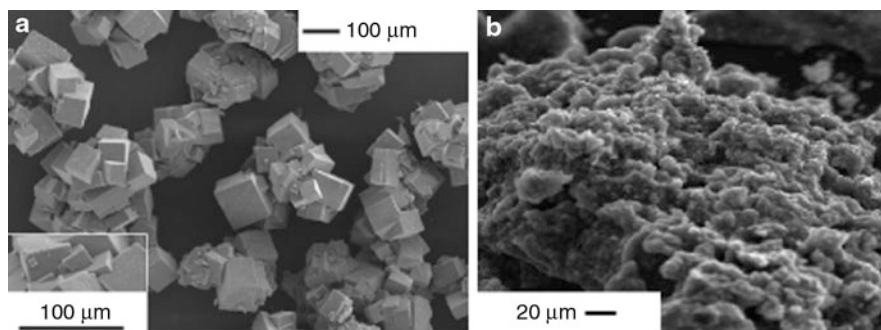


specific surface area or chemical bonding (chemisorption) in covalent and ionic compounds is becoming more attractive [55]. Nanostructured materials improve the thermodynamics and kinetics of hydrogen adsorption and dissociation by decreasing the required diffusion length and increasing the diffusion rate [40]. Carbonaceous nanomaterials such as activated carbon, graphene nanofibers, carbon nanotubes, or a group of synthetic materials called as metal oxide frameworks (MOF) are considered for storage by physisorption [29]. The hydrogen is adsorbed onto the surface by weak van der Waals interaction between the hydrogen molecules and the surface of the adsorbent material [15, 39]. MOF-5, shown in Fig. 10a, with a formula unit of $Zn_4O(C_{24}H_{12})_{12}$ or $Zn_4O(\text{benzene decarboxylate})_3$, is one such extensively studied MOF material [42]. A hydrogen storage capacity of 4 wt.% at 78 K (-195°C) and 1 wt.% at room temperature has been reported for different MOFs [40]. However, due to the highly porous nature of these materials, volumetric capacity may still be a significant issue [40]. The other frequently studied physisorption-based materials include zeolites [11], clathrate hydrates [45, 48], and hollow glass microspheres (HGM) [2]. Usually hydrogen storage by physisorption-based materials, such as MOF, HGM, CNT, etc., has poor capacity at normal

operating conditions [29]. Figure 10b shows another class of materials – a complex hydride, borohydride ($\text{LiBH}_4 + 1/2 \text{ZnCl}_2$) doped with nano-Ni which stores hydrogen by chemisorptions – used in storage of hydrogen [40]. These hydrides along with other chemisorption materials, e.g., magnesium-based hydride, metal nitrides, and amides imides, are very expensive, and in some cases the hydrogen absorption/desorption process is irreversible [29, 31, 32, 44]. The huge capacity values of various materials published in respected technical and scientific reports should be accepted with caution as many values were wrongfully reported due to gross experimental measurement errors [29].

Challenges

There is a profound effect of spatial boundaries and contribution of surfaces due to small particle size on many of the properties of materials; this challenges researchers to develop new theories or at least adapt and develop theories that have been established for bulk materials [46]. Understanding how materials store and transport charges at electrode-electrolyte interfaces is critically important and will require a fundamental understanding of charge transfer and transport mechanisms. High electrolyte/electrode surface areas may lead to more significant side reactions



Nanomaterials for Electrical Energy Storage Devices, Fig. 10 Different types of materials used to store hydrogen. (a) SEM image of metal-organic framework – 5 (MOF-5) which stores hydrogen by physisorption on its surface [42], and (b)

SEM image of the complex borohydride ($\text{LiBH}_4 + 1/2 \text{ZnCl}_2$) doped with nano-Ni which stores hydrogen by chemisorptions [40]

with the electrolyte, and more difficulty maintaining interparticle contact. There is also a challenge to synthesize nanoparticles having controlled dimensions with a cost effective process. The characterization of functional and structural properties of these nanomaterials also challenges the scientific community in modifying the existing tools and/or developing the new tools.

Cross-References

► Aging

References

- Ahlberg, E., Palmquist, U., Simic, N., Sjoval, R.: Capacity loss in Ni–Cd pocket plate batteries. The origin of the second voltage plateau. *J. Power Sources* **85**(2), 245–253 (2000)
- Akunets, A.A., Basov, N.G., Bushuev, V.S., Dorogotvtsev, V.M., Gromov, A.I., Isakov, A.I., Kovylnikov, V.N., Merkul'ev, Yu.A., Nikitenko, A.I., Tolokonnikov, S.M.: Super-high-strength microballoons for hydrogen storage. *Int. J. Hydrogen Energ.* **19**, 697–700 (1994)
- Anonymous: Basic Research Needs for Electrical Energy Storage. Office of Basic Energy Sciences, U.S. Department of Energy, Washington, DC (2007)
- Becker, H. E.: Low voltage electrolytic capacitor. U.S. Patent 2,800,616 (1957)
- Broka, K., Ekdunge, P.: Oxygen and hydrogen permeation properties and water uptake of Nafion[®] 117 membrane and recast film for PEM fuel cell. *J. Appl. Electrochem.* **27**, 117–123 (1997)
- Bruce, P.G., Scrosati, B., Tarascon, J.-M.: Nanomaterials for rechargeable lithium batteries. *Angew. Chem. Int.* **47**, 2930–2946 (2008)
- Carrette, L., Friedrich, K.A., Stimming, U.: Fuel cells: principles, types, fuels and applications. *Chem. Phys. Chem.* **1**, 162–193 (2000)
- Costamagna, P., Srinivasan, S.: Quantum jumps in the PEMFC science and technology from the 1960s to the year 2000. Part I. Fundamental scientific aspects. *J. Power Sources* **102**, 242–252 (2001)
- Costamagna, P., Srinivasan, S.: Quantum jumps in the PEMFC science and technology from the 1960s to the year 2000. Part II. Engineering, technology development and application aspects. *J. Power Sources* **102**, 253–269 (2001)
- Deabate, S., Fourgeot, F., Henn, F.: Electrochemical behaviour of the $\beta(\text{II})\text{-Ni}(\text{OH})_2/\beta(\text{III})\text{-NiOOH}$ redox couple upon potentiodynamic cycling conditions. *Electrochim. Acta.* **51**(25), 5430–5437 (2006)
- Dong, J., Wang, X., Xu, H., Zhao, Q., Li, J.: Hydrogen storage in several microporous zeolites. *Int. J. Hydrogen Energ.* **32**, 4998–5004 (2007)
- Drezen, T., Kwon, N.-H., Bowenb, P., Teerlinck, I., Isono, M., Exnar, I.: Effect of particle size on LiMnPO_4 cathodes. *J. Power Sources* **174**, 949–953 (2007)
- Fan, Q., Chupas, P.J., Whittingham, M.S.: Characterization of amorphous and crystalline tin-cobalt anodes. *Electrochem. Solid-State Lett.* **10**, A274–A278 (2007)
- Ferg, E., Gummow, R.J., de Kock, A., Thackeray, M.M.: Spinel anodes for lithium-ion batteries. *J. Electrochem. Soc.* **141**, L147–L150 (1994)
- Felderhoff, M., Weidenthaler, C., Helmolt, R., von Eberle, U.: Hydrogen storage: the remaining scientific and technological challenges. *Phys. Chem. Chem. Phys.* **9**, 2643–2653 (2007)
- Goodenough, J.B.: Cathode materials: a personal perspective. *J. Power Sources* **174**, 996–1000 (2007)
- Haile, S.M.: Fuel cell materials and components. *Acta Mater.* **51**, 5981–6000 (2003)
- Hossain, S., Kim, Y.-K., Saleh, Y., Loutfy, R.: Comparative studies of MCMB and C-C composite as anodes for lithium-ion battery systems. *J. Power Sources* **114**, 264–276 (2003)

19. Jacobson, A.J.: Materials for solid oxide fuel cells. *Chem. Mater.* **22**, 660–674 (2010)
20. Jansen, A.N., Kahalan, A.J., Kepler, K.D., Nelson, P.A., Amine, K., Dees, D.W., Vissers, D.R., Thackeray, M.M.: Development of a high-power lithium-ion battery. *J. Power Sources* **81**, 902–905 (1999)
21. Kanno, R., Kawamoto, Y., Takeda, Y., Ohashi, S., Imanishi, N., Yamamoto, O.: Carbon fiber as a negative electrode in lithium secondary cells. *J. Electrochem. Soc.* **139**, 3397–3404 (1992)
22. Kong, F., Kostecki, R., Nadeau, G., Song, X., Zaghib, K., Kinoshita, K., McLarnon, F.: In-situ studies of SEI formation. *J. Power Sources* **97–98**, 58–66 (2001)
23. Kostecki, R., McLarnon, F.: Degradation of $\text{LiNi}_{0.8}\text{Co}_{0.2}\text{O}_2$ cathode surfaces in highpower lithium-ion batteries. *Electrochem. Solid State Lett.* **5**, A164–A166 (2002)
24. Kostecki, R., McLarnon, F.: Local probe studies of degradation of composite $\text{LiNi}_{0.8}\text{Co}_{0.15}\text{Al}_{0.05}\text{O}_2$ cathodes in high-power lithium-ion batteries. *Electrochem. Solid State Lett.* **7**, A380–A383 (2004)
25. Kötz, R., Carlen, M.: Principles and application of electrochemical capacitors. *Electrochim. Acta.* **45**, 2483–2498 (2000)
26. Kreuer, K.D.: On the development of proton conducting polymer membranes for hydrogen and methanol fuel cells. *J Membr. Sci.* **185**, 29–39 (2001)
27. Léger, C., Tessier, C., Ménétrier, M., Denage, C., Delmas, C.: Investigation of the second discharge plateau of the $\beta(\text{III})\text{-NiOOH}/\beta(\text{II})\text{-Ni(OH)}_2$ system. *J. Electrochem. Soc.* **146**, 924–932 (1999)
28. Li, Y., Tan, B., Wu, Y.: Mesoporous Co_3O_4 nanowire arrays for lithium ion batteries with high capacity and rate capability. *Nano Lett.* **8**, 265–270 (2008)
29. Lim, K.L., Kazemian, H., Yaakob, Z., Daud, W.R.W.: Solid-state materials and methods for hydrogen storage: a critical review. *Chem. Eng. Technol.* **33**, 213–226 (2010)
30. Liu, Y., Compson, C., Liu, M.: Nanostructured and functionally graded cathodes for intermediate temperature solid oxide fuel cells. *J. Power Sources* **138**, 194–198 (2004)
31. Liu, Y., Xiong, Z., Hu, J., Wu, G., Chen, P., Murata, K., Sakata, K.: Hydrogen absorption/desorption behaviors over a quaternary Mg–Ca–Li–N–H system. *J. Power Sources* **159**, 135–138 (2006)
32. Luo, K., Liu, Y., Wang, F., Gao, M.I., Pan, H.: Hydrogen storage in a Li–Al–N ternary system. *Int. J. Hydrogen Energ.* **34**, 8101–8107 (2009)
33. Mancier, V., Willmann, P., Metrot, A.: A semi theoretical approach of the second plateau appearing during the discharge of aged nickel oxyhydroxide electrodes. *J. Power Sources* **85**, 181–185 (2000)
34. Miller, J.R., Burke, A.F.: Electrochemical capacitors: challenges and opportunities for real-world applications. *Electrochem. Soc. Interf.* **17**, 53–57 (2008)
35. Miller, J.R., Simon, P.: Electrochemical capacitors for energy management. *Science* **321**, 651–652 (2008)
36. Minh, N.Q.: Ceramic fuel cells. *J Am. Cer. Soc.* **78**, 563–588 (1993)
37. Nagpure, S.C., Bhushan, B.: In: Bhushan, B., Fuchs, H. (eds.) *Applied Scanning Probe Methods – Biomimetics and Industrial Applications*, vol. 13, pp. 203–233. Springer, Heidelberg (2009)
38. Nagpure, S. C., Babu, S. S., Bhushan, B., Kumar, A., Mishra, R., Windl, W., Kovarik, L., Mills, M.: Local electronic structure of LiFePO_4 nanoparticles in aged Li-ion batteries (2011, submitted)
39. Nalwa, H.S.: *Nanomaterials for Energy Storage Application*, 1st edn. American Scientific Publishers, Stevenson Ranch (2009)
40. Niemann, M.U., Srinivasan, S.S., Phani, A.R., Kumar, A., Goswami, D.Y., Stefanakos, E.K.: Nanomaterials for hydrogen storage applications: a review. *J. Nanomater.* **2008**, 1–9 (2008)
41. Ormerod, R.M.: Solid oxide fuel cells. *Chem. Soc. Rev.* **32**, 17–28 (2003)
42. Panella, B., Hirscher, M., Pütter, H., Müller, U.: Hydrogen adsorption in metal–organic frameworks: Cu-MOFs and Zn-MOFs compared. *Adv. Funct. Mater.* **16**, 520–524 (2006)
43. Peled, E.: The electrochemical behavior of alkali and alkaline earth metals in nonaqueous battery systems-The solid electrolyte interphase model. *J. Electrochem. Soc.* **126**, 2047–2051 (1979)
44. Reule, H., Hirscher, M., Weißhardt, A., Kronmüller, H.: Hydrogen desorption properties of mechanically alloyed MgH_2 composite materials. *J. Alloys Compd.* **305**, 246–252 (2005)
45. Rovetto, L.J., Strobel, T.A., Hester, K.C., Dec, S.F., Koh, C.A., Miller, K.T., Sloan, E.D.: *Molecular Hydrogen Storage in Novel Binary Clathrate Hydrates at Near-Ambient Temperatures and Pressures*, pp. 629–632. Basic Energy Sciences, U.S. Department of Energy, Washington, DC (2006)
46. Simon, P., Gogotsi, Y.: Materials for electrochemical capacitors. *Nat. Mater.* **7**, 845–854 (2008)
47. Singhal, S.C.: Solid oxide fuel cells for stationary, mobile, and military applications. *Solid State Ionics* **152–153**, 405–410 (2002)
48. Sloan, E.D.: *Clathrate Hydrates of Natural Gases*, 2nd edn. CRC Press, Boca Raton (1998)
49. Takahashi, K., Limmer, S.J., Wang, Y., Cao, G.: Synthesis and electrochemical properties of single-crystal V_2O_5 nanorod arrays by template-based electrodeposition. *J. Phys. Chem. B* **108**, 9795–9800 (2004)
50. Vetter, J., Novák, P., Wagner, M.R., Veit, C., Möller, K.–C., Besenhard, J.O., Winter, M., Wohlfahrt-Mehrens, M., Vogler, C., Hammouche, A.: Ageing mechanisms in lithium-ion batteries. *J. Power Sources* **147**, 269–281 (2005)
51. Wang, Y., Xia, Y.: Electrochemical capacitance characterization of NiO with ordered mesoporous structure synthesized by template SBA-15. *J. Electrochim. Acta.* **51**, 3223–3227 (2006)
52. Whittingham, M.S.: Electrical energy storage and intercalation chemistry. *Science* **192**, 1126–1127 (1976)
53. Whittingham, M.S.: Materials challenges facing electrical energy storage. *MRS Bulletin* **33**, 411–419 (2008)
54. Whittingham, M.S.: Materials challenges facing electrochemical energy storage: batteries and capacitors. In: Ginley, D.S., Cohen, D. (eds.) *Fundamentals of Materials for Energy and Environmental Sustainability*. Materials Research Society, Warrendale (2011) (in press)
55. Züttel, A.: Materials for hydrogen storage. *Mater. Today* **6**, 24–33 (2003)

Nanomaterials for Excitonic Solar Cells

Mónica Lira-Cantú and Irene González-Valls
Laboratory of Nanostructured Materials for
Photovoltaic Energy, Escola Técnica Superior
d'Enginyeria (ETSE), Centre d'Investigació en
Nanociència i Nanotecnologia (CIN2, CSIC),
Bellaterra (Barcelona), Spain

Synonyms

Excitonic solar cell; Dye sensitized solar cells; Hybrid solar cells; Nanostructured solar cells; Organic solar cells; Quantum dot solar cells

Definition

An excitonic solar cell (XSC) is characterized by the formation of an *exciton* that is produced from the absorbed photons from sunlight. An exciton is a tightly bound electron-hole pair which must be split for charge generation. The exciton dissociation takes place at the interface between the constituent semiconductors. Nanostructuring of interfaces is a crucial approach toward highly efficient devices.

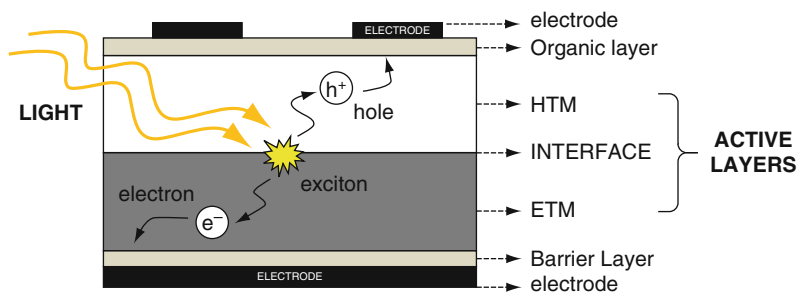
Overview

Among the many energy resources, the conversion of sunlight into electricity has emerged as one of the most promising alternatives. The evolution of photovoltaic systems initiated several years ago with the silicon-based solar cells (SiSCs). The need to lower energy cost and to improve device versatility (e.g., flexible, lightweight devices) has resulted on the maturity of the technology toward inorganic thin film photovoltaics (TFSCs) and, more recently, into a new generation called excitonic solar cells (XSCs) [1–3]. The most representative examples of XSCs are dye-sensitized solar cells [1, 4], hybrid solar cells [3, 5–9], and organic polymer solar cells [2, 3, 10–13]. The main difference between inorganic solar cells (SiSCs and TFSCs) and XSCs is their photoconversion mechanism. In an inorganic solar cell, the absorption of light leads to the direct formation of free

electron-hole pairs and the direct production of charge carriers. In XSCs, the absorption of photons from sunlight produces an exciton, a tightly electron-hole pair which must be separated for charge generation. The exciton dissociation takes place at the interface of the constituent semiconductors. While inorganic solar cells rely only on inorganic semiconductors, XSCs apply organic semiconductor materials as part of the device active layer. The presence of organic semiconductors (e.g., dyes, small molecules, organic polymers) implies weak intermolecular interactions and low dielectric constants which leads to greater localization of the photoexcited states. The dissociation of the exciton can only be possible if the band offset energy at the junction is greater than the exciton binding energy in the material from which it was generated. In comparison, excitons in SiSCs have binding energies of about 20 meV, and, as a consequence, the absorption of photons generates free charge carriers in the bulk. The electron and hole formed segregates to the electrodes due to the build-up potential generated at the p-n junction. These different mechanisms indicate that each type of solar cell shows a high level of dependence on the crystallinity of the materials applied. Therefore, inorganic solar cells require highly pure and highly crystalline inorganic semiconductors (e.g., Silicon), while XSCs can utilize less crystalline and less pure compounds (e.g., TiO_2). The application of less crystalline materials together with organic semiconductors indicates that the interface between constituents must be carefully controlled, but also permits the fabrication of XSCs by low-cost and scalable printing techniques. The key for highly efficient XSCs is the control of the interfaces through materials nanostructuring.

Classification of Excitonic Solar Cells (XSCs)

XSCs can be classified by the type of semiconductor material applied or by the type of p-n junction used in the final device [7]. Both types of classification are strongly interconnected and sometimes generate confusion. A good example is given by the terms *bulk heterojunction solar cell* and *organic solar cell*, which are usually employed when a blend between a polymer and an electron transport material (e.g., P3HT and PCBM) is used as the active layer. In this case, the term *organic solar cell* indicates that the device is being labeled depending on the materials employed (two organic semiconductors). In the case



Nanomaterials for Excitonic Solar Cells, Fig. 1 A general schematic representation of an excitonic solar cell (XSC) showing the active layers made of the hole transport material (HTM)

and the electron transport material (ETM). The possible extra layers that can be found in XSCs are also indicated. The interface plays a crucial role in exciton dissociation

of the term *bulk heterojunction*, the solar cell is defined by the type of p-n junction, since the organic semiconductors employed for the fabrication of the device (donor and acceptor) are mixed heterogeneously in the bulk.

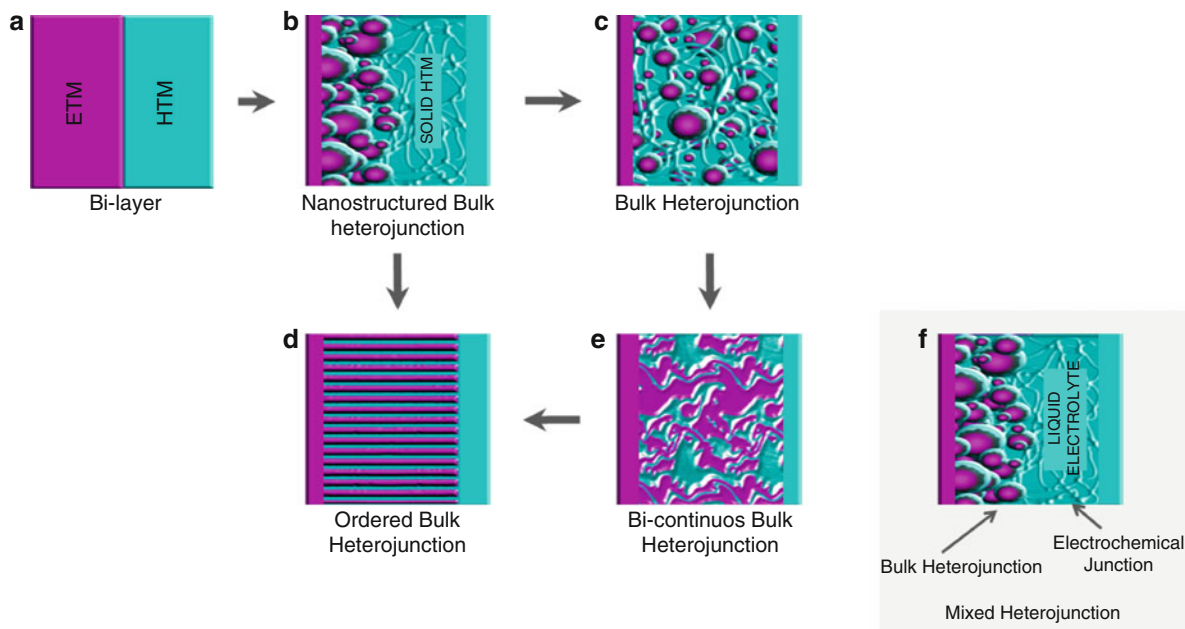
Before a general classification can be defined, the identification of the basic components of XSCs must be made. The main parts of these systems are the electron transport material (ETM) and the hole transport material (HTM) that are in contact through an *interface* (Fig. 1). In some cases, the ETM and the HTM are separated by a thin layer of a light-harvesting molecule (LHM), like the dye in a dye-sensitized solar cell, or inorganic quantum dots (QD) in some hybrid solar cells. In other cases, the light-harvesting material has a dual purpose acting also as the HTM, for example, in an oxide/polymer hybrid solar cell, where the polymer is the responsible for light harvesting and hole transport at the same time. The ETM, LHM, and HTM are “sandwiched” between two electrically conductive electrodes, accountable for charge collection. Thus, in a working device, the LHM is the responsible for photon absorption, the exciton that is formed is then dissociated at the *interface* between the HTM and ETM. Both the ETM and the HTM are responsible for charge transport (electrons and holes, respectively) through the current collectors. Thus, the LHM, the ETM, and the HTM are the main components of XSCs for which the device is being classified. Other components of the solar cell offer benefits to the whole device, for example, oxide buffer layers in DSCs or PEDOT and oxide layers in OSCs, among others. These extra layers are used for multiple purposes, especially to facilitate electron or hole transport or to improve interface connection, but are not the main

active layer of the solar cell. Thus, in order to identify the type of XSCs under study, a clear differentiation of these extra layers from the main active layer of the XSCs should be made. An easy rule of thumb is that the XSCs will never function without the main active layers, but they can still show photovoltaic response (perhaps not optimal) when the “extra” layers are not present. Figure 1 shows a schematic representation of a XSCs indicating each layer within the device.

Classification by the Type of p-n Junction

XSCs can be classified by the type of junction used [3]. However, it has to be clear that the p-n junction is composed by the ETM and the HTM. Examples of ETM are semiconductor oxides like TiO₂ or ZnO, inorganic semiconductors like Sb₂S₃ or CdS, or C₆₀ or C₇₀ derivatives. Conducting organic polymers usually act as HTM, but some of them can also act as the light-harvesting material (LHM) like polythiophenes. There are also HTMs made of small molecules like the *spiro-OMeTAD*. In the case of LHM, a dye molecule is an excellent example. These dyes are usually applied in DSC, but some conducting polymers and inorganic quantum dots can also act as the LHM.

In an XSC, the p-n junction is a critical parameter, since the exciton diffusion length is limited to 10–20 nm in polymers and up to 1 μm in high-quality small molecule films. Thus, in the device structure, the p-n junction must maximize the volume of the light-harvesting material at the interface of the constituent semiconductors, contributing in this way to an optimum charge generation and providing mobility pathways to enhance charge collection. A schematic representation of some of the possible p-n junctions is illustrated in Fig. 2. The most basic p-n junction is



Nanomaterials for Excitonic Solar Cells, Fig. 2 Schematic representation of p-n junctions found in XSCs: (a) bilayer, (b) nanostructured bilayer, (c) nanostructured bulk heterojunction,

(d) ordered bulk heterojunction, (e) bicontinuous bulk heterojunction, and (f) mixed bulk heterojunction and electrochemical junction

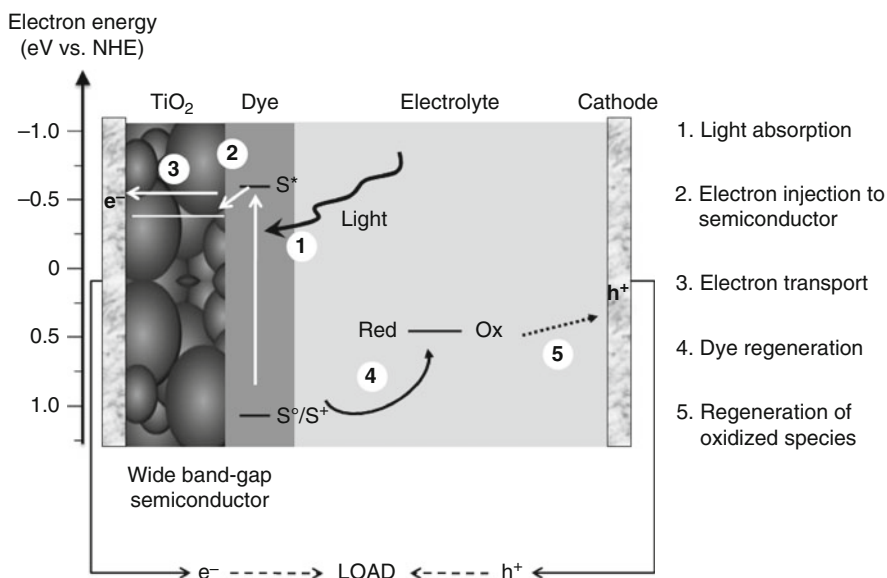
the bilayer (Fig. 2a) where a smooth and flat interface is made between the HTM and the ETM. The example in Fig. 2b represents a nanostructured bulk heterojunction where the surface area of the p-n junction has been enhanced by the application of nanoparticles increasing the amount of LHM present at the interface, if compared to bilayer flat devices represented in Fig. 2a. This type of heterojunction can be well ordered, as represented in Fig. 2d, by the application of vertically aligned nanowires, nanorods, nanotubes, or similar. If the nanoparticles in the electrode shown in Fig. 2b, are distributed heterogeneously within a polymer matrix, the final electrode will look like as in Fig. 2c. The latter configuration is now called a “bulk heterojunction.” In this junction, the electrons and holes must be transported through the bulk and reach the current collector; thus an interconnection between materials constituents (especially between molecules, nanoparticles, or similar) must be made. The latter is achievable usually by the application of thermal treatments (or similar), and the resulting p-n junction is a bicontinuous network represented in Fig. 2e. A bicontinuous bulk heterojunction is usually observed in organic solar cells (OSC) where the ETM, a C₆₀ derivative like the PCBM, is well dispersed

within a polymer matrix, like the P3HT. The same heterojunction can also be found in, for example, an HSC of the type polymer/oxide where different nanoforms of ZnO, can be dispersed on a polymer matrix like P3HT polymers. The bicontinuous bulk heterojunction is characterized by the random dispersion of the active constituents. A drawback is observed if some of the materials used stay isolated after thermal treatment, forming islands that cannot account for exciton dissociation, and an unoptimized charge generation is observable from the solar cell. Thus, optimization of the bulk heterojunction to enhance charge separation at the interface is possible if these interfaces can be arranged, for example, by the application of well-ordered nanostructured electrodes. Nanostructuration provides order, and order means control over the electron transport and charge collection. Figure 2d represents an example of such electrode: a well-ordered nanostructured bulk heterojunction made of vertically aligned nanorods, nanowires, or similar (other different nanoforms are also possible).

Two or more heterojunctions can be part of a single device. For example, Fig. 2f represents a device where a bulk heterojunction and an electrochemical junction

Nanomaterials for Excitonic Solar Cells,

Fig. 3 Schematic representation of a DSC. The photovoltaic process initiates with the absorption of photons from sunlight on the light-harvesting dye (S). The different processes taking place are light absorption (1), electron injection into the TiO₂ (2), electron transport through the TiO₂ (3), dye regeneration (4), and electrolyte regeneration (5) (Adapted from Ref. [1])



are used. This example is representative of a DSC where the bulk heterojunction is made between the interaction of the TiO₂ electrode and the light-harvesting dye, and the electrochemical junction is made by the TiO₂/dye electrode and the liquid electrolyte. In a solid-state dye-sensitized solar cell (ss-DSC), where a solid organic semiconductor replaces the liquid electrolyte, both junctions are bulk heterojunctions [3].

Classification by the Type of Material

Classification of XSCs can be made by the type of semiconductor applied: organic (polymers, dyes, small organic molecules, etc.) or inorganic (e.g., ZnO, TiO₂, CdSe, etc.). The outcome is three basic solar cell types: organic solar cells (OSCs), hybrid solar cells (HSCs), and dye-sensitized solar cells (DSCs). In OSCs, both materials are purely organic semiconductors, for example, the blend between two conducting organic polymers or a mixture between a C₆₀ derivative and a conjugated polymer. In the case of HSC, the device is characterized for the presence of, at least, one phase of an inorganic semiconductor (e.g., Sb₂S₃, TiO₂) and a second phase of an organic semiconductor, like a conducting organic polymer. Although DSC can be classified also as HSC (due to the application of organic and inorganic materials), it is considered

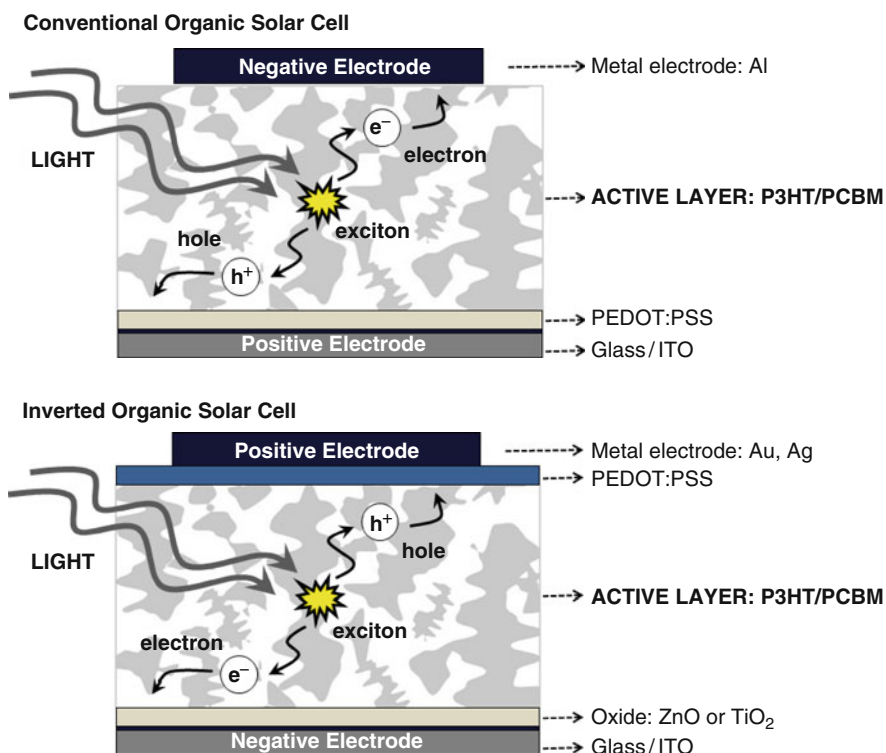
a different type of XSC, due to the presence of an electrolyte (semiliquid, gel, solid, etc.).

Dye-Sensitized Solar Cell (DSC)

A DSC is characterized by an electrode made of a semiconductor oxide, usually TiO₂ but others like ZnO can also be employed. A charge transfer dye is anchored to the surface of the semiconductor oxide, and finally, a liquid electrolyte containing a redox couple, usually iodide/triiodide in a solvent, completes the device (Fig. 3). The photovoltaic process initiates when photons from sunlight are absorbed by the dye. The exciton is formed and dissociated at the interface between the dye molecule and the semiconductor oxide. The corresponding electron travels through the semiconductor oxide until reaching the current collector (a transparent electrode), and the hole travels through the liquid electrolyte. The regeneration of the dye is possible when it accepts an electron from the electrolyte. The iodine is regenerated by the reduction of the triiodide at the counter electrode (Pt), and the circuit is complete by the migration of the electron through an external load. The voltage generated from the DSC under illumination corresponds to the difference between the Fermi level of the

Nanomaterials for Excitonic Solar Cells,

Fig. 4 Schematic representation of organic solar cells (OSC) showing the two configurations: conventional and inverted. Charge generation is possible through three general steps: photon absorption, charge separation, and charge extraction



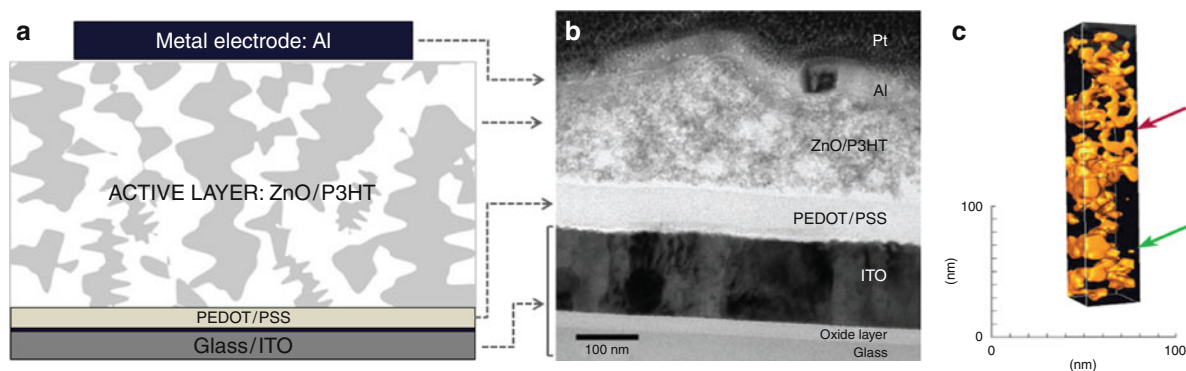
electron in the oxide and the redox potential of the electrolyte [1, 4].

A variation of the DSC is the solid-state dye-sensitized solar cell (ss-DSC) that rose in order to overcome the problems observed in DSC due to the application of liquid electrolyte (solvent evaporation, electrode corrosion, among others). Thus solid, semi-solid, gel-type, or nanoclay-based hole conductors have been employed to replace it.

Organic Solar Cells (OSC)

Organic solar cells are characterized for the presence of organic semiconductors in a bulk heterojunction configuration as the main active layer [2, 10–13]. The best combination is the blend between a C60 derivative, PCBM, acting as the ETM and a conducting organic polymer as the HTM, yet other combinations of organic semiconductors are possible (e.g., the blend between two polymers) [14, 15]. In general, the photovoltaic energy conversion in OSCs is shown schematically in Fig. 4. Photons from sunlight are absorbed

by the light-harvesting component, usually a polymer. An exciton is formed at the interface between the HTM and the ETM (or donor/acceptor interface) and dissociates in electron and hole. The electron travels through the ETM and the hole through the HTM until reaching the electrodes and generate charge. Since the active materials that form the bulk heterojunction are distributed on the whole active layer, thin films used, made of organic PEDOT:PSS or inorganic semiconductor oxides, determine the direction at which the electrons and holes migrate, as shown in Fig. 4. Thus, two types of configurations can be found in OSC: the conventional and the inverted. In a conventional OSC, a thin film of PEDOT:PSS is placed between the transparent conducting oxide (TOC) and the active material. Thus, the holes generated at the interface will be transported toward the PEDOT:PSS electrode and the electrons toward the opposite direction. In an inverted device, a thin-film layer of a semiconductor oxide is placed between the TOC and the active layer, while in the opposite side of the solar cell, a layer of the PEDOT:PSS polymer is used. This configuration allows for the electrons to travel toward the side of



Nanomaterials for Excitonic Solar Cells, Fig. 5 (a) Schematic representation of an HSC made of ZnO/P3HT bulk heterojunction. (b) SEM image of transmission electron micrograph of a cross section of the photovoltaic cell. The platinum layer is deposited later to allow preparation of the TEM sample.

(c) The *lower arrow* indicates an isolated ZnO domain. The *upper arrow* indicates a ZnO domain that is connected to the top, but not through a strictly rising path (Copyright Nature Publishing Group for images (b) and (c) [16])

the oxide thin film, while holes are transported in the opposite direction toward the PEDOT:PSS layer. Although lower photovoltaic properties have been found for the inverted solar cell configuration, the lifetime of the device is being enhanced.

Hybrid Solar Cells (HSC)

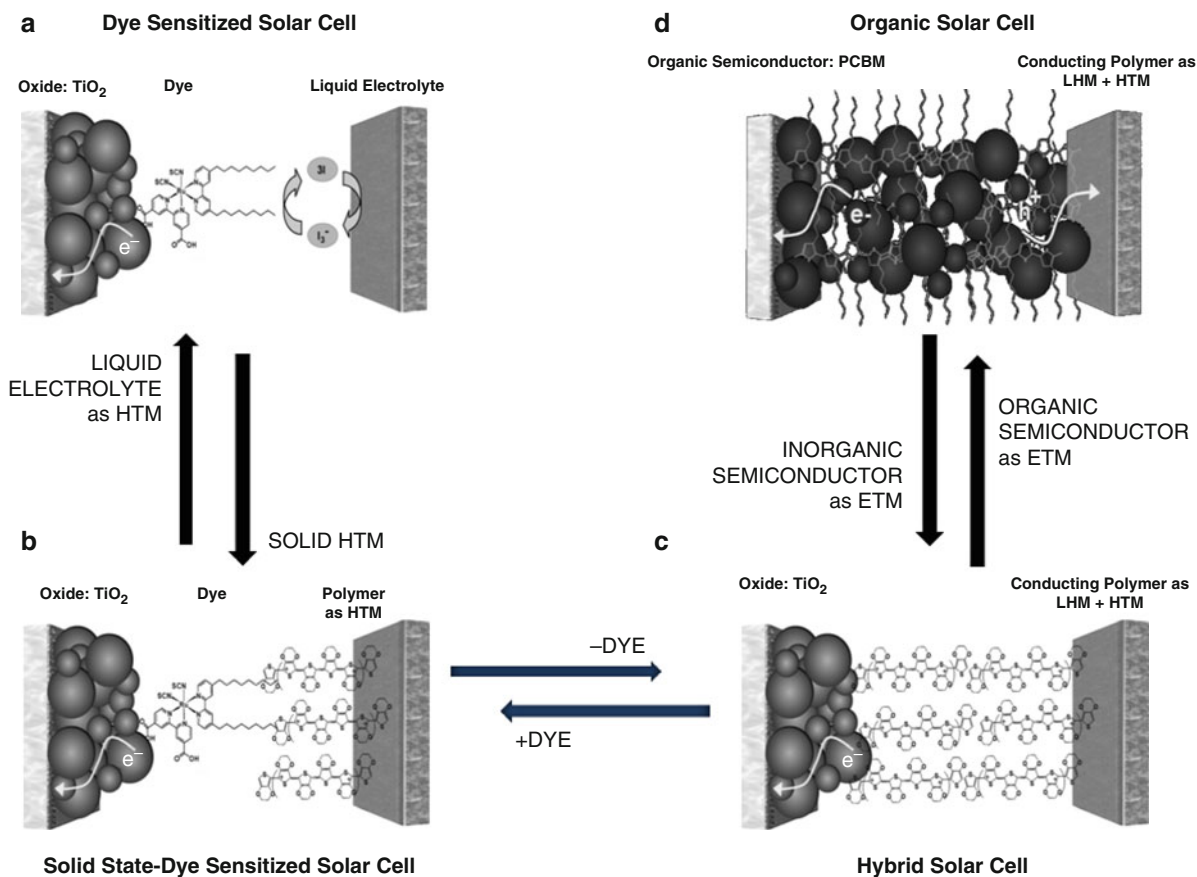
Hybrid solar cells present very similar characteristics and mechanisms with OSC; the main difference is that at one of the main components is an inorganic semiconductor [5–9, 16, 17]. The ETMs in an HSC are, for example, semiconductor oxides like TiO_2 , ZnO, etc., or inorganic semiconductors like CdS, CdSe, and Sb_2S_3 . In both cases, the inorganic semiconductor can be used in any form: as a dense layers, nanoparticles, nanorods, nanowires, etc. Except for the case of dense layers, all other nanoforms can be mixed within a conducting organic polymer to form a bulk heterojunction configuration. Figure 5 shows an example of an HSC. The active layer is composed of nanoparticles of ZnO (the ETM) in direct contact with a conducting organic polymer like P3HT (as the HTM) acting as the main active layer. The presence of a PEDOT:PSS layer on the ITO side is used as the hole collector, and the Al metal electrode as the electron collector. The position of these electrodes indicates that the device is in its conventional configuration.

Figure 5c shows a SEM-3D tomography of a section of the active layer of an HSC made of nanoparticles of

ZnO and a conducting polymer, P3HT. In the image, the polymer P3HT is represented as transparent and the ZnO nanoparticles in gray. The arrows show two of the main problems observed in bulk heterojunctions: the presence of islands or isolated material (useless as charge carriers) and the connection of some ZnO domains connected only by the top but not through a correct path for charge transport. The latter reduces the efficient electron transport lowering overall device efficiency.

Device configuration

All XSCs share similar device configuration, active materials, and p-n junctions. To illustrate the latter, in Fig. 6 representation of different XSCs is shown, where the same device configuration is maintained though all the examples and only the materials employed have been changed depending on the type of XSC. Figure 6a represents the DSC device made with a nanoparticulated electrode (usually TiO_2) with a Ru-based dye anchored to it. A redox liquid electrolyte completes the device. All the components are “sandwiched” between two electrically conductive electrodes, FTO and Pt respectively, where one of these electrodes must be transparent for light irradiation. A modification of the DSC has been made by the replacement of the liquid electrolyte, with the purpose of minimizing its disadvantages like solvent evaporation, electrode corrosion, etc. Thus, the replacement of



Nanomaterials for Excitonic Solar Cells, Fig. 6 Schematic representation of XSCs by the type of material applied and the relation between solar cell types: (a) dye-sensitized solar cell

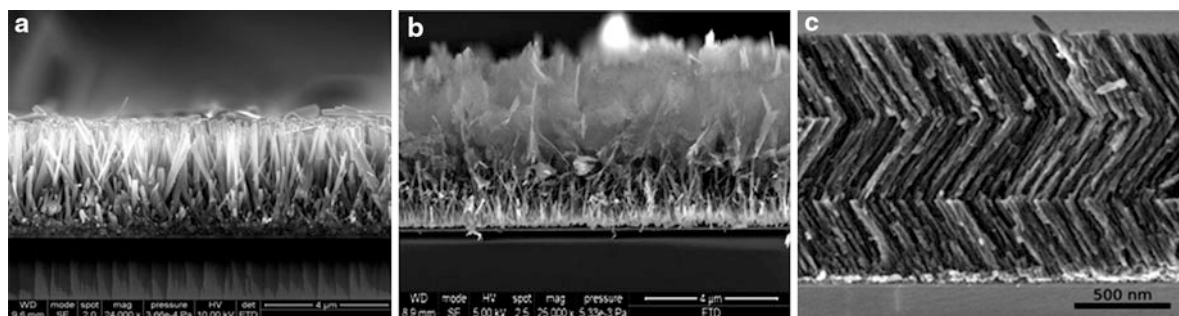
(DSC), (b) solid-state dye-sensitized solar cell (ss-DSC), (c) hybrid solar cell (HSC), and (d) organic polymer solar cell (OSC)

the liquid electrolyte in a DSC by a solid HTM, like a conducting organic polymer or organic molecule like *spiro*-OMeTAD, results in what is called a solid-state dye-sensitized solar cell, ss-DSC (Fig. 6b). The light-harvesting dye and the solid hole conductor can be replaced by a single material, for example, a polymer like P3HT which is characterized by a dual function: to collect photons from sunlight and to work as a hole transport. In this simplified version of the ss-DSC, where the ETM is an inorganic semiconductor (e.g., TiO₂, CdS, or similar) and the HTM is an organic polymer (P3HT, MEH-PPV, etc.), the solar cell is known as a hybrid solar cell, HSC (Fig. 6c). The term hybrid is used since the device is made of both organic and inorganic compounds. In this HSC, replacing the *inorganic* ETM semiconductor (e.g., TiO₂, CdS, or similar) by an *organic* ETM semiconductor (e.g., C₆₀, C₇₀, or similar) without modifying the

HTM (same organic semiconductor) would result in an organic solar cell (Fig. 6d).

The Benefit of Nanostructuring in XSCs

The development of new materials tailored at the nanoscale has become a major research area due to the unique and novel properties that can be obtained. Nanostructured excitonic solar cells have shown the advantages of controlling the interfaces at the nanoscale level. For example, in a bulk heterojunction made of inorganic nanoparticles and a conducting organic polymer, the inorganic semiconductor oxide nanoparticle is dispersed within the conjugated polymer matrix. The resulting oxide nanoparticle/polymer blend must be annealed to improve nanoparticle connection and to favor electron pathway through the



Nanomaterials for Excitonic Solar Cells, Fig. 7 Well-ordered semiconductor oxide (ZnO and TiO₂)-based electrodes applied in XSCs: (a) vertically aligned ZnO NANORODS [7]

(b) vertically aligned ZnO nanotrees, and (c) well-ordered TiO₂ nanocolumns with tailored inclination angle (From Ref. [18], copyright 2011, Royal Society of Chemistry)

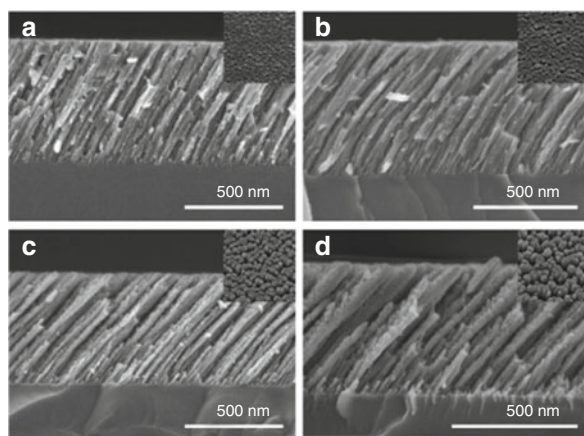
electrode. Nevertheless, these devices are still limited by incomplete exciton dissociation due to the isolation of some nanoparticles from each other after thermal treatment.

A nanostructured thin film, for example, well-crystalline and vertically aligned nanowires can provide higher interfacial area between the donor and the acceptor material (polymer/oxide, respectively) with highly efficient electron transport pathways. In the case of DSCs, the replacement of the nanoparticles electrode by vertically aligned nanostructures emerged also as a possibility to obtain faster electron transport thus improving solar cell efficiency [7]. Moreover, the need to replace problematic liquid electrolytes in DSC by solid hole conductors with slower kinetics impels the application of faster electron transport materials like in vertically aligned nanostructures. Figure 7 shows SEM images of different oxide nanostructures, from nanorods (a), nanotrees (b), nanocolumns (c). The electron transport in an electrode made of nanoparticles is limited by the grain boundaries encountered between each nanoparticle. The application of nanorods, especially if well ordered, paves the way for an efficient electron transport through the crystalline structure. The nanostructure can also be tailored to enhance light absorption without eliminating the advantage of the nanorods. The nanotrees (Fig. 7b) were synthesized with a dual purpose: to enhance the light-harvesting properties of the nanorods by increasing the surface area and by maintaining the good electron transport properties that characterize nanorods.

A practical approach of electrode nanostructuration where high transparency is needed has been recently reported. In this case, increasing the NR length will

increase light scattering and decrease the transparency of the DSSC device. A singular example to overcome these problems is the application of nanostructured thin films [7, 18]. Transparent TiO₂ thin films were grown by the physical vapor deposition method at glancing incidences (PVD-GLAD). This method of synthesis furnishes aligned nanocolumnar layers at well-controlled oblique angles. This nanostructure is advantageous for solar collection, as it has already been proved that tilting the nanocolumns improves light entrapment efficiency. Zig-zag and other complex morphologies can be also prepared by this method. The resulting thin-film electrodes present very open and porous microstructure where nanometer size columns with a high internal porosity are separated by wide pores that extend from the electrode substrate up to the film surface. Many are the factors to be considered when synthesizing these nanostructures: the effect of electrode thickness, TiO₂ crystallinity, topology when arranged into multilayers, the pore structure, or the distribution and concentration of dye molecules within the film thickness. Figure 8 shows SEM images (left) of ordered nanocolumns of TiO₂ deposited by the GLAD technique at different angles. Each sample is 500 nm thick and the angles at which were deposited vary between 60° and 85°. Analyses of the samples as electrode in DSC demonstrated that an optimal deposition angle exists at 70°.

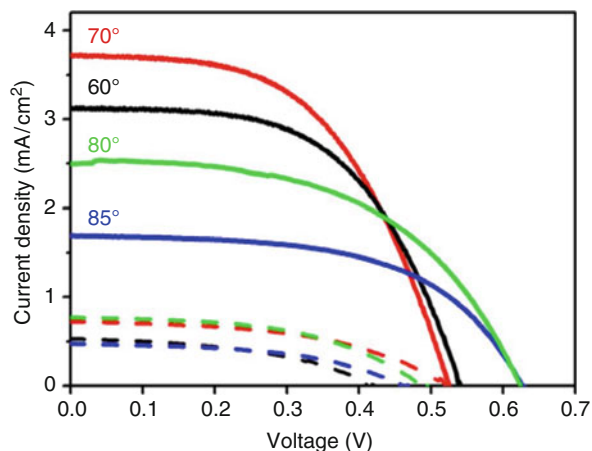
The enhancement observed has been attributed to the increase in the dye-loading capacity of the thin films. Setting the deposition angle at around 70°, at which a 500-nm-thick electrode solar cell presents a maximum efficiency of 1.04%. A zig-zag nanostructure of the TiO₂ thin films made of stacks of 500-nm layers deposited at the optimum 70° angles, was



Nanomaterials for Excitonic Solar Cells, Fig. 8 Left: Ordered nanocolumns of TiO_2 deposited by the GLAD technique at different angles (From Ref. [18], copyright 2011, Royal

applied to grow thicker films. This configuration ensured good mechanical stability and high transparency for the films. Power conversion efficiency was observed to increase with electrode thickness, reaching a conversion efficiency of 2.78% for 3- μm electrodes of quite good transparency [18]. The high transparency of the obtained cells supports its implementation onto colored glass and other related applications.

There are many possible methods to fabricate nanomaterials, and the observed properties depend a great extent on the preparation conditions. Nanostructure synthesis can be found as simple nanoparticles or as nanorods, nanobelts, branched nanorods, nanowires, ultranarrow nanobelts, hierarchical nanostructures, nanocombs, nanosprings, nanospirals, seamless nanorings, core-shell nanostructures, nanocages, nanoflowers, among many others. It is, however, also difficult to establish the optimal synthesis methodology for solar cell applications because the efficiency of an XSC depends on many aspects and not only on the material nanostructure, like thin film thickness, nanostructure dimension or surface area, but also on overall device preparation like polymer or dye applied, back contacts, active area, etc. However, for XSCs, the most attractive syntheses techniques due to the ease of its process, the low temperature employed, and the possibility to obtain the oxide nanostructures on a large area by continuous methods and in different nanostructures are the hydrothermal methods, among them the chemical bath deposition. The SILAR method is also an alternative



Society of Chemistry), and **Right:** IV curves of the DSC applying the TiO_2 electrodes

which interest is growing due to the ease of its application. These methods permit the synthesis of materials not only as nanoparticles but also as well-aligned nanostructures or core-shell nanostructures. These synthesis methods have been applied for the synthesis of many different inorganic semiconductors like TiO_2 , ZnO , CdSe , etc. The synthesis involves precursor baths with the ions of the materials to be synthesized and the technique can be applied for the synthesis of isolated materials (e.g., nanoparticles), the fabrication of nanostructured electrodes, or the sensitization of molecules on different substrates. One of the most attractive examples applying the hydrothermal method is the synthesis of vertically aligned nanostructures. The reason behind the interest in these materials is due to the excellent transport properties if compared to nanoparticles. An example is the synthesis of ZnO nanorods that can be made at low temperature, on almost any substrate, and in a continuous mode [19]. The latter are attractive conditions for large scale production. In general, the method to obtain vertically aligned ZnO nanorods is based on a two-step technique involving the coating of a substrate with ZnO seeds of less than 10 nm in diameter. Every ZnO seed works as a nucleation site for the formation of ZnO nanowires under mid-wet conditions. The synthesis conditions, like zinc concentration in solution, growth time, or growth temperature, determine the final nanorod dimensions and quality. Also the stability of the precursor ZnO seeds (nanoparticles) or the substrates to be applied are important factors affecting the final ZnO

nanostructures with reproducible results. As an example, the variation of ZnO nanorod length and diameter can be adjusted depending on the additives added during synthesis. Polyethylenimine (PEI) has been used for large surface area electrodes, so higher amount of dye absorption can be obtained when applied in a DSC. Nanorods as long as 30 μm were observed when PEI additive was used; without additive the length of the nanorods was not more than 5 μm , and their diameters increased also with time. Yet, the longest nanorods, 20–25 μm , show only one fifth of the surface area of an electrode made of nanoparticles [7]. Vertically aligned nanostructures of ZnO have been extensively applied in XSCs. Nevertheless, in terms of upstanding and well-ordered alignment, the current synthesis methodologies applied for their fabrication produce irreproducible nanostructures with dimensions that are too far away from the ideal electrode characteristics. Synthesis condition must be controlled carefully; the minimal change in synthesis condition or even the roughness of the underlying substrate can affect nanorod dimension. Thus, a careful design of the nanostructured electrode made of vertically aligned nanorods should be made in order to fulfill the best requirements for the type of solar cell where it will be applied. For example, one of the most problematic issues observed in XSCs is dye or polymer infiltration (wettability); a solvent-soluble dye could reach the whole surface area of the nanostructured electrode more easily than a viscous polymer solution which has to penetrate into dense packed nanowire network.

Vertically aligned nanocolumns (nanorods, nanowires, nanotips, etc.) made of semiconductor oxides have been proved to enhance electron transport and reduce the recombination probability in different excitonic solar cells (organic, dye sensitized, or hybrid). Nevertheless, the power conversion efficiencies of these modified solar cells are still far below that obtained with nanoparticle-based electrodes. Their lower efficiency has been attributed to the relatively low surface area of nanocolumnar electrodes and some inefficiency of light absorption since some photons fall on the gap between adjacent vertical nanorods. Moreover, the possibility to obtain highly porous ZnO nanomaterials with good crystalline structure is possible and it suggests that there is room for improvement on solar cell efficiency. In principle, the efficiency of these type of solar cells could be enhanced by

increasing the thickness of the aligned nanocolumnar thin films. The solar cells can be nanostructured to get higher surface areas and longer nanocolumns. Nevertheless, recent results show that even NR lengths as large as 40 μm result in low power conversion efficiency. Most recent results show that the modification of the synthesis methodology by the application of additives or the tuning of the synthesis conditions (temperature, time, etc.) reduces traps found on the semiconductor surface. The challenge nowadays resides in the interplay between large-area printing technology and high-quality nanostructuring. The control of both parameters will balance final device cost and functionality.

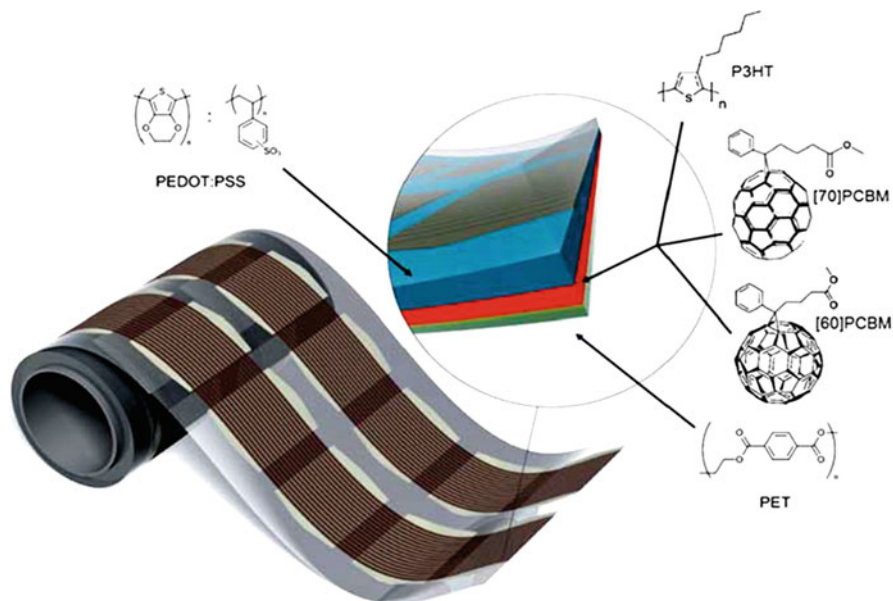
Future Trends: From Nanomaterials to Large-Scale Printed Devices

XSCs are promising devices due to the possibility to be fabricated by scalable and low-cost printing techniques (as example see Fig. 9), unachievable nowadays for photovoltaic systems based on crystalline silicon. Application of printing methods, like screen printing, roll-to-roll, gravure, offset, lithography, inject, etc., will also permit the integration of other printing electronic devices (lighting, batteries, etc.) to constitute a complete system [10, 12, 20–22]. A challenge in printing technology is to achieve well-ordered nanostructuring to the main active layer. For the latter, friendly and low-temperature processes (synthesis, pre-/post treatments, etc.) are required.

The development of liquid precursors or “inks,” is the key for the scale-up of laboratory devices to commercialization of large-scale printed photovoltaics. Inks can be made from solutions or suspensions. Solution-based inks present advantages like uniform composition and phase, and the possibility of modification of the material (e.g., doping). Nevertheless, the ink solubility may depend on factors such as additives, solvents, humidity, and temperature, among others. All these parameters must be carefully controlled in order to obtain a final useful ink. The preparation of suspensions is another alternative which allows the application of nanostructured solids designed and synthesized at the nanoscale level, with good results [20–22]. Moreover, some recent synthesis methods allow for certain degree of control of the final nanoforms [22].

Nanomaterials for Excitonic Solar Cells, Fig. 9

A roll of the printed modules showing how two parallel sets of modules were prepared simultaneously. The zoom-in shows the layer stack which is PET-ITO-ZnO-P3HT:PCBM-PEDOT:PSS-printed silver. The device stack was encapsulated using roll-to-roll lamination post production (From Ref. [20], copyright 2011 Royal Society of Chemistry)



In the suspension, the problem of the elimination of organic precursors or additives after treatments is a challenge, but the benefits range from tunable electrical and optical properties to the enhancement of mechanical characteristics.

Cross-References

- ▶ [Hybrid Solar Cells](#)
- ▶ [Organic Photovoltaics: Basic Concepts and Device Physics](#)
- ▶ [Plasmonic Structures for Solar Energy Harvesting](#)
- ▶ [Toward Bioreplicated Texturing of Solar-Cell Surfaces](#)

References

1. O'Regan, B., Grätzel, M.: A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO₂ films. *Nature* **353**(6346), 737–740 (1991)
2. Sun, S.-S., Sariciftci, N.S.: *Organic Photovoltaics: Mechanisms, Materials, and Devices* (Optical Science and Engineering). CRC Press, Boca Raton (2005). ISBN: 0-8247-5963-X
3. Nelson, J.: *The Physics of Solar Cells*, p. 360. Imperial College Press, London (2003)
4. Hagfeldt, A., Boschloo, G., Sun, L., Kloo, L., Pettersson, H.: Dye-sensitized solar cells. *Chem. Rev.* **110**(11), 6595–6663 (2010)
5. Beek, W.J.E., Janssen, R.A.J.: Hybrid polymer-inorganic photovoltaic cells. In: Merhari, L. (ed.) *Hybrid Nanocomposites for Nanotechnology: Electronic, Optical, Magnetic and Biomedical Applications*, pp. 321–386. Springer, New York (2009)
6. Lira-Cantu, M., Gomez-Romero, P.: Multifunctional hybrid materials based on conducting organic polymers. Nanocomposite systems with photo-electro-ionic properties and applications. In: Gomez-Romero, P., Sanchez, C. (eds.) *Functional Hybrid Materials*, pp. 210–269. Wiley-VCH, Weinheim (2004)
7. Gonzalez-Valls, I., Lira-Cantu, M.: Vertically-aligned nanostructures of ZnO for excitonic solar cells: A review. *Energy Environ. Sci.* **2**(1), 19–34 (2009)
8. Lira-Cantu, M., Gomez-Romero, P.: Next-generation hybrid nanocomposite materials based on conducting organic polymers: Energy storage and conversion devices. In: Merhari, L. (ed.) *Hybrid Nanocomposites for Nanotechnology: Electronic, Optical, Magnetic and Biomedical Applications*, pp. 289–320. Springer, New York (2009)
9. Nozik, A.J., Beard, M.C., Luther, J.M., Law, M., Ellingson, R.J., Johnson, J.C.: Semiconductor quantum dots and quantum dot arrays and applications of multiple exciton generation to third-generation photovoltaic solar cells. *Chem. Rev.* **110**(11), 6873–6890 (2010)
10. Krebs, F.C.: *Polymeric Solar Cells: Materials, Design, Manufacture*. DEStech, Lancaster (2010)
11. Günes, S., Neugebauer, H., Sariciftci, N.S.: Conjugated polymer-based organic solar cells. *Chem. Rev.* **107**(4), 1324–1338 (2007)
12. Brabec, C., Dyakonov, V., Scherf, U.: *Organic Photovoltaics: Materials, Device Physics and Manufacturing Technologies*. Wiley-VCH, Weinheim (2009)

13. Clarke, T.M., Durrant, J.R.: Charge photogeneration in organic solar cells. *Chem. Rev.* **110**(11), 6736–6767 (2010)
14. Segura, J.L., Martín, N., Guldi, D.M.: Materials for organic solar cells: The C60/ π -conjugated oligomer approach. *Chem. Soc. Rev.* **34**(1), 31–47 (2005)
15. Martín, N., Giacalone, F., Prato, M.: Fullerene Polymers: Synthesis, Properties and Applications, p. 305. Wiley-VCH, Weinheim (2009)
16. Oosterhout, S.D., Wienk, M.M., van Bavel, S.S., Thiedmann, R., Koster, L.J.A., Gilot, J., Loos, J., Schmidt, V., Janssen, R.A.J.: The effect of three-dimensional morphology on the efficiency of hybrid polymer solar cells. *Nat. Mater.* **8**, 818–824 (2009)
17. Kamat, P.V., Tvrđy, K., Baker, D.R., Radich, J.G.: Beyond photovoltaics: Semiconductor nanoarchitectures for liquid-junction solar cells. *Chem. Rev.* **110**(11), 6664–6688 (2010)
18. González-García, L., González-Valls, I., Lira-Cantu, M., Barranco, A., Gonzalez-Elipe, A.G.: Aligned TiO₂ nanocolumnar layers prepared by PVD-GLAD for transparent dye sensitized solar cells. *Energy Environ. Sci.* (2001). doi:10.1039/C0EE00489H
19. Morkoc, H., Ozgur, U.: Zinc Oxide: Fundamentals, Materials and Device Technology. Wiley-VCH, Weinheim (2009)
20. Krebs, F.C., Fyenbo, J., Jorgensen, M.: Product integration of compact roll-to-roll processed polymer solar cell modules: methods and manufacture using flexographic printing, slot-die coating and rotary screen printing. *J. Mater. Chem.* **20**, 8994–9001 (2010)
21. Arias, A.C., MacKenzie, J.D., McCulloch, I., Rivnay, J., Saello, A.: Materials and applications for large area electronics: Solution-based approaches. *Chem. Rev.* **110**(1), 3–24 (2010)
22. Habas, S.E., Platt, H.A.S., van Hest, F.A.M., Ginley, D.S.: Low-cost inorganic solar cells: From ink to printed device. *Chem. Rev.* **110**(11), 6571–6594 (2010)

Nanomaterials for Sensors

- ▶ [Nanostructured Materials for Sensing](#)

Nano-mechanical Machining

- ▶ [Ultraprecision Machining \(UPM\)](#)

Nanomechanical Microcantilever Chemical and Biological Sensors

- ▶ [Microcantilever Chemical and Biological Sensors](#)

Nanomechanical Properties of Nanostructures

Bharat Bhushan

Nanoprobe Laboratory for Bio- & Nanotechnology and Biomimetics, The Ohio State University, Columbus, OH, USA

Synonyms

[Bending strength](#); [Creep](#); [Fatigue strength](#); [Fracture toughness](#); [Hardness](#); [Indentation](#); [Nanohardness](#); [Relaxation](#); [Young's modulus](#)

Definition

Structural integrity is of paramount importance in all devices. Load applied during the use of devices can result in component failure. Cracks can develop and propagate under tensile stresses, leading to failure. Knowledge of the mechanical properties of nanostructures is necessary for designing realistic micro-/nanoelectromechanical systems (MEMS/NEMS) and biological micro-/nanoelectromechanical systems (BioMEMS/BioNEMS) devices. Elastic and inelastic properties are needed to predict deformation from an applied load in the elastic and inelastic regimes, respectively. The strength property is needed to predict the allowable operating limit. Some of the properties of interest are hardness, elastic modulus, bending strength, fracture toughness, and fatigue strength. Many of the mechanical properties are scale dependent; therefore, these should be measured at relevant scales. Atomic force microscopy and nanoindenters can be used satisfactorily to evaluate the mechanical properties of micro/nanoscale structures.

Overview

A wide variety of MEMS, including Si-based devices, chemical and biological sensors and actuators, and miniature non-silicon structures (e.g., devices made from plastics or ceramics) have been fabricated with

dimensions in the range of a couple to a few thousand microns (see e.g., [4, 5, 9, 10, 15–17, 21–23, 29, 32, 40]). A variety of NEMS have also been produced (see e.g., [13, 14, 18, 33, 38, 43]). MEMS/NEMS technology and fabrication processes have found a variety of applications in biology and biomedicine, leading to the establishment of an entirely new field known as BioMEMS/BioNEMS [2, 3, 9, 12, 19, 24, 30, 31, 37, 44]. The ability to use micro/nanofabrication processes to develop precision devices that can interface with biological environments at the cellular and molecular level has led to advances in the fields of biosensor technology [42, 44], drug delivery [34, 36], and tissue engineering [25]. The miniaturization of fluidic systems using micro/nanofabrication techniques has led to new and more efficient devices for medical diagnostics and biochemical analysis [39]. The largest “killer” industrial applications of MEMS include accelerometers (over a billion US dollars a year in 2004), pressure sensors for manifold absolute pressure sensing for engines (more than 30 million units in 2004), and tire pressure measurements, inkjet printer heads (more than 500 million units in 2004), and digital micromirror devices (about US \$700 million revenues in 2004). BioMEMS and BioNEMS are increasingly used in commercial applications. Largest applications of BioMEMS include silicon-based disposable blood pressure sensor chips for blood pressure monitoring (more than 25 million units in 2004), and a variety of biosensors.

Structural integrity is of paramount importance in all devices. Load applied during the use of devices can result in component failure. Cracks can develop and propagate under tensile stresses leading to failure [6, 8]. Friction/stiction and wear limit the lifetimes and compromise the performance and reliability of devices involving relative motion [4, 5, 7, 9, 10]. Most MEMS/NEMS applications demand extreme reliability. Stress and deformation analyses are carried out for an optimal design. MEMS/NEMS designers require mechanical properties on the nanoscale. Elastic and inelastic properties are needed to predict deformation from an applied load in the elastic and inelastic regimes, respectively. The strength property is needed to predict the allowable operating limit. Some of the properties of interest are hardness, elastic modulus, creep, bending strength (fracture stress), fracture toughness, and fatigue strength. Micro/nanostructures have some surface topography and local scratches

dependent upon the manufacturing process. Surface roughness and local scratches may compromise the reliability of the devices, and their effect needs to be studied.

Most mechanical properties are scale dependent [5, 6, 8–10]. Several researchers have measured the mechanical properties of silicon and silicon-based milli- to microscale structures including tensile tests and bending tests, resonant structure tests for the measurement of elastic properties, fracture toughness tests, and fatigue tests. A few researchers have measured the mechanical properties of nanoscale structures using atomic force microscopy (AFM) and nanoindentation. For stress and deformation analyses of simple geometries and boundary conditions, analytical models can be used. For analysis of complex geometries, numerical models are needed. Conventional finite element method (FEM) can be used down to a few tens of nanometer scale, although its applicability is questionable at nanoscale. FEM has been used for simulation and prediction of residual stresses and strains induced in MEMS devices during fabrication, to perform fault analysis in order to study MEMS faulty behavior, to compute mechanical strain resulting from doping of silicon, analyze micromechanical experimental data, and nanomechanical experimental data. FEM analysis of nanostructures has been performed to analyze the effect of types of surface roughness and scratches on stresses in nanostructures.

Commonly used materials for MEMS/NEMS are single-crystal silicon and silicon-based materials (e.g., SiO₂ and polysilicon films deposited by low pressure chemical vapor deposition (LPCVD) process) [9, 10]. An early study showed silicon to be a mechanically resilient material in addition to its favorable electronic properties. Single-crystal 3C-SiC (cubic or β -SiC) films, deposited by atmospheric pressure chemical vapor deposition (APCVD) process on large-area silicon substrates, are produced for high-temperature micro/nanosensor and actuator applications. Amorphous alloys can be formed on both metal and silicon substrates by sputtering and plating techniques, providing more flexibilities in surface-integration. Electroless deposited Ni-P amorphous thin films have been used to construct microdevices, especially using the so-called LIGA techniques. Micro/nanodevices need conductors to provide power as well as electrical/magnetic signals to make them functional. Electroplated gold films have found wide

applications in electronic devices because of their ability to make thin films and process simplicity.

As the field of MEMS/NEMS has progressed, alternative materials, especially polymers, have established an important role in their construction [9, 10]. This trend has been driven by the reduced cost associated with polymer materials. Polymer microfabrication processes, including micromolding and hot embossing techniques, can be orders of magnitude less expensive than traditional silicon photolithography processes. Many polymers are biocompatible, so these can be integrated into biomedical devices with minimal detrimental effects to the host or on the biofluids. An improvement in device functionality (relative to silicon) is also possible due to the mechanical properties of the polymer used, which have properties much closer to biological tissues. Polymer BioMEMS structures involving microbeams have been designed to measure cellular forces. Polymer materials most commonly used for biomedical applications include poly(methyl methacrylate) (PMMA), poly(dimethylsiloxane) (PDMS), and polystyrene. Another material of interest due to ease of fabrication is poly(propyl methacrylate) (PPMA), which has lower glass transition temperature (T_g) (35–43°C) than PMMA (104–106°C), which permits low temperature thermal processing.

Experimental Techniques for Measurement of Mechanical Properties of Nanostructures

Indentation and Scratch Tests Using Micro/Nanoindenters

A nanoindenter is commonly used to measure hardness, elastic modulus, and fracture toughness, and to perform micro/nanoscratch studies to get a measure of scratch/wear resistance of materials [5, 9–11].

Hardness and Elastic Modulus

The nanoindenter monitors and records the dynamic load and displacement of the three-sided pyramidal diamond (Berkovich) indenter during indentation with a force resolution of about 75 nN and displacement resolution of about 0.1 nm. Hardness and elastic modulus are calculated from the load-displacement data [5, 10, 11]. The peak indentation load depends on the mechanical properties of the specimen; a harder material requires a higher load for reasonable indentation depth.

Fracture Toughness

The indentation technique for fracture toughness measurement of brittle samples on the microscale is based on the measurement of the lengths of median-radial cracks produced by indentation. A Vickers indenter (a four-sided diamond pyramid) is used in a microhardness tester. A load on the order of 0.5 N is typically used in making the Vickers indentations. The indentation impressions are examined using an optical microscope with Nomarski interference contrast to measure the length of median-radial cracks, c . The fracture toughness (K_{IC}) is calculated by the following relation [26]:

$$K_{IC} = \alpha \left(\frac{E}{H} \right)^{1/2} \left(\frac{P}{c^{3/2}} \right) \quad (1)$$

where α is an empirical constant depending on the geometry of the indenter, H and E are hardness and elastic moduli, and P is the peak indentation load. For Vickers indenters, α has been empirically found based on experimental data and is equal to 0.016 [5]. Both E and H values are obtained from the nanoindentation data. The crack length is measured from the center of the indent to the end of crack using an optical microscope. For one indent, all crack lengths are measured. The crack length c is obtained from the average value of several indents.

Indentation Creep

For indentation creep test of polymer samples, the test is performed using a continuous stiffness measurements (CSM) technique [11]. In a study by Wei et al. [45], the indentation load was typically 30 μ N, and the loading rate was 3 μ N/s. The tip was held typically for 600 s after the indentation load reached 30 μ N. To measure the mean stress and contact stiffness, during the hold segment the indenter was oscillated at a peak-to-peak load amplitude of 1.2 μ N and a frequency of 45 Hz.

Scratch Resistance

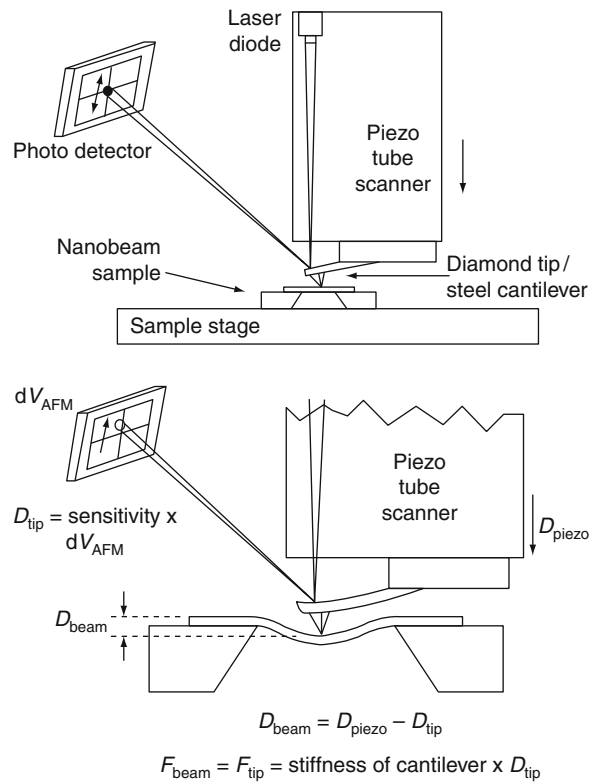
In micro/nanoscratch studies, in a nanoindenter, a conical diamond indenter having a tip radius of about 1 μ m and an included angle of 60° is drawn over the sample surface, and the load is ramped up until substantial damage occurs [9–11]. The coefficient of friction is monitored during scratching. In order to obtain scratch depths during scratching, the surface

profile of the sample surface is first obtained by translating the sample at a low load of about 0.2 mN, which is insufficient to damage a hard sample surface. The 500 μm long scratches are made by translating the sample while ramping the loads on the conical tip over different loads dependent upon the material hardness. The actual depth during scratching is obtained by subtracting the initial profile from the scratch depth measured during scratching. In order to measure the scratch depth after the scratch, the scratched surface is profiled at a low load of 0.2 mN and is subtracted from the actual surface profile before scratching.

Bending Tests of Nanostructures Using an AFM

Quasi-static bending tests of fixed nanobeam arrays in the normal direction are carried out using an AFM [41]. A three-sided pyramidal diamond tip (with a radius of about 200 nm) mounted on a rectangular stainless steel cantilever is used for the bending tests. The beam stiffness is selected based on the desired load range. The stiffness of the cantilever beams for application of a normal load up to 100 μN is about 150–200 N/m.

For the bending test, the tip is brought over the nanobeam array with the help of the sample stage of the AFM and a built-in high magnification optical microscope (Fig. 1) [41]. For the fine positioning of the tip over a chosen beam, the array is scanned in contact mode at a contact load of about 2–4 μN , which results in negligible damage to the sample. After scanning, the tip is located at one end of a chosen beam. To position the tip at the center of the beam span, the tip is moved to the other end of the beam by giving the X-piezo an offset voltage. The value of this offset is determined after several such attempts have been made in order to minimize the effects of piezo drift. Half of this offset is then applied to the X-piezo after the tip is positioned at one end of the beam, which usually results in the tip being moved to the center of the span. Once the tip is positioned over the center of the beam span, the tip is held stationary without scanning, and the Z-piezo is extended by a known distance, typically about 2.5 μm , at a rate of 10 nm/s, as shown in Fig. 1. During this time, the vertical deflection signal (dV_{AFM}), which is proportional to the deflection of the cantilever (D_{tip}), is monitored. The displacement of the piezo is equal to the sum of the displacements of the cantilever and the nanobeam. Hence the



Nanomechanical Properties of Nanostructures, Fig. 1 Schematic showing the details of a nanoscale bending test using an AFM. The AFM tip is brought to the center of the nanobeam and the piezo is extended over a known distance. By measuring the tip displacement, a load–displacement curve of the nanobeam can be obtained [41]

displacement of the nanobeam (D_{beam}) under the point of load can be determined as

$$D_{\text{beam}} = D_{\text{piezo}} - D_{\text{tip}} \quad (2)$$

The load (F_{beam}) on the nanobeam is the same as the load on the tip/cantilever (F_{tip}) and is given by

$$F_{\text{beam}} = F_{\text{tip}} = D_{\text{tip}} \times k \quad (3)$$

where k is the stiffness of the tip/cantilever. In this manner, a load–displacement curve for each nanobeam can be obtained.

The photodetector sensitivity of the cantilever needs to be calibrated to obtain D_{tip} in nm. For this calibration, the tip is pushed against a smooth diamond sample by moving the Z-piezo over a known distance. For the hard diamond material, the actual deflection of

the tip can be assumed to be the same as the Z-piezo travel (D_{piezo}), and the photodetector sensitivity (S) for the cantilever setup is determined as

$$S = D_{piezo}/dV_{AFM} \text{ nm}/V \quad (4)$$

In the measurements, D_{tip} is given as $dV_{AFM} \times S$.

Since a sharp tip would result in an undesirable large local indentation, Sundararajan and Bhushan [41] used a diamond tip which was worn (blunt). Indentation experiments using this tip on a silicon substrate yielded a residual depth of less than 8 nm at a maximum load of 120 μN , which is negligible compared to displacements of the beams (several hundred nanometers). Hence one can assume that negligible local indentation or damage is created during the bending process of the beams and that the displacement calculated from Eq. 2 is from the beam structure.

Elastic Modulus and Bending Strength

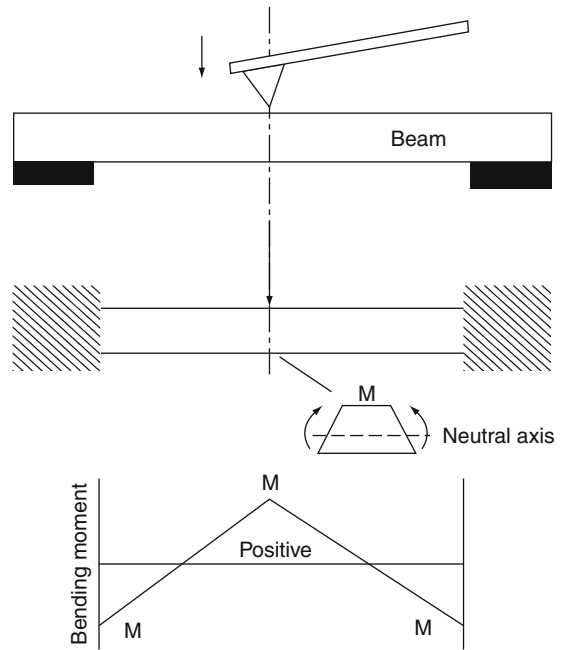
Elastic modulus and bending strength (fracture stress) of the beams can be estimated by equations based on the assumption that the beams follow the linear elastic theory of an isotropic material. This is probably valid since the beams have high length-to-width ℓ/w and length-to-thickness ℓ/t ratios and also since the length direction is along the principal stress direction during the test. For a fixed elastic beam loaded at the center of the span, the elastic modulus is expressed as

$$E = \frac{\ell^3}{192I} m \quad (5)$$

where ℓ is the beam length, I is the area moment of inertia for the beam cross section, and m is the slope of the linear region of the load–displacement curve during bending [46]. The area moment of inertia for a beam with a trapezoidal cross section is calculated from the following equation:

$$I = \frac{4w_1^2w_1w_2 + w_2^2}{36(w_1 + w_2)} t^3 \quad (6)$$

where w_1 and w_2 are the upper and lower widths, respectively, and t is the thickness of the beam. According to linear elastic theory, for a centrally loaded beam, the moment diagram is shown in Fig. 2. The maximum moments are generated at the ends



Nanomechanical Properties of Nanostructures, Fig. 2

A schematic of the bending moments generated in the beam during a quasi-static bending experiment, with the load at the center of the span. The maximum moments occur under the load and at the fixed ends. Due to the trapezoidal cross section, the maximum tensile bending stresses occur at the top surfaces at the fixed ends

(negative moment) and under the loading point (positive moment) as shown in Fig. 2. The bending stresses generated in the beam are proportional to the moments and are compressive or tensile about the neutral axis (line of zero stress). The maximum tensile stress (σ_b , which is the bending strength or fracture stress) is produced on the top surface at both the ends and is given by [46]

$$\sigma_b = \frac{F_{max}\ell e_1}{8I} \quad (7)$$

where F_{max} is the applied load at failure, and e_1 is the distance of the top surface from the neutral plane of the beam cross section and is given by [46]

$$e_1 = \frac{t(w_1 + 2w_2)}{3(w_1 + w_2)} \quad (8)$$

Although the moment value at the center of the beam is the same as at the ends, the tensile stresses at

the center (generated on the bottom surface) are less than those generated at the ends (per Eq. 7) because the distance from the neutral axis to the bottom surface is less than e_1 . This is because of the trapezoidal cross section of the beam, which results in the neutral axis being closer to the bottom surface than the top (Fig. 2).

In the preceding analysis, the beams were assumed to have fixed ends. However, in the nanobeams used by Sundararajan and Bhushan [41], the underside of the beams was pinned over some distance on either side of the span. Hence a finite element model of the beams was created to see if the difference in the boundary conditions affected the stresses and displacements of the beams. It was found that the difference in the stresses was less than 1%. This indicates that the boundary conditions near the ends of the actual beams are not that different from that of fixed ends. Therefore the bending strength values can be calculated from Eq. 7.

Fracture Toughness

Fracture toughness is another important parameter for brittle materials such as silicon. In the case of the nanobeam arrays, these are not best suited for fracture toughness measurements because they do not possess regions of uniform stress during bending. Sundararajan and Bhushan [41] developed a methodology, and its steps are outlined schematically in Fig. 3a. First, a crack of known geometry is introduced in the region of maximum tensile bending stress, i.e., on the top surface near the ends of the beam. This is achieved by generating a scratch at high normal load across the width (w_1) of the beam using a sharp diamond tip (radius < 100 nm). A typical scratch thus generated is shown in Fig. 3b. By bending the beam as shown, a stress concentration will be formed under the scratch. This will lead to failure of the beam under the scratch once a critical load (fracture load) is attained. The fracture load and relevant dimensions of the scratch are input into the FEM model, which is used to generate the fracture stress plots. Figure 3c shows an FEM simulation of one such experiment, which reveals that the maximum stress does occur under the scratch.

If one assumes that the scratch tip acts as a crack tip, a bending stress will tend to open the crack in Mode I. In this case, the stress field around the crack tip can be described by the stress intensity parameter K_I (for Mode I) for linear elastic materials [20]. In particular

the stresses corresponding to the bending stresses are described by

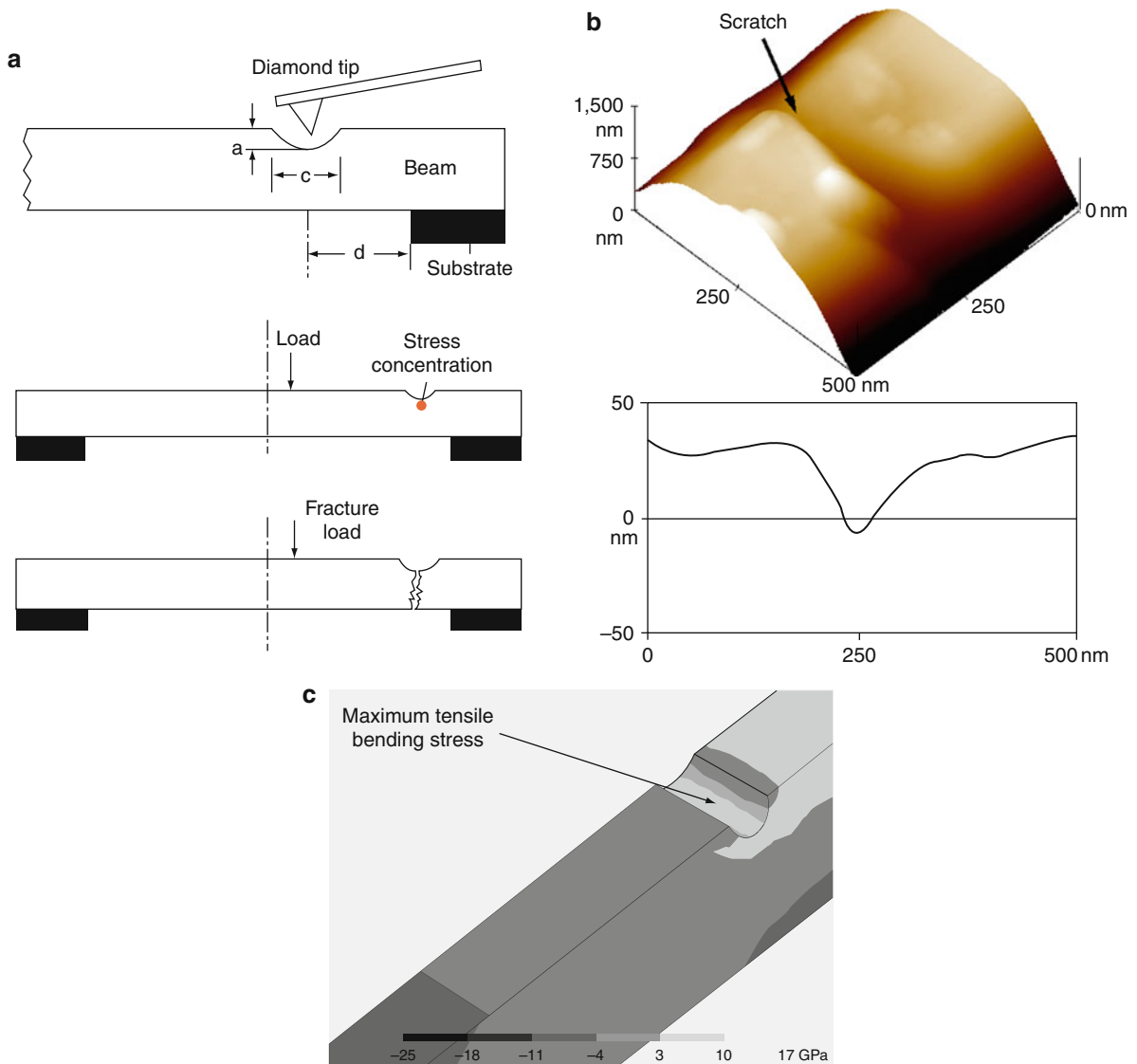
$$\sigma = \frac{K_I}{\sqrt{2\pi r}} \cos\left(\frac{\theta}{2}\right) \left[1 + \sin\left(\frac{\theta}{2}\right) \sin\left(\frac{3\theta}{2}\right) \right] \quad (9)$$

for every point $p(r, \theta)$ around the crack tip as shown in Fig. 4. If the fracture stress (σ_f) is substituted into the left hand side of Eq. 9, then the K_I value can be substituted by its critical value, which is the fracture toughness K_{IC} . Now, the fracture stress can be determined for the point ($r = 0, \theta = 0$), i.e., right under the crack tip as explained above. However, $r = 0$ cannot be substituted in Eq. 9. The alternative is to substitute a value for r which is as close to zero as possible. For silicon, a reasonable number is the distance between neighboring atoms in the (111) plane, the plane along which silicon exhibits the lowest fracture energy. This value was calculated from silicon unit cell dimensions of 0.5431 nm [1] to be 0.4 nm (half of the face diagonal). This assumes that Si displays no plastic zone around the crack tip, which is reasonable since in tension, silicon is not known to display much plastic deformation at room temperature. Sundararajan and Bhushan [41] used values $r = 0.4$ – 1.6 nm (i.e., distances up to 4 times the distance between the nearest neighboring atoms) to estimate the fracture toughness for both Si and SiO₂ according to the following equation:

$$K_{IC} = \sigma_f \sqrt{2\pi r} \quad r = 0.4 \text{ to } 1.6 \text{ nm} \quad (10)$$

Fatigue Strength

In addition to the properties mentioned so far that can be evaluated from quasi-static bending tests, the fatigue properties of nanostructures are also of interest. This is especially true for MEMS/NEMS involving vibrating structures such as oscillators and comb drives and hinges in digital micromirror devices. To study the fatigue properties of the nanobeams, Sundararajan and Bhushan [41] applied monotonic cyclic stresses using an AFM, Fig. 5a. Similar to the bending test, the diamond tip is first positioned at the center of the beam span. In order to ensure that the tip is always in contact with the beam (as opposed to impacting it), the piezo is first extended by a distance D_1 , which ensures a minimum stress on the beam. After this extension, a cyclic displacement of amplitude D_2 is applied

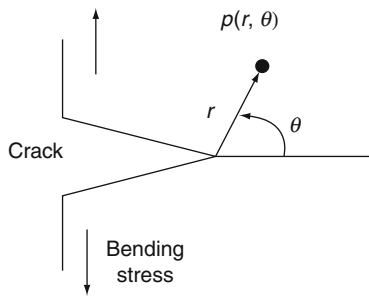


Nanomechanical Properties of Nanostructures, Fig. 3 (a) Schematic of technique to generate a defect (crack) of known dimensions in order to estimate fracture toughness. A diamond tip is used to generate a scratch across the width of the beam. When the beam is loaded as shown, a stress concentration is

formed at the bottom of the scratch. The fracture load is then used to evaluate the stresses using FEM. (b) AFM 3-D image and 2-D profile of a typical scratch. (c) Finite element model results verifying that the maximum bending stress occurs at the bottom of the scratch [41]

continuously until failure of the beam occurs. This results in the application of a cyclic load to the beam. The maximum frequency of the cyclic load that could be attained using the AFM by Sundararajan and Bhushan [41] was 4.2 Hz. The vertical deflection signal of the tip is monitored throughout the experiment. The signal follows the pattern of the piezo input up to failure, which is indicated by a sudden drop in the signal. During initial runs, piezo drift was observed

that caused the piezo to gradually move away from the beam (i.e., to retract), resulting in a continuous decrease in the applied normal load. In order to compensate for this, the piezo is given a finite extension of 75 nm every 300 s as shown in Fig. 5a. This results in keeping the applied loads fairly constant. The normal load variation (calculated from the vertical deflection signal) from a fatigue test is shown in Fig. 5b. The cyclic stress amplitudes (corresponding to D_2) and



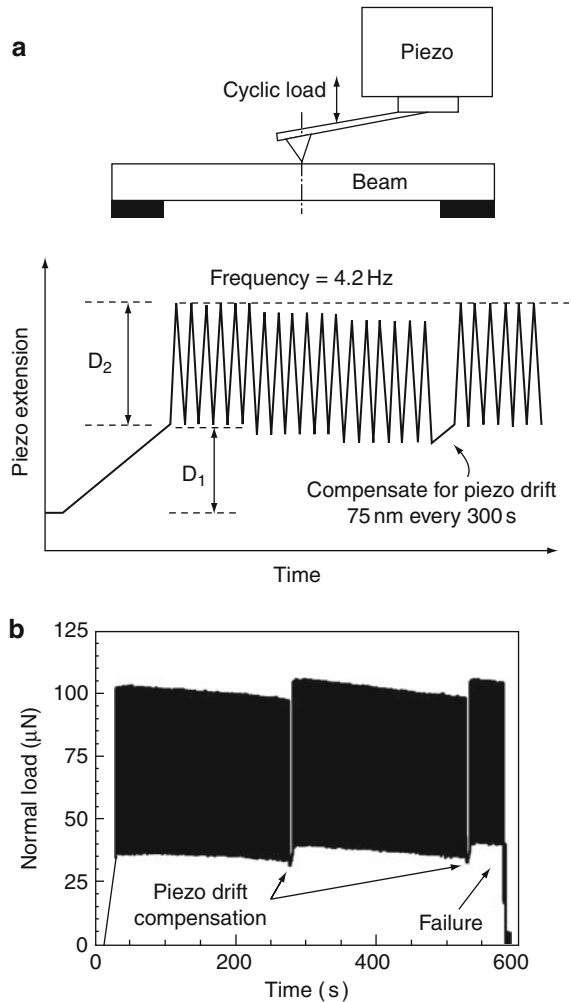
Nanomechanical Properties of Nanostructures, Fig. 4

Schematic of crack tip and coordinate systems used in Eq. 9 to describe a stress field around the crack tip in terms of the stress intensity parameter, K_I [41]

fatigue lives are recorded for every sample tested. Values for D_1 are set such that minimum stress levels are about 20% of the bending strengths.

Bending Tests of Micro/Nanostructures Using a Nanoindenter

Quasi-static bending tests of micro/nanostructures in both normal and lateral directions are carried out using a nanoindenter (Fig. 6) [27, 28, 35, 45]. The advantage of a nanoindenter is that loads up to about 400 mN, higher than that in AFM (up to about 100 μ N), can be used for structures requiring high loads for experiments. Bending experiments in the normal direction have been carried out on the suspended beams using the nanoindenter fitted with a conical tip that has a 1 μ m radius of curvature and a 60° included angle. Figure 6a shows the schematic of the beam bending experimental setup. To avoid the indenter tip pushing into the specimen, a blunt tip is used in the bending and fatigue tests. For ceramic and metallic beam samples, Li et al. [28] used a diamond conical indenter with a radius of 1 μ m and an included angle of 60°. For polymer beam samples, Wei et al. [45] and Palacio et al. [35] reported that the diamond tip penetrated the polymer beams easily and caused considerable plastic deformation during the bending test, which led to significant errors in the measurements. To avoid this issue, the diamond tip was dip-coated with PMMA (about 1–2 μ m thick) by dipping the tip in the 2% PMMA (wt/wt) solution for about 5 s. Load position used was at the center of the span for the bridge beams and at 10 μ m off from the free end of cantilever beams. An optical microscope with a magnification of 1500 x or an in situ AFM is used to locate the loading position.



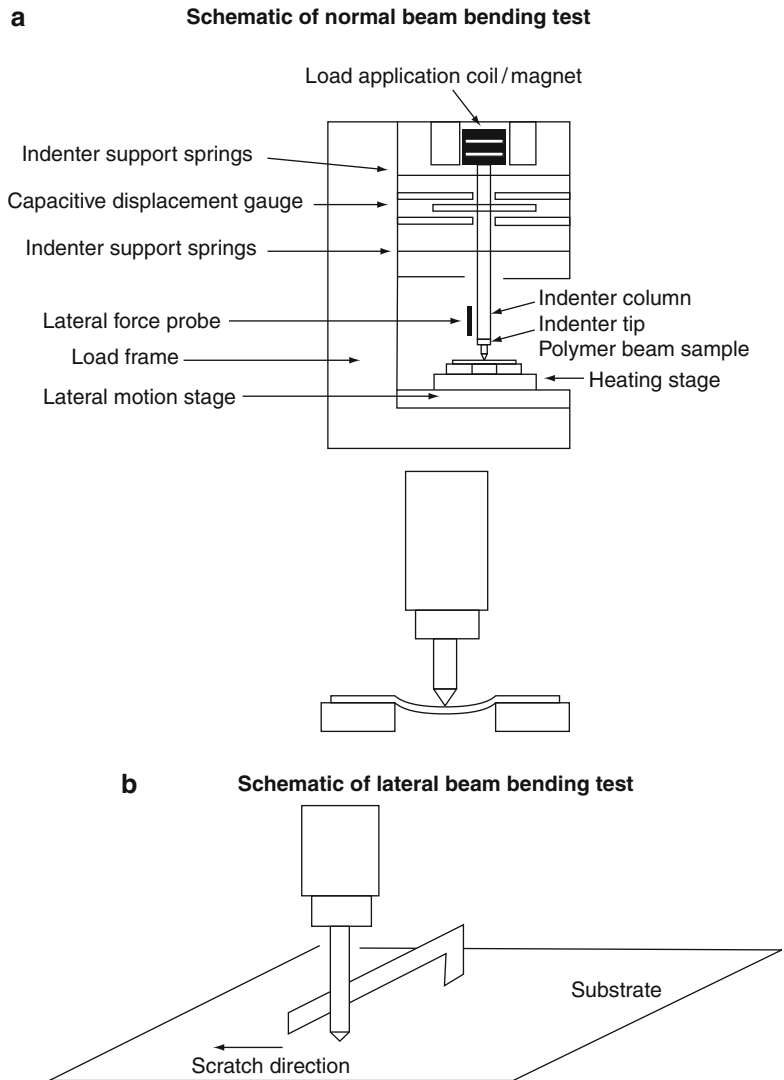
Nanomechanical Properties of Nanostructures, Fig. 5 (a)

Schematic showing the details of the technique to study fatigue behavior of the nanobeams. The diamond tip is located at the middle of the span and a cyclic load at 4.2 Hz is applied to the beam by forcing the piezo to move in the pattern shown. An extension is made every 300 s to compensate for the piezo drift to ensure that the load on the beam is kept fairly constant. (b) Data from a fatigue experiment on a nanobeam until failure. The normal load is computed from the raw vertical deflection signal. The compensations for piezo drift keep the load fairly constant [41]

Then the specimen is moved under the indenter location with a resolution of about 200 nm in the longitudinal direction and less than 100 nm in the lateral direction. Using the analysis presented earlier, elastic modulus and bending strength of the beams can be obtained from the load–displacement curves [28, 35, 45]. For fatigue tests, an oscillating load is applied and

Nanomechanical Properties of Nanostructures,

Fig. 6 Schematic of micro/nanoscale bending experimental setup using a nanoindenter (a) in the normal direction, and (b) in the lateral direction



contact stiffness is measured during the tests. A significant drop in the contact stiffness during the test is a measure of the number of cycles to failure [27].

A schematic of the lateral bending experimental setup is shown in Fig. 6b. The nanoindenter was fitted with a conical tip which has a $1\ \mu\text{m}$ radius of curvature and a 60° included angle. The tip was positioned at $200\text{--}300\ \mu\text{m}$ away from the beam anchor and $20\ \mu\text{m}$ from the edge of the beam length. The indenter was programmed to perform a scratch at a constant load of $400\ \mu\text{N}$, and the stage was moved at a rate of $10\ \mu\text{m/s}$. Since the lateral force contribution of the substrate surface is constant, any additional force recorded

comes from the cantilever beam response. From the linear elastic regime in the measured lateral force as a function of lateral displacement, the elastic modulus can be evaluated. For a cantilever beam with one end clamped, it is expressed as [46]

$$E = \frac{\ell^3}{3I}m \quad (11)$$

where, as described earlier, ℓ is the beam length, I the area moment of inertia for the beam cross section, and m is the slope of the linear region of the force-displacement curve.

Closure

Mechanical properties of nanostructures are necessary in designing realistic MEMS/NEMS and BioMEMS/BioNEMS devices. These can be measured using an AFM or nanoindenter. Most mechanical properties are scale dependent.

Cross-References

► Nanotechnology

References

- Anonymous: Properties of silicon. In: EMIS Datareviews Series No. 4, INSPEC, Institution of Electrical Engineers, London.
- Becker, H., Lacascio, L.E.: Polymer microfluidic devices. *Talanta* **56**, 267–287 (2002)
- Beebe, D.J., Mensing, G.A., Walker, G.M.: Physics and applications of microfluidics in biology. *Annu. Rev. Biomed. Eng.* **4**, 261–286 (2002)
- Bhushan, B.: *Tribology Issues and Opportunities in MEMS*. Academic, Dordrecht (1998)
- Bhushan, B.: *Handbook of Micro/Nanotribology*, 2nd edn. CRC Press, Boca Raton (1999)
- Bhushan, B.: *Principles and Applications of Tribology*. Wiley, New York (1999)
- Bhushan, B.: Macro- and microtribology of MEMS materials. In: Bhushan, B. (ed.) *Modern Tribology Handbook*, pp. 1515–1548. CRC Press, Boca Raton (2001)
- Bhushan, B.: *Introduction to Tribology*. Wiley, New York (2002)
- Bhushan, B.: *Springer Handbook of Nanotechnology*, 3rd edn. Springer, Heidelberg (2010)
- Bhushan, B.: *Nanotribology and Nanomechanics I – Measurement Techniques and Nanomechanics, II – Nanotribology, Biomimetics, and Industrial Applications*, 3rd edn. Springer, Heidelberg (2011)
- Bhushan, B., Li, X.: Nanomechanical characterization of solid surfaces and thin films. *Int. Mater. Rev.* **48**, 125–164 (2003)
- Cheng, J., Kricka, L.J. (eds.): *Biochip Technology*. Harwood Academic, Philadelphia (2001)
- Dresselhaus, M.S., Dresselhaus, G., Avouris, Ph: *Carbon Nanotubes – Synthesis, Structure, Properties and Applications*. Springer, Berlin (2001)
- Drexler, K.E.: *Nanosystems: Molecular Machinery, Manufacturing and Computation*. Wiley, New York (1992)
- Elwenspoek, M., Wiegerink, R.: *Mechanical Microsensors*. Springer, Berlin (2001)
- Fujimasa, I.: *Micromachines: A New Era in Mechanical Engineering*. Oxford University Press, Oxford, UK (1996)
- Gad-el-Hak, M.: *The MEMS Handbook*. CRC Press, Boca Raton (2002)
- Goddard, W.A., Brenner, D.W., Lyshevski, S.E., Iafate, G. J. (eds.): *Handbook of Nanoscience, Engineering, and Technology*. CRC Press, Boca Raton (2002)
- Heller, M.J., Guttman, A. (eds.): *Integrated Microfabricated Biodevices*. Marcel Dekker, New York (2001)
- Hertzberg, R.W.: *Deformation and Fracture Mechanics of Engineering Materials*, 3rd edn, pp. 277–278. Wiley, New York (1989)
- Hierlemann, A.: *Integrated Chemical Microsensor Systems in CMOS Technology*. Springer, Berlin (2005)
- Hsu, T.R.: *MEMS and Microsystems: Design and Manufacture*. McGraw-Hill, Boston (2002)
- Kovacs, G.T.A.: *Micromachined Transducers Sourcebook*. WCB McGraw-Hill, Boston (1998)
- Lai Poh San, C., Yap, E.P.H. (eds.): *Frontiers in Human Genetics*. World Scientific, Singapore (2001)
- Lanza, R.P., Langer, R., Vacanti, J. (eds.): *Principles of Tissue Engineering*. Academic, San Diego (2000)
- Lawn, B.R., Evans, A.G., Marshall, D.B.: Elastic/Plastic indentation damage in ceramics: the median/radial system. *J. Am. Ceram. Soc.* **63**, 574 (1980)
- Li, X., Bhushan, B.: Fatigue studies of nanoscale structures for MEMS/NEMS applications using nanoindentation techniques. *Surf. Coat. Technol.* **163–164**, 521–526 (2003)
- Li, X., Bhushan, B., Takashima, K., Baek, C.W., Kim, Y.K.: Mechanical characterization of micro/nanoscale structures for MEMS/NEMS applications using nanoindentation techniques. *Ultramicroscopy* **97**, 481–494 (2003)
- Madou, M.: *Fundamentals of Microfabrication: The Science of Miniaturization*, 2nd edn. CRC Press, Boca Raton (2002)
- Manz, A., Becker, H. (eds.): *Microsystem Technology in Chemistry and Life Sciences. Topics in Current Chemistry*, vol. 194. Springer, Heidelberg (1998)
- Mastrangelo, C. H., Becker, H. (eds.) *Microfluidics and BioMEMS*. In: *Proceedings of SPIE*, vol. 4560. SPIE, Bellingham, Washington (2001)
- Muller, R.S., Howe, R.T., Senturia, S.D., Smith, R.L., White, R.M.: *Microsensors*. IEEE, New York (1990)
- Nalwa, H.S. (ed.): *Nanostructures Materials and Nanotechnology*. Academic, San Diego (2002)
- Oeberg, P.A., Togawa, T., Spelman, F.A.: *Sensors in Medicine and Health Care*. Wiley, New York (2004)
- Palacio, M., Bhushan, B., Ferrell, N., Hansford, D.: Nanomechanical characterization of polymer beam structures for BioMEMS applications. *Sensors Actuators A* **135**, 637–650 (2007)
- Park, K. (ed.): *Controlled Drug Delivery: Challenges and Strategies*. American Chemical Society, Washington, DC (1997)
- Poole, C.P., Owens, F.J.: *Introduction to Nanotechnology*. Wiley, Hoboken (2003)
- Rietman, E.A.: *Molecular Engineering of Nanosystems*. Springer, New York (2001)
- Schulte, T.H., Bardell, R.L., Weigl, B.H.: Microfluidic technologies in clinical diagnostics. *Clin. Chim. Acta* **321**, 1–10 (2002)
- Senturia, S.D.: *Microsystem Design*. Academic, Boston (2000)
- Sundararajan, S., Bhushan, B.: Development of AFM-based techniques to measure mechanical properties of nanoscale structures. *Sensors Actuators A* **101**, 338–351 (2002)

42. Tang, W.C., Lee, A.P.: Defense applications of MEMS. *MRS Bull.* **26**, 318–319 (2001)
43. Timp, G. (ed.): *Nanotechnology*. Springer, New York (1999)
44. van der Berg, A. (ed.): *Lab-on-a-Chip: Chemistry in Miniaturized Synthesis and Analysis Systems*. Elsevier, Amsterdam (2003)
45. Wei, G., Bhushan, B., Ferrell, N., Hansford, D.: Microfabrication and nanomechanical characterization of polymer MEMS for biological applications. *J. Vac. Sci. Technol. A* **23**, 811–819 (2005)
46. Young, W.C., Budynas, R.G.: *Roark's Formulas for Stress and Strain*. McGraw-Hill, New York (2002)

Nanomechanical Resonant Sensors and Fluid Interactions

Rustom B. Bhiladvala
Department of Mechanical Engineering,
University of Victoria, Victoria, BC, Canada

Definition

Nanomechanical resonators (or nanoresonators, for this entry) are typically simple nanoscale objects such as rods, beams of rectangular cross-section, grids, or hollow tubes, set into mechanical vibration around a resonance frequency. This entry focuses on ambient fluid (gas or liquid) interaction with nanoresonator-based sensors.

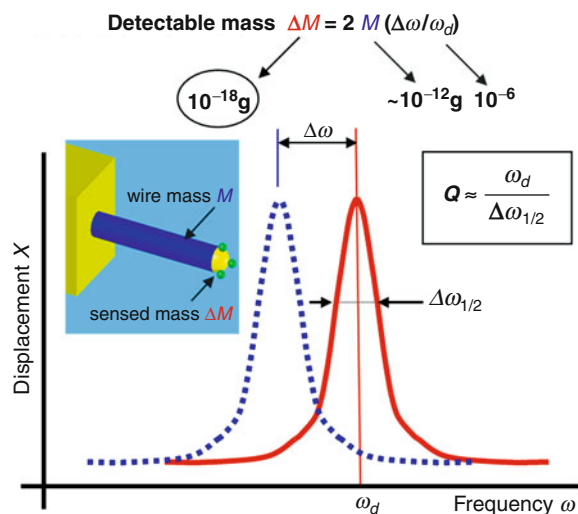
Introduction

Nanoresonators, often referred to as examples of NEMS (Nano Electro Mechanical Systems) devices, are of technological interest because of their potential for extraordinary sensitivity as mass and force measurement devices. To convey this sensitivity to the nonspecialist, a relevant unit for measurable added mass is a zeptogram ($1 \text{ zg} = 10^{-21} \text{ g}$), the mass of about 33 water molecules, or a billionth part of one trillionth of a typical 0.1 ml water droplet seen in a household faucet; a relevant unit for measurable force is an attonewton ($1 \text{ aN} = 10^{-18} \text{ N}$), equivalent to the weight of one millionth part of one billionth of such a droplet. In the last half-decade, published reports of zeptogram mass resolution [1] and

attonewton force resolution [2], using nanomechanical structures of submicron dimensions, now allow us to anticipate new roles for nanoresonators in healthcare, biological research, surface science, and several interdisciplinary areas across physics, chemistry, and engineering.

How has the scientific community moved to use the extraordinary capabilities of nanomechanical resonators to measure such small forces and mass? Perhaps the most widespread interest expressed has been in molecular diagnosis – screening for, or quantifying, small concentrations of biomarker molecules of devastating disease states such as cancer and neurodegeneration, with the goal of early intervention to improve prognosis – well before disease symptoms appear. Cantilevered tips of less than 100 nm width have been batch fabricated for metrology of biomolecules and Magnetic Resonance Force Microscopy (MRFM) with 4 nm resolution for imaging single viruses [2]. The role of mechanics in the intriguing and complex world of living cells and tissues is being investigated by cell stiffness and tissue elasticity measurements using small-scale resonators for force microscopy in liquids [3, 4]. In the physical sciences, the quest to correlate the structure and behavior of matter at small scales, occupies a significant fraction of research with nanoresonators. Examples are the quantum behavior of mesoscopic mechanical systems, the understanding of elastic and dissipative processes in solids with a large surface to volume ratio, the description of fluid-solid interaction at small scales, high frequencies and in particular, in gases at moderate rarefaction.

Two measured quantities of interest in nanoresonator-fluid interaction are the resonance peak frequency ω_d , and the quality factor or Q -factor, (symbol Q), as shown in Fig. 1. Nanowire (NW) resonators, tenths of a micron in diameter and a few microns long, typically have resonance frequencies in the MHz-range and small masses (pg), which allow attogram-level resolution of added mass. Measurement of damping force uses the quality factor, Q , as defined in Fig. 1. The Q -factor is the most relevant quantity for fluid-nanoresonator interaction studies. Why is nanoresonator interaction with fluids – liquids or gases – of interest? There are two points of view. The first view to emerge was that interaction with fluids is a nuisance in many nanoresonator applications and it is necessary to learn enough about it to reduce or correct



Nanomechanical Resonant Sensors and Fluid Interactions,

Fig. 1 Schematic of a cylindrical nanoresonator made by nanowire assembly. The left-shift $\Delta\omega$ of the damped resonance peak frequency ω_d , upon binding of mass, allows attogram-level mass resolution. The quality factor, Q , as defined, may be obtained from the spectrum by measuring the full width at half-power, $\Delta\omega_{1/2}$

for its effects. Resonance experiments showing great promise for sensor performance in a high-vacuum environment, quickly encountered the sobering reality of high fluid damping in air and water. Measurement attempts showed that high fluid damping could reduce accuracy of frequency peak location, decrease signal-to-noise ratio, and, in liquids, make resonance spectra difficult or impossible to detect. Efforts aimed at addressing these problems constitute the bulk of published research in fluid-nanoresonator interaction. The second and more recent view to emerge is that fluid interaction with nanoresonators is worthy of study in its own right. Such studies aim to answer questions of broad applicability in fluid mechanics of external oscillatory flows, in which viscous forces dominate over inertial forces (i.e., flows at low Reynolds numbers, Re). Examples are the material behavior of fluids at very high forcing frequencies, where the timescales can be small compared to relaxation time of the fluid, and flow in moderately rarefied gases, where the derivation of governing equations and boundary conditions to describe the flow are a subject of active research. The fluid mechanics of internal (channel) flows of liquids at micro and nanoscale is acquiring some maturity, with several books now available in

microfluidics and nanofluidics. The study of oscillatory flows around nanostructures is presently seeing a period of vigorous early growth in research publication.

This entry is organized in four sections. The section “Nanoresonator Applications Involving Fluid Interaction”, divides nanoresonator-fluid interaction applications into those in which fluid interaction is primarily a performance concern to be reckoned with or is proposed as a sensor principle and those applications where research toward formulating and answering questions in fluid mechanics is the primary goal. As the scope and limitations for both sets of applications is coupled strongly to choices of materials and transduction methods, this coupling is first introduced in the section “Interdisciplinary Development of Nanoresonator Sensor Platforms”, but plays a significant role throughout the article. The section “Measurement Relations” deals with the development of the simplest relations required to interpret measurements. The role of geometry, materials, and fabrication schemes in nanoresonator-fluid interaction measurement is discussed in detail in the section “Fabrication and Materials”, and a summary and a few reviews with complementary material are presented in the section “[Summary and Related Reviews](#)”.

Nanoresonator Applications Involving Fluid Interaction

While the potential field of nanoresonator applications is large, this entry is restricted to applications of nanomechanical resonators that involve fluid interaction. It is useful to divide these into two cases: (1) Nanoresonators for technological applications involving fluid interaction and (2) Nanoresonators as fluid damping sensors for research in fluid mechanics. In case (1), an accurate value of the Q -factor is not important, provided it is large enough ($Q \gg 1$) and does not change significantly in the course of measurement. A significant decrease in Q -factor as a result of increasing ambient gas pressure can reduce accuracy and signal-to-noise ratio and constitutes a decrease in performance in this case. In case (2), an accurate value of the Q -factor is required, as it constitutes the raw data for the measurement of fluid damping. In this case, a larger change in Q -factor for a given change in fluid damping is desirable, as this increases the dynamic range for fluid damping measurement.

Technological Applications Involving Fluid Interaction

The extraordinary mass and force sensitivity of nanoresonators and their size, commensurate with length scales of large macromolecules in living cells, makes them well suited to cellular and molecular studies. For microcantilevers, Lavrik et al. [5] provide a comprehensive review of studies using resonance mode detection, static deflection, or a combination of both for gas and vapor phase analyte sensing. Molecular diagnosis/screening can be enabled by functionalizing nanoresonators for specific binding of target biomarker molecules and using the frequency shift following target binding. Accurate measurements of small mass require a resonant spectrum with adequate signal-to-noise ratio and a sharp resonant peak (high quality factor) for the resonance curve. Since the aqueous environment is native to most cells and biomolecules, this would be most preferred for operating nanoresonator biosensors – but it is also the environment with highest fluid damping, which degrades the overall performance of the resonators. Demonstration of ideas such as the use of higher harmonics of the fundamental resonance frequency [6] or high pretension in nanobeams [7] for viable resonance in fluids are an established stream of research literature focused mainly on reducing the deleterious effects of fluid damping on sensor performance in measurement schemes. Use of the change in damping of a single nanoresonator or the resonance properties of a group of fluid-coupled nanoresonators by the binding of biomolecular analytes have also been proposed [8]. The development of scanned probe microscopies for mapping and understanding the structure of biomolecules is an important research area. Q -control techniques have been developed to offset the deleterious effects of fluid damping in scanned force microscopies.

Sensing for Fluid Mechanics Research

Nanoresonator interaction with gases of varying density can span rarefaction regimes from continuum flow to highly rarefied (free-molecular) flow. While both these limits and quantitative approaches to them are familiar, neither free-molecular nor standard continuum calculation methods are valid [9] in the intermediate transition regime of rarefied gas flow. NWs can serve as measuring instruments of gas damping forces from atmospheric pressure to

a modest vacuum $\sim 10^{-3}$ atm and spanning Knudsen numbers (gas molecule mean free path/wire diameter) in the range of $0.1 < Kn < 100$ which includes the transition and free-molecular flow regimes of rarefied gas dynamics. Neither the continuum fluid approach nor simple models for molecular momentum transport can be trivially extended to work in the transition regime – the development of theoretical and computational models in this regime is currently a research area of significant interest [9–11]. Gas damping force measurements with nanowires can provide data for these models.

A theoretical quantitative description allowing accurate calculation in the transition regime must blend smoothly into known solutions for the free-molecular and continuum regimes. For fluid which may be regarded as a continuum, what is the effect of forcing at high frequencies, or times small compared to any relaxation (i.e., adjustment) time within the fluid? For a smooth blend at the continuum end, the large span in frequencies (and Reynolds numbers, Re) can require either steady or unsteady forms of the Stokes equation, and for a few exceptional high-frequency cases, the usually appropriate neglect of inertial effects represented by the nonlinear term in the Navier-Stokes equations is not valid (i.e., $Re > 1$). How should the physics of unsteadiness and inertia be modeled in the transition regime of rarefaction? How should transition regime calculations for geometries more complex than beams be approached? What are the appropriate length scales to be used in determining the boundaries of rarefaction regimes? How do the results scale with Mach number (nanoresonator velocity/gas molecule velocity)? Fundamental insights into fluid mechanics at small scales and high frequencies can be gained by comparing results from proposed models with those from tabletop experiments using oscillatory nanoresonator flows.

While early experimental demonstrations of viable micro or nanoresonator operation in fluids [5, 6] have provided some data suitable for a study of resonator-fluid interaction, experimental studies focused solely on understanding fluid interaction are now beginning to emerge [12, 13]. Better answers to some of the questions above are being proposed, as members of computational, analytical, and experimental subcommunities undertake collaborative work [10, 14] to improve understanding of fluid mechanics at small scales, high frequencies, and in regimes of gas

rarefaction that are not yet well described. These fundamental studies aim to create the analytical and computational framework, which in turn will enable a diverse range of useful applications. Such applications range from the very small – for example, sensitive nanoscale instruments for understanding biology and the physical sciences, building sensors for healthcare, engineering and technology – to the macroscale – for example, drag force and trajectory calculations of space vehicle or satellite fragments following atmospheric entry.

The potential to create extraordinarily sensitive instruments from nanoresonators is easy to see in principle, but the path to realization of such sensor platforms is paved with challenges. Some of these challenges are introduced and explained with an example in the following section.

Interdisciplinary Development of Nanoresonator Sensor Platforms

The challenge in measurements is that the high surface to volume ratio of nanostructures can cause an unusual and strong coupling of physical effects. Sensor platform development must typically reckon with such coupling between transduction (vibration detection and driving) schemes, intrinsic stresses from fabrication, and nonlinear geometric effects [15]. While the focus here is on nanoresonators and fluid mechanics, note that the actual conduct of work for technological applications requires making several choices in other areas that are strongly coupled and cannot be ignored.

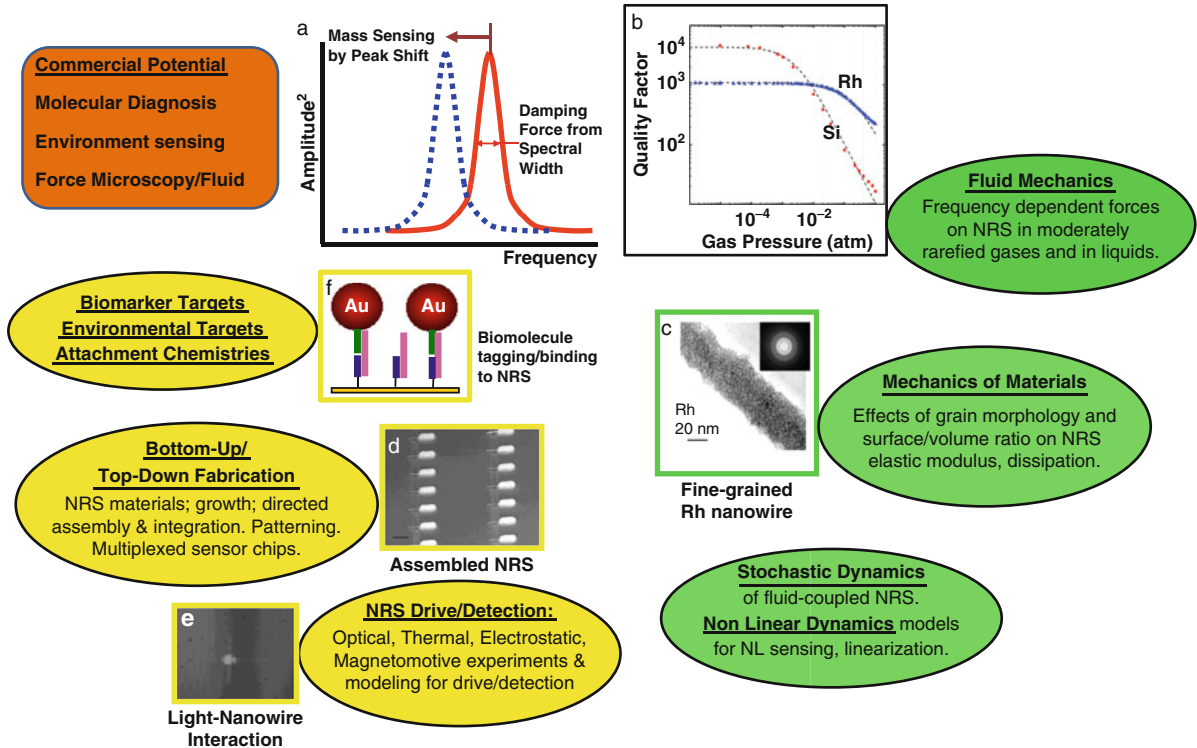
A representative view of these areas is shown in Fig. 2. A few partially realized examples of commercial potential are listed at the top left of the figure. Experimental research areas for engineering and technological development are shown in the three ellipses on the left. What are the sensor targets – for example, which gases, biological macromolecules, or cell components? How can attachment chemistry be developed to help functionalize a nanomechanical resonator (NR) – that is, to give it surface properties that will enable specific capture of only the selected target(s)? The attachment chemistry is dependent on the choice of NR material. The material choice may be limited by the fabrication approach used. Choice of the fabrication approach, in turn, may alter the ability for creating multiplexed sensors (capable of sensing different targets on the

same chip). What schemes are suitable for driving and detecting resonance for these small high-frequency devices? Each method listed in Fig. 2 is associated with advantages, and frequently, with measurement artifacts. Experiments and theoretical or computational analysis are often required to quantify the artifacts. Fundamental understanding in the disciplines of Fluid Mechanics, Mechanics of Materials, and Nonlinear Dynamical Systems (right half of figure) guides this engineering and technological work. Many aspects of this understanding need to be improved. A strong application for nanoresonators in fluids is thus to help improve the framework of areas of Fluid Mechanics that are not well understood. For example, what happens when a shear disturbance (vorticity) is introduced in a fluid at time scales far smaller than the time taken for viscosity to smoothen it out (viscous relaxation time)? Nanoresonator-fluid experiments are especially suitable for this question – they can provide very small forcing timescales (very high frequencies) and are typically associated with small fluid inertia (low Reynolds number) so that the nonlinear inertial effects are associated negligible.

As an example of iterative work due to interdisciplinary coupling shown in Fig. 2, consider a long, thin, doubly-clamped resonant nanobeam made from a metal film, with high intrinsic stress from a top-down fabrication process. Suppose the metal was chosen because the attachment chemistry for molecules enabling specific capture of a desired detection target molecule is well understood. During optical measurement of the motion of the doubly-clamped nanobeam in a high damping gas environment, one may turn up the power of the measuring laser to improve signal strength at low vibration amplitude. Experiments and modeling (nonlinear dynamics, mechanics of materials) have shown this can cause periodic heating of the vibrating beam and a thermomechanical coupling which has been shown to introduce softening and hardening nonlinearity as a function of laser power [15]. To circumvent this problem arising from intrinsic axial stress in the metal film, a similar long, thin, cantilevered metal beam (no axial stress) will be a necessary replacement. Such beams are difficult to make with a top-down technique, spurring efforts in the use of alternate fabrication techniques to make such structures.

This example shows that the conduct of work as well as the correct interpretation of data needs

NANOMECHANICAL RESONANT SENSORS (NRS) & FLUIDS



Nanomechanical Resonant Sensors and Fluid Interactions, Fig. 2 One view of the scope of research for nanomechanical resonant sensors (NRS) involving fluid interaction. Commercial potential (*top left*) and experimental methods

development (three ellipses at *left, yellow*) have served as drivers. Fields of engineering science which support and are enriched by such research, are shown on the right (three ellipses, *green*). (c), (d) reproduced with permission, from Refs. [25], [22]

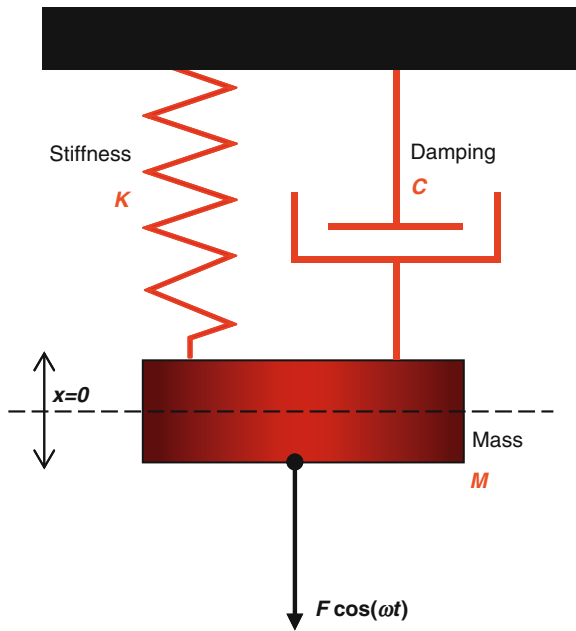
continual interaction between different disciplines to sort out artifacts arising from the coupling of physical effects, materials, fabrication schemes, and transduction methods. It is important to be mindful of this coupling while approaching the following section of relations used in the design of experiments and the interpretation of data.

Measurement Relations

This section presents different definitions of the Q -factor (Fig. 1) based on material properties, energy, and on spectral or time-domain measurements of resonator amplitude. The distinctions between definitions are usually and justifiably ignored if the Q -factor is very large compared to unity, but can be significant for fluid-nanoresonator interaction with high damping.

These relations and the relations for the resonance frequency also provide a scaling guide for geometry and material choice to tune the range of Q -factor for a given range of fluid damping.

When the modes of vibration are sufficiently separated, a single degree of freedom, harmonically driven, linear resonator model is reasonable for understanding key features of fluid-nanoresonator interaction. The damped resonance peak frequency ω_d and the dimensionless quality factor (Q -factor, or Q), which allows the measurement of dissipation through the sharpness of the resonance peak, are the most widely used quantities. Most instruments based on resonators avail of simplifications that are possible when the quality factor is large ($Q \gg 1$). High fluid damping produces systems with low Q -factor, leading to errors from the use of these simplifications. It is helpful to gather the concepts and relations necessary to maintain the accuracy of calculations for resonators where Q



Nanomechanical Resonant Sensors and Fluid Interactions, Fig. 3 Single degree of freedom model for resonator

may not be much larger than 1. A summary is included later in this section.

The displacement with time, $\{x(t)\}$, of a resonator with mass M , damping C , and stiffness K , driven harmonically with a forcing amplitude F_0 at a frequency ω , as seen in Fig. 3, is

$$M\ddot{x} + C\dot{x} + Kx = F_0 \cos(\omega t) \quad (1)$$

A surrogate form of the Q -factor, defined using material properties alone, is the dimensionless ratio Q_p , given by $Q_p = \sqrt{MK}/C = M\omega_0/C$, where $\omega_0 = \sqrt{K/M}$ is the undamped resonance frequency. Q_p is related to the commonly used damping ratio ζ , through $\zeta = C/C_c = 1/(2Q_p)$, where $C_c = 2\sqrt{MK}$ is the critical value of damping. Damping decreases the resonance frequency ω_d , from the undamped resonance value ω_0 , as given by the relation,

$$\omega_d = \omega_0 \left[1 - 1/(2Q_p^2) \right]^{\frac{1}{2}} = \omega_0 \left[1 - 2\zeta^2 \right]^{\frac{1}{2}}. \quad (2)$$

For steady-state resonance, the mean power supplied by the driving force balances the energy dissipation rate, and the total stored kinetic energy and spring potential energy, averaged over many cycles,

does not change with time. With displacement amplitude x_0 , the cycle-averaged value for the displacement-squared, $\langle x^2 \rangle = (x_0^2/2)$ and velocity-squared $\langle \dot{x}^2 \rangle = (\omega^2 x_0^2/2)$. The mean power dissipated, $\langle P \rangle$, the average energy dissipated per cycle $\langle E_{dissip} \rangle$ and the stored energy $\langle E_{stored} \rangle$ are related [16] by

$$\begin{aligned} \langle E_{dissip} \rangle &= \langle P \rangle (2\pi/\omega) = C \langle \dot{x}^2 \rangle \left(\frac{2\pi}{\omega} \right) \\ &= \pi C \omega x_0^2 \end{aligned} \quad (3a)$$

$$\langle E_{stored} \rangle = M(\omega^2 + \omega_0^2) x_0^2/4. \quad (3b)$$

If the average energy stored in the resonator is very large compared to the energy dissipated (= energy supplied, for steady-state) per cycle, one may anticipate lower noise operation when driven, and a slower decay of the oscillation amplitude envelope when the driving force is switched off, allowing for a time-domain measurement of damping. These are measures of the quality of a resonator, expressed by the magnitude of the quality factor, defined [16] as

$$Q = 2\pi \frac{\langle E_{stored} \rangle}{\langle E_{dissip} \rangle} \quad (4)$$

Using Eqs. 3a and 3b, this yields

$$Q = \frac{M(\omega_0^2 + \omega^2)}{2C\omega} = \frac{K + M\omega^2}{2C\omega}. \quad (5)$$

For forcing at the damped resonance frequency ω_d , using Eq. 5 with the definition for Q_p , $\omega = \omega_d$, with ω_d given by Eq. 2, the expression for the quality factor reduces to

$$Q = \frac{Q_p^2 - (1/4)}{\sqrt{Q_p^2 - (1/2)}}. \quad (6)$$

Most applications aim for quality factors $Q \gg 1$, for which ω_d reduces to ω_0 and Q to Q_p , as seen from Eqs. 2 and 6. It is also easy to show [16] that for high Q -factors, the measured resonance spectrum, expressed as a plot of amplitude-squared against forcing frequency, has a full width at half maximum (FWHM) height, as seen in Fig. 1, equal to damping

Nanomechanical Resonant Sensors and Fluid Interactions,**Table 1** Significant deviations at low Q_p

Q_p	ζ	$(Q - Q_p)/Q_p$	$(\omega_d - \omega_0)/\omega_0$
1,000	0.0005	3.12E-14	-2.50E-07
100	0.0050	3.13E-10	-2.50E-05
10	0.0500	3.14E-06	-2.50E-03
5	0.1000	5.10E-05	-1.01E-02
2	0.2500	2.22E-03	-6.46E-02
1	0.5000	6.07E-02	-2.93E-01

coefficient divided by mass, C/M , which allows for a simple measurement of damping.

The ringdown method is an alternative to the spectral technique above. It is a time-domain technique for measuring damping. In this method, the forcing for a driven nanoresonator is turned off, and the amplitude of the oscillation begins to decay with time t as $\exp(-t/\tau)$, where the time constant $\tau = C/M = \omega_0/Q_p$ [16]. Q_p can be obtained from τ , the time taken for the measured amplitude envelope to reduce to $1/e$ of the previous value.

Nanoresonator operation in a fluid environment is typically associated with high damping (C). Some consequences of low Q_p due to high damping may now be examined. Low Q -factor spectra, $Q \sim 1$, have been encountered in measurement attempts in liquids. For low Q -spectra, the spectral width (FWHM) may be hard to obtain and the spectra are noisier – the signal-to-noise ratio, $\text{SNR} \sim Q^{1/2}$. Using the best-fit Lorentzian curve solution of Eq. 1 to low- Q experimental spectra is preferable for the evaluation of the damping. As seen in Table 1, obtained using Eqs. 2 and 6, the deviations of Q from Q_p and ω_d from ω_0 , become significant and increase rapidly as Q_p approaches a value of 1. As an example, the fractional shift of the resonance peak frequency due to damping, changes from 0.25 ppm (parts per million) at Q_p of 1,000 to 65,000 ppm at Q_p of 2. Sensitive mass detection aims for resolution of frequency shifts of the order of 10 ppm or less, between mass-loaded and unloaded measurement states. For $Q_p \sim 1$, small variations in fluid damping between these states could themselves cause frequency shifts that will exceed the frequency shift from small added mass. The experimental evaluation of the Q -factor (sketch in Fig. 1) shows that the accuracy of any sensing scheme dependent on location of the peak frequency will also degrade significantly for broad peaks (low Q). In liquids and dense gases,

there are cases where fluid damping can be high enough to prevent the detection of the resonance peak altogether. Ringdown measurements also become less accurate or impossible, if Q is low enough to reduce the displacement amplitude to zero in a time comparable to one oscillation time period. Finally, when an electrostatic or piezoelectric drive is used for exciting and driving the nanostructure through resonance, the large range of swept frequency required to capture low Q -factor spectra, has the potential to create nonconstant forcing F_0 . This arises due to changes in the amplitude of the driving signal, due to significant frequency-dependent changes in the impedance of the driving circuit (parasitics).

Techniques for enhancing the Q -factor Q_p of nanoscale resonators in fluids could include the use of resonance at higher frequency harmonics (higher ω_0), and an increase of effective string-like stiffness (higher K , ω_0), by using tensioned beams [7]. These have demonstrated viable resonance with high quality factor in some fluid measurements, and should prove useful for many applications. However, their reduced sensitivity to fluid damping makes them less suitable as measurement devices for well-resolved and accurate measurement of the value of fluid damping, when that is the object of the study.

To complete this section, the experimental steps used for the evaluation of gas damping force on nanomechanical resonators are reviewed. A few results from mechanics of materials and vibration of continuous beams will be useful to see the interplay of geometry, material characteristics, and measurement artifacts related to nanoresonator fabrication and transduction schemes that are characteristic of nanoresonator measurements.

Consider fluid interactions with the nanoresonator that are linear – that is, they are not functions of products of the displacement x or its time derivatives. The nanoresonator motion may then be represented by the equation governing the displacement x ,

$$(M_s + M_f)\ddot{x} + (C_s + C_f)\dot{x} + (K_s + K_f)x = F_0 \cos(\omega t). \quad (7)$$

In Eq. 7, subscripts s, f denote structural intrinsic (solid material) and fluid quantities respectively. The fluid force on the beam in phase with the velocity is represented using fluid damping C_f ; the force on the

beam due to the fluid in phase with the acceleration can be significant in liquids, and is represented by the fluid added mass term M_f ; a restoring spring force component is represented by a fluid stiffness term K_f . A measurement at high vacuum, where fluid molecules have no appreciable effect, will involve only the intrinsic quantities. In nanoresonators, Eq. 2, with typical $Q_{p,s} \gg 1$, based on intrinsic damping C_s , shows that C_s is not large enough to cause an appreciable shift of the resonance peak frequency from the undamped value.

The undamped resonance frequency for transverse flexural vibrations of a solid beam having l, M_s, K_s, ρ_s, E , and I as the length, mass, stiffness, density, Young's modulus, and cross-sectional moment of inertia, is [17]

$$\omega_s \approx \omega_0 = \sqrt{K_s/M_s} = \beta_n^2 \sqrt{EI/M_s l^3} \quad (8)$$

$$= \beta_n^2 \sqrt{(1/12)^* (E/\rho_s)^* (t/l^2)} \quad (8a)$$

$$= \beta_n^2 \sqrt{(1/16)^* (E/\rho_s)^* (d/l^2)} \quad (8b)$$

where Eq. 8a is for a rectangular beam of width w , and thickness t parallel to the direction of motion and Eq. 8b is for a circular section beam (nanowire) of diameter d .

For rigid clamping, β_0 is 1.88 for cantilever beams and 4.73 for doubly-clamped beams vibrating in the fundamental mode ($n = 0$). The surrogate Q -factor for the beam in vacuum is given by,

$$Q_{p,s} = \sqrt{M_s K_s / C_s} = M_s \omega_0 / C_s = (1/C_s)^* \beta_n^2 \sqrt{M_s EI / l^3} \quad (9)$$

$$= (1/C_s)^* \beta_n^2 \sqrt{(E \rho_s / 12)^* (w t^2 / l)} \quad (9a)$$

$$= (1/C_s)^* \beta_n^2 \sqrt{(E \rho_s / 16)^* (\pi d^3 / 4l)} \quad (9b)$$

with Eq. 9a for the rectangular beam and (Eq. 9b) for the circular section nanowire. If squeeze-film damping is negligible, the effective stiffness in the fluid remains unchanged from its intrinsic value in vacuum, so that $K_{s+f} = K_s = M_s \omega_0^2$. Substituting for effective mass and damping, $M_{s+f} = M_s + M_f$ and $C_{s+f} = C_s + C_f$, yields the damped resonance frequency ω_{s+f} and quality factor $Q_{p,s+f}$ for the beam in fluid,

$$\omega_{s+f} = \omega_0 \sqrt{1 - 1/(2Q_{p,s+f}^2)} \quad (10)$$

$$Q_{p,s+f} = \sqrt{(M_{s+f} K_s) / C_{s+f}} \\ = Q_{p,s} \left[\frac{\{1 + (M_f/M_s)\}^{0.5}}{1 + (C_f/C_s)} \right] \quad (11)$$

For rarefied gas damping experiments with nanoresonators, the situation is usually simplified with fluid added mass M_f being negligible. A first measurement may be made at high vacuum, with no gas damping ($C_f = 0$), for which the Q -factor value, Q_s , depends only on the structural properties of the solid nanowire, M_s, C_s , and K_s . Following this, the ambient gas pressure in the chamber is increased in steps, and the frequency ω_{s+f} and quality factor Q_{s+f} are measured at each step. In experiments, we assume the surrogate property-related Q -factor (Q_p) relations hold true for the measured Q -factor values. This yields

$$Q_s = \frac{M_s \omega_0}{C_s} \quad (12a)$$

$$Q_{s+f} = \frac{M_s \omega_0}{C_s + C_f}. \quad (12b)$$

Using Eq. 12a and mass M_s from known density and geometry of nanoresonator, the value of intrinsic damping C_s may be obtained from measurement at high vacuum. Subsequent measurements of the Q -factor Q_{s+f} , at each value of ambient gas pressure, are related using Eqs. 12a and 12b to obtain

$$Q_{s+f} = Q_s \left(1 + \frac{C_f}{C_s} \right)^{-1}, \quad (13)$$

from which the gas damping C_f can be calculated. For liquids or dense gases, the added fluid mass M_f is not negligible, and the full Eq. 11 must be considered (see [9]), in place of its subset Eq. 13.

These relations provide a summary for the conduct of experimental work and are useful in a number of ways. For measurements in vacuum, they establish the expected scaling of frequency and Q -factor with parameters that depend on geometry, material, mode shape, and type of clamping. They allow evaluation of the rigidity and repeatability of the clamps (through the value of mode constant β_n in Eq. 9) if the geometric parameters and material properties are accurately known. Most importantly, they allow experimental

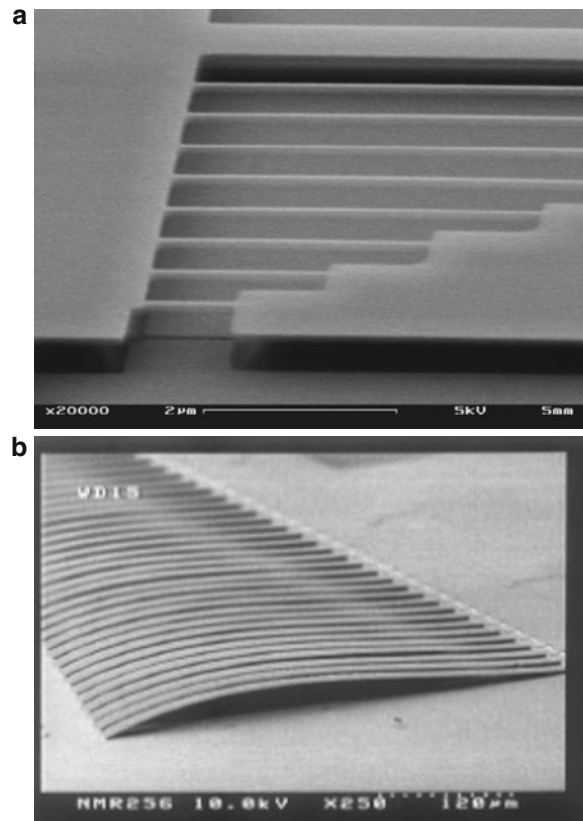
testing of models for fluid damping C_f and fluid added mass M_f . These relations also show how nanoresonator materials may be chosen to increase or decrease the Q -factor range for a given range of fluid damping. This is discussed in the following section detailing the coupling of fabrication, materials, and transduction choices.

Fabrication and Materials

Nanomechanical resonators are suspended structures – most often long and prismatic, clamped at one end (cantilevered) or both ends (doubly-clamped). How does the choice of geometry, material, fabrication, and transduction schemes influence the operation of nanomechanical resonant sensors in fluids? Earlier discussion of Fig. 2 showed the strong coupling between these choices and the necessity of using the framework of mechanics of fluids, mechanics of materials, and nonlinear/stochastic dynamics as guides to make sensor design choices. Fabrication choice plays a dominant role and it is worth briefly reviewing two fabrication paradigms – termed top-down and bottom-up fabrication and an example of the synergy that may be obtained by hybrid methods combining features of each.

Top-down Methods

Top-down methods are named for the processes of patterning and etching from larger planar sheets of material (top-level) to obtain smaller desired structures (bottom-level), typically as large arrays. No assembly is required as the structures are built in the desired locations. Typically, a sacrificial layer is sandwiched between the substrate and a film of the intended nanoresonator material (device layer). An etch mask, usually made from a photosensitive chemical that resists etching (photoresist) is used while etching through the device layer, leaving it in the shape of the desired resonator array. A second etch targeted at the sacrificial layer material is then used to remove material beneath the devices and create suspended resonators. Doubly-clamped resonators created by this method are shown in Fig. 4. Optical lithography allows transferring micrometer scale features on to large (upto centimeter-scale) areas of photoresist for patterning thousands of devices using a single exposure through a patterned mask. This well-established technique, long used in

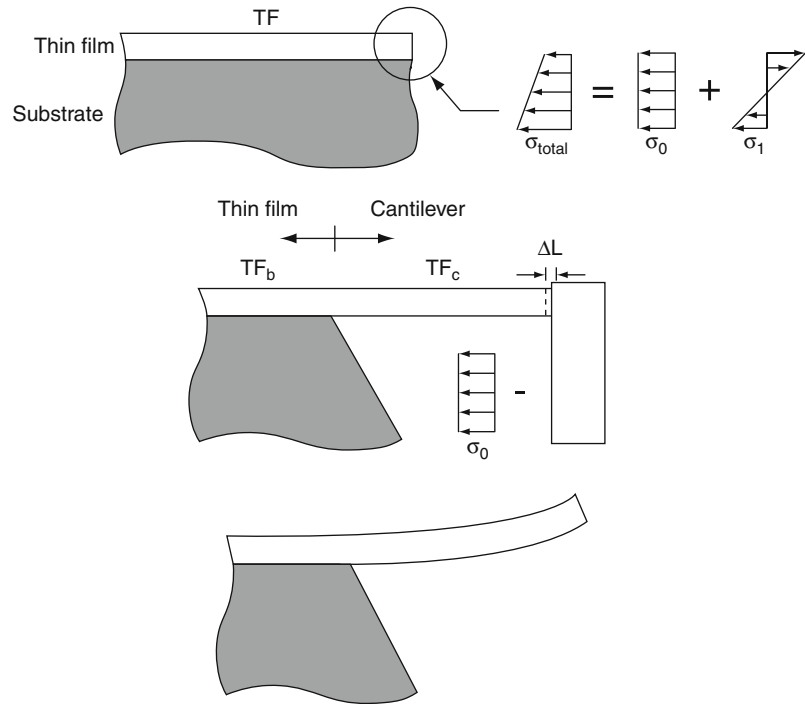


Nanomechanical Resonant Sensors and Fluid Interactions, Fig. 4 (a) Doubly-clamped nanoresonator beams of different lengths with the same undercut. Figure reprinted from Ref. [18], with permission. (b) Distortion of long thin metal cantilevers attempted using a top-down process (Figure reprinted from Ref. [19], with permission)

microelectronics to make integrated circuits (ICs), is easy to implement, and is almost always preferred at microscale. Minimum feature sizes for ICs, limited by the wavelength of light used, have been extended to submicron scale with deep-UV light sources and phase-shift masks. Nanoresonators, involving smaller features, from tens to hundreds of nanometers, typically have been made by electron-beam (e-beam) lithography. Here, no mask is used, but an electron beam is rastered across a film of e-beam resist to create the pattern for nanoscale structures. Time-consuming coverage of large areas required to make commercial sensor arrays, the cost of procuring and operating e-beam equipment, and the fabrication sensitivity to small perturbations do not allow inexpensive production of large-area sensor arrays by e-beam lithography. An alternative pattern transfer technique of imprinting,

Nanomechanical Resonant Sensors and Fluid Interactions, Fig. 5

Film stress and stress gradients create residual axial stresses in doubly-clamped beams and distort cantilevered beams, during their release from the substrate, by etching (Figure reprinted from Ref. [20], with permission)



molding, and stamping known as microcontact printing, or soft lithography, is similar to the use of an inked rubber stamp. This technique, somewhat surprisingly, has been used to demonstrate the ability to transfer pattern features below 100 nm in size. The master pattern in this case may be made by e-beam lithography and transferred to one or more soft “stamp pads,” thus reducing the requirements for e-beam lithography, with the promise of lowered costs. Mature techniques, relatively well-controlled shape, size and position, the ability to make devices with well-ordered internal structure through the use of single-crystal layers, are important strengths of top-down fabrication [18].

There are limitations of top-down fabrication for nanoresonator sensor arrays. Besides the cost of large-area coverage, the limited choice of local material selection (e.g., difficult to have resonator material different from clamp material) can restrict functionality. Geometries such as smooth rods or hollow tubes, with a high cross-sectional moment of inertia allowing for a higher stiffness per unit mass, are prohibitively difficult to produce in large numbers by top-down techniques. Undercuts, or significant underetching of supports during the removal of the sacrificial layer [18] are seen as a lighter-gray region around the periphery of the harp-like structure of Fig. 4a. Undercuts depend

on the sacrificial layer etch time and are of the same depth for all structures released on a chip. They can cause significant increase in the compliance, particularly for short, stiff, nanoresonators seen in Fig. 4a. Energy losses and cross talk between devices through undercuts can cause significant changes in frequencies and Q -factors. During release by etching of the sacrificial layer, shape distortion (see Fig. 4b) in the case of cantilevered structures [19] or residual axial stress in doubly-clamped nanoresonators (see Fig. 5) typically arises due to stress gradients in the device layer [20]. For long thin metal film structures, these effects are particularly difficult to control. Higher residual axial stress adds an increasing string-like component to the restoring force (normally flexural), changing the simple frequency scaling presented in Eq. 9 and makes the nanoresonator more prone to geometric stiffness nonlinearities. High axial stress and string stiffness can, however, be used to increase the effective stiffness K and the peak frequency, to enable a higher Q -factor value for given damping, provided a sufficient linear range exists. Distorted cantilevers also cannot be described by Eq. 9 – furthermore, it may be difficult or impossible to drive and detect resonance in such structures. Film stress gradient effects can be controlled in device materials such as single-crystal

silicon, silicon nitride or carbide, using silicon oxide as a sacrificial layer. These materials are frequently used for laboratory experiments. However, stress gradient effects (Fig. 5) are very pronounced for metal films. It is therefore difficult for top-down methods to yield long, thin metal cantilevers which are free of distortion – a “convex-up” example [19] of such distortion for titanium cantilevers is seen in Fig. 4b. Though long thin cantilevered resonators made of metals may have desirable surface functionalization attributes, they are difficult to achieve using top-down methods. Finally, for multiplexing, or the sensing of multiple analytes on the same chip, functionalization of different devices to capture different targets is tedious, and one may anticipate the value of nanostructures that can be made and functionalized off-chip, then directed to predetermined areas of a chip and integrated to form nanoresonators.

While top-down methods remain the primary choice for most microcantilever fabrication, in spite of their obvious strengths, many of the shortcomings of established top-down techniques discussed above have led to an interest in alternatives that bottom-up fabrication could provide. This is particularly true for nanoscale resonators, when long thin cantilevers made of metals are required for their attachment chemistries, or when undercut reduction or multiplexed sensor arrays are important.

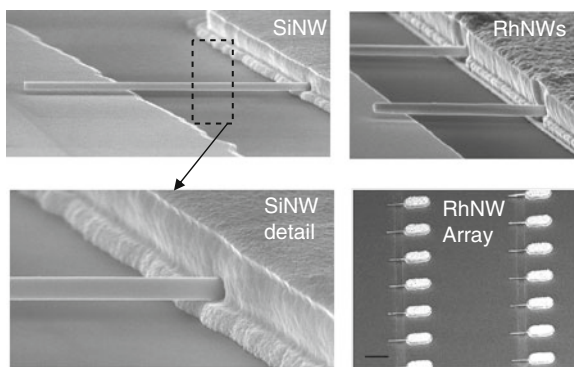
Bottom-Up Methods

Bottom-up methods are identified by their use of assembly of smaller units into larger structures, for example, single nanoresonator elements assembled and integrated into arrays for sensing. While the possibilities for improved control of individual element geometry and clamping are attractive, making arrays of a few tens of thousands of elements by an individual pick and place technique is unfeasible. Bottom-up methods are thus defined by the use of physical or chemical forces for self-assembly or directed assembly of nanostructures. A recent review of the directed assembly of nanowires [21] discusses assembly by molecular forces, electrostatic interactions, shear forces, magnetic fields, and dielectrophoresis. A significant advantage over top-down methods is that special vapor or solution phase synthetic chemistry can be used to tailor the shape, composition, and crystallinity of nanoscale elements, before their assembly and integration on a sensor chip. This greatly widens the scope of material properties (e.g., nanowires

made of magnetic, piezoelectric, semiconductor, or multiple-metal segments) and available geometries (e.g., cylindrical or hexagonal section rods, hollow tubes, single and multi-walled carbon nanotubes) that can be put to use. The elements may also be functionalized off-chip to capture different sensing targets on the same chip (multiplexing), with each target-type assembled in a predetermined area on the chip. This requires the development of assembly and integration techniques that do not destroy the capability of molecules used for functionalization.

Results Using a Hybrid Fabrication Strategy for Nanoresonator Arrays

Hybrid strategies combine the ease of patterning simple planar structures, borrowed from top-down methods, with the flexibility of tailoring individual nanostructures provided by bottom-up methods, while partially overcoming the difficulties of each method. As an example of this synergy, consider the hybrid bottom-up method reported for nanoresonator array assembly and integration, and designed to accommodate a multiplexed biosensor chip [22]. In this study, both metallic rhodium and semiconducting silicon were synthesized off-chip. Polycrystalline rhodium nanowires were grown by electrodeposition in nanoporous aluminum oxide membranes, while predominately single-crystal silicon nanowires were grown by the Au-catalyzed vapor–liquid–solid (VLS) technique. Bottom-up directed assembly was used to position single nanowires at lithographically defined locations on a silicon chip for fabricating multiplexed arrays. To do so, patterned arrays of wells in a sacrificial insulating photoresist layer covering metal guiding electrodes were defined on the chip surface. Dielectrophoretic forces were used to align the nanowires within the wells. Thick metal clamps were then built over the guiding electrodes by electrodeposition. The photoresist was dissolved to leave arrays of freestanding single-NW cantilevers as shown in scanning electron micrograph images of Fig. 6. The figures show the electrodeposited metal in intimate contact with the nanowire at each nanowire clamp. The clamps were shown to be of repeatable rigidity, with no undercuts. Besides securing the nanowire to the chip, the choice of gold for the clamp material, irrespective of the nanowire material, enables a good electrical connection between the chip and conductive nanowire, which can be used for electrical



Nanomechanical Resonant Sensors and Fluid Interactions, Fig. 6 Silicon and rhodium nanowire resonators made by a hybrid scheme using top-down patterning of electrodes and a bottom-up directed assembly scheme. Scale bar is 10 μ m, for picture at lower right. (Figure reprinted from Ref. [22], with permission)

drive and detection of single nanowires in the array. Figure 6 shows that the distortion found in metal cantilevered nanowires from thin film stress evolution in the top-down process is absent. The photoresist used as a sacrificial layer was removed without the need for etchants that could damage the surface functionalization for biosensing. The example hybrid method above eliminates expensive e-beam lithography (reduces cost), shows the viability of long cantilevered metal nanoresonators. It uses lithographic patterning to enable directed assembly and nanoresonator clamping, allows choice of different materials for clamps and nanoresonators, and shows the potential to use different nanoresonator cross-section shapes, materials, and sensor functionalization chemistries for multiplexing on a single chip.

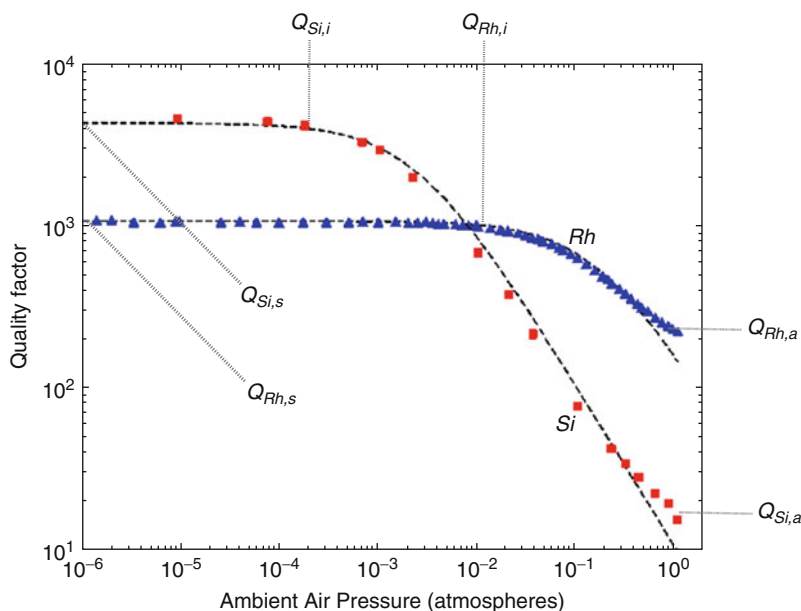
As seen above, the choice of fabrication technique plays a strong role in determining the nanoresonator properties available for sensing. In particular, flexible methods such as the hybrid technique described above, allow a wider choice of nanoresonator materials and cross sections. The density, Young's modulus and intrinsic damping within the nanowire material (C_s) can serve to provide some measure of tuning of the Q -factor for a given fluid damping, depending on the application desired. Figure 7 provides an example of this. It shows the Q -factor change from moderate vacuum to atmospheric pressure for a rhodium and a silicon nanowire of roughly comparable dimensions and fluid damping C_f . Although fine-grained, polycrystalline Rh should be expected to have a much higher

intrinsic damping C_s compared to the single-crystal silicon nanowire, its higher elastic modulus and density compared to silicon compensate (Eq. 9), giving it a vacuum Q -value ($Q_{Rh,s}$) as high as a fourth of that of the silicon nanoresonator ($Q_{Si,s}$). The experimental data [22] of Fig. 7 further agrees with a prediction that may be made from Eq. 13 – the Q -factor for the silicon nanowire, with smaller intrinsic damping C_s , should begin to deviate from its vacuum value at a lower pressure value (point $Q_{Si,i}$ vs. $Q_{Rh,i}$) and show a sharper fall-off with pressure, than the rhodium nanowire. The Q -factor of the rhodium nanowire at atmospheric pressure ($Q_{Rh,a}$) is seen to be higher than that of the silicon nanowire ($Q_{Si,a}$). Gas damping of the nanoresonator motion increases as ambient gas pressure is raised from high vacuum to atmospheric pressure. Eq. 13 shows that the high intrinsic damping in the rhodium nanowire serves as a buffer to delay the onset of Q -factor reduction and to limit its rate of change. This makes it more suitable for applications such as small mass sensing where reduction in Q -factor would adversely affect performance. However, the value of the Q -factor for the silicon nanowire begins to decrease at a lower pressure and then shows a larger change in measured Q -factor over the same pressure range, compared to the rhodium nanowire. As an instrument for measuring gas damping force, the silicon wire begins to show response at lower pressure and has a greater dynamic range than the rhodium nanowire. Silicon nanowires would thus serve as better instruments for studying gas damping as a function of pressure than rhodium nanowires.

In this section, the link between fabrication technique and nanoresonator material choice has been reviewed. Well-established top-down methods remain a strong choice for most batch-fabricated resonant sensors at microscale. For nanoresonators with dimensions far smaller than the wavelength of UV light sources, the costs of patterning nanostructures by rastering an electron beam, increase with chip area. Bottom-up methods, which rely on assembly of nanostructures made off-chip, begin to look more favorable for such structures. A larger variety of materials and shapes may be synthesized and assembled by these techniques, but precise location and individual addressing of assembled structures can be a challenge. An example of a hybrid fabrication scheme that borrows simple top-down photolithography steps for patterning relatively large microscale structures to enable

Nanomechanical Resonant Sensors and Fluid Interactions,

Fig. 7 Measured quality factor variation with ambient gas pressure for silicon and rhodium nanowires of comparable size. (Figure reprinted from Ref. [22], with permission)



spatially directed assembly and clamping to create nanoresonator arrays was reviewed. Nanoresonators of two different materials –polycrystalline rhodium and single-crystal silicon, illustrated how this method may be used to match material with different performance requirements for two different nanoresonator applications. The data show rhodium to be a more favorable nanoresonator material for a sensitive mass detector operating in air at near atmospheric pressure. The material properties of silicon make it a better choice for a nanoresonant sensor of gas damping force over a large range of pressure, spanning the transition and free molecular flow regimes of rarefied gas dynamics.

Summary and Related Reviews

This section provides a summary of the choice of topics presented in this entry and points to kindred reviews of complementary material, which could not be accommodated here. Nanoscale resonant sensors were chosen due to their potential for creating new instrument systems of unprecedented sensitivity – the promise of their application to biosensing has been included in a recent comprehensive review of mechanical biosensors [23]. Table 1 of Ref. [23] indicates detection limits of 300 fM (femto-molar) and 1.5 fM

concentration have been achieved with mechanical microcantilever resonators with and without mass-amplifying labels, respectively, compared to $\sim 1,000$ fM for the nonmechanical standard test ELISA (enzyme-linked immunosorbent assay). Bio-molecular target capture kinetics, target depletion upon capture from small analyte volumes, target sequestration of low-abundance molecules by abundant nontarget molecules are a few other areas where a research effort is focused on extending the current boundaries of knowledge for molecular diagnosis with resonant cantilevers. Related work may be found in the entries on “Cantilever Biosensors” and “Nanowire Biosensors” in this Encyclopedia.

Work in nanoresonators constitutes a large field and the focus in this entry has been restricted to nanoresonators interacting with fluids – either gases or liquids. The goal has been to establish notions and the simplest useful quantitative framework to interrogate fluid interaction with nanoresonators. Damping in fluid environments can be strong and lead to low Q -factor operation, associated with a divergence in values calculated from different definitions of the Q -factor. Low Q -factor operation demands large frequency sweeps, which can bring about nonconstant forcing due to frequency-dependent circuit parasitics. Figure 2 provides an example summary of the coupled interdisciplinary work required to deal with the

interdependencies between material properties desired for an application, fabrication schemes available, and transduction artifacts. Such work both requires and can extend knowledge in mechanics of materials, and fluid mechanics of continuum and rarefied gas flows at small scales [9, 14]. Details of transduction artifacts in nanoresonator measurements are worthy of a separate review and have not been included here. Ekinci [24] provides a useful review of nanoresonator transduction schemes. While the field of micro/nanofluidics, dealing with small-scale internal flows has acquired some maturity, small-scale external flows at high frequencies realized with nanoresonators hold promise of both technological application and an improved understanding of fluid mechanics.

Cross-References

- ▶ [AFM in Liquids](#)
- ▶ [Electric-Field-Assisted Deterministic Nanowire Assembly](#)
- ▶ [Nanomechanical Properties of Nanostructures](#)

References

1. Jensen, K., Kim, K., Zettl, A.: An atomic-resolution nanomechanical mass sensor. *Nat. Nanotechnol.* **3**, 533–537 (2008)
2. Hickman, S.A., et al.: Batch-fabrication of cantilevered magnets on attonewton-sensitivity mechanical oscillators for scanned-probe nanoscale magnetic resonance imaging. *ACS Nano* **4**, 7141–7150 (2010)
3. Radmacher, M.: Studying the mechanics of cellular processes by atomic force microscopy. In: Wang, Y.L., Discher, D.E. (eds.) *Cell Mechanics*, pp. 347–372. Elsevier, Amsterdam (2007)
4. Engler, A.J., Rehfeldt, F., Sen, S., Discher, D.E.: Microtissue elasticity: measurements by atomic force microscopy and its influence on cell differentiation. In: Wang, Y.L., Discher, D.E. (eds.) *Cell Mechanics*, pp. 521–545. Elsevier, Amsterdam (2007)
5. Lavrik, N.V., Sepaniak, M.J., Datskos, P.G.: Cantilever transducers as a platform for chemical and biological sensors. *Rev. Sci. Instrum.* **75**, 2229–2253 (2004)
6. Naik, T., Longmire, E.K., Mantell, S.C.: Dynamic response of a cantilever in liquid near a solid wall. *Sensor. Actuator. A* **102**, 240–254 (2003)
7. Verbridge, S.S., Craighead, H.G., Parpia, J.M.: A megahertz nanomechanical resonator with room temperature quality factor over a million. *Appl. Phys. Lett.* **92**, 013112 (2008)
8. Arlett, J.L., et al.: BioNEMS: nanomechanical systems for single-molecule biophysics. In: Nobel Symposium 131, Bäckaskog Slott, Sweden, 31 Aug 2005
9. Bhiladvala, R.B., Wang, Z.J.: Effect of fluids on the Q factor and resonance frequency of oscillating micrometer and nanometer scale beams. *Phys. Rev. E* **69**, 036307 (2004)
10. Ramanathan, S., Koch, D.L., Bhiladvala, R.B.: Non-continuum drag force on a nanowire vibrating normal to a wall: simulations and theory. *Phys. Fluid* **22**, 103101 (2010)
11. Zheng, Y.S., Reese, J.M., Struchtrup, H.: Comparing macroscopic continuum models for rarefied gas dynamics: a new test method. *J. Comput. Phys.* **218**, 748–769 (2006)
12. Karabacak, D.M., Yakhot, V., Ekinci, K.L.: High-frequency nanofluidics: an experimental study using nanomechanical resonators. *Phys. Rev. Lett.* **98**, 254505 (2007)
13. Svitelskiy, O., et al.: Pressurized fluid damping of nanoelectromechanical systems. *Phys. Rev. Lett.* **103**, 244501 (2009)
14. Ekinci, K.L., et al.: High frequency nanofluidics: a universal formulation of the fluid dynamics of MEMS and NEMS. *Lab Chip* **10**, 3013–3025 (2010)
15. Sahai, T., Bhiladvala, R.B., Zehnder, A.T.: Thermomechanical transitions in doubly-clamped micro-oscillators. *Int. J. Non Linear Mech.* **42**, 596–607 (2007)
16. Feynman, R.P., Leighton, R.B., Sands, M.: *The Feynman Lectures on Physics*. Addison Wesley Longman, Reading (1970)
17. Weaver, W., Timoshenko, S.P., Young, D.H.: *Vibration Problems in Engineering*. Wiley, New York (1990)
18. Carr, D.W., et al.: Measurement of mechanical resonance and losses in nanometer scale silicon wires. *Appl. Phys. Lett.* **75**, 920–922 (1999)
19. O'Mahony, C., et al.: Titanium as a micromechanical material. *J. Micromech. Microeng.* **12**, 438–443 (2002)
20. Fang, W., Wickert, J.A.: Determining mean and gradient residual stresses in thin films using micromachined cantilevers. *J. Micromech. Microeng.* **6**, 301–309 (1996)
21. Wang, M.C.P., Gates, B.D.: Directed assembly of nanowires. *Mater. Today* **12**, 34–43 (2009)
22. Li, M.W., et al.: Bottom-up assembly of large-area nanowire resonator arrays. *Nat. Nanotechnol.* **3**, 88–92 (2008)
23. Arlett, J.L., Myers, E.B., Roukes, M.L.: Comparative advantages of mechanical biosensors. *Nat. Nanotechnol.* **6**, 213–215 (2011)
24. Ekinci, K.L.: Electromechanical transducers at the nanoscale: actuation and sensing of motion in nanoelectromechanical systems (NEMS). *Small* **1**, 786–797 (2005)
25. Tian, M., Wang, J., Kurtz, J., Mallouk, T.E., Chan, M.H.W.: Electrochemical growth of single-crystal metal nanowires via a two-dimensional nucleation and growth mechanism. *Nano Lett.* **7**, 919–923 (2003)

Nanomechanics

- ▶ [Reliability of Nanostructures](#)

Nanomedicine

- ▶ [Acoustic Nanoparticle Synthesis for Applications in Nanomedicine](#)

Nanomotors

- ▶ [Molecular Modeling on Artificial Molecular Motors](#)

Nano-optical Biosensors

- ▶ [Nanophotonic Structures for Biosensing](#)

Nano-optics Biosensing

- ▶ [Nanophotonic Structures for Biosensing](#)

Nano-optics, Nanoplasmonics

- ▶ [Optical Properties of Metal Nanoparticles](#)

Nano-optomechanical Systems (NOMS)

- ▶ [Optomechanical Resonators](#)

Nanoparticle

- ▶ [Toxicology: Plants and Nanoparticles](#)

Nanoparticle Cytotoxicity

Nastassja A. Lewinski
Department of Bioengineering, Rice University,
Houston, TX, USA

Synonyms

[Cellular toxicity](#); [In vitro toxicity](#)

Definition

Nanoparticle cytotoxicity is defined as the extent to which the interaction of nanoparticles with cells disrupts cellular structures and/or processes essential for cell survival and proliferation. Cytotoxicity assays are a quick and simple way to perform initial acute toxicity assessments. Coupling nanoparticle cytotoxicity data with other safety testing data can help predict the biocompatibility of nanoparticles.

Overview

When evaluating the biocompatibility of a chemical, often the first step is to screen for adverse effects in cultured cells before using animal models. Compared to animal studies, *in vitro* testing is less ethically ambiguous, easier to control and reproduce, and less expensive. Conducting preliminary evaluations *in vitro* also provides a basis for prioritizing chemical testing and reducing the number of animals used for *in vivo* toxicity testing. Generally, the first endpoint tested in chemical biocompatibility assessments is cell lethality. Cytotoxicity assays are used to measure the extent of cell death and inhibition of cell proliferation after exposure to a chemical. Since toxicity is often dose-dependent, a range of concentrations is tested to determine the LC_{50} or the concentration at which 50% cell death occurs. The LC_{50} serves as a point of comparison of the toxicity of chemicals and is used to group chemicals into toxicity classes ranging from extremely toxic (e.g., botulinum toxin) to relatively harmless (e.g., water). Knowledge of dose is essential as many chemicals that are beneficial at low doses can be toxic at high doses.

When evaluating the toxicity of a new chemical, characterizing the composition, purity, stability, and solubility is necessary before testing. In the case of nanoparticles, more extensive characterization, including particle size and size distribution, shape, crystal structure, surface charge, and surface chemistry, is warranted since these factors may also contribute to the observed effects. Generally size, shape, and agglomeration state are thought to be governing properties; however, surface composition and structure also influence the way nanoparticles interact with biological environments. Careful attention should be paid to how the various physiochemical properties of

Nanoparticle Cytotoxicity, Table 1 Summary of common cytotoxicity assays

Assay	Mechanism of action	Indication	Limitations in testing NPs
Trypan blue	Dye exclusion	Dead cells stained blue	
Neutral red	Lysosomal retention	Live cells stained red, abs = 540 nm	Dye adsorption to NPs in the absence of cells results in absorbance readings
Calcein AM	Dye retention	Live cells fluoresce green, em = 520 nm	
Fluorecein diacetate	Dye retention	Live cells fluoresce green, em = 520 nm	
Ethidium homodimer	DNA marker	Dead cells fluoresce red, em = 620 nm	
Propidium iodide	DNA marker	Dead cells fluoresce red, em = 620 nm	Dye adsorption to NPs quenches fluorescence
LDH	LDH release	LDH presence converts dye, abs = 490 nm	
MTT	Mitochondrial reduction	Live cells stained purple, abs = 570 nm	Substrate adsorption to NPs reduces absorbance
Alamar blue	Mitochondrial reduction	Live cells fluoresce red, em = 590 nm	Substrate adsorption to NPs quenches fluorescence
Luciferin/luciferase	ATP-dependent oxidation	Live cells luminesce, em = 560 nm	
Annexin V-FITC	Phosphatidylserine binding	Apoptotic cells fluoresce green, em = 520 nm	

NPs nanoparticles, *abs* absorbance, *em* emission

nanoparticles modify their toxicity. Although a detailed assessment of the various particle characterization techniques is beyond the scope of this definition, the reader is referred to review articles on this topic [1, 2]. Given the combinatorial nature of nanoparticle testing, *in vitro* assays facilitate quick screening of nanoparticle formulations to identify which nanoparticles elicit adverse effects. Depending on their intended use, the results from initial *in vitro* tests can be used to prioritize them for further development and/or *in vivo* testing. The following sections will describe some of the most common cytotoxicity assays, the potential mechanisms of nanoparticle toxicity that have been determined to date, and some challenges for future nanoparticle cytotoxicity research.

Common Cytotoxicity Assays

Cytotoxicity assays can be categorized by the endpoints they evaluate. Markers of cell injury and death include changes in membrane integrity, mitochondrial function, and cell proliferation. While some of these endpoints can be identified by microscopy, visual inspection of cellular changes provides only

qualitative information. Instead, a majority of published nanoparticle cytotoxicity studies present measurements of cell death using colorimetric methods. A compilation of the assays discussed is presented in Table 1 toward the end of section.

Membrane Integrity

The cell membrane serves as a physical barrier separating the cell from its surrounding environment, with traffic into and out of the cell regulated by transporters, receptors, and secretion pathways. Exposure to nanoparticles could compromise the cell membrane, making it “leaky” which is the basis for the following assays. Membrane integrity is typically assessed by measuring either (1) the uptake of a cell impermeable dye, (2) the retention of a cell permeable dye, or (3) the release of a cell biomolecule into the extracellular media.

Cell impermeable dyes can only enter cells with compromised membranes; therefore, injured or dead cells are stained while intact or live cells remain unstained. Examples of dyes normally excluded from intact cells that are utilized for cytotoxicity assays include trypan blue, ethidium bromide (EtBr), and propidium iodide (PI). Trypan blue is an absorbing dye, staining dead cells blue which can be manually

counted using a hemocytometer. Although trypan blue assay results have been published by a few groups, it is generally considered a qualitative assay as it is less sensitive and less reliable than other assays. EtBr and PI are red fluorescent dyes, whose peak emission intensities at 620 nm are greatly enhanced upon binding to DNA and can be measured using a flow cytometer or spectrofluorometer. Since quantitative measurements can be derived from these fluorescent dye based assays, they are more widely used. In addition, EtBr and PI can be used to identify necrotic cells and when co-stained with annexin V-FITC differentiate them from apoptotic cells. Refer to the *Apoptosis Versus Necrosis* section for further discussion.

A complementary technique to dye exclusion uses cell permeable dyes, which tend to accumulate within cells, to assess membrane integrity. If the cell membrane is damaged, the dye can leak out of cells or its uptake is decreased, with fluorescence intensity differentiating live and dead cells. Examples of commonly used dyes include neutral red, calcein acetoxymethyl (calcein AM), and fluorescein diacetate (FDA). Neutral red absorbs at 540 nm and can be measured with a spectrophotometer. Calcein AM and FDA are nonfluorescent dyes, which are converted by intracellular esterases to their fluorescent counterpart once within the cell. Cytotoxicity can be quantified by measuring the fluorescence intensity at 520 nm using a flow cytometer or spectrofluorometer. While the fluorescent dyes can be used individually, they are more commonly used in a double “live/dead” assay. The cell sample is exposed to fluorescent dye pairs (e.g., calcein AM/EtBr, FDA/PI) where live cells fluoresce green and dead cells fluoresce red.

Another membrane integrity assay detects lactate dehydrogenase release. Lactate dehydrogenase (LDH) is an enzyme normally present in the cytosol and can only be detected outside the cell after cell damage has occurred. In this assay, LDH released from damaged cells oxidizes lactate to pyruvate which promotes the conversion of tetrazolium salt INT to formazan. The amount of LDH released, determined by measuring the absorption of formazan at 490 nm, is proportional to the number of cells damaged or lysed. Significant LDH release from the cytosol is associated with necrosis.

Mitochondrial Function

Changes in cell function due to toxicity can occur before serious cell damage, such as compromised membrane integrity, are detected. Indicators of

metabolic activity can be used to detect early cell injury. Considered the “gold standard” for cytotoxicity, the MTT assay is a colorimetric assay that measures the activity of mitochondrial enzymes. MTT, or 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide, is yellow in solution but is converted into an insoluble purple formazan product in living cells. Absorbance measurements at 570 nm can be made after solubilizing the formazan with DMSO or isopropanol, and the absorbance intensities can be correlated to cell viability. In addition, the absorbance is also representative of cell number and can be used to measure cell proliferation. While MTT is less expensive and more common, other tetrazolium dyes (MTS, XTT, or WST) similarly produce a formazan that absorbs around 570 nm in live cells but have greater sensitivity with less reaction time. Similar to the MTT assay, the alamar blue assay is another colorimetric assay that measures metabolic enzyme activity. The assay is based on the reduction resazurin, a nonfluorescent dye, to resorufin, a pink fluorescent dye, mainly by acting as an electron acceptor for enzymes such as NADP and FADH during oxygen consumption. The converted dye fluoresces at 590 nm and can be detected using a fluorometer. In addition, alamar blue shows little to no cytotoxic effects in tested cells, and thus has been used to perform several tests or kinetic measurements on the same set of cells.

Measurement of ATP levels and oxygen consumption rate are less commonly used methods of assessing nanoparticle cytotoxicity. This could be due to the assays being more expensive or requiring specialized equipment. Briefly, ATP bioluminescence assay measures the concentration of ATP, which is proportional to the number of metabolically functional cells, using a luciferin-luciferase system. Healthy cells maintain ATP concentrations through a balance of production and consumption pathways. Since the oxygen consumption rate is disrupted upon cell death, cells lose the ability to produce ATP, and since ATPases can continue to consume ATP after cell death, the concentration of ATP decreases. The decrease in ATP concentration corresponds to a loss of luciferin–luciferase luminescence. The oxygen consumption rate of cells can be measured using electron paramagnetic resonance oximetry, Clark-type polarographic oxygen electrode, and fluorescent oxygen sensors. However, to date, little or no nanoparticle cytotoxicity studies use these approaches.

Proliferation

If nanoparticle exposure induces cell senescence but not cell death, this could be considered a cytotoxic effect as cell proliferation has been disturbed. Comparing the number of cells present at different exposure time points can indicate effects on cell proliferation. Cell number can be measured by direct cell counting, nuclei counting, or by the measurement of total cell protein or DNA content, which are proportional to the number of cells. Fluorescence from dyes that label DNA, such as Hoechst or PI, is directly proportional to the sample DNA content. Compared to the control, higher fluorescence indicates proliferation and lower fluorescence signals cytotoxicity. This method is relatively cheap and fast. In comparison, markers such as ^3H -thymidine and 5-bromo-2'-deoxyuridine (BrdU) incorporation into DNA can also be used to measure cell proliferation; however, both methods require significant sample processing. As mentioned previously, MTT can also be used to measure cell proliferation since the dye absorbance is proportional to cell number.

Apoptosis Versus Necrosis

Differentiating the type of cell death can elucidate information about the mechanism of toxicity. For example, the plasma membrane of apoptotic cells remains largely intact, and therefore cell death via may not be accounted for when using assays based on compromised membrane integrity, such as PI staining. Apoptosis involves a series of signaling cascades that are initiated by the cell after disruption of intracellular homeostasis. During apoptosis, phosphatidylserine is transported from the inner surface of the cell membrane to the outer surface. Annexin V binds specifically to phosphatidylserine and therefore can be used to label apoptotic cells. In contrast, necrosis is passive and results from irreversible cell damage, such as disruption of the plasma membrane. As mentioned previously, PI is a nuclear stain that only fluoresces in cells with compromised plasma membranes. Combined cell staining with annexin V-FITC (fluorescein isothiocyanate) and PI allows for the discrimination of intact cells (no staining), early apoptotic (FITC staining only), and late apoptotic or necrotic cells (both PI and FITC staining). Staining can be observed using fluorescence microscopy or quantified using flow cytometry.

Additional Study Design Considerations

Choice of Cell Type

Most nanoparticle cytotoxicity data is generated from monocultures of cells representative of specific organs of the body. Recently co-culture in vitro systems have been developed to better simulate the cellular environment in vivo [3]. Cell lines recommended in the ASTM E2526-08 and ISO 10993-5 cytotoxicity testing standard protocols include LLC-PK3 porcine kidney cells, HepG2 liver cancer cells, and 3T3 mouse embryo fibroblasts. Other cell lines often used in nanoparticle cytotoxicity studies include A549 lung cancer cells, HEK293 human embryonic kidney cells, human epidermal keratinocytes, human dermal fibroblasts, and HeLa human epithelial cancer cells. Epithelial cells are often used because they form the outer barrier layer of organs and are usually the first cells to come into contact with nanoparticles during exposure. Regardless of which cell lines chosen, it is important to recognize that there may be differences in sensitivity between cell types [4]. In addition, if primary cells are chosen, these cells may have greater sensitivity than established cell lines.

Suitability for High-Throughput Screening

A considerable amount of time and money can be lost on nanoparticle development when they fail late in the process due to toxicity issues. The growing number of nanoparticles needing to be screened places new demands on cytotoxicity tests, with more attention being drawn to the use of high throughput methods. High-throughput screening (HTS) utilizes automated systems to rapidly conduct in vitro tests on microtiter plates. If multiple measurements are obtained from each well, then the process is referred to as high content screening. Several of the established cytotoxicity assays mentioned have been successfully adapted to 96-well and 384-well plates for use in high-throughput plate reader and flow cytometer setups [5]. Since the advantage of in vitro cytotoxicity screening is the ability to evaluate the toxicity of a large number of compounds, testing nanoparticle libraries with HTS is one strategy toward generating the volume of data needed to identify toxic nanoparticle formulations, elucidate nanoparticle structure-activity relationships, and better prioritize nanoparticles for further in vivo testing.

Nanoparticle Cytotoxicity, Table 2 Toxicological effects of some widely used NPs

NP	Toxicity observations	Reference
C ₆₀	Unmodified C ₆₀ more toxic than functionalized (hydroxylated and carboxylated) C ₆₀ , residual THF solvent can contribute to toxicity observed from unmodified C ₆₀ cluster (nC ₆₀) toxicity, photocatalytic behavior contributes to toxicity	[6]
CNTs	Unpurified as-prepared CNTs exhibit toxicity possibly due to presence of catalytic iron, single-walled CNTs elicit more toxic response than multiwalled CNTs	[7]
Au	Relatively "nontoxic," cationic AuNPs more toxic than anionic AuNPs, CTAB and citrate coatings contribute to toxicity but can be mitigated with PEG	[8]
Ag	Silver ion release contributes to toxicity, ROS generation and oxidative stress are two major mechanisms, different surface chemistries induce different DNA damage responses	[9]
QDs	Heavy-metal (e.g., Cd) leaching contributes to toxicity, shell and surface coating can reduce toxicity by containing core and modulating surface charge, toxicity also influenced by oxidative & photolytic stability	[10]
TiO ₂	Anatase form ~10 times more cytotoxic than rutile form, photoactivity enhances toxicity, oxidant driven inflammation and protein nitration also associated with TiO ₂ NP exposure	[11]

CNTs carbon nanotubes, NPs nanoparticles, THF tetrahydrofuran, CTAB cetyl trimethylammonium bromide, PEG polyethylene glycol

Potential Mechanisms of Nanoparticle Toxicity

Critical determinants of nanoparticle-induced cytotoxicity include exposed cell type, composition, and size of nanoparticles. In addition, through one-by-one evaluation of nanoparticles, structure-activity relationships specific to certain nanoparticle formulations have been observed. Several reviews provide detailed discussions on general and nanoparticle-type specific cytotoxicity, and recent examples are included in the References section. A summary of the current understanding on the mechanisms behind toxicity observed from six widely used nanoparticles with reference to recent reviews is presented in Table 2. The following section elaborates on some of the effects that are common across nanoparticle types.

Aggregation and Agglomeration

Nanoparticle stability can vary in different dispersion solutions, such as deionized water and cell culture media. Aggregation, or the strong binding of nanoparticles together, and agglomeration, or the clustering of nanoparticles held together by relatively weak forces, can change the concentration and physicochemical properties of nanoparticles in solution. The properties affected are their size, size distribution, surface area, and therefore their surface reactivity. Since these parameters play a role in the toxicity of nanoparticles, it is important to keep in mind the potential contribution of aggregation and agglomeration during study design and analysis. Methods of reducing nanoparticle

aggregation and agglomeration as well as homogeneously dispersing them in liquids include surface modifications and sonication before exposure. While surface modifications may improve nanoparticle stability, they may also influence their bioactivity. The durability of the surface coatings once inside biological environments is another issue, and release of the surface coatings may also contribute to the toxicity of nanoparticles.

Cellular Uptake

Uptake of nanoparticles into cells through endocytosis or phagocytosis processes is size-dependent and can vary greatly among cell types. The amount of nanoparticle internalization by cells can contribute to the observed toxicity as internalized nanoparticles may interact with cell organelles and disrupt cellular equilibrium or homeostasis. Transmission electron microscopy has been the preferred method of studying the cellular uptake of nanoparticles due to its high resolution, resolving the intracellular location of nanoparticles. Most studies have found internalized nanoparticles localize in endosomes or lysosomes. The low pH in endosomes and lysosomes has been suggested to dissolve or degrade nanoparticles releasing their core atoms. This can result in cytotoxicity, especially with metal nanoparticles, since metal ions can disrupt calcium homeostasis and/or generate reactive oxygen species.

Oxidative Stress and Inflammation

In vitro studies have shown that nanoparticles of various compositions generate free radicals and reactive

oxygen species (ROS) more than larger particles, due to their higher surface area. ROS can damage cells by inducing oxidative stress, which involves a redox imbalance within cells often due to increased intracellular ROS and decreased antioxidants. The oxidative stress associated with nanoparticle exposure can result from:

- The presence of transition metals (ex. iron, copper, nickel) catalyzing Fenton reactions.
- Generation of ROS at the particle surface.
- Phagocytosis of nanoparticles by inflammatory cells, such as alveolar macrophages and neutrophils.
- Physical damage due to nanoparticles interacting and/or entering mitochondria, altering mitochondrial function.

The cellular damage resulting from ROS generation and activation of inflammatory cells can also lead to inflammation. Markers of inflammation include the pro-inflammatory cytokines IL-1b, IL-6, and TNF- α plus the chemokine IL-8. Cytokines can be quantified using the enzyme-linked immunosorbant assay (ELISA), and ELISA results have been reported for metal oxides, carbon-based nanoparticle, silica, and quantum dots. However, carbon nanoparticles and metal oxides have also been reported to adsorb cytokines IL-8 and IL-6, interfering with the ELISA assay. These observations should be considered when interpreting results from nanoparticle inflammation testing.

DNA Damage

Of the cellular events that can occur in the absence of cell death, DNA damage is of critical concern as it can initiate disease development, such as cancer, and genetic abnormalities in reproductive cells. Both the nanoparticles themselves as well as nanoparticle-elicited events, such as overload and persistent inflammation, have been suggested to play a role in nanoparticle-induced genotoxicity. However, to date, limited genotoxicity testing of nanoparticles has been conducted. Genotoxicity may result from nanoparticle exposure due to their (a) small size which may facilitate their transport across cell and nuclear membranes where they may interact directly with DNA, (b) high surface area and/or chemical components (e.g., heavy metals) could generate ROS, which damage DNA. The comet assay has been used to detect DNA damage in individual cells using gel electrophoresis. Cells with damaged DNA appear as “comets” with intact DNA residing in

the head portion and broken DNA pieces migrating away, forming the tail. A DNA-specific dye such as PI is used to read the gel, and the amount of DNA found in the tail is proportional to the amount of DNA damage. The comet assay is the most widely used genotoxicity test in published nanoparticle studies; however, it only nonspecifically indicates the presence of DNA strand breaks. The most investigated oxidative lesion caused by hydroxyl radical attack, 8-OHdG, has been shown to be induced in vitro by nanoparticles, including carbon nanotubes, CeO₂, gold, and TiO₂ nanoparticles [12]. Isotope-dilution gas chromatography–mass spectrometry (GC-MS) can be used to detect and quantify DNA damage mediated by ROS produced from nanoparticle interactions in cells. This method is more robust compared to the commonly used comet assay since it has been shown to differentiate among more than 20 DNA base lesions. Since oxidatively modified DNA lesions have different mutagenic potential, it is important to identify the types of DNA modifications that occur due to nanoparticle exposure.

Challenges for Future Nanoparticle Cytotoxicity Testing

It is important to consider that in vitro dose–response data serves as a starting point for evaluating the cytotoxicity of nanoparticles. Often in vitro data cannot be extrapolated to nor is representative of what would occur in vivo. For example, cells exposed to nanoparticles suspended in cell culture media poorly reflects inhalation exposure conditions, where lung cells at the air–liquid interface are exposed to aerosols of nanoparticles. In addition, in vitro cell culture systems are controlled environments, often containing a single cell line at low cell density, which may not accurately model normal cell–cell interactions or tissue injury signaling cascades. However, these shortcomings arise from the inherent differences between in vitro and in vivo systems. Cytotoxicity assays also presents unique challenges for nanoparticle testing since particulates differ from soluble chemicals. The following section describes areas for improvement in nanoparticle cytotoxicity testing.

Dose Metric

Compared to soluble chemicals, nanoparticles can agglomerate, aggregate and, as a consequence, settle

out of solution depending on their size, concentration, and surface chemistry. These processes can affect the cellular dose, making quantification of the exposure concentration more complicated and less comparable across nanoparticle types. As a result, questions have been raised on the applicability of standard cytotoxicity assay protocols, developed to test soluble chemicals, for nanoparticles. Nanoparticle–cell interactions may be enhanced due to nanoparticles settling to the cell surface, leading to potential nanoparticle overloading of the cells. Resulting cytotoxicity may therefore be due to the higher exposure level from poor stability. It is important to account for this phenomenon when conducting nanoparticle *in vitro* dose–response assessment. In addition, the exposure dose is often inconsistently reported. There are several dose metrics that can be used, such as mass per volume, particles per volume, mass per surface area, or mass per cell number. To aid study interpretation, enough data should be reported so that the reported values can be converted to different dose metric units.

Most *in vitro* studies assess dosimetry by observing the dose–response relationship after external introduction of different concentrations of nanoparticles. Yet the internalized concentration of nanoparticles may better correlate to cytotoxicity [13]. A more informative representation of exposure dose would be to quantify the amount of nanoparticles externally attached or internalized by the cells (delivered dose) in addition to the concentration of nanoparticles added to the media (administered dose). Teegarden et al. and more recently Wittmaack provide thoughtful discussions on how the particulate nature of nanoparticles introduces issues such as gravitational settling that impact dosimetry [14, 15]. Despite these examples, few studies make a distinction between administered and delivered dose. Knowing the external and internal fractions of the delivered dose could also give insights into the mechanisms of nanoparticle toxicity; however, experimental methods to resolve these concentrations need to be developed. Future research should work to correlate the deposited and internalized nanoparticle concentrations to observed cytotoxicity.

Nanoparticle Interference

Due to their high surface area, and in some cases optical properties, the presence of nanoparticles in the test solutions can interfere with cytotoxicity assay results. The increased adsorption capacity and

increased catalytic activity resulting from their high surface area can either bind or interact with assay dyes leading to false readings. For example, carbon nanomaterials, especially single-walled carbon nanotubes (SWCNTs), have been reported to interact with cytotoxicity indicator dyes, including MTT, WST-1, Coomassie blue, alamar blue, and neutral red. Adsorption of the dyes to the carbon nanotubes results in either false negative data for the absorbing dyes or false positive data for the fluorescent dyes. Fluorescence quenching by carbon nanomaterials has also been reported with adsorption of the dye suggested as a reason for the quenching response. Other ways nanoparticles can contribute optical interference include physically blocking the emitted light, scattering the excitation light, and adding their own fluorescence signal (e.g., quantum dots or polystyrene beads). The reader is also referred to a review that describes in more detail these and other assay interferences derived from nanoparticles [16, 17]. Future developments could include designing label-free cytotoxicity assays, which would eliminate the issue of nanoparticle–dye interactions. As a safeguard, conducting multiple tests is advantageous to ensure valid conclusions are drawn.

Inter-Laboratory Consistency

While many studies using two or more different cytotoxicity assays have reported reasonable agreement in the results, large discrepancies have also been reported. Lanone et al. showed that for highly toxic or not toxic nanoparticles, inter-laboratory reproducibility was good, with LC_{50} values very similar for toxic nanoparticles [4]. However, their data also indicated that the reproducibility for nanoparticles with intermediate toxicity was lower in comparison. Since the cell lines and methodology were the same between labs, the variation between labs found suggests that there may be other factors unaccounted for in protocol development to ensure reliability in performance. These could include differences in nanoparticle storage methods, age of nanoparticle solution at testing, instrument calibration, and user performance may have contributed to the variability observed. Until data generated in different labs come within an accepted margin of error, inconsistencies in study findings and questions on methodology appropriateness during nanoparticle cytotoxicity assessment will remain.

Conclusion

Cytotoxicity testing constitutes just one of many steps required for nanoparticle biocompatibility and human health risk assessment. Traditionally, *in vitro* toxicity testing focuses on whether or not exposure to a potentially toxic agent results in cell death. However, changes in cellular function may serve as early signs of cytotoxicity when no cell damage or death may be apparent. As not all disruptive effects result in cell death, more extensive toxicity studies have attempted to determine the sublethal effects of nanoparticles. Given the variety of assays developed for cytotoxicity testing, it is important to verify that the endpoints of the assay used are appropriate for the hypothesis being tested. In addition, many of these assays were developed for testing soluble chemicals and may not be suitable for nanoparticle testing. Evidence of nanoparticle interferences with the indicator dyes has raised questions on the accuracy of currently available nanoparticle cytotoxicity data. Researchers should take care when interpreting the available literature as well as when conducting their own nanoparticle cytotoxicity testing. Ultimately, the goal of nanoparticle cytotoxicity assessments is to provide initial toxicity data to guide further risk assessment, which will inform decision making to protect human health.

Cross-References

- ▶ [Genotoxicity of Nanoparticles](#)
- ▶ [Quantum-Dot Toxicity](#)
- ▶ [In Vitro and In Vivo Toxicity of Silver Nanoparticles](#)

References

1. Card, J.W., Magnuson, B.A.: A method to assess the quality of studies that examine the toxicity of engineered nanomaterials. *Int. J. Toxicol.* **29**, 402–410 (2010)
2. Powers, K.W., Palazuelos, M., Moudgil, B.M., Roberts, S.M.: Characterization of the size, shape, and state of dispersion of nanoparticles for toxicological studies. *Nanotoxicology* **1**, 42–51 (2007)
3. Stone, V., Johnston, H., Schins, R.P.F.: Development of *in vitro* systems for nanotoxicology: methodological considerations. *Crit. Rev. Toxicol.* **39**, 613–626 (2009)

4. Lanone, S., Rogerieux, F., Geys, J., Dupont, A., Maillot-Marechal, E., Boczkowski, J., Lacroix, G., Hoet, P.: Comparative toxicity of 24 manufactured nanoparticles in human alveolar epithelial and macrophage cell lines. *Part. Fibre Toxicol.* **6**, 14 (2009)
5. Riss, T., Moravec, R., Niles, A.: Assay development for cell viability and apoptosis for high-throughput screening. In: Chen, T. (ed.) *A Practical Guide to Assay Development and High-Throughput Screening in Drug Discovery*. CRC Press, Hoboken (2009)
6. Johnston, H.J., Hutchison, G.R., Christensen, F.M., Aschberger, K., Stone, V.: The biological mechanisms and physicochemical characteristics responsible for driving fullerene toxicity. *Toxicol. Sci.* **114**, 162–182 (2010)
7. Cui, H.-F., Vashist, S.K., Al-Rubeaan, K., Luong, J.H.T., Sheu, F.-S.: Interfacing carbon nanotubes with living mammalian cells and cytotoxicity issues. *Chem. Res. Toxicol.* **23**, 1131–1147 (2010)
8. Alkilany, A.M., Murphy, C.J.: Toxicity and cellular uptake of gold nanoparticles: what we have learned so far? *J. Nanopart. Res.* **12**, 2313–2333 (2010)
9. Ahamed, M., AlSalhi, M.S., Siddiqui, M.K.J.: Silver nanoparticle applications and human health. *Clin. Chim. Acta* **411**, 1841–1848 (2010)
10. Pelley, J.L., Daar, A.S., Saner, M.A.: State of academic knowledge on toxicity and biological fate of quantum dots. *Toxicol. Sci.* **112**, 276–296 (2009)
11. Johnston, H.J., Hutchison, G.R., Christensen, F.M., Peters, S., Hankin, S., Stone, V.: Identification of the mechanisms that drive the toxicity of TiO₂ particulates: the contribution of physicochemical characteristics. *Part. Fibre Toxicol.* **6**, 33 (2009)
12. Petersen, E., Nelson, B.: Mechanisms and measurements of nanomaterial-induced oxidative damage to DNA. *Anal. Bioanal. Chem.* **398**, 613–650 (2010)
13. Lewinski, N., Colvin, V., Drezek, R.: Cytotoxicity of nanoparticles. *Small* **4**, 26–49 (2008)
14. Teeguarden, J.G., Hinderliter, P.M., Orr, G., Thrall, B.D., Pounds, J.G.: Particokinetics *in vitro*: dosimetry considerations for *in vitro* nanoparticle toxicity assessments. *Toxicol. Sci.* **95**, 300–312 (2007)
15. Wittmaack, K.: Novel dose metric for apparent cytotoxicity effects generated by *in vitro* cell exposure to silica nanoparticles. *Chem. Res. Toxicol.* **24**(2), 150–158 (2011)
16. Doak, S.H., Griffiths, S.M., Manshian, B., Singh, N., Williams, P.M., Brown, A.P., Jenkins, G.J.: Confounding experimental considerations in nanogenotoxicology. *Mutagenesis* **24**, 285–293 (2009)
17. Kroll, A., Pillukat, M.H., Hahn, D., Schneckeburger, J.: Current *in vitro* methods in nanoparticle risk assessment: limitations and challenges. *Eur. J. Pharm. Biopharm.* **72**, 370–377 (2009)

Nanoparticle Self-Assembly

- ▶ [Electric Field-Directed Assembly of Bioderivatized Nanoparticles](#)

Nanoparticle Tracking Analysis

Matthew Wright

NanoSight Limited, Amesbury, Wiltshire, UK

Synonyms

LM10; LM20; NS500; NTA

Definition

Nanoparticle Tracking Analysis (NTA) is an analytical methodology which facilitates the direct and real-time visualization and characterization of nanomaterials sized between 10 and 1,500 nm in liquid suspension.

The aim of this entry is to address the need to characterize different properties of particles in the nanoscale, explain the principles of operation of Nanoparticle Tracking Analysis, explore the concentrations and size ranges measurable, look at other parameters characterized by Nanoparticle Tracking Analysis (NTA) as well as functions and applications of the technique.

The use of nanoscale materials is widespread throughout many different industries. Whether this is in pigments for the paint industry, or viruses used in the development and manufacture of vaccines, materials in this size range play a large part in the effectiveness and constancy of many products and processes. As differences in not only the size and concentration, but also shape and both magnetic and electrical charge of these materials can make a substantial difference to the end product, it is clear there is a need for processes and techniques that have the ability to characterize the different properties of these materials. There are numerous techniques that allow different aspects of this combination of properties to be determined, each with their own sets of benefits and limitations. These techniques include: Dynamic Light Scattering, also known as Photon Correlation Spectroscopy, Analytical Disk Centrifuge, Flow Cytometry, Transmission Electron Microscopy, and Scanning Electron Microscopy.

NTA is a multiparameter methodology that allows the direct and real-time visualization, sizing, and counting of materials between 10 nm and 1.5 μm in

liquid suspension. As populations of particles can be simultaneously analyzed on an individual basis, it is ideally suited for analysis of polydisperse systems ranging from 10–20 nm to 1–1.5 μm (depending on particle type). Concurrent measurements also allow users to acquire information on concentration, zeta potential, and relative intensity of light scattered and also to analyze fluorescently labeled particles as discussed by Carr et al. [1].

Principles of Operation and Analysis

NTA obtains particle size distributions of samples by utilizing properties of both light scattering and Brownian motion.

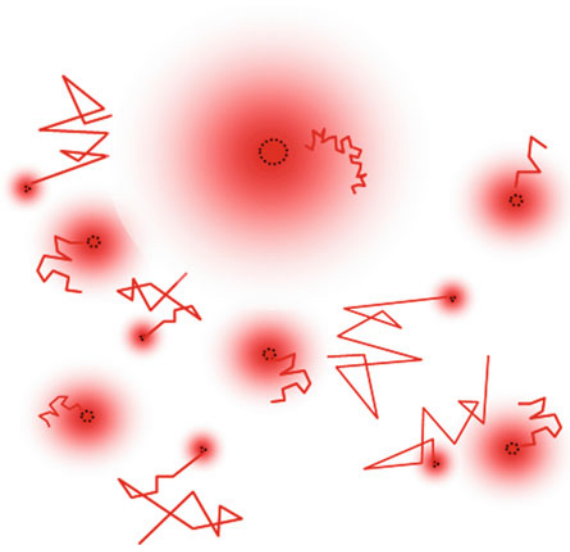
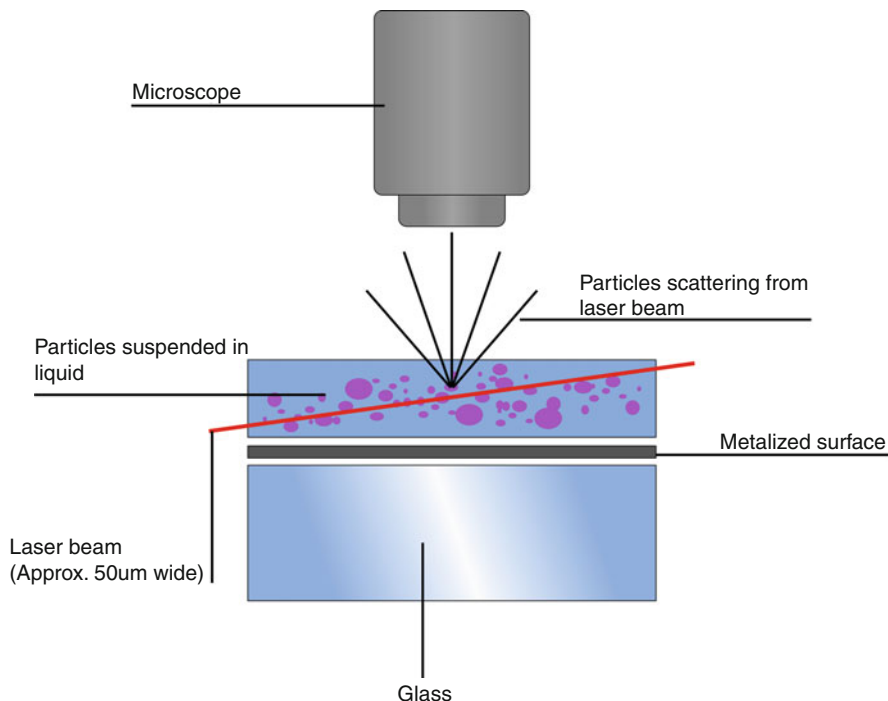
A finely focused laser beam (usually from laser diodes operating at 635, 532, 488, or 404 nm) is passed through an optical flat with a beveled edge within a sample chamber. The angle that the laser is introduced to the flat is such that when the beam reaches the interface between the flat and the liquid layer of solvent containing the sample above it, it is refracted almost flat. This compression of the laser to such a low profile gives a very high power density. As such, a region of intense illumination is created within which particles can be visualized by utilizing the light scattered from them by Rayleigh Scattering. [Figure 1](#) shows a simplified version of the configuration of laser (running from left to right), glass flat and objective lens used by systems using NTA. This allows particles to be visualized using light scattered.

A conventional microscope arrangement with a $\times 20$ magnification objective fitted is used in order to visualize the particles in suspension. Coupled with this microscope arrangement, is a video camera (typically a CCD, emCCD, or CMOS). This camera, operating at 30 frames a second, is used to capture a video field of approximately $106 \mu\text{m} \times 80 \mu\text{m} \times 6 \mu\text{m}$, within which particles in the sample can be seen moving under Brownian motion. [Figure 2](#) shows the movement of a system of polydisperse nanoparticles under Brownian motion. On average, smaller particles move at a greater rate than larger particles, as can be seen by the relative step lengths of the trails following the particles (each change in direction or “step” indicating one video frame).

Once a field of view containing particles in suspension moving under Brownian motion has been

Nanoparticle Tracking Analysis,

Fig. 1 Configuration of components used by NTA in order to visualize nanoparticles (Figure courtesy: NanoSight Ltd.)



Nanoparticle Tracking Analysis, Fig. 2 Showing the movement of polydisperse nanoparticles under Brownian motion (Figure courtesy: NanoSight Ltd.)

established, the NTA software is used to record a video file of the movement, typically 30–60 s in duration. This video file is then analyzed by the software which automatically finds the center of each scattering point

using a proprietary version of a Centroid tracking algorithm to give x and y coordinates for each particle.

Eq. 1: Centroid tracking algorithm

$$Cx = \frac{\sum_{i=1}^n \sum_{j=1}^m (x_i \cdot I_{ij})}{\sum_{i=1}^n \sum_{j=1}^m I_{ij}}$$

Where

x_i = Coordinate of pixel on the x-axis

I_{ij} = Pixel intensity [2]

The center of the same scattering point is then found in the next frame, and the distance that the particle has moved in that time is calculated .

Eq. 2: Finding displacement of particles between frames (r)

$$r^2 = \Delta x^2 + \Delta y^2$$

By tracking this movement over several consecutive frames, the software determines the mean displacement and hence the diffusion coefficient of each particle within the field of view. Provided that the temperature (T) and the viscosity (η) of the system are known, the Stokes-Einstein equation (Eq. 3) is then used to relate the diffusion coefficient to the

hydrodynamic sphere equivalent particle diameter of the material. The Stokes-Einstein equation is as follows:

Eq. 3: The Stokes-Einstein equation

$$Dt = \frac{K_B T}{3\pi\eta d_h}$$

where

Dt = Diffusion coefficient

K_B = Boltzmann constant

T = Temperature

η = Viscosity

d_h = Sphere equivalent hydrodynamic diameter particle size

As Brownian motion occurs in three dimensions, a modified version of the Stokes-Einstein equation is needed for NTA as this technique only observes the motion in two dimensions. However, it is possible to determine Dt from the mean squared displacement of a particle in one, two, or three dimensions (Eq. 4a–c respectively).

Eq. 4: (a, b and c): Tracking Brownian motion in different numbers of dimensions

$$\overline{(x)^2} = \frac{2TK_B T}{3\pi\eta d} \quad (a)$$

$$\overline{(x, y)^2} = \frac{4TK_B T}{3\pi\eta d} \quad (b)$$

$$\overline{(x, y, z)^2} = \frac{2TK_B T}{\pi\eta d} \quad (c)$$

So in the case of Nanoparticle Tracking Analysis where measurements are made in two dimensions:

Eq. 5: Tracking Brownian motion in two dimensions to determine particle size

$$\frac{\overline{(x, y)^2}}{4} = Dt = \frac{K_B T}{3\pi\eta d_h}$$

Equation 5 is applied to the trajectories of each of the particles within the field of view allowing the user to obtain a reading of particle size (hydrodynamic diameter, in nm) against concentration (absolute in particles per ml).

Other Measurable Parameters

Concentration

The field of view interrogated by NTA systems is of known dimensions and therefore known volume which means that determining the concentration of particles in terms of particles per ml can easily be achieved using the NTA software. As the software not only tracks and sizes every particle within the field of view but also counts the number of particles identified, a concentration can be achieved by extrapolation of this number.

The field of view measures $106 \mu\text{m} \times 80 \mu\text{m} \times 6 \mu\text{m}$. This leads to a volume of $50,880 \mu\text{m}^3$. Equation 6 shows how this is used in order to determine particle concentration.

Eq. 6: Extrapolation of particle count to concentration in particles per ml

$$\text{Concentration} = P_{\text{count}} \times \left(\frac{10^{12} \mu\text{m}^3}{50,880 \mu\text{m}^3} \right)$$

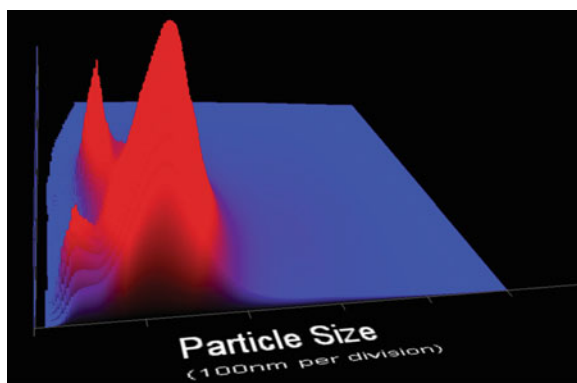
$$\text{Concentration} = P_{\text{count}} \times 1.97 \times 10^7$$

where P_{count} = Number of particles within the field of view

In a 2010 study Filipe et al. investigated the ability for NTA to measure concentration and concluded that the capability of NTA to determine approximate concentration ranges for submicron particles offers an advantage over other techniques for particle sizing [3]. A study conducted by Du et al. found that concentrations obtained using NTA correlated very well with those they calculated using other methodologies [4].

Scatter Intensity

While the size of particles is determined by the NTA software using Brownian motion, the system also measures the relative amount of light scattered by each particle in the scattering volume. These two properties can be plotted as a function of one another allowing differentiation of particles not only by size, but also by composition. For a mixture of similar sized particles with significantly different refractive indices (for example, gold particles of 100 nm and polystyrene particles of 100 nm), both would appear in the same size distribution, but can be separated when size is plotted against relative refractive index. Figure 3



Nanoparticle Tracking Analysis, Fig. 3 Illustrating differentiation of particle material by comparing particle size (x-axis), relative intensity (y-axis), and number of particles (z-axis) (Figure courtesy: NanoSight Ltd.)

shows a distribution of 30-nm and 60-nm gold nanoparticles mixed with a population of 100-nm polystyrene particles, separated by both particle size and the relative amount of light scattered. Particle size is given on the x-axis, with relative intensity on the y-axis and number of particle on the z-axis.

These populations may have merged with one another on a simple particle size versus concentration graph. However, by employing the third dimension of relative intensity of light scattered, it is easy to define the three discrete populations of materials.

Fluorescence

Being a direct visualization and tracking method, NTA does not rely on interference from coherent light sources. This means that fluorescently labeled or inherently fluorescent materials can be analyzed without affecting the results that are obtained. In order to do this, typically a green laser (532 nm) or a blue/violet laser (404 nm) is used in place of the usual red laser used within the system allowing the fluorescent excitation of particles with absorption spectra that span these wavelengths. By comparing profiles generated with and without a filter removing the excitation wavelength present in the optical path of the system, the user is able to determine the effectiveness of their labeling process, or the proportion of labeled to unlabelled material within their sample, and whether or not the presence or absence of fluorescence is in any way related to the hydrodynamic sphere equivalent particle size.

Suspension Viscosity

NTA can also be used to determine unknown viscosities of fluids. The system uses the Stokes-Einstein equation in order to relate the diffusion coefficient of the material to the sphere equivalent particle size. However, if the size of the particles in question is already well characterized, then it is possible to rearrange the Stokes-Einstein equation as shown in Eq. 7 and use the known parameters of diffusion coefficient, particle size, and temperature in order to use the collected data to determine the viscosity of the system. This will be briefly discussed in the section on drug delivery studies.

Eq. 7: Stokes-Einstein equation to find viscosity

$$\eta = \frac{K_B T}{3\pi D_t d_h}$$

where

D_t = Diffusion coefficient

K_B = Boltzmann constant

T = Temperature

η = Viscosity

d_h = Sphere equivalent hydrodynamic diameter particle size

Zeta Potential

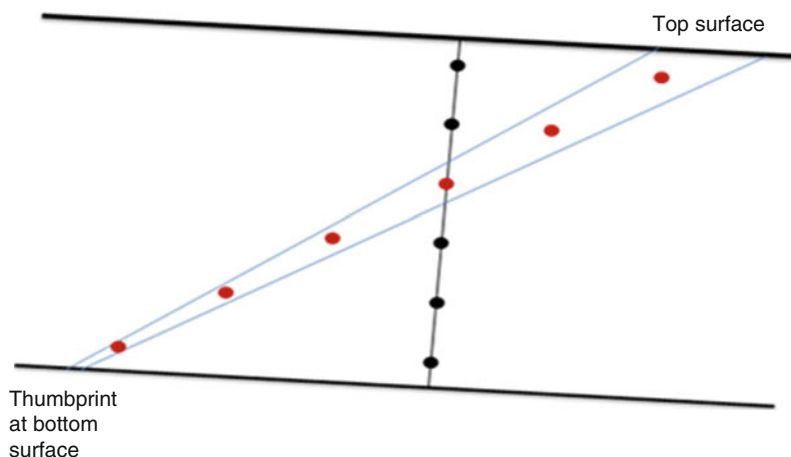
The zeta potential of a system is a measure of charge stability and has an effect on all particle-particle interactions within a system. Understanding zeta potential is of critical importance in controlling dispersion and determining the stability of a nanoparticle suspension. The zeta potential is the electric potential at the slip plane between the bound layer of solvent molecules surrounding the particle, and the overall solution. This can be closely linked to the particle's surface charge in simple systems but is also heavily dependent on the properties of the solvent in which the particles are suspended. A higher level of zeta potential results in greater electrostatic repulsion between the particles, which results in less aggregation and flocculation.

In order to carry out investigations into the zeta potential of a system, the hardware used for NTA is adapted so that the sample chamber is fitted with platinum electrodes through which an electric field can be applied to the cell.

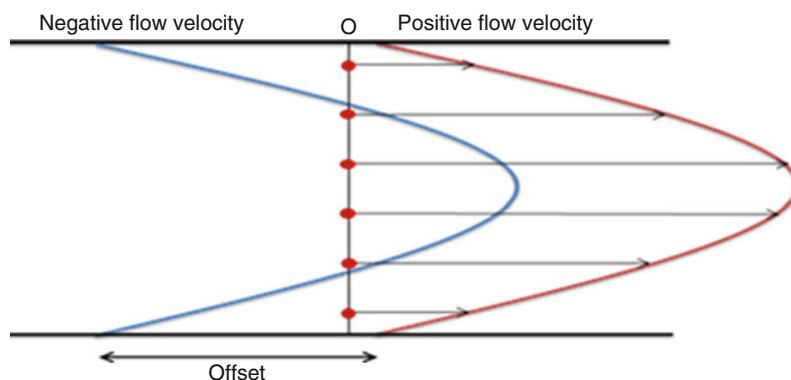
This field causes motion of both the sample particles (electrophoresis) and the diluent (electroosmosis). The NTA software records the apparent drift velocity

Nanoparticle Tracking

Analysis, Fig. 4 Diagram showing profile map positions where six standardized depth locations intersect the laser beam (Figure courtesy: NanoSight Ltd.)

**Nanoparticle Tracking**

Analysis, Fig. 5 Total velocity profile measured using NTA (*red*) and electroosmotic velocity profile (*blue*) inferred by subtraction of a contact offset to obtain a velocity profile that sums to zero over the channel depth (Figure courtesy: NanoSight Ltd.)



for each tracked particle, which will be a superposition of both electrophoresis and electroosmosis. This total velocity is measured at different depths within the sample chamber (as shown in Fig. 4, each dot corresponding to a depth of measurement), making it possible to separate these components and obtain a measurement of the electrophoretic velocity (due to the force impinging directly on the particles), and hence the zeta potential of the particles.

By employing the assumption that the net electroosmotic flow summed over the chamber depth will be zero, as should be the case for a closed system, the electroosmotic profile can be inferred, as shown in Fig. 5.

Using the calculated flow profile, the velocity component caused by electroosmosis is removed from the tracked velocities for each particle. The corrected drift velocities then give a measure of the electrophoretic mobility of individual particles.

The electric field strength (E) within the flow cell is determined using the Voltage (V) applied through the sample by the electrodes, and the distance between the electrode surfaces (d) (Eq. 8).

Eq. 8: Determining electric field strength

$$E = \frac{V}{d}$$

The particle velocities can then be converted into electrophoretic mobilities (velocity divided by electric field strength).

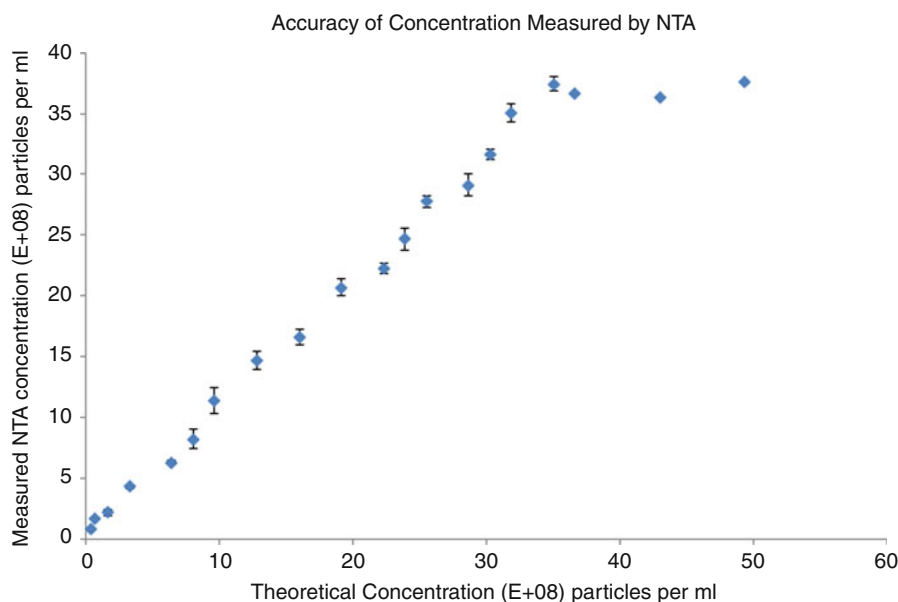
By application of the Henry equation using the Smoluchowski approximation (appropriate for aqueous diluent media with moderate electrolyte concentration, seen in Eq. 9) the zeta potential (ZP) for each particle can be calculated:

Eq. 9: Calculation of zeta potential

$$ZP = \frac{\mu\eta}{\epsilon_0\epsilon_r}$$

Nanoparticle Tracking

Analysis, Fig. 6 Relationship between absolute concentration (x) versus that measured by NTA(y) (Figure courtesy: NanoSight Ltd.)



where μ is the electrophoretic mobility of the particle, ϵ_0 is the permittivity of free space, ϵ_r is the relative sample solution permittivity, and η is the sample solution viscosity.

This particle-by-particle approach allows different charge populations within a sample to be determined rather than being lost in the “noise” of an ensemble technique.

Concentration and Size Ranges Measurable

The ability of NTA to analyze polydisperse samples comes from the ability to track and analyze each particle in the system individually, irrespective of the others around it. It is therefore important that a sufficient number of particles are analyzed during analysis to ensure that the data produced is statistically viable. Repeatable and accurate results can be obtained when analyzing particles in concentrations between 10^7 and 10^9 particles per ml. Below this concentration, the number of particles within the field of view is so low that the typical analysis time of 30–60 s will only allow a very small number of particles to be analyzed, leading to inaccuracies of a statistical nature. Above 10^9 particles per ml, the field of view becomes very crowded. As Brownian motion is a random process, when two particles cross over, the tracking software is unable to determine which particle is which and must

therefore stop tracking both to avoid erroneous tracking. When the field of view is crowded, the rate at which crossover of particles occurs is so great that there is insufficient time to determine diffusion coefficient and therefore size.

A high concentration of particles also leads to erroneous results when trying to identify solution concentration. Figure 6 shows the relationship between the concentration of particles in a solution and the concentration as measured using NTA.

It is clear that the correlation between the actual and measured concentration remains stable until a concentration of around 3.5×10^9 particles per ml whereupon the accuracy of the NTA concentration reading begins to deteriorate.

The typical size range between which NTA operates is from 10–20 nm up to 1.5 μm . The lower limit of this figure is mainly dependent upon the amount of light scattered by the particles. This value in turn varies in relation to particle size, shape, refractive index (real and imaginary), illumination power, wavelength, power, and polarization as well as solvent refractive index. The amount of light falling on the detector also fluctuates according to many variables including efficiency of optics, sensitivity, and spectral response of the camera.

In their 1983 work, Bohren et al. comprehensively describe the theory of light scatter and the formula for Rayleigh scattering of small particles [5]:

Eq. 10: Rayleigh scattering of light by small particles

$$\frac{I}{I_{in}} = \frac{16\pi^4 a^6}{r^2 \lambda^4} \left(\frac{n^2 - 1}{n^2 + 2} \right) \sin^2 \psi$$

where

a = Particle radius

n_1 = Refractive index of particle

n_2 = Refractive index of solvent

λ = Wavelength of light

I_{in} = Incident power per unit area

I = The scattered power per unit area

r = Distance from the scattering region and

ψ = Angle between input polarization and scattering direction

The total scattering (P_{scat}) into an aperture of collection angle θ (numerical aperture $NA = \sin \theta$) is then:

Eq. 11: Total light scattered into an aperture

$$P_{scat} = \frac{64\pi^4 a^6}{\lambda^4} \left(\frac{n^2 - 1}{n^2 + 2} \right) \eta_0 I_{in}$$

where $\eta_0 = \frac{(1 - \cos \theta)}{4} + \frac{(1 - \cos^3 \theta)}{12}$

As the system has fixed laser wavelength and power as well as a fixed aperture and detection angle, the variables that affect the lower limit of detection consist primarily of the particle size and the relative refractive index difference between the particles and the solvent in which they are suspended.

As light scattering at these scales varies as a function of radius [6], as particles reduce in size, they scatter considerably less light. The functionality of NTA relies on the ability to visualize the particles, and so if the amount of light scattered is not sufficient that the particles can be seen, then it is not possible to size them. For strongly scattering materials such as colloidal gold, the lower limit of detection is around 10 nm. For those materials with a weaker scattering potential, such as biological materials, the lower limit of detection is more likely to be 20–40 nm. The upper size limit of the system is determined by the rate of Brownian motion. When particles reach a certain size the rate of Brownian motion becomes so low that inherent centering errors in the software start to negate the tracking of the movement leading to inaccuracies in sizing. For dense, high Ri particles such as gold, their rapid sedimentation rate and excessive scattering

restricts the upper size at which they can be analyzed to approximately <500 nm. This limit is typically 1–2 μm .

Applications of Nanoparticle Tracking Analysis

Owing to the varied nature of work at the nanoscale, NTA has been found to be of use across a variety of different applications.

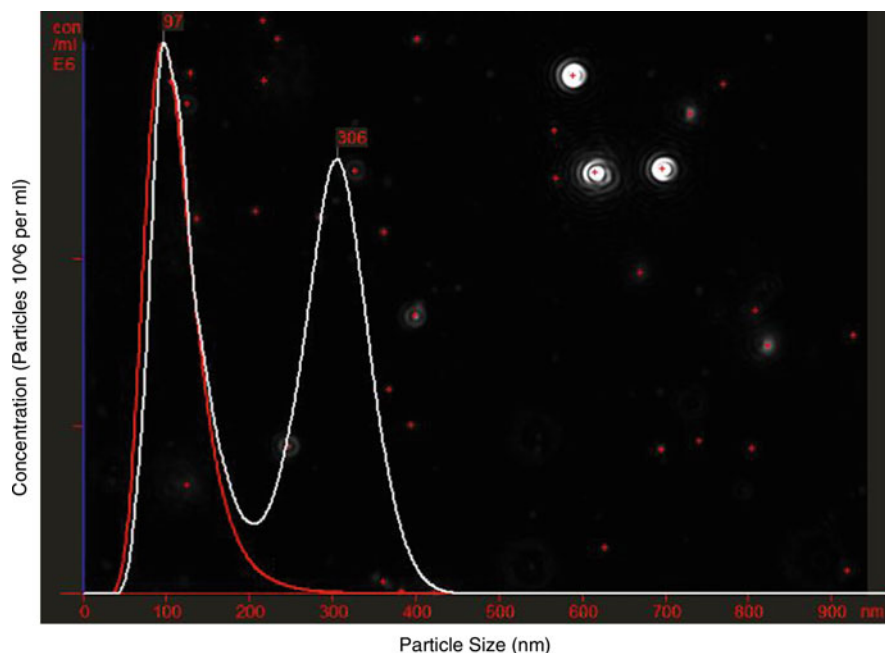
Virology and Vaccine Production Due to the nature in which NTA tracks, sizes, and counts particles on an individual basis, the methodology allows users to obtain an immediate and direct estimation of product purity and concentration. This also allows researchers and manufacturers to monitor the degree and rate of formation of aggregates within a virus preparation, complementing the process of product optimization and purification. An example of the use of NTA for virus purification studies can be seen in Fig. 7. This shows the difference between a partially purified virus preparation (white line) and the same sample having undergone a successful purification protocol, removing all the aggregated material or contaminants (red line). In this figure, the x-axis represents the particle size and the y-axis represents the particle concentration in virus particles per ml.

Exosomes Originally, it was thought that microvesicles (typically in the 100 nm to 1 μm range) and nanovesicles (also known as exosomes, with a size range of around 30–100 nm) were cellular debris. However, it is now appreciated that these cell-derived particles are of significant importance and play an important role in cell to cell communication and cell signaling.

Recent research suggests that exosomes are present in significantly elevated concentrations in samples taken from patients with a number of diseases including preeclampsia, diabetes, coronary artery disease, and cancer. NTA is not only able to visualize, size, and count these materials in liquid suspension, but when used with specific fluorescent labels, for example, antibody-mediated fluorescent labeling, it can be used to selectively analyze specific populations within a complex sample.

In their evaluation of systems for the analysis of exosomes in liquid suspension Harrison et al. concluded that NTA's single particle tracking and sizing

Nanoparticle Tracking Analysis, Fig. 7 Illustrating the use of NTA in virus purification processes with purified (red line) against non-purified (white line) product shown (Figure courtesy: NanoSight Ltd.)



offers advantages over bulk scattering measurements found in other light scattering methodologies [6].

Nanoparticle Toxicology and Ecotoxicology Current research suggests that the biological response to nanoparticles in living organisms, and thus any potential health implications, can be affected by particle size, composition, shape, aggregation state, and exposure route. Before any studies into these effects can be carried out, it is imperative that the particles used are as well characterized as possible, in particular their size distribution and concentration in appropriate test media. Researchers are also interested in what happens when engineered nanoparticles are released into the natural environment.

Protein Aggregation Biological activity and immunogenicity of a biopharmaceutical product are both highly susceptible to the state of protein aggregation. These aggregates span a broad range of sizes, from the small oligomers (a few nanometers) to insoluble micron-sized aggregates that can contain millions of monomer units. Determining and understanding the state of aggregation in these systems can not only help in improving product stability, but can also lead to efficiency gains, and aid in process optimization. Individual protein monomers are far smaller in size than the lower resolving limit of the Nanoparticle Tracking Analysis methodology. When a solution containing

only protein monomers is viewed using this technique, a background “haze” is seen within which individual monomers cannot be resolved. However, when aggregates start to form within the system, and reach around 25–30 nm in size, it is possible to differentiate them from the background signal, and count and size them as the technique normally would. This, combined with the possibility of controlling the temperature within the sample cell, from 10° below ambient to 60°C, can allow characterization of the parameters that cause aggregation within the product. A full critical evaluation of the NTA technique by Filipe et al. concluded that the technique was suitable for the analysis of protein aggregates; however, care should be taken that sample preparation does not affect the aggregate distribution.

Drug Delivery By utilizing Nanotechnology, it is possible to deliver drugs to a specific area of the body. Localized drug delivery allows the required drug volume to be reduced, along with any associated negative health implications. Both poorly soluble and fragile drugs can be carried using drug delivery vectors such as degradable polymeric nanoparticles or liposomes. It is clear that both the particle size and concentration will affect the payload, and so Nanoparticle Tracking Analysis provides a unique way to monitor these properties in liquid suspension. The location of drug delivery can also have an impact on the effectiveness of the

drug. If particles are unable to pass through biological barriers, then these can act as natural filters, preventing the drug from reaching the desired location. By using Nanoparticle Tracking Analysis to compare concentrations of particles before and after passing through membranes and barriers researchers are able to study the ability for the drug delivery vector to reach the desired location.

Nanobubbles *Nanobubbles* are stable pockets of gas in liquid suspension which can be applied to drug delivery, cleaning, and sterilization. As Nanoparticle Tracking Analysis allows direct visualization, sizing, and counting of the nanobubbles in liquid suspension, users are able to characterize their product in the final liquid state. This could prove to be difficult when using other techniques such as electron microscopy, due to the preparation required in order for electron microscopy to function.

Inks and Pigments In the field of inkjet applications, the particle size and concentrations of pigment particles are very important. The particle size has a great effect on the strength of the color generated, and aggregation of the particles could lead to the blockage of nozzles. The high-resolution size distribution that Nanoparticle Tracking Analysis can provide allows both studies into the relationship between size and color to be analyzed, as well as the conditions that lead to aggregation of pigment particles.

In conclusion, Nanoparticle Tracking Analysis is an analytical technique which allows the direct visualization, sizing, and counting of materials between 10 and 1,500 nm in liquid suspension. The technique works by tracking the Brownian motion of each particle within the field of view, irrespective of the others, and relating this movement to the sphere equivalent hydrodynamic diameter of the particle. This parameter is calculated for each particle within the field of view, and a particle size distribution is produced. In addition, particle concentration, light scattering intensity, zeta potential, and fluorescent properties can also be measured. The technique is of use in a variety of different applications, both as a competing and complementary technique to other particle sizing methodologies.

Cross-References

- ▶ [Chitosan Nanoparticles](#)
- ▶ [Detection of Nanoparticulate Biomarkers in Blood](#)

- ▶ [Dye-Doped Nanoparticles in Biomedical Diagnostics](#)
- ▶ [Liposomes](#)
- ▶ [Nanoencapsulation](#)
- ▶ [Nanomedicine](#)
- ▶ [Nanoparticles](#)
- ▶ [Optical Properties of Metal Nanoparticles](#)
- ▶ [Perfluorocarbon Nanoparticles](#)
- ▶ [Toxicology: Plants and Nanoparticles](#)

References

1. Carr, B., Hole, P., Malloy, A., Nelson, P., Wright, M., Smith, J.: Applications of nanoparticle tracking analysis in nanoparticle research- a mini review. *Eur. J. Parenter. Pharm. Sci.* **14**(2), 45–50 (2009)
2. Cheezum, M.K., Walker, W.F., Guilford, W.H.: Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys. J.* **81**, 2378–2388 (2001)
3. Filipe, V., Hawe, A., Jiskoot, W.: Critical evaluation of nanoparticle tracking analysis (NTA) by nanosight for the measurement of nanoparticles and protein aggregates. *Pharm. Res.* (2010). doi:10.1007/s11095-010-0073-2
4. Du, S., Kendall, K., Morris, S., Sweet, C.: Measuring number-concentrations of nanoparticles and viruses in liquids on-line. *J. Chem. Technol. Biotechnol.* (2010). doi:10.1002/jctb.2459
5. Bohren, C., Huffman, D.: *Absorption and Scattering of Light by Small Particles*. Wiley, New York (1983)
6. Harrison, P., Dragovic, R., Albanyan, A., Lawrie, A., Murphy, MF and Sargent, I.: Application of dynamic light scattering to the measurement of microparticles; In: 55th Annual Meeting of the Scientific and Standardization Committee of the ISTM, The XXII Congress of The International Society on Thrombosis and Haemostasis, Boston, USA, 11–18 July 2009 (in press)

Nanoparticles

Paola Martino¹, Paolo Allia² and Alessandro Chiolerio³

¹Politronica inkjet printing technologies S.r.l., Torino, Italy

²Materials Science and Chemical Eng. Department, Politecnico di Torino, Torino, Italy

³Physics Department, Politecnico di Torino, Torino, Italy

Synonyms

[Nanopowders](#); [Synthesis of nanoparticles](#)

Definition

Ultrafine particles featuring a size distribution in the nanometer range.

Introduction

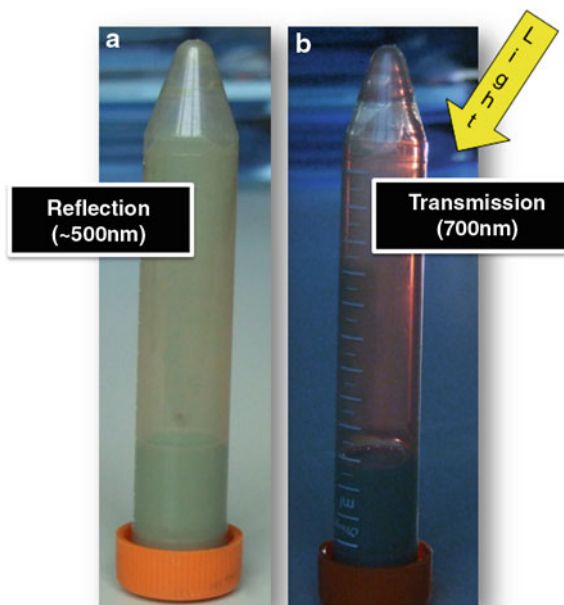
Studies of nanostructured materials have been steadily developing in recent years; in particular, metallic, magnetic, and oxide nanoparticles (NPs) have become a subject of increasing scientific interest. At the present time they hold a place of great importance that covers a wide range of potential applications including biomedical, optical, and electronics.

Although NPs are generally considered an invention of modern science, they actually have a very long history: centuries ago, NPs of silver, gold, and copper were used by artisans to decorate pottery; the first examples date from the IX century A.D. and were found in Mesopotamia [1]. An important archeological finding is a Roman cup, currently at British Museum in London, called the Lycurgus cup (fourth century A.D.). It contains tiny amounts of colloidal gold and silver, which give it unusual optical properties: when illuminated from outside, it appears green; when retro-illuminated by an inner source, it glows red. An example of this property is well represented in Fig. 1 (a–b) that shows a special silver colloidal NPs ink (by Politronica[®] Inkjet Printing Technologies S.r.l.) in emitted (green) and transmitted (red) light.

In addition, NPs exist widely in the natural world, for example, as photochemical by-products generated by plants and algae. They have also been created for thousands of years as by-products of combustion and cooking, and more recently from industrial and vehicle exhausts [2].

Definition of Nanoparticles

The term “nanoparticles” (NPs) is a very general one and mainly identifies ultrafine particles in the size of a nanometer. The definition of NPs differs depending upon the materials, fields, and application one refers to. In a narrow sense, they are regarded as the particles smaller than 10–20 nm but, generally speaking, all particles in the 1 nm–1 μm range may still be defined as NPs [3].



Nanoparticles, Fig. 1 Silver colloidal NPs ink (by Politronica[®] Inkjet Printing Technologies S.r.l.) in reflected light (Fig. 1a – green) and in transmitted light (Fig. 1b – red)

The principal parameters of NPs are their shape (including aspect ratios where appropriate), size, and the morphological substructure of the substance.

“Size Effect” and Size-Dependent Properties

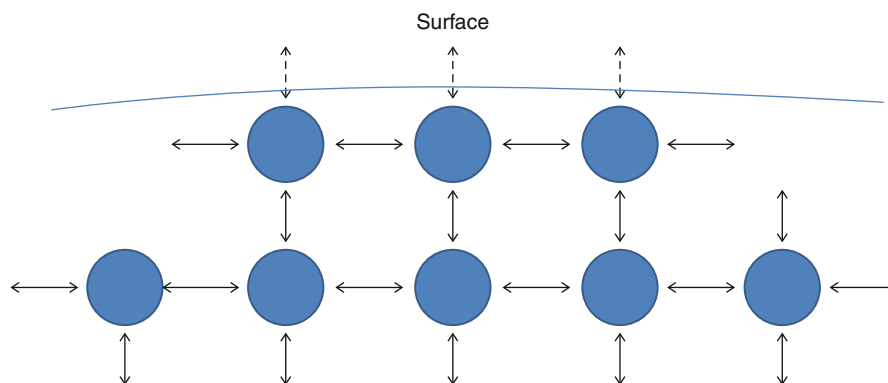
NPs represent a bridge between bulk materials and atomic, ionic, or molecular structures. Bulk materials should have constant physical properties regardless of its size, whereas at the nanoscale, size-dependent properties are often observed.

The changes in the physical properties with particle size (from bulk to NPs) are called “size effects”: by decreasing particle size, the solid particles generally tend to show different properties from the bulk material due to the increased importance of surface upon bulk. Even physical properties like melting point and dielectric constant may change. In addition, it is important to underline that the “size effect” not only allows one to discriminate between physical properties of the bulk when compared to NPs, but also to distinguish among NPs having different size distributions.

For bulk materials larger than one micrometer, the fraction of atoms, ions, or molecules at the surface is insignificant in relation to the fraction corresponding to

Nanoparticles,

Fig. 2 Schematic representation of surface of a NP. The bonds on the solid surface are broken up

**a**

Atom colorless
1Å

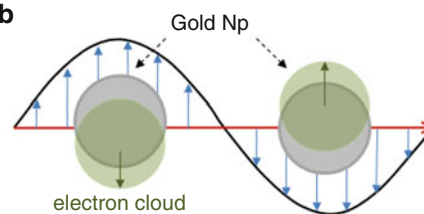
Gold Np < 1 nm
orange

Gold Np : 3–30 nm
red



Gold Np : 30–500 nm
purple to blue

Gold bulk

**b****Nanoparticles, Fig. 3**

(a) Size-dependent changes in color of gold, from gold bulk to atomic size; (b) schematic drawing of plasmon oscillation for a gold NP, showing the displacement of the conduction electron charge cloud relative to the nuclei

its volume. In NPs the fraction of atoms, ions, molecules located at the surface becomes greater and, therefore, more active because it is subject to asymmetric interaction forces due to a different bonding state, the bonds at the interface between solid and vacuum (or liquid or atmosphere) are broken (Fig. 2). This affects either the properties of the particle itself or its interaction with other materials.

One of the most important size-dependent consequences of the “size effect” of NPs is the drastic reduction of the *melting point* of ultrafine particles if compared to that of bulk material: the high surface area to volume ratio of NPs provides a big driving force for diffusion. NPs, in fact, tend to be able to move easier at lower temperatures. For example, the melting point of gold is 1,336 K, but if the size of material decreases to 2.5 nm, it becomes about 570 K [3].

A second size-dependent feature of NPs is the *dielectric constant*. The use of NPs embedded in electronic devices allows one to improve the product performance making smaller devices with high dielectric constants. PbTiO_3 represent a significant example of this feature because its dielectric constant tends to increase considerably as the particles become smaller than about 20 nm [3].

NPs often possess unexpected *optical properties* as they are small enough to absorb the light with a specific wavelength: a coherent oscillation of the conduction band electrons is induced by interaction with the optical electromagnetic field. Consequentially, a transmitted light with different color depending upon the kind of metal and particle size is obtained. This resonance is called Surface Plasmon Resonance (SPR) (Fig. 3) [4].

Another size-dependent feature of NPs are *mechanical properties*. Nanocomposites with nano-scale dispersed phases and nanocrystalline materials provide quite different mechanical properties when compared to conventional microstructures. The hardness of the crystalline materials generally increases with the decreasing of crystallite size and the mechanical strength of the metal and ceramic materials considerably increases by decreasing grain size down to the nanoscale or by making composites at the nanoscale [5].

The NPs are seldom used by themselves; they are dispersed instead in other materials or combined with them. The dispersion process is a crucial point for the NPs technology as well as their preparation methods, because the performance of final products is affected by their dispersion conditions. In this way, it is expected with great possibility to develop various new materials and applications by producing and processing the NPs, which have different properties from the bulk material due to size effects as mentioned above [3].

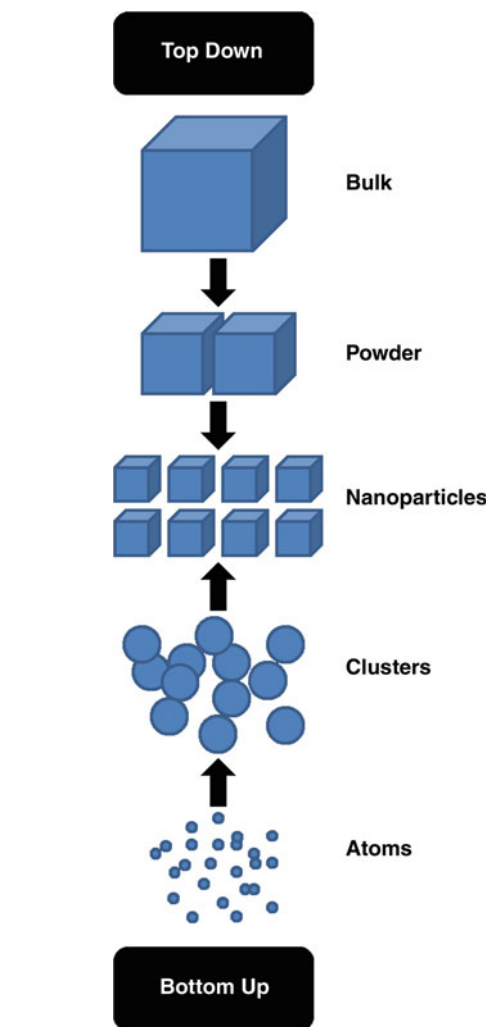
Synthesis Methods

Synthesis methods for NPs are typically grouped into two categories: *top-down* and *bottom-up* approaches (Fig. 4). The first approach involves the breaking down of the bulk material into nano-sized structures or particles. This approach may address milling or attrition, chemical methods, and volatilization of a solid followed by condensation of the volatilized components. Metal NPs so produced are usually large in size (close to 1 μm) and have wide size distribution. High-energy ball milling is shown below as an example of such a method.

The alternative approach, *bottom-up*, refers to the buildup of a material from the bottom: atom-by-atom, molecule-by-molecule, or cluster-by-cluster in a gas phase or in solution. The latter approach is far more popular and less expensive in the synthesis of NPs and can be identified in two sub-categories: (1) Vapor (Gas) Phase Methods (e.g., pyrolysis, inert gas condensation); (2) Liquid Phase Methods (e.g., solvothermal reaction, sol-gel, microemulsion method).

Top-Down: High-Energy Ball Milling

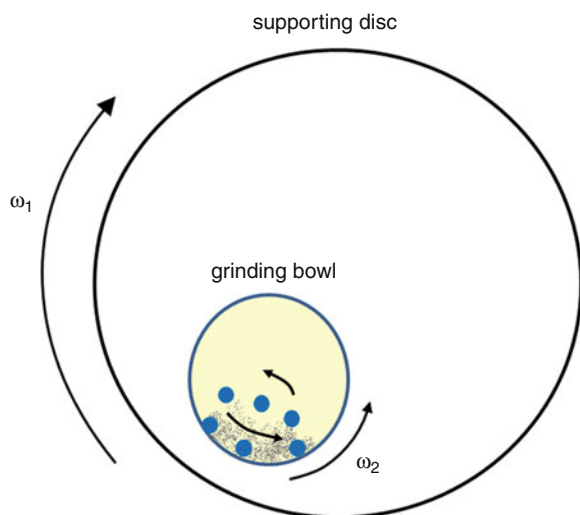
Within top-down technologies, one of the most representative techniques is High-Energy Ball Milling (HEBM). This technique is well known as a simple and



Nanoparticles, Fig. 4 Schematic representation of the *bottom-up* and *top-down* synthesis processes of nanomaterials

inexpensive method for the production of nanostructured materials, in particular it allows the preparation of alloys and composites, which cannot be synthesized via conventional routes. Different types of milling devices have been developed for different purposes including tumbler mills, attrition mills, shaker mills, vibratory mills, planetary mills, etc. In HEBM, particles of the starting material are subject to heavy deformation, cold work hardening, and subsequent fragmentation [6].

In a HEBM system, the grinding balls and the material in the grinding bowl are thus acted upon by centrifugal forces, which constantly change in direction and intensity. The grinding bowl and the



Nanoparticles, Fig. 5 Schematic drawing of a high-energy ball mills

supporting disc rotate in opposite directions, so that the centrifugal forces alternatively act in the same and opposite directions (Fig. 5). The process requires a long milling time to activate and complete the structural changes and chemical reactions, as well as a control of milling atmosphere and temperature which are crucial to create the desired structural changes or chemical reactions [7].

Bottom-Up: Vapor (Gas) Phase Methods

Vapor (Gas) phase synthesis is a well-known chemical manufacturing technique for an extensive variety of nano-sized particles. The transformation of a gaseous precursor to the final particulate is a complex physical process involving nucleation of the particulate phase, condensation of gasses onto particles, and coagulation and coalescence between particles.

Vapor phase fabrication processes are typically conducted in vacuum under elevated temperatures. The mechanism for particle formation starts with the precursor vaporization (typically involves a catalyst) followed by the nucleation and growth stage.

Synthesis of NPs via high-temperature processing is a promising technique that offers a good route to production of high-purity NPs with specifically tailored chemical and physical properties. However, NPs generated by gas phase processes are usually in the form of aggregates, due to their Brownian coagulation at high temperatures in the reactor: coagulation is an

intrinsic mechanism which inevitably occurs at high particle concentrations and therefore in all industrial aerosol processes [8].

Bottom-Up: Spray Pyrolysis Methods

Spray methods, such as spray pyrolysis, spray drying, and the flame spray method, have the advantage of being continuous, low-cost, and single-step methods for the preparation of NPs and nanocomposites (which includes metals, metal oxides, non-oxide ceramics, superconducting materials, and nanophase materials) that are spherical, chemically homogeneous, and comprised of multiple materials. A representative spray system is shown schematically in Fig. 6.

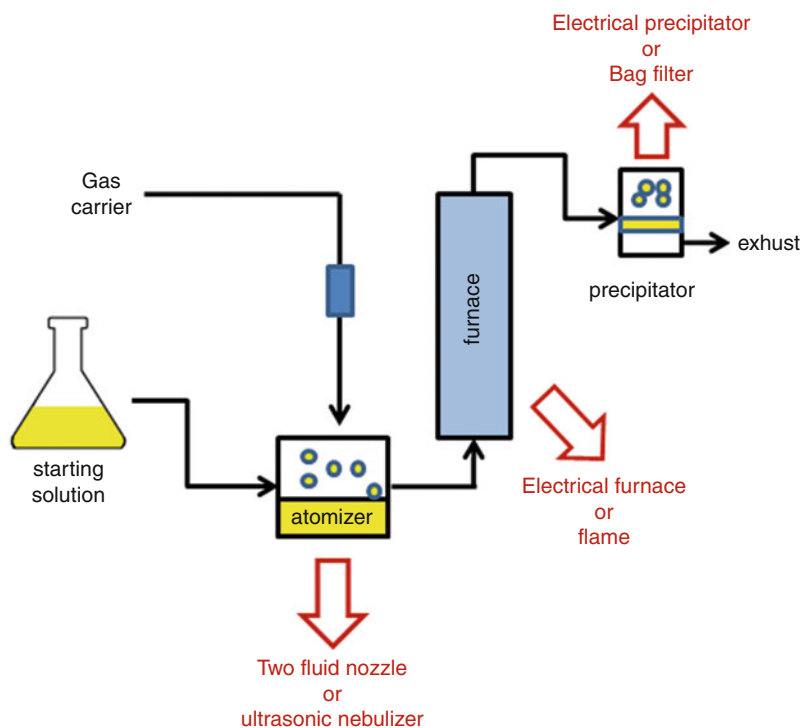
In spray pyrolysis, the starting liquid solution is dispersed as droplets into a gas carrier and then sprayed into a drying chamber (furnace) that is heated above the solvent vaporization temperature. The remaining solutes undergo precipitation, thermal decomposition, and intraparticle reactions to form NPs. The solid particles are then collected by an electric field precipitator or a bag filter.

This technique is flexible as it allows the use of a wide range of precursors, solvents, and process conditions and thus provides control over particle size and composition [9].

Bottom-Up: Inert Gas Condensation

Inert Gas Condensation (IGC) process entails the evaporation of a source substance in an inert gas atmosphere to generate the nanopowder which is convectively transported and collected on a cold substrate. The NPs develop in a zone just above the evaporative source, due to interactions between the hot vapor species and the much colder inert gas atoms in the chamber. A feature of the process is the ability to generate non-agglomerated nanopowders, which are sinterable at relatively low temperatures. IGC process is generally classified as a Physical Vapor Deposition (PVD) method, but if a catalytic process is involved in addition, reference is made to it as *Chemical Vapor Deposition (CVD)* method. CVD is one of the most important synthesis processes for fabricating metal, semiconductor, and magnetic NPs. In the CVD process, the reaction or the thermal decomposition of gas phase takes place in a reaction chamber maintained under vacuum at high temperature ($>1,173$ K).

In common CVD reactors, source molecules react and produce both molecular and cluster

Nanoparticles,**Fig. 6** Schematic drawing of a spray pyrolysis system

intermediates in the gas phase. The intermediates become supersaturated and cause the nucleation in the gas phase due to the low vapor pressure. Nucleated NPs collide vigorously with each other and form agglomerated particles [10].

Growth and agglomeration of the particles are mitigated via rapid expansion of the two-phase gas stream at the outlet of the reaction chamber. Subsequent heat treatment of the synthesized nanopowders in various high-purity gas streams allows compositional and structural modifications, including particle purification and crystallization, as well as transformation to a desirable size, composition, and morphology. CVD is widely used to produce carbon nanotubes [11].

Bottom-Up: Liquid Phase Methods

Liquid phase methods include the set of processes that involve a wet chemistry route such as *Hydrothermal (solvo)thermal synthesis*, *sol-gel synthesis*, and *microemulsion synthesis*. The particle formation mechanism is the same as in the vapor phase process. While in gas condensation it is not easy to control grain size and crystal shape, and furthermore its product yield is low, the wet chemistry route provides reasonable control over growth by utilizing organic ligands that act as growth confining agents.

Bottom-Up: Hydrothermal Synthesis

Hydrothermal synthesis of NPs provides precise control over particle size, shape distribution, and crystalline fraction.

This method takes advantage of the solubility of almost all inorganic substances in water at elevated temperatures and pressures and subsequent crystallization of the dissolved material from the fluid.

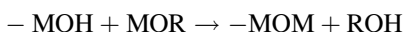
Water at high temperatures plays an essential role in the precursor material transformation because the vapor pressure is much higher and the structure of water at elevated temperatures is different from that at room temperature. The properties of reactants, including their solubility and reactivity, also change at high temperatures. The solvent is not limited to water but includes also other polar or nonpolar solvents, such as benzene, and the process is more appropriately called *solvothermal synthesis* in different solvents [12].

Bottom-Up: Sol-gel Synthesis

The sol-gel process is a recent technique widely used in the field of material science and ceramic engineering for the generation of colloidal NPs from liquid phase. Sol-gel method is well adapted for oxide NPs and

composite nanopowder synthesis starting from a chemical solution (*sol*) which is followed by the formation of a *gel*. It is based on inorganic polymerization reaction including hydrolysis, polycondensation, drying and thermal decomposition.

Typical starting materials are inorganic precursors: metal or nonmetal alkoxides (MOR) dispersed in an appropriate solvent. Precursors hydrolyze with water or alcohols, according to the reaction scheme:



or

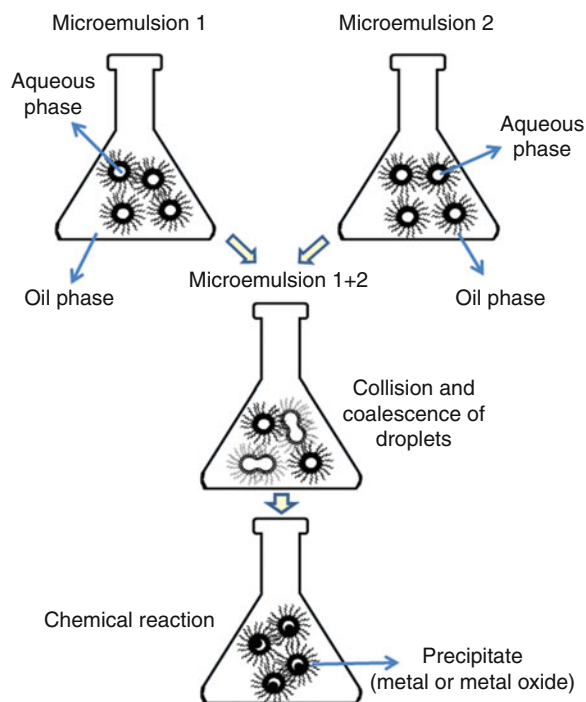


The rates of hydrolysis and condensation are important parameters that affect the properties of final products. Slower and more controlled hydrolysis typically leads to smaller particle size and more unique properties. After solution condensation to gel, the solvent should be removed. This requires a drying process, which is typically accompanied by a significant amount of shrinkage and densification. Generally, higher calcination temperatures are needed to decompose the organic precursors [11].

Bottom-Up: Microemulsion Method

A microemulsion is a thermodynamically stable dispersion of at least three components: two immiscible fluids and a surfactant with amphiphilic properties (stabilizer).

A surfactant is a molecule that possesses both polar and nonpolar moieties. In very diluted water (or oil) solutions, it dissolves and exists as monomer, but when its concentration exceeds a certain minimum critical micelle concentration (CMC), the stabilizer molecules associate spontaneously to form aggregates-micelles [13]. The formation of microdroplets can be in the form of oil-swollen micelles dispersed in the aqueous phase as the oil/water (O/W) microemulsion or water-swollen micelles dispersed in oil as the water/oil (W/O) microemulsion (reverse microemulsion) [12]. The formation of O/W or W/O micelles is driven by strong hydrophobic interactions of the hydrophobic tail of the surfactant molecule (O/W micelle) or by hydrophilic interactions of the polar head of the surfactant molecule (W/O micelle).



Nanoparticles, Fig. 7 Proposed mechanism for the formation of metal particles within micelle “nanochamber”

Micelles in these systems can be described as “nanoreactors” providing a suitable environment for controlled nucleation and growth. Figure 7 illustrates the mixing of two microemulsions containing the appropriate reactants (metal salt and reducing agent) during the collision of the water droplets. This intermicellar interchange process is very fast (from 10 μs to 1 ms) and it is the rate-limiting step for particle growth, but it is slow when compared to diffusion of reagents inside the nanoconfinement (“nanoreactor”). The attractive interaction between the droplets, which is responsible for the percolation process, is of great importance for the kinetics [13].

Characterization of Nanoparticles

NPs are generally characterized in particle size distribution, morphology, charge determination, and surface hydrophobicity. More specific characterizations depend on the properties of NPs synthesized such as magnetic characterizations for NPs featuring magnetic properties (e.g., Fe, Fe_3O_4) and electric characterizations for conductive or dielectric NPs (e.g., Ag, Cu, Au, SiO_2 , BaTiO_3 , etc.). Different parameters versus principal characterization methods used for NPs are given in Table 1.

Nanoparticles, Table 1 Parameters versus principal characterization methods for Nps

Parameters	Characterization methods
Particle size, size distribution, morphology	Scanning Electron Microscopy (SEM), Transmission Electron Microscopy (TEM), Atomic Force Microscopy (AFM), X-ray Diffraction (XRD), Laser diffractometry
Charge determination	Zeta potentiometer
Surface hydrophobicity	Water contact angle measurements, hydrophobic interaction chromatography
Magnetic features	Vibrating Sample Magnetometry (VSM), Alternating gradient force magnetometry (AGFM), Magnetic Force Microscopy (MFM)
Electronic properties	Percolation threshold measure (both current-voltage DC and AC), dielectric characterization

Magnetic Nanoparticles

Magnetic NPs containing pure 3D transition metals – or some mixture of their oxides – are one of the most studied nanomaterials in view of the prospective, ubiquitous applications in quite different areas, the most notable being biomedicine, sensor technology, and magnetic recording [14]. According to their usage, magnetic NPs are often either embedded in a (typically diamagnetic) solid [15], or dispersed in a fluid; in some cases they are surrounded by an outer shell of a diamagnetic material.

This entry deals with particles having sizes not exceeding a few tens of nanometers. The lower limit is taken to be a few (>2) nanometers. As a result, the number of magnetic atoms per NP typically ranges between hundreds and a few millions. Smaller aggregates should be more properly termed magnetic clusters; their magnetic properties are closer to those of magnetic surfaces and are described in a quite different way. On an even smaller scale, magnetic molecules (mainly organic molecules containing a small number, about 10, of transition metal ions) are the subject of much study even if their applicability is still controversial.

Nowadays, magnetic NPs are routinely prepared by a variety of techniques which can be termed as physical (e.g., laser-assisted deposition, sputtering, etc.) or chemical (e.g., wet chemical routes) [14, 16].

The main feature of bona fide NPs made of a particular element or compound is that their magnetic properties are quite different from those of the corresponding bulk material. In a simplified picture, an ideal magnetic NP can be figured out as provided of a highly symmetrical body (e.g., spherical, ellipsoid, acicular, etc.) and characterized by a homogeneous magnetization provided by the ferro- (or ferri-) magnetic alignment of elementary magnetic moments, giving rise to a mesoscopic permanent magnetic moment associated to the particle. Of course this picture can be improved to account for the role played by the NP surface, whose chemical and magnetic properties can be quite different from those at the particle core. Strictly speaking, the locution “core-shell nanoparticle” [17] means a particle whose chemical or stoichiometric properties are substantially different when moving from the center to the surface (e.g., a metallic core surrounded by an oxide shell). In a sense, however, all magnetic NPs fall in this category at least when their magnetic properties are considered, even when surface and core do not substantially differ in their chemistry.

The simplified approach is instrumental in understanding the reason why magnetic properties of NPs change with respect to bulk counterparts. Any macroscopic magnetic material is usually divided into magnetic domains in order to minimize the magnetostatic energy; the number of domains, as well as their shape, is dictated by the interplay and balance between magnetic energies, such as exchange, anisotropy, and domain-wall energy. When the size of a piece of magnetic material is ideally shrunk down to the sub-micrometer range without changing its aspect ratio, all the associated magnetic energies decrease accordingly; however, they decrease following different laws. Both magnetostatic energy and anisotropy energy are volume energies; however, domain wall energy being a surface energy, it has a slower reduction rate with reducing particle size. This makes the nucleation of magnetic domains unfavorable when the size of a particle becomes smaller. Below a critical size which strongly depends on the material’s magnetic properties and morphology, the total energy at remanence is no longer decreased by magnetic domains, and the NP remains in the single-domain state. The magnetic moment points along one of the magnetically “easy” directions corresponding to the minima of the anisotropy energy. In order to have it switched away

from the easy direction, one must apply a magnetic field; on removing it, the magnetic moment will be again aligned along any one of the easy directions. This is the so-called single-domain (magnetically) blocked state [18]. However, on further decreasing the particle size, the anisotropy energy keeps decreasing as the particle volume, and eventually becomes so small that it falls below $k_B T$, where k_B is the Boltzmann's constant and T the absolute temperature, that is, the typical thermal energy steadily exchanged at equilibrium by the NP with the thermal bath. As a consequence, the magnetization is no longer "blocked" along an easy direction, being continuously affected by thermal (random) torques of thermal origin, and the magnetic moment vector behaves in a sense as a free rotator which can be aligned by an external magnetic field, very much as paramagnetic substances do. However, the NP's magnetic moment is usually much larger than that of a single atom or ion, so that it can be termed as a "supermoment," and the particular magnetic behavior of such a NP is known as "superparamagnetism" [19]. The size below which a NP ceases to be in the magnetically blocked state at a given temperature to enter the superparamagnetic (SP) region, and the temperature above which a NP of given size begins to follow the SP behavior are termed critical size (temperature) for the onset of superparamagnetism, respectively. Both depend on NP size, shape, and the leading anisotropy term.

The transition from blocked to SP state is not a thermodynamic phase transition, because the SP behavior emerges only when the magnetic moment's direction randomly fluctuates at a rate high enough that during the time needed to perform a single measurement it virtually takes all directions in space, so that its time average is zero at zero applied field. The observation time plays a crucial role in determining whether a NP at a given temperature is in the SP state or not. From this viewpoint, the blocked-SP transition resembles, in a sense, that between an undercooled liquid and a glass.

SP particles exhibit a reversible magnetization versus field behavior, as paramagnetic substances do. Therefore, no hysteresis loop appears when an alternating magnetic field is applied to a system of genuine SP particles. The material's coercivity and remanence disappear. On the other hand, blocked single-domain NPs usually exhibit coercivity and

remanence values which can be much higher than the ones measured in a macroscopic specimen of the same material. This means that a system containing blocked NPs made of a magnetically soft material acts as a magnetically hard material [18]. Therefore, two quite opposite magnetic behaviors are expected at the same temperature, strictly depending on NP size (and shape) alone.

Magnetic NPs in the SP phase are attractive for biomedical applications, because they can be easily driven by a magnetic field (an oscillating field makes them their "free" moments to oscillate; an inhomogeneous field exerts a force on them, displacing them toward a target point). Oscillating magnetic moments dissipate energy as heat, and may effectively contribute in the local heating of organic tissue, for example, malignant tissue. Magnetic NPs may be properly functionalized with diamagnetic, organic shells in order to act as the portable inner core of a larger, multi-layered particle aimed to a specific intra-body target where it can release specific drugs carried by the organic shell (drug delivery). Other applications of functionalized particles include cell separation and DNA reconnaissance.

Blocked NPs are interesting because they can be exploited as low-cost substitutes of more expensive high-coercivity materials in permanent magnet industry and in magnetic data storage [16]. Hard-disk magnetic memories are often based upon nanoparticulate media where the information can be safely stored owing to the high coercivity of the medium. In this case, superparamagnetism is detrimental because it causes a loss of information. The SP limit introduces a lower boundary to the size of the NPs which a magnetic memory can be composed of, and poses a major, still partially unsolved technological problem in the development of higher-density magnetic memories.

A major drawback of almost all techniques proposed so far to produce magnetic NPs (at least, all of the low-cost ones) is the lack of full control on their size, which results in a (more or less) broad size distribution. Considering how size matters in determining the magnetic properties of these magnetic nanomaterials, it is apparent why many efforts of both fundamental and applied research are now aimed to develop preparation techniques resulting in monodisperse nanoparticle systems.

Application of Nanoparticles

Here it is proposed a selection of the application fields of NPs, in addition to those of magnetic NPs which were summarized in the previous paragraph. As references, the same literature previously cited was kept.

Drinking water purification is a new and promising application, performed making use of Fe, Ni, Au, Ag, or Cu NPs, for example for pesticide removal and organic contaminant detection. The NPs feature an enhanced reactivity and selectivity toward some dangerous elements (such as heavy metals) and organic substances (pesticides); when combined with a MEMS structure, this property is used to detect the chemicals by monitoring certain physical properties of the system (resonance frequency, impedance, etc.).

The enhanced chemical activity of metal NPs is also well known in catalysis, where transition metal NPs are used in the synthesis of organic compounds. Catalytic NPs may be surrounded by a dendritic polymeric structure (dendrimer) in order to allow catalyst recovery at the end of the chemical reaction. A peculiar form of catalysis, called water splitting, is performed using semiconductor NPs (Cu_2O , Fe_3O_4), where sunlight is directly converted in a current (either electrons or holes, depending on the semiconductor properties) and used to split the water molecule into oxygen and hydrogen, to produce an environmentally friendly fuel.

Cellular imaging, intended as optical microscopy performed on cells using fluorescence, Raman spectroscopy and SERS, is normally enhanced by uptake of NPs into the cell body. For this purpose are often used Au and, TiO_2 NPs, and so-called quantum-dots, where this definition includes semiconductor NPs, as CdSe and ZnS.

Photoluminescence of semiconductor NPs (quantum dots) is well known and studied in CdTe, Au/CdTe NPs assemblies, PbS, CdS, etc.

Biomedical applications of noble metal NPs include targeted drug delivery and targeted heat release; similar applications are more commonly approached by using ferromagnetic NPs (e.g., hyperthermia treatment on malignant cells). Also polymeric NPs, which are not a matter of discussion here, are currently used as vectors for drug delivery.

Inhibition of microorganisms and antimicrobial effects are obtained using Ag and/or TiO_2 NPs.

For solar cells optimization noble metal (Au, Ag) NPs have been proposed as electromagnetic field

enhancement media due to their SPR, even though TiO_2 and SiO_2 NPs possess more interesting antireflection properties.

Flexible electronics: Metallic NP tracks (Au, Ag, Cu, Sn), deposited by means of inkjet printing technology, are sintered toward electrical percolation at very low temperatures, allowing the realization of conductive paths on flexible polymeric substrates [20]. Currently, research is going on, considering possible solutions to assemble NPs and realize functional devices.

Besides the direct application of NPs, they are currently added to either metallic or ceramic or polymeric matrices to realize nanocomposite materials featuring enhanced mechanical, electronic, thermal, magnetic, and optical properties, which are not a matter of discussion in this entry.

Cross-References

- ▶ [Chemical Vapor Deposition \(CVD\)](#)
- ▶ [Chitosan Nanoparticles](#)
- ▶ [In Vitro and In Vivo Toxicity of Silver Nanoparticles](#)
- ▶ [Nanostructures for Energy](#)
- ▶ [Physical Vapor Deposition](#)

References

1. ESRF newsletter n°38. December 2003 and references therein
2. Patel, J.K., Patel, D.J., Pandya, V.M.: An overview: nanoparticles. *Intl. J. Pharmaceut. Sci. Nanotechnol.* **1**(3), 215–220 (2008), and references therein
3. Hosokawa, M., Nogi, K., Naito, M., Yokoyama, T.: *Nanoparticle Technology Handbook*, pp. 5–48. Elsevier, Oxford (2007), and references therein
4. Ung, T., Liz-Marzánand, L.M., Mulvaney, P.: Redox catalysis using $\text{Ag}@\text{SiO}_2$ colloids. *J. Phys. Chem. B.* **103**(32), 6770–6773 (1999), and references therein
5. Scientific Committee on emerging and newly identified health risks (SCENIHR): The appropriateness of existing methodologies to assess the potential risks associated with engineered and adventitious products of nanotechnologies (2007) and references therein
6. Munóz, J.E., Cervantes, J., Esparza, R., Rosas, G.: Iron nanoparticles produced by high-energy ball milling. *J. Nanopart. Res.* **9**, 945–950 (2007), and references therein
7. Chen, Y., Pui Li, C., Chen, H., Chen, Y.: One-dimensional nanomaterials synthesized using high-energy ball milling and annealing process. *Sci. Technol. Adv. Mater.* **7**, 839–846 (2006), and references therein

8. Gutsch, A., Krämer, M., Michael, G., Mühlenweg, H., Pridöhl, M., Zimmermann, G.: Gas-phase production of nanoparticles. *KONA* **20**, 24–37 (2002), and references therein
9. Iskandar, F.: Nanoparticle processing for optical applications – a review. *Adv. Powder Technol.* **20**, 283–292 (2009), and references therein
10. Adachi, M., Tsukui, S., Okuyama, K.: Nanoparticle formation mechanism in CVD reactor with ionization of source vapor. *J Nanopart Res* **5**, 31–37 (2003), and references therein
11. Kumar, M., Ando, Y.: Chemical vapor deposition of carbon nanotubes: a review on growth mechanism and mass production. *J. Nanosci. Nanotechnol.* **10**(6), 3739–3758 (2010), and references therein
12. Tavakoli, A., Sohrabi, M., Kargari, A.: A review of methods for synthesis of nanostructured metals with emphasis on iron compounds. *Chem. Pap.* **61**(3), 151–170 (2007), and references therein
13. Capek, I.: Preparation of metal nanoparticles in water-in-oil (w/o) microemulsions. *Adv. Coll. Interf. Sci.* **110**, 49 (2004), and references therein
14. Reiss, G., Hütten, A.: Magnetic nanoparticles: applications beyond data storage. *Nat. Mater.* **4**, 725–726 (2005)
15. Allia, P., Tiberto, P., Coisson, M., Chiolerio, A., Celegato, F., Vinai, F., Sangermano, M., Suber, L. and Marchegiani, G.: Evidence for magnetic interactions among magnetite nanoparticles dispersed in photoreticulated PEGDA-600 matrix. *J. Nanopart. Res.* (2011), in press, doi:10.1007/s11051-011-0249-7
16. Gubin, S.P., Koksharov, Y.A., Khomutov, G.B., Yurkov, G.Y.: Magnetic nanoparticles: preparation, structure and properties. *Russian Chem. Rev.* **74**(6), 489–520 (2005)
17. Iglesias, Ò., Labarta, A., Batlle, X.: Exchange bias phenomenology and models of core/shell nanoparticles. *J. Nanosci. Nanotechnol.* **8**, 2761–2780 (2008)
18. Tannous, C., Gieraltowski, J.: The Stoner–Wohlfarth model of ferromagnetism. *Eur. J. Phys.* **29**, 475–487 (2008)
19. Knobel, M., Nunes, W.C., Socolovsky, L.M., De Biasi, E., Vargas, J.M., Denardin, J.C.: Superparamagnetism and other magnetic features in granular materials: a review on ideal and real systems. *J. Nanosci. Nanotechnol.* **8**, 2836–2857 (2008)
20. Chiolerio, A., Maccioni, G., Martino, P., Cotto, M., Pandolfi, P., Rivolo, P., Ferrero S., Scaltrito, L.: Inkjet printing and low power laser annealing of silver nanoparticle traces for the realization of low resistivity lines for flexible electronics. *Microelectron. Microeng* **88**, 2481–2483 (2011) and references therein

Nanoparticles Toxicity Examined In Vivo in Animals

► [In Vivo Toxicity of Titanium Dioxide and Gold Nanoparticles](#)

Nanopatterned Substrata

► [Nanopatterned Surfaces for Exploring and Regulating Cell Behavior](#)

Nanopatterned Surfaces for Exploring and Regulating Cell Behavior

Jason P. Gleghorn¹ and Celeste M. Nelson^{1,2}

¹Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA

²Department of Molecular Biology, Princeton University, Princeton, NJ, USA

Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
AFM	Atomic force microscopy
ECM	Extracellular matrix
ES	Embryonic stem cell
MSC	Mesenchymal stem cell
RGD	Argine-glycine-aspartic acid
SAM	Self-assembled monolayer

Synonyms

[Nanopatterned substrata](#)

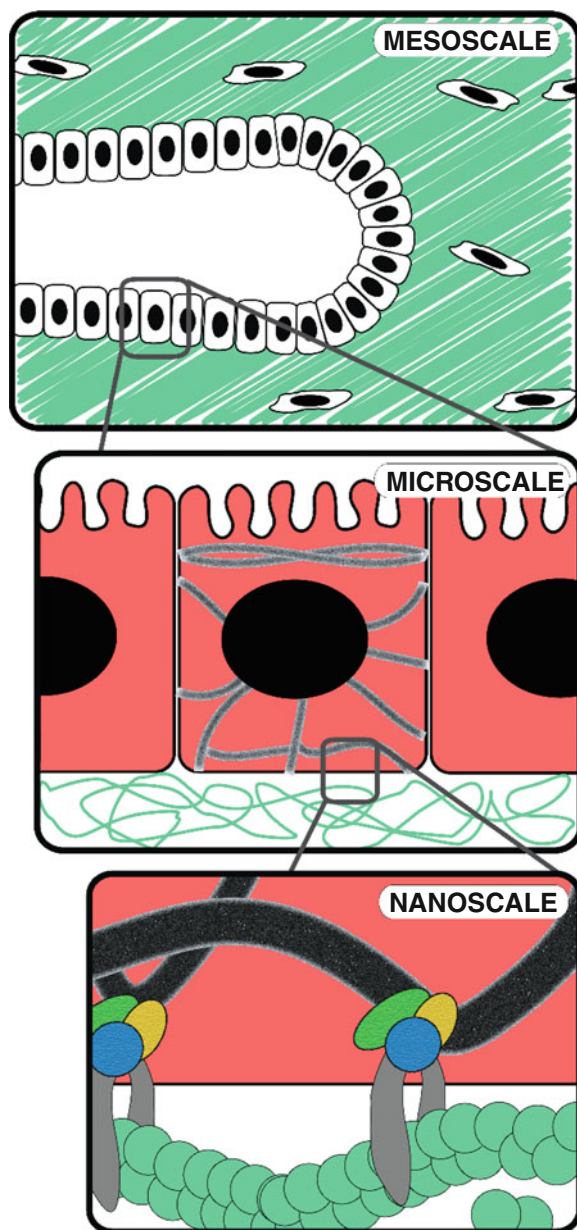
Definition

Surfaces (substrata) that are fabricated to have chemically or topographically nanometer-sized features, which are used to investigate cell morphology, phenotype, and/or behavior.

Overview

Cells Sense and Respond to Their Environment

Cellular environments consist of complex mechanical, chemical, electrical, and topological gradients ranging



Nanopatterned Surfaces for Exploring and Regulating Cell Behavior, Fig. 1 Cell interaction and behavior are governed over multiple length scales. Cell-cell and cell-matrix interactions collectively form tissue architecture and function at the meso-scale. Microscale interactions guide cell shape and connectivity, whereas nanoscale interactions influence cell processes such as adhesion, motility, and gene expression

in size from the nanoscale (1–100 nm) to the macro-scale (>1 mm) (Fig. 1). Cell behavior is governed by these multiscale environmental properties, which, when disrupted, lead to aberrant tissue function and

disease. At the mesoscale (100–1,000 μm), aggregate cell behavior resulting from cell-cell interactions, coupled with properties of the extracellular matrix (ECM), create the tissue architecture and sculpt the mechanical and chemical microenvironment. Heterogeneity of the tissue architecture at the mesoscale produces unique microenvironments. Cell-cell and integrin-dependent cell-ECM interactions at the microscale (1–100 μm) regulate cell shape and adhesion, and consequently affect cell survival, growth, differentiation, and migration. At the nanoscale, molecular and subcellular processes such as integrin activation, focal adhesion formation, actin polymerization, and cytoskeletal organization enable the cell to be physically coupled to its ECM. Additionally, membrane-bound receptors signal through intracellular pathways to modulate gene expression and protein synthesis in response to chemical signals such as morphogens.

Cells interact with and regulate their surroundings in a relationship referred to as dynamic reciprocity [1]. Through dynamic and reciprocal crosstalk, cell behavior is modulated by cell-cell and cell-ECM interactions, and that behavior sculpts and defines the microenvironment. One of the many ways in which the cell senses and reacts to its environment is through integrin-dependent signaling. Transmembrane integrins attach and mechanically couple to several ECM ligands, including the prototypic sequence arginine-glycine-aspartic acid (RGD). Integrin engagement leads to recruitment and clustering to form focal adhesions. Intracellularly, integrins within the focal adhesion associate with several proteins including vinculin, talin, tensin, α -actinin, paxillin, Src, and focal adhesion kinase. These proteins couple the focal adhesion to the actin cytoskeleton. This physical linkage enables the cell to adhere to, move along, and interact mechanically with the ECM. In addition, the focal adhesion activates several kinase cascades that biochemically transduce information about the microenvironment into the nucleus, which leads to changes in gene expression. These gene expression changes can result in the production and secretion of soluble signaling molecules, enzymes, and ECM components that shape and maintain the extracellular environment. The formation, maturation, and disassembly of focal adhesions are not only key for cell spreading and migration, but also appear to be central modulators of many cellular functions including proliferation and differentiation.

Engineering Model Substrata

The tissue microenvironment is complex and the responses of cells to microenvironmental interactions can occur through individual and synergistic mechanisms. As a result, model substrata are often employed to decouple and isolate the effects of various microenvironmental factors. These model surfaces typically fit into one of two categories: chemically or topologically patterned. Chemically patterned surfaces contain regions of active and inactive (inert) biomolecules and thus present spatial patterns of chemical ligands to the cells of interest. ECM proteins such as fibronectin, vitronectin, laminin, and collagen or their active binding motifs are often used. In contrast, topologically patterned surfaces alter the surface topology that is presented to the cells. In these systems, etched pits, posts, grooves, and gratings are used to characterize cell behavior. Regardless of the type of model substratum used, advances in micro- and nanotechnology have enabled the creation of such tailored surfaces to investigate the interactions between cells and their microenvironment at multiple length scales.

Because of the existence of several simple, well-established techniques, micropatterned surfaces have become a popular tool to interrogate cell behavior. Micrometer-scale topology can be created using standard photolithography coupled with traditional etching processes; however, at the micrometer scale most studies have focused on chemically patterned substrata. For chemically patterned surfaces, photolithographic techniques can be used to spatially deposit or remove different materials and thus define regions for chemical functionalization. Microcontact printing is the most common method of microscale chemical functionalization. This technique employs the use of micropatterned elastomeric stamps that are “painted” with the ECM ligand of interest. The coated stamps are then pressed onto a treated surface, which deposits the cell-adhesive ECM ligands onto the surface in the same spatial pattern as that of the stamp.

Micrometer-sized features can be used to control cell position and connectivity; and thus micropatterning is often employed to investigate both mesoscale and microscale cell behavior. Micropatterned surfaces have been used to coculture fibroblasts and hepatocytes to maintain normal hepatocyte phenotype [2]. Additional mesoscale behavior has been investigated with epithelial sheets grown on micropatterned islands of varying size and shape.

Patterns of cell proliferation emerge that result from cell-cell and cell-ECM interactions within the epithelial sheet [3]. Chemical micropatterning has also been used with single cells to identify the role of cell shape in determining cell fate [4]. Geometric control of cell growth and viability was demonstrated by varying the size of chemically patterned islands of ECM.

Although micropatterning is a powerful tool to investigate cell behavior, particularly at the mesoscale, this approach has its limitations. Micropatterning techniques do not afford the ability to control the surface density of the patterned ligand. Additionally, the minimum feature sizes achieved with these techniques do not permit the exposure of multiple different ligands to a single cell. These drawbacks make it difficult to probe cell-ECM interactions at the scale of integrin engagement or focal adhesion formation. Additionally, cells *in vivo* are not exposed to flat two-dimensional (2D) surfaces but rather to a complex topological microenvironment. The ECM consists primarily of interwoven protein fibers ranging from 10 to 300 nm in diameter. Basement membranes consist of complex mixtures of nanoscale pits, pores, and fibers ranging in size from 5 to 200 nm with peaks and valleys of approximately 100 nm in height [5]. The investigation of cell behavior in response to these native nanotopologies cannot be explored using micropatterned surfaces.

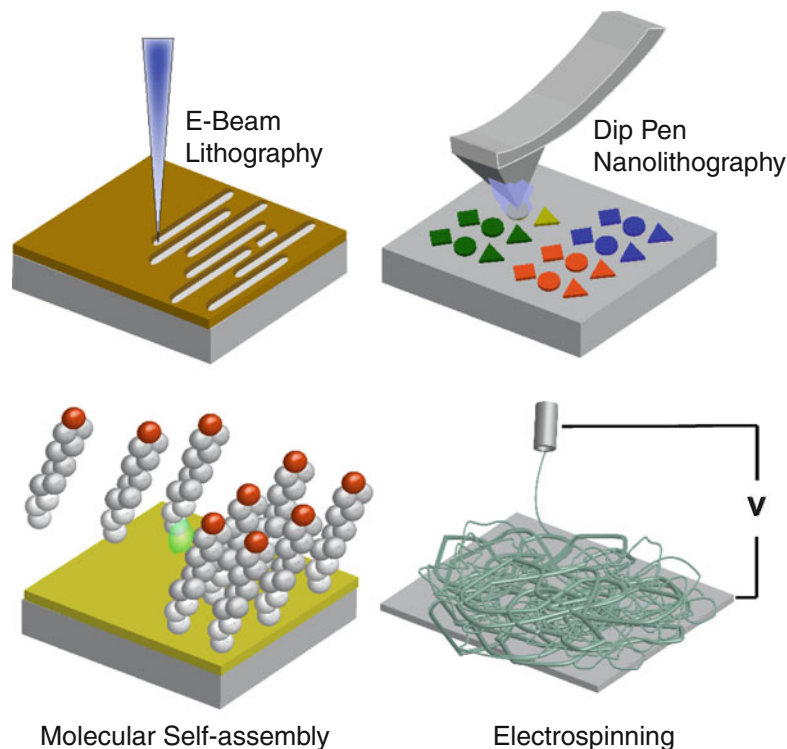
The advent of new tools and fabrication techniques has enabled the creation of spatial patterns in chemistry and topology that can vary greatly over the length scale of a cell. Nanoscale differences in surface roughness, fiber diameter, and ligand spacing can be produced using micropatterned substrata to more closely recapitulate the native tissue microenvironment. Thus, the goal of nanoscale patterning is to investigate molecular mechanisms and subcellular processes that determine fundamental cell behaviors including adhesion, proliferation, gene expression, and differentiation.

Fabricating Nanopatterned Substrata

Unlike micropatterned surfaces, fabrication of nanopatterned surfaces large enough for multicellular studies is technically challenging as well as time and labor intensive. As a result, studies to date have focused primarily on the behavior of single cells. The types of fabrication methodologies discussed herein are by no means comprehensive, but rather

Nanopatterned Surfaces for Exploring and Regulating Cell Behavior,

Fig. 2 Cartoon representations of various fabrication modalities used to create nanopatterned surfaces



a sampling of commonly used techniques to create chemically and topologically defined nanopatterned surfaces. Fabrication technologies adapted from the semiconductor industry as well as novel nanobiotechnology methods are rapidly changing, enabling faster and simpler creation of nanopatterned features. For a more comprehensive review of fabrication techniques, the reader is directed to specialized reviews [6].

Topologically nanopatterned substrata are created using several techniques. These surfaces can be categorized as unordered or ordered topologies. Unordered topologies typically arise spontaneously as a result of processing methods including chemical etching and polymer demixing. Surfaces with unordered topologies have randomly patterned features and lack orientation and geometrical control, but the fabrication processes are relatively simple, fast, and inexpensive. Ordered topologies, on the other hand, are created with precise control of feature pattern, orientation, and geometry. Electron-beam (e-beam) lithography (Fig. 2) is the most commonly used approach to create nanoscale ordered topologies. E-beam lithography uses an electron beam to develop a layer of resist

coated on the substratum. Following e-beam lithography, further processing of the substratum using standard etching techniques results in nanoscale topologies including nanoscale grooves, gratings, or pits.

Dip-pen nanolithography (Fig. 2) is a method to create chemically nanopatterned surfaces using an atomic force microscopy (AFM) tip [7]. The AFM tip is “dipped” into a reservoir, coating the tip with the ligand of interest, and is then used to repeatedly deposit small amounts of the ligand onto the substratum to form chemically patterned regions with an approximate spatial resolution of 5 nm. This technology enables the placement of different adhesive ligands in close proximity in defined shapes with nanoscale precision and dimensions. Additionally, this technology can be multiplexed to create a higher throughput chemical patterning system by combining multiple AFM tips in parallel.

In addition to the “top-down” fabrication strategies used to deposit or remove material from a surface as described above, newer “bottom-up” technologies of molecular self- or templated-assembly (Fig. 2) have been used to create larger chemically nanopatterned surfaces. Molecular self-assembly is a broad category

of techniques wherein (supra)molecular building blocks or colloids are spontaneously self-organized as a result of their interactions to form nanoscale topology and/or spatial patterns of cell-adhesive ligands [8]. Similarly, templated-assembly creates the same types of surfaces with the use of a reusable template to define an organization for the molecular building blocks. These newer fabrication techniques (e.g., self-assembled monolayers (SAMs) and phase-separated block copolymers) have enabled the reliable creation of large patterned areas.

Whereas all of the fabrication techniques described above can be combined to create surfaces with heterogeneously patterned features that vary across length scales, other techniques such as electrospinning [9] do exist to create patterned surfaces that simultaneously expose cells to topologies with widely different dimensions. In electrospinning, an electric field is used to produce polymer fibers from a liquid solution. As the fiber orientation and diameter are easily controlled with simple experimental parameters, fibrous mats containing both nanoscale and microscale fibers can be created with aligned or randomly oriented topologies. Additionally, multiple fibers and complex fiber geometries can be spun simultaneously from different materials or conjugated with different cell-adhesive ligands. This produces substrata that more closely mimic the topology of tissues *in vivo*: complex fibrous meshes with nano- to micro-sized fibers and randomly sized micropores/pits. Thus, electrospun substrata are widely used to determine how cell behavior is regulated by fiber alignment and multiscale distributions of fiber size and surface ligands.

Key Research Findings

Chemically nanopatterned substrata have been primarily used to investigate integrin engagement and focal adhesion formation in an effort to dissect out cell adhesion properties. The majority of these studies have been carried out on self-assembled surfaces containing RGD-functionalized gold nanoparticles [10] (Fig. 3). Integrin receptors are approximately 10 nm wide and these “nanodots” can be fabricated such that they bind to single integrins. As a result, spacing of bound integrins can be altered by changing the spacing of the nanodots on the substrata. These nanopatterned surfaces have revealed that cell

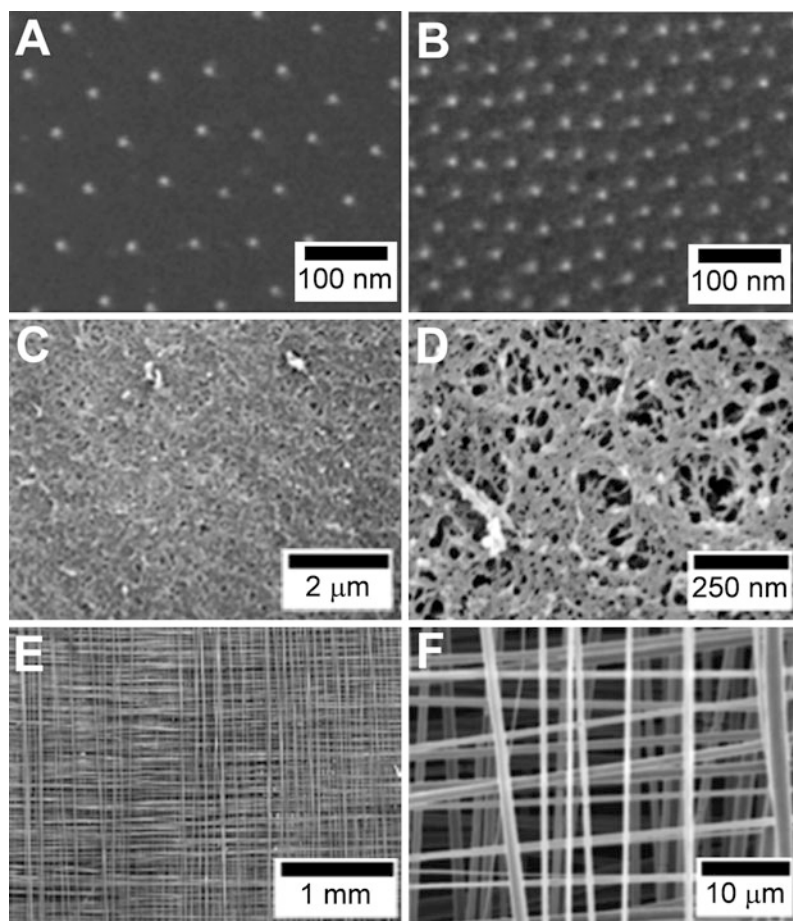
behavior is regulated by the spacing of adhesive ligands. Results from early studies showed that cell motility was affected on surfaces with RGD ligand spacings in the range of 6–300 nm and that spacings from 14 to 25 nm influenced cell-adhesion strength. Subsequent studies explored nanodot spacings between 28 and 85 nm [11] and found that RGD ligand spacings from 58 to 73 nm are optimal for integrin clustering and activation, whereas ligand separation greater than 73 nm causes significantly less cell adhesion, spreading, and integrin-mediated focal contact formation in osteoblasts. Additional studies have measured cell detachment forces from these substrata and demonstrated that spacings greater than 90 nm inhibited focal adhesion formation and decreased detachment forces compared to spacings less than 50 nm [12]. Overall, the data obtained using these chemically nanopatterned surfaces indicate that RGD ligand spacings should be less than approximately 70 nm to stimulate collective cell functions. Additionally, these stimulatory effects tend to increase as the ligand spacing decreases down to approximately 10 nm.

Studies using topologically nanopatterned substrata can be grouped into two categories, ordered or unordered, based on the layout of the surface features. Ordered surfaces consist of repeated features that are arranged with consistent height, spacing, and orientation. Such surfaces include those with uniform posts or pits and gratings formed by regularly spaced grooves. Conversely, surfaces with randomly oriented, varying sized features, such as electrospun meshes, are classified as unordered. The mechanisms by which nanotopographic cues regulate cell behavior are not well understood. The emerging picture is that changes in cytoskeletal organization and structure may drive some of these cell behaviors, potentially in response to the underlying geometry or feature size of the surface. Regardless, it is clear that there is a range of different responses, or differing levels of response that vary by cell type.

Morphology and proliferation are altered when cells are cultured on surfaces with ordered nanotopography [14] (Fig. 4). Human mesenchymal stem cells (MSCs) grown on 350 nm-wide grooves have an aligned cytoskeleton along the direction of the features. Likewise, stem-cell-derived osteoblasts not only aligned and spread in response to nanogrooves of polystyrene, but also showed alignment in actin and

Nanopatterned Surfaces for Exploring and Regulating Cell Behavior,

Fig. 3 Example substrata. (a, b) Ligand-coated gold nanodots can be spaced at varying distances to form chemically nanopatterned substrata (Reproduced with permission from Selhuber-Unkel et al. [12]). (c, d) The basement membrane of the cornea has topographical variations with both micrometer- and nanometer-sized features (Reproduced with permission from Abrams et al. [5]). Electrospun meshes can be formed to create both unordered and ordered topographies (e) and fiber diameters can vary within individual samples (f) (Reproduced with permission from Carnell et al. [13])



mineralized matrix. Additional studies with human embryonic stem cells (ES) cultured on 600 nm polydimethylsiloxane (PDMS) ridges have demonstrated an altered organization of cytoskeletal components including F-actin, vimentin, γ -tubulin, and α -tubulin. The alterations to morphology and proliferation were eliminated when the ES cells were exposed to actin-disrupting drugs [15]. Generally, cells seem to be more sensitive to the depth of the groove as opposed to the groove pitch or spacing. Additionally, cell orientation increases with increasing groove depth on grated surfaces but decreases as the groove width or pitch increases.

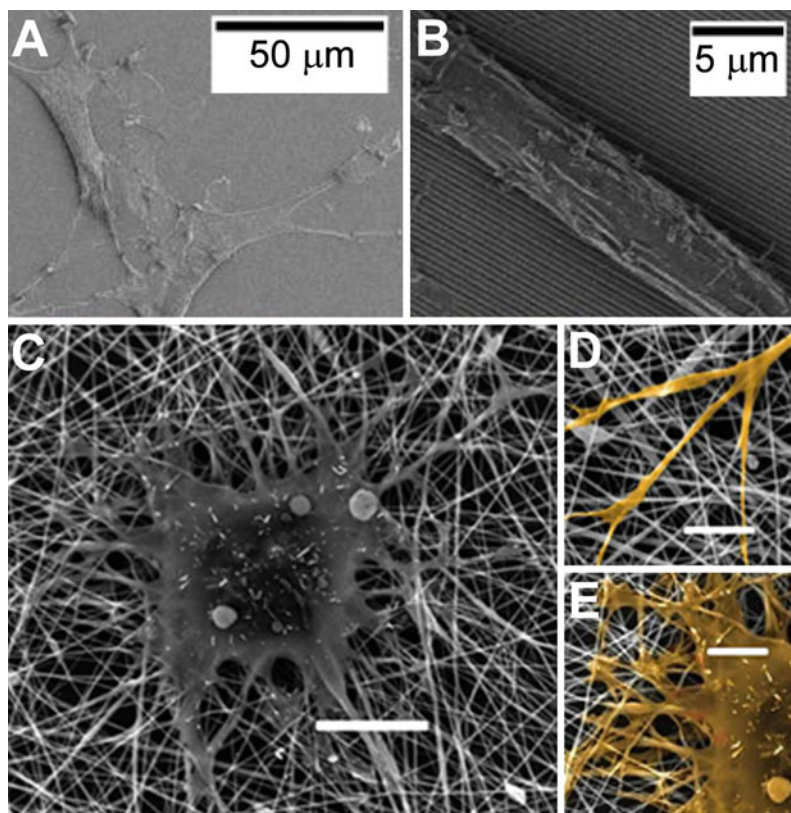
Whereas the phenomenon of cell and cytoskeletal alignment is fairly conserved over a wide range of nanogratings of varying dimensions, cell behavior is more varied in response to ordered surfaces with pits and posts. The adhesion of osteoblasts was significantly increased when cultured on 11 nm-high posts

compared to either 85 nm-high posts or flat polystyrene [16]. Similarly, fibroblasts also demonstrated increased adhesion and spreading on 13 nm posts compared with 95 nm-high posts. However, endothelial cells displayed less spreading and lower cytokine production when cultured on 100 nm-high posts as compared to flat control surfaces. In general, cells cultured on surfaces that are uniformly or randomly textured with nanometer-sized pits and pillars do not display aligned or preferential orientations.

Although comparing different unordered topologies is difficult due to differences in feature size, shape, uniformity, and chemistry, a trend does seem to emerge. In electrospun fiber meshes ranging from 283 to 1,425 nm in diameter, rat hippocampus-derived adult neural stem cells differentiate and proliferate in response to fiber diameter [17]. For fiber meshes smaller than 283 nm, cells did not demonstrate preferential patterns of spreading; however, as the fiber

Nanopatterned Surfaces for Exploring and Regulating Cell Behavior,

Fig. 4 Examples of cell responses to nanotopography. Uniform cell spreading is observed on flat substrata (a), whereas cells spread and align along the direction of nanogrooved substrata (b) (Reproduced with permission from Yim et al. [14]). Cells spread, attaching to and around fibers on an unordered electrospun mesh (c–e). The cell is highlighted in yellow and scale bars are 10 μm (c) and 5 μm (d, e) (Reproduced with permission from Christopherson et al. [17])



diameter increased, cells extended along the length of the fiber. In contrast, proliferation increased as fiber diameter decreased. In another study investigating cell adhesion, fibroblasts demonstrated a higher level of adhesion to surfaces with unordered topology as compared to either flat surfaces or those with ordered topology [18]. In summary, the general trend is that cell behaviors such as proliferation, adhesion, and spreading are positively stimulated as the feature size decreases from ~ 100 to 10 nm.

While almost all of the studies to date have focused exclusively on either chemically or topologically patterned substrata, these parameters are difficult to decouple. Chemically nanopatterned surfaces inevitably have 5–10 nm variations in height across the substratum due to the islands of conjugated ligands. In topographically nanopatterned systems, differences in protein transport and adsorption due to structural features may result in differences in the presentation of adhesive areas and thereby alter cell behavior. Few studies have investigated the competitive or synergistic effects of both chemical and topographic cues [16].

When continuous chemically patterned strips of cell-adhesive ligand were orthogonally overlaid onto nanogrooves, fibroblasts preferentially oriented along the chemical patterns. However, when fibronectin was orthogonally overlaid discontinuously along the top of a nanogroove pattern, osteoblasts preferentially aligned along the topographic pattern. From these limited data it is clear that a variety of systems involving both chemical and topological patterning over varying nanometer-sized scales need to be employed to fully understand interaction effects.

Applications

The basic science studies highlighted herein that seek to understand how nanoscale environmental cues influence cell behavior are fundamental to a range of applications. In particular, these data can be used to engineer biomaterial surface characteristics for improved performance. Materials that interact with the circulatory system, such as those used for stents or

cardiac catheters, could be designed with a nanopatterned surface to prevent cell adhesion and clotting. Conversely, orthopedic implants, which need to integrate strongly with the surrounding tissue, could be designed with a nanopatterned bioactive surface to encourage cell adhesion, proliferation, and motility into the surface. Thus, understanding the nano-environmental cues will enable engineers to recreate physiologically desirable microenvironments for various biomaterial-tissue applications.

Nanopatterning also has applications for diagnostics and cell biology studies such as cell-based biosensors and cell sorting and culture systems. Integration of micro- and nanotechnologies with cell biology has yielded many new and exciting approaches. Use of microfluidics has created cells-on-a-chip or lab-on-a-chip designs where differential cell adhesion based on cell type is critical. Introduction of nanopatterned surfaces can enable a passive physical modification to these systems that would save fabrication time and complexity. Nanopatterned surfaces can be used as passive cell sorting systems that leverage the fact that different cell types respond differently to the same nanotopology. Thus a heterogeneous mixture of cells could be exposed to a nanopatterned surface and only the desired cell population would adhere. Additionally, as understanding of stem cell behavior on nanopatterned substrata grows, one could envision designing substrata that control stem cell adhesion, proliferation, and differentiation – cell behaviors that are difficult to control and that are currently modulated using complex media formulations and cell handling procedures.

Finally, studies with nanopatterned substrata can be integrated into the fields of tissue engineering and regenerative medicine. It is clear that the native in vivo environments and the chemical and physical cues to which cells are exposed are vastly different from those presented by traditional in vitro surfaces such as tissue culture plastic. Examination of signaling cascades within cells on nanopatterned surfaces can reveal how cells detect and respond to chemical and topographic signals in terms of short- and long-term cell functions, such as changes in gene expression and protein synthesis. These details of scaffold architecture and ligand presentation can improve tissue engineering technologies, and provide insight on disease processes where microenvironmental regulation is disrupted, such as in cancer. While some examples of

nanopatterned tissue scaffolds exist, particularly with nanoporous materials for bone tissue engineering, the link between cell behavior in 2D and 3D is not straightforward. Cells in 3D environments adhere differently than those in 2D environments and contain different cell-ECM adhesions [19]. A recent study has even demonstrated that cells cultured on 1D patterns (lines) of adhesive ligands behave more similarly to cells in 3D matrices than to those on 2D chemically patterned islands [20]. Thus, while the link between cell behaviors in 2D and 3D is not certain, it is clear that further understanding gained using well-controlled nanopatterned surfaces can contribute to improved 3D tissue models and translational tissue engineering technologies.

Cross-References

- ▶ [Dip-Pen Nanolithography](#)
- ▶ [Electron-Beam-Induced Deposition](#)
- ▶ [Electrospinning](#)
- ▶ [Nanoimprint Lithography](#)
- ▶ [Self-assembly](#)

References

1. Bissell, M.J., Hall, H.G., Parry, G.: How does the extracellular matrix direct gene expression? *J. Theor. Biol.* **99**(1), 31–68 (1982)
2. Bhatia, S.N., Balis, U.J., Yarmush, M.L., Toner, M.: Probing heterotypic cell interactions: hepatocyte function in microfabricated co-cultures. *J. Biomater. Sci. Polym. Ed.* **9**(11), 1137–1160 (1998)
3. Nelson, C.M., Jean, R.P., Tan, J.L., Liu, W.F., Sniadecki, N.J., Spector, A.A., Chen, C.S.: Emergent patterns of growth controlled by multicellular form and mechanics. *Proc. Natl. Acad. Sci. U.S.A.* **102**(33), 11594–11599 (2005)
4. Chen, C.S., Mrksich, M., Huang, S., Whitesides, G.M., Ingber, D.E.: Geometric control of cell life and death. *Science* **276**(5317), 1425–1428 (1997)
5. Abrams, G.A., Goodman, S.L., Nealey, P.F., Franco, M., Murphy, C.J.: Nanoscale topography of the basement membrane underlying the corneal epithelium of the rhesus macaque. *Cell Tissue Res.* **299**(1), 39–46 (2000)
6. Norman, J.J., Desai, T.A.: Methods for fabrication of nanoscale topography for tissue engineering scaffolds. *Ann. Biomed. Eng.* **34**(1), 89–101 (2006)
7. Lee, K.B., Park, S.J., Mirkin, C.A., Smith, J.C., Mrksich, M.: Protein nanoarrays generated by dip-pen nanolithography. *Science* **295**(5560), 1702–1705 (2002)

8. Zhang, S.: Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* **21**(10), 1171–1178 (2003)
9. Pham, Q.P., Sharma, U., Mikos, A.G.: Electrospinning of polymeric nanofibers for tissue engineering applications: a review. *Tissue Eng.* **12**(5), 1197–1211 (2006)
10. Lim, J.Y., Donahue, H.J.: Cell sensing and response to micro- and nanostructured surfaces produced by chemical and topographic patterning. *Tissue Eng.* **13**(8), 1879–1891 (2007)
11. Arnold, M., Cavalcanti-Adam, E.A., Glass, R., Blummel, J., Eck, W., Kanteleiner, M., Kessler, H., Spatz, J.P.: Activation of integrin function by nanopatterned adhesive interfaces. *Chemphyschem* **5**(3), 383–388 (2004)
12. Selhuber-Unkel, C., Erdmann, T., Lopez-Garcia, M., Kessler, H., Schwarz, U.S., Spatz, J.P.: Cell adhesion strength is controlled by intermolecular spacing of adhesion receptors. *Biophys. J.* **98**(4), 543–551 (2010)
13. Carnell, L.S., Siochi, E.J., Holloway, N.M., Stephens, R. M., Rhim, C., Niklason, L.E., Clark, R.L.: Aligned mats from electrospun single fibers. *Macromolecules* **41**(14), 5345–5349 (2008)
14. Yim, E.K., Pang, S.W., Leong, K.W.: Synthetic nanostructures inducing differentiation of human mesenchymal stem cells into neuronal lineage. *Exp. Cell Res.* **313**(9), 1820–1829 (2007)
15. Gerecht, S., Bettinger, C.J., Zhang, Z., Borenstein, J.T., Vunjak-Novakovic, G., Langer, R.: The effect of actin disrupting agents on contact guidance of human embryonic stem cells. *Biomaterials* **28**(28), 4068–4077 (2007)
16. Lim, J.Y., Hansen, J.C., Siedlecki, C.A., Runt, J., Donahue, H.J.: Human foetal osteoblastic cell response to polymer-demixed nanotopographic interfaces. *J. R. Soc. Interface* **2**(2), 97–108 (2005)
17. Christopherson, G.T., Song, H., Mao, H.Q.: The influence of fiber diameter of electrospun substrates on neural stem cell differentiation and proliferation. *Biomaterials* **30**(4), 556–564 (2009)
18. Curtis, A.S., Casey, B., Gallagher, J.O., Pasqui, D., Wood, M.A., Wilkinson, C.D.: Substratum nanotopography and the adhesion of biological cells. Are symmetry or regularity of nanotopography important? *Biophys. Chem.* **94**(3), 275–283 (2001)
19. Cukierman, E., Pankov, R., Stevens, D.R., Yamada, K.M.: Taking cell-matrix adhesions to the third dimension. *Science* **294**(5547), 1708–1712 (2001)
20. Doyle, A.D., Wang, F.W., Matsumoto, K., Yamada, K.M.: One-dimensional topography underlies three-dimensional fibrillar cell migration. *J. Cell Biol.* **184**(4), 481–490 (2009)

Nano-patterning

► BioPatterning

Nanophotonic Structures for Biosensing

Emiliano Descrovi¹, Mirko Ballarini² and Francesca Frascella¹

¹DISMIC – Dipartimento di Scienza dei Materiali e Ingegneria Chimica, Politecnico di Torino, Torino, Italy

²DIFIS – Dipartimento di Fisica, Politecnico di Torino, Torino, Italy

Synonyms

[Nano-optical biosensors](#); [Nano-optics biosensing](#)

Definition

Nanophotonic structures for biosensing: one-dimensional, two-dimensional, or three-dimensional structures made of suitable materials, having dimensions ranging from few nanometers to few hundreds of nanometers, said structures showing a well-defined interaction with light providing an optical transduction mechanism for detecting/revealing specific biomolecules in close proximity to the structure itself.

Overview

Optical biosensors [1] constitute powerful detection and analysis tools with wide applications in the biomedical domain. They can provide parallel detection within a single device. Generally speaking, there are two detection methods that are implemented in optical biosensing: fluorescence-based detection and label-free detection. Thanks to the recent advances of technological capabilities for the fabrication of nanometer-sized structures, most of the conventional techniques can be improved, and new original techniques are become feasible.

In fluorescence-based detection, target molecules are labeled with fluorescent tags, typically organic dyes and/or chemically functionalized/decorated quantum dots; the overall intensity of fluorescence indicates the presence of the target molecules. Some quantification can also be inferred, but in general quantitative analysis is challenging due to the fluorescence

signal bias, as the number of fluorophores on each molecule cannot be precisely controlled. While fluorescence-based detection is extremely sensitive, with the detection limit down to a single molecule, it suffers from laborious labeling processes that may also interfere with the biological/chemical function of the biomolecule and/or its interaction with the environment.

In label-free techniques, target molecules are not labeled or altered, and are detected in their natural forms. This type of detection is relatively easy and cheap to perform, and allows for kinetic measurement of molecular interaction.

Despite all these differences between fluorescence-based and label-free detection, both protocols can benefit of the use of proper nanostructures and provide complementary information regarding interactions among biomolecules. In this framework, the use of dielectric, metallic, and metallo-dielectric structures obtained by means of a number of methods, for example, optical, electronic or ionic lithographic processes, physical/chemical deposition/sputtering, chemical synthesis, or self-assembling can improve both fluorescence-based and label-free biosensing. In the following, a broad variety of structures are described with particular focus to their respective main applications, but additional original combinations are still under investigation.

Label-free detection

Label-free techniques include refractive index (RI) detection, optical absorption detection, and Raman spectroscopic detection. In a typical label-free biosensing application a mean should be provided in order to uniquely identify the desired target biomolecules. There are basically two strategies that are implemented: a chemical functionalization of the optically sensitive area of the photonic nanostructure and a spectroscopic measurement. In the first case, the specificity of the photonic biosensor is demanded to the chemical functionalization, therefore allowing the specific detection of desired target molecules on the unique molecular recognition mechanism embodied by the chemical functionalization. In the second case, the recognition of the target molecules is performed by detecting the corresponding spectroscopic fingerprint (either the IR absorption or the Raman emission spectra). Unfortunately, the low Raman activity of the majority of biomolecules makes them little suitable for Raman analysis. Furthermore, a detectable Raman

spectrum cannot be attained by simply increasing the power of the laser excitation radiation, because irreversible damages to the biological matter can be induced. Photonic metallic nanostructures producing strong field enhancements by means of Localized Surface Plasmons (LSP) can overcome this limitation, thus allowing a Surface-Enhanced Raman Scattering (SERS) or a Tip Enhanced Raman Scattering spectroscopic detection. Nanostructures sustaining LSP can also provide for Surface-Enhanced Infrared Absorption (SEIRA) spectroscopic detection.

Surface-enhanced Raman spectroscopy (SERS) was originally discovered in the 1970s where it was found that submonolayers of small organic molecules, when adsorbed onto the surface of silver nanoparticles (and a few other metals such as copper and gold) would exhibit greatly enhanced Raman intensities (enhancements of 10^6) [2]. These results have always held the promise for using this technique to observe very low concentrations of molecules on nanoparticles and nanostructured surfaces, but only recently has this promise started to be fulfilled in a predictable way.

Thanks to exciting advances in techniques for making nanoparticles (such as nanosphere e-beam lithography and Focused Ion Beam), to characterize the surfaces using electron and scanning probe microscopies, to functionalizing the surfaces of the particles using self-assembled monolayers with attached chemical receptors, and to the laser and optics technology associated with measuring the Raman spectra, a number of important applications have been reported recently. SERS has almost exclusively been associated with three metals, silver (by far the most important), gold, and copper. The generalization to other metals and other materials has been explored since the discovery of SERS, but only in the last few years have the tools been available to generate well-characterized experiments where the reported enhancements are reliable.

Light incident on the nanoparticles induces the conduction electrons in them to oscillate collectively with a resonant frequency that depends on the nanoparticles' size, shape, and composition. As a result of these LSPR modes, the nanoparticles absorb and scatter light so intensely that single nanoparticles are easily observed by eye using dark-field (optical scattering) microscopy. The shape of the nanoparticle extinction and scattering spectra, and in particular the peak wavelength λ_{\max} , depends on nanoparticle

composition, size, shape, orientation, and local dielectric environment [3]. The LSPR can be tuned during fabrication by controlling these parameters with a variety of chemical syntheses and lithographic techniques. Although silver and gold are the most commonly used materials, LSPR is theoretically possible in any metal, alloy, or semiconductor with a large negative real dielectric constant and small imaginary dielectric constant.

When molecules adsorb onto a plasmonic nanoparticle or move to within a few nanometres of its surface, the local electromagnetic fields around the nanoparticle can enhance the Raman scattering by a factor of 10^6 – 10^8 for an ensemble of molecules and by as much as 10^{14} – 10^{15} for single molecules. Two principal enhancement mechanisms are generally thought to explain the large SERS signals. First, an electromagnetic enhancement factor arises because LSPR modes in the metal nanoparticles focus the incident light energy at the nanoparticle surface and also increase the density of states at Stokes-shifted wavelengths. The electromagnetic enhancement factor is typically 10^5 – 10^8 for nanoprisms, but simulations indicate that it can be as large as 10^{13} for structures such as arrays of nanoprism dimers. Second, chemical enhancement factors arise from changes in the molecular electronic state or resonant enhancements from either existing molecular excitations or newly formed charge transfer states.

Wet chemical synthesis methods have now made it possible to fabricate plasmonic nanoparticles having a variety of shapes (for example spheres, triangles, prisms, rods, and cubes) with controllable sizes and narrow size distributions. Furthermore, metallic–dielectric, and metallic–metallic core–shell nanoparticles and mixed metallic–alloy nanoparticles with different shapes (for example, nanoshells and nanorice) have been prepared. Fabrication of this large variety of different structures makes a variety of applications possible, as the spectral position of the surface plasmon resonance depends on both the shape and the size of the nanoparticle [4].

An alternative possibility of exploiting the localized near-field generated in close proximity to metallic nanostructures is to employ a metallic nanoprobe or tip where LSP are coupled. Such a structure allows a well localized source of highly confined electromagnetic field to be raster scanned on the surface of a given sample, in such a way that a high spatial resolution Raman mapping of the surface can be performed [5].

Tip Enhanced Raman Scattering (TERS) has been observed for a number of molecules adsorbed at various substrates, including single-crystalline metal surfaces, showing thereby a more than million-fold enhancement of the Raman scattering [6].

In order to excite LSPs at the apex of a metallic nanoprobe, a further structuring can be performed. For example, gratings can be lithographed on the side surface of a gold cantilevered tip to ease the excitation of plasmon polaritons upon external lateral illumination of the tip. Alternatively, resonant nanoantennas can be fabricated directly on the top of the tip.

An effect quite similar to SERS occurs in the mid-infrared region: molecules on metal surfaces show infrared absorption 10–1,000 times more intense than would be expected from conventional measurements without the metal. This effect is referred to as Surface-Enhanced Infrared Absorption (SEIRA) to emphasize the analogy to SERS [7]. Vacuum-evaporated thin metal films are used most frequently for SEIRA experiments. Thin metal films that show strong SEIRA are not continuous but consist of metal islands smaller than the wavelength of the incident light. The average dimension of the islands is few tens of nanometers. The density, shape, and size of the islands depend on the mass thickness, the evaporation conditions, and the chemical nature of the substrate. As the mass thickness increases, the islands grow in size, contact each other, and eventually form a continuous film. Connection of the islands significantly reduces the enhancement, suggesting that the small metal islands play an important role in the enhancement. The metal islands are polarized by the incident photon field through the excitation of collective electron resonance, or localized plasmon, modes, and the dipole p induced in an island generates a local electromagnetic field stronger than the incident photon field around the island. Experimental results suggest that the collective electron resonance is excited even in the mid-infrared region and that the enhanced electromagnetic field enhances the infrared absorption of adsorbed molecules.

As previously stated besides those label-free techniques based on spectroscopic measurements, there are quantitative methods based on the measurements of refractive index changes close to the nanostructured surface of the photonic biosensor. Among these, the Surface Plasmon Resonance (SPR) technique and the waveguide- (or optical fiber-) based technique play a major role [8].

SPR is based on the coupling of propagating or stationary plasmons on silver/gold/copper thin films. The potential of surface plasmon resonance (SPR) for characterization of thin films and monitoring processes at metal interfaces was recognized in the late seventies. Generally, an SPR optical sensor comprises an optical system, a transducing medium which interrelates the optical and (bio)chemical domains, and an electronic system supporting the optoelectronic components of the sensor and allowing data processing. The transducing medium transforms changes in the quantity of interest into changes in the refractive index which may be determined by optically interrogating the SPR. The sensor sensitivity, stability, and resolution depend upon properties of both the optical system and the transducing medium [9].

SPR can be used in a prism-coupler configuration or with a grating-coupler configuration. If a metal-dielectric interface is periodically distorted, the incident optical wave is diffracted forming a series of beams directed away from the surface at a variety of angles. If the total component of momentum along the interface of a diffracted order is equal to that of the plasmon, the optical wave may couple to the plasmon. Grating-based optical SPR sensors have been demonstrated which use the measurement of the light intensity variations at SPR and the wavelength interrogation. A high-sensitivity SPR grating-based gas sensor using silver as an SPR active metal has attained sensitivity $1,000 \text{ nm RIU}^{-1}$ in wavelength interrogation mode, in angular interrogation mode, the system's sensitivity would be about $100^\circ \text{ RIU}^{-1}$. Gold-based SPR grating sensors have been used for monitoring biomolecular interactions in aqueous environments, with estimated refractive index sensitivity of $30^\circ \text{ RIU}^{-1}$ and $900\% \text{ RIU}^{-1}$ in the angular interrogation and intensity measurement modes, respectively. The grating coupler has been also used for the excitation of an SPW in an optoelectronic SPR-sensing device based on a Schottky-barrier semiconductor structure. Periodic and quasiperiodic structuring of smooth metal films sustaining plasmons can be used to create Bragg mirrors (Stationary plasmons) and Fabry-Perot cavities.

Electromagnetic modes propagating at the interface between a homogeneous medium and a truncated periodic structure like a dielectric monodimensional photonic crystal (1DPC), also referred to as Bloch Surface Waves (BSW), have been proposed as an alternative to

surface plasmon polaritons (SPP). The possibility to excite simultaneously several different BSW modes in the same 1DPC allows for self referencing sensing [10]. In analogy to surface plasmons on smooth thin films, a surface periodic patterning of 1DPC sustaining BSWs have been performed in order to fold BSW dispersion curves, open photonic band gaps in the dispersions, and further enhance the sensitivity of BSW upon surface perturbations.

Dielectric micro and nanostructures can be used for biosensing based on either refractometric or absorption measurements. Such sub-wavelength structures include Bragg mirrors, gratings, microcavities, microrings and optical fibers and waveguides based on Photonic Crystals (PC). Provided that some of the photonic structures mentioned above produce strong field localization that can be exploited for field-enhancement effects, a large portion of their application consists in measuring shifts of the resonance conditions (light incident angle and/or wavelength) upon a perturbation of the refractive index of the external medium, containing the target biomolecules [11, 12].

Bragg mirrors consist of a periodic stack of dielectric materials acting like a lossless mirror for a given wavelength band. Periodic multilayers can be fabricated, for example, by sputtering, thermal evaporation, and Chemical Vapor Deposition (CVD) techniques. When one or more defect layers (spacers) are introduced therein, a defect (cavity) mode is obtained. In such a simple case, the cavity is a one-dimensional Fabry-Perot cavity that has been widely used for refractometric measurements. If the material is porous in nature, such as porous silicon electrochemically etched, target molecules can penetrate inside the porous photonic structure (pores size for few nanometers to tens of nanometers), thus providing a much stronger perturbation and lowering the detection threshold. If fluorescent labels are used, one-dimensional microcavities can also be used to improve the detection of fluorescence radiation, as described in the following.

Photonic crystal surfaces can be designed to provide a wide range of functions that are used to perform biochemical and cell-based assays. Detection of the optical resonant reflections from photonic crystal surfaces enables high sensitivity label-free biosensing, while the enhanced electromagnetic fields that occur at resonant wavelengths can be used to enhance the detection sensitivity of any surface-based fluorescence

assay. Fabrication of photonic crystals from inexpensive plastic materials over large surface areas enables them to be incorporated into standard formats that include microplates, microarrays, and microfluidic channels. The periodic modulation of refractive index within a PC, along with incorporation of intentional line or point “defects” in the PC can be used to concentrate and direct the electromagnetic fields associated with light to produce efficient wavelength selective reflectors, waveguides, optical circuits, beam steering devices, and optical multiplexors. PCs can be designed to interact strongly with particular optical wavelengths through selection of their materials and the period of their modulation. At the wavelengths of “optical resonance” (also known as “guided mode resonance”), light will couple strongly to the PC structure for a particular incident angle, resulting in electric fields inside the PC that can be many times higher than the electric field of incident radiation. These optical resonances can be observed by an external observer simply by illuminating the PC at normal incidence with a broad band of and observing a narrow band of wavelengths that are reflected back with nearly 100% efficiency.

Two-dimensional photonic crystals used for sensing applications are typically microfabricated periodic nanostructures on a two-dimensional planar silicon on insulator (SOI) chip. Such structures can be designed to exhibit a two-dimensional photonic band gap (PBG) that can be tuned by varying the structural parameters. Photonic cavity states within this PBG are formed by systematically introducing defects into the perfect hexagonal lattice during the fabrication process. In the case of nanocavities, the modal volume is very small, of the order of $V < 1 \mu\text{m}^3$ and furthermore, the strongly localized electric field in the cavity is highly sensitive to local changes of the refractive index of the environment.

The photonic crystal concept is fruitfully applied to optical fibers, in which the transverse cross section of the fiber is patterned with channels specifically arranged, wherein the refractive index contrast provided by the nanostructuring results in multiple effects such as optical field localization, multimodal guiding, etc.

By combining the technology available for waveguide fabrication and the working principle of Fourier-Transform-based spectrometers, a suitable microdevice based on a high-density sample of the

standing-wave interferogram produced inside a single waveguide has been obtained [13].

All the above mentioned techniques can be combined for obtaining devices and structures aimed at biosensing based on one or more of the described principles. For example, optical fibers presenting metal coatings or hosting metallic nanostructures can combine guiding and plasmonic effects, fiber grating can ease light coupling/decoupling, or resonance conditions, etc.

Fluorescence-based detection

As already mentioned, dielectric micro and nanostructures can provide a mean to improve the detection of labeled target molecules in close contact or impregnated within the photonic nanostructure. An example of such a structure is represented by one-dimensional porous silicon microcavities [14], in which the fluorescence emitted by labeled molecules (e.g., proteins) is enhanced upon diffusion of said molecules inside the porous structures of the cavity [14]. Additional one-dimensional, two-dimensional, and three-dimensional (e.g., artificial opals) dielectric structures can be exploited for such enhancement effects, reaching very high Q factor of the order of 10^3 or more [15].

In recent years, there has been a growing interest in the interactions of fluorophores with metallic surfaces or particles. Near-field interactions are those occurring within a wavelength distance of an excited fluorophore. The spectral properties of fluorophores can be dramatically altered by near-field interactions with the electron clouds present in metals. These interactions modify the emission in ways not seen in classical fluorescence experiments and can be controlled by properly structuring the surface of plasmonic substrates [16].

While all metal–fluorophore interactions are based on the same physics [17], the effects can be different based on the geometry of the metal structure. Three possibilities are considered, in which a fluorophore is interacting with silver particles, a smooth metal surface and a metal surface with a regular pattern. Metal particles can be typically used to increase the fluorescence intensities. This increase occurs by a combination of enhanced fields around the metal and rapid and efficient plasmophore emission. These effects are usually called metal-enhanced fluorescence (MEF), and typically result in increased intensities and decreased lifetimes.

In the case of a fluorophore interacting with a smooth metal film, it is typically about 40 nm thick silver or gold. In this case, the fluorophore creates plasmons which radiate at a defined angle into the substrate. Typically the intensities and lifetimes are not dramatically changed. This phenomenon is called surface plasmon-coupled emission (SPCE). Here, the emission spectrum is the same as the fluorophore but the polarization properties indicate that the plasmon is radiating. Finally, an exemplary fluorophore can be placed over regular patterns (nanorings, grooves, gratings, etc.). In this case, the emission at certain wavelengths is expected to show well-defined beaming into the substrate (or air), while other wavelengths are deflected from the normal. These effects are due to a combination of interactions with a smooth surface and with the sub-wavelength features (according to a diffraction phenomenon). This more general effect is called plasmon-controlled fluorescence (PCF). Plasmon-controlled fluorescence provides an opportunity to create ultra-bright probes based on complexes of metal particles with fluorophores. The basic idea is to create metal-fluorophore complexes in which the metal particle increases the brightness of the bound fluorophores. This can be accomplished by trapping the fluorophores in metal shells or by coating metal particles with fluorophores [18]. Theory has predicted that fluorophores in metal shells can be 100-fold brighter than the isolated fluorophore, even after consideration of the transfer efficiency of incident light into the shell and radiation out of the shell. Such probes may have the advantage over quantum dots because of the low toxicity of silver particles, even if the silver surface is completely exposed.

Instead of organic fluorophores such as dyes, solid state semiconductor nanoparticles (quantum dots) can be used. The luminescent nanostructures are becoming more and more popular in labeling bio-assay [19].

Cross-References

- ▶ [Active Plasmonic Devices](#)
- ▶ [Biosensors](#)
- ▶ [Chemical Vapor Deposition \(CVD\)](#)
- ▶ [Detection of Nanoparticulate Biomarkers in Blood](#)
- ▶ [Field Electron Emission from Nanomaterials](#)
- ▶ [Microfabricated Probe Technology](#)
- ▶ [Nanostructured Materials for Sensing](#)

- ▶ [Nanostructures for Photonics](#)
- ▶ [Optical Properties of Metal Nanoparticles](#)
- ▶ [Optical Techniques for Nanostructure Characterization](#)
- ▶ [Polymer Coatings](#)
- ▶ [Scanning Near-Field Optical Microscopy](#)
- ▶ [Self-Assembled Monolayers](#)

References

1. Fan, X., White, I.M., Shopova, S.I., Zhu, H., Suter, J.D., Sun, Y.: Sensitive optical biosensors for unlabeled targets: a review. *Anal. Chim. Acta* **620**, 8–26 (2008)
2. Kneipp, K., Moskovits, M., Kneipp, H.: *Surface Enhanced Raman Scattering*. Springer, Heidelberg (2006)
3. Anker, J.N., Hall, W.P., Lyandres, O., Shah, N.C., Zhao Van Duyne, J.R.P.: Biosensing with plasmonic nanosensors. *Nat. Mater.* **7**, 442–453 (2008)
4. Lal, S., Link, S., Halas, N.J.: Nano-optics from sensing to waveguiding. *Nano Lett.* **1**, 641–648 (2007)
5. Steidtner, J., Pettinger, B.: Tip-enhanced Raman spectroscopy and microscopy on single dye molecules with 15 nm resolution. *Phys. Rev. Lett.* **100**, 236101 (2008)
6. Alkire, R.C., Kolb, D.M., Lipkowski, J., Ross, P.N.: Tip-enhanced Raman spectroscopy – recent developments and future prospects. In: Pettinger, B. (ed.) *Advances in Electrochemical Science and Engineering: Diffraction and Spectroscopic Methods in Electrochemistry*, vol. 9. Wiley, Weinheim (2008)
7. Osawa, M.: Surface-enhanced infrared absorption. In: Kawata, S. (ed.) *Near-Field Optics and Surface Plasmon Polaritons*. Springer, Heidelberg (2001)
8. Le Ru, E.C., Etchegoin, P.G.: *Surface-Enhanced Raman Spectroscopy and Related Plasmonic Effects*. Elsevier, Amsterdam (2009)
9. Homola, J.: *Surface Plasmon Resonance Based Sensors*. Springer, Berlin (2006)
10. Konopsky, V.N., Alieva, E.V.: A biosensor based on photonic crystal surface waves with an independent registration of the liquid refractive index. *Biosens. Bioelectron.* **25**, 1212–1216 (2010)
11. Zourob, M., Lakhatakia, A.: *Optical Guided-Wave Chemical and Biosensor II*. Springer, Heidelberg (2010)
12. Arregui, F.J., Matias, I.R., Claus, R.O.: Fiber optic biosensors. In: *Encyclopedia of Optical Engineering*. Taylor & Francis, London (2004)
13. le Coarer, E., et al.: Wavelength-scale stationary-wave integrated Fourier-transform spectrometry. *Nat. Photon.* **1**, 473–478 (2007)
14. Sciacca, B., Frascella, F., Venturolo, A., Rivolo, P., Descrovi, E., Giorgis, F., Geobaldo, F.: Doubly resonant porous silicon microcavities for enhanced detection of fluorescent organic molecules. *Sens. Act.* **B 137**, 467–470 (2009)
15. Dorfner, D., Zabel, T., Hürlimann, T., Hauke, N., Frandsen, L., Rant, U., Abstreiter, G., Finley, J.: Photonic crystal nanostructures for optical biosensing applications. *Biosens. Bioelectron.* **24**, 3688–3692 (2009)

16. Lakowicz, J.R.: Plasmon-controlled fluorescence: a new paradigm in fluorescence spectroscopy. *Analyst* **133**, 1308–1346 (2008)
17. Lakowicz, J.R.: *Principles of Fluorescence Spectroscopy*, 3rd edn. Springer, New York (2006)
18. Burns, A., Ow, H., Wiesner, U.: Fluorescent core–shell silica nanoparticles: towards “Lab on a Particle” architectures for nanobiotechnology. *Chem. Soc. Rev.* **35**, 1028–1042 (2006)
19. Viste, P., Plain, J., Jaffiol, R., Vial, A., Adam, P.M., Royer, P.: Enhancement and quenching regimes in metal – semiconductor hybrid optical nanosources. *ACS Nano* **4**, 759–764 (2010)

Nano-piezoelectricity

► [Piezoelectric Effect at Nanoscale](#)

Nano-pillars

► [Nano-sized Nanocrystalline and Nano-twinned Metals](#)

Nanopolaritonics

Daniel Neuhauser¹, Christopher Arntsen¹ and Kenneth A. Lopata²

¹Department of Chemistry and Biochemistry, UCLA, Los Angeles, CA, USA

²William R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA

Synonyms

[FDTD](#); [Maxwell–Bloch simulations or equations or studies](#); [Maxwell–Schrödinger simulations \(or equations\)](#); [Molecular plasmonics](#); [Near-field molecular optics](#)

Definition

Molecular nanopolaritonics is the study of nanoscale control of near-field energy transport via combined plasmon–molecule excitations.

Introduction

Absorption and emission of light (from near-IR to UV) on the nanoscale is qualitatively associated with one of two types of entities: excitons and plasmons. Excitons are associated with transition of electrons between occupied and unoccupied orbitals, i.e., a particle-hole excitation. Plasmons, on the other hand, are classical-like collective excitation of electrons in metals that are analogous to tidal waves. Specifically, when an electric field excites a plasmon, all the valence electrons in the metal shift, in tandem, creating a surface dipole; the remaining background exerts a restoring force and the resulting oscillation is in the IR to UV range, depending on geometry.

The coupling of an exciton and a plasmon is known as a polariton. Polaritons are usually associated with surface excitons. They have been applied in various fields, and at low temperatures even exhibit Bose–Einstein condensates [1]. Here, the combination of excitons and plasmons on the nanoscale is discussed, where the plasmon is part of a relatively small metal structure and even more importantly, the exciton is associated with a molecule or a cluster. Such a combination is labeled as nanopolariton [2], and its properties and approaches for simulation are discussed here. The greatest emphasis here is on the manipulation of transport of nanopolaritons by varying the molecular properties, as detailed below.

Broadly speaking plasmon–molecule interactions come in two forms: the influence of plasmonic fields on molecules, and the influence of molecules on plasmonic fields. The first, which is the typical focus of plasmon–molecule studies, is where plasmons focus light into molecules driving them into highly excited states, such as in surface-enhanced Raman spectroscopy (see SERS – Raman Spectroscopy). SERS relies on the fact that classically, an electric field near a metal surface is very strong, especially near the plasmon resonance. In SERS, the massive electric fields due to plasmons on a metal lead to huge enhancements in Raman scattering, allowing for the detection of single molecules [3]. Other examples of plasmons acting on molecules include plasmon-enhanced fluorescence, near-field optical microscopy, and enhancements in FRET [4].

The second direction, where single molecules have a significant effect on plasmons, is more subtle and less understood. Molecules have a much smaller transition

dipole moment than metal nanostructures (as they have few electrons, versus hundreds to million and more electrons in metal structures) and therefore molecules will generally have a negligible effect on overall light absorption of a polaritonic system, except possibly at very low temperatures. However there is one place where plasmonic interactions are weak: in the transport of plasmons from one structure to another. As this transport involves comparatively weaker coupling between plasmon modes, under certain circumstances it can be influenced by nearby molecules. While this article focuses primarily on molecule-mediated plasmon transport, linked DNA molecules have been shown to alter the dielectric properties of arrays of metal nanoparticles [5].

The length scales relevant for nanopolaritonics are significantly smaller than even UV wavelengths. At such scales, the electromagnetic fields are essentially “near fields,” i.e., the field due to a dipole is closer to its electrostatic form, so it is strong and decays rapidly with distance; the far-field-radiated part, proportional to the acceleration of the dipole, is much weaker over these scales (the ratio between the far and near fields for a distance scale R is approximately R^2/λ^2 , where λ is the wavelength). The simulation of the near-field combined plasmon–molecule excitations (polaritons) involve a complicated interplay of distinct mechanisms, including quantum excitations, classical electrodynamics, and charge transfer, which involve a range of length and energy scales. This makes concise treatment of these systems very difficult: molecules on the order of a nanometer must be treated quantum mechanically, whereas plasmonic materials with sizes of up to hundreds of nanometers are usually too large to treat quantum mechanically in a computationally efficient way, and thus must be treated classically in most cases. Therefore, a primary thrust of theoretical molecular nanopolaritonics is and will be the development of methods which can effectively treat and provide an interface between these classical (metal) and quantum (molecular) regions.

Nanopolaritonics: Methodology

Overview

There are several techniques used to model behavior in nanopolaritonic systems, but in general they involve coupling classical fields to a quantum mechanical

molecule. Some examples are described in greater detail below. Briefly, the simplest method evolves the spatial charge density of the metal in time and couples plasmons to a molecular density matrix via dipole interactions. A more sophisticated treatment models classical regions with the finite-difference time-domain (FDTD) method (see the entry in this volume on ► [Finite-Difference Time-Domain Technique](#)), and couples the classical electric field to the molecular density matrix again via dipole interactions. This particular technique has the advantage that the classical description can easily be interfaced with essentially any type of quantum mechanical treatment of a molecule, including time-dependent density functional theory (TDDFT), time-dependent Hartree–Fock (TDHF), or semiempirical methods, as described later.

Plasmon Surface Modes Coupled to Electronic Density Matrices

In the simplest treatment, appropriate for molecules near metallic structures, a set of modes is first used to represent the plasmons via surface modes on the metals [6]. The charge density on the metals is written as

$$\sigma(\mathbf{r}, t) = \sum_J C_J(t) s_J(\mathbf{r}), \quad (1)$$

where $s_J(\mathbf{r})$ is a surface-localized charge mode on a metal particle (e.g., dipole, quadrupole, etc., for a spherical metal particle, or a more general function in other cases), and the time-dependent coefficients which determine the evolution of the surface modes were introduced. The dynamics of these surface modes is associated with plasmons; the collective oscillation of the charges causes a time-dependent surface charge.

The surface charges are then coupled via dipole interaction to specific molecules; the resulting equation for the motion of the modes, including interaction with other modes and with specific molecules is then [2]:

$$\begin{aligned} \sum_K M_{JK} \ddot{C}_K(t) = & - \sum_K V_{JK} C_K(t) - \sum_K M_{JK} \frac{\dot{C}_K(t)}{\tau_K} \\ & - \sum_{ij} \rho_{ij}(t) q_{ij,J} - v_J(t). \end{aligned} \quad (2)$$

The first three terms are purely plasmon related, including mass terms (M_{JK}) associated with the kinetic

energy of the surface modes, potential terms (V_{JK}) associated with coulomb interactions between the surface modes, and Drude damping terms ($1/\tau_K$) responsible for the dephasing and damping of the surface modes. The next term contains a dipole integral $q_{ij,J}$ coupling the plasmon modes to the molecule; the molecular density is represented as a density matrix with a constant (pre-interaction) part, ρ_{ij}^0 , and a time-dependent part induced by interaction with the plasmons, $\rho_{ij}(t)$. Finally, the last term is the interaction with the external field, $v_J(t)$.

The other relevant equation is the time-dependence of the density matrix for the molecules. For each molecule, schematically, apply the Liouville equation:

$$i \frac{d\rho_{ij}(t)}{dt} = [F(\rho), \rho]_{ij} - i\gamma_{ij}(\rho_{ij}(t) - \rho_{ij}^0) + \left[\sum_J C_J(t) q_J, \rho \right]_{ij}, \quad (3)$$

where the following variables are introduced: the Fock matrix ($F(\rho)$), which depends itself on the density matrix; the damping coefficients (γ_{ij}) which determine the return of ρ to the equilibrium density matrix, ρ_0 ; and the coupling to the classical fields of the plasmons via the dipole coefficients $q_{ij,J}$. Details of the different possibilities for a Fock operator are presented later.

The simplest way to view Eqs. 2–3 is when the excitation is not very large, as then they can be linearized in the density matrix. Then, the linearized equations have a simple form:

$$i \frac{d\mathbf{f}}{dt} = \mathbf{L}\mathbf{f}, \quad (4)$$

where \mathbf{f} is the nanopolariton, i.e., the combined vector describing both the plasmon modes and the density matrix. Schematically, \mathbf{f} is made from four parts: the surface charge coefficients and their time derivatives, as well as the real and imaginary part of the excitation of the density matrix away from the ground state:

$$\mathbf{f}^T = \left(C_J, \frac{dC}{dt}, \text{Re}(\rho - \rho^0), \text{Im}(\rho - \rho^0) \right). \quad (5)$$

In the generalized Liouville coupling matrix, L is straightforwardly shown to be

$$L = \begin{pmatrix} 0 & -\mathbf{M}^{-1} & 0 & 0 \\ -\mathbf{V} & -\tau^{-1} & -q & 0 \\ 0 & 0 & -\gamma & \Delta \\ -q & 0 & -\Delta - \mathbf{Z}^{2B} & -\gamma \end{pmatrix}.$$

The specific terms in the Liouville matrix result from a linear-response analysis of Eqs. 2 and 3.

Maxwell–Schrödinger Simulations

The plasmonic-mode model is appropriate for qualitative studies, especially because of its very modest computational requirements. However, a quantitative investigation requires a more realistic treatment of the electric field distribution. The best-known approach for simulating the electromagnetic fields is FDTD, where one propagates the Maxwell equations [7]. Assuming for simplicity no magnetic susceptibility effects, these read:

$$\begin{aligned} \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} &= \frac{1}{\epsilon_{\text{eff}}(\mathbf{r})} (\nabla \times \mathbf{H}(\mathbf{r}, t) - \mathbf{J}(\mathbf{r}, t)) \\ \frac{\partial \mathbf{H}(\mathbf{r}, t)}{\partial t} &= -\frac{1}{\mu_0} \nabla \times \mathbf{E}(\mathbf{r}, t), \end{aligned} \quad (6)$$

where the electric and magnetic fields and the frequency-independent part of the susceptibility are introduced.

The current, J , is made from two parts, $\mathbf{J} = \mathbf{J}_p + \mathbf{J}_m$; the first term is the usual plasmonic contribution, which in FDTD is simulated as a sum of terms,

$$\mathbf{J}_p(\mathbf{r}, t) = \sum_n \mathbf{J}_{p,n}(\mathbf{r}, t), \quad (7)$$

where the individual terms fulfill

$$\frac{d\mathbf{J}_{p,n}(\mathbf{r}, t)}{dt} = \alpha_n(\mathbf{r}) \mathbf{J}_{p,n}(\mathbf{r}, t) + \beta_n(\mathbf{r}) \mathbf{E}(\mathbf{r}, t). \quad (8)$$

The sum typically includes 1–5 terms, and the individual terms ($\alpha_n(\mathbf{r})$ and $\beta_n(\mathbf{r})$) in the sum are usually fitted to give the correct dielectric function in the frequency range of interest. Other, somewhat more sophisticated time evolution techniques for Eq. 8 have also been suggested.

The second part of the current is the molecular contribution, which for a single molecule will be

$$\begin{aligned} \mathbf{J}_M(\mathbf{r}, t) &= \delta(\mathbf{r} - \mathbf{r}_M) \frac{d}{dt} \text{Tr}(\rho \mathbf{x}) \\ &= -i\delta(\mathbf{r} - \mathbf{r}_m) \text{Tr}([F, \rho] \mathbf{x}) \end{aligned} \quad (9)$$

where \mathbf{x} is the electronic 3-D coordinate, and the delta function in the position of the molecule is represented numerically on the FDTD grid. The molecular current is obtained from a time-dependent study of the density matrix (or alternatively from propagating the molecular orbitals). In the simplest version, where only a dipole electric field is assumed, these take a form similar to Eq. 3 but now the coupling is directly due to the electric field:

$$i \frac{d\rho_{ij}(t)}{dt} = [F_{\text{tot}}(\rho), \rho]_{ij} - i\gamma_{ij}(\rho_{ij}(t) - \rho_{ij}^0), \quad (10)$$

where

$$F_{\text{tot}}(\rho) = F(\rho) - \mu \cdot \mathbf{E}. \quad (11)$$

Molecular Simulation Methodologies

Throughout, the molecules are simulated in real time. For the present purposes, the molecules will be treated by a Liouville equation or a time-dependent orbital equation, where the initially occupied orbitals are propagated as:

$$i \frac{\partial \psi_n}{\partial t} = F(\rho) \psi_n. \quad (12)$$

There are different possible Fock operators; the most established are TDDFT and TDHF. Both methods, however, could be numerically expensive, so a simple alternate approach is to propagate in time a semiempirical approach; an example is TD-PM3 (time-dependent Parameter-Method 3) where the Fock operator has the generic TDHF form:

$$F_{ij} = H_{0,ij} + \sum_{kl} V_{ijkl} \rho_{kl}, \quad (13)$$

but with a very sparse set of matrix elements, V_{ijkl} . TDPM3 generally gives excitation energies with a $\sim 15\%$ accuracy [8], so is very suitable for the

qualitative questions which arise in the study of nanopolaritonics. As the method uses a minimal basis set and scales well with system size, it is ideal for simulating large dye molecules.

Splitting Fields

It turns out that when there is one molecule, the delta function associated with having a localized charge is problematic and leads to numerical singularities. The reason is that this strong delta function source due to the molecule acts back on the molecule, creating an artificial self-energy problem. More precisely, for a single electron this is a self-energy problem as the electron interacts with itself; for a general molecule, the interaction with the molecule-induced current induces a double-counting problem, since the Fock operator already includes the electron–electron interaction and the interaction with the molecular current has essentially the same role.

This difficulty was alleviated through the development of a new formalism, in which the electric field is divided into two parts, a molecular and a plasmonic part [9]. The molecule is only affected by the external part, while the total field acts on the plasmon. Specifically, for the electric field the resulting equations are:

$$\begin{aligned} \frac{\partial \mathbf{E}_p}{\partial t} &= \frac{1}{\epsilon_{\text{eff}}} \nabla \times \mathbf{H} + \left(\frac{1}{\epsilon_{\text{eff}}} - \frac{1}{\epsilon_0} \right) \nabla \times \mathbf{H}_m \\ &\quad - \frac{\mathbf{J}_p}{\epsilon_{\text{eff}}}, \end{aligned} \quad (14)$$

$$\frac{\partial \mathbf{E}_m}{\partial t} = \frac{1}{\epsilon_0} (\nabla \times \mathbf{H}_m - \mathbf{J}_m), \quad (15)$$

while the magnetic fields equations are simpler

$$\frac{\partial \mathbf{H}_p}{\partial t} = -\frac{1}{\mu_0} \nabla \times \mathbf{E}_p, \quad (16)$$

$$\frac{\partial \mathbf{H}_m}{\partial t} = -\frac{1}{\mu_0} \nabla \times \mathbf{E}_m. \quad (17)$$

As mentioned, only the electric field that acts on the molecule enters into the evolution equation for the density matrix. Only the plasmonic field affects now the molecule, so that the Fock operator for the evolution of the density matrix is now:

$$F_{\text{tot}}(\rho) = F(\rho) - \mu \cdot \mathbf{E}_p. \quad (18)$$

The advantage in this approach is that the decoupling prevents numerical singularities. The disadvantage is that a separate electric field is required for each molecule. Therefore, future approaches will mix the FDTD treatment of the plasmonic field with a dipole and quadrupole treatment of the molecules, i.e., write the total electric field as

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_p(\mathbf{r}, t) + \sum_j \mu(\mathbf{r}, t') \nabla \frac{1}{|\mathbf{r} - \mathbf{r}_j|} + \text{Quad.}, \quad (19)$$

where “Quad.” refers to quadrupolar interaction with nearest neighbors, and the molecular dipole is introduced. The magnetic field is written similarly in terms of the molecular current and the plasmonic field, and the equation of evolution of the total field can be derived directly from Eq. 14. This mixed field-dipole approach will allow the treatment of a host of molecules interacting with an external field.

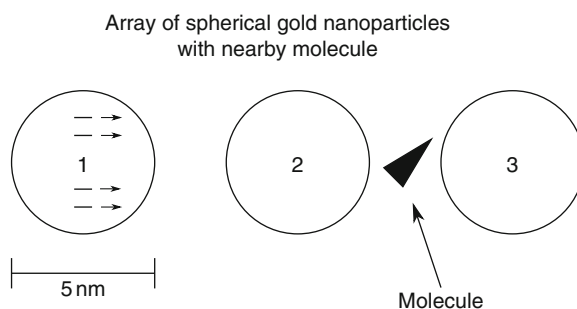
In the extreme limit that a continuum of molecules are required, the FDTD equations will be solved directly; then, the effect of each molecule will be relatively small so that the separation between plasmonic and molecular fields will not be needed, and the resulting, single electric field equations are the familiar Maxwell–Bloch equations, i.e., there will be no need to split the electric fields.

Key Findings

The fundamental conclusion from plasmon mode and Maxwell–Schrödinger (Sect. Nanopolaritronics: Methodology) simulations is that under certain circumstances single molecules have a marked effect on the near-field transport of plasmons, as detailed below.

Molecule-Mediated Plasmon Transfer in Waveguides

An illustrative example of the type of system studied is shown in Fig. 1, where a single two-level dipolar molecule is located near a linear array of gold nanoparticles. In the absence of a molecule, an initial x -polarized plasmon excitation on the leftmost (input) nanoparticle will propagate down the array, which after a few femtoseconds results in an x -polarized plasmon on the rightmost (output) nanoparticle. Since



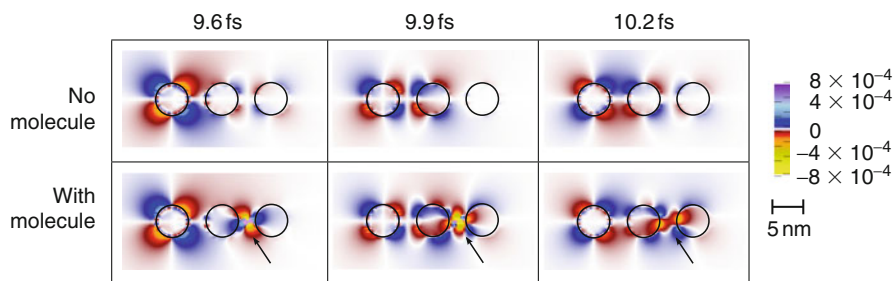
Nanopolaritronics, Fig. 1 Schematic of the computed setup. Three small metal spheres are placed close to each other. The leftmost sphere is excited; the middle sphere mediates this transport. The molecule is placed between the middle and the final, rightmost sphere, and mediates this transport. The direction of the molecule influences the final polarization of the molecule

the x , y , and z plasmon modes are decoupled in this linear configuration, no y or z excitation is induced on the output nanoparticle. Noble metal plasmonic waveguides of this type have been extensively studied both theoretically and experimentally [10].

Perhaps surprisingly, the addition of a single molecule to the waveguide can have a drastic effect. Take, as a simple example, an xy -oriented two-level molecule with a strong resonance at $\omega_m = 2.55$ eV, which is close to the plasmon frequency of the array (2.60 eV). Figures 2 and 3 show results of a FDTD/linear-response TDDFT simulation. Figure 2 shows the clear transmission effect of a molecule in a time-domain simulation; the presence of the molecule causes significant enhancement in the amount of transmitted light in the nonnative initial polarization.

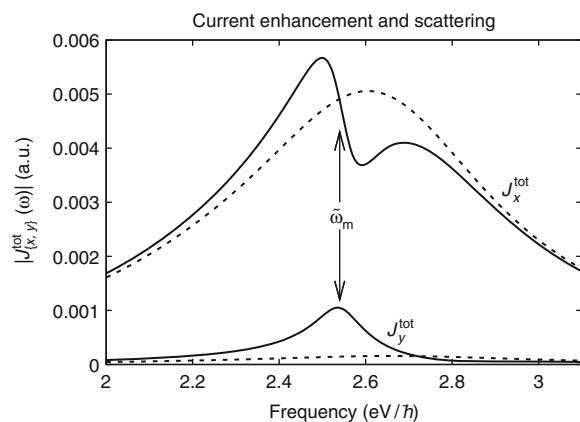
The frequency-domain picture is even more revealing. Figure 3 shows that the Fourier transform of the total current on the output nanoparticle displays a Fano-like resonance. As a reminder, Fano resonance are generally the result of coupling of single (or a few, in the generalized case) “dark” quantum level(s) to a continuum of states. Their most obvious mark is a sine-like modification of the observable (spectrum or transmission profile) which in the original Fano resonance case will actually lead to a dip of zero absorption due to destructive interference caused by the presence of the quantum level.

In the present simulation case, a clear Fano-like transmission profile results due to the presence of the molecule; the excitation of the x -polarization is drastically modified. Specifically, the molecule enhances the x -polarized plasmon transmission over a range of



Nanopolaritonics, Fig. 2 Three snapshots in time of the y -component of the electric field with (*bottom*) and without (*top*) a molecule, for an excitation starting on the left sphere. The

dipolar molecular field serves to scatter x -oriented energy (not shown) into y -oriented energy (From Ref. [9])



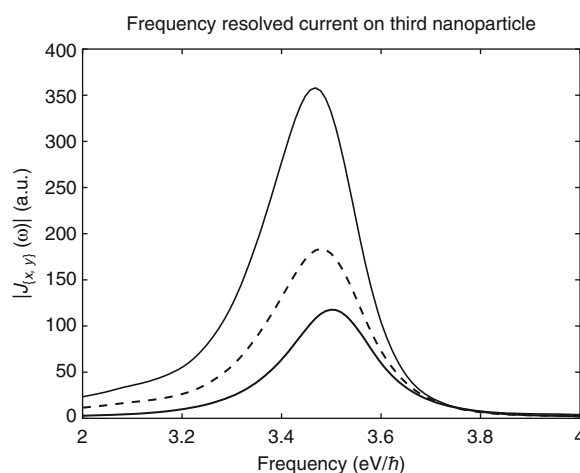
Nanopolaritonics, Fig. 3 Fourier transform of the currents on the output nanoparticle of a linear waveguide, showing that a single two-level xy -oriented molecule rotates incident plasmon polarization from $x \rightarrow y$ (From Ref. [9])

frequencies below ω_m and decreases the transport for a range above. Further, the molecule rotates the excitation from the x -polarization to the y -polarization, resulting in a significant y -current (approximately 20% of the x -current). Physically, the molecule, which has a transition dipole moment aligned along the xy axis, scatters x -polarized near-field light and reemits y -polarized light.

Despite the fact that the molecule does not affect the overall absorption of the array, the intense molecule–plasmon resonance in the gap between nanoparticles strongly influences plasmon transport, to the point of significantly scattering plasmons between otherwise orthogonal modes of excitation.

Realistic Simulations

To study this problem in a more rigorous manner, i.e., by using a more realistic quantum-chemistry



Nanopolaritonics, Fig. 4 Frequency-resolved current on the third silver nanoparticle of the system shown in Fig. 1 with 2-nm nanoparticles with (*bold lines*) and without (*thin line*) a molecule in the x -direction (*solid*) and y -direction (*dashed*) (From Ref. [11])

treatment, the FDTD approach has also been combined with a time-dependent semiempirical method, TDPM3, mentioned above, to show rotation of current for large molecules with a significant number of possible excitations [11]. The system examined used silver nanoparticles, and thus tartrazine, a yellow dye molecule having a strong transition dipole moment near the plasmon resonance of silver, was chosen. The realistic simulations verify that molecules with an excitation energy near the plasmon resonance of the metal scatter the x -oriented current into the y -direction, as observed with the two-level molecule. The results are shown in Fig. 4. These results are encouraging as they confirm that even a molecule with a broad range of excitations can influence the transport of plasmons.

These studies offer insight into near-field light transport in heterogeneous nanosystems. Single molecules, although too weak to play a role in overall light absorption, can have drastic effects on transport if strategically placed in “hot spots,” where they can modulate the comparatively lower energy coupling between nearby metal structures.

Since this is a resonance effect, it is highly dependent on the materials involved. The molecule, in order to scatter current, must have a strong optical transition at or near the coupling (i.e., plasmon resonance) of the waveguide. On one hand, this limits the diversity of the systems that can be used; however, this specificity is actually advantageous for sensing applications and tunable frequency-dependent wave guiding.

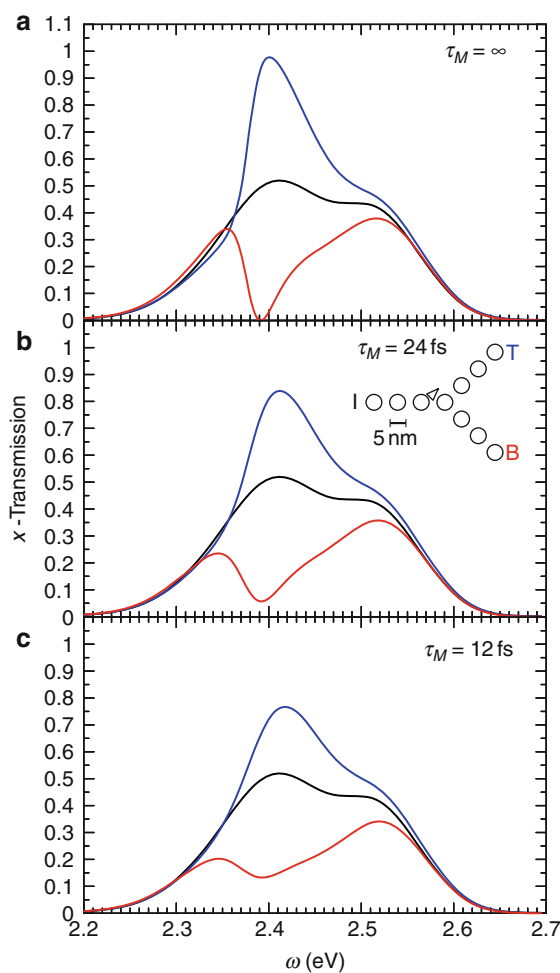
Controlling Plasmon Transport with Single Molecules

The strong effect of a single molecule on nanopolariton transport can be exploited to selectively control plasmon transport on the nanoscale. This has implications in the field of plasmonics, where molecule–plasmon interactions have promise as device components such as switches or higher harmonic generators. A molecule placed near a fork-like arrangement of gold nanoparticles, e.g., can selectively divert incoming plasmon waves into either the top or bottom paths, as demonstrated in Fig. 5, which shows transmission through the fork calculated using the plasmon mode approach. Beyond being sensitive to the molecule’s location, this effect is also highly dependent on its orientation: When the molecule points upward, the plasmons are directed into the upper branch of the fork (*blue curve* in Fig. 5) much more than to the lower branch (*red curve*). In the absence of a molecule (*black curve*), the transition probabilities of the two branches are equal.

The controllability can be extended to use in molecular switches: Molecules can potentially be used to gate the transfer of near-fields in a specific and deliberate manner. Likewise, a similar configuration can act as a type of highly specific molecular sensor, sensitive to a molecule’s identity, orientation, and position.

Limitations

There are several important caveats to these findings. First, compared to metals, molecules have a much smaller capacity to absorb and radiate light, as they are unable to further absorb light (of a given frequency) once already in their excited state. FDTD/



Nanopolaritonics, Fig. 5 Simplified-model study of control of radiation transfer through a fork by placing a molecule aligned parallel to the top branch of the fork. *Blue* top branch, *red* bottom, *black* either branch without a molecule. The frames differ by the amount of dephasing on the excited state of the molecule. The gold dots have the experimentally fitted dephasing time, ~ 6 fs (From Ref. [2])

TDDFT simulations of nanopolaritonic systems which treat the molecule beyond linear response have shown that under increasing field intensities the polarization rotation effects saturate rapidly; this transition region is especially interesting, since it can be used as a nonlinear gating mechanism.

Future Directions

Nanopolaritonics is still in its infancy. Although a variety of simulations have demonstrated proof of

concept for controlling near-field light with molecules, there is much untapped potential for novel plasmon–molecule applications, as well as many outstanding theoretical and experimental questions. This last section briefly overviews future directions for both experiments and theory.

Transport-Based Sensing

Among the most promising potential applications is transport-based sensing, where extremely dilute concentrations of molecules are detected and uniquely identified based on how they alter the plasmon transport properties of a nanostructure system. In contrast to a SERS-like device, which is used to enhance and probe the response of a molecule, here the plasmon transport is measured, which is predicted to be extremely sensitive to the presence of select molecules. A pair of nanoscale metal tips, e.g., can act as a sensor for molecules which drift between them. Alternatively, one can build a parallel detector using a chip containing a series of different plasmonic junctions, each with a slightly different plasmon resonance, which are tuned to different target analyte molecules and excited by different frequencies of light.

Plasmonic Metamaterials

Another potential application of nanopolaritonics is in the development of more sophisticated metamaterials. Beyond simply mediating wave guiding, molecules could also mediate coupling between elements of a plasmonic metamaterial. This could lead to, e.g., tunable metamaterials where the presence of different molecules shifts the resonant frequency of the device.

Need for Experimental Studies

There is a pressing need for high-quality experimental measurements of plasmon transport in the presence of strongly interacting molecules. Whereas there have been a wealth of experiments concerning the effects of intense fields on single molecules, studies on the effects of molecules on plasmons are lagging.

Improved Multiscale Simulation Methodologies

From a theoretical standpoint, progress needs to be made on efficiently and accurately capturing plasmon–molecule interactions, especially in the “chemical” regime where molecules are bound to the surface of a metal and share electrons. In reality both charge and energy are shared on a surface, and

developing real-time embedding approaches capable of describing this is necessary to drive the field forward. This is a daunting task as it necessitates treating the quantum mechanical and classical regions as open systems. Moreover, seamless integration of the two regions requires careful choice of either boundary conditions or embedding potentials.

Cross-References

- ▶ [Finite-Difference Time-Domain Technique](#)
- ▶ [Nanoparticles](#)

References

1. Kasprzak, J., Richard, M., Kundermann, S., Baas, A., Jeambrun, P., Keeling, J.M.J., Marchetti, F.M., Szymanska, M.H., Andre, R., Staehli, J.L., Savona, V., Littlewood, P.B., Deveaud, B., Dang, L.S.: Bose-Einstein condensation of exciton polaritons. *Nature* **443**, 409–414 (2006)
2. Neuhauser, D., Lopata, K.: Molecular nanopolaritonics: Cross manipulation of near-field plasmons and molecules. I. Theory and application to junction control. *J. Chem. Phys.* **127**, 154715 (2007)
3. Nie, S., Emory, S.R.: Probing single molecules and single nanoparticles by surface-enhanced Raman scattering. *Science* **275**, 1102–1106 (1997)
4. Govorov, A.O., Lee, J., Kotov, N.A.: Theory of plasmon-enhanced Förster energy transfer in optically excited semiconductor and metal nanoparticles. *Phys. Rev. B.* **76**, 125308 (2007)
5. Lazarides, A.A., Schatz, G.C.: DNA-linked metal nanosphere materials: Fourier-transform solutions for the optical response. *J. Chem. Phys.* **112**, 2987–2993 (2000)
6. Prodan, E., Nordlander, P.: Plasmon hybridization in spherical nanoparticles. *J. Chem. Phys.* **120**, 5444–5454 (2004)
7. Gray, S.K., Kupka, T.: Propagation of light in metallic nanowire arrays: Finite-difference time-domain studies of silver cylinders. *Phys. Rev. B.* **68**, 045415 (2003)
8. Bartell, L.A., Wall, M.R., Neuhauser, D.: A time-dependent semiempirical approach to determining excited states. *J. Chem. Phys.* **132**, 234106 (2010)
9. Lopata, K., Neuhauser, D.: Multiscale Maxwell–Schrödinger modeling: A split field finite-difference time-domain approach to molecular nanopolaritonics. *J. Chem. Phys.* **130**, 104707 (2009)
10. Maier, S.A., Atwater, H.A.: Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures. *J. Appl. Phys.* **98**, 011101 (2005)
11. Arntsen, C., Lopata, K., Bartell, L.A., Wall, M.R., Neuhauser, D.: Modeling molecular effects on plasmonic transport: Silver nanoparticles with tartrazine. *J. Chem. Phys.* **134**, 084101 (2011)

Nanopores

- ▶ [Nanochannels for Nanofluidics: Fabrication Aspects](#)

Nanopowders

- ▶ [Nanoparticles](#)

Nanoribbons

- ▶ [Graphene](#)

Nanorobotic Assembly

- ▶ [Robot-Based Automation on the Nanoscale](#)

Nanorobotic Manipulation of Biological Cells

- ▶ [Robot-Based Automation on the Nanoscale](#)

Nanorobotic Spot Welding

Lixin Dong¹, Xinyong Tao², Zheng Fan¹, Li Zhang³, Xiaobin Zhang⁴ and Bradley J. Nelson⁵

¹Electrical and Computer Engineering, Michigan State University, East Lansing, MI, USA

²College of Chemical Engineering and Materials Science, Zhejiang University of Technology, Hangzhou, China

³ETH Zurich, Institute of Robotics and Intelligent Systems, Zurich, Switzerland

⁴Department of Materials Science and Engineering, Zhejiang University, Hangzhou, China

⁵Institute of Robotics and Intelligent Systems, ETH Zurich, Zurich, Switzerland

Synonyms

[Nanobonding](#); [Nanointerconnection](#); [Nanojoining](#); [Nanosoldering](#)

Definition

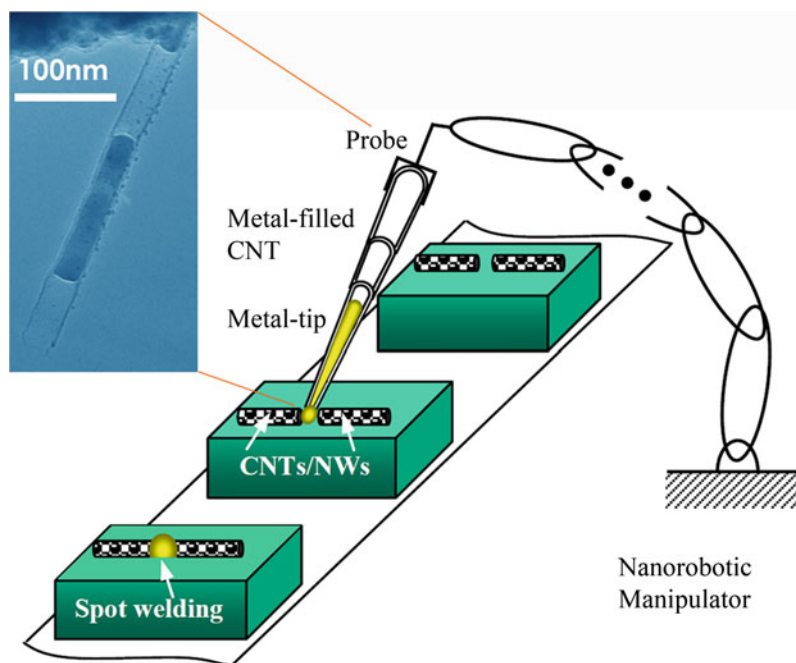
Spot welding, a widely used resistance welding process, was originally developed by Elihu Thompson in the early 1900s [1]. The most common application of spot welding is in the automobile industry, where it is widely used and most often performed automatically by preprogrammed industrial robots on assembly lines. With the continuing development of bottom-up nanotechnology fabrication processes, the nanoscale version of robotic spot welding, or nanorobotic spot welding [2], may likewise play an important role in interconnecting carbon nanotubes (CNTs), nanowires, nanobelts, nanohelices, and other nanomaterials and structures for the assembly of nanocircuits and nanoelectromechanical systems (NEMS).

Overview

Van der Waals forces [3], electron-beam-induced deposition (EBID) [4], focused-ion-beam chemical vapor deposition (FIB-CVD), high-intensity electron-beam welding [5], Joule-heating-induced joining [6], and nanomechanical bonding [7] are experimentally demonstrated interconnection strategies, although all have limitations. Van der Waals forces are generally very weak, Joule-heating-induced joining and nanomechanical bonding are promising but not yet mature, and the other methods involve high-energy electron or ion beams, which significantly limits their applications. Another interconnection approach is to use CNTs, with their hollow cores and large aspect ratios, as possible conduits for nanoscale amounts of various materials that can be used to fuse CNTs together.

The concept of nanorobotic spot welding is schematically shown in Fig. 1. A metal-filled carbon nanotube is positioned with a nanorobotic manipulator to the interconnection site and the metal is then deposited to solder the nano-building blocks, such as CNTs and nanowires, together, or to weld them onto electrodes.

To facilitate the deposition, the encapsulated metallic materials may need to be melted or evaporated, which means only materials with a lower melting point than that of CNTs can be applied.

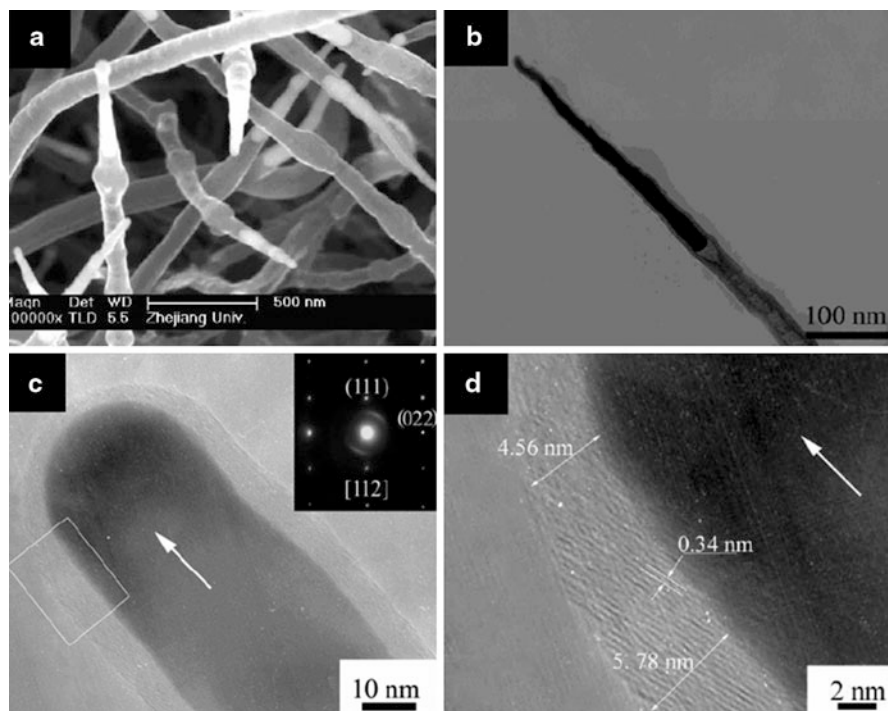
Nanorobotic Spot Welding,**Fig. 1** Concept of nanorobotic spot welding**Nanorobotic Spot Welding, Table 1** Melting point of pure metals and alloy solders

Materials	Melting point (°C) ^a	Possibility to be filled into CNTs
W	3,410	
CNTs in vacuum	2,800 [8]	
Fe	1,535	Produced by pyrolysis of ferrocene powder [9]
Y	1,523	Yttrium carbide synthesized from yttrium oxide powder mixed with isopropanol [10]
Cu	1,083	Synthesized using an alkali doped Cu catalyst by a thermal CVD method [11]
Au	1,064.43	Capillary filling [12]
Ag	961.93	Capillary filling [13]
CNTs in air	750 [8]	
Pb	327.5	Capillary filling [14]
Bi	271.3	Capillary filling
Sn	231.9	Synthesized by catalytic deposition of acetylene using nanocrystalline SnO ₂ as a catalyst [11]
Sn ₅₀ Pb ₅₀	183–212	
Ga	29.78	Synthesized in a vertical radiofrequency furnace from a homogeneous mixture of Ga ₂ O ₃ and pure, amorphous, active carbon under a flow of pure N ₂ gas [15]

^aData source for melting points not specified: <http://www.chemicalelements.com/show/meltingpoint.html>

Fortunately, most pure metals have a lower melting point than that of CNTs in vacuum and a large portion of them have a lower melting point than that of CNTs in air (Table 1). Alloy solders routinely used in the larger scale for electronics and silversmithing may also serve as excellent candidates of nanosolders and they

typically have lower melting points than the pure metals. These include soft solders (melting range: 90–450°C, typically tin-lead (Sn-Pb) alloys), lead-free solders (melting point 5–20°C higher than soft solders, alloys contain tin, copper (Cu), silver (Ag), bismuth (Bi), indium (In), zinc (Zn), antimony (Sb),



Nanorobotic Spot Welding, Fig. 2 (a–d) Copper-filled CNTs. (a) FESEM images of Cu-filled CNTs. Observation shows that all the CNTs have sharp tips filled with metal nanoneedles. These CNTs are up to 5 μm long with outer diameters in a range of 40–80 nm. (b) A typical copper-filled CNT synthesized for 30 min. (c) HRTEM image reveals that the Cu nanoneedles are encapsulated in graphite walls approximately 4–6 nm thick. The inset is the corresponding SAED pattern of the Cu nanoneedle along the zone axis, showing that

the Cu nanoneedle is single crystalline. The appearance of a pair of arcs in the SAED pattern indicates some orientation of the (002) planes in the carbon tubes. (d) The magnified image of the rectangular region in (c). The *large arrows* indicate the growth direction of the CNTs. The interlayer spacing of carbon nanotube is about 0.34 nm, consistent with the (002) plane lattice parameter of graphite. It can also be seen that the graphite layers are not parallel to the tube axis

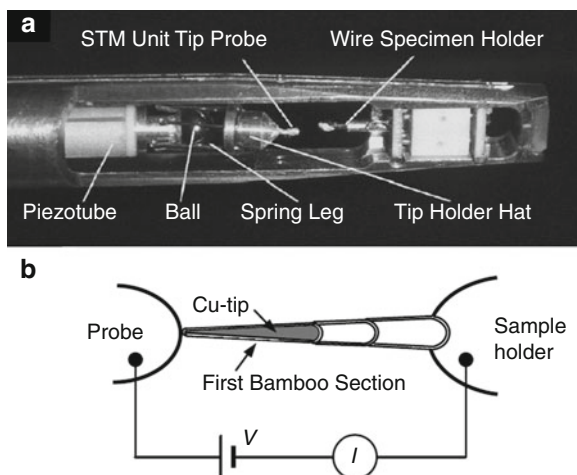
and traces of other metals), and hard solders (melting point $>450^\circ\text{C}$, alloys of copper with either zinc or silver).

Figure 2 shows CNT samples synthesized using an alkali-doped Cu catalyst by a thermal CVD method [16]. Figure 2a–d show their structures imaged by field emission scanning electron microscopy (FESEM, Sirion, FEI), transmission electron microscope (TEM, JEM-2010, 200 kV), selected area electron diffraction (SAED), and high-resolution TEM (HRTEM). It can be seen from these observations that the yield of the Cu-filled CNTs was high. The CNTs were up to 5 μm long with outer diameters in a range of 40–80 nm. The single-crystalline Cu nanoneedles were encapsulated in graphite walls approximately 4–6 nm thick in the first sections of the bamboo-structured CNTs. The graphite layers were not parallel to the tube axis.

Key Research Findings

An experimental investigation of controlled melting and flowing of single crystalline copper from CNTs has been performed using nanorobotic manipulation [3], and its application in spot welding of nanotubes has been demonstrated using this copper. Because copper is a good conductor of heat and electricity and has a very low binding energy (0.1–0.144 eV/atom) when bound to carbon, copper-filled CNTs were used in the investigation.

The experiments were performed in a CM30 TEM equipped with a scanning tunneling microscope (STM) built in a TEM holder (Nanofactory Instruments AB, ST-1000) serving as a manipulator, as shown in Fig. 3a or schematically in Fig. 3b. The material consisting of a CNT bundle was attached to a 0.35 mm thick Au wire



Nanorobotic Spot Welding, Fig. 3 (a, b) Nanorobotic manipulation system in a TEM

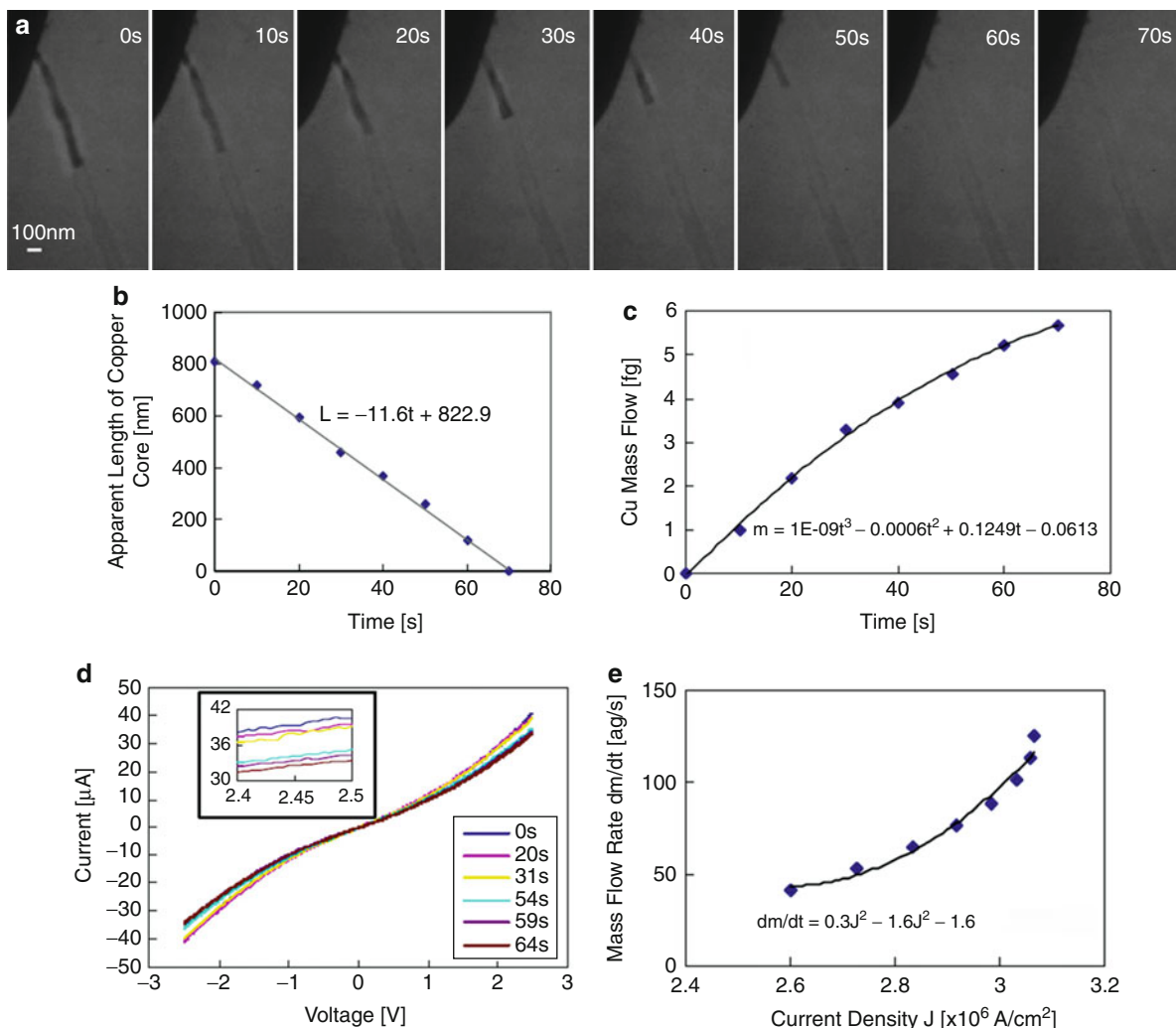
using silver paint, and the wire was held in the specimen holder. The probe was an etched 10 μm thick tungsten wire with a tip radius of approximately 100 nm (Picoprobe, T-4-10-1 mm). The probe can be positioned in a millimeter-scale workspace with subnanometer resolution with the STM unit actuated by a three-degree-of-freedom piezo-tube, making it possible to select a specific CNT. Physical contact can be made between the probe and the tip of a nanotube. Applying a voltage between the probe and the sample holder establishes an electrical circuit through a CNT and injects thermal energy into the system via Joule heating. By increasing the applied voltage, the local temperature can be increased past the melting point of the copper encapsulated in a tube. The process was recorded by TEM images and real-time video.

In the experiment, the tip of a Cu-filled CNT was first brought into contact with the tungsten probe. Then a bias voltage was applied on the two ends of the CNT with the tungsten probe serving as the anode. The voltage was slowly increased from 0 mV with 100 mV steps. When the voltage reached 1,500 mV, a vacant section inside the carbon shells appeared, indicating that the copper has begun melting. Transportation of the copper to the probe-tip contact moved the vacant section to the root of the first bamboo section of the Cu-filled CNT. After the vacant section reached the root, the copper core started to flow to the tip of the CNT when bias was increased to 2,500 mV.

Figure 4a is a series of time-resolved TEM images taken from video frames showing the flowing process. The copper core started to flow inside the carbon shell from the bottom to the tip of the first bamboo section as the bias voltage reached 2.5 V. The entire process continued for about 70 s. The flow rate was found to be 11.6 nm/s according to the change of apparent length of the copper core (Fig. 4b). Accordingly, the mass change was calculated as shown in Fig. 4c, and the mass flow rate can then be determined by fitting the data to the curve $3 \times 10^{-9}t^2 - 1.2 \times 10^{-3}t + 0.12$, yielding approximately 120 ag/s, which is strikingly slow and well controllable, allowing precise delivery of mass at the attogram scale for time-based control that can readily reach sub-second precision.

Figure 4d shows time-resolved current versus voltage characteristics under a constant positive bias of 2.5 V. The current density under 2.5 V when flowing occurred was then calculated according to the cross-sectional area as $2.60 - 3.07 \times 10^6 \text{ A/cm}^2$. This is comparable to the observed value for electromigration of iron in CNTs (ca. $7 \times 10^6 \text{ A/cm}^2$) [17]. The difference can be a result of the lower binding energy of copper to the carbon shells (0.1–0.144 eV/atom) than that of irons to carbon shells (0.3 eV/atom). The high current densities employed here will lead to resistive heating. Temperatures as high as 2,000–3,000°C have been estimated according to the lattice spacing change in electric breakdown experiments on non-filled multiwalled nanotubes (MWNTs) [18] at a slightly higher bias (3 V) than those used here. The current density J and the mass flow rate \dot{m} were then correlated as shown in e. The relation $\dot{m} = 0.3135J^2 - 1.6206J + 2.1373$ suggests that a real positive value of $\dot{m}(\geq 42.9 \text{ ag/s})$ can only be given when the current density J surpasses $2.5847 \times 10^6 \text{ A/cm}^2$. The existence of this threshold also implies that the mechanism of the observed flowing is most possibly electromigration [17]. Under a negative bias, that is, when the tungsten probe served as a cathode, the flow was observed in the opposite direction.

Other possible mechanisms for flow can be excluded. Capillary force can induce filling/flowing, but the direction should be opposite to the observed flow, that is, from the tip to the bottom of the carbon shells. Thermal expansion can enable flow, but the flow should be isotropic, heading toward both the tip and the bottom. A recent investigation showed that the



Nanorobotic Spot Welding, Fig. 4 (a) Time-resolved TEM images from video frames showing the flowing process. The copper core started to flow inside the carbon shells from the root to the tip as the bias voltage reached up to 2.5 V. The whole process was continued for about 70 s. (b) The flow rate has been found to be 11.6 nm/s according to the change of apparent length of the copper core. The tungsten probe has been positively

biased. (c) The mass changes along the time. The mass flow rate can be then drawn out from the fitting curve as approximately 0.12 fg/s. (d) Time-resolved current versus voltage characteristics under a constant bias of 2.5 V. The current density under 2.5 V as flowing occurred is about $2.60 - 3.07 \times 10^6$ A/cm². (e) Correlation of the current density J and the mass flow rate dm/dt

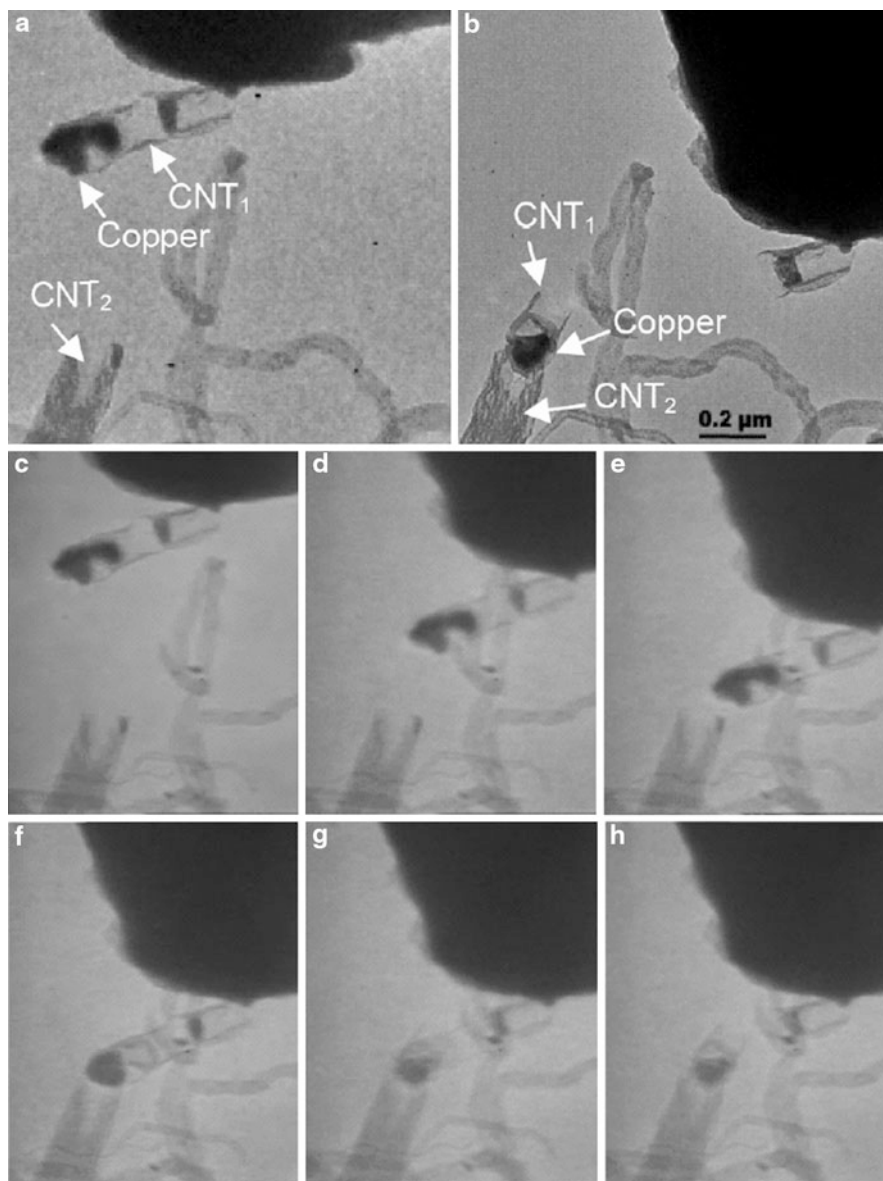
irradiation of MWNTs can cause a large pressure buildup within the nanotube core that can plastically deform, extrude, and break encapsulated solid material [19]. In the experiments, however, no contraction of the carbon shells was observed.

Although there is no way of measuring the actual temperature of the nanotubes or the copper core, this process provides the possibility for estimating the temperature according to the resistivity change if

Matthiessen's rule is applied on temperature dependence of resistivity, provided that the resistivity was known at a certain temperature. Unfortunately, the latter is unknown. However, it has been noted that the temperature for CNT growth from 100-nm scale Cu particles in the synthesis (700°C) is far lower than the melting point of bulk Cu (1,083°C) [16]. It is well known that the surface-to-volume ratio with respect to a nanosized particle can affect the melting point.

Nanorobotic Spot Welding,

Fig. 5 Self-soldering of CNTs with copper encapsulated in a CNT. (a) A copper-filled tube, CNT₁, was attached to a tungsten probe. (b) A section of CNT₁ was soldered onto CNT₂ by the melted copper. (c–h) Video frames showing the soldering process. A copper-filled tube, CNT₁, was first attached to a probe, and brought into contact with another tube, CNT₂. The probe (cathode) has a -10 V bias. (c–e) Three different positions as the probe is approaching CNT₂. (f) Showing how contact has been made. The shape change of the copper suggested a melting process has occurred. With a higher bias (-15 V), CNT₁ was broken (g) and its end section remains soldered to the tip of CNT₂ (h)

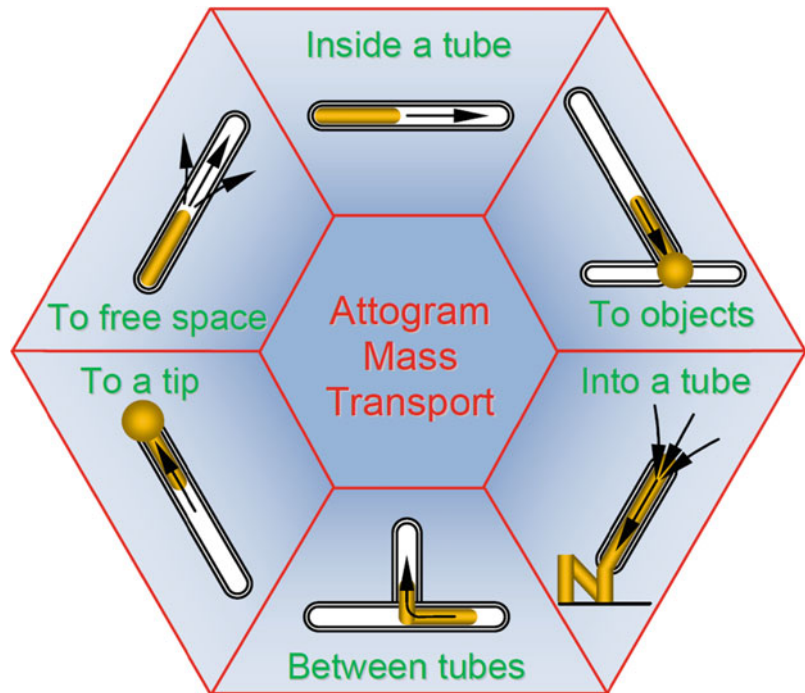


Although the exact reason for melting at such a low temperature is not yet known, a rough estimation of the melting point of the copper core can be made, that is, around 700°C . This is comparable to the in situ electron microscope observations of the melting point of the encapsulated 20–60 nm diameter Cu nanocrystals in multilayer graphitic carbon spheres, which was reported at 802°C [20]. It is also known according to the experiment that the melting point of the carbon shells is much higher than that of the copper core. This provides further evidence of molten copper inside carbon shells.

Examples of Application

The application of such controlled transportation of the copper core was then investigated. Self-soldering of CNTs with copper encapsulated in a CNT is shown in Fig. 5. A copper-filled tube, CNT₁, was first attached to a tungsten probe (Fig. 5a). A section of CNT₁ was then attached to and soldered onto CNT₂ by the melted copper (Fig. 5b). Figure 5c–h are video frames showing the soldering process. A copper-filled tube, CNT₁, was attached to a probe, and brought into contact with another tube, CNT₂. The probe has a -10 V bias. Parts

Nanorobotic Spot Welding,
Fig. 6 Related investigations
 and applications



c–e of Fig. 5 show three different positions as the probe was approaching CNT₂, and Fig. 5f shows how the contact has been made. The shape change of the copper suggests a melting process has occurred. It has been found from the video frames (25 fps) that the melting of the copper happened in a very short interval (<70 ms). With a higher bias (–15 V), CNT₁ was broken (Fig. 5g) and its end section remained soldered to the tip of CNT₂ (Fig. 5h).

Compared with the other interconnection processes previously investigated, electrically driven spot welding has several interesting aspects: (1) A very low current can induce the melting and drive the flow; which is much more efficient than irradiation-based techniques involving high-energy electron beams [4, 5, 19, 21, 22], FIB, or lasers. Combined with dielectrophoretic assembly, it is possible to solder the tubes onto electrodes for batch fabrication of NEMS. (2) The welding site can be readily selected using nanorobotic manipulation, which enables 3D position and orientation control for continuous mass delivery and will potentially enable 3D prototyping and assembly. (3) The melting occurs rapidly (at least at the millisecond level), several orders of magnitude faster than that using high-intensity electron-beam or FIB, which is generally on the order of a minute

[4, 5, 19, 22, 23]. (4) Because both the rate and direction of mass transport depends on the external electrical drive, precise control of an ultras-small mass delivery is possible. Time-based control will allow the delivery of attograms of mass. (5) Copper has good compatibility in the conventional semiconductor industry. The experiments show that it will likewise play an important role for scaled-down systems. Carbon shells provide an effective barrier against oxidation and consequently ensure a long-term stability of the copper core, which also facilitates the conservation of the material than conveying mass on the external surface of nanotubes.

Future Directions for Research

The investigation on nanorobotic spot welding has so far shown that controlled melting and flowing of copper inside nanotube shells can be realized by applying a low bias voltage (~1.5–2.5 V). The melting can facilitate the flowing and it is a result of Joule heating, whereas the flowing is caused by electromigration. The mass flow rate can be controlled to the scale of attogram per second based on time: a typical value is around 1 atom/μs.

As an emerging technique, there are still a lot of unsolved problems with nanorobotic spot welding. (1) The materials investigated are limited to copper and tin. More efforts need to be put into the filling of conventional alloy solders and silver-based solders to broaden the adaptability of this technique. Universal and high yield filling techniques need to be developed. (2) The controllability of melting and flowing is related to the bias profile, which is determined mainly by the diversity of the resistivity of CNTs. Uniform CNTs in a batch are strongly required. (3) TEM provided an ideal environment for the in situ investigations. To lower the cost for practical applications, it would be a better solution to combine the spot welder with an AFM. The impact of the change of the environment from a high-vacuum chamber to the air needs to be evaluated. Obviously, solders with lower melting points will be required. (4) For automatic batch soldering, continuous feeding, arrays of solders (can be individually filled with heteromaterials), automatic detection of the soldering sites, and path planning are some examples of critical techniques.

Nanorobotic spot welding is a fundamental technology related to the mass transport inside, between, and into nanochannels (Fig. 6). Mass transport systems at this scale provide a platform for the investigation of nanofluidics and can serve as components for prototyping nanostructures. Mass flowing out from a nanotube can generate a variety of results based on the targets. Evaporation to a free space is a result of discharging, ionization, or heating. Deposition on a surface can be obtained if placing the nanotube tip against such a target. Nanorobotic spot welding is based on mass transport to the interface between two objects. Free-standing nanostructure can be formed on the tip of the injector nanotube if moving it back from the target. As an emerging additive nanolithography technique, electromigration-based deposition (EMBD) is a more general term representing the mass transport of metallic materials against a surface. The investigations of all these systems are correlated and can benefit from the development of each other.

Cross-References

► [Electron-Beam-Induced Deposition](#)

References

1. Cary, H.B., Helzer, S.C.: *Modern Welding Technology*, 6th edn. Prentice-Hall, Upper Saddle River (2005)
2. Dong, L.X., Tao, X.Y., Zhang, L., Zhang, X.B., Nelson, B. J.: Nanorobotic spot welding: controlled metal deposition with attogram precision from copper-filled carbon nanotubes. *Nano Lett.* **7**(1), 58–63 (2007)
3. Dong, L.X., Arai, F., Fukuda, T.: Destructive constructions of nanostructures with carbon nanotubes through nanorobotic manipulation. *IEEE-ASME Trans. Mechatron.* **9**(2), 350–357 (2004)
4. Dong, L.X., Arai, F., Fukuda, T.: Electron-beam-induced deposition with carbon nanotube emitters. *Appl. Phys. Lett.* **81**(10), 1919–1921 (2002)
5. Xu, S.Y., Tian, M.L., Wang, J.G., Xu, J., Redwing, J.M., Chan, M.H.W.: Nanometer-scale modification and welding of silicon and metallic nanowires with a high-intensity electron beam. *Small* **1**(12), 1221–1229 (2005)
6. Hirayama, H., Kawamoto, Y., Ohshima, Y., Takayanagi, K.: Nanospot welding of carbon nanotubes. *Appl. Phys. Lett.* **79**(8), 1169–1171 (2001)
7. Dong, L.X., Arai, F., Fukuda, T.: Nanoassembly of carbon nanotubes through mechanochemical nanorobotic manipulations. *Jpn. J. Appl. Phys. Part 1 – Regul. Pap. Short Notes Rev. Pap.* **42**(1), 295–298 (2003)
8. Collins, P.G., Avouris, P.: Nanotubes for electronics. *Sci. Am.* **283**(6), 62–69 (2000)
9. Svensson, K., Olin, H., Olsson, E.: Nanopipettes for metal transport. *Phys. Rev. Lett.* **93**(14), 145901 (2004)
10. Seraphin, S., Zhou, D., Jiao, J., Withers, J.C., Loufy, R.: Yttrium carbide in nanotubes. *Nature* **362**(6420), 503 (1993)
11. Tao, X.Y., Zhang, X.B., Cheng, J.P., Luo, Z.Q., Zhou, S.M., Liu, F.: Thermal cvd synthesis of carbon nanotubes filled with single-crystalline Cu nanoneedles at tips. *Diam. Relat. Mater.* **15**(9), 1271–1275 (2006)
12. Chu, A., Cook, J., Heesom, R.J.R., Hutchison, J.L., Green, M.L.H., Sloan, J.: Filling of carbon nanotubes with silver, gold, and gold chloride. *Chem. Mater.* **8**(12), 2751–2754 (1996)
13. Ugarte, D., Chatelain, A., de Heer, W.A.: Nanocapillarity and chemistry in carbon nanotubes. *Science* **274**(5294), 1897–1899 (1996)
14. Ajayan, P.M., Iijima, S.: Capillarity-induced filling of carbon nanotubes. *Nature* **361**(6410), 333–334 (1993)
15. Gao, Y.H., Bando, Y.: Carbon nanothermometer containing gallium – gallium’s macroscopic properties are retained on a miniature scale in this nanodevice. *Nature* **415**(6872), 599 (2002)
16. Tao, X.Y., Zhang, X.B., Cheng, J.P., Luo, Z.Q., Zhou, S.M., Liu, F.: Thermal cvd synthesis of cnts filled with single-crystalline cu nanoneedles at tips. *Diam. Relat. Mater.* **15**, 1271–1275 (2006)
17. Svensson, K., Olin, H., Olsson, E.: Nanopipettes for metal transport. *Phys. Rev. Lett.* **93**(14), 145901 (2004)
18. Huang, J.Y., Chen, S., Jo, S.H., Wang, Z., Han, D.X., Chen, G., Dresselhaus, M.S., Ren, Z.F.: Atomic-scale imaging of wall-by-wall breakdown and concurrent transport measurements in multiwall carbon nanotubes. *Phys. Rev. Lett.* **94**(23), 236802 (2005)

19. Sun, L., Banhart, F., Krasheninnikov, A.V., Rodriguez-Manzo, J.A., Terrones, M., Ajayan, P.M.: Carbon nanotubes as high-pressure cylinders and nanoextruders. *Science* **312**, 1199–1202 (2006)
20. Schaper, A.K., Phillipp, F., Hou, H.: Melting behavior of copper nanocrystals encapsulated in onion-like carbon cages. *J. Mater. Res.* **20**(7), 1844–1850 (2005)
21. Madsen, D.N., Molhave, K., Mateiu, R., Rasmussen, A.M., Brorson, M., Jacobsen, C.J.H., Boggild, P.: Soldering of nanotubes onto microelectrodes. *Nano Lett.* **3**(1), 47–49 (2003)
22. Yokota, T., Murayama, M., Howe, J.M.: In situ transmission-electron-microscopy investigation of melting in submicron Al-Si alloy particles under electron-beam irradiation. *Phys. Rev. Lett.* **91**, 265504 (2003)
23. Matsui, S., Kaito, T., Fujita, J., Komuro, M., Kanda, K., Haruyama, Y.: Three-dimensional nanostructure fabrication by focused-ion-beam chemical vapor deposition. *J. Vac. Sci. Technol. B* **18**(6), 3181–3184 (2000)

Nanorobotics

Soichiro Tsuda

Exploratory Research for Advanced Technology,
Japan Science and Technology Agency, Osaka, Japan

Synonyms

[Microrobotics](#)

Definition

Nanorobots are robotic devices operating at micro- and nanoscale.

Overview

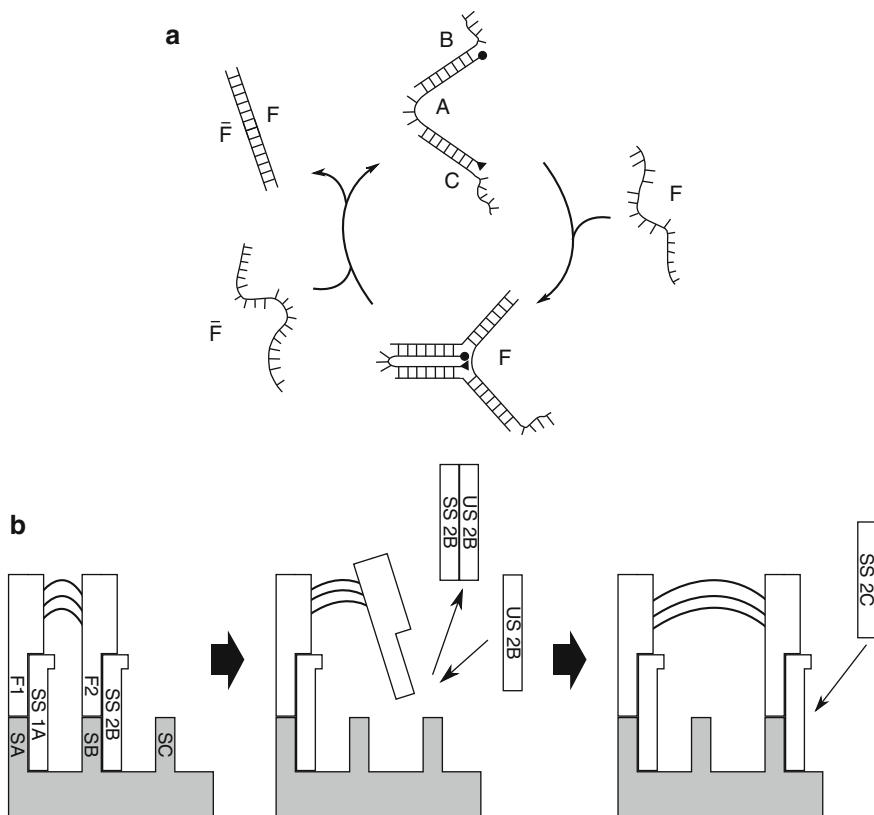
Nanorobotics is an emerging subfield of nanotechnology that aims to build robotic devices operating at micro- and nanoscale. The envisioned applications of nanorobotics are quite wide: ranging from medical and environmental sensing to space and military applications [1, 2]. Devices at micron scale that handle micron objects (e.g. bacterial robots), which are sometimes termed as “microrobots,” are also referred to as nanorobots here. Nano manipulators that control nanoscale objects using microscope

devices are also referred to as nanorobots, but it is not covered in this article.

While nanotechnology developed so far is concerned with the construction of static nanoscale structures and manipulation of nanoscale objects, the field of nanorobotics studies more dynamic and active structures. It includes the design and development of nanoscale machines that are capable of actuation, sensing, signaling, information processing, and intelligent behavior at nanoscale [1]. Being extremely small (potentially at atomic scale), nanorobots are unlikely to be equipped with many functions as in their macroscopic counterparts. A possible solution to implement more complex functionality would be to use swarm intelligence and cooperative behavior, in which a swarm of nanorobots communicate each other and accomplish tasks collectively. Another issue in regard to nanorobotics is the basic law that governs the dynamics of nanorobots. Conventional robots at macroscopic scale are governed by Newtonian mechanics, whereas nanorobots are governed mostly by quantum mechanics. The design principle of nanorobots is thus bound to be different from macroscale robots. For example, Brownian motion would need to be taken into serious consideration at such scale. Molecules constituting nanoscale devices are constantly exposed to collisions with other smaller molecules (e.g., water molecules) and not in a rigid, static structure [3]. Therefore, novel schemes of design, manufacturing, programming, and control for nanorobots are needed because schemes for macroscale robots cannot be applied.

Although nanorobots in practical use are yet to be developed, several prototypes of nanorobots have been built in recent years. In particular, they have been explored in the context of biomolecular machines inspired by “nature’s way of doing things” [1, 2]. Biomolecules are of particular interest in the field of nanorobotics because they operate at extraordinarily high efficiencies and readily available in nature [4]. In addition, they have several functions which are difficult to achieve with artificial molecular machines, for example, self-assembly and self-replication. As it is impossible to precisely fabricate nanomachines atom by atom, nature employs self-organization to build natural molecular machines. Proteins, the most widely used nature’s nanomachines, are constructed using DNA as blueprint. A DNA sequence is translated into a chain of amino acids via messenger RNA, and the

Nanorobotics, Fig. 1 (a) A schematic diagram of DNA tweezers. Open–close operations are controlled by additional fuel strands F and \bar{F} . Black circle and triangle indicate TET (fluorescent dye) and TAMRA (quencher), respectively. After [5]. (b) DNA walker walks on the steps constructed with single-stranded DNA. F1 and F2 indicate feet of the DNA walker, SA, SB, and SC are steps for DNA walker. Gaits of the walker are precisely controlled by external DNA strands, SS (Set Strand) and US (Unset Strand). See text for details. After [6]



amino acid chain will self-assemble into a three-dimensional structure through free-energy minimization. By separating replication process from the system itself (i.e., protein), self-replication becomes easier, and over one billion possible types of proteins are possible out of 20 amino acids and four nucleic acids. Potentially, nanorobots could be fabricated in a similar way that biological devices are assembled. In fact, several prototypes of nanorobots, such as DNA-based robots, rely on self-assembling nature of biomolecules for their operation. The following section gives a brief overview of nanorobotic devices implemented on various materials.

DNA-Based Nanorobots

Yurke and coworkers developed a molecular machine that takes DNA strands as fuels, called DNA tweezers [5]. The machine itself is also consists of three single-strands of DNA (Fig. 1a). Two of these stands (Strand B and C) have complementary sequence with another

strand (Strand A). In particular, half of Strand B is complementary to half of Strand A and half of Strand C is to another half of Strand A. Thus, when they are bound together, the DNA tweezers has a single-stranded DNA sequence on each end, which are parts of Strand B and C. Strand A is labeled with fluorescent dye TET (5' tetrachloro-fluorescein phosphoramidite) and TAMRA (carboxy-tetramethylrhodamine) on each end. Fluorescence emission from TET reports the positions of DNA tweezers. When TET is excited with an 514.5 nm argon ion laser, it emits fluorescent light with a peak wavelength of 536 nm. TAMRA works as a quencher of TET due to resonant intramolecular energy transfer when they come close together. When no fuel DNA (Strand F) is present, the DNA tweezers is “open” and TET fluoresces. If Strand F is added, it binds to the sticky ends of DNA tweezers because it has complementary sequences to single-stranded ends of the tweezers (parts of Strand B and C). As a result, they take a Y shape and the DNA tweezers are in “closed” position. This brings together TET and TAMRA attached to Strand A, and the light emission

from TET is quenched. The tweezers can be reopened by adding Strand \bar{F} , a complementary DNA to Strand F. It cleaves the hybridization between the DNA tweezers and Strand F and forms a double-stranded DNA with Strand F. This open-close cycle can be repeated several times by removing fuel DNA added in the previous cycle. They have estimated the force exerted by the closing as 15 pN, which is at the upper range of measured forces exerted by biomolecular motors, such as kinesin and myosin.

Sherman and Seeman developed a simple mobile robot based on DNA strands, called DNA walker [6]. The walker has two feet (Foot 1 and 2 in Fig. 1b), each of which is comprised of double-stranded and single-stranded DNA sequences. Two DNA feet are connected with three single-stranded DNA. The DNA walker walks on a line of “stepping stones” that are also made of single-stranded DNA sequences. The DNA walker is set to “stand” on stepping stones by another DNA strands (Set Strand) that are complementary to a Foot and a Step. For example, to set Foot 1 on Step A, a Set Strand 1A, which is complementary to the sequences of Foot 1 and Step A, is added to a solution, and it hybridizes with the foot and step to connect them together. The molecular machine walks on the stepping stones by addition and removal of Set Strands: If feet are set on neighboring stones Step A and B (linked by Set Strand 1A and 2B), the DNA walker lifts up one of legs by adding a complementary DNA strand to Set Strand 2B (called Unset Strand 2B). The strand cleaves the binding of Set Strand 2B with the foot and step to form a double-stranded DNA, and thus the Foot 2 is released from Step B. Foot 2 can be set on Step C by adding new strand Set Strand 2C. The same process can be applied to Foot 1 in order to step forward from Step A to B. Hence, the DNA walker can walk on the steps by repeating the process.

Based on the DNA walker, Lund and colleagues developed a DNA-based molecular robot that walks on two-dimensional prescribed landscapes [7]. The robot, called molecular spider, consists of a head and three legs, all of which are made of DNA. The way the molecular spider walks is basically similar to the DNA walker: It proceeds on a “track” because each leg (a single-stranded DNA) hybridizes and dehybridizes with another single-stranded DNA with complementary sequences to the leg. A big difference with the spider is autonomous behavior. It carries out sequences of actions, such as “start,” “follow,” “turn,” and “stop”

without any external controls. This autonomous action relies on catalytic activity of the spider legs. A DNA sequence with catalytic action, called DNA enzyme, constitutes each leg, and a track on the 2D DNA origami landscape [8] is constructed with three lines of oligodeoxynucleotide substrates (i.e., stepping stones). The locomotion of DNA spider is carried out as follows: First, a leg binds to a substrate by hybridization. DNA enzyme cleaves the single ribonucleotide linkage in the substrate strand into two shorter products. Due to a lower affinity of the products for the DNA enzyme, the leg will be dissociated from the site. It will quickly find a nearby substrate to reattach because multiple legs binding to substrates prevent the molecular spider from being completely dissociated from the track. Although the leg could attach to a site visited before (i.e., stepping back), it is likely to stay longer on new sites because the substrate has higher affinity for the enzyme than the products. Once a molecular spider is set on the starting point of a track, three legs bind and cleave the substrates in parallel, and the robot moves forward as a result. It has been shown that the spider can not only walk on a straight track, but also turns the direction (30° and 90°). After it reaches the end of the track, it would perform random Brownian walks because all the substrates are converted to the product and, therefore, any sites on the track have same affinity for the legs. To prevent this, “stop” sites, which are non-chimeric and uncleavable analogues of the substrate, can be introduced at the end of a track. When a leg steps on the site, cleavage of DNA strand does not occur, and, therefore, the robot stays in the stop sites and never goes back.

Cell-Based Nanorobots

Another approach towards nanorobots is to employ whole living cells as machines. As mentioned above, natural biomolecular machines are quite efficient in many aspects, and it is hard to fabricate competing artificial robotic elements from scratch. Cell-based nanorobotics aims at making use of biological functionality of living organisms for man-made tasks, such as cargo transportation.

Behkam and Sitti developed a bacterial actuator for propulsion of polystyrene microbeads [9]. The cell-based actuator employs flagella motors of *Serratia*

marcescens. Using a plasma-based bacteria patterning technique, the flagellated chemotactic bacteria have been randomly attached to half of a 10 μm microbead. This configuration of bacteria attachment propels the microbead, although the direction of motion is stochastic. It has been shown that the more bacteria are attached, the faster the bead is propelled. The on-off motion of the flagella motors can be controlled by introducing copper ions (Cu^{2+}) and ethylenediaminetetraacetic acid (EDTA). Steager and colleagues [10] showed the bacteria can thrust SU-8 microfabricated structures (larger than the microbeads) and the movement of the structures can be halted and resumed by UV exposure.

Martel and colleagues has been taking a different approach to the development of bacterial microrobots [11]. They point out several drawbacks of bacteria actuators using *Serratia marcescens*: First, the directional control is difficult. Second, the chemical control of bacterial motion is not appropriate for applications in many environments, especially when applications for operations in the human body are in focus. Third, chemical control is not compatible with electronic circuits, i.e., digital computers. To tackle these problems, they use Magnetotactic Bacteria (MTB). The bacteria, marine magnetotactic coccus strain MC-1, incorporate a chain of nanoparticles called magnetosomes which are used to orient the bacteria in geomagnetic fields. On the other hand, the direction of magnetotactic bacteria can be controlled by external magnetic field by inducing a torque on a magnetosome. A bacterium can produce a thrust force exceeding 4 pN, and a collective of MC-1 cells assembled by computer-controlled magnetic field can generate larger force. It has been shown that the collective of approximately 5,000 flagellated magnetotactic bacteria can not only propel microstructures, but also complex micro-assembly tasks, such as building a pyramid-like structure [12].

Other Types of Nanorobots

There are several approaches toward nanorobots apart from the DNA and cell-based implementations. Another example is protein-based nanorobotic system. Protein is the most common natural molecular robotic devices and abundant in living cells. Central to the protein-based nanorobotics is the conformational change of three-dimensional structure because proteins are quite

sensitive to the environmental changes. It has been proposed to use viral proteins as motors (Viral Protein Linear (VPL) motor). This has been inspired by the fact that the families of retroviruses, such as influenza virus and HIV-1, are able to sensitively detect a drop in pH as a signal that its future host is close. The drop of pH is detected by a protein on the viral membrane, which changes the conformation from loosely connected structure to a distinctive α -helical structure. Artificially synthesized protein-like viral proteins could potentially used to produce forces of the order of μN .

Inorganic nanorobots are also pursued (cf. [13]). What has been proposed is controllable nanomotor devices using catalytic reaction of bimetal. When cylindrical asymmetrical alloys (e.g., gold and platinum) are suspended in a solution with hydrogen peroxide as fuel, the platinum segments catalyze the oxidation of hydrogen peroxide (i.e., $\text{H}_2\text{O}_2 \rightarrow \text{O}_2 + 2\text{H}^+ + 2\text{e}^-$) whereas the gold segments enhance the opposite reaction. As a result, electrons in the nanomotor flows from the platinum to gold side along with migration protons in the double layer on the motor, and the nanomotor shows random but non-Brownian movements. Precise directional control for performing complex tasks is also possible by applying magnetic field. Ferromagnetic nickel incorporated in the nanomotors allows the aligning of the motor and modulation of magnetic field strength enables start and stop control of the motion.

Cross-References

- ▶ [3D Micro/Nanomanipulation with Force Spectroscopy](#)
- ▶ [Nanorobotic Assembly](#)
- ▶ [Nanorobotic Manipulation of Biological Cells](#)
- ▶ [Nanorobotic Spot Welding](#)
- ▶ [Nanorobotics for Bioengineering](#)
- ▶ [Nanorobotics for NEMS Using Helical Nanostructures](#)
- ▶ [Robot-Based Automation on the Nanoscale](#)

References

1. Ummat, A., Dubey, A., Mavroidis, C.: Bionanorobotics: a field inspired by nature, Chapter 7. In: Bar-Cohen, Y. (ed.) Biomimetics: Biologically Inspired Technologies, pp. 1–227. CRC, Boca Raton (2006)

2. Ummat, A., Dubey, A., Sharma, G., Mavroidis, C.: Bio-nano-robotics: state of the art and future challenges, Chapter 19. In: Yarmush, M.L. (ed.) *The Biomedical Engineering Handbook*, vol. 3, 3rd edn. CRC, Boca Raton (2006)
3. Yanagida, T.: Fluctuation as a tool of biological molecular machines. *Biosystems* **93**(1–2), 3–7 (2008)
4. Weir, N.A., Sierra, D.P., Jones, J.F.: A review of research in the field of nanorobotics. Technical Report SAND2005-6808. Sandia National Laboratories (2005)
5. Yurke, B., Turber Jr., A.J., Mills, A.P., Simmel, F.C., Neumann, J.L.: A DNA-fuelled molecular machine made of DNA. *Nature* **406**(August), 605–608 (2000)
6. Sherman, W.B., Seeman, N.C.: A precisely controlled DNA biped walking device. *Nano Lett.* **4**(7), 1203–1207 (2004)
7. Lund, K., Manzo, A.J., Dabby, N., Michelotti, N., Johnson-Buck, A., Nangreave, J., Taylor, S., Pei, R., Stojanovic, M.N., Walter, N.G., Winfree, E., Yan, H.: Molecular robots guided by prescriptive landscapes. *Nature* **465**(7295), 206–210 (2010)
8. Rothmund, P.W.: Folding DNA to create nanoscale shapes and patterns. *Nature* **440**(7082), 297–302 (2006)
9. Behkam, B., Sitti, M.: Effect of quantity and configuration of attached bacteria on bacterial propulsion of microbeads. *Appl. Phys. Lett.* **93**(22), 223901 (2008)
10. Steager, E., Kim, C., Patel, J., Bith, S., Naik, C., Reber, L., Kim, M.: Control of microfabricated structures powered by flagellated bacteria using phototaxis. *Appl. Phys. Lett.* **90**(26), 263901 (2009)
11. Martel, S., Mohammadi, M., Felfoul, O., Lu, Z., Pouponneau, P.: Flagellated magnetotactic bacteria as controlled mri-trackable propulsion and steering systems for medical nanorobots operating in the human microvasculature. *Int. J. Rob. Res.* **28**(4), 571–582 (2009)
12. Martel, S., Mohammadi, M.: Using a swarm of self-propelled natural microrobots in the form of flagellated bacteria to perform complex micro-assembly tasks. In: 2010 IEEE International Conference on Robotics and Automation (ICRA), Anchorage, May 2010, pp. 500–505 (2010)
13. Wang, J.: Can man-made nanomachines compete with nature biomotors? *ACS Nano* **3**(1), 4–9 (2009)

Nanorobotics for Bioengineering

Sylvain Martel

NanoRobotics Laboratory, Department of Computer and Software Engineering, and Institute of Biomedical Engineering, École Polytechnique de Montréal (EPM), Montréal, QC, Canada

Synonyms

Macroscale platforms capable of biological studies and manipulations at nanoscale to bio-inspired and bacterial nanorobotics

Definition

Nanorobotics for bioengineering is the use of robotic platforms capable of nanometer resolution, nanoscale robots, or microscale robots relying on nanoscale components and phenomena being applied to living systems.

Overview

Robotics

The field of nanorobotics bridges nanotechnology with robotics. As such, to understand the fundamentals of nanorobotics, one must first know the basic elements being associated with robotics and nanotechnology.

While robotics would be defined as the science or study of the technology associated with the design, fabrication, theory, and application of robots, a robot would typically be referred to as a machine or device that operates automatically or by remote control. As such, a robot must have an appropriate structure suitable for its working environment while being able to hold all the essential embedded components, and will include a power source (embedded, wired, or wireless), actuation (e.g., for displacement, grasping, etc.), sensing (vision, pressure, temperature, etc.), and control or computation (embedded or remote). Depending on the tasks and requirements, communication may also be involved.

Nanotechnology

Nanotechnology on the other hand is often defined as the understanding and control of matter at dimensions of roughly 1–100 nanometers (nm), where unique phenomena enable novel applications. Nanotechnology involves imaging, measuring, modeling, and manipulating matter at this length scale. It is important to note that at the nanoscale, the physical, chemical, and biological properties of materials differ from the properties of matter at the macroscale, and this fact has motivated research and development in nanotechnology to be directed toward understanding and creating improved materials, devices, and systems that exploit these new properties.

Nanorobotics

So far, researches in nanorobotics are being done within a range delimited by two extremes. The

bottom one is devoted to the design of robots with overall dimensions in the nanometer scale while the other involves interactions with nanoscale objects using macroscopic instruments such as scanning probe microscopes (SPMs) used for automated nanomanipulation. Indeed, nanotechnology and nanoscience began in the early 1980s in great part due to the invention of the scanning tunneling microscope (STM) [1] that allowed scientists to observe atomic structures. Later other forms of SPMs such as the well-known atomic force microscope (AFM) [2] appeared allowing more tasks to be performed at the atomic and molecular scales. Soon, engineers put efforts at making SPM-based tasks remotely [3] and later automatically [4] with more precision by integrating robotic principles and hence, these efforts have led among the first known applications of nanorobotics.

Although intensive research efforts were done in the various disciplines required to enhance the functionality and performance of these macroscale nanorobotic platforms, other efforts aimed at designing robots with overall dimensions in the nanometer scale. But although some experimental principles in the form of primitive molecular machines have been shown, the field of creating real practical machines or robots with overall dimensions in the nanometer range (e.g., 1–100 nm) is still beyond present technological limits.

As such, in an attempt to design the smallest practical robots, some research efforts concentrated on nanorobots with overall dimensions in the micrometer (μm) scale. These microscale nanorobots typically differ from the more classic microrobots by the fact that unlike the latter, the formers exploit the unique properties of nanotechnology by embedding the appropriate nanoscale components to integrate within the robots, the required functionalities to allow them to accomplish their assigned tasks.

Previous experimental results have also shown that the implementation of these so-called microscale nanorobots can be entirely artificial (synthetic) including biology-inspired nanorobots, natural including bacterial nanorobots (e.g., bacteria controlled by computer to perform the same tasks as futuristic artificial microscale nanorobots), or hybrid (also sometimes referred to as bio-nanorobots), i.e., built with a mix of artificial or synthetic and biological components.

Bioengineering

Although several applications for nanorobotics have been envisioned including but not limited to nano-assembly tasks within nanofactories, by observing the more recent experimental results, it becomes obvious that the field of nanorobotics has great potential in providing significant contributions for instrumentation in bioengineering.

Bioengineering [5] can be defined as the application of engineering design and technology to living systems. As such, bioengineers are concerned with the application of engineering sciences, methods, and techniques to problems mainly encountered in medicine and biology. Although macroscale nanorobotics as defined earlier can play a great role for biology, microscale nanorobots due to their overall dimensions could be particularly suited for *in vivo* applications in medicine. These two main approaches in nanorobotics applied to bioengineering are described in more details in the following sections.

Macroscale Nanorobotics in Bioengineering

Macroscale nanorobotics platforms can be generally categorized as contact-mode nanomanipulation systems using SPM, and non-contact-mode systems such as the ones using techniques such as optical trapping.

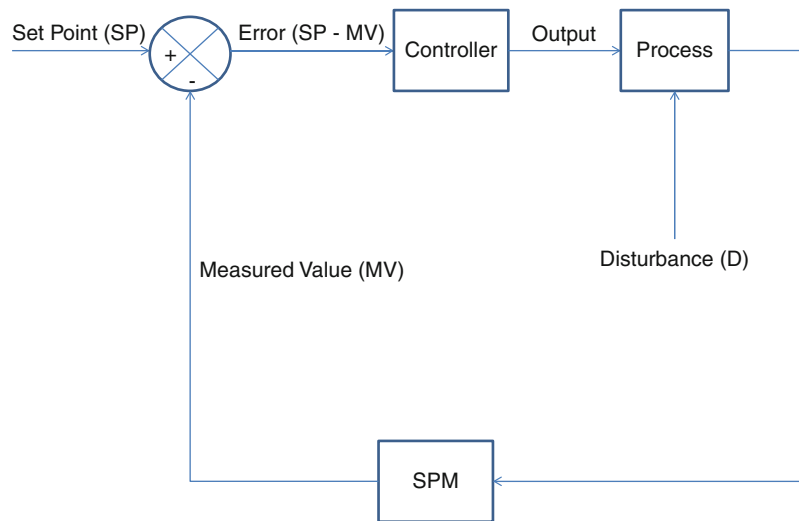
SPM-Based Macroscale Nanorobotics

In recent years, SPMs were used to image and manipulate biological samples and as such, it had a tremendous impact on the field of biology. Indeed, SPMs can be used to gather intrinsic properties of individual molecules as opposed to just bulk properties of larger samples. Therefore, since SPM technology has become an invaluable tool for the understanding of biological structures and processes at the nanoscale, it was adopted by the nanorobotics community for providing nanoscale sensing and manipulation in their platforms. The basic principle of these macroscale nanorobotic platforms is depicted in Fig. 1.

Figure 1 shows a closed-loop control system based on negative feedback. Here, the controller, implementing a given control algorithm, computes an output signal that will typically be used to actuate or move a SPM tip (i.e., the interfacing point between the SPM and the sample) at a desired location or set point (SP). As such, the controller corrects for any drifts

Nanorobotics for Bioengineering,

Fig. 1 Block diagram showing the basic principle of SPM-based macroscale nanorobotic platforms



resulting from the difference between the measured value (MV) from the SPM and the set point SP. These drifts or errors may be caused by the SPM itself (e.g., hysteresis and nonlinear characteristics of the piezoelectric (PZT) actuators used to move the SPM tip) and by various disturbances in the process. Such closed-loop control process is repeated at a rate sufficiently high to avoid excessive drifts beyond an acceptable error margin. The output computed by the digital controller is converted to an analog form using a digital-to-analog converter (DAC) and amplified to drive the actuator acting on the SPM which in most cases will be piezoelectric (PZT) actuators which are capable of atomic scale resolution.

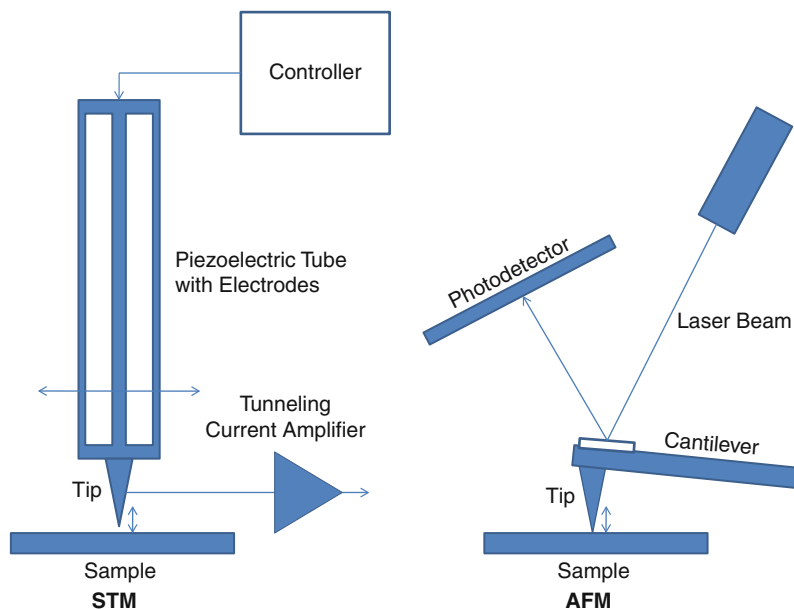
Although there is a huge variety of SPMs, so far most development projects of macroscale SPM-based nanorobotic systems used STM or AFM and as such, they will be emphasized here. The type of SPM and the mode used for biological samples is of critical importance. For instance, AFM has been particularly popular in biology since it has the unique capability of imaging biological samples with molecular resolution in buffer solution. Not only topographical images of surfaces with nanometer- to angstrom-scale resolution can be achieved, but the forces between single molecules and mechanical properties of biological samples can be investigated as well. Since the measurements can be made in buffer solutions, it allows researchers to maintain biological samples alive within a physiological-like environment. This allows temporal changes in nanostructures to be measured such as comparing

structural changes before and after the addition of chemical reagents. These advantages distinguish AFM from conventional imaging techniques of comparable resolution, including electron microscopy (EM) which will be discussed later. Thus, the attractive potential of SPMs and particularly AFM for biologists is the ability to visualize molecular processes under natural or physiological conditions. As such, AFM offers the resolution of most commercial electron microscopes but under the experimental conditions familiar with the much lower resolution of optical microscopes. For these reasons, AFM has been the most popular SPM in nanorobotics for bioengineering applications. The basic principle of AFM compared to STM is depicted in Fig. 2.

As shown in Fig. 2, the method used for AFM differs from STM. For STM, a controller generates voltage outputs to electrodes on a piezo-tube in order to deflect the tip in any directions. When the tip is positioned close enough to a sufficiently conductive sample, electrons from the atoms on the surface of the sample are attracted to the positively charged tip resulting in a tunneling current which is then amplified to create an output analog signal that will be converted to digital form by an analog-to-digital converter (ADC). This digital signal is then fed back to the digital controller that will compute the voltage outputs to move the tip in order to perform a scan of the surface of the sample using closed-loop control as depicted in Fig. 1. The changes in the magnitude of the tunneling current can be used to create an image of the atomic

Nanorobotics for Bioengineering,

Fig. 2 Basic functional principles of STM and AFM



structure or the tip can be positioned to maintain the tunneling current constant. In the latter case, the position of the tip is controlled as depicted in Fig. 1 and is used to create the image of the atomic structure.

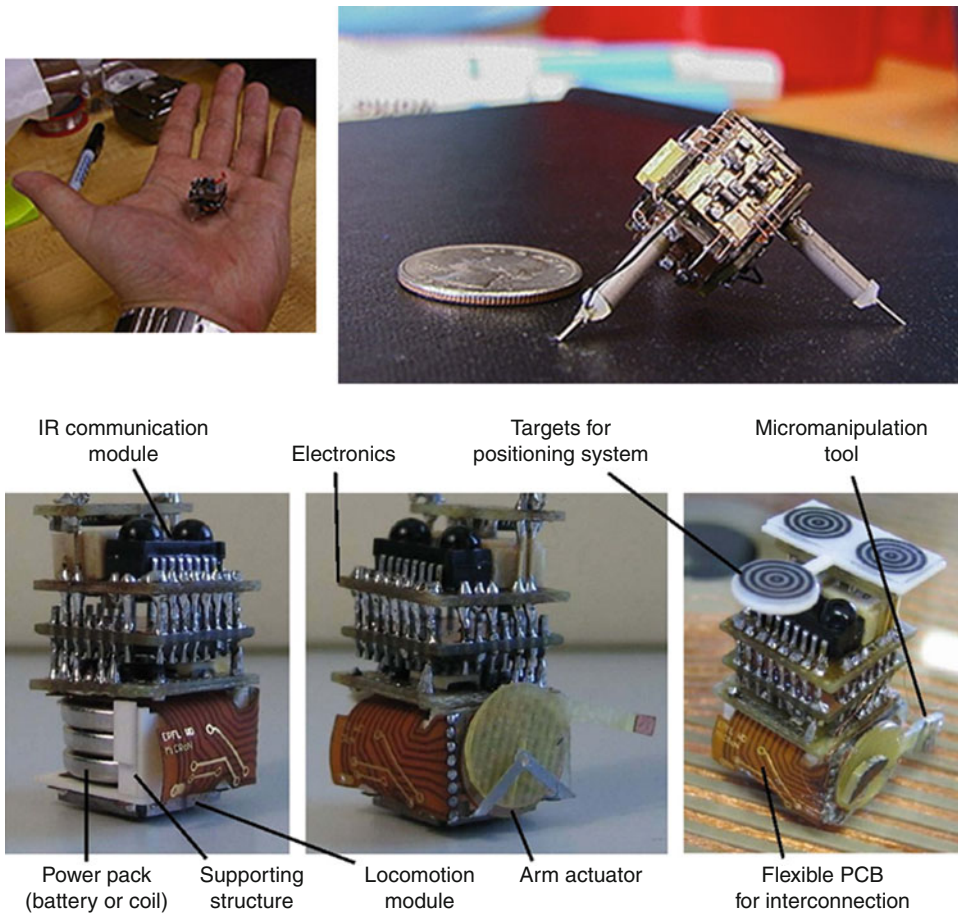
For AFM, a cantilever with a tip is used instead. When the cantilever is deflected due to the interaction between the tip and the sample, a laser beam is also deflected along a photodetector. The position of the laser beam on the photodetector provides information about the interaction between the tip and the sample that can be used by a controller to perform closed-loop control in a macroscale nanorobotic platform. Such information can not only be used for generating an image point by point but it also provides information such as the force involved, etc. The latter would prove to be useful to characterize the mechanical properties of biological samples and in performing nanomanipulation tasks under robotic closed-loop control such as the nanomanipulation of single DNA molecules for instance [6].

Research and development of macroscale nanorobotic systems have taken two main approaches that could be referred to as desktop SPM-based platforms and platforms based on miniature robots carrying an SPM. Instead of bringing the samples to the instrument as the former suggests, the latter proposes to bring the instruments to the samples, allowing the potential for higher throughput through parallelism

and enhanced scheduling or dispatching of each instrumented miniature robots. The concept of a nanorobotic platform relying on the dispatching of SPM-based miniature robots to the samples was first proposed and introduced with a robot known as the NanoWalker [7, 8]. Photographs of the first prototypes of the NanoWalker are being shown in the top section of Fig. 3.

The NanoWalker project began in the late 1990s. Equipped with three piezo-legs mounted in a pyramidal fashion with the apex pointing upward, the miniature robot could move with nanometer precision using approximately 4,000 steps per second (being equivalent to the resonant frequency of the robot). The maximum recording displacement speed was approximately 200 mm/s. The final design of the NanoWalker robot was equipped with an onboard computer with infrared (IR) wireless communication. It was powered through the legs when in contact with a special power floor. Although it was designed to have several types of SPMs, the initial design included only an embedded STM. With an STM, the NanoWalker robot although it could later be upgraded with an AFM was not initially suited for bioengineering applications.

This issue was corrected later with the MiCRoN robot [9] (see bottom of Fig. 3) capable of maximum displacement speed of 0.4 mm/s while being powered with batteries or wirelessly using inductive coupling.



Nanorobotics for Bioengineering, Fig. 3 *Top:* Photographs showing two different prototypes of the NanoWalker robot; *Bottom:* Photographs of the MiCRoN robot

The goal of this project was to make automated manipulation with nanometer precision of micrometer-scaled objects possible by these robots. Bioengineering applications included closed-loop force control for an embedded AFM tool in cell manipulation [10].

But to date, demonstrations of bioengineering applications for both types of macroscale nanorobotic platforms are still relatively few. On the other hand, efficient automatic micrometer scale manipulations in bioengineering have been demonstrated for desktop versions especially in autonomous cell injection [10] where micromanipulation strategies required for manipulating individual biological cells with a microrobotic system capable of conducting automatic embryo pronuclei DNA injection (cell injection) were described. One of the main reasons explaining that micromanipulation instead of nanomanipulation has generally been applied more successfully in

robotics is the fact that autonomous nanoscale operations are much more challenging due in great part to the accuracy required for successful operations coped with the difficulty in gathering real-time visual feedback information for the essential closed-loop control used in robotics.

To improve performance in gathering real-time visual feedback information, SPM-based microscale mobile nanorobots were placed in the vacuum chamber of a scanning electron microscope (SEM) [11]. The SEM electron beam was used to provide real-time 3D positional information about the tool used for nanomanipulation. Applications in bioengineering were limited since the samples needed to be electrically conductive and placed in a vacuum environment. Living cells, tissues, and soft-bodied organisms usually required chemical fixation to preserve and stabilize their structure.

Non-Contact-Mode Macroscale Nanorobotics

Non-contact nanomanipulation is presently dominated by optical trapping using optical tweezers. Optical tweezers use the forces exerted by a strongly focused beam of light to trap and move objects ranging in size from tens of nanometers to tens of micrometers [12]. Since its introduction in 1986, the optical tweezer has become an important tool for research in the fields of physical chemistry, soft condensed matter physics, and biology. Since optical tweezers apply loads in the order of picoNewtons (pN) in a non-contact mode, it is widely used in bio-nanomanipulation and mechanical studies of very fragile biological samples. For instance, single molecule manipulation combined with interaction study of double-stranded DNA [13] is only one example. Nanorobotic platforms relying on optical tweezers generally operate in a manner to assist a human operator. Furthermore, unlike SPMs, optical tweezers are generally more difficult to integrate in miniature robots and, hence, they are typically more appropriate for desktop macroscale nanorobotic platforms.

Microscale Nanorobotics in Bioengineering

Because of their overall dimensions, research and development efforts for microscale nanorobotics applied to bioengineering have been focus primarily toward in vivo target therapeutic interventions and diagnostics. Nanorobotic research and development efforts for in vivo applications were initiated in great part following the first demonstration of the MR-Sub (Magnetic Resonance Submarine) project where an untethered object was successfully navigated automatically for the first time in the blood vessels of a living animal [14] with the experimental data published in 2007. Although this field is relatively new, substantial advanced experimental results have been recently published, providing important insights about the future possibilities for medical interventions in the human body. But as shown in the following sections, several types of microscale nanorobots have been and still investigated for such bioengineering applications.

Artificial and Synthetic Microscale Nanorobotics: An Introduction

Although both artificial and synthetic robots refer to man-made robots, synthetic robots although artificial

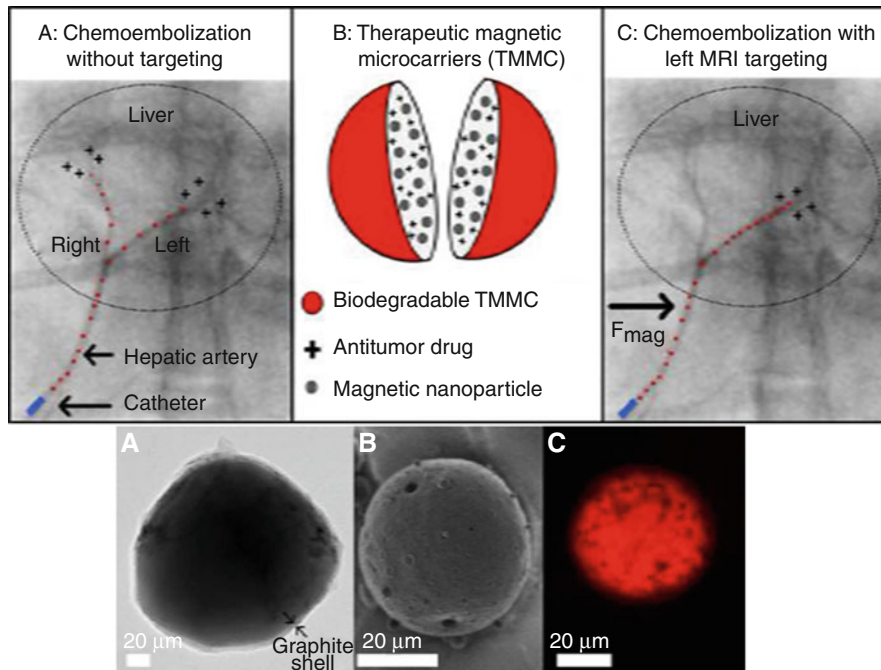
often imply, unlike non-synthetic artificial robots, the use of a chemical process to produce such robots. Hence, synthetic microscale nanorobots can typically be produced in large quantities at a relatively lower cost compared to the production of non-synthetic artificial (later simply referred to as artificial) microscale nanorobots.

The concept of integrating processes based on self-assembly for the synthesis of synthetic microscale nanorobots is an important yet critical aspect since for many targeted nano-medical interventions, large quantities of the so-called synthetic microscale nanorobots produced at a relatively low cost would typically be required, e.g., in targeted drug delivery through the vasculature in cancer therapy. On the other hand, one or a relatively small number of artificial microscale nanorobots could be useful in other particular medical interventions.

But at these scales, one of the primary challenges shared by both artificial and synthetic microscale nanorobotics is power. Traditional onboard power supplies such as batteries cannot be scaled at these levels and therefore, two main approaches have been considered, namely, relying on an external source of energy or harvesting energy from the environment. For artificial and synthetic microscale robots, the former has been preferred where in most cases a magnetic field gradient generated by electric coils and induced on the microscale robots was used to create motion.

Synthetic Microscale Nanorobotics: Therapeutic Magnetic Microcarriers (TMMCs)

The first untethered examples based on nanometer-scale components and depicted in Fig. 4 that were navigated through the blood vessels toward a predefined target in a living animal using the same method previously reported in Reference [14] with a larger untethered device being propelled by magnetic gradients have been recently described in Ref. [15]. Referred to as therapeutic magnetic microcarriers (TMMCs), these untethered microscale agents were synthesized to act like synthetic microscale nanorobots as defined previously. Indeed, each TMMC had an appropriate structure with a diameter of approximately 40 μm which was the most suitable dimension for targeted liver chemoembolization in live rabbits in these particular experiments. The structure was built with a biocompatible and biodegradable polymer that was able to hold all embedded components, namely,



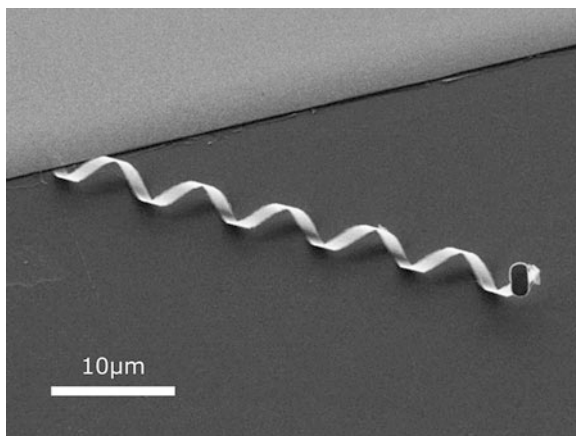
Nanorobotics for Bioengineering, Fig. 4 Unlike simple passive micro-entities that would be injected in the blood streams without control, TMMCs (**B-upper**) are synthesized to be compatible and integrated within a robotic environment allowing them to be controlled precisely as required to perform a given task such as the precise controlled navigation and delivery of therapeutic agent to a specific targeted location (**C-upper** versus

A-upper). In the lower section, photographs of the TMMC (**B**), with image of the loaded drug (**C**), are shown with one of the MNPs (**A**) allowing closed-loop MRN to be performed by inducing the correct directional propelling force on the embedded MNPs from MR-tracking information gathered from distortions in the MRI homogeneous magnetic field caused by the same MNPs (Reprinted from Ref. [15] with permission)

the magnetic nanoparticles (MNPs) and the therapeutic agent. The embedded MNPs allowed propulsion and navigation of the TMMCs in the blood vessels by inducing propelling force using magnetic gradients generated by an upgraded clinical magnetic resonance imaging (MRI) system, a method referred to as magnetic resonance navigation (MRN). The same MNPs were also acting as MRI contrast agents and as such, they enabled the TMMCs to be tracked by MRI, providing the possibility for feedback information required to execute closed-loop navigation control. Like other forms of robotic systems such as the ones relying on a central computer to remotely control a robotic arm or manipulator, this approach also relies on computation to remotely control special agents designed to perform a particular task. While the larger and more familiar robots often use mechanical means to embed functionalities (e.g., mechanical grippers), microscale nanorobots for bioengineering would not typically use mechanical parts and as such, they would not appear as conventional robots as we know since

they rely more on less familiar fields such as nanotechnology and biochemistry to implement embedded functionalities (e.g., with the use of antibodies instead of mechanical grippers).

But with MRN, the propelling force induced decreases with the quantity of magnetic material embedded in the microscale nanorobots and hence, larger magnetic gradients must be generated to navigate smaller nanorobots efficiently in tinier blood vessels. In an upgraded clinical MRI platform, an approximate tenfold increase in magnetic gradients is technologically feasible from conventional scanners, increasing the typical 40 m-Teslas (mT) per meter (m) 3D gradients of many conventional MRI systems up to approximately $400 \text{ mT} \cdot \text{m}^{-1}$ in upgraded MRI scanners (referred to as MRN systems, i.e., upgraded MRI scanners for MRN operations). Despite the increase in magnetic gradients, miniaturization of the microscale nanorobots conceived for MRN-based interventions is still limited to a few tens of micrometers in diameters. To navigate in the microvasculature and in the tiniest



Nanorobotics for Bioengineering, Fig. 5 Photograph of an artificial bacterial flagellum about half as long as the thickness of a human hair. This one can swim at a speed of up to one body length per second. (Image: Institute of Robotics and Intelligent Systems/ETH Zurich) (Image reprinted with permission)

blood vessels found in humans and required to target deeper inside tumors, microscale nanorobots with an overall diameter of approximately $2\ \mu\text{m}$ are required. Therefore, in order to compensate for the limitation of generating larger magnetic gradients for interventions in humans, other approaches have been investigated.

Biologically Inspired Artificial Microscale Nanorobotics: Artificial Bacterial Flagella (ABF)

One of the best known approaches to compensate for the limitation in generating larger magnetic gradients is referred to as artificial bacterial flagella (ABF) where a good up-to-date review is given in Ref. [16]. ABF (one example is depicted in Fig. 5) are good examples of biological-inspired artificial devices that can act under computer closed-loop control. Indeed, in nature, microorganisms swim in a low Reynolds number regime using a variety of techniques that differ from macroscale swimmers. The most studied experimental technique in micro- and nanorobotics due in great part to the possibility in micro- and nanofabrication technologies has been flagellated propulsion systems.

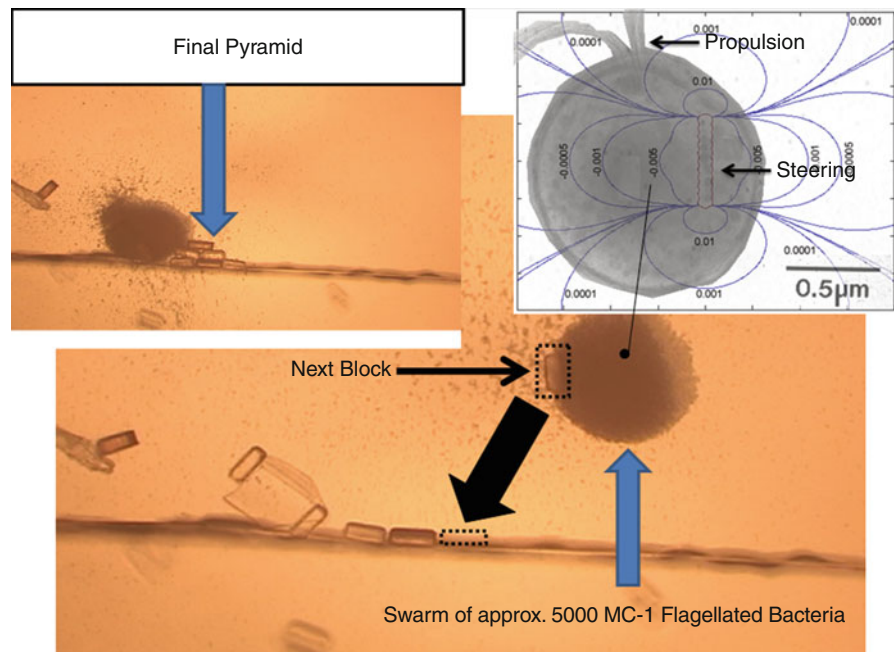
ABF typically rely on a rotating magnetic field creating a torque instead of a pulling force as for MRN to move the robot forward. Since creating a torque requires much lower magnitudes of magnetic gradients compared to pulling as done during MRN

operations, the gradients required could be decreased to a few $\text{mT}\cdot\text{m}^{-1}$ instead of a few hundred $\text{mT}\cdot\text{m}^{-1}$. But to date, no known experiments have been done yet in the blood vessels using ABF. Although the approach has advantages, several issues would have to be addressed before being adopted for targeted interventions through the human's vasculature. Two critical aspects are the biodegradation and response to the immune system since once located deep in the human's body, possibility for recovery will be unlikely. For the TMMCs described earlier, these issues have already been resolved by using materials that are biodegradable and biocompatible, and MNPs with dimensions that already prove to be compatible with the human's immune system.

Another consideration for robotic operations is the possibility of gathering positional feedback data for closed-loop navigational control. The spatial resolution of all medical imaging modalities is too low for detecting a robot or an aggregation of robots with overall diameters of only a few micrometers or a few tens of micrometers, respectively, due to the space available in the microvasculature. X-ray systems have among the best spatial resolution to image inside humans (approximately $300\ \mu\text{m}$), but they will fail to detect and hence track microscale nanorobots operating in the microvasculature such as in capillaries. Although MRI has typically a slightly lower spatial resolution than X-ray, the microscale nanorobots can be made such that the embedded magnetic particles would create a magnetic field distortion much larger than the robot itself and large enough to be visible with MRI. Past studies have shown that a single magnetic object as small as $10\ \mu\text{m}$ could be detected inside the human's body using MRI. But because MRI scanners have a high homogeneous magnetic field referred to as the B_0 field (1.5 T or 3 T in most clinical MRI scanners) along a given direction, the shape of the robots must ideally be isotropic such as the spherical shapes of the TMMCs depicted in Fig. 4. For instance, the B_0 field would induce a directional torque on ABF if operated inside a MRI scanner such that they will always point to the same direction during the entire intervention and, hence, making directional control or steering for targeting purpose impossible. On the other hand, because of the spherical shape of each TMMC, steering in all directions is not an issue. But one fact remains true is that, up to now, TMMCs are still too large to be navigated in the microvasculature and, as

Nanorobotics for Bioengineering,

Fig. 6 Computer control of a swarm of approximately 5,000 magnetotactic bacteria is demonstrated by the coordinated construction of a tiny pyramidal structure using blocks with length approximately equivalent to a human's hair thickness



such, the idea of using a propulsion method suitable for low Reynolds hydrodynamic conditions, such as flagella-based propulsion allowing for the implementation of smaller untethered devices, is still a valid and promising approach.

Natural Microscale Nanorobotics: Bacterial Nanorobots

One strategy that was recently considered in nanorobotics was to harness what nature provides instead of mimicking it using modern technologies. An example that showed its potential through experiments conducted in the blood vessels of living mice is the use of MC-1 flagellated magnetotactic bacteria (MTB) acting as bacterial nanorobots or natural microscale nanorobots [17]. Although most people would see this bacterium as a simple living microorganism, in an engineering point of view, this microorganism can be considered as a sophisticated robot with a very effective embedded propulsion system and control interface. The possibilities of computer closed-loop control of these MTBs compared to potential far-futuristic artificial nanorobots can be appreciated by observing the example of a tiny pyramid being built one component at a time by a swarm of approximately 5,000 MC-1 flagellated bacteria (see Fig. 6).

As depicted in the upper-right corner of Fig. 6, the cell of the MC-1 bacterium is spherical in shape and measures between 1 and 2 μm in diameter, making it ideal for navigation in the tiniest blood vessels found in humans. Each cell has two flagella bundles providing a thrust force exceeding 4 pN, making it a serious candidate for operations in the human's microvasculature. This value is also relatively high compared to other known species of flagellated bacteria characterized with a typical thrust force in the range of 0.3–0.5 pN. This allows the MC-1 bacteria to swim in water and human's blood at speeds often exceeding 200 $\mu\text{m}/\text{s}$. This is a very high speed when we know that the swimming speeds of other well-known species of flagellated bacteria in the same conditions are ~ 30 $\mu\text{m}/\text{s}$. Maximum displacement speeds of ~ 300 $\mu\text{m}/\text{s}$ and corresponding to ~ 150 times its cell's length for a relatively large proportion of the MC-1 bacteria samples have also been recorded experimentally.

Unlike most bacteria that are based on chemotaxis to detect nutrient gradients and hence influence their motility, the direction of displacement of MTB with their chain of magnetosomes, which are membrane-based MNPs that act like an embedded nano-compass, can be influenced by a directional magnetic field, hence allowing for magnetotaxis-based control by

inducing a directional torque on the same chain acting as a nano-steering system. Since this magnetic nano-compass is extremely sensitive, an extremely low magnitude magnetic field is required to influence their directional motions.

Since the same MNPs used for directional control also act as MRI contrast agents, it was shown that a sufficiently large aggregate of MTBs can be detected using a clinical MRI scanner. But as for the ABFs, MTBs cannot change direction while operating in the MRI scanner. But unlike artificial or synthetic microscale nanorobots, these microorganisms do not need real-time closed-loop navigation control but rather computer guidance for determining their positions using MRI at relatively long time intervals in order to guide them toward the targeted site. Indeed, MTBs when encountering an obstacle will not require real-time closed-loop navigation control to direct them around the obstacle as an artificial robot would, but they will typically swim without any external control around obstacles and toward the direction indicated through magnetotaxis using a special platform named a magnetotaxis system, a platform located outside the MRI scanner. This self-contact avoidance capability is a significant advantage compared to artificial versions since real-time closed-loop navigation control is not possible in these regions considering that the microvasculature cannot be imaged due to a lack of spatial resolution in existing medical imaging modalities. Nonetheless, although these bacteria can swim efficiently in the microvasculature, they are ineffective in larger blood vessels such as the arteries and the arterioles where blood flow velocities are too high. Hence, a mean of transport is needed to bring such bacterial nanorobots from the injection location in an artery to the entrance of the microvascular network.

Hybrid Microscale Nanorobotics: Bio-nanorobots

One solution that is under study for the transport of bacterial nanorobots toward the microvasculature is the use of microscale liposomal transporters. Already, such transporters with diameters of a few tens of micrometers have been synthesized with free swimming MC-1 bacteria inside. Although not tested yet in a living animal, these liposomal microscale transporters can be synthesized with MNPs acting as propulsion engines and for MRI-based tracking for MRN operations. This is a perfect example of hybrid microscale nanorobots where biochemistry

and biology are used in their implementations. Since biological components are also used, they could also be referred to as bio-nanorobots. Bio-nanorobots could also refer to artificial or synthetic robots relying on nanostructures such as MNPs being functionalized with biomolecules. Indeed, often in targeted therapeutic or diagnostic applications, microscale nanorobots would be functionalized, i.e., they will be coated with molecular components, e.g., antibodies, that have specificity for the targeted cells such as specific cancerous cells. Functionalized hybrid or bio-nanorobots could not only help achieving more retention when coming in contact with the targeted cells for targeted therapies, but functionalized microscale nanorobots could also be potentially useful in targeted diagnostics as well.

Conclusions

The two main forms of nanorobotic platforms for practical applications in bioengineering are the macroscale nanorobotic platforms and the microscale nanorobotic platforms which are mainly dedicated to biology and in vivo medical interventions, respectively. Macroscale nanorobotic systems rely mainly on AFM for contact-mode nanomanipulation and studies of biological samples while optical tweezers are the mostly used for non-contact applications. Microscale nanorobots can be artificial, synthetic, or natural. Synthetic versions called TMMCs based on biodegradable polymer and MNPs and capable of releasing therapeutic agents have already been navigated to target sites in live animals using an upgraded clinical MRI scanner but the miniaturization of these synthetic carriers is limited to a few tens of micrometers in diameter due to limitation in increasing magnetic gradients beyond approximately half a Tesla for human subjects. To compensate for such limitation, artificial versions called ABF using a rotary magnetic field to induce a torque to mimic the flagellum of bacteria have been proposed. Although the magnitude of the magnetic gradients could be reduced significantly, several issues would need to be resolved before ABFs can be used for in vivo interventions in the blood vessels. Instead of a bio-mimetic artificial version, another strategy relies on harnessing natural bacteria and more specifically, the MC-1 flagellated magnetotactic bacteria. But the latter would need to be encapsulated in special

microscale carriers for many target medical interventions. As such, hybrid microscale nanorobots are likely to play an important role in medical interventions conducted deeply in the vasculature, especially hybrid functionalized nanorobots.

This chapter only provides a quick overview of the present state of the art for nanorobotics in bioengineering, and for more details, the readers should consult adequate literature. Although macroscale nanorobotics platforms are presently more mature than microscale nanorobotic systems and techniques, the latter field is progressing at a very fast pace and research in this field may soon look at developing nanoscale nanorobots as needed to transit through the blood–brain barrier (BBB).

Cross-References

- ▶ [Angiogenesis](#)
- ▶ [Atomic Force Microscopy](#)
- ▶ [Biomimetics](#)
- ▶ [Biosensors](#)
- ▶ [CMOS MEMS Biosensors](#)
- ▶ [Dielectrophoresis](#)
- ▶ [DNA Manipulation Based on Nanotweezers](#)
- ▶ [Liposomes](#)
- ▶ [Nanogrippers](#)
- ▶ [Nanomechanical Properties of Nanostructures](#)
- ▶ [Nanomedicine](#)
- ▶ [Nanoparticles](#)
- ▶ [Nanorobotic Assembly](#)
- ▶ [Nanorobotic Manipulation of Biological Cells](#)
- ▶ [Nanorobotics](#)
- ▶ [Nanotechnology](#)
- ▶ [Optical Tweezers](#)
- ▶ [Robot-Based Automation on the Nanoscale](#)
- ▶ [Scanning Tunneling Microscopy](#)

References

1. Binnig, G., Rohrer, H.: Scanning tunneling microscopy. *IBM J. Res. Dev.* **30**, 4 (1986)
2. Binnig, G., Quate, C.F., Gerber, Ch: Atomic force microscope. *Phys. Rev. Lett.* **56**, 930–933 (1986)
3. Sitti, M., Hashimoto, H.: Tele-nanorobotics using atomic force microscope. In: *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Victoria, pp. 1739–1746, Oct 1998
4. Mokaberi, B., Yun, J., Wang, M., Requicha, A.A.G.: Automated nanomanipulation with atomic force microscopes. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'07)*, Rome, pp. 1406–1412, 10–14 Apr 2007
5. Rege, K., Medintz, I.L.: *Methods in Bioengineering: Nanoscale Bioengineering and Nanomedicine*, p. 332. Artech House, Boston (2009). 978-1-59693-410-8
6. Hu, J., Zhang, Y., Li, B., Gao, H.B., Hartmann, U., Li, M.Q.: Nanomanipulation of single DNA molecules and its applications. *Surf. Interface Anal.* **36**(2), 124–126 (2004)
7. Martel, S., Madden, P., Sosnowski, L., Hunter I., Lafontaine S.: NanoWalker: a fully autonomous highly integrated miniature robot for nanoscale measurements. In: *Proc. SPIE* **3825**, p. 111 (1999); doi:10.1117/12.364292
8. Martel, S., Hunter, I.: Nanofactories based on a fleet of scientific instruments configured as miniature autonomous robots. *J. Micromechatronics* **2**(3–4), 201–214 (2004)
9. Brufau, J., Puig-Vidal, M., López-Sánchez, J., Samitier, J., Driesen, W., Breguet, J.-M., Snis, N., Simu, U., Johansson, S., Gao, J., Velten, T., Seyfried, J., Estaña, R., and Woern, H.: MICRON: Small autonomous robot for cell manipulation applications. In: *Proceedings of 2005 IEEE International Conference on Robotics and Automation*, Barcelona, 18–22 Apr 2005
10. Sun, Y., Nelson, B.: Biological cell injection using an autonomous microrobotic system. *Int. J. Robot. Res.* **21**(10–11), 861–868 (2002)
11. Fatikow, S.: *Automated Nanohandling by Microrobots*. Springer Series in Advanced Manufacturing, p. 346. Springer, London (2008)
12. Grier, D.G.: A revolution in optical manipulation. *Nature* **424**, 810–816 (2003)
13. Bennink, M., Schaver, O., Kanaar, R., et al.: Single molecule manipulation of double-stranded DNA with RecA and Yoyo-1. *Cytometry* **36**, 200–208 (1999)
14. Martel, S., Mathieu, J.-B., Felfoul, O., Chanu, A., Aboussouan, E., Tamz, S., Pouponneau, P.: Automatic navigation of an untethered device in the artery of a living animal using a conventional clinical magnetic imaging system. *Appl. Phys. Lett.* **90**, 114105 (2007)
15. Pouponneau, P., Leroux, J.-C., Soulez, G., Gaboury, L., Martel, S.: Coencapsulation of magnetic nanoparticles and doxorubicin into biodegradable microcarriers for deep tissue targeting by vascular MRI navigation. *Biomaterials* **32**(13), 3481–3486 (2011)
16. Zhang, L., Peyer, K.E., Nelson, B.J.: Artificial bacterial flagella for micromanipulation. *Lab Chip* **10**, 2203–2215 (2010)
17. Martel, S., Mohammadi, M., Felfoul, O., Lu, Z., Pouponneau, P.: Flagellated magnetotactic bacteria as controlled MRI-trackable propulsion and steering systems for medical nanorobots operating in the human microvasculature. *Int. J. Robot. Res.* **28**(4), 571–582 (2009)

Nanorobotics for MEMS and NEMS

- ▶ [Robot-Based Automation on the Nanoscale](#)

Nanorobotics for NEMS Using Helical Nanostructures

Didi Xu¹, Li Zhang¹, Lixin Dong² and Bradley J. Nelson¹

¹Institute of Robotics and Intelligent Systems, ETH Zurich, Zurich, Switzerland

²Electrical and Computer Engineering, Michigan State University, East Lansing, MI, USA

Synonyms

Helical nanobelt; Nanocoil; Nanohelix; Nanospring; Rolled-up nanostructure; Scrolled nanostructure

Definition

Three-dimensional helical nanostructures, such as carbon nanotube coils (CNTs) [1], zinc oxide helical nanobelts [2], and rolled-up helical nanostructures [3–7] (Fig. 1 [5, 8–10]) have attracted intensive research interest because of their potential applications in nanoelectromechanical systems (NEMS) as springs, electromagnets, inductors, resonators, sensors, and actuators [11–14]. These NEMS will serve as the tools for fabricating future nanorobots, and in the meantime basic components constructing those nanorobots. Three-dimensional helical nanostructures could be manipulated and characterized by the nanorobotic tools, which could be a good candidate for fascinating applications, such as, mass sensors in femto-gram ranges, force sensors at pico-Newton scales, and resonators at GHz ranges.

Physical and Chemical Properties

Bottom-up Fabrication of Helical Nanostructures

Top-down and bottom-up nanofabrication strategies are being independently investigated by various researchers to realize helical nanostructures. Bottom-up strategies are assembly-based techniques, resulting in controllable nanopatterns at large scales, such as carbon nanotube coils [1], silica nanosprings [15], and SiC nanohelices [16]. For example, helical

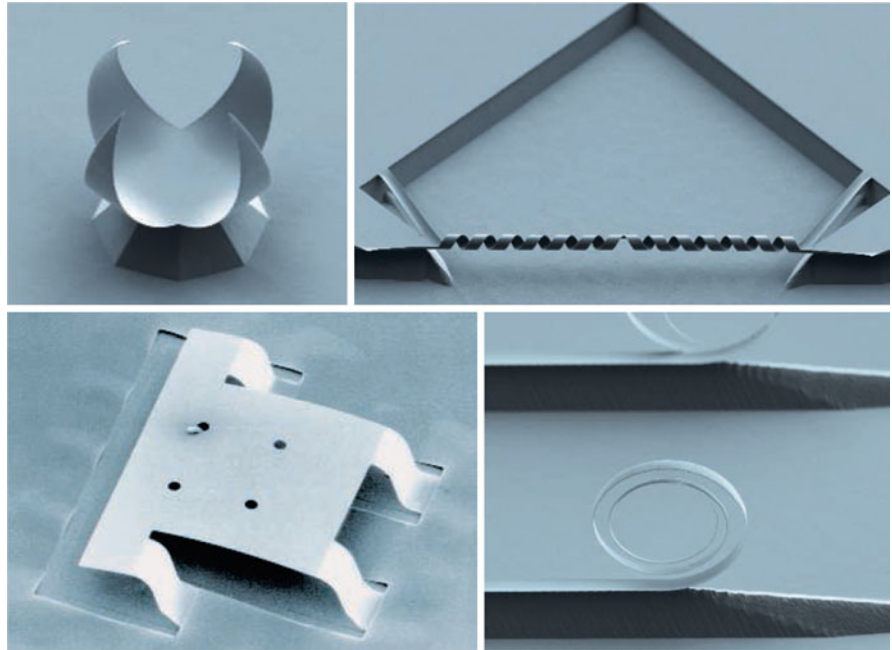
nanostructures of semiconducting zinc oxide (ZnO) were grown using a vapor-solid process. A rigid structural alteration, as a result of superlattice formation [17], gives rise to the formation of a helical structure, either left-handed or right-handed, as shown in Fig. 2. The diameters, widths, and pitch distances of the nanohelix range between 300–700, 100–500, and 500–2,500 nm, respectively. The length of the as-fabricated ZnO nanohelix can be as long as 100 μm and the thickness of the helical ribbon is less than 20 nm [17].

Top-down Fabrication of Helical Nanostructures

Top-down approaches are based on conventional fabrication process, including nanolithography, nanoimprinting, and chemical etching. Nanocoils are created through a top-down fabrication process in which a strained nanometer-thick heteroepitaxial layer of semiconductor compounds curls up to form 3-D structures [4]. Using GaAs-InAs bilayer as an example, when the basic bilayer film is released from an InP substrate by selective etching, the elastic forces of two layers are oppositely directed, generating a nonzero moment of forces, therefore its top layer contracts and its bottom layer expands, causing the film to curl and roll up [4]. The preferential direction for the film to roll up depending on the orientation of a mesa stripe on the substrate, either tube-like or helical nanobelts, can be formed [5]. A helical geometry based on the bilayer SiGe/Si can be achieved through this process (Fig. 3) [9]. The gray arrows indicate the $\langle 100 \rangle$ directions, that is, the smallest Young modulus direction on a $\langle 001 \rangle$ substrate, in which the bilayer will start to roll up when it has been freed from the substrate. Several techniques have been proposed to improve the control over the fabrication process in terms of the resulting length, shape, and orientation of structures. It has been demonstrated that the misaligned angle between the stripe and etching direction and the edge effect of the helical-shaped ribbon are crucial to determine the chirality and pitch of a helical nanobelt [5, 9]. Moreover, various materials, such as metal, dielectrics, and polymers, can be integrated into the helical nanobelt by thin film deposition and lithographic patterning [9]. It is worthy of note that nanohelices with a minimum diameter of 7 nm were obtained using 6-monolayer-thick InGaAs/GaAs bilayers [4].

Nanorobotics for NEMS Using Helical Nanostructures,

Fig. 1 Various helical nanostructures. *Bottom left* is reproduced with permission from ref. [8], (Copyright 2003 APEX/JJAP. The others are in courtesy of L. Zhang and D. Grützmacher)



Key Research Findings

Nanorobotic manipulation [18] enables property characterization to be performed after intermediate processes, and in situ characterization can be performed using manipulation rather than conventional static observations. Nanorobotic approach extends the lower limit of robotic manipulation to the dimension of nanometers, and it paves the way for the design and fabrication of nanocomponents having applications in NEMS.

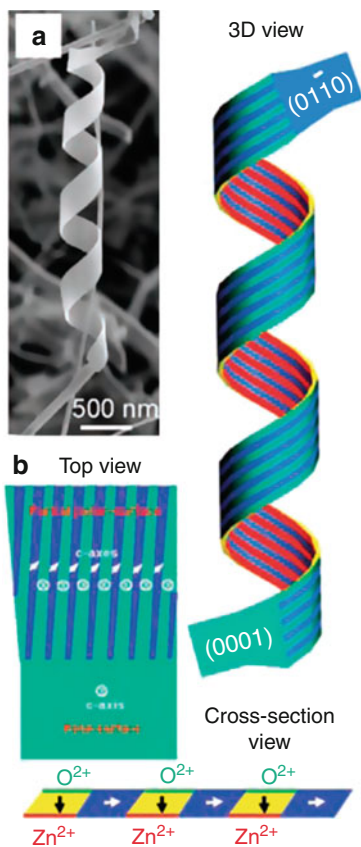
Mechanical Properties of Helical Structures

By in situ manipulation using a nanoprobe, mechanical properties of various helical structures have been investigated, such as superlattice-structured ZnO nanohelices, carbon nanocoils, and rolled-up nanosprings [11, 12, 14]. It is found that the superelasticity (shape memory) of superlattice-structured ZnO nanohelices enables the nanohelix to recover its shape after an extremely large axial stretching and compressing, which could be of great interest for fabricating nanoscale elastic energy storage in microelectromechanical systems (MEMS) and NEMS. The manipulation experiments of rolled up helical nanostructures can be performed by a nanomanipulator and an atomic force microscope (AFM) cantilever inside a scanning electron microscope (SEM). The nanohelix was picked up by welding its one end using platinum (Pt) deposition onto the tip of a tungsten nanoprobe, and the

mechanical deformation was recorded via extending or compressing the nanohelix using the nanoprobe. As a result, its spring constant can be increased continuously for up to 300–800%.

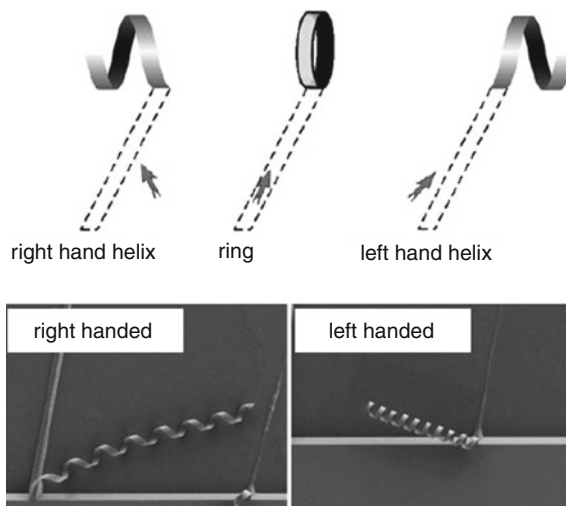
In Fig. 4a, b, it is evident that the welded nanohelix can be pulled to an almost straightened shape, with its twist cycles preserved. Next, the welded nanohelix is released and started to restore its original shape, as shown in Fig. 4c. With comparison to the initial nanohelix dimensions in Fig. 4a, it indicates that the nanohelix has an almost identical dimensionality including pitch and radius, suggesting a complete elastic recovery from stretching. Similar mechanical behavior, that is, the superelasticity (shape memory) behavior, has also been observed when applying a compression load on the welded nanohelix (Fig. 4d, e). The spring constant of ZnO nanohelices is in a range of 0.07–4.2 N/m, and the transverse fracture force is in the micro-Newton range [12, 17].

Nanomanipulators can also be used to carry out the mechanical characterization of rolled-up helical structure. A probe picked up an InGaAs/GaAs nanospring and attached it to the AFM tip (Fig. 5a, b). Then, a tensile force was applied to the nanospring, while continuous SEM images were taken to detect the AFM tip displacement and the nanospring deformation. When the tensile force was increased dramatically, the attachment between the nanospring and the AFM cantilever



Nanorobotics for NEMS Using Helical Nanostructures, Fig. 2 Initiation and formation process of a nanohelix. (a) Low magnification SEM image of a ZnO nanohelix showing its starting point and finishing end. (b) A three-dimensional (3-D) schematic model of a nanohelix, showing its initiating point and finishing end. The periodicity of the superlattices may result in the formation of periodic piezoelectric domains (Reproduced with permission from ref. [17], Copyright 2005 AAAS)

broke, and the nanosprings returned to their initial shape at the zero-displacement position. From the displacement data and the known stiffness of the AFM cantilever, the tensile force acting on the nanosprings versus the nanospring displacement was plotted (Fig. 5c). For the nanospring, an overall elastic strain of 48% was measured when the attachment to the AFM tip broke. Ultra-flexible rolled-up Si-based nanospring with a spring constant of 0.003 N/m was also reported which exhibits long-range linear elasticity (with ca. 189% relative elongation) and excellent mechanical stability [19]. The rolled-up structure is promising for use in electromechanical sensors, such as force or pressure sensors, because of intrinsic piezoresistive and piezoelectric properties of these materials.



Nanorobotics for NEMS Using Helical Nanostructures, Fig. 3 Illustration of the effect of the angle between the mesa stripe and preferred $\langle 100 \rangle$ etch direction (gray arrows) on the formation of rings or helices with right- and left-handed chirality. The thickness of the SiGe/Si/Cr ribbons and the diameter of the helices are 40 nm and 3 μm , respectively (Reproduced with permission from ref. [9], Copyright 2006 ACS. Reproduced with permission from ref. [5], Copyright 2005 IOP)

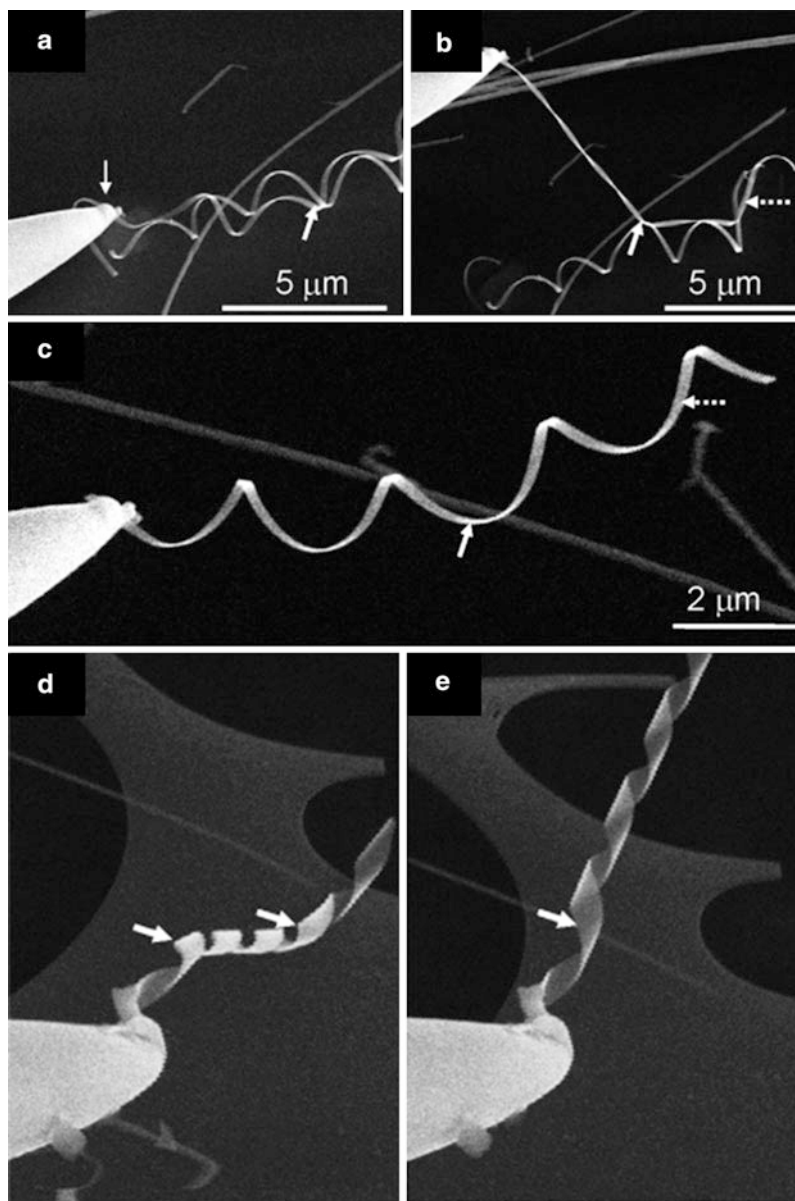
Linear-to-Rotary Motion Converter

The conversion between various forms of motions will play an important role in future NEMS applications. Among various nano building blocks, 3-D helical nanostructures, for example, dual-chirality helical nanobelts (DCHNBs), could realize conversion between linear and rotary motion [20]. As schematically shown in Fig. 6a, the motion converter consists of a DCHNB with a left-handed and a right-handed part. By linearly stretching both ends of the DCHNB, the central part was mobilized in rotary motion; allowing for linear-to-rotary motion conversion, which has not been previously demonstrated in other nanostructures. With an extended arm, the output can be linear (small displacement) or rotary (large displacement) motion, as seen in Fig. 6b. When a tensile force F is applied to the DCHNB, it elongates and rotates about the unwinding direction.

To implement and characterize motion converters, experimental investigations have been performed in the SEM using a nanomanipulator equipped with a tungsten probe and an AFM cantilever. Manipulation of an as-fabricated DCHNB was performed by cutting the lower end of the DCHNB shown in Fig. 7a.

Nanorobotics for NEMS Using Helical Nanostructures,

Fig. 4 Manipulation process of a nanohelix during axial stretching and compressing. (a) One end of a nanohelix was welded with Pt onto a tungsten nanoprobe. (b) A SEM image shows the extremely stretched nanohelix. (c) A complete restoring of the nanohelix shape after releasing. (d, e) Compressed deformation process of a nanohelix induced by the nanoprobe. Both stretching and compressing suggest the superelasticity behavior (Reproduced with permission from ref. [12], Copyright 2006 ACS)



A “sticky” probe was then prepared by dipping a tungsten probe (Picoprobe, T-4-10-1 mm) into a silver tape and attached on the lower end of the DCHNB. Rotation was then generated in the central part by moving the probe downward (Fig. 7b–e).

Examples of Application

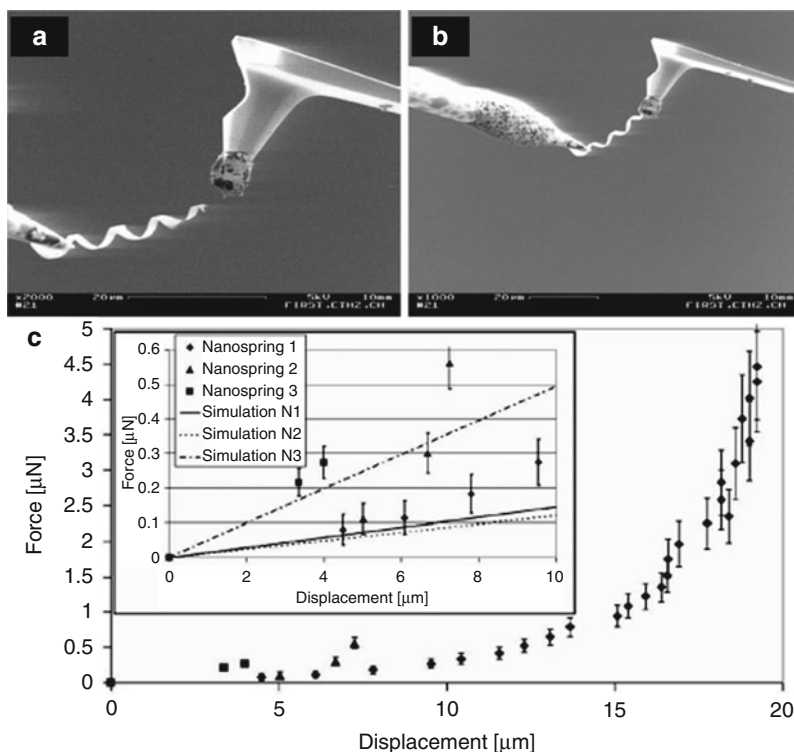
A fundamental application of the rolled up nanohelix is the 3-D microscopy based on the linear-to-rotary

motion conversion [20], in which samples are rotated so that different views are exposed to light, electron beams, or focused-ion beams (FIBs). Due to their extremely compact size, the dual-chirality helical nanobelts are well suited to serve as rotary scanners for creating 3-D scanning probe microscope (SPMs), rotary stages for a microgoniometer for observing an object from different crystal surfaces, and components of nanomachines.

Figure 8 shows 3-D imaging of a pollen grain. The pollen grain is picked up by a Picoprobe and attached

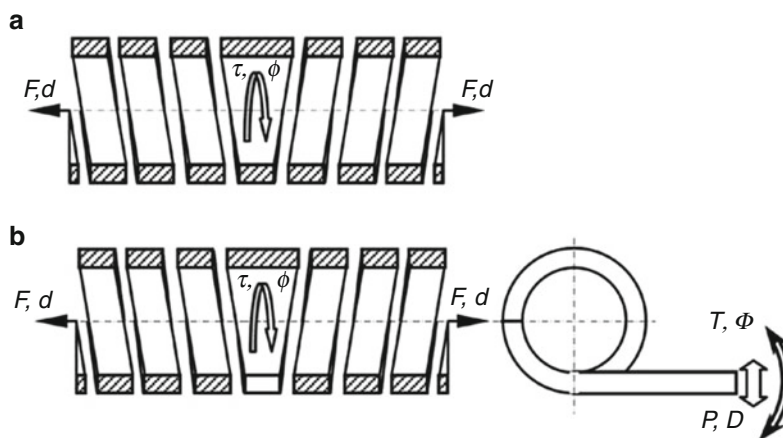
Nanorobotics for NEMS Using Helical Nanostructures,

Fig. 5 Mechanical characterization. (a, b) Break and pick up nanospring and attach it to AFM tip. (c) Simulation and experimental results of axial force versus nanospring displacement (Reproduced with permission from ref. [14], Copyright 2006 ACS)



Nanorobotics for NEMS Using Helical Nanostructures,

Fig. 6 Motion converter based on the dual-chirality helical nanobelts (DCHNB). (a) Linear-to-rotary motion converter using a DCHNB. (b) Transmission converting linear motion to extended linear (small displacement)/ rotary (large displacement) motion (Reproduced with permission from ref. [20], Copyright 2009 IEEE)

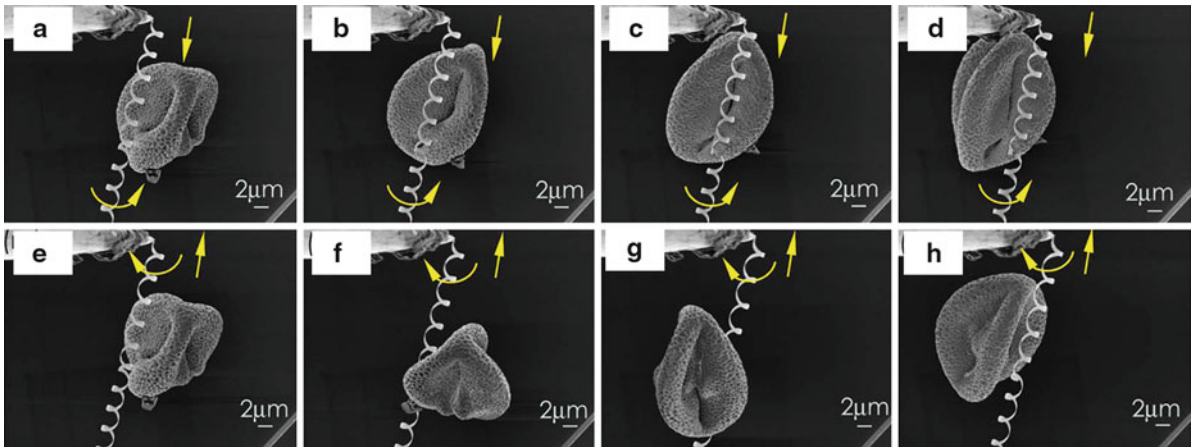
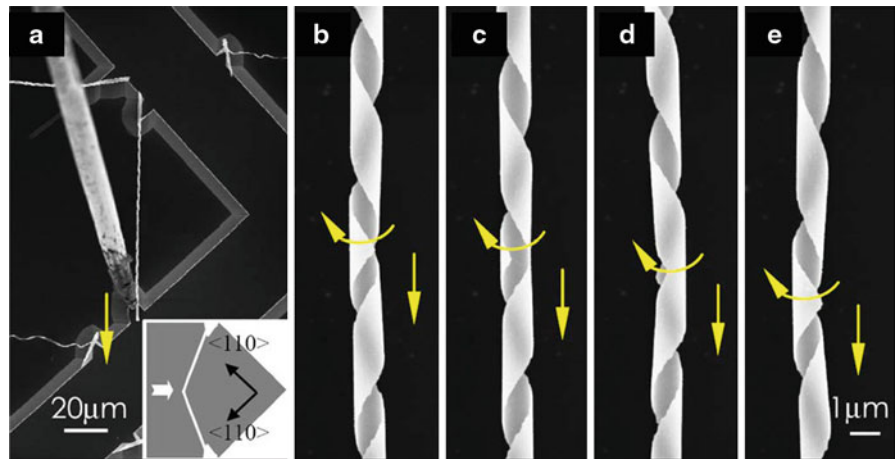


to the DCHNB. The motion converter has been characterized before loading the sample. Figure 8a–d shows that the sample rotates counterclockwise (top view) for approximately 180° when releasing the extended DCHNB. Figure 8e–h shows further stretching of the DCHNB, and the sample rotates clockwise (top view) for another 180° . It can be seen that the different aspects of the sample can be exposed

to the electron beam simply by extending or releasing the DCHNB with a small displacement. Further possibilities for this converter include goniometry for a transmission electron microscope (TEM), tomography for a SEM or a TEM, and 3-D SPM. By integrating in the fabrication processes of DCHNBs, rolled-up spirals can serve as claws for holding relatively large samples [6].

Nanorobotics for NEMS Using Helical Nanostructures, Fig. 7

In situ characterization of motion conversion using nanorobotic manipulation. (a) One end of an as-fabricated DCHNB is cut and attached to a “sticky” probe. (b, e) Rotation of the central part while moving the probe downward (Reproduced with permission from ref. [20], Copyright 2009 IEEE)



Nanorobotics for NEMS Using Helical Nanostructures, Fig. 8 Three-dimensional microscopy of a pollen grain. (a–d) When releasing the extended DCHNB, the sample rotates

counterclockwise (*top view*). (e–h) When stretching further the DCHNB, the sample rotates clockwise (*top view*) (Reproduced with permission from ref. [20], Copyright 2009 IEEE)

References

- Amelinckx, S., Zhang, X.B., Bernaerts, D., Zhang, X.F., Ivanov, V., Nagy, J.B.: A formation mechanism for catalytically grown helix-shaped graphite nanotubes. *Science* **265**(5172), 635–639 (1994)
- Kong, X.Y., Wang, Z.L.: Spontaneous polarization-induced nanohelices, nanosprings, and nanorings of piezoelectric nanobelts. *Nano Lett.* **3**(12), 1625–1631 (2003)
- Schmidt, O.G., Eberl, K.: Nanotechnology – thin solid films roll up into nanotubes. *Nature* **410**(6825), 168–168 (2001)
- Prinz, V.Y., Seleznev, V.A., Gutakovsky, A.K., Chehovskiy, A.V., Preobrazhenskii, V.V., Putyato, M.A., Gavrilova, T.A.: Free-standing and overgrown InGaAs/GaAs nanotubes, nanohelices and their arrays. *Physica E* **6**(1–4), 828–831 (2000)
- Zhang, L., Deckhardt, E., Weber, A., Schonenberger, C., Grutzmacher, D.: Controllable fabrication of SiGe/Si and SiGe/Si/Cr helical nanobelts. *Nanotechnology* **16**(6), 655–663 (2005)
- Zhang, L., Dong, L.X., Bell, D.J., Nelson, B.J., Schonenberger, C., Grutzmacher, D.: Fabrication and characterization of freestanding Si/Cr micro- and nanospirals. *Microelectron. Eng.* **83**(4–9), 1237–1240 (2006)
- Schmidt, O.G., Schmarje, N., Deneke, C., Muller, C., Jin-Phillipp, N.Y.: Three-dimensional nano-objects evolving from a two-dimensional layer technology. *Adv. Mater.* **13**(10), 756–759 (2001)
- Kubota, K., Fleischmann, T., Saravanan, S., Vaccaro, P.O., Aida, T.: Self-assembly of microstage using micro-origami technique on GaAs. *Jpn. J. Appl. Phys* **42**(6B), 4079–4083 (2003). Part 1 – Regular Papers Short Notes & Review Papers
- Zhang, L., Ruh, E., Grutzmacher, D., Dong, L.X., Bell, D.J., Nelson, B.J., Schonenberger, C.: Anomalous coiling of SiGe/Si and SiGe/Si/Cr helical nanobelts. *Nano Lett.* **6**(7), 1311–1317 (2006)

10. Cho, A.: News focus: nanotechnology: pretty as you please, curling films turn themselves into nanodevices. *Science* **313**, 164–165 (2006)
11. Chen, X.Q., Zhang, S.L., Dikin, D.A., Ding, W.Q., Ruoff, R.S., Pan, L.J., Nakayama, Y.: Mechanics of a carbon nanocoil. *Nano Lett.* **3**(9), 1299–1304 (2003)
12. Gao, P.X., Mai, W.J., Wang, Z.L.: Superelasticity and nanofracture mechanics of ZnO nanohelices. *Nano Lett.* **6**(11), 2536–2543 (2006)
13. Bell, D.J., Sun, Y., Zhang, L., Dong, L.X., Nelson, B.J., Grützmacher, D.: Three-dimensional nanosprings for electromechanical sensors. *Sens. Actuator A-Phys.* **130–131**, 54–61 (2006)
14. Bell, D.J., Dong, L.X., Nelson, B.J., Golling, M., Zhang, L., Grützmacher, D.: Fabrication and characterization of three-dimensional InGaAs/GaAs nanosprings. *Nano Lett.* **6**(4), 725–729 (2006)
15. Zhang, H.F., Wang, C.M., Buck, E.C., Wang, L.S.: Synthesis, characterization, and manipulation of helical SiO₂ nanosprings. *Nano Lett.* **3**(5), 577–580 (2003)
16. Zhang, D.Q., Alkhateeb, A., Han, H.M., Mahmood, H., McIlroy, D.N., Norton, M.G.: Silicon carbide nanosprings. *Nano Lett.* **3**(7), 983–987 (2003)
17. Gao, P.X., Ding, Y., Mai, W.J., Hughes, W.L., Lao, C.S., Wang, Z.L.: Conversion of zinc oxide nanobelts into superlattice-structured nanohelices. *Science* **309**(5741), 1700–1704 (2005)
18. Dong, L.X., Nelson, B.J.: Robotics in the small, part II: nanorobotics. *IEEE Robot. Autom. Mag.* **14**(3), 111–121 (2007)
19. Dai, L., Zhang, L., Dong, L.X., Shen, W.Z., Zhang, X.B., Ye, Z.Z., Nelson, B.J.: Long-range linear elasticity and mechanical instability of self-scrolling binormal nanohelices under a uniaxial load. *Nanoscale* **3**, 4301–4306 (2011)
20. Dong, L.X., Zhang, L., Kratochvil, B.E., Shou, K.Y., Nelson, B.J.: Dual-chirality helical nanobelts: linear-to-rotary motion converters for three-dimensional microscopy. *J. Microelectromech. Syst.* **18**(5), 1047–1053 (2009)

Nanorods

- ▶ [Nanostructures for Energy](#)
- ▶ [Physical Vapor Deposition](#)³

Nanoscaffold

- ▶ [Ligand-Directed Gold-Phage Nanosystems](#)

Nanoscale Drug Vector

- ▶ [Fullerenes for Drug Delivery](#)

Nanoscale Fluid Mechanics

- ▶ [Computational Micro/Nanofluidics: Unifier of Physical and Natural Sciences and Engineering](#)

Nanoscale Heat Transport

- ▶ [Thermal Conductivity and Phonon Transport](#)

Nanoscale Particle

- ▶ [Gas Phase Nanoparticle Formation](#)

Nanoscale Printing

Timothy J. Merkel¹ and Joseph M. DeSimone^{1,2,3,4,5,6,7,8}

¹Department of Chemistry, University of North Carolina, Chapel Hill, NC, USA

²Department of Pharmacology, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

³Carolina Center of Cancer Nanotechnology Excellence, University of North Carolina, Chapel Hill, NC, USA

⁴Institute for Advanced Materials, University of North Carolina, Chapel Hill, NC, USA

⁵Institute for Nanomedicine, University of North Carolina, Chapel Hill, NC, USA

⁶Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA

⁷Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC, USA

⁸Sloan–Kettering Institute for Cancer Research, Memorial Sloan–Kettering Cancer Center, New York, NY, USA

Synonyms

[Dip-pen nanolithography](#); [Imprint lithography](#); [Microcontact printing](#); [Molecular printing](#);

Nanoimprint lithography; Nanotransfer printing; PRINT[®]; Replica molding; Soft lithography

new industries, especially in the life sciences and energy sciences fields.

Definition

Nanoscale printing refers to the patterning of surfaces in two or three dimensions with at least one feature on the submicron length scale. The printing techniques may resemble macroscopic printing, in which a material, or ink, is transferred to a substrate by a stamp or a stylus. However, nanoscale printing may also refer to embossing or molding processes, in which topographical features are transferred from a three-dimensional mold to a surface by either elevating or depressing areas or by the addition of material. For the case of patterning features in three dimensions, two distinct outcomes are possible: the features generated can remain connected by an underlying layer of material, yielding an embossed film or, alternately, the molded features can be free of the connecting layer, yielding discrete particles. Arrays of features have been used to direct the assembly of materials for applications in molecular electronics, genetic analysis, and disease detection. Particles generated by these methods have found a myriad of uses, including those in drug delivery and biomedical imaging.

Overview

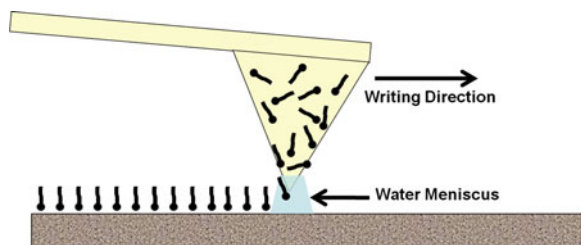
Methods used to generate nanoscale structures are generally divided into “top-down” and “bottom-up” approaches. In bottom-up approaches, features are generated from the atomic or molecular level to the supermolecular or macroscopic level and utilize interactions between molecules or particles to achieve two- or three-dimensional structures [1]. For top-down methods, materials are processed at the desired size scale, with external forces directing the material to form the desired architecture. The development of nanoscale printing methods for drug delivery and imaging applications have been enabled by the remarkable developments in the semiconductor industry over the last 50 years to reduce feature sizes in order to increase the number of transistors in a microchip [2]. Now these top-down technologies are poised to be exploited for benefit in a variety of

Patterning Surfaces

Two complementary nanoscale printing techniques have emerged for the transfer of reagents, both chemical and biological to a surface: techniques based on scanning probe microscope-based lithography (SPL) and soft lithography-based techniques. Based on the development of scanning probe microscopy in the 1980s, SPL nanofabrication techniques have emerged that exploit the use of the tip of the scanning probe to manipulate matter on a surface. These methods offer the advantage of extremely high resolution as these probes are capable of manipulating individual atoms or molecules. However, this typically requires high vacuum and temperatures which approach absolute zero, and is not readily scalable for patterning of large areas [2]. While these methods represent powerful tools for high-resolution patterning over a small area, probe tips were not used in a printing method until the advent of dip-pen nanolithography (DPN) in 1999 [3].

Patterning with Dip-Pen Nanolithography

Using an “ink”-coated atomic force microscope (AFM) tip as a stylus, DPN is a direct-write method in which soft or hard materials can be deposited to a surface with sub 50 nm resolution [3]. Developed in the Mirkin laboratory at Northwestern University in 1999, DPN relies on humidity to cause the spontaneous formation of a meniscus between the AFM tip and the surface (Fig. 1). The ink is deposited along this meniscus, with the deposition rate dependent upon factors such as the nature of the ink, the contact time between the tip and the substrate, and the coverage of the ink on the AFM tip. The first example of DPN deposited an alkanethiol onto a gold substrate, with the strong interactions between the gold substrate and the thiol overcoming the weaker interaction between the tip and the thiol to drive the formation of patterns on the surface. A variety of metallic, semiconducting and insulating substrates have been used with DPN [3]. The technique has proven to be quite versatile with respect to the composition of the ink; structures have been formed from organometallics, polymers, colloidal nanoparticles, metal ions, and a range of biologicals



Nanoscale Printing, Fig. 1 A schematic showing the dip-pen nanolithography (DPN) technique. In DPN, ink molecules are transferred from the tip of an atomic force microscope stylus to a surface through a water meniscus

including DNA, proteins, peptides, and affinity templates which allowed for the subsequent immobilization of viruses [2].

A primary application of DPN techniques is the generation of biological nanoarrays for the study of cellular genetics and rapid screening of diseases. DPN-generated nanoarrays of an antibody have been used to screen for the human immunodeficiency HIV-1 virus antigen in serum samples. This study demonstrated an increase in detection limits of orders of magnitude over conventional enzyme-linked immunosorbent assays [3]. Nanoarrays generated in this way have advantages of holding 10^4 – 10^5 more features than conventional microarrays, offering the opportunity for either decreased sample volume or for the screening of a larger number of targets per experiment.

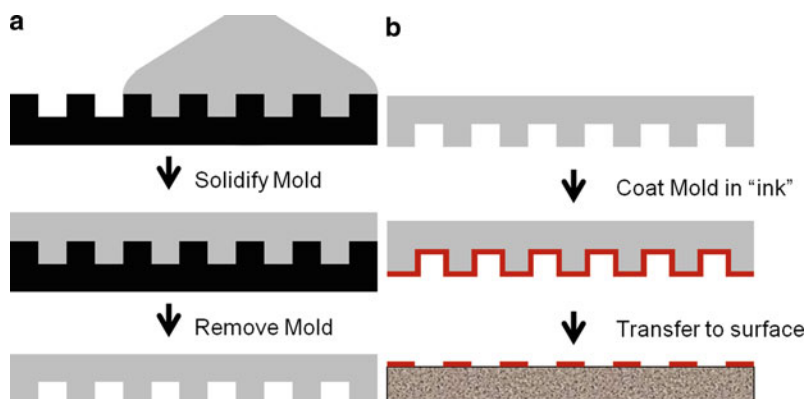
While printing with a single AFM tip over large areas would be impractically slow, massively parallel arrays of cantilevers have been fabricated to increase the throughput of the technique. Arrays with as many as 1.3 million cantilevers have been microfabricated toward this end, increasing the throughput by four orders of magnitude. A number of variations to DPN have emerged to broaden the applicability of the technique, with modification of the tips a common theme. Coating the AFM tip with a silicone polymer increases the amount of ink that can be loaded to the tip, allowing for printing of larger areas. The inclusion of microfluidic channels to deliver ink to the tips, termed nanofountain pen (NFP) probes, ensures a constant supply of ink to the surface. Features with greater than 50 nm resolution have been printed with NFP probes. The technique also serves to keep the ink molecules solvated, an aspect which can be critical for biological applications.

Heating of the AFM tip (thermal DPN) can allow for deposition of insoluble ink molecules that are solids at room temperatures.

Patterning with Soft Lithography Techniques

Pioneered in the 1990s in the Whitesides laboratory at Harvard University, soft lithography begins with the generation of a rigid template which is used to generate an elastomeric mold or stamp which replicates the three-dimensional shape of the original template. Master templates are typically prepared by photolithographic techniques, which were pioneered in the electronics industry. The molds of these master templates are typically generated from poly (dimethyl siloxane) (PDMS), although polyurethanes, polyamides, perfluoropolyethers, and natural polymers such as agarose have been used for this purpose. To fabricate the mold, the desired polymer is poured onto the master template and cured, generating a relief structure with the raised portions of the stamp defining the pattern to be transferred.

Microcontact printing (μ CP) is the process by which the mold is used as a stamp after it is coated with an ink, and subsequently brought into contact with a substrate to transfer the ink (Fig. 2). The ink transfers from the stamp by chemisorption or physisorption to the surface. While alkanethiol inks transferred to a gold substrate are, perhaps, the most common usage of μ CP, the technique has also been used to pattern proteins, DNA, cells, colloids, and polymers [4] onto a variety of surfaces. The flexibility of the elastomeric stamp allows for conformal atomic level contact to be achieved on flat and curved surfaces, and over large stamp areas. The μ CP technique has advantages as it is low-cost, is high-throughput and is easy to use which has earned it popularity as a tool among researchers. While the low cost of the elastomeric molds has rendered μ CP accessible to researchers in many fields, there are some limitations to the technique. Patterns are determined by the original master template, so that each stamp can only produce one pattern. The mechanical properties of the stamp can limit the flexibility of the design; features spaced too widely (> 20 times the feature height for typical PDMS) or with aspect ratios outside of the 0.2–2 range can cause defects from deformation of the stamp [2]. Stamping of features smaller than 150 nm is a challenge using conventional methods, though features as small as 80 nm have



Nanoscale Printing, Fig. 2 A schematic for microcontact printing (μ CP). (a) The fabrication of a mold (*top to bottom*). The mold material (*gray*) is poured onto a master template with topographical features (*black*). The liquid mold materials solidifies and is removed from the master template, producing a mold

with the inverse of the features found on the master template. (b) The μ CP process (*top to bottom*). The elastomeric mold is coated with a chemical ink (*red*). Ink is transferred to the surface by contact to the raised features of the mold

been prepared by using DPN to modify the surface of a flat stamp to passivate surrounding areas with a perfluorinated monolayer.

Nanotransfer printing (nTP) is the process of transferring a thin solid film from a nanoscale-featured stamp to a substrate. In this case, the stamp can be either rigid or flexible and the material transferred is typically a metal. In a typical process, the stamp is coated with a thin film (e.g., 20 nm of gold) and brought into contact with the desired surface. The film can transfer from the stamp by a number of methods, including binding to reactive self-assembled monolayers (SAMs) like alkanethiols, cold welding between two metal layers, or condensation between silanol or titanol groups [5]. Applications for nTP-derived structures are typically electronic in nature, including organic thin film transistors, capacitors, and electrostatic lenses. Features as small as 70 nm have been fabricated with nTP.

The recent innovation of polymer pen lithography (PPL) in 2008 combines many of the advantages of μ CP and DPN by using an elastomeric array of tips to print a pattern. With PPL, a PDMS stamp consisting of an array of tips is affixed to a piezo scanner and delivers an ink with spot sizes from 65 nm to over 10 μ m determined by varying the force and time over which the ink is delivered [6]. The elastomeric arrays of PPM pens, which can contain as many as 10 [8] pyramid-shaped tips, are made from a master prepared by traditional photolithographic techniques. The cost

of preparing new arrays of tips is greatly reduced for PPL compared to DPN, with the cost of a new PPL array estimated at less than \$1 after the fabrication of the initial master template [2]. The resulting arrays are transparent, allowing for registration of the cantilevers with previously printed features. The master template used to generate the arrays can be filled by inject printing, allowing it to be used as an array of inkwells with which the array of tips is perfectly aligned. The entire elastomeric array absorbs ink, acting as a reservoir and allowing for printing over large areas without reinking.

Molding and Embossing Methods

The molding process involves pouring a liquid precursor onto a nanopatterned template and curing or solidifying it into a solid which replicates the original features. Molding features on the nanoscale can be performed with both hard and soft molds. This is the basis for techniques such as step-and-flash imprint lithography (SFIL), replica molding (RM), and pattern replication in non-wetting templates (PRINT). Embossing or imprinting methods use a mold to transfer features to an initially flat polymer film. Embossing methods include nanoimprint lithography (NIL), solvent-assisted micromolding (SAMIM), and PRINT.

Hard molds have been fabricated out of quartz, silicon, and metals, though quartz and silicon are the most commonly used. Features as small as 10 nm have been transferred into silicon, while quartz molds have a minimum size of 20 nm [4]. Though soft molds are

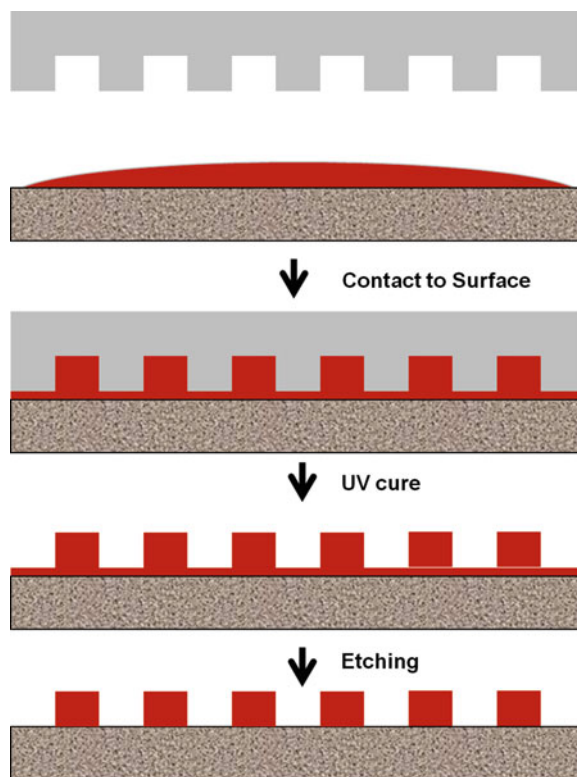
less costly and easier to fabricate, rigid molds offer several advantages over softer molds. Rigid molds will show little deformation over large areas with the pressures required for embossing. Hard molds are stable at the temperatures required for cross-linking polymer precursors and are chemically inert to them. Release of the mold from the patterned substrate can be difficult, depending on the surface properties of the mold and the molded polymer, and release coatings of fluoroalkyl silanes are typically linked to the mold surface to facilitate release. Quartz substrates are transparent to UV light and visible light, while silicon is not, allowing for photochemical cross-linking of prepolymer with quartz molds.

Step-and-Flash Imprint Lithography

Step-and-flash imprint lithography replicates the topography of a quartz mold using a photocurable prepolymer solution as the molded material at room temperature using low applied pressure. The prepolymer solution consists of a monomer and a photoinitiator with hardening of the material proceeding by free-radical polymerization upon exposure to UV light. An important consideration for complete filling of the mold is the viscosity of the prepolymer, or monomer, solution; low molecular weight monomers are used due to their low viscosity ($<5\text{cPs}$) which allows for complete filling of the mold features. The quartz mold is pressed down onto a substrate which is coated with a liquid prepolymer and cured with UV light. While the mold features fill completely, they remain connected by a thin residual layer of cured monomer. This layer can be removed by an etching process (such as reactive ion etching), yielding discrete features on the substrate (Fig. 3).

The quartz molds used in SFIL are typically coated with a release layer to facilitate release of the mold from the cured polymer and to increase the lifetime of the mold. Failure of the mold to release cleanly can result in broken mold features, which adds to the cost of this technique. The first reported coatings were fluorinated silanes, which provided a mold lifetime of less than 100 uses. Newer coatings have extended this lifetime to the patterning of over 1,500 substrates [4].

In addition to the generation of patterned surfaces, SFIL can be used to generate polymeric particles. In this case, after the etching step removes the connecting layer from between the particles, a subsequent step provides release of the cured polymer from the



Nanoscale Printing, Fig. 3 A schematic for step-and-flash imprint lithography (SFIL). From *top to bottom*: A UV transparent quartz mold (*gray*) is brought into contact with a thin coating of a prepolymer solution (*red*) on a surface. The prepolymer is exposed to UV light through the quartz mold, leaving a patterned polymer surface after removal of the mold. An etching step removes the thin layer of polymer which connected the raised features

substrate. Particles can be removed by physical agitation or by dissolution of a sacrificial surface coating (i.e., poly vinyl alcohol) on the initial substrate. While the etching process, which involves exposure to an oxygen plasma, could be considered quite harsh, the biological activity of a protein encapsulated in SFIL particles was maintained after 20 s of exposure to plasma [7]. SFIL particles have potential for use in drug delivery, as particles containing an enzymatically degradable peptide sequence and carrying a DNA plasmid cargo were fabricated and demonstrated successful cargo release in an *in vitro* experiment [8].

Nanoimprint Lithography

Nanoimprint lithography (NIL) involves the transfer of a pattern from a rigid stamp to a polymer film via application of pressure at temperatures above the

glass transition temperature of the polymer. The typical mold material is silicon with features as small as 5 nm patterned. While NIL is typically used to pattern polymers, biomolecules have also been successfully patterned. High pressures and temperatures are required with NIL, putting stress on the molds and limiting their lifetime to ~ 50 imprints. The high viscosity of many of the polymers selected for patterning is another issue with this technique. Low molecular weight polymers may be desirable due to their lower viscosities and relative ease of processing [9].

Imprint Lithography

Soft, or imprint, lithography generates elastomeric molds from rigid substrates and uses the molds to replicate the original topographical features. Imprint lithography methods were first seen as a low-cost, high-resolution alternative to photolithographic techniques for the replication of patterns in the micro- and nanoscales. The typical mold materials have included cross-linked poly (dimethyl siloxane) (PDMS or silicone) or perfluorinated polyethers which have low toxicity, are extremely flexible, and have low surface energies, all of which make them excellent materials for molding applications.

Replica Molding

The replica molding (RM) process consists of transferring the topographical pattern of a master template to an elastomeric mold (typically made from poly (dimethyl siloxane) (PDMS)) which is in turn used to transfer the pattern to a liquid which is cured, or solidified, in contact with the mold. Molded materials can be photo or thermally curable, though there are limitations. Many low molecular weight monomers display some solubility in PDMS, causing the mold to swell and limiting the molding process. Molded materials include polymers, gels, precursors to ceramics and carbons, luminescent phosphors, salts, and colloids [5]. The use of the soft PDMS mold allows for molding of non-flat surfaces, an advantage over hard template molding techniques. While molds may only remain viable for approximately 20 replications, the mold material is inexpensive and many molds can be prepared from a single master template. Replication of the high-cost master templates in a low-cost material makes this technique accessible to a broad range of researchers.

Replica molding in PDMS has transferred features as small as 3 nm wide structures with 0.5 nm vertical deflections. Based on the physics of van der Waals interactions, the theoretical ultimate size limit for molding with PDMS is less than 0.5 nm. Another advantage of PDMS is the low surface energy of the material, which promotes facile release from the master template as well as the patterned surface. The PDMS mold surface can be modified to tailor the hydrophilicity; treatment with oxygen plasma and with fluorosilane have been explored to increase or decrease hydrophilicity, respectively. Changing the surface properties of the mold has implications in wetting and filling of the mold as well as mold release [8].

One limitation of RM techniques is difficulty in eliminating the thin layer of material which connects the patterned features, called a flash layer. While this layer can be etched away generating isolated structures or free particles, as in the case of SFIL, exposure of biological or otherwise sensitive materials to this harsh condition is thought to be undesirable due to the risk of degradation. However, there are many examples of particle replication using PDMS molds. Particles were generated from thermoplastics, thermosets, and hyperbranched polymers, though these were all on the micron size range and the generation of nanoparticles remains a challenge.

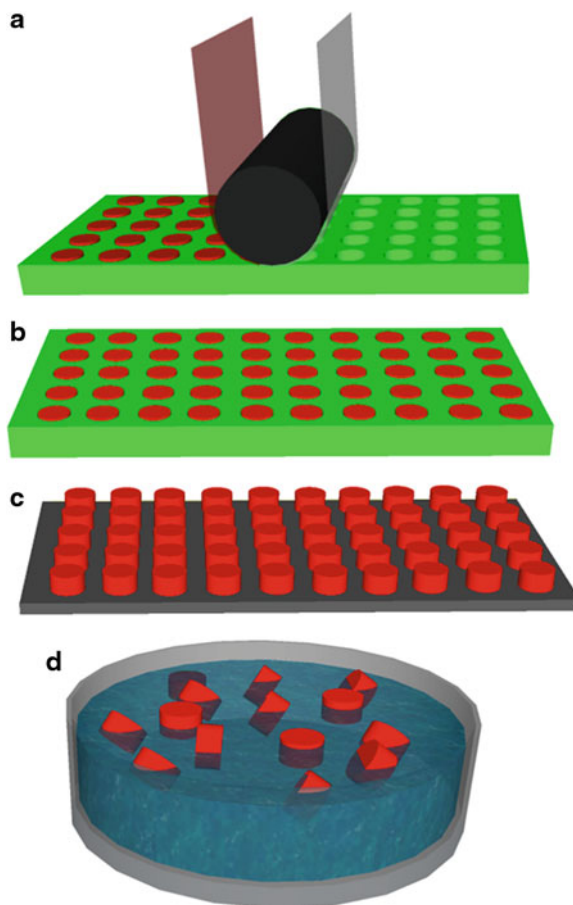
PRINT

The particle (or pattern) replication in non-wetting templates (PRINT[®], Liquidia Technologies) technique, developed in the DeSimone laboratory at the University of North Carolina at Chapel Hill and commercialized by Liquidia Technologies, has extensively focused on the use of a perfluoro polyether (PFPE) elastomer as the mold material to replicate the features of a rigid master template. Though similar to other soft lithography techniques, the use of PFPE offers several advantages over PDMS. First, the PFPE is not as responsive to solvents as PDMS, which tends to swell when exposed to organics. This allows for the patterning of a wide variety of substrates with varied properties. Second, PFPE has a lower surface energy than PDMS which is advantageous for particle making applications. Due to the non-wetting nature of the PFPE mold, the depressions in the mold can be filled without wetting the landmass in between, yielding

distinct particles without the troublesome flash layer. Thirdly, the properties of the PFPE allow for facile release of the particles from the mold without the modification of the mold surface as is required with PDMS molds. Lastly, PFPE forms a less flexible mold than PDMS, which is advantageous for high fidelity molding since the PFPE molds are less prone to unwanted deformation than their PDMS counterparts [10].

A drawback of the PRINT technique is the cost associated with PFPE when compared to the relatively inexpensive PDMS. The amount of PFPE can be minimized with the use of an inexpensive backing material coated with a very thin film of the PFPE molding material. Molds produced in this way have facilitated scale-up and commercialization of the PRINT technique, with a rate of 1,000 ft of mold per day [7]. Scalability is crucial to the application of PRINT, and other particle fabrication techniques, in nanomedicine due to the high material demands of clinical trials.

PFPE has been used to replicate the topography of a variety of nanoscale objects, including carbon nanotubes with diameters of 2 nm and soft objects such as block copolymer micelles and viruses [8], which were then replicated on a substrate by an embossing process. To generate particles, the PRINT mold is filled with a liquid preparticle solution, with the excess solution removed with a high surface-energy polymer sheet via a wicking process. The preparticle solution is solidified and particles removed from the mold by physical agitation or by transfer to a surface through use of an adhesive which is subsequently dissolved, yielding free particles (Fig. 4). Particles have been generated from a variety of materials, including cross-linked PEG hydrogels, linear biocompatible polyesters, proteins, therapeutics, and block copolymer micelles. A variety of cargos have been encapsulated in PRINT particles, including proteins, anticancer drugs, nucleic acids, as well as a variety of imaging agents such as iron oxides, gadolinium chelates, and ^{64}Cu chelates [10]. PRINT has evolved to a roll-to-roll process and includes the use of a very important film splitting technique that allows the flash-free filling of the cavities in the mold to generate the individual particles. Such a roll-to-roll process is amenable to a wide range of materials to be processed using PRINT including the use of an increasing variety



Nanoscale Printing, Fig. 4 A schematic for particle replication in non-wetting templates (PRINT). (a) A preparticle solution (*red*) is distributed evenly into an elastomeric mold made from perfluoro polyether (*green*) by a roller (*black*) covered by a high-surface-energy sheet (*gray*) which removes excess solution. (b) The particle solution is solidified in the mold. (c) Particles are removed from the mold with a harvesting film (*gray*). (d) The film is dissolved, freeing the particles

of mold materials that are being employed which are tailored to the chemistries being used to fill the mold [7].

Particles generated by nanoscale printing methods like PRINT and SFIL are of growing interest to researchers in the areas of drug delivery and medical imaging. These processes offer precise control over the size, shape, and composition of the particle, with broad capabilities with respect to the loading of imaging agents and/or therapeutics. PRINT particles of various shapes and sizes were used to probe cellular uptake

kinetics and mechanisms with respect to particle size and shape.

Statement of disclosure: Joseph DeSimone is a founder, member of the board of directors, and maintains a financial interest in Liquidia Technologies. Liquidia was founded in 2004 to commercialize PRINT technology and other discoveries of Professor Joseph DeSimone and colleagues at the University of North Carolina at Chapel Hill.

Cross-References

- ▶ [Atomic Force Microscopy](#)
- ▶ [Dip-Pen Nanolithography](#)
- ▶ [Microcontact Printing](#)
- ▶ [Nanoimprint Lithography](#)
- ▶ [Nanoimprinting](#)

References

1. Claridge, S.A., Castleman, A.W., Khanna, S.N., Murray, C.B., Sen, A., Weiss, P.S.: Cluster-assembled materials. *ACS Nano* **3**, 244 (2009)
2. Braunschweig, A.B., Huo, F.W., Mirkin, C.A.: Molecular printing. *Nat. Chem.* **1**, 353 (2009)
3. Salaita, K., Wang, Y.H., Mirkin, C.A.: Applications of dip-pen nanolithography. *Nat. Nanotechnol.* **2**, 145 (2007)
4. Gates, B.D., Xu, Q., Stewart, M., Ryan, D., Willson, C.G., Whitesides, G.M.: New approaches to nanofabrication: molding, printing, and other techniques. *Chem. Rev.* **105**, 1171 (2005)
5. Gates, B.D.: Nanofabrication with molds and stamps. *Materials Today* **8**, 44 (2005)
6. Zheng, Z., Daniel, W.L., Giam, L.R., Huo, F., Senesi, A.J., Zheng, G., Mirkin, C.A.: Multiplexed protein arrays enabled by polymer pen lithography: addressing the inking challenge. *Angewandte Chemie* **121**, 7762 (2009)
7. Merkel, T.J., Herlihy, K.P., Nunes, J., Orgel, R.M., Rolland, J.P., DeSimone, J.M.: Scalable, shape-specific, top-down fabrication methods for the synthesis of engineered colloidal particles. *Langmuir* **26**, 13086 (2010)
8. Canelas, D.A., Herlihy, K.P., DeSimone, J.M.: Top-down particle fabrication: control of size and shape for diagnostic imaging and drug delivery. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **1**, 391 (2009)
9. Zankovych, S., Hoffmann, T., Seekamp, J., Bruch, J.U., Torres, C.M.S.: Nanoimprint lithography: challenges and prospects. *Nanotechnology* **12**, 91 (2001)
10. Jeong, W., Napier, M.E., DeSimone, J.M.: Challenging nature's monopoly on the creation of well-defined nanoparticles. *Nanomedicine-Uk* **5**, 633 (2010)

Nanoscale Properties of Solid–Liquid Interfaces

Lucio Colombi Ciacchi and Susan Köppen
Hybrid Materials Interfaces Group, Faculty of
Production Engineering and Bremen
Center for Computational Materials Science,
University of Bremen, Bremen, Germany

Synonyms

[Solid–liquid interfaces](#)

Definition

A solid–liquid interface is the boundary region separating a solid phase from a liquid phase. While its extension is strongly dependent on the particular phases, for poorly soluble solids and small-molecule liquids such as water, it spreads over a thickness of the order of a few nm. The properties of solid–liquid interfaces can be described at a macroscopic, continuum level, but are ultimately governed by atomistic details. These are the subject of the present entry, which is written with a particular emphasis on results stemming from molecular dynamics simulations and focuses especially on interfaces between inorganic solids and water.

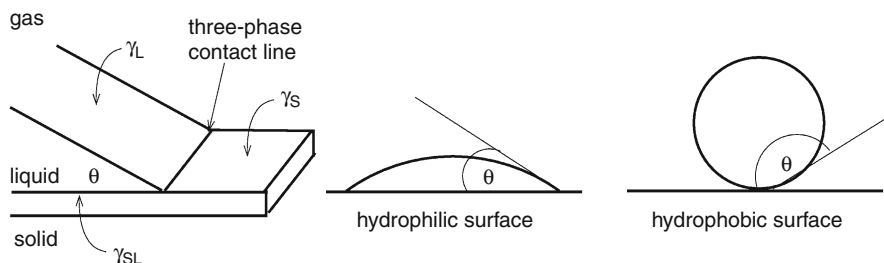
Macroscopic Description of Solid–Liquid Interfaces

The macroscopic interaction between a liquid phase (l) in contact with a solid surface (s) is described by the *interface energy* γ_{sl} , defined as the (positive) change in free energy associated with a unitary change of the interface area. The spreading of a liquid droplet over a solid surface can be quantified by the equilibrium *contact angle* θ at the droplet's edge, which is related to γ_{sl} and the surface energies of the two separated phases, γ_s and γ_l , via the Young equation

$$\gamma_{sl} + \gamma_l \cos \theta = \gamma_s \quad (1)$$

Nanoscale Properties of Solid–Liquid Interfaces,

Fig. 1 Scheme of the three-phase contact of a liquid droplet with a contact angle θ on a solid surface and typical water droplet shapes on a hydrophilic and a hydrophobic surface

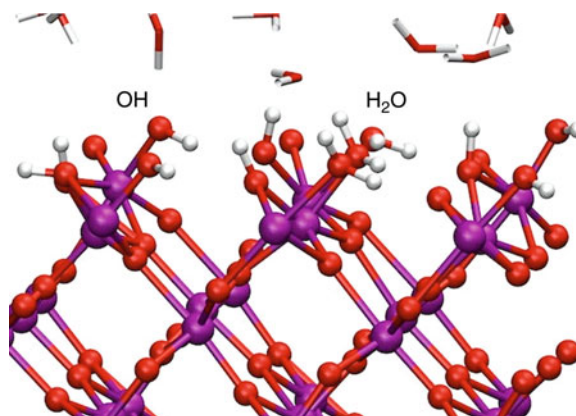


Small values of the contact angle indicate easy spreading of the liquid phase and thus wetting of the solid surface, while large values indicate only partial spreading, up to the limit of complete dewetting for $\theta = 180^\circ$. In the case of water, the solid is considered *hydrophilic* or *hydrophobic* depending on whether θ is smaller or larger than 90° (Fig. 1).

Chemical Details of Solid–Water Interfaces

The interface energy and the contact angle defined above capture effects at the macroscopic level that are ultimately driven by the microscopic interactions of liquid molecules with the superficial atomic layers of the underlying solid substrate.

Water molecules interact with hydrophobic surfaces via weak *van der Waals forces*. These include a repulsive contribution at very small distances arising from the non-allowed mixing of molecular orbitals with electronic surface states, and attractive contributions due to *London dispersion* and *induced polarization* effects. Polarization is especially important in the case of polar liquids with a high dielectric constant, such as water, close to non-reactive metallic surfaces, such as gold or defect-free graphene sheets. It is important to note that all other metal surfaces are reactive in a water environment and spontaneously form ultrathin *native oxide* layers (with a thickness of the order of 1 nm), which confer them a hydrophilic character. Border-line cases are noble metals such as platinum, to which water is able to chemisorb through relatively weak chemical bonds, but not to trigger the formation of superficial oxides under normal conditions. Non-metallic hydrophobic surfaces of technological relevance include for instance hydrogen-terminated silicon or carbon (typically produced via etching of the natively oxidized surfaces with fluoridic acid), hydrocarbon or fluorocarbon polymers such as polyethylene and Teflon, and, to a lesser extent, silicones.



Nanoscale Properties of Solid–Liquid Interfaces,
Fig. 2 Chemisorbed water layer on a rutile $\text{TiO}_2(100)$ surface, as obtained in first-principles molecular dynamics simulations. Note the mixed dissociative adsorption mode with both terminal OH and H_2O groups

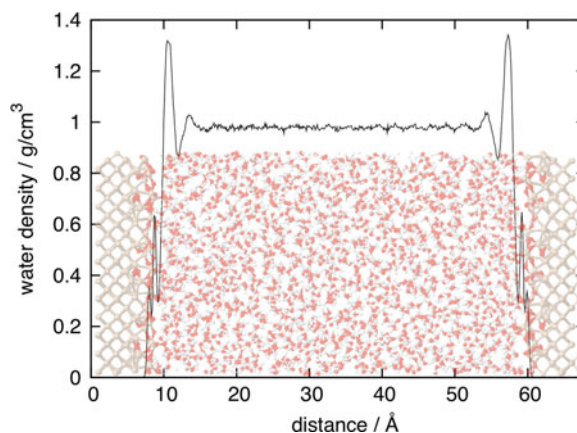
The surfaces of ceramic oxides, natively oxidized metals, and polar polymers such as polyethylene oxide are hydrophilic. Water molecules bind to them strongly via hydrogen bonds, with typical bond distances in the range 1.5–2.0 Å, which are smaller than typical van der Waals distances by about 1.5 Å. Covalent bonds are readily formed upon reaction between water molecules and anhydrous inorganic solid surfaces. Depending on the bulk composition and the surface crystallographic direction, the reaction can lead either to molecularly chemisorbed water, or to dissociative chemisorption with formation of terminal OH groups (Fig. 2). Mixed molecular and dissociative adsorption is common and the resulting hydroxylation equilibrium on the surface can be determined with the help of the so-called *Multi Site Complexation Model* [1]. In analogy with the standard nomenclature of organic alcohols, terminal hydroxyl groups can be classified as: *geminal* (two groups bound to the same site), *vicinal* (two groups bound to nearest-neighbor sites), or *isolated* (otherwise).

After chemisorption, the terminal groups may either donate or accept protons to or from the water solution, depending on their equilibrium dissociation constants and the solution pH. This results in the formation of charged surface sites, such as O^- , OH^- , H_2O^+ , and an average surface charge density σ . At the macroscopic level, charged surfaces in contact with electrolyte solutions can be described by well-known theories such as the *Gouy-Chapman electrical double layer model* [2]. The protonation/deprotonation equilibrium at the interface can be shifted by varying the pH of the solution. The pH value at which the surface is neutral is defined as the *point of zero charge* (pzc), an intrinsic property of the solid surface. This is conceptually different, but technically often difficult to distinguish, from the *isoelectric point* (IEP), defined as the pH for which the electrokinetic ζ -potential is zero. The latter is defined as the electrostatic potential at the edge of the layer of fluid molecules that remain immobile (either chemisorbed or otherwise strongly adsorbed) over the surface when subjected to a hydrodynamic flow. The ζ -potential at a given pH, and thus the IEP upon continuous pH variation, can be determined, e.g., from electrophoretic measurements of solid particle suspensions.

Molecular Ordering in Surface Proximity

Liquid molecules close to a solid–liquid interface experience a substantial reordering with respect to their arrangement in bulk, induced by the interaction with the surface atoms. This manifests itself with a typical increase of the liquid density in proximity of the surface, followed by density oscillations with amplitude of the order of the molecular diameter, which may reach up to several nm within the liquid phase. The density profile of water at a natively oxidized Si(100) surface, as obtained in molecular dynamics (MD) simulations [3], is reported in Fig. 3, clearly showing the oscillatory behavior described above. Noteworthy is the presence of water molecules deep inside the crevices of the atomic-scale roughness intrinsic of the native oxide layer, as indicated by the small features preceding the first main density peak. These features are characteristic of amorphous oxide phases such as glasses or superficial native oxides.

In contrast, hydrophobic surfaces are separated from the liquid phase by a vacuum gap with thickness



Nanoscale Properties of Solid–Liquid Interfaces, Fig. 3 Density profile of water in contact with a natively oxidized Si(100) surface, schematically shown in the background, as obtained in molecular dynamics simulation [3]

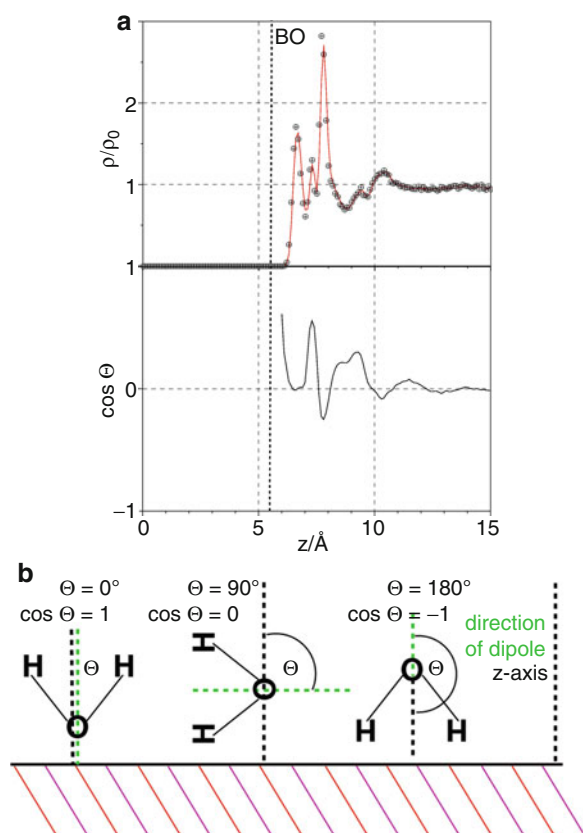
of the order of a van der Waals distance, known as *depletion layer*. Knowing the density profiles $\rho_s(z)$ and $\rho_l(z)$ and the corresponding bulk values ρ_s^B and ρ_l^B of the solid and liquid phases, the thickness δ of the depletion layer can be calculated as

$$\delta = \int_{-\infty}^{+\infty} (1 - \rho_s(z)/\rho_s^B) - (\rho_l(z)/\rho_l^B) dz \quad (2)$$

Note that this definition is general for any interface between two phases, in particular also for hydrophilic, rough, or porous surfaces in contact with water, where negative values of δ can be expected. For ideally flat surfaces (where δ cannot be negative), MD simulations have shown that the degree of hydrophobicity, as measured by the contact angle θ , is related to the thickness of the depletion layer according to [4]

$$\delta^{-2} \propto (1 + \cos \theta) \quad (3)$$

In the water layers closer to the interface, water molecules assume, on average, a preferential orientation that strongly depends on the surface charge, the available terminal groups, the presence of adsorbed counterions, and the surface polarizability. This effect can be quantified by an analysis of the average angle between the water molecular axis (i.e., the molecular dipole) and the surface plane, as obtained in MD simulations. An example is shown in Fig. 4 for the case of a TiO_2 surface in contact with a saline water solution.



Nanoscale Properties of Solid–Liquid Interfaces,

Fig. 4 Profiles of the density and molecular dipole orientation of liquid water at a rutile $\text{TiO}_2(100)$ surface in contact with a 0.15 M NaCl solution at pH 9. Note the non-trivial structuring induced both by negatively charged surface sites and by adsorbed positive counterions. The angle Θ is defined as in the scheme here above [5]. The label BO indicates the position of the outermost O atoms (bridging oxygen) of the surface

Consistently with the variation of preferential water orientation, also the profile of the electrostatic potential in the liquid phase shows a characteristic oscillatory behavior [5]. Moreover, the higher molecular density, preferential orientation, and reduced mobility (see section [Self-Diffusion at Solid–Liquid Interfaces](#) below) of the water layers in surface proximity are at the origin of a considerably reduced *relative dielectric constant* with respect to bulk liquid water. While the latter is about 80, values around 10 describe correctly the dielectric properties of the water (and ion) layers in the interface region [6].

Slip Length at Solid–Liquid Interfaces

The presence of a depletion layer between hydrophobic surfaces and liquid water is connected with the existence of a positive *slip length* characterizing the flow of liquid at the interface. The slip length b is defined as the ratio between the relative velocity of liquid molecules at the interface with respect to the surface and the *shear rate* at the surface, or

$$v(0) - v_{surf} = b\dot{\gamma}(0) \quad (4)$$

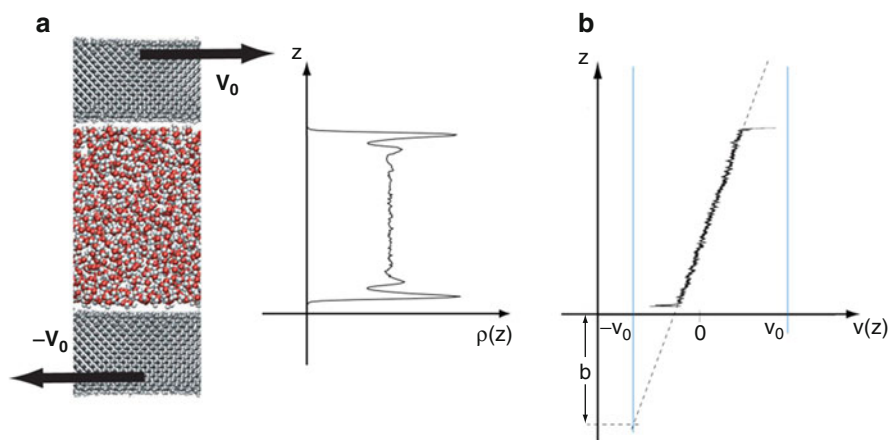
The slip length can be computed theoretically in MD simulations under shear boundary conditions, in which a bulk liquid is confined between two surfaces moving at equal and opposite velocities v_0 (Fig. 5). At the solid–liquid interface, the molecules slip over the depletion layer and move with a velocity different from v_0 . The slip length is obtained by extrapolation of the linear velocity profile (see Fig. 5b).

By varying the strength of the molecule–surface interactions (i.e., the degree of hydrophobicity), Netz and collaborators have demonstrated that the slip length is related to the contact angle θ and the depletion layer thickness δ by the following relationships:

$$b \propto (1 + \cos \theta)^{-2}; \quad b \propto \delta^4 \quad (5)$$

These, combined, lead to the relationship between δ and the contact angle described in the previous section.

In the case of hydrophilic surfaces, the water–surface hydrogen bonds effectively anchor the interfacial molecules to the surfaces, and no-slip conditions are attained. Simulations under shear boundary conditions have revealed that in this case, the shear rate at the interface is smaller close to the surface than in the bulk, in a region defined roughly by the first main density peak of the layered water structure. This smaller shear rate corresponds to an increase of water viscosity, the extent of which is about half an order of magnitude and is strongly dependent on the degree of surface hydrophilicity. The increased velocity gradient in the bulk region leads, by extrapolation, to negative slip lengths for hydrophilic surfaces.



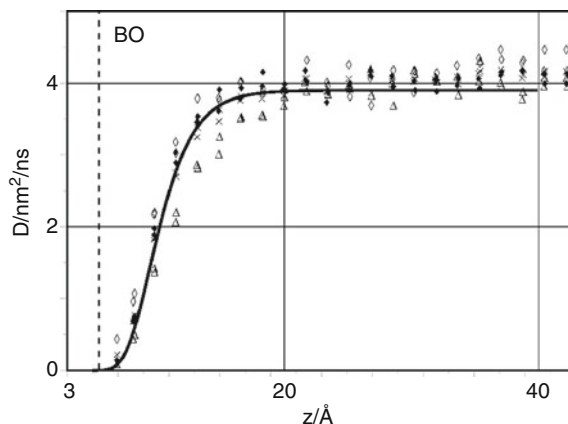
Nanoscale Properties of Solid–Liquid Interfaces, Fig. 5 Water confined between two H-terminated C(100) surfaces under shear boundary conditions, the surfaces moving at equal and opposite velocities v_0 . Density profile of the water

molecules (cf. Fig. 3) between the surfaces (a). Profile of the average molecular velocity between the surfaces (b). Note the non-zero relative velocity at the solid–liquid interface, defining a slip length b (Reproduced with permission from Ref. [4])

Self-Diffusion at Solid–Liquid Interfaces

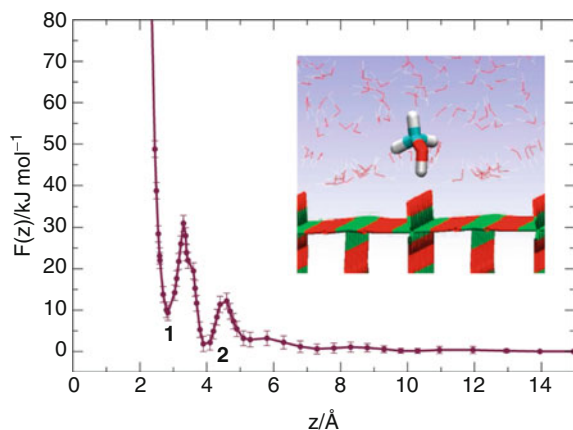
The layering of the molecular structure of liquids in proximity of a solid surface and the corresponding increase of viscosity are at the origin of a reduced mobility of the molecules close to the interface. In particular, in the case of anhydrous oxide surfaces in contact with water, the first water layer is chemisorbed to reactive surface sites and thus almost completely immobile, except for the intrinsic equilibrium exchange between adsorbed and desorbed molecules. Correspondingly, the self-diffusion coefficient of water is vanishingly small at hydrophilic surfaces and increases at increasing distances from the surface. (Note that in this discussion the chemisorbed water is considered to be part of the liquid phase, but it could be considered equally well as belonging to the solid.) Quantitative investigations of the self-diffusion coefficient in MD simulations are difficult due to a number of reasons (see e.g., Ref. [7] and references therein). In particular, the long residence time of water molecules close to the interface makes difficult the convergence of the velocity autocorrelation function from which the diffusion coefficient is typically calculated.

Nevertheless, the available literature results show that the diffusion coefficient of water at polar surfaces qualitatively follows a sigmoidal function, indicative of strong water trapping in the near-surface region.



Nanoscale Properties of Solid–Liquid Interfaces, Fig. 6 Variation of the self-diffusion coefficient of water near a hydrophilic surface, calculated in an MD simulation for a 0.15 M NaCl solution in contact with a $\text{TiO}_2(100)$ surface terminated according to the hydroxylation equilibrium at pH 7.4. The different symbols correspond to different directions used to compute the velocity autocorrelation function, indicating a roughly isotropic diffusive behavior [5]. The label BO is defined as in Fig. 4

Simulations of a 0.15 M saline water solution in contact with the $\text{TiO}_2(100)$ surface (Fig. 6) reveal that the bulk diffusive behavior of the water solvent is reached for distances of about 1.5 nm from the solid–liquid interface.



Nanoscale Properties of Solid–Liquid Interfaces, Fig. 7 Free energy profile of a methanol molecule interacting with a rutile $\text{TiO}_2(110)$ –water interface. The oscillatory behavior reflects the water layering in surface proximity (cf. Fig. 3). Note the lack of a net driving force for adsorption but the presence of relatively deep local minima that can kinetically trap the molecule close to the surface (Adapted and reproduced with permission from [9])

Molecular Adsorption at Solid–Liquid Interfaces

The adsorption of molecules, in particular organic residues and peptides, at solid–liquid interfaces has gained large attention in the past few years, due to its importance in a variety of applications such as biomedical implants, biosensors, pharmaceutical packaging, and many others. Atomistic details of the adsorption can be successfully investigated in MD simulations, provided that the force fields employed in the simulations realistically capture the chemical details of the interface. It is now well established that the water layering at the interface has an important influence both on the structural and energetic details of the adsorption. This holds for small molecules as well as for larger peptides such as collagen fragments [8]. In the latter case, for instance, a helical molecular configuration has been observed to be stabilized at hydrophobic surfaces in part by a striking matching between the water density oscillations in surface proximity and the ordered hydrophilicity/hydrophobicity pattern of the adsorbed peptide.

Correspondingly, the free energy profile relative to the approach or retraction of a molecule from solid–liquid interfaces often presents a complex structure as a direct consequence of the oscillatory behavior of the

solvent layers. This can be observed in MD simulations that sample the molecular *potential of mean force* near the interface by means of constrained methods such as *umbrella sampling* (Fig. 7).

It is interesting to note that an influence of the interfacial water structuring on free energy adsorption profiles has been observed both for hydrophilic surfaces such as TiO_2 [9] and for rather hydrophobic surfaces such as gold [10]. As mentioned above, in the latter case, an important contribution of the interaction across the solid–liquid interface is the electronic charge redistribution in the metal surface when polar molecules approach it (induced polarization).

Cross-References

- ▶ [Applications of Nanofluidics](#)
- ▶ [Surface Energy and Chemical Potential at Nanoscale](#)
- ▶ [Wetting Transitions](#)

References

1. van Hiemstra, T., Riemsdijk, W.H.: A surface structural approach to ion adsorption: the charge distribution (CD) model. *J. Colloid Interface Sci.* **179**, 488–508 (1996)
2. Israelachvili, J.: *Intermolecular and surface forces*. Academic, London (1991)
3. Cole, D.J., Csányi, G., Payne, M.C., Spearing, S.M., Colombi Ciacchi, L.: Development of a classical force field for the oxidised Si surface: application to hydrophilic wafer bonding. *J. Chem. Phys.* **127**, 204704 (2007)
4. Huang, D.M., Sendner, C., Horinek, D., Netz, R.R., Bocquet, L.: Water slippage versus contact angle: a quasiuniversal relationship. *Phys. Rev. Lett.* **101**, 226101 (2008)
5. Köppen, S., Langel, W.: Simulation of the interface of (100) rutile with aqueous ionic solution. *Surf. Sci.* **600**, 2040–2050 (2006)
6. Butt, H.-J., Graf, K., Kappl, M.: *Physics and chemistry of interfaces*. Wiley-VCH, Weinheim (2006)
7. Sendner, C., Horinek, D., Bocquet, L., Netz, R.R.: Interfacial water at hydrophobic and hydrophilic surfaces: slip, viscosity, and diffusion. *Langmuir* **25**, 10768–10781 (2009)
8. Cole, D.J., Payne, M.C., Colombi Ciacchi, L.: Water structuring and collagen adsorption at hydrophilic and hydrophobic silicon surfaces. *Phys. Chem. Chem. Phys.* **11**, 11395 (2009)
9. Monti, S., Walsh, T.R.: Free energy calculations of the adsorption of amino acid analogues at the aqueous titania interface. *J. Phys. Chem. C* **114**, 22197–22206 (2010)
10. Hoefling, M., Iori, F., Corni, S., Gottschalk, K.-E.: Interaction of amino acids with the Au(111) surface: adsorption free energies from molecular dynamics simulations. *Langmuir* **26**, 8347–8351 (2010)

Nano-scale Structuring

► [Optical and Mechanical Nanostructured Coatings for Future Large-Scale Manufacturing](#)

Nanoscale Thermal Analysis

► [Nanocalorimetry](#)

Nanoscale Water Phase Diagram

Michael Nosonovsky

Department of Mechanical Engineering, University of Wisconsin-Milwaukee, Milwaukee, WI, USA

Definition

The phase diagram of water shows conditions (temperature and pressure) at which thermodynamically distinct phases can occur at equilibrium at the nanoscale.

Theoretical Background

Physical properties of many materials at the nanoscale are different from their properties at the macroscale due to the so-called scale effect. The causes of the scale effect include geometrical reasons (such as high surface-to-volume ratios at the nanoscale) and physical reasons (different physical mechanisms acting at the nanoscale compared to those at the macroscale).

Capillary Bridges and the Kelvin Equation

Besides high surface-to-volume ratios, an important feature of a small mechanical system is that capillary forces are often present. When contact of two solid bodies occurs in air, water tends to condense near the contact spots. This is because a certain amount of vapor is always present in air. Even under low relative humidity (RH), it is not possible to completely eliminate condensation in the form of water capillary bridges forming menisci near the contact spots, such as the tips of the asperities of rough solid surfaces. The

menisci are usually concave and, therefore, the pressure inside them is reduced compared with the pressure outside in accordance with the Laplace theory. This leads to the attractive capillary force between the contacting bodies that is proportional to the foundation area of the meniscus. For small systems, this capillary force may become very significant, dominating over other forces such as the van der Waals adhesion or electrostatic forces. Therefore, understanding the properties of water in capillary bridges (menisci) is very important.

A flat water–air interface is in thermodynamic equilibrium with a certain amount of vapor in air, the partial pressure of which is referred to as “the saturated vapor pressure,” P_{sat} , so that evaporation and condensation between the flat interface and vapor occurs at the same rate. If the partial pressure of vapor in air, P_V , is lower than P_{sat} , evaporation prevails over condensation, whereas in the opposite case ($P_V > P_{sat}$) condensation prevails over the evaporation. The ratio of the two values is the relative humidity, $RH = P_V/P_{sat}$. For a concave interface, the equilibrium occurs at $P_V/P_{sat} < 1$, whereas for a convex interface is occurs at $P_V/P_{sat} > 1$. The exact value of P_V/P_{sat} for a meniscus of a given curvature is given by the Kelvin equation. The pressure drop of water with density ρ risen for the height h in a capillary tube is $\Delta P = \rho gh$. Using the Laplace equation and the hydrostatic formula for vapor pressure $P = P_0 \exp(-gh/RT)$ – where R is the gas constant, P_0 is the pressure at the surface ($h = 0$), and T is the temperature – yields the Kelvin equation:

$$\gamma_{LV} \left(\frac{1}{R_1} + \frac{1}{R_2} \right) = \rho RT \ln \frac{P_V}{P_{sat}} \quad (1)$$

where R_1 and R_2 are the principle radii of curvature of the meniscus ($R_k = (1/R_1 + 1/R_2)^{-1}$ is referred to as the Kelvin radius). [Equation 1](#) relates the interface curvature at the thermodynamic equilibrium with the ratio of actual and saturated vapor pressure, P_V/P_{sat} (relative humidity). According to the Kelvin theory, a concave meniscus with a negative curvature given by [\(3\)](#) may form at any relative humidity [1].

Water at the Nanoscale

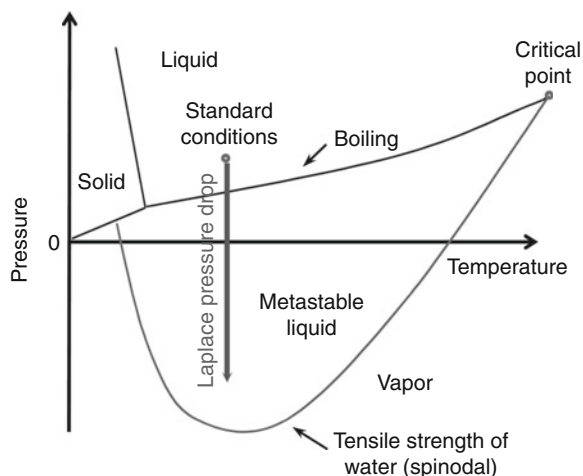
Normally, the phase state of water is uniquely characterized by its pressure and temperature. The phase diagram of water shows whether water is in its solid, liquid, or vapor state at a given temperature and

pressure. However, at the nanoscale, the situation may be different. Both molecular dynamics (MD) simulations and experimental studies with the atomic force microscope (AFM) and the surface force apparatus (SFA) show that the state of nanoscale volumes of water at a given pressure and temperature is not always the same as prescribed by the macroscale water phase diagram. In particular, confined small water volumes demonstrate quite unusual properties, such as the melting point depression. Several effects that can be detected by AFM are related to this unusual phase behavior of the nanoscale capillary bridges.

First is the metastability of small volumes of water. When pressure is below the liquid–vapor transition line of the water phase diagram, vapor is the most stable state. However, the liquid–vapor transition involves energy barriers, and thus a metastable liquid state is possible [1]. At the macroscale, random fluctuations are normally large enough to overcome the barriers. The metastable states are very fragile, and thus they are not normally observed, except for special circumstances (e.g., superheated water). However, at the nanoscale, the barriers are large compared with the scale of the system, and metastable states can exist for long intervals of time [1].

Second, it was argued that ice could form at room temperature inside the capillary bridges. The water in such a bridge can be neither liquid, nor solid, but form a liquid–ice condensate. This explains the slow rearrangement of the bridges between tungsten AFM tips and graphite surfaces reported in recent experiments [2]. The presence of ice-like structured water adsorbed at the silicon oxide surface was suggested to be responsible for the large *RH* dependence of the adhesion force in the single-asperity contact between silicon oxide surfaces in AFM experiments.

Third, pressure inside the capillary bridge is lower than the pressure outside. At the reduced pressure, the water boiling point is lower than 100°C. There is experimental evidence that boiling in the capillary bridges indeed happens in concave water menisci. Investigating nanoscale water phase transitions using AFM data can lead to modifications of the conventional water phase diagram by introducing a scale dependence. Such a diagram should depend on the characteristic size of the system, in addition to the pressure and temperature dependencies [3].



Nanoscale Water Phase Diagram, Fig. 1 Schematic of a phase diagram of water showing solid, liquid, vapor, and metastable liquid phases. While the energy barriers separating the metastable states are small, at the nanoscale these barriers are very significant

Stretched Water

When a phase transition line in the phase diagram is crossed, it is normally expected that the substance would change its phase state. However, such a change would require additional energy input for nucleation of seeds of the new phase. For example, the liquid–vapor transition requires nucleation of vapor bubbles (this process is called cavitation), while the liquid–solid transition requires nucleation of ice crystals. If special measures are taken to prevent nucleation of the seeds of the new phase, it is possible to postpone the transition to the equilibrium state phase leading to supercooled, superheated, or stretched water (Fig. 1).

Two factors affect a vapor bubble inside liquid, needed as a nucleation center for vapor in superheated or stretched water: the pressure and the interfacial tension. While the negative pressure (tensile stress) is acting to expand the size of the bubble, the interfacial tension is acting to collapse it. For a small bubble, the interfacial stress dominates; however, for a large bubble, the pressure dominates. Therefore, there is a critical radius of the bubble, and a bubble with a radius larger than the critical radius would grow, whereas one with a smaller radius would collapse. The value of the critical radius is given by:

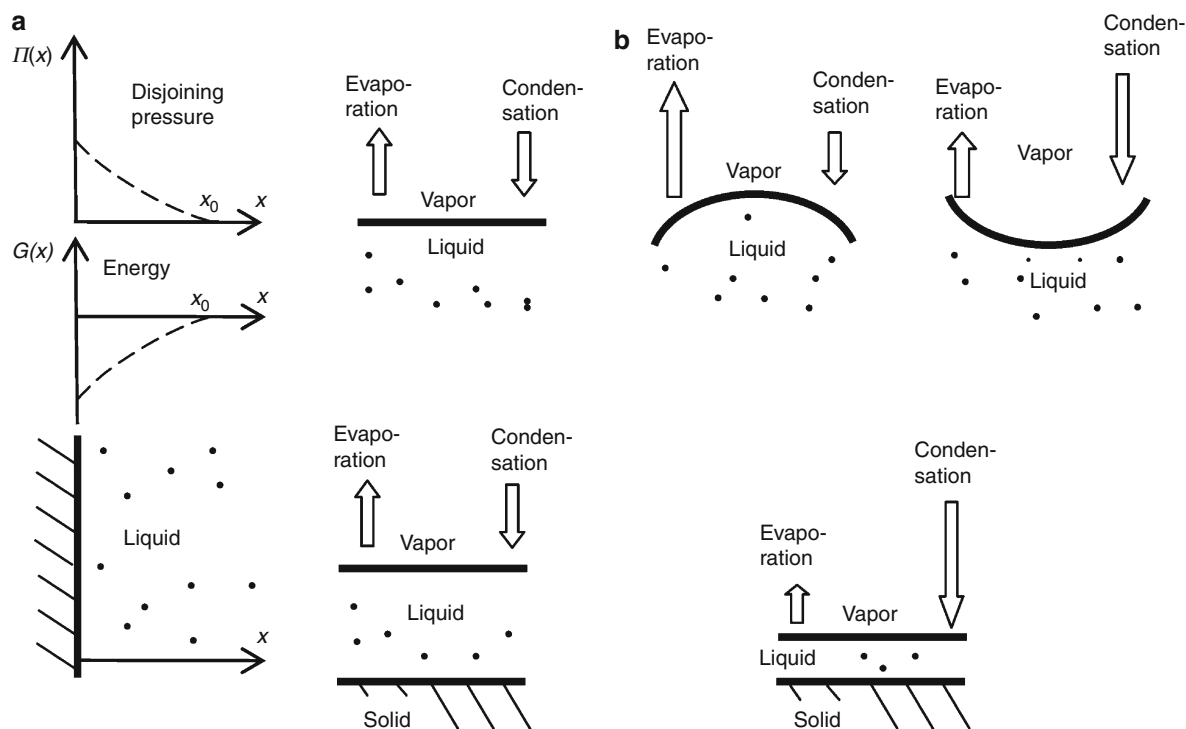
$$R_C = \frac{2\gamma_{LV}}{P_{sat} - P} \quad (2)$$

where γ_{LV} is the liquid–vapor surface tension, P_{sat} is the saturated vapor pressure, and P is the actual liquid pressure [1].

With the further decrease of pressure in the negative region, the so-called spinodal limit can be reached. At that pressure, the critical cavitation radius becomes of the same order as the thickness of the liquid–vapor interface. In this case, there is a lower energy path of nucleation connecting the liquid to the gas phase by expansion of a smoothly varying density profile. In the phase diagram, the line that corresponds to the spinodal limit is expected to go all the way to the critical point. This critical spinodal pressure at a given temperature effectively constitutes the tensile strength of liquid water. Various theoretical considerations, experimental observations, and MD simulation results have been used to determine the value of the spinodal limit. At room temperature, the spinodal pressure is between -150 and -250 MPa [1].

Disjoining Pressure

Another important effect is the so-called disjoining pressure. The adhesion force between the solid and water has a certain range x_0 and decreases with the distance x from the interface. As a result, water in the layer next to the interface is subjected to higher pressure $P = P_0 + \Pi_d(x)$ than the bulk pressure, where $\Pi_d(x)$ is the so-called disjoining pressure. For a thin liquid layer of thickness $H < x_0$, the evaporation/condensation equilibrium will be shifted toward condensation, since an additional attractive force acts upon water molecules from the solid surface. As a result, a thin water layer can be in equilibrium with undersaturated vapor, $P_V/P_{sat} < 1$, so the effect of the disjoining pressure on the thermodynamic equilibrium is similar to the effect of concave meniscus [4]. The influence of the disjoining pressure and the related wetting films in narrow confinement have to be considered as they may significantly change the meniscus curvature (Fig. 2). A typical effect of nanoscale confinement is the appearance of capillarity-induced negative pressure [3].



Nanoscale Water Phase Diagram, Fig. 2 (a) Disjoining pressure near a solid surface (b) The effect of meniscus curvature (*flat, concave, and convex*) and liquid layer thickness on the evaporation–condensation equilibrium

Key Findings

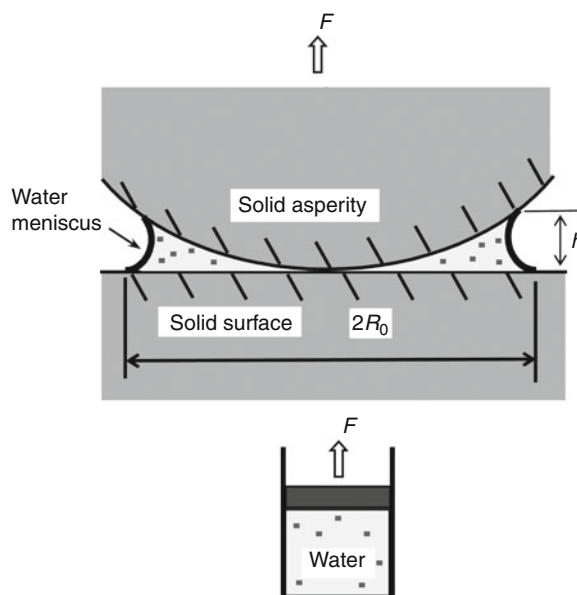
Negative Pressure in Water Capillary Bridges

It is very difficult to obtain deeply negative pressure at the macroscale because of bubble nucleation. The pressure values that have been reported constitute -19 MPa with the isochoric cooling method, -17.6 MPa with the modified centrifugal method, and -25 MPa with the acoustic method. The situation is different at the micro- and nanoscale. The pressure of -140 MPa in the microscopic aqueous inclusions in quartz crystals was reported. Water plugs in hydrophilic nanochannels can be at a significant absolute negative pressure due to tensile capillary forces [1].

Negative pressure of -160 MPa was reported in nanoscale liquid capillary bridges. The capillary component of the adhesion force was determined by the comparison of the adhesion force in air (at different RH levels) and in ultra-high vacuum (UHV), where capillary bridges are absent. Very small values of the Atomic force microscope (AFM) tip radius (on the order of 10 nm) correspond to a small size of meniscus (meniscus foundation area on the order of 10^{-16} m²) and thus even a small force of 10 nN corresponds to huge pressure differences inside and outside the meniscus on the order of 100 MPa or $1,000$ atm (Fig. 3). The reason for the existence of such large tension limits lies in the strength of the hydrogen bonds which must be broken in order to form a cavity that is large enough to act as a stable nucleus for cavitation. The metastable state of the capillary bridges is observed because the size of the bridges is smaller than the critical cavitation radius given by (2) and, therefore, no cavitation occurs [1].

Room Temperature Ice in Capillary Bridges

Most AFM results show that the capillary force has a strong dependence on RH : first it increases with increasing relative humidity up to approximately 30% and then decreases with a further increase of the relative humidity. For low humidity, the pressure difference inside and outside the bridge is large; however, the size of the bridge is small due to high curvature of the meniscus. For very low humidity, the size of the bridge may be insufficient to generate a capillary force. On the other hand, for high humidity, the pressure difference is small; however, the meniscus is large (but, of course, its size is limited by the radius of the



Nanoscale Water Phase Diagram, Fig. 3 A water capillary bridge between an AFM tip and a flat sample. The Laplace pressure inside the bridge is reduced because the bridge is concave. The bridge is under tension, similarly to water in a sealed tube with a piston with the applied tensile force F

tip). Thus, it is expected that in the limiting cases of $RH = 0\%$ and $RH = 100\%$ the capillary force would vanish, since at $RH = 0\%$ there is no capillary bridge and at $RH = 100\%$ the Laplace pressure drop is zero. The capillary force achieves its maximum at a certain value of RH [3].

It has been argued that the capillary force alone cannot explain the observed magnitude of the humidity dependence. The origin of the large RH dependence is due to the presence of ice-like structured water adsorbed at the silicon oxide surface at room temperature. Thus, the force is due to the rupture of an ice-ice bridge at the center of the contact region with the structure, thickness, and viscoelastic behavior of the adsorbed water layer influencing the adhesion force [2].

Water often forms adsorbed layers at the surface of many materials, and these layers may have an ice-like structure and be connected to the capillary bridge, as indicated by the study of the molecular configuration of adsorbed water as a function of relative humidity using attenuated total reflection (ATR) infrared spectroscopy [4]. A completely hydrogen-bonded ice-like network of water grows up to three layers as the

relative humidity increases from 0% to 30%. In the relative humidity range of 30–60%, the liquid water structure starts appearing, while the ice-like structure continues growing to saturation. Above 60% relative humidity, the liquid water configuration grows on top of the ice-like layer.

During the contact between a tungsten tip and a graphite surface it was found that water in the nanoscale capillary bridges may include ice or be a mixture of liquid water with ice [2]. Such a mixture would have high viscosity and may act like glue, demonstrating elastic response.

Low-Temperature Water Boiling

Another important consequence of the water pressure reduction in the capillary bridges is the possibility of boiling at temperatures lower than 100°C. If a capillary bridge is large enough, so that cavitation occurs, the pressure inside the bridge is reduced compared with the ambient, and the boiling point is depressed by the amount ΔT . The Clausius–Clapeyron equation provides the change of pressure $\Delta P = P_2 - P_1$ as related to the change of the boiling temperature $\Delta T = T_2 - T_1$ as

$$\ln(P_2/P_1) = \Delta h/R(1/T_1 - 1/T_2) \quad (3)$$

where Δh is the change of specific enthalpy [54]. Note that ΔP is given by (1).

For example, for a meniscus between an AFM cantilever and a flat surface with a gap between them of 3 μm and the principal radii of curvature of the meniscus $R_1 = R_2 = -2 \mu\text{m}$, the Kelvin radius is $R_k = (1/R_1 + 1/R_2)^{-1} = -1 \mu\text{m}$ and the pressure drop, based on (6) is $\Delta P = 72.9 \text{ kPa}$ (or 0.72 atm). Thus, the pressure inside the meniscus is $P = 28.1 \text{ kPa}$ (or 0.28 atm), which corresponds to the boiling point of about 333 K [3].

Nanobubbles

There is no adequate theoretical explanation why ice can be present in the capillary bridges at room temperature; however, it is clear that this effect is related to the adsorption of water to solid surfaces. The ability of water to form thin layers on hydrophobic surfaces is associated with the disjoining pressure caused by the spatial dependence of the van der Waals adhesion force. Due to the dependence of the adhesion force of the distance (on the scale of nanometers), the energy of the water molecules in the vicinity of the solid surface

is lower than the energy of those far away. Thus, it is often not only energetically profitable for water to cover the solid, but also to create a thin layer.

It is noted that not only the solid phase can be adsorbed at a surface. Very thin (5–80 nm) gas films could exist for a long time at the interface between a hydrophobic solid and water. Despite certain thermodynamic objections for nanobubble existence (such as expected high pressures in bubbles leading to their dissolution), some experimental studies with the AFM and other equipment have demonstrated the existence of nanobubbles at the solid/water interface [5,6]. Nanobubbles can play a significant role in a number of processes, such as the long-range hydrophobic attractive force.

Water in Confinement

A number of studies are devoted to water phase behavior in confinement. MD simulations show that water confined to carbon nanotubes of a critical size under ambient conditions (1 bar, 300 K) can undergo a transition into a state having ice-like mobility with an amount of hydrogen bonding similar to that in liquid water. Maeda [7] investigated experimentally phase transitions of capillary-held liquids in a slit-like pore. When the surfaces were placed at a separation very close to the Kelvin length, it was possible to detect a stage in which the system was in an apparent kinetic equilibrium between two physical states – with and without the liquid connecting the two surfaces. Giovambattista et al. [8] used MD simulation to investigate the phase behavior of water confined between two nanoscale hydrophobic or hydrophilic plates at room temperature, varying the pressure ($-0.15 \text{ GPa} \leq P \leq 0.2 \text{ GPa}$) and plate separation ($0.4 \text{ nm} \leq d \leq 1.6 \text{ nm}$). With hydrophobic plates, capillary evaporation occurred between the plates at low P . The threshold value of d at which this transition occurs decreases with P (e.g., 1.6 nm at $P = 0.05 \text{ GPa}$, 0.5 nm at $P = 0.1 \text{ GPa}$), until, at high P , no capillary evaporation occurred. For $d = 0.6 \text{ nm}$ and $P > 0.1 \text{ GPa}$, the system crystallizes into a bilayer ice. They developed a P – d phase diagram showing the vapor, liquid, and bilayer ice. With hydrophilic plates, water always remained liquid even at the P at which bulk water cavitates. This indicates that size is important for a phase diagram due to the finite range of the adhesion force in addition to the critical cavitation radius and curvature of the meniscus, which were discussed in the preceding sections.

Cross-References

- ▶ [AFM in Liquids](#)
- ▶ [Disjoining Pressure and Capillary Adhesion](#)
- ▶ [Wetting Transitions](#)

References

1. Yang, S.-H., Nosonovsky, M., Zhang, H., Chung, K.-H.: Negative pressure in water capillary bridges at nanocontacts. *Chem. Phys. Lett.* **451**, 88–92 (2008)
2. Jinesh, K.B., Frenken, J.W.M.: Capillary condensation in atomic scale friction: how water acts like a glue. *Phys. Rev. Lett.* **96**, 166103 (2006)
3. Nosonovsky, M., Bhushan, B.: Phase behavior of capillary bridges: towards nanoscale water phase diagram. *Phys. Chem. Chem. Phys.* **10**, 2137–2144 (2008)
4. Asay, D.B., Kim, S.H.: Effects of adsorbed water layer structure on adhesion force of silicon oxide nanoasperity contact with humid ambient. *J. Chem. Phys.* **124**, 174712 (2006)
5. Steitz, R., Gutberlet Hauss, T., Klosgen, B., Krastev, R., Schemmel, S., Simonsen, A.C., Findenegg, G.H.: Nanobubbles and their precursor layer at the interface of water against a hydrophobic substrate. *Langmuir* **19**, 2409–2418 (2003)
6. Mashl, R.J., Joseph, S., Aluru, N.R., Jakobsson, E.: Anomalous immobilized water: a new water phase induced by confinement in nanotubes. *Nano Lett.* **3**, 589–592 (2003)
7. Maeda, N.: Phase transitions of capillary-held liquids in a slit-like pore. *J. Phys. Chem. B* **110**, 25982–25993 (2006)
8. Giovambattista, N., Rossky, P.J., Debenedetti, P.G.: Effect of pressure on the phase behavior and structure of water confined between nanoscale hydrophobic and hydrophilic plates. *Phys. Rev. E* **73**, 1–14 (2006)

Nanoscratch Tester

- ▶ [Nanoindentation](#)

Nano-sized Nanocrystalline and Nano-twinned Metals

Dongchan Jang
Department of Applied Physics and Materials Science,
California Institute of Technology, Pasadena,
CA, USA

Synonyms

[Electroplating](#); [Nanocrystalline materials](#); [Nano-pillars](#); [Nano-twinned Materials](#)

Definition

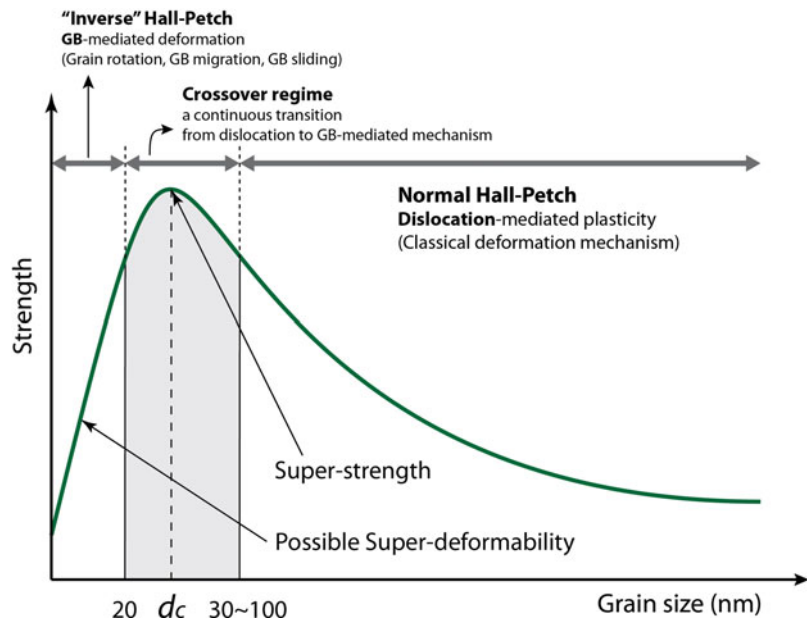
- Nano-pillars are pillar-shaped specimens at the nanoscale, whose aspect ratio (height/diameter) is generally within $3 \sim 7$.
- Nanocrystalline material refers to a class of polycrystalline materials whose grain size is smaller than ~ 100 nm.
- Nano-twin refers to a class of crystalline materials that contain a number of parallel coherent twin boundaries with the inter-boundary distance at the nanoscale.

Overview

Traditionally, the most fundamental principle in materials science is that the properties of materials are strongly influenced by their internal microstructure [1]. During the plastic deformation of many metallic systems, dislocations carry most of the plasticity and hence their interactions with the internal microstructure, such as grain boundaries, solute atoms, precipitation or even other dislocations, are the most influential in determining mechanical responses of those materials [2]. However, recent studies found that the microstructure is not the only factor that affects the mechanical properties, but the external dimension of the specimen also significantly changes the mechanical behavior when it is reduced to the submicron and nanometer scale [3]. Therefore, when the effects from those two factors, microstructure and specimen size at the nanoscale, are combined, the materials often exhibit very unique properties distinguished from both of the bulk samples containing same kind of microstructure and nano-samples not containing the corresponding microstructure [3]. Especially, the homogeneous internal interfaces, such as ordinary grain boundary (GB) and coherent twin boundary (TB), are one of the most interesting microstructures as their interaction to dislocations is a key parameter in understanding plasticity of metallic materials. In this entry, the review of the recent studies investigating the mechanical properties of nano-sized metallic samples containing ordinary GBs and coherent TBs is presented. The main focus is placed on the results from uniaxial tension/compression experiments on the nanometer-sized pillar-shaped specimen (nano-pillars). Fabrication techniques to produce the nano-pillars with various internal interfaces are also introduced.

Nano-sized Nanocrystalline and Nano-twinned Metals, Fig. 1

Strength of polycrystalline materials as a function of grain size: Hall-Petch relation and transition to inverse Hall-Petch (Reprint with permission from [3])



Mechanical Properties

Nanocrystalline Nano-pillars

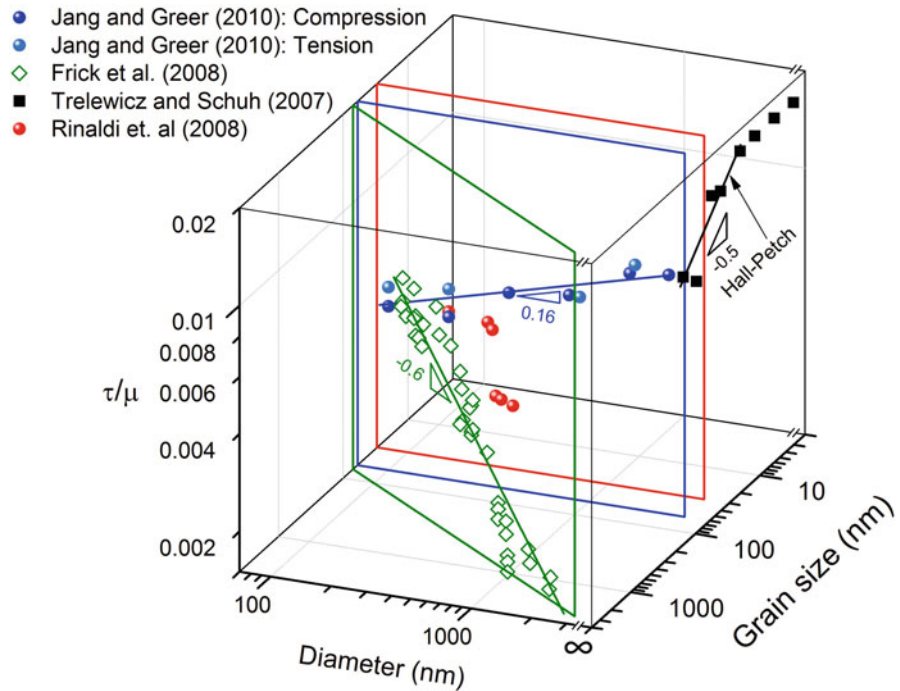
GB is the interface between differently oriented grains in the polycrystalline materials. Generally, there is atomic mismatch across the GB as the crystallographic orientation of one grain abruptly changes to that of the adjacent one [1]. In conventional bulk polycrystalline materials, GBs usually act as a barrier to the dislocation motion across the boundary because of the discontinuity of atomic planes at the GB [1]. As a result, the strength of the bulk polycrystalline materials is inversely correlated with the grain size because total GB area increases with decreasing grain size, so dislocations are impeded more often with smaller grain size. For many conventional polycrystalline metals, the strength is inversely proportional to the square root of grain size, i.e., $\sigma_y = \sigma_o + k_y d^{-1/2}$, where σ_y and d are the yield strength and grain size, respectively, and σ_o and k_y are constants for a particular material. This expression is called Hall-Petch relation [1, 4, 5]. However, this Hall-Petch relation cannot be extrapolated down to infinitesimally small grain sizes. Instead, it is reported to break down when grain size is reduced below a critical value (usually 10 ~ 20 nm). Below this critical size, the strength is often saturated or decreases with further reduction of grain size. This reversed relation between the strength and grain size is called inverse Hall-Petch equation. The deviation

from the Hall-Petch relation is likely due to the transition in plastic deformation mechanism from dislocation-driven in Hall-Petch to GB-mediated in inverse Hall-Petch. The dependence of strength on the grain size and the corresponding deformation mechanism change are schematically summarized in Fig. 1. Further details of bulk nanocrystalline materials are available in the entry “► [Mechanical Properties of Nanocrystalline Metals.](#)”

As briefly stated in the Overview section, the specimen with external dimensions at the nanoscale shows different mechanical responses from those of their bulk counterpart. For example, the strengths of single-crystalline nano-pillars drastically increase when their diameters decrease below submicron scale, showing the well-known “Smaller is stronger” trend (See the entry “► [Size-Dependent Plasticity of Single Crystalline Metallic Nanostructures](#)” for further details). In nanocrystalline nano-pillars, where both of the external specimen and internal grain sizes are in the nanometer range, the mechanical properties are clearly distinguishable from those of bulk nanocrystalline materials and single-crystalline nano-pillars. Figure 2 shows the 3D plot, where the strengths for nearly the same materials, Ni, are plotted as functions of nano-pillar diameter (D) and grain size (d). Quantities τ and μ in this plot are the shear strength and shear modulus, respectively. When $D \gg d$, the sample dimension approaches to the bulk

Nano-sized Nanocrystalline and Nano-twinned Metals,

Fig. 2 3D representation of strength of Ni as functions of nano-pillar diameter (D) and grain size (d). The quantities τ and μ are the shear strength and shear modulus, respectively (Reprint with permission from [3])



poly-/nanocrystalline materials, and the strength becomes independent of the sample size, D . Therefore, the strength is the function of grain size, d , only and follows the ordinary Hall-Petch and inverse Hall-Petch relations of the bulk nanocrystalline Ni. In Fig. 2, the data points (solid squares) on the vertical plane perpendicular to the D axis at $D \sim \infty$ show these relations. When $D = d$, there is only one grain across the diameter, and the resultant microstructure is single crystal (another possibility is bamboo structure, where the GBs are orthogonal to the pillar axis, but this structure is not discussed in this entry). The single-crystalline nano-pillars display “Smaller is stronger” trend, where the strength increases as decreasing the pillar diameter, D . In Fig. 2, the green plane satisfies the condition of $D = d$, and the data points on that plane (open diamonds) represent the strength of single-crystalline Ni nano-pillars as a function of pillar diameter, D (or equivalently as a function of d). The linear regression line in this log-log scale plot shows the power-law slope of ~ 0.6 . Further details on the single-crystalline nano-pillars are available in the entry “Plastic Properties of Single Crystalline at the Nanoscale” or in Refs. [3, 6]. While both of bulk nanocrystalline Ni and single-crystalline Ni nano-pillars exhibit “Smaller is stronger” trend, where the strength increases as decreasing either d or D (in bulk nanocrystalline

samples, this statement is valid only down to the critical grain size of ~ 20 nm), the strength of nanocrystalline nano-pillars shows different aspects from both of them. In Fig. 2, the vertical plane perpendicular to the d axis at $d = 60$ nm displays the strength of nanocrystalline Ni nano-pillars with fixed grain size of 60 nm as a function of D ranged from 1,600 to 100 nm. Unlike bulk nanocrystalline Ni or single-crystalline Ni nano-pillars, this nanocrystalline Ni nano-pillars show “Smaller is weaker” trend, where the strength decreases with decreasing D . Possible explanation to this trend can be due to the contribution of the free surface, whose relative fraction increases with decreasing the pillar diameter. In general, the strength of an individual grain embedded within the surrounding ones (i.e., a grain in bulk polycrystalline materials) is higher than that of the same but completely freed grain. It is because for plastic deformation in bulk polycrystalline materials, all grains must deform cooperatively by matching the displacement across the GB, and it applies additional constraints to the plastic deformation. However, when the adjacent grains are removed, those constraints are released, and the grain can deform more easily. In nanocrystalline nano-pillars, the grains located at the free surface (surface grains) have less constraint than those in the pillar (inner grains) since the surface grains

face the empty space at least to one direction. Therefore, the strength of the surface grains is lower than that of the inner grains, dividing the whole nano-pillars into hard inner core and soft outer shell. The thickness of the outer shell is approximately same with the average grain size of the nano-pillar. Therefore, when they have a fixed average grain size, the relative fraction of the soft outer shell increases as the diameter of the nanocrystalline nano-pillars decreases. As a result, the strength of the nanocrystalline nano-pillar lowers [3].

Another important observation on the nanocrystalline nano-pillars is the onset of GB-mediated plasticity at larger grain size than that of bulk nanocrystalline samples. In bulk nanocrystalline metals, activation of GB-mediated plasticity is usually correlated with transition from the Hall-Petch to inverse Hall-Petch relation, and this transition grain size for Ni is around 20 nm, as can be seen in Fig. 2. However, recent study reveals that reduction of the nano-pillar diameter to ~ 100 nm facilitates the activation of GB-mediated process, such as grain boundary sliding and grain rotation, at 60-nm grain size, which is significantly larger than the transition size of bulk samples, 20 nm [3].

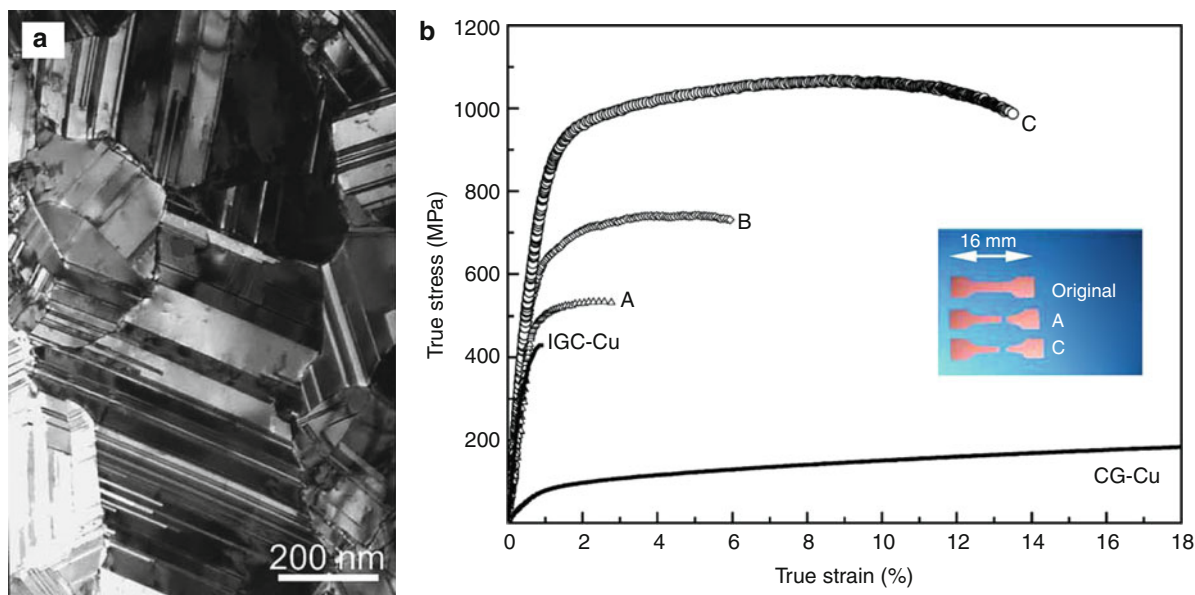
Nano-twinned Nano-pillars

In general, the misorientation angles between the crystallographic planes across the ordinary GBs described in the previous section are randomly distributed, and atomic structure of those GBs is incoherent in that the lattice planes lose the crystallographic registry across the interface. However, there is a different type of internal coherent and homogeneous interface, called twin boundary (TB). Twinned microstructure usually refers to a special type of relative adjacent grain orientations, where one of the grains has the mirror-symmetric crystallographic structure of the other, and the mirror plane is called a twin boundary [1]. In this structure, one of the grains is called matrix and the other twin. Particularly, when the lattice planes do not lose the crystallographic registry across the boundary, the TB is called coherent. Because the TBs are also a special type of the GB, their strengthening mechanism shares many common factors with the ordinary GBs, i.e., TBs also interfere with the glide of dislocations. However, their influence on strengthening is known to be less pronounced because they are less effective at resisting the penetration of the moving dislocations, especially when the distance between the neighboring TBs is in micron range [7].

Over the last decade, the so-called nano-twinned materials (usually metals) have attracted a lot of scientific and industrial interest as promising candidates for the next generation structural and/or electrical material because of their simultaneous attainment of high strength, ductility, and electrical conductivity. In these nano-twinned metals, the nano-twinned area consists of alternative pileup of nanometer-thick planar matrix and twin grains (nano-twin lamellas). The transmission electron microscope (TEM) image of a typical nano-twinned Cu and the corresponding stress-strain curve is presented in Fig. 3 [8].

An important difference between ordinary GB and coherent TB in their roles during plastic deformation is that while the dislocation can hardly move along or across the ordinary GBs because of their disordered or discontinued atomic structure, the coherent TBs relatively easily allow dislocations to glide along and across the boundary [7]. The consequence of this difference is that dislocation mobility is highly limited in the presence of GBs, resulting in increased strength with sacrificing the ductility in regular nanocrystalline metals. On the other hand, because of its coherent nature, the TBs are capable of accommodating more various interactions with moving dislocations other than simply blocking their glide. For example, a moving dislocation can be absorbed to the TB, leaving a sessile dislocation on the interface plane, or can be dissociated into two gliding partial dislocations on the TB plane, leaving the stacking fault, or can transmit across the TB, leaving a step at the TB [7]. All of these interactions gradually degrade coherency of the TBs, hinder subsequent dislocations in interacting in the similar manners, and hence change the characteristic of coherent TB close to that of ordinary incoherent GBs. As a result, it gradually requires higher applied stress to keep the dislocations moving, and the material is work-hardened [7]. Usually, the increase of work-hardening coefficient correlates with higher resistance to instability, such as necking [4, 9], so it can contribute to the effective increase of ductility.

Most of nano-twinned samples fabricated to date are polycrystalline in the sense that the nano-twin lamellas are embedded within polycrystalline grains as sub-grain structures as can be seen in Fig. 3. Properties measured in the polycrystalline nano-twinned samples are convoluted by the effects from many different microstructures like GBs, and differently orientated TBs, etc. Therefore, the specific role of TBs and



Nano-sized Nanocrystalline and Nano-twinned Metals, Fig. 3 (a) TEM observations of typical microstructure in nano-twinned Cu sample. (b) Typical tensile stress-strain curves for the as-deposited Cu samples with average nano-twin thickness

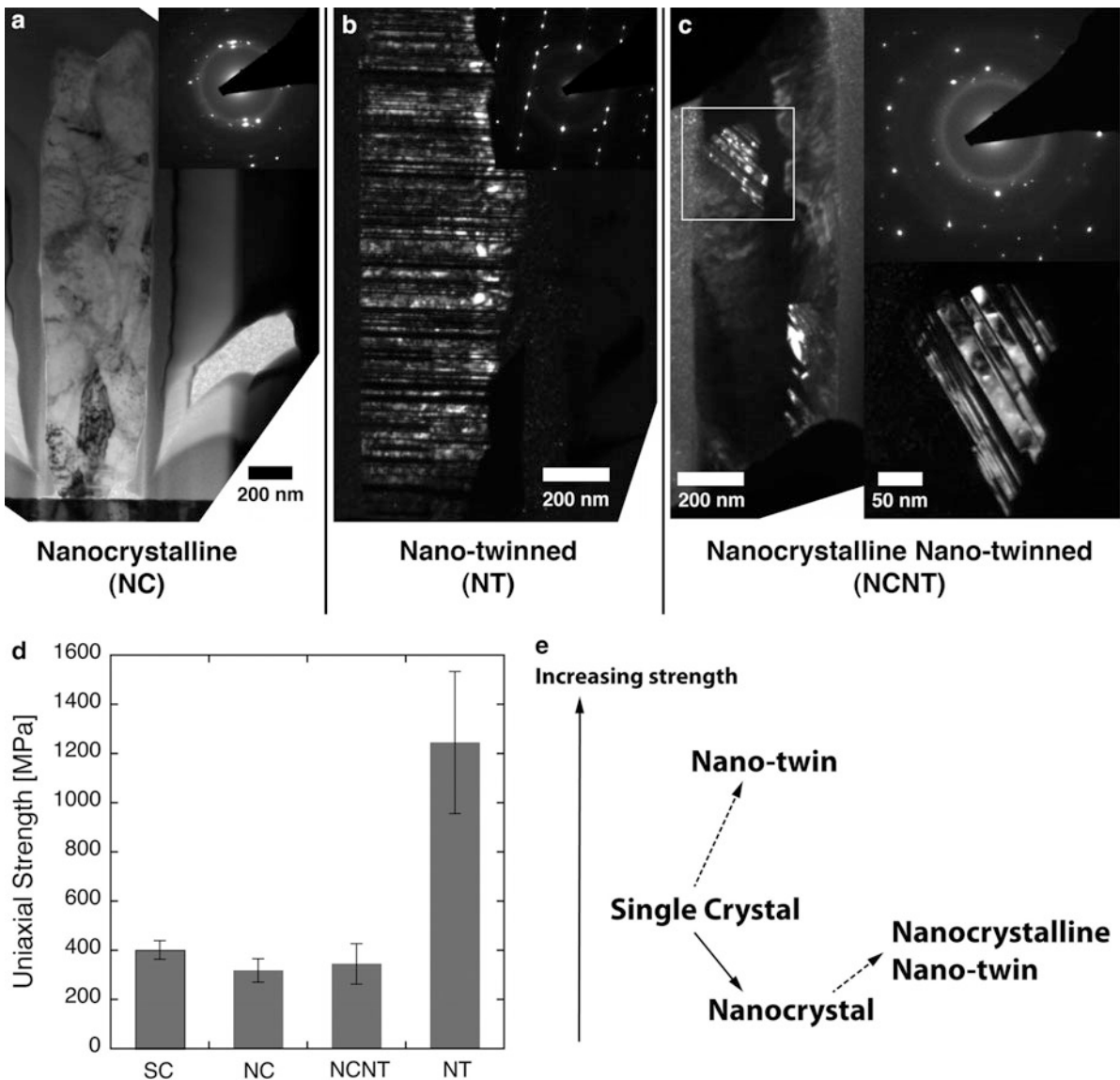
of 93, 30, and 15 nm for samples A, B, and C, respectively. For comparison, the stress-strain curves for coarse-grained polycrystalline Cu sample (CG-Cu) and nanocrystalline Cu (IGC-Cu) are also displayed (Reprint with permission from [8])

nano-twinned lamellae in the enhanced mechanical properties is far from being understood. However, recent studies on the nano-twinned nano-pillars, where the specimen contains well-aligned TBs only, offered an insight into understanding the specific role of TB during plastic deformation. The TEM images in Fig. 4a–c show the Cu nano-pillars with 500-nm diameters, whose microstructures are nanocrystalline (NC), purely nano-twinned (NT), and nanocrystalline nano-twinned (NCNT), respectively [10]. NC nano-pillars only contain the randomly distributed ordinary GB with average grain size of ~ 160 nm, while NT nano-pillars have only coherent TBs orthogonal to the pillar axis. Average twin thickness of NT nano-pillars is ~ 8.8 nm. NCNT nano-pillars contain both of GBs and TBs, where nano-twin lamellae are embedded within the regular grains. Average grain size and twin thickness of NCNT nano-pillars are 250 and 6.7 nm, respectively. The strengths of those Cu nano-pillars measured by uniaxial nano-compression experiments are shown in Fig. 4d, where SC stands for single-crystalline Cu nano-pillars. These results suggest that in Cu nano-pillars with 500-nm diameter, the two different interfaces, ordinary GB and coherent TB, appear to affect the strengths differently. Given that

strength of NC is lower than that of SC by $\sim 20\%$, the ordinary GBs seem to lower the strength of the nano-pillars. On the other hand, strength of NT is higher than that of SC by a factor of ~ 3 , implying that orthogonally aligned TBs drastically increase the strength of nano-pillars. Comparing the strengths of NC and NCNT, TBs also increase the strength even in the polycrystalline case. The influence of GBs and TBs on the strength of 500-nm Cu nano-pillars is summarized in the diagram in Fig. 4.

Fabrication of Nano-pillars via Electroplating

Traditionally, electroplating is a widely used metallic surface coating technique on a conductive surface. During electroplating, the metallic ions in the electrolyte migrate to the surface of the cathode under externally applied electrical voltage between the cathode and anode, and deposit a thin metallic layer on the cathode. Basically, the fabrication of nano-pillars shares the same physics and chemistry with this traditional electroplating technique. However, rather than forming a metallic layer of uniform thickness, it requires the ions to deposit only onto prescribed

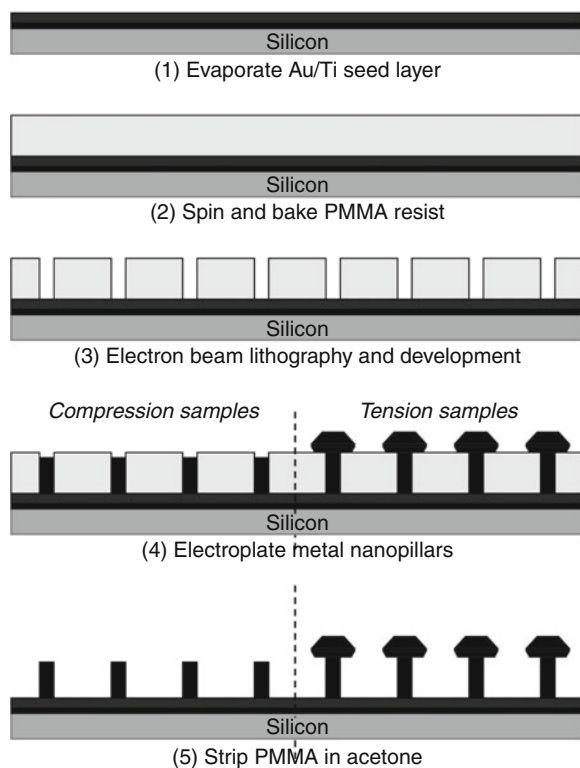


Nano-sized Nanocrystalline and Nano-twinned Metals, Fig. 4 (a) TEM bright-field image and diffraction pattern (*inset*) of a typical nanocrystalline (NC) Cu nano-pillar. TEM dark-field images and diffraction patterns (*upper-right insets*) of typical (b) nano-twinned (NT) and (c) nanocrystalline nano-twinned (NCNT) Cu nano-pillars. The *lower-right inset*

in (c) is the zoomed-up image of the region within the *white box*. (d) The average yield strength of each microstructure. (e) Illustration showing the effect of grain and nano-twin boundaries on the strength of 500-nm Cu nano-pillars (Reprint with permission from [10])

locations. This can be achieved by using the pre-patterned template as the cathode. The template consists of the chemically etchable (e.g., polymethylmethacrylate (PMMA)) patterned nonconductive sacrificial layer on top of a conductive metallic thin film deposited on an insulating substrate (Si, SiO₂ or sapphire). A pattern of desired shapes can be written

via e-beam lithography, and the resist is subsequently developed to generate arrays of through-holes to the underlying conductive metallic film. The underlying thin metallic layer provides an electric-conducting path for the electroplating, and the ions deposit onto the metallic layer through the open spaces in the patterned template. After electroplating, the PMMA



Nano-sized Nanocrystalline and Nano-twinned Metals, Fig. 5 Schematic representation of template fabrication and electroplating (Reprint with permission from [11])

sacrificial layer can be stripped off in acetone. Therefore, the resulting electroplated structure becomes a negative of the pattern. The nonconductivity of the substrate prevents the ions from depositing to the backside of the template. The schematic illustration of this process is presented in Fig. 5.

The specific details of the template fabrication are available in [11]. For example, a 100-nm thick Au seed layer on top of the 20-nm thick Ti adhesion layer is deposited on the Si wafer by electron beam evaporation. The Au layer provides the electric-conducting path for the electroplating and Ti layer enhances the adhesion of the Au layer to the Si substrate. The PMMA resist diluted in anisole is spin-coated on top of the Au seed layer. The thickness of the PMMA resist determines the maximum height of the nano-pillars, as illustrated in Fig. 5. Table 1 shows the list of some examples of the resist parameters and corresponding layer thicknesses [11].

The fabricated template can be used as the part of the cathode during the electroplating. The cathode is

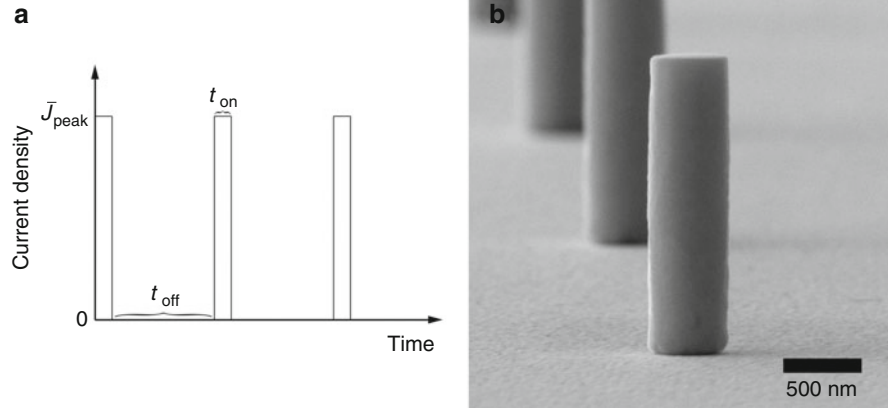
Nano-sized Nanocrystalline and Nano-twinned Metals, Table 1 Polymethylmethacrylate (PMMA) resist and electron beam exposure parameters for template fabrication [11]. The quantity d is the diameter of the circular patterns, which is effectively the same as the diameter of the electroplated nanopillars, and L is the thickness of PMMA (Reprint with permission from [11])

Pillar dimensions		Resist parameter		Exposure parameters Dose ($\mu\text{C}/\text{cm}^2$)
d (nm)	L (nm)	Dilution (%)	Speed (rpm)	
25	100	4	2,500	2,600
50	200	4	2,500	1,600
100	400	7	3,000	1,450
150	600	7	1,500	1,600
200	800	9	4,500	1,550
250	1,000	9	2,000	1,500
500	2,000	11	2,500	1,600
750	3,000	11	1,500	1,750

split into two parts: the template and a Cu-coated dummy chip. The purpose of the dummy chip is to precisely control the current density at the cathode (\bar{J}), which is defined as the total current divided by the cross-sectional area of the cathode. With the template alone, the cross-sectional area, which is the sum of the area of the patterns on the template, is extremely small (on the order of nm^2 or μm^2) so that the current density can sharply change by a small error in the area measurement. By adding the dummy chip whose cross-sectional area is a few orders of magnitude larger than the patterns, the total cross-sectional area of the anode becomes insensitive to the small error in the area measurement, so that the current density at the cathode can be reliably controlled. As the anode material, either insoluble platinized titanium mesh or soluble metallic plate can be used, depending on the material to be plated. Microstructure of the electroplated nanopillars is highly dependent on the plating condition, including composition of electrolyte, waveform of the applied current, and current density at the cathode. For example, in order to fabricate the Cu nano-pillars, the platinized titanium mesh was used as the anode material, and rectangular pulsed current was applied between the cathode and anode, as shown in Fig. 6a, where the current density is maintained at \bar{J}_{peak} during on time, t_{on} and becomes zero during off time, t_{off} [10]. Electrolyte composition is also an important factor to control the microstructure of the Cu nano-pillars. In Table 2, the electroplating parameters to create the NT,

Nano-sized Nanocrystalline and Nano-twinned Metals, Fig. 6

(a) The waveform of the pulsed electroplating current. (b) SEM image of a typical as-fabricated Cu nano-pillar (Reprint with permission from [10])



Nano-sized Nanocrystalline and Nano-twinned Metals, Table 2 The electroplating condition for the fabrication of 500-nm Cu nano-pillars for each microstructure (Reprint with permission from [10])

Microstructure	Electrolyte composition	J_{on} (A/cm ²)	t_{on} (ms)	t_{off} (ms)	Feature sizes	
					Grain size (nm)	Twin thickness (nm)
Nano-twinned (NT)	125 g/L of CuSO ₄ ·5H ₂ O + 50 g/L of H ₂ SO ₄	0.4	1	100	N/A	8.8
Nanocrystalline (NC)	100 g/L of C ₆ H ₈ O ₇ ·H ₂ O + 28 g/L of CuSO ₄ ·5H ₂ O + 50 g/L of (NH ₄) ₂ SO ₄	1.2	1	100	160	N/A
Nanocrystalline Nano-twinned (NCNT)	50 g/L of C ₆ H ₈ O ₇ ·H ₂ O + 28 g/L of CuSO ₄ ·5H ₂ O + 50 g/L of (NH ₄) ₂ SO ₄	0.72	1	100	250	6.7

NC, and NCNT Cu nano-pillars are summarized, including the electrolyte composition, J_{peak} , t_{on} , and t_{off} [10]. The Scanning Electron Microscope (SEM) image of a typical as-fabricated pillar is shown in Fig. 6b.

Fundamentally, all conventional electroplatable metals can be also fabricated into the nano-pillar forms utilizing the method in the previous paragraph. However, determining electroplating parameters for specific microstructure, especially the waveform of the applied current, can be challenging. While specific details have to be developed for each material, the general guideline is that (1) the grain size inversely scales with the plating overpotential, and hence reducing forward current density often yields the larger grain size. (2) Sometimes, using pulsed current is efficient in producing larger grain size as it periodically reduces the depletion region near the cathode, and consequently lowers the overpotential. The recipes to create single crystalline Cu [11], polycrystalline Au [11], and polycrystalline In [12] nano-pillars are available elsewhere.

Cross-References

- ▶ [Mechanical Properties of Nanocrystalline Metals](#)
- ▶ [Nanomechanical Properties of Nanostructures](#)
- ▶ [Plasticity Theory at Small Scales](#)
- ▶ [Size-Dependent Plasticity of Single Crystalline Metallic Nanostructures](#)

References

1. Callister, W.D.: Materials Science and Engineering: An Introduction. Wiley, New York (1997)
2. Hirth, J.P., Lothe, J.: Theory of Dislocations. Wiley, New York (1982)
3. Greer, J.R., De Hosson, J.T.M.: Plasticity in small-sized metallic systems: intrinsic versus extrinsic size effect. Prog. Mater. Sci. **56**, 654–724 (2011)
4. Courtney, T.H.: Mechanical Behavior of Materials. McGraw-Hill, Boston (2000)
5. Meyers, M.A., Mishin, Y., Benson, D.J.: Mechanical properties of nanocrystalline materials. Prog. Mater. Sci. **51**, 427–556 (2006)

6. Zhu, T., Li, J.: Ultra-strength materials. *Prog. Mater. Sci.* **55**, 710–757 (2010)
7. Lu, K., Lu, L., Suresh, S.: Strengthening materials by engineering coherent internal boundaries at the nanoscale. *Science* **324**, 349–352 (2009)
8. Shen, Y.F., Lu, L., Lu, Q.H., Jin, Z.H., Lu, K.: Tensile properties of copper with nano-scale twins. *Scripta Mater.* **52**, 989–994 (2005)
9. Dieter, G.E.: *Mechanical Metallurgy*. McGraw-Hill, New York (1988)
10. Jang, D., Cai, C., Greer, J.R.: Influence of homogeneous interfaces on the strength of 500 nm diameter Cu nanopillars. *Nano Lett.* **11**, 1743–1746 (2011)
11. Burek, M.J., Greer, J.R.: Fabrication and microstructure control of nanoscale mechanical testing specimens via electron beam lithography and electroplating. *Nano Lett.* **10**, 69–76 (2010)
12. Lee, G., Kim, J., Budiman, A.S., Tamura, N., Kunz, M., Chen, K., Burek, M.J., Greer, J.R., Tsui, T.Y.: Fabrication, structure and mechanical properties of indium nanopillars. *Acta Mater.* **58**, 1361–1368 (2010)

Nano-sized Particle

- ▶ [Gas Phase Nanoparticle Formation](#)

Nanoslits

- ▶ [Nanochannels for Nanofluidics: Fabrication Aspects](#)

Nanosoldering

- ▶ [Nanorobotic Spot Welding](#)

Nanospring

- ▶ [Nanorobotics for NEMS Using Helical Nanostructures](#)

Nanostructure

- ▶ [Structure and Stability of Protein Materials](#)

Nanostructure Field Effect Transistor Biosensors

Jason Li¹, Steve To², Lidan You¹ and Yu Sun³

¹Department of Mechanical and Industrial Engineering and Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

²Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

³Department of Mechanical and Industrial Engineering and Institute of Biomaterials and Biomedical Engineering and Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

Synonyms

[Bio-FET](#); [Chem-FET](#); [CNT biosensor](#); [CNT-FET](#); [DNA FET](#); [Nano-FET](#); [Nanowire FET biosensor](#)

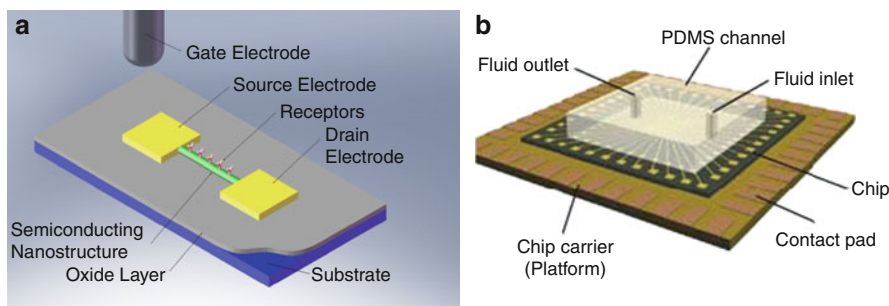
Definition

Nanostructure field effect transistor (nano-FET) biosensors are nanoscale electronic devices used to detect and measure the presence and concentration of specific biological molecules within a given sample solution. These sensors utilize a quasi-one-dimensional semi-conducting nanostructure as the active sensing element to detect the presence of local electric fields produced by charged biological molecules.

Overview

Biomolecular Detection Technologies

Methods for detecting and measuring the presence, abundance, and activity of biological molecules, such as proteins, DNA, and RNA, enable the investigation of basic biological function, molecular diagnosis of diseases, and development of therapeutic treatments. The combination of polyacrylamide gel-based electrophoresis separation (and related blotting techniques) and absorption-based chromogenic dye staining (e.g., Coomassie brilliant blue or silver staining) has traditionally formed the core technologies for protein and



Nanostructure Field Effect Transistor Biosensors, Fig. 1 (a) Schematic of a single nano-FET biosensor. (b) Schematic of a nano-FET biosensor chip with integrated microfluidic

sample delivery system. The nano-FET biosensor array is housed within the microfluidic system (Adapted with permission from Macmillan Publishers Ltd: [5], copyright 2006)

nucleic acid detection. Subsequent development of fluorescence- and luminescence-based labeling stains offered the opportunity for multicolor labeling, making multiplexed analysis possible. However, these methods were limited to the detection of biological molecules on gels and blots. Immuno-labeling of fluorescent and luminescent probes allowed specific protein detection and facilitated the emergence of immunoassays able to detect and quantify the abundance of a target protein molecule in solution. These optical detection methods suffer from large sample volume consumption, high reagent costs, and long assay times which, when combined with the drawback of using labeling molecules, preclude the application of these technologies to many laboratory and clinical applications (e.g., real-time protein quantification and point-of-care clinical diagnostics).

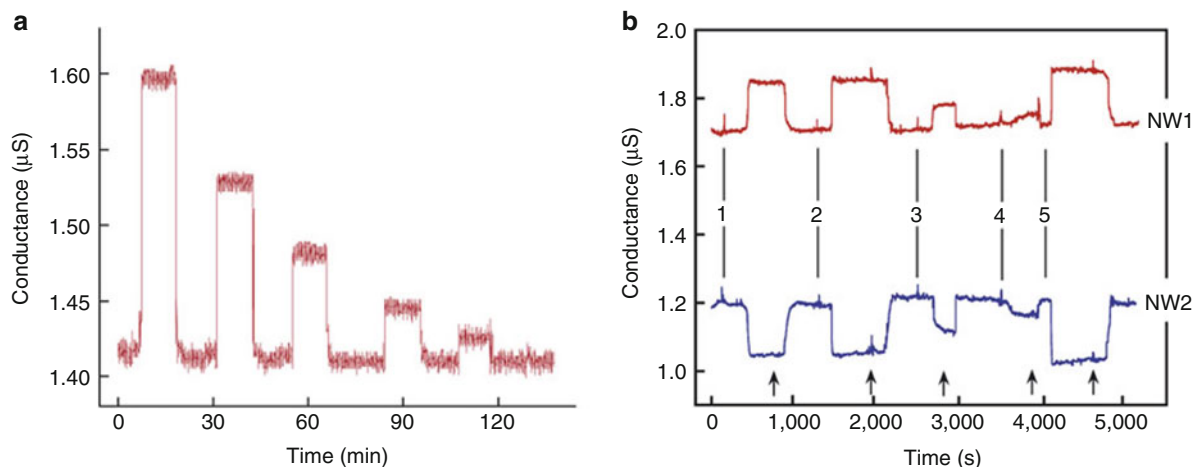
A number of novel label-free protein-detection strategies have since been developed in an effort to overcome some of the aforementioned limitations. Many of these strategies leverage nonoptical transduction modalities to circumvent the limitations imposed by optical probes. Examples include micro-/nano-surface plasmon resonance sensors [1], micro- and nano-cantilevers that translate biomolecular interactions into mechanical deformations [2], nanostructured field effect transistors (FETs) that measure intrinsic biomolecular charge [3], and electrochemical sensors that translate biomolecular adsorption to changes in redox current [4]. Of these technologies, nano-FET biosensors offer unique device characteristics including ultrasensitive, real-time, and label-free measurement capability, compact physical form,

CMOS-compatible and simple on-chip integration with circuitry and microfluidic systems, direct electrical readout, and multiplexed detection capability.

Structure

Nano-FET biosensors are composed of a semiconducting quasi-one-dimensional nanostructure (e.g., nanowire or nanotube) bridging the source and drain electrodes (Fig. 1). The nanostructure serves as the transistor's channel and functions as the sensing element of the device. The biosensor structure deviates slightly from that of the transistor in the placement and structure of the gating electrode. While the gate electrode in a nano-FET is typically located directly above the channel, separated by a thin dielectric layer, the nano-FET biosensor uses a solution gate electrode that is typically composed of platinum or Ag/AgCl and is immersed within the aqueous solution being measured. Furthermore, the nanostructure surface is functionalized with a layer of analyte-specific receptors to impart a degree of specificity to the sensor. This may be accomplished via chemical bonding using bifunctional linker molecules, or through physical interaction due to adsorption, van der Waals forces, and weak electrostatic forces.

While the specific nanostructure employed may vary in shape, size, and composition, it is important that these structures be quasi-one-dimensional and exhibit semiconductive material properties. Typical nanostructures used in the construction of nano-FETs include single and multi-walled carbon nanotubes, graphene, and silicon nanowires. Devices constructed from other semiconductive materials such as



Nanostructure Field Effect Transistor Biosensors, Fig. 2 Real-time nanowire FET sensing results. (a) Conductance versus time data recorded following alternate delivery of prostate-specific antigen (PSA) and pure buffer solution (1 μM phosphate (potassium salt) containing 2 μM KCl, pH 7.4). Subsequent PSA concentrations were 5 ng ml^{-1} , 0.9 ng ml^{-1} , 9 pg ml^{-1} , 0.9 pg ml^{-1} , and 90 fg ml^{-1} , respectively.

(b) Complementary sensing of PSA using p-type (NW1) and n-type (NW2) nanowire FET devices. Vertical lines correspond to addition of PSA solutions of (1, 2) 0.9 ng ml^{-1} , (3) 9 pg ml^{-1} , (4) 0.9 pg ml^{-1} , and (5) 5 ng ml^{-1} . Arrows on the bottom represent the injections of sensing buffer solution (Adapted with permission from Macmillan Publishers Ltd: [5], copyright 2006)

metal-oxides and nanostructures of various shapes including nanotubes, nanobelts, nanorods, nanoribbons, and nanobars have also been reported.

Physics

Nano-FET biosensors utilize semiconductive nanostructures to detect the presence of charged molecules in the immediate vicinity around the nanostructure surface. The sensing mechanism is based on the observation that an externally applied electric potential is capable of modulating the number of free charge carriers within a semiconductive material, thus leading to a measurable change in electrical conductance. For a cylindrical semiconducting nanostructure with radius R and length L , the charge carrier concentration may be described as a function of the applied potential:

$$n_0 = \frac{C(V_G - V_t)}{q\pi R^2 L} \quad (1)$$

where C is the gate capacitance, q is the elementary charge, V_G is the gate voltage, and V_t is the threshold voltage [6]. As the conductance of the semiconducting nanowire is described by:

$$G = nq\mu \left(\frac{\pi R^2}{L} \right) \quad (2)$$

the current passing through the nanowire (I_D) for a given source-drain voltage (V_{DS}) is therefore:

$$I_D = \frac{\pi R^2}{L} n_0 q \mu V_{DS} \quad (3)$$

where μ is the carrier mobility. Charged functional groups present on the surface of bound biological molecules such as proteins and nucleic acids provide the necessary electric potential required to induce a conductance change in the semiconductive nanostructure, thus enabling the direct electrical detection of such molecules using nano-FET devices (Fig. 2). The strength of a molecule's surface charge is a function of electrolyte pH and is characterized by the molecule's isoelectric point.

The simplified analysis qualitatively illustrates the physical sensing mechanism behind nano-FET biosensors. However, this treatment does not accurately describe the complexities of the nano-FET biosensor system. Equations (1–3) describe the electrical behavior of a cylindrical semiconducting crystal subjected to a uniform charge density over the entire nanostructure surface. This assumption fails to take into consideration the nonuniform nanostructure surface charge distribution resulting from discrete biomolecular binding events. The discrete gating mechanism by individual

molecules in solution introduces several additional phenomena that significantly influence the sensor's response including depletion of analyte from solution, analyte transport mechanisms to the nanostructure surface, and analyte-receptor binding and unbinding kinetics. These effects become particularly important in the detection of rare analytes.

Equation (3) also neglects the influence of electrostatic screening – the spatial arrangement of mobile ions around a charged molecule to effectively neutralize that charge. Importantly, target analytes must be sufficiently close to the sensor surface in order to elicit a conductance change due to electrostatic screening by mobile charges around the analyte molecule (e.g., ions in solution and charge carriers in the semiconducting crystal). The Debye screening length, the distance away from a charged molecule at which the electrostatic potential is diminished to zero due to screening, defines the required proximity.

Debye length (aqueous electrolyte):

$$k^{-1} = \frac{1}{\sqrt{8\pi\lambda_B N_A I}} \quad (4)$$

where I is the ionic strength of the electrolyte (mole/ m^3), N_A is the Avogadro number, and λ_B is the Bjerrum length of the medium (0.7 nm for water).

Debye length (silicon):

$$L_D = \sqrt{\frac{\epsilon_0 \epsilon_r k_B T}{q^2 N_D}} \quad (5)$$

where ϵ_r is the dielectric constant of the semiconducting material, k_B is the Boltzmann's constant, T is the absolute temperature in Kelvin, q is the elementary charge, and N_D is the density of donors in a substrate. The implications of charge screening and a finite Debye length make several important predictions on nano-FET biosensing performance. Namely, these relations suggest a conductance dependence on charge carrier concentrations (ionic strength of an electrolyte solution and doping density of a semiconductor), size of the semiconducting crystal, strength of the molecule's net surface charge, and the nanostructure-analyte separation distance. Inclusion of these phenomena into a unified analytical model that accurately describes experimental nano-FET biosensor response is an active area of research.

Performance Parameters

Nano-FET Sensitivity

Nano-FET sensitivity is the induced change in device conductance (ΔG_{sd}), upon exposure to a certain biomolecular stimulus for a constant source-drain voltage (V_{ds}). Sensitivity is defined as:

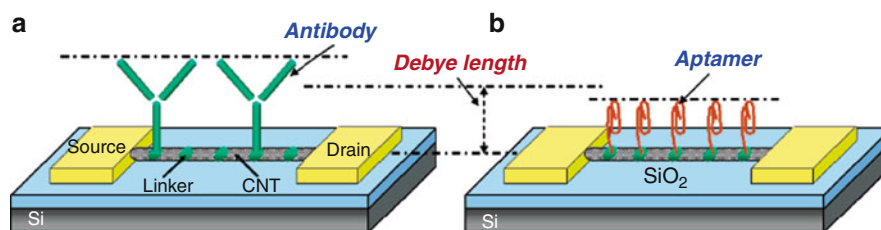
$$\text{Sensitivity} = \frac{\Delta G_{sd}}{G_0} \quad (6)$$

where ΔG_{sd} is the direct conductance change observed in sensing experiments and G_0 is the initial device conductance [7]. Because ΔG_{sd} depends on the specific nanowire parameters (e.g., diameter, mobility, etc.) this value does not reflect the intrinsic sensitivity. Instead, it is more meaningful to characterize the sensitivity using a dimensionless parameter (6). Normalization enables comparison across devices with different physical dimensions.

Sensitivity is closely related to three additional parameters: *detection limit* (the smallest analyte concentration that can be measured, which is dictated by the system signal-to-noise ratio); *measurement resolution* (the smallest detectable change in analyte concentration, which is dictated by noise levels); and *measurement range* (the measurable analyte concentrations, with the upper and lower bounds defined by nanostructure saturation and the detection limit, respectively). A higher sensitivity directly translates into lower detection limit, finer measurement resolution, and a broader measurement range. Typical protein and DNA detection limits reported in literature are in the pM (picomolar) to fM (femtomolar) range [8, 9]. Single virus detection has also been achieved [10].

Specificity

Specificity is a measure of a biosensors relative responsiveness to the target analyte as compared to all other molecules present within the sample solution. Very specific sensors should possess high sensitivity toward the target analyte and low sensitivity toward all other biomolecules. This attribute is imparted onto a nano-FET biosensor by functionalizing the nanowire surface (either covalently or noncovalently) with a confluent monolayer of analyte-specific receptors/ligands that bind specifically to the molecule of interest.



Nanostructure Field Effect Transistor Biosensors, Fig. 3 Schematic representation of nano-FET biosensors functionalized with (a) antibodies and (b) aptamers. Aptamers (and antibody fragments) are smaller than whole antibodies. The use of these capture probes thus results in increased nano-FET

sensitivity. For certain experimental conditions, aptamers may serve as the ideal probe candidate as they are shorter than the system Debye length while antibodies are not (Reprinted with permission from [14]. Copyright 2007 American Chemical Society)

Response Time

Response time is the time it takes for the biosensor system to obtain a stable biomolecular concentration measurement upon introduction of the sample solution to the sensor surface. As with any dynamic system, this characteristic is captured in the step response of the system by the *settling time*. For a nano-FET biosensor, this parameter is influenced by the analyte-receptor binding and unbinding kinetics, transport of analyte molecules to the sensor surface, concentration and degree of mixing of the analyte solution, and the electrical response behavior of the measurement system. The typical response time of nano-FET biosensors is on the order of minutes.

Key Research Findings

Active areas of research in the field of nano-FET biosensing include: (1) optimization of biosensor performance based on the systematic examination of how fundamental device parameters affect sensitivity, specificity, and response time; (2) development of robust large-scale fabrication strategies for the manufacturing of biosensor arrays; and (3) development of an analytical model capable of accurately describing and predicting nano-FET biosensor performance.

Performance Optimization

Minimizing the Effects of Charge Screening

As the effect of charge screening acts to minimize the total nanostructure volume gated by surface charges, thus reducing device sensitivity, strategies that increase screening length are expected to improve device sensitivity. The maximum device sensitivity

may be obtained in a situation where the effective screening length is much larger than the radius of the nanowire. Improved device sensitivity can thus be obtained through careful design and optimization of several device parameters that influence the effective screening length:

Diameter Reduction of the nanostructure diameter will dramatically increase the surface-to-volume ratio, thus improving device sensitivity. This is the primary advantage of decreasing the channel size and explains why nanostructure FET sensors exhibit significantly higher performance as compared to traditional planar FET devices.

Dielectric Thickness Silicon nanowires form a native oxide layer ($\sim 1\text{--}2$ nm) on their surface that serves as a dielectric layer reducing the electric field strength acting on the nanowire itself. This layer also increases the separation distance between surface-bound molecules and the semiconductive core of the nanowire. Thus, a thinner oxide layer is expected to improve sensitivity. Complete removal of the native oxide layer surrounding silicon nanowires has been shown to increase device sensitivity [11, 12].

Functionalization Scheme and Receptor Size

Decreasing the separation distance between the receptor-bound analyte molecule and the nanostructure surface minimizes the effect of electrolyte screening. While antibodies are commonly used to selectively bind analyte molecules to the biosensor surface, the use of smaller capture probes such as antibody fragments or aptamers in place of antibodies can increase device sensitivity without the loss of selectivity (Fig. 3) [13, 14]. Similarly, utilizing shorter

bifunctional linker molecules to attach the capture probes to the nanostructure surface may also increase device sensitivity. The choice of functionalization strategy may also dictate the capture probe density and coverage on the nanostructure surface. A poor capture probe density can result in the nonspecific adsorption of molecules onto unmodified areas of the nanowire surface and may also limit the upper detection limit of the sensor due to capture probe saturation.

Electrolyte Ion Concentration and pH The ionic strength of the electrolyte solution influences the efficiency of charge screening as the density of free ions is directly related to size of the electric double layer formed around a charged molecule. Higher ionic strength electrolytes are thus able to screen a charged molecule over a shorter distance. It is therefore advantageous to perform biosensing experiments in low-ion conditions to maximize sensitivity [15]. However, this is not always feasible as many sensing applications are performed with physiological samples (e.g., serum, whole blood, etc.).

The pH of a solution also influences device sensitivity by changing the effective electrical charge on a biological molecule. Biological molecules such as proteins and nucleic acids contain both acidic and basic functional groups that may be positively or negatively charged depending on the availability of protons (H^+) present in solution. The pH for which the negative and positive charges on the molecule are balanced, resulting in zero net charge, is defined as the isoelectric point (pI). The net charge on a molecule is therefore a function of the electrolyte pH and can become increasingly more positively or negatively charged as the term $|pI-pH|$ increases. While tuning this parameter may be useful in specific applications, the pH of a sample is usually fixed at physiological levels ($pH = 7.4$).

Carrier Concentration The screening length within a semiconductive nanowire dictates the depth of penetration into the nanostructure for which surface-bound charged molecules are able to gate. For example, the screening length in a silicon nanowire operating with a typical carrier concentration of 10^{18} – 10^{19} cm^{-3} is ~ 1 – 2 nm, suggesting that only a small portion of the total nanowire volume (a ~ 1 – 2 -nm thick layer at the nanowire surface) is

affected by surface charges. Decreasing the charge carrier concentration will therefore increase device sensitivity by increasing the screening length and the total gated volume. This can be achieved by varying either of two device parameters: the nanostructure doping density or the bias voltage applied through a gate electrode on the nano-FET biosensor. In terms of semiconductors, lower doping densities result in longer screening lengths and therefore produce higher sensitivity devices [16]. It is instructive to note that the optimum biosensor sensitivity is obtained when the device is operating in the subthreshold regime, as under these conditions, the transconductance is at a maximum [7, 17, 18].

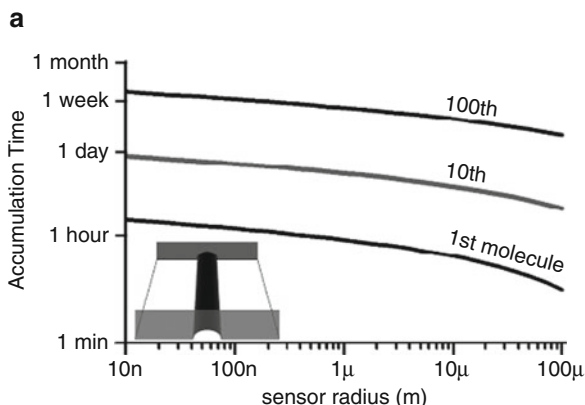
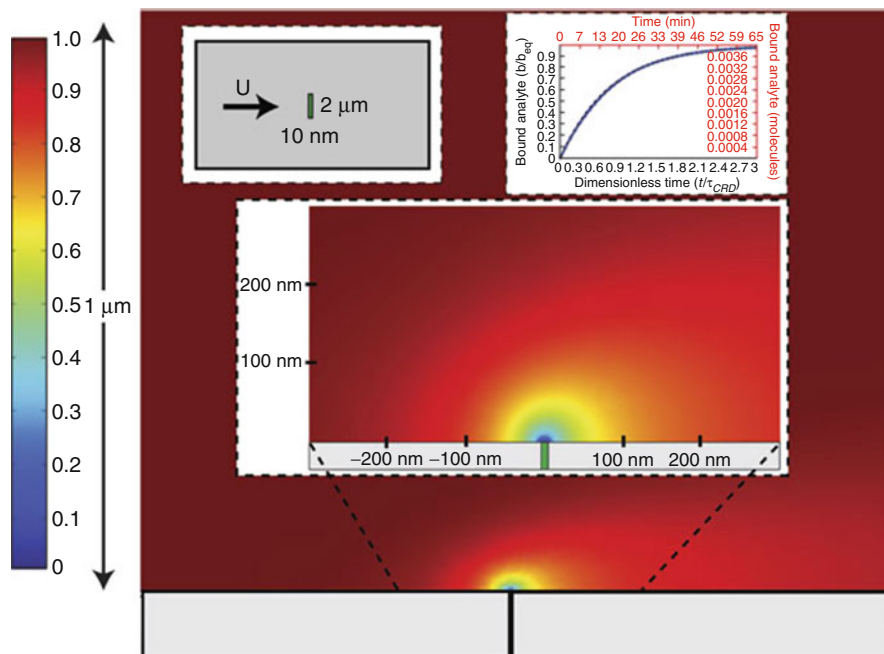
Optimization of Analyte Delivery Efficiency

Biomolecular detection using nano-FET sensors requires the delivery of analyte to the sensor surface. The efficiency of this process can therefore limit the performance of nano-FET biosensors. Several phenomena that influence this process include diffusive and convective molecular transport mechanisms of analyte to the biosensor surface, depletion of free analyte from solution, and nonspecific analyte adsorption and binding.

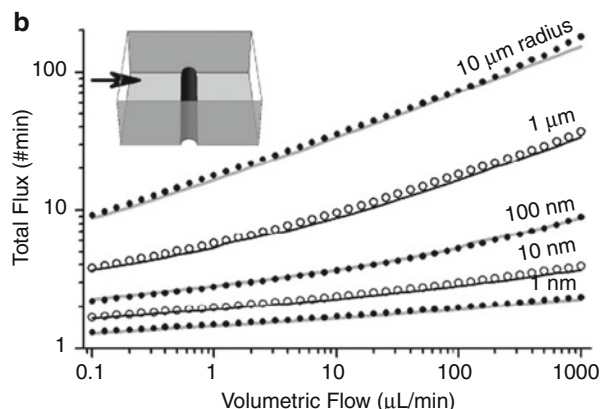
Molecular Transport As with many surface-based biosensor systems, the transport of analyte in solution to the sensor surface plays a crucial role in governing binding kinetics and ultimately sensor performance. This is especially true when analyzing samples with dilute concentrations of target molecules and/or employing microfluidic systems for efficient and automated handling of small sample solution volumes. A number of competing physical processes influence target transport to the sensor surface. Analyte molecules suspended in solution may diffuse randomly within the solution, be convected along with flowing fluid, bind to adjacent surface-bound receptors, or subsequently unbind to reenter solution. The binding (or collection) of biomolecules onto the sensor surface simultaneously depletes the surrounding solution to form a so-called depletion region around the sensor (Fig. 4). As the depletion region grows in size, the analyte diffusive flux decreases and thus the collection rate becomes slower. Through finite element analysis, Sheehan and Whitman determined that the size and shape of the sensor profoundly affects the total analyte flux to the sensor surface (Fig. 5a) and further argues

Nanostructure Field Effect Transistor Biosensors,

Fig. 4 The steady-state depletion zone around a nanowire sensor from a pure mass-transport problem. The depletion zone is thick compared to the nanowire, but is substantially thinner than the channel itself ($\delta_s \sim LPe_s^{-1/2} = 100$ nm). Reaction-limited binding onto the nanowire follows the Langmuir binding curve (inset *top right*) (Adapted with permission from Macmillan Publishers Ltd. Nature Biotechnology [20], copyright 2006)



Nanostructure Field Effect Transistor Biosensors,
Fig. 5 (a) Theoretical calculation of the time required for a 10- μ m-long hemicylindrical sensor to accumulate 1, 10, and 100 molecules due to diffusive molecular transport. The sensor lies at the *bottom* of a channel whose width is equal to the sensor's length and which is filled with a 1 fM analyte solution.



(b) Finite element analysis (*points*) and theoretical calculation (*lines*) of the total flux of molecules to a hemicylindrical sensor in a microchannel (800 μ m wide, 100 μ m high). The total flux plotted is the steady-state value at the given volumetric flow rate for a 1 fM analyte concentration (Adapted with permission from [19]. Copyright 2005 American Chemical Society)

that without directed transport of biomolecules, individual nanoscale sensors would be limited to picomolar order sensitivity for practical time scales (hours to days) [19].

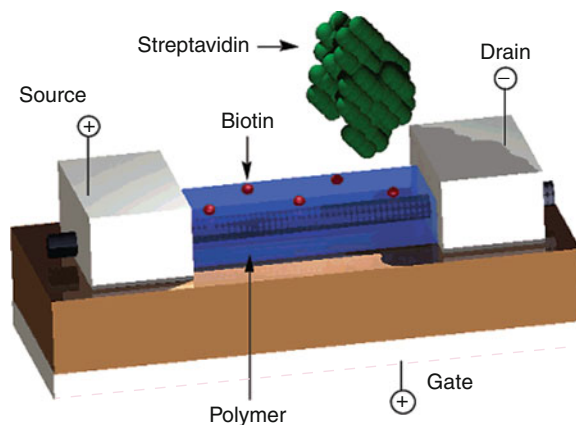
While the diffusive depletion zone grows indefinitely, albeit at an ever-decreasing rate, introduction of convective flow (assumed to be laminar) halts this

growth resulting in a steady-state depletion zone with a length scale defined by the balance between the convective analyte flux delivered to the depletion zone boundary and the diffusive flux out of the depletion zone. By adjusting the convective flow rate through the channel, it is possible to tune the size of the depletion layer and thus control the rate at which

molecules are delivered to the sensor surface (Fig. 5b). However, Squires et al. argue that the total mass transport varies only weakly with the flow rate as flow rates must be increased 1,000-fold to enhance flux by a factor of 10 [20]. The treatment thus far has assumed instantaneous analyte-receptor coupling upon transport of target molecules to the sensor surface. This represents one of two sensing regimes that can be described as being “transport-limited” as the availability of analyte adjacent to the sensor surface limits sensor response.

A second biosensing regime arises when considering the kinetics of analyte-receptor association and dissociation (assumed to be first-order Langmuir kinetics) [20]. If delivery of target molecules to the sensor surface occurs at a faster rate than the net analyte-receptor association rate, then the transport is said to be “reaction limited.” In general, the reaction kinetics is determined by the fidelity of the immobilized reagents. While each of the two regimes imposes an independent maximum target collection rate, the “reaction-limited” condition is most desirable for nano-FET operation as, for a given device, biomolecular sensing is purely a function of analyte concentration. The sensor size and flow rate should thus be adjusted to ensure nano-FET operation within this regime. Alternatively, introduction of turbulent flow or fluid mixing through innovative microfluidic design may also be employed to overcome the limitations imposed by convective and diffusive molecular transport [8].

Depletion of Free Analyte As silicon and metal-oxide nanostructures share similar surface chemistries with several popular microfabrication substrates (e.g., silicon wafers, glass, quartz), surface modification techniques used to attach receptor molecules to the nanostructure surface can also functionalize the device substrate. In this situation, the total device surface area exposed to the sample solution during biomolecular detection experiments is proportional to the number of receptors available for reaction. Analyte binding events with substrate-bound receptors do not induce a change in nanostructure conductance. Rather, these reactions serve only to deplete free analyte molecules from solution, thus reducing the number of analyte molecules available for reaction with nanostructure-bound receptors. Minimizing the total surface area exposed to sample solutions is, therefore, important



Nanostructure Field Effect Transistor Biosensors, Fig. 6 Schematic of a nanotube field effect transistor coated with a polymeric functional layer. Co-functionalization with a molecular receptor (biotin) and a nonreactive polymer (e.g., polyethylene glycol or polyethylene oxide) prevents nonspecific adsorption of molecules onto the nanotube surface (Adapted with permission from [22]. Copyright 2003 American Chemical Society)

for maximizing device sensitivity. This effect may be circumvented using nanostructure-specific surface modification chemistries.

Nonspecific Adsorption Random adsorption of biological molecules onto the nano-FET sensor produces a false-positive measurement signal. This effect may be minimized by grafting a dense layer of long nonreactive polymer chains (e.g., polyethylene glycol) onto the nanostructure surface alongside receptor molecules (Fig. 6). The polymer chains prevent molecular adsorption directly onto the nanostructure surface by acting as a spacer while the co-functionalized receptors are able to bring analyte molecules into intimate contact with the nanostructure through molecular binding. The length of the polymer chains used for this application should exceed the screening length of the system.

Nonspecific adsorption of random molecules is especially problematic when working with samples containing numerous types of biological molecules such as whole blood. One strategy that enables detection of a specific analyte within a whole blood sample involves the use of a sample pre-purification procedure before nano-FET measurement [21]. In this method, the whole blood is introduced into a chamber in which analyte-specific receptors have been immobilized on to the chamber walls. Upon binding of free analytes to the

surface-bound receptors, the whole blood solution is flushed out of the chamber and discarded. The receptors are then detached from the chamber surfaces and transported to the nano-FET biosensor for analysis.

Large-Scale Methods of Fabrication

Reliable large-scale fabrication of nano-FET biosensors remains to be a significant technological challenge, impeding the adoption of this technology in medical diagnostic and biomolecular detection applications. Unlike traditional microfabrication methods for producing micrometer-level structures, the difficulties encountered in nano-FET biosensor fabrication arise from the process of establishing electrical contact with nano-sized structures in a high-throughput and reproducible manner.

Early methods for manufacturing nano-FETs involve (1) the growth of individual nanostructures using a chemical vapor deposition process, (2) transfer of as-grown nanostructures from the growth substrate onto a device substrate (a process which results in the random placement and orientation of the deposited nanostructures), and (3) identification of a suitable nanostructure on the device substrate followed by the patterning of metallic electrodes onto the ends of that nanostructure using a mask-less nano-lithographic tool (e.g., electron beam lithography or focused ion beam lithography), which has a low throughput and is only suitable for proof-of-principle use. Recently devised batch-fabrication strategies for the wafer-scale production of nano-FET arrays can be broadly categorized into “bottom-up” or “top-down” fabrication methods.

Bottom-Up Fabrication Methods

One strategy for bottom-up fabrication is to make use of pre-synthesized freestanding nanostructures using epitaxial growth methods such as chemical vapor deposition. This strategy affords several advantages over top-down and other bottom-up approaches including (1) strict control over nanostructure dimensions and electrical properties, (2) the ability to assemble nano-FET devices onto flexible and transparent substrates, and (3) the ability to incorporate nanostructures with unique chemical compositions and architecture into nano-FET devices. Successful incorporation of these nanostructures into nano-FET devices relies primarily on the ability to transfer, align, and position the as-grown nanostructures onto a device

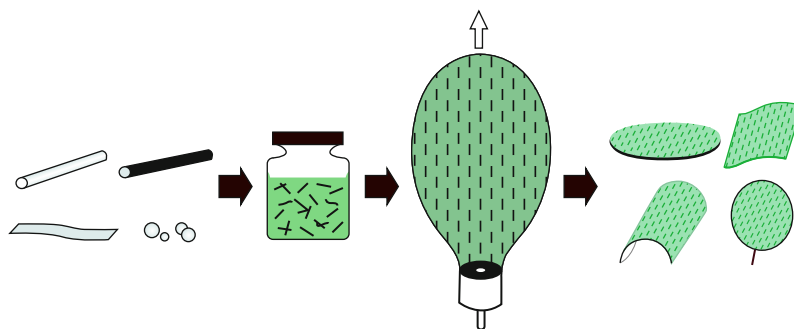
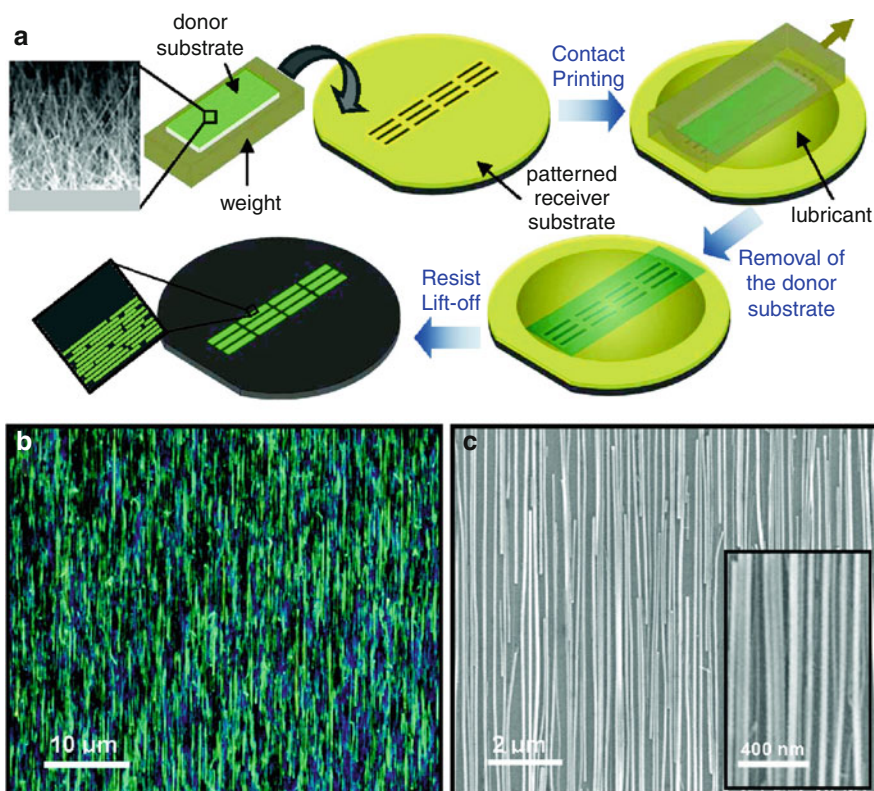
substrate. Contacting electrodes may then be patterned using conventional photolithography to form individual sensors or sensor arrays.

Flow Alignment Flow alignment is a method that makes use of flowing fluid to align suspended nanostructures along a single orientation in the direction of flow [23]. By passing a nanostructure suspension through a microfluidic structure formed between a poly(dimethylsiloxane) mold and the device substrate, free-flowing nanostructures assemble onto the device substrate due to surface forces and remain relatively aligned parallel to the direction of flow. The resulting degree of alignment is a function of the flow rate as increased flow velocities result in a narrower angular distribution among deposited nanostructures. Furthermore, varying the total flow duration dictates the nanostructure deposition density, with longer flow durations favoring higher nanostructure densities. This method is useful in controlling the degree of nanostructure alignment and the deposition density, thus allowing subsequent electrode formation. Nano-FET fabrication can therefore be achieved by patterning a pair of electrodes parallel to the flow direction using conventional microfabrication processes. For this physical arrangement, minimizing the electrode pair separation distance increases the likelihood of a bridging nanostructure between the electrode pair; however, the exact number of bridging nanostructures resulting from this process is probabilistic.

Nanostructure Contact Printing Nanostructure contact printing is a method that provides similar results to the flow alignment strategy. This method, however, relies on mechanical shear forces to detach freestanding nanostructures from the growth substrate, and friction to align the deposited nanowires onto the device substrate [24]. In this method, the growth substrate is inverted and placed on top of the device substrate such that the nanostructures are sandwiched in between (Fig. 7). Weights are added on top of the growth substrate to control the applied pressure between the two substrates in order to optimize the deposited nanostructure density. The growth substrate is then pulled over the device substrate at a continuous velocity in order to deposit and align the nanostructures onto the device substrate. The shear velocity and the amount of friction between the two

Nanostructure Field Effect Transistor Biosensors,

Fig. 7 Contact printing of nanowires. (a) Schematic of the contact printing process for producing well-aligned nanowire arrays. (b) Dark-field optical and (c) SEM images of 30-nm nanowires printed on a Si/SiO₂ substrate showing highly dense and aligned monolayer of nanowires. The self-limiting process limits the transfer of nanowires to a single layer, without significant nanowire bundling (Adapted with permission from [24]. Copyright 2008 American Chemical Society)



Nanostructure Field Effect Transistor Biosensors,
Fig. 8 Blown bubble film approach. Nanostructures are dispersed in a polymer solution. A volume of solution is expanded as a bubble using a die to produce well-aligned nanostructures

suspended within a thin film. The film is then contacted with a substrate to deposit well-aligned nanostructures onto the substrate ([25] Reproduced with permission of The Royal Society of Chemistry)

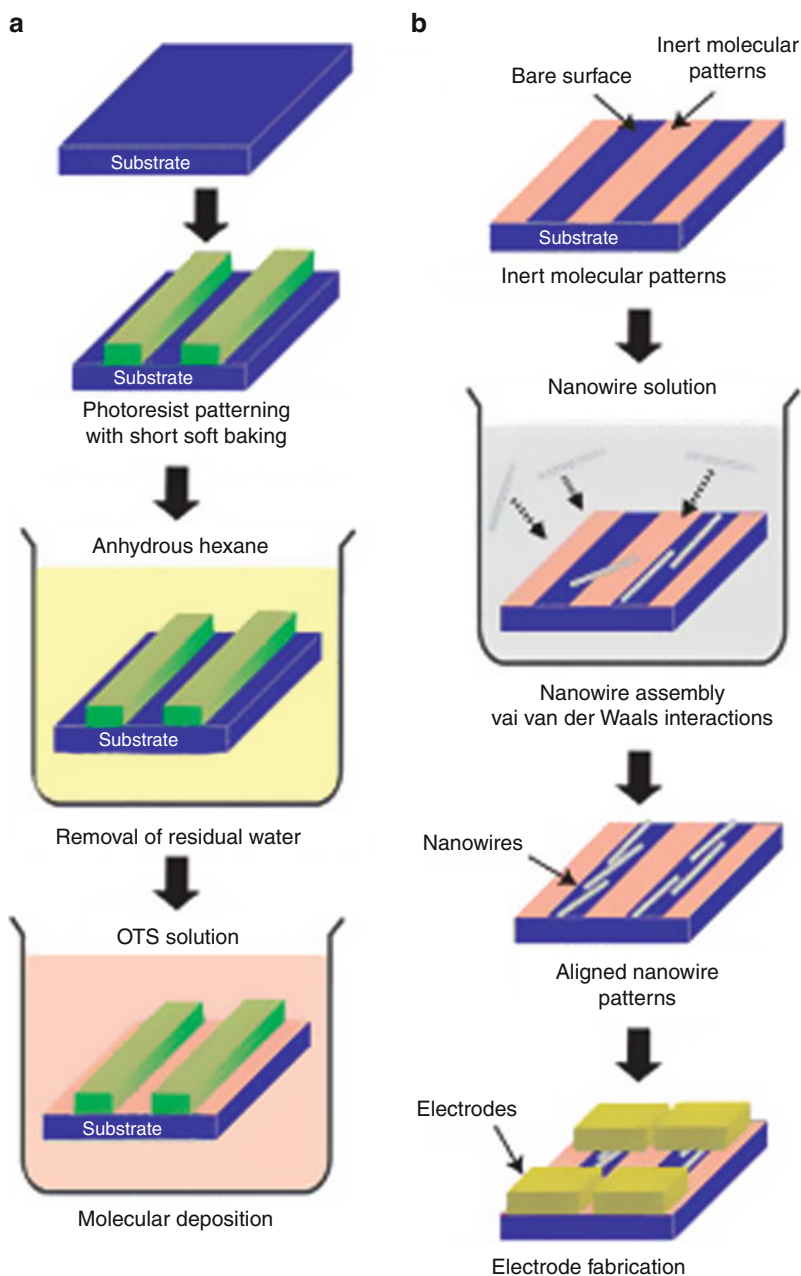
substrates influence the resulting degree of nanostructure alignment. The latter may be adjusted with the use of a lubricant.

Thin-Film Nanowire Suspension Thin-film nanowire suspension is a method that makes use of the observation that nanostructures suspended within a liquid thin film will align along a single direction

when that film is elongated. Therefore, this method involves the formation of a large balloon from a nanostructure suspension (Fig. 8) [25]. The aligned nanostructures within the thin film can then be transferred to a substrate by placing the substrate in contact with the balloon. Nano-FET fabrication would subsequently proceed as in the case of flow alignment.

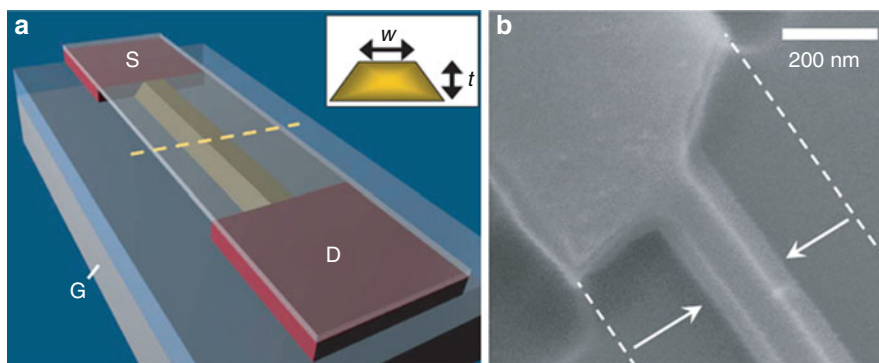
Nanostructure Field Effect Transistor Biosensors,

Fig. 9 Linker-free directed self-assembly. (a) The patterning process of octadecyltrichlorosilane self-assembled monolayer on a solid substrate via photolithography. (b) Self-assembly suspended nanostructures onto the bare surface regions on the substrate. Subsequent photolithography and metallization can be used to pattern electrodes onto the ends of the deposited nanostructures (Adapted with permission from Macmillan Publishers Ltd. Nature Nanotechnology [26]: copyright 2006)



Directed Self-Assembly Directed self-assembly relies on the minimization of surface energy to assemble suspended nanostructures into predefined locations on a substrate [26]. In this method, surface modification strategies are used to render certain portions of the device substrate either hydrophobic or hydrophilic (Fig. 9). Upon immersion of the substrate into a nanostructure suspension, individual nanostructures will

self-assemble onto the hydrophobic portions of the substrate. This strategy can be used to deposit individual nanostructures at specific locations over a large area on a substrate. A similar strategy using analyte-ligand interactions has also been employed in which suspended nanostructures are functionalized with biotin molecules. Upon immersion of a substrate with bound streptavidin molecules at predefined

**Nanostructure Field Effect Transistor Biosensors, Fig. 10**

(a) Schematic of nano-FET device after anisotropic etching. The silicon-on-insulator active channel (yellow, width w and thickness t) is undercut etched, whereas degenerate leads (red) are

etch-resistant. The source (S), drain (D), and underlying backgate (G) are labeled. (b) Scanning electron micrograph of a complete device (Adapted with permission from Macmillan Publishers Ltd: Nature [26], copyright 2007)

locations, biotin-streptavidin conjugation forces the nanostructures to assemble onto the device substrate at predefined locations.

Dielectrophoresis Dielectrophoresis (DEP) utilizes electrical fields to manipulate nanostructures. In this method, metallic contact electrodes are patterned on the device substrate prior to the assembly of nanostructures onto the device. By applying a biased alternating voltage across the electrodes, a local nonuniform electric field is produced, exerting a force on suspended semiconducting nanostructures. This force causes the nanostructures to assemble across the two electrodes. This process is easily amenable to the large-scale production of nano-FET arrays and can be used to fabricate single nanostructure FET devices; however, typical DEP trapping processes yield networked multi-nanostructure FET devices [27]. This process is highly susceptible to the experimental conditions including the size, shape, and properties of the nanostructure to be manipulated, and parameters of the electrical signals, as well as the electrical properties of the surrounding medium.

In Situ Growth of Nanostructures In situ growth of nanostructures is a method that utilizes patterned catalysts to selectively synthesize nanostructures at certain regions on the device substrate. These catalysts (e.g., iron or gold particles) are required to initiate and/or propagate nanostructure growth. In some cases, the direction of nanostructure growth can be influenced with the application of an electric field, therefore enabling control of the

resulting nanostructure orientation. This method usually produces numerous nanostructures and is, therefore, useful in fabricating nano-FET sensors that utilize a network of nanostructures.

Top-Down Fabrication Method

Anisotropic Lateral Wet Etching Top-down methods are typically based on the anisotropic lateral wet etching of nanometer-thin SOI (silicon-on-insulator) wafers to produce well-defined nanostructures from the device layer (Fig. 10) [8]. Micro- or nano-sized etch masks are patterned via conventional or e-beam photolithography and anisotropically time-etched to produce nanometer-wide nanostructures. Source and drain electrodes are degenerately doped, rendering them conductive and unaffected by the etchant. This approach enables wafer-scale formation of semiconducting nanostructures at precise locations on a wafer-scale, making subsequent electrode fabrication relatively straightforward using conventional photolithography. As this method relies on time-controlled anisotropic etching for the removal of material, it is highly susceptible to small changes in processing conditions such as etch time, processing temperature and mixing, and device-layer doping density and crystal orientation.

Future Directions

At this point in time, the development of nano-FET biosensors is still in the proof-of-concept phase. The promise of highly sensitive, label-free electrical

sensors with real-time measurement capability remains attractive for numerous medical and basic science research applications. While great progress has been made in nano-FET fabrication and detection, continued research into large-scale fabrication methods for batch-manufacturing nano-FET sensors will be essential for the ultimate commercial success of this technology and its application to clinical and research applications. In addition to low cost and high yield, these fabrication methods must produce devices with consistent sensing performance across the nano-FET sensors such that measurements made with these devices can be standardized. Another critical hurdle that must be overcome is the issue of poor analyte specificity and nonspecific binding. While some strategies have been put forth to remedy these issues, these solutions must be further improved upon before nano-FET biosensors are able to reliably analyze whole blood, serum, and other specimens. Lastly, efforts in the area of nano-FET biosensor integration with microfluidic and lab-on-chip devices will facilitate point-of-care diagnostic applications and real-time closed-loop drug delivery systems. Successful lab-on-chip integration will also enable the creation of multiplexed nano-FET biosensor arrays, which could hold great promise in the areas of basic chemical and biological research, high-throughput screening systems for drug development, and novel in vitro biology experimentation. The future will likely see many point-of-care biosensors using electrical-based detection, and nano-FET biosensing technologies can possibly play an important role.

Cross-References

- ▶ [Biosensors](#)
- ▶ [Carbon Nanotubes](#)
- ▶ [Chemical Vapor Deposition \(CVD\)](#)
- ▶ [Dielectrophoresis](#)
- ▶ [Dielectrophoretic Nanoassembly of Nanotubes onto Nanoelectrodes](#)
- ▶ [Electric-Field-Assisted Deterministic Nanowire Assembly](#)
- ▶ [Functionalization of Carbon Nanotubes](#)
- ▶ [Nanophotonic Structures for Biosensing](#)
- ▶ [Nanostructured Materials for Sensing](#)
- ▶ [Nanotechnology](#)
- ▶ [Self-assembly of Nanostructures](#)

References

1. Lee, S.W., et al.: Highly sensitive biosensing using arrays of plasmonic Au nanodisks realized by nanoimprint lithography. *ACS Nano* **5**, 897–904 (2011)
2. Fritz, J.: Cantilever biosensors. *Analyst* **133**, 855–863 (2008)
3. Cui, Y., Wei, Q., Park, H., Lieber, C.M.: Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science* **293**, 1289–1292 (2001)
4. Soleymani, L., Fang, Z., Sargent, E.H., Kelley, S.O.: Programming the detection limits of biosensors through controlled nanostructuring. *Nat. Nanotechnol.* **4**, 844–848 (2009)
5. Patolsky, F., Zheng, G., Lieber, C.M.: Fabrication of silicon nanowire devices for ultrasensitive, label-free, real-time detection of biological and chemical species. *Nat. Protoc.* **1**, 1711–1724 (2006)
6. Fan, Z., Lu, J.G.: Gate-refreshable nanowire chemical sensors. *Appl. Phys. Lett.* **86**, 123510–123513 (2005)
7. Gao, X.P., Zheng, G., Lieber, C.M.: Subthreshold regime has the optimal sensitivity for nanowire FET biosensors. *Nano Lett.* **10**, 547–552 (2010)
8. Stern, E., et al.: Label-free immunodetection with CMOS-compatible semiconducting nanowires. *Nature* **445**, 519–522 (2007)
9. Zheng, G., Patolsky, F., Cui, Y., Wang, W.U., Lieber, C.M.: Multiplexed electrical detection of cancer markers with nanowire sensor arrays. *Nat. Biotechnol.* **23**, 1294–1301 (2005)
10. Patolsky, F., et al.: Electrical detection of single viruses. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14017–14022 (2004)
11. Bunimovich, Y.L., et al.: Quantitative real-time measurements of DNA hybridization with alkylated nonoxidized silicon nanowires in electrolyte solution. *J. Am. Chem. Soc.* **128**, 16323–16331 (2006)
12. Zhang, G.J., et al.: Highly sensitive measurements of PNA-DNA hybridization using oxide-etched silicon nanowire biosensors. *Biosens. Bioelectron.* **23**, 1701–1707 (2008)
13. Zhang, G.J., et al.: DNA sensing by silicon nanowire: charge layer distance dependence. *Nano Lett.* **8**, 1066–1070 (2008)
14. Maehashi, K., et al.: Label-free protein biosensor based on aptamer-modified carbon nanotube field-effect transistors. *Anal. Chem.* **79**, 782–787 (2007)
15. Stern, E., et al.: Importance of the Debye screening length on nanowire field effect transistor sensors. *Nano Lett.* **7**, 3405–3409 (2007)
16. Nair, P.R., Alam, M.A.: Design considerations of silicon nanowire biosensors. *Electron Devices, IEEE Trans.* **54**, 3400–3408 (2007)
17. Heller, I., Mannik, J., Lemay, S.G., Dekker, C.: Optimizing the signal-to-noise ratio for biosensing with carbon nanotube transistors. *Nano Lett.* **9**, 377–382 (2009)
18. Lu, M.P., Hsiao, C.Y., Lai, W.T., Yang, Y.S.: Probing the sensitivity of nanowire-based biosensors using liquid-gating. *Nanotechnology* **21**, 425505 (2010)
19. Sheehan, P.E., Whitman, L.J.: Detection limits for nano-scale biosensors. *Nano Lett.* **5**, 803–807 (2005)

20. Squires, T.M., Messinger, R.J., Manalis, S.R.: Making it stick: convection, reaction and diffusion in surface-based biosensors. *Nat. Biotechnol.* **26**, 417–426 (2008)
21. Stern, E., et al.: Label-free biomarker detection from whole blood. *Nat. Nanotechnol.* **5**, 138–142 (2010)
22. Star, A., Garbiel, J.P., Bradley, K., Gruner, G.: Electronic detection of specific protein binding using nanotube FET devices. *Nano Lett.* **3**, 459–463 (2003)
23. Huang, Y., Duan, X., Wei, Q., Lieber, C.M.: Directed assembly of one-dimensional nanostructures into functional networks. *Science* **291**, 630–633 (2001)
24. Fan, Z., et al.: Wafer-scale assembly of highly ordered semiconductor nanowire arrays by contact printing. *Nano Lett.* **8**, 20–25 (2008)
25. Yu, G., Li, X., Lieber, C.M., Cao, A.: Nanomaterial-incorporated blown bubble film for large-area aligned nanostructures. *J. Mater. Chem.* **18**, 728–734 (2008)
26. Lee, M., et al.: Linker-free directed assembly of high-performance integrated devices based on nanotubes and nanowires. *Nat. Nanotechnol.* **1**, 66–71 (2006)
27. Vijayaraghavan, A., et al.: Ultra-large-scale directed assembly of single-walled carbon nanotube devices. *Nano Lett.* **7**, 1556–1560 (2007)

Nanostructured Antireflective Surfaces Arrays

► Moth-Eye Antireflective Structures

Nanostructured Functionalized Surfaces

Lorenzo Lunelli¹, Cristina Potrich¹,
Laura Pasquardini² and Cecilia Pederzoli²
¹Biofunctional Surfaces and Interfaces, FBK-CMM
Bruno Kessler Foundation and CNR-IBF, Povo,
TN, Italy
²Biofunctional Surfaces and Interfaces, FBK-CMM
Bruno Kessler Foundation, Povo, TN, Italy

Synonyms

Functionalized nanomaterials

Definition

Nanostructured functionalized surfaces are materials that contain microscale and nanoscale features produced in a well-controlled manner with nanocoating,

film deposition, or in situ surface modifications at nanoscale level, with the aim to improve the material performances. The potential applications of nanostructured functionalized surfaces are wide; some examples are in vivo imaging and diagnostics (nanosized contrast agents for medical imaging, miniaturized devices); in vitro miniaturized diagnostics (lab-on-a-chip, point-of-care measurements); drug development, delivery, and therapy (nanoparticles, nanobody, affibodies); tissue engineering; and implants of drug dispenser/factory.

Overview

Nanostructures or nanomaterials have peculiar physical and/or chemical properties, different from their bulk material. Main differences rely on topography or chemistry or on both aspects of the outer layer of the material which becomes “nanostructured” when these topographical and/or chemical modifications are made at a nanometric level. Nanostructured surfaces can be produced by surface modification of materials conventionally interacting with biological systems such as human body (biomaterials). Nanofunctionalization techniques can be divided in two main categories: (a) nanocoating and film deposition and (b) in situ surface nanofunctionalization. These two types of techniques are often combined to produce surfaces with hybrid nanostructures such as a coating/film and a nanostructured zone. Typical techniques belonging to the first category are, for example, plasma spraying, plasma immersion ion implantation and deposition, chemical or physical vapor deposition, cold spraying, and self-assembly, whereas the second category is represented by methods like laser etching, shot blasting, acid and alkali treatments, anodic oxidation, micro-arc oxidation, ion implantation, and similar [1].

The nanofunctionalization of the surface can improve the performances of biomaterials, helping their use in clinical applications. In fact, cells life in complex tissues is controlled through modulation of cell binding to the extracellular matrix, which is composed by nanostructures. Cells respond to micro- and nanofeatures with different chemistries and topographies, resulting in changes in cell alignment, polarization, elongation, migration, proliferation, and gene expression. Through micro- and nanopatterning techniques it is possible to precisely position selected

biomolecules on a substrate for driving cell growth and, possibly, regulating cell functions [2].

While flat nanostructured surfaces are mainly suitable for cells growing in different conditions depending on chemical and topographical patterning of substrates, other nanostructures are also deeply studied for various applications. Nanostructures such as metal or metal oxide nanoparticles, quantum dots, silica, carbon nanotubes, C₆₀, hydrogel, block copolymer, and dendrimer are reported [3]. Main applications of these different types of nanostructures are imaging (both biolabeling and biosensing, comprising diagnostics purposes) and drug or gene delivery (for instance nanobodies or affibodies against cancer cells). Nanostructured surfaces, such as possibly functionalized nanoparticles or even nanotubes, are also included in micro- and nanodevices aimed at diagnostics and/or medical treatments.

Various methods for the fabrication of nanostructured functionalized surfaces are available, either adapted from macro-scale technologies or ad hoc developed, while fewer are the techniques suitable for the production of nanoarrays containing multiple features such as a range of different proteins or small ligands. Topographical and chemical nanopatterning of surfaces by methodologies substantially developed for fabrication in the microelectronic field are not reported in this chapter, whereas scanning probe-based techniques utilized to obtain functional structures at nanoscale (such as nanoarray) with different features are described in the following paragraphs. They are based on, for example, dip pen nanolithography, SNOM lithography, and nanografting.

Flat Nanostructured Surfaces

Flat nanostructured functionalized surfaces are surfaces with defined regions of different chemical functionality to achieve site-specific attachment of one species in some areas while minimizing unwanted surface interactions in other areas. Two surface patterning types can be distinguished, a structural topography and a chemical patterning, but at nanometric scale these features can be mixed. Moreover, a topographical pattern can be performed maintaining the surface chemistry, but the majority of chemical techniques induce a change in topography. Different strategies to achieve such patterned surfaces have been

proposed, mainly using techniques developed in the microelectronics field to chemically modify some parts of the surface, for example, via electron irradiation or extreme UV interference lithography or via microcontact printing. Alternatively, the combination of different substrate materials for the creation of patterned surfaces, commonly via photolithography, is also reported [4]. The production of patterns of low cost and high complexity (three or more addressable chemistries) over large areas was recently described [5]. The method proposes the creation of a long-range-ordered array of microscale “cups” with multiple material regions by a combination of colloidal lithography and the bottom-up deposition techniques of plasma polymerization and evaporation as well as sputtering via physical vapor deposition. Plasma polymerization possesses indeed the unique practical advantages that offer single-step coating with nanometer-scale control over thickness and a wide range of functional groups can be presented with control over density.

Patterned surfaces gained great attention since cells respond to micro- and nanofeatures with different chemistries and topographies. The influence of various topographic features such as grooves, ridges, steps, pores, wells, and nodes in microscale or nanoscale has been analyzed by culturing a wide variety of cells in the presence of such patterned nanosurfaces. Fibroblasts are greatly studied with nanotopography, showing as a general behavior that the adhesion increases with decreasing feature size or height. Morphology and cytoskeleton arrangement also change. The details of the nanotopography, and hence randomness and pattern, play an important role in eliciting different cell responses [6]. Muscle cells, endothelial cells, blood cells, hepatocytes, and many other cell lines were also studied, as revised by Yim et al. [6], concluding that cells vary their adhesion, proliferation, migration, and gene expression properties if grown on nanostructured substrates. This behavior could be ascribed to the fact that cells in their natural environment interact with extracellular matrix components which are in the nanometer scale. For example, the basement membrane of many tissues displays features of pores, fibers, and ridges in the nanometer range.

Moreover, surface nanofunctionalization was demonstrated to significantly affect cellular and subcellular functions, as reviewed by Liu et al. [1], stating that clinical applications of biomaterials can be improved

by producing a nanostructured surface. Apart from physical or chemical treatments aimed at the creation of nanostructures on a material surface, the modification of a substrate with specific ligand which improves cell attachment and differentiation is also well documented. Bell et al. [7], for example, reported the use of oligopeptides to coat titanium surfaces for osteointegration of orthopedic implants by favoring attachment of osteoprogenitor cells and promoting osteoblastic differentiation. Stem cell differentiation to desired phenotypes was also deeply studied in response to nanostructured growing substrates [2]. In particular, it was demonstrated that it is possible to drive cell adhesion, migration, proliferation, and even differentiation of stem cells, by the design and fabrication of growing surfaces or scaffolds with controlled topography and chemistry at the micro- and nanoscale. Topography acts on cell shape and alignment through mechanical interaction of the plasma membrane and the specific ligands present on it with the nanostructures on the substrate surface. However, the most important parameter at micro- and nanometer level seems the order or randomness characteristic of the surface. Chemical patterning regulates cell functions by mechanical interaction (e.g., adhesive microislets) or by direct receptor-ligand interaction with cell-membrane proteins. Topography and chemistry can also be combined to have synergistic effects and mimic the natural cell micro-environment.

Apart from eukaryotic cells, bacteria can also colonize surfaces possibly responding to nanotopography. Main difference is however the fact that bacteria live on surfaces as a community within a species- and strain-specific extracellular polymer, forming the so-called biofilm. Little is known about the response of bacteria toward nanostructures and only indirect experimental results support the idea that bacteria are able to use molecular features of their cell membrane as sensors and implement intracellular signaling pathways to sense the surface and to react to the stimuli created by the surface [8].

Functionalized Nanostructures

Various nanostructures or nanomaterials sensitive to environmental or biological parameters have been reported. Usually they are made of metals, metal oxides, semiconductors, silica, carbon nanotubes, C_{60} , hydrogel,

block copolymer, and dendrimer [3]. Nanometer-sized metallic structures, such as platinum, gold, and silver nanoparticles, exhibit unique optical properties. In fact, the characteristics of the plasmons excited by the light on their surface are depending on the dielectric constant of the media that constitute the interface. The adsorption of molecules to the nanoparticle changes this parameter, leading to detectable optical changes of the plasmon resonance. Moreover, these nanostructures can be functionalized to specifically interact, for example, with cells for site-specific drug delivery or imaging purposes. It was clearly demonstrated that differently decorated nanoparticles interact in different extent with cell membranes and cell uptake is indeed greatly influenced [9]. Iron oxide magnetic nanoparticles have been used widely in biological detection because of their good biocompatibility and ability to be manipulated under an external magnetic field. Quantum dots are also extensively used for biolabeling and biosensing, since they exhibit advantages in signal brightness, photostability, multicolor-light emission, and size-tunable properties.

Carbon-based nanomaterials are recently gaining interest for biomedical applications. Carbon nanotubes have been used as drug or gene carriers and for biosensing, whereas certain fullerene C_{60} derivatives and fluorescent carbon dots have been developed as nontoxic alternatives to semiconductor quantum dots for biomedical imaging [10].

Self-assembly has been also extensively explored to functionalize the material surface and build nanostructures or nanomaterials, including ordered mono- and multilayers, liposomes or vesicles, nanoparticles, nanotubes, and nanofibers. Moreover, self-assembled nanostructures from block copolymers have also frequently been reported. They consist of two or more covalently bounded blocks with different physical and chemical properties, resulting in a supramolecular assembly that can respond to changes in temperature, pH, and redox potential. In fact, they can generate various microdomain morphologies by means of intramolecular forces.

Since active selectivity is often desirable, targeted drug delivery carriers are functionalized with antibodies or antibody fragments to provide such active localization. The conjugation of multiple antibodies to each nanocarrier enhances their specificity and ability to detect their targets; nanocarriers can thus be surface-functionalized with multiple distinct antibodies to overcome tumor heterogeneity. Functional, single-

domain heavy-chain antibodies, referred to as nanobodies, have been raised against cancer targets with the purpose to either antagonize receptor function or deliver an enzyme for prodrug activation. Affibodies against a variety of cancer-related targets have been developed and are now commercially available, including: EGFR, HER2, and transferrin [11]. Nanoparticles have also been capped with various biocompatible polymers such as PVA, dextran, silica, and inorganic capping agents like biotin, citric acid, avidin, carbodiimide, SiO₂, and chitosan-PEG. These molecules then act as linkers for the coupling of cytotoxic drugs or target antibodies. A step forward is, for example, the design of luminomagnetic nanocarriers functionalized with the ligand folic acid [12]. This approach combines the luminescent and magnetic properties of the luminomagnets with the specificity of folic acid making them feasible for folate receptor-mediated molecular imaging and targeted drug delivery. Besides these, the small size of these luminomagnetic particles (about 8 nm) can be exploited for their easy elimination from the body after delivery of the drugs.

Furthermore, micro- and nanostructures are included in the development of smaller medical devices or systems aimed to drug and gene delivery, tissue engineering, biosensors, and diagnostic systems. The final goal relies in developing implantable sensing and/or treatment devices to collect information for diagnosing disease and administering treatment *in vivo* [3].

Finally, another class of nanostructured materials is often mentioned as potentially suitable for bio-applications. Hydrogels and molecular imprinted polymers can be considered either as flat nanofunctionalized surfaces or nanosystems similar to those described in this paragraph. Hydrogels are 3D high-molecular-weight networks composed of a polymer backbone, water, and a cross-linking agent. They can be stimuli sensitive, thus being of great interest in drug delivery, cell encapsulation, and tissue engineering [3]. Molecular imprinted polymers are also described as powerful nanotools for specific applications. They are tailor-made biomimetic receptors often used for functionalizing the surface of nanomaterials with sensing or separation purposes. Molecular imprinted polymers are produced by polymerization in the presence of molecular templates that are removed afterward, thus creating complementary nanostructured surfaces. These new surfaces can bind selectively the target (or template) used to prepare them [3].

Nanoscale Patterns Fabrication

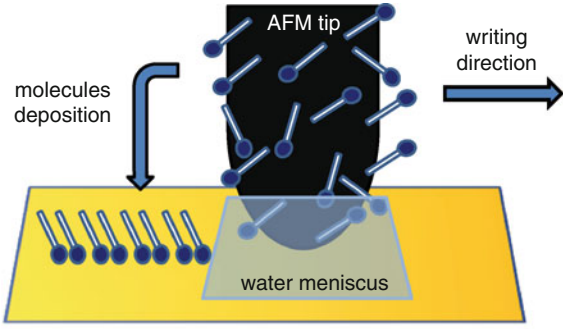
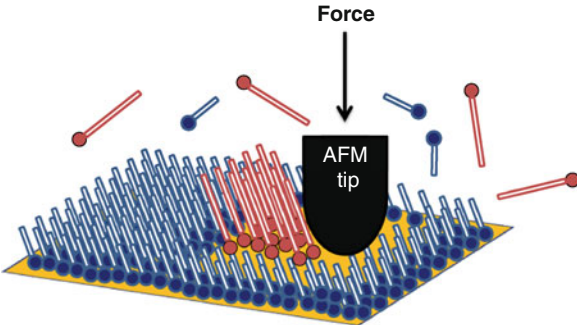
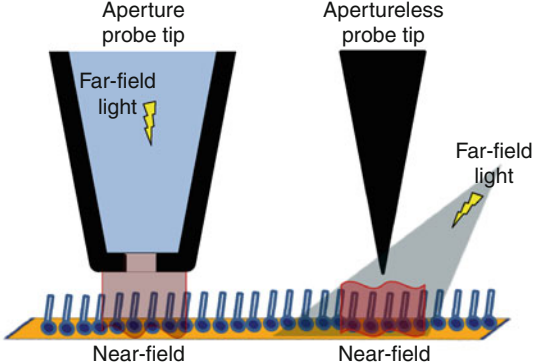
Nanoscale patterns are generally quite difficult to work with, both in terms of preparation and characterization. Besides downscaling the fabrication methods originally developed for microconditions to nanoscale (such as photolithography, nanoimprint lithography), dedicated fabrication methods have been proposed for these applications. The following paragraphs describe the scanning probe-based techniques most commonly utilized to obtain functional structures at nanoscale such as nanoarrays containing different features.

Dip Pen Nanolithography

A scanning probe technique that has been developed for the production of nanoarrays is the ► **Dip-Pen Nanolithography** (DPN), first reported by Mirkin's group in 1999 [13]. DPN makes use of one or several AFM tips that are dipped into an ink and subsequently used to write a nanoscale pattern on a surface. The method was presented as a powerful technique to deposit molecules on a substrate from AFM tips at resolutions comparable to those achieved with much more expensive lithographic methods. The principle takes advantage of the water meniscus that naturally forms between the ink-coated probe tip and the substrate when the scanning was done in contact mode under ambient laboratory conditions. The ink moves on the substrate by capillary transport through the meniscus, as shown in Fig.1, Panel A. A key point is the correct choice of the ink and the substrate, which have to possess reciprocal chemical affinities, thus favoring the chemical adsorption of the ink molecules onto the substrate. An useful application of this technique is related to the precise functionalization of nanoscale devices previously prepared by more conventional lithographic methods. The resolution of DPN depends on several parameters such as the grain size of the substrate, the molecule diffusion rate, the tip-substrate contact time and the scan speed, and finally on the environmental conditions such as humidity. The minimum feature sizes producible by DPN are in the range of 15 nm [8].

Nanografting

An effective way to introduce chemical and functional features with nanoscale lateral resolution down to few nanometers has been developed by Liu and coworkers starting from 1997 [14, 15] using an AFM (► **AFM**) tip.

<p>a DIP PEN LITHOGRAPHY</p> <p>The patterning is obtained directly depositing the molecules.</p>		<p><u>Lateral Resolution</u></p> <p>Down to 15 nm</p>
<p>b NANOGRAFTING</p> <p>The patterning is obtained replacing some molecules with others.</p>		<p>3–5 nm</p>
<p>c SNOM LITHOGRAPHY</p> <p>The patterning is obtained partially removing some molecules or photoactivating specific areas.</p>		<p><i>Aperture tips:</i> down to 50 nm</p> <p><i>Apertureless tips:</i> down to 10 nm</p>

Nanostructured Functionalized Surfaces, Fig.1 Summary of scanning probe-based techniques for functional surface nanopatterning. A brief description of the principle, an illustrative cartoon, and the lateral resolution of each technique are reported

This method is based on the local modification of a self-assembled monolayer (► [Self-assembled monolayers](#)) surface, previously deposited starting from an alkanethiol solution, on a flat gold substrate. The surface is then mounted in a liquid environment AFM cell, allowing its observation and nanomanipulation with the AFM tip. While scanning the surface in a low load regime (< 1 nN) results in the three-dimensional imaging of the surface, increasing the load over a threshold value one obtains the removal of the previously deposited thiolated molecules under the tip, locally disrupting the monolayer and exposing the gold surface. If thiolated

molecules (usually different from that used in the initial SAM deposition) are present in solution, a newly deposited SAM is obtained in areas where the tip is removing the old one, locally changing morphology and chemistry of the surface. Patches with different height with respect to the surrounding surface can be deposited and detected in topography AFM images. Nanografted structures can also introduce characteristic groups in specific areas that show a chemical contrast in AFM friction images. The principle is illustrated in Fig.1, Panel B.

Nanografting can be applied even with a solution of the same molecule previously used for the initial SAM

formation (homogeneous nanografting), obtaining in this case the removal of defects that were possibly present in the SAM [15]. Having obtained a nanopatterned surface endowed with specific chemical groups in well-defined patches, one can then exploit this method to create nanoarrays of biomolecules. Liu and coworkers [16] showed how to take advantage of both non-covalent and covalent attachments to form such protein nanoarrays, down to a lateral scale of tens of nanometers. This method generally leads to non-oriented protein patches, as demonstrated by the AFM analysis, that shows the presence of proteins with different heights with respect to the surface, corresponding to the different protein orientation. The nanografting technology can be also successfully applied to the realization of DNA nanoarrays, when thiolated single-stranded DNAs are used in the solution during the nanografting process. DNA lines as narrow as 10 nm can be indeed fabricated. Moreover, nanografting allows a controlled density deposition of DNA strands, depending on the number of tip passages over the same area [17]. Successful DNA hybridization has been demonstrated on such nanoarrays.

It is also possible to extend such a technique to realize protein nanoarrays: Patches of single-stranded DNAs of different sequences are deposited using nanografting in specific places of a surface previously covered with a SAM able to discourage aspecific protein adhesion. Then, proteins linked with DNA sequences that are complementary to that of the patch where the protein deposition is desired are incubated over the surface. Thanks to the DNA complementarity, a specific protein localization is then obtained. The effectiveness of this methodology for the realization of nanoscale protein sensors has been demonstrated using human serum enriched with antibodies against specific proteins that were previously immobilized on nanografted patches, showing that antibodies bind to the patches where their targets are present [18].

A method that in similar fashion takes advantage of the DNA hybridization properties to direct the localization of nanosized objects with the help of an AFM probe (although not based on nanografting) has been demonstrated by Kufer and coworkers in 2008 [19]. Their method is based on the fabrication of specific areas on a glass surface, namely, depot areas – where DNA-coupled nanoobjects are picked up by a single-stranded functionalized AFM tip, and target areas – where the nanoobjects are delivered with a precision of some nm to form the desired pattern.

More than 5,000 object transfers with the same tip have been demonstrated.

SNOM Lithography

In 1983, it was proved the Scanning Near-field Optical Microscopy (SNOM), shortly after the invention of the first proximal probe technology, the Scanning Tunneling Microscope (STM). These two techniques are based on a similar principle, but in the case of STM the detection is performed on the electron tunneling currents, while in the case of SNOM the detection is performed on the tunneling of photons. This characteristic makes the SNOM suitable for the analysis of nonconductive samples [20].

By illuminating the sample with near-field light, the probe tip is held very close to the sample without touching. An evanescent light field is created in very close proximity of surface. The evanescent light through the aperture is incident on the sample, exciting the atoms in the surface to reradiate propagating waves. The same probe can then collect the propagating light and transmit it to an appropriate detector. The probe tip can be designed in two different ways, an aperture-based tip and an apertureless-based tip (as illustrated in Fig.1, Panel C). In an aperture-based SNOM, however, the nanometer-scale aperture limits the light throughput and the resolution. To overcome these limitations an apertureless SNOM was developed that uses an apertureless probe, similar to an STM or AFM tip, to scan within the near field of the sample. In this case the sample is first irradiated from above by far-field light in order to produce the evanescent field on the surface of the sample. The evanescent photons then excite the atoms in the probe tip, which reradiate the propagating photons.

An interesting application of this technique is related to the lithography-like process. The SNOM probe is used as an energy source for modifying a variety of materials at the nanoscale level, including conventional photoresists and non-conventional electro-magneto-optical materials as well as ferroelectric and SAM. The photoresists layer deposited on a substrate may be changed for instance by laser pulses emitted from the SNOM probe, offering resolutions unachievable by conventional optical and laser systems (well beyond the diffraction limit). Patterns are generated by scanning the probe over these target surfaces and both aperture and apertureless probe tips are used, obtaining different resolution limits (50 nm for aperture probe and down to 10 nm for the apertureless probe).

Cross-References

- ▶ [AFM](#)
- ▶ [Chitosan Nanoparticles](#)
- ▶ [Dip-Pen Nanolithography](#)
- ▶ [Functionalization of Carbon Nanotubes](#)
- ▶ [Microfabricated Probe Technology](#)
- ▶ [Nanomedicine](#)
- ▶ [Nanoparticles](#)
- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)
- ▶ [Self-assembled Monolayers](#)
- ▶ [Self-assembly of Nanostructures](#)

References

1. Liu, X., Chu, P.K., Ding, C.: Surface nano-functionalization of biomaterials. *Mat. Sci. Eng. R* **70**, 275–302 (2010)
2. Martínez, E., Lagunas, A., Mills, C.A., Rodríguez-Seguí, S., Estévez, M., Oberhansl, S., Comelles, J., Samitier, J.: Stem cell differentiation by functionalized micro- and nanostructured surfaces. *Nanomedicine (Lond.)* **4**(1), 65–82 (2009)
3. He, J., Qi, X., Miao, Y., Wu, H.L., He, N., Zhu, J.J.: Application of smart nanostructures in medicine. *Nanomedicine (Lond.)* **5**(7), 1129–1138 (2010)
4. Briand, E., Humblot, V., Landoulsi, J., Petronis, S., Pradier, C.M., Kasemo, B., Svedhem, S.: Chemical Modifications of Au/SiO₂ Template Substrates for Patterned Biofunctional Surfaces. *Langmuir* **27**(2), 678–685 (2011)
5. Ogaki, R., Cole, M.A., Sutherland, D.S., Kingshott, P.: Microcup arrays featuring multiple chemical regions patterned with nanoscale precision. *Adv. Mater.* **23**(16), 1876–1881 (2011)
6. Yim, E.K.F., Leong, K.W.: Significance of synthetic nanostructures in dictating cellular response. *Nanomed. Nanotechnol. Biol. Med.* **1**, 10–21 (2005)
7. Bell, B.F., Schuler, M., Tosatti, S., Textor, M., Schwartz, Z., Boyan, B.D.: Osteoblast response to titanium surfaces functionalized with extracellular matrix peptide biomimetics. *Clin. Oral Impl. Res.* **22**(8), 865–872 (2011)
8. Anselme, K., Davidson, P., Popa, A.M., Giazzon, M., Liley, M., Ploux, L.: The interaction of cells and bacteria with surfaces structured at the nanometre scale. *Acta Biomater.* **6**, 3824–3846 (2010)
9. Verma, A., Uzun, O., Hu, Y., Han, H.S., Watson, N., Chen, S., Irvine, D.J., Stellacci, F.: Surface-structure-regulated cell-membrane penetration by monolayer-protected nanoparticles. *Nat. Mater.* **7**, 588–595 (2008)
10. Xing, Y., Dai, L.: Nanodiamonds for nanomedicine. *Nanomedicine* **4**(2), 207–218 (2009)
11. Murday, J.S., Siegel, R.W., Stein, J., Wright, J.F.: Translational nanomedicine: status assessment and opportunities. *Nanomedicine* **5**(3), 251–273 (2009)
12. Dutta, R.K., Sharma, P.K., Pandey, A.C.: Design and surface modification of potential luminomagnetic nanocarriers for biomedical applications. *J. Nanopart. Res.* **12**, 1211–1219 (2010)
13. Piner, R.D., Zhu, J., Xu, F., Hong, S., Mirkin, C.A.: “Dip-Pen” Nanolithography. *Science* **283**, 661–663 (1999)
14. Xu, S., Liu, G.: Nanometer-scale fabrication by simultaneous nanoshaving and molecular self-assembly. *Langmuir* **13**(2), 127–129 (1997)
15. Xu, S., Miller, S., Laibinis, P.E., Liu, G.: Fabrication of nanometer scale patterns within self-assembled monolayers by nanografting. *Langmuir* **15**(21), 7244–7251 (1999)
16. Wadu-Mesthrige, K., Xu, S., Amro, N.A., Liu, G.: Fabrication and imaging of nanometer-sized protein patterns. *Langmuir* **15**(25), 8580–8583 (1999)
17. Mirmomtaz, E., Castronovo, M., Grunwald, C., Bano, F., Scaini, D., Ensafi, A.A., Scoles, G., Casalis, L.: Quantitative study of the effect of coverage on the hybridization efficiency of surface-bound DNA nanostructures. *Nanoletters* **8**(12), 4134–4139 (2008)
18. Bano, F., Fruk, L., Sanavio, B., Glettenberg, M., Casalis, L., Niemeyer, C.M., Scoles, G.: Toward multiprotein nanoarrays using nanografting and DNA directed immobilization of proteins. *Nanoletters* **9**(7), 2614–2618 (2009)
19. Kufer, S.K., Puchner, E.M., Gump, H., Liedl, T., Gaub, H.E.: Single-molecule cut-and-paste surface assembly. *Science* **319**, 594–596 (2008)
20. Tseng, A.A.: Recent developments in nanofabrication using scanning near-field optical microscope lithography. *Opt. Laser Technol.* **39**, 514–526 (2007)

Nanostructured Hydrogels

- ▶ [Nanoengineered Hydrogels for Cell Engineering](#)

Nanostructured Materials

- ▶ [Nanomaterials for Electrical Energy Storage Devices](#)

Nanostructured Materials for Sensing

Mariangela Lombardi
 IIT - Italian Institute of Technology @ POLITO -
 Centre for Space Human Robotics, Torino, Italy

Synonyms

[Nanomaterials for sensors](#)

Definition

Materials having nanometric dimensions are able to improve their functional properties, which allow their exploitation in highly performing sensors.

Overview

Materials with at least one dimension smaller than 100 nm are classified as nanostructured. Compared to micro-sized ones, nanomaterials are characterized by a higher surface to bulk ratio, usually leading to an enhancement of optical, mechanical, electrical, structural, and magnetic properties [1].

In the last few years, research focused its efforts on the development of multifunctional composite materials, able to perform multiple non-structural functions, sometimes combined with a suitable mechanical response. If nanostructured, they can also imply additional functional properties with respect to conventional composites. Multifunctional materials can be used for developing highly performing sensors, particularly in the case of single-phased and composite materials containing carbon nanotubes and piezoelectric phases [2].

Piezoelectric Materials

Piezoelectric materials are able to directly convert a mechanical signal into an electrical one and vice versa. For this reason piezoelectric materials produce a charge when subjected to a stress (direct effect) or a strain when an electric field is applied (converse effect). These materials can then be used as actuators or sensors, by exploiting their converse or direct piezoelectric effect, respectively. Generally piezoelectric sensors are used for measuring several parameters, like acceleration, angular rate, pressure, force, temperature, mass, magnetic field, and amount of several chemical substances [3].

Piezoelectricity can be expressed by several materials, such as single crystals, ceramics, polymers, and composites [4, 5]. Even if they naturally present this behavior, piezoelectric single crystals, as quartz and Rochelle salt, cannot be easily used being isotropic, e.g., they have different properties depending on the

cut and the orientation directions. Most piezoelectric ceramics applied in commercial sensing devices are made of polycrystalline ferroelectric ceramics [6], which present a higher strength and an easier fabrication with respect to single crystals [4]. Ferroelectrics are characterized by a spontaneous polarization which can be orientated under a realizable electric field. In fact, they initially contain randomly oriented, polarized regions, named domains. These polarized regions form upon cooling at a temperature called the Curie temperature to minimize the total elastic energy of the ceramic structure. In this state, they do not present piezoelectric properties. When an electric field is applied, the domains align; this process is known as poling. When the electric field is removed, most domains remain aligned and piezoelectricity is induced.

Generally the piezoelectric ceramics present a composition that implies a structural instability such as polymorphic phase transition (PPT) or morphotropic phase boundary (MBP) [7]. PPT-based systems are characterized by a large temperature dependence of piezoelectric properties and a rapid depoling, due to temperature changes close to PPT temperature. In the case of MBP-based materials, piezoelectric features are slightly influenced by temperature, but their working temperatures approach a half of Curie temperature.

Among polycrystalline ferroelectric ceramics, the first system studied for its piezoelectric properties was barium titanate (BaTiO_3) in 1943 and it remained the primary electroceramic material until the discovery of the lead zirconate titanate or PZT ($\text{PbZr}_x\text{Ti}_{1-x}\text{O}_3$) in 1954 [4]. The former is a PPT-based, high-performance piezoelectric ceramic, the latter presents a MBP from tetragonal to rhombohedral phase at about 48 mol% PbTiO_3 content.

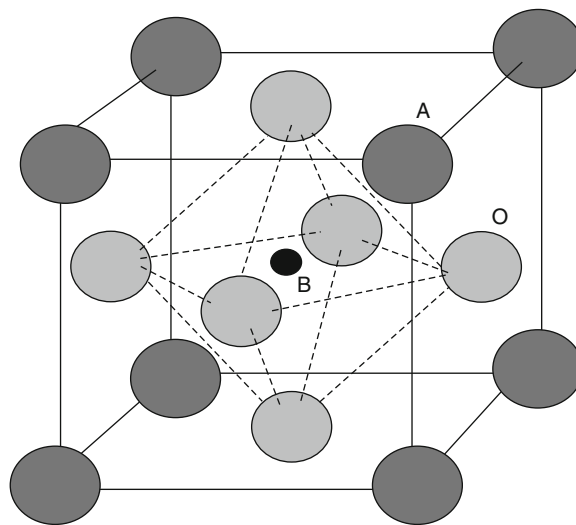
For a long time PZT was the most used piezoceramic for its very attractive behavior in sensor applications [5, 7]. Notwithstanding this, in 2003 the European Union claimed PZT as a hazardous substance and research efforts were then focused on lead-free piezoceramics having high performances. For this reason, in the last decade, alternative systems were deeply studied modifying the composition of the traditional ferroelectric ceramics. In particular, this *compositional engineering approach* involved perovskite-based materials, exploiting their solid solutions with other ferroelectric systems or their doping, e.g.,

the introduction of different types of cations within their structure. The perovskite structure ABO_3 consists in a face-centered cubic array, in which the large cations A are placed at the corners of the unit cell, the smaller cations B in the body center, and the oxygen atoms in the center of each face, as illustrated in Fig. 1 [5–7].

Even if the high symmetry of the perovskite structure could lead to improved poling and consequently higher piezoelectric performances, currently no innovative lead-free ceramics are able to compare the piezoelectric and thermal behavior of PZT-based materials.

On the other hand, a *structural engineering approach*, based on template grain growth, grain size optimization, and domain engineering, is able to enhance the electromechanical performances of existing piezoelectric materials. For instance, in the case of $BaTiO_3$, a microstructural refinement can double the piezoelectric coefficient. In this case, the finer grain size implies a decrease of the domain size and an increase in density of domain walls, which are able to influence the piezoelectric properties [5, 8]. Moreover, a fourfold increase of the piezoelectric coefficient can be reached synergically combining the effect of submicrometric domain sizes and grain orientation obtained by template grain growth technique. The positive effect of the grain orientation was also demonstrated in the binary system $(Bi_{1/2}K_{1/2})TiO_3$ – $BaTiO_3$ (BKT–BT) [5].

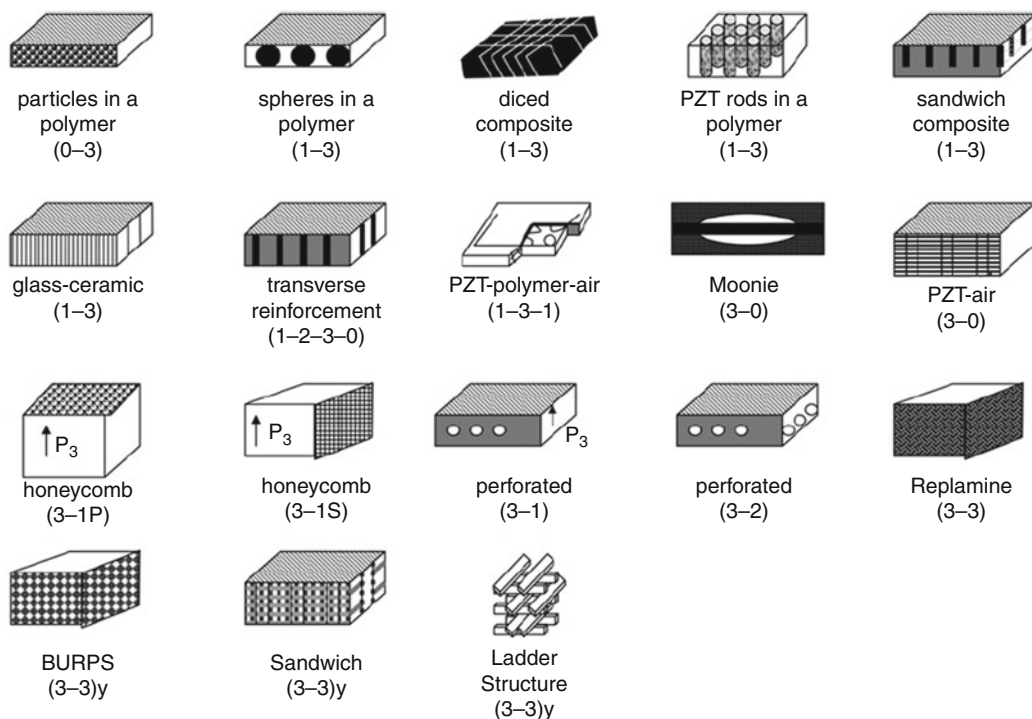
Generally piezoelectric ceramics can be applied in sensors as bulk components or thin films, by using micromachining techniques [2, 3]. Thin films from one tenth to several tens of microns in thickness can present different properties with respect to the corresponding bulk material. This is due to the dependence of the piezoelectric features to the morphology and density of the film, as well as to the compositional aspects related to the stoichiometry and presence of impurities. Moreover, the piezoelectric behavior can also be influenced by the orientation and nature of the substrate material. All the deposition methods are based on three fundamental steps; firstly, the atomic or molecular species are yielded by chemical or physical processes. In the former case chemical vapor deposition or sol-gel techniques can be used; in the latter one, sputtering or pulsed laser deposition can be employed. After this step, deposition on the substrate is achieved through transport and condensation



Nanostructured Materials for Sensing, Fig. 1 The perovskite structure common in most ceramic piezoelectric materials (With permission from [5])

mechanisms and finally, the film is obtained by crystallization and annealing [3].

In order to achieve a long-term stability of piezoelectricity at room temperature and to develop flexible devices, piezoelectric polymers can be employed [9]. Generally ferroelectric materials based on polyvinylidene fluoride (PVDF), odd nylon, vinylidene cyanide (VDCN), and aromatic and aliphatic polyurea families are exploited [9, 10]. In the case of PVDF-based polymers, the residual polarization after poling is due to the alignment of CF_2 dipoles present in their structure $(CH_2-CF_2)_n$ and trapped charges. Dipoles are oriented in the crystalline phases inside lamellae and in the interfaces between crystalline and noncrystalline phases, while the charge traps are present at the surface of crystallites [10]. The electromechanical performances of these polymeric materials can be improved by the polymerization of PVDF with other fluorinated polymers, which implies an increase of the crystallinity [9]. In the case of odd-nylon and polyurea families the field-induced orientation of N–H and C–O dipoles is stabilized by the hydrogen bonding present between adjacent polymer chains. On the other hand, in the case of amorphous polymers such as VDCN-based materials, an easier orientation of the C–CN dipoles is obtained by poling at temperatures near to the glass transition



Nanostructured Materials for Sensing, Fig. 2 Possible connectivity in bicomponent piezoelectric composites (With permission of [5])

temperature and freezing the residual polarization, through the cooling down of the system under the electric field [9, 10].

Among piezoelectric polymers, PVDF presents the best electromechanical behavior at room temperature, but its performances decrease at higher temperature [9]. This drawback can be partially limited through the copolymerization of PVDF [11]. Other ferroelectric polymers can be used, but they are characterized by less piezoelectric properties [9, 11].

Compared to the pure polymer performances, piezoelectric composites offer significantly higher functional properties. It is in fact possible to synergically couple the best features of the ceramic and polymeric phases while minimizing the poorest properties [4, 11]. Composites can be classified on the basis of the connectivity between the constituent phases according to Newnham notation [4]. This notation describes the number of dimensions each phase is physically in contact with itself. When just two components are considered, ten configurations are possible: 0-0, 1-0, 2-0, 3-0, 1-1, 2-1, 3-1, 2-2, 3-2, or 3-3, illustrated in Fig. 2 [5].

Among them, 0-3, 1-3, and 2-2 connectivity have been recently developed, in order to tailor the piezoelectric response through the exploitation of anisotropic composites having different ceramic phase contents [11].

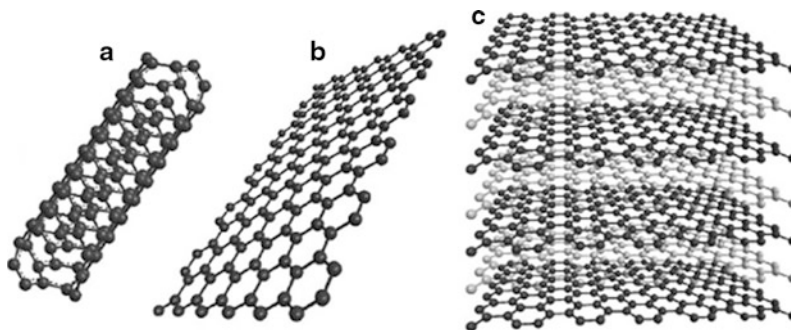
In the last few years, miniaturization of devices based on piezoelectric behavior involves the exploitation of carbon nanotubes as functional materials [12].

Carbon Nanomaterials

Multidimensional carbon nanomaterials include graphite, graphene, and carbon nanotubes. They are made of carbon atoms in an exagonal configuration, bonded each other with sp^2 bonds [13]. Graphene is a two-dimensional sheet of carbon atoms and represents the repetitive unit of the graphite. Carbon nanotubes (CNT) consist in graphene sheets rolled to form cylinders, as shown in Fig. 3. In particular, carbon nanotubes can be classified as single-walled (SWCNT) or multi-walled (MWCNT); the former, consisting in

Nanostructured Materials

for Sensing, Fig. 3 Structure of the multidimensional carbon nanomaterials: (a) CNT, (b) graphene, and (c) graphite (With permission of [13])



a single graphene cylinder, the latter, being made of concentric cylinders with an interlayer spacing of 0.34 nm. CNT nanostructure, atomic arrangement, diameter, and length strongly affect their mechanical, electric, and thermal properties [12].

Carbon nanotubes show an electromechanical behavior that, unlike to piezoelectric ceramics, is due to the phonon frequency shifts as a consequence of a charge injection [12]. This feature is mainly exploited in carbon nanotubes composites, which are characterized by a change in resistance when subjected to a chemical, thermal, or mechanical loading. Among them, the most promising sensing devices based on carbon nanotubes are the chemiresistive-type ones. In fact, their high sensitivity, small size, high surface area, and high aspect ratio imply the ability to detect smaller concentrations of analyte than traditional sensors.

In comparison to carbon nanotubes, the homogeneous distribution of electrochemically active sites at a nanometer scale and the higher purity could make graphene an ideal material for sensing and detection. Notwithstanding this, no consolidated studies have been carried out about its application in sensing devices to date [14].

Cross-References

- ▶ [Active Carbon Nanotube-Polymer Composites](#)
- ▶ [Carbon Nanotubes](#)
- ▶ [Graphene](#)
- ▶ [NEMS Piezoelectric Switches](#)
- ▶ [Piezoelectric Effect at Nanoscale](#)
- ▶ [Piezoelectric MEMS Switch](#)
- ▶ [Sol-Gel Method](#)
- ▶ [Synthesis of Carbon Nanotubes](#)
- ▶ [Synthesis of Graphene](#)

References

1. Fauchais, P., Montavon, G., Lima, R.S., Marple, B.R.: Engineering a new class of thermal spray nano-based microstructures from agglomerated nanostructured particles, suspensions and solutions: an invited review. *J. Phys. D: Appl. Phys.* **44**, 53 (2011). 093001
2. Gibson, R.F.: A review of recent research on mechanics of multifunctional composite materials and structures. *Compos. Struct.* **92**, 2793–2810 (2010)
3. Tadigadapa, M., Mateti, K.: Piezoelectric MEMS sensors: state-of-the-art and perspectives. *Meas. Sci. Technol.* **20**, 30 (2009). 092001
4. Tressler, J.F., Alkoy, S., Newnham, R.E.: Piezoelectric sensors and sensor materials. *J. Electroceram.* **2**, 257–272 (1998)
5. Cook-Chennault, K.A., Thambi, N., Sastry, A.M.: Powering MEMS portable devices—a review of non-regenerative and regenerative power supply systems with special emphasis on piezoelectric energy harvesting systems. *Smart Mater. Struct.* **17**, 33 (2008). 043001
6. Rodel, J., Jo, W., Seifert, K.T.P., Anton, E.M., Granzow, T., Damjanovic, D.: Perspective on the development of lead-free piezoceramics. *J. Am. Ceram. Soc.* **92**, 1153–1177 (2009)
7. Leontsev, S.O., Eitel, R.E.: Progress in engineering high strain lead-free piezoelectric ceramics. *Sci. Technol. Adv. Mater.* **11**, 13 (2010). 044302
8. Setter, N., Damjanovic, D., Eng, L., Fox, G., Gevorgian, S., Hong, S., Kingon, A., Kohlstedt, H., Park, N.Y., Stephenson, G.B., Stolitchnov, I., Taganstev, A.K., Taylor, D.V., Yamada, T., Streiffer, S.: Ferroelectric thin films: review of materials, properties, and applications. *J. Appl. Phys.* **100**, 46 (2006). 051606
9. Eberle, G., Schmidt, H., Eisenmenger, W.: Piezoelectric polymer electrets. *IEEE T. Dielect. El. In.* **3**, 624–646 (1996)
10. Fukada, E.: History and recent progress in piezoelectric polymers. *IEEE T. Ultrason. Ferr.* **47**, 1277–1290 (2000)
11. Lushcheikin, G.A.: New polymer-containing piezoelectric materials. *Phys. Solid State* **48**, 1023–1025 (2006)
12. Li, C., Thostenson, E.T., Chou, T.W.: Sensors and actuators based on carbon nanotubes and their composites: a review. *Compos. Sci. Technol.* **68**, 1227–1249 (2008)

13. Pumera, M., Ambrosi, A., Bonanni, A., Chng, E.L.K., Poh, H.L.: Graphene for electrochemical sensing and biosensing. *Trac-Trend. Anal. Chem.* **29**, 954–965 (2010)
14. Lü, P., Feng, Y.Y., Zhang, X.Q., Li, Y., Feng, W.: Recent progresses in application of functionalized graphene sheets. *Sci. China Tech. Sci.* **53**, 2311–2319 (2010)

Nanostructured Solar Cells

► Nanomaterials for Excitonic Solar Cells

Nanostructured Thermoelectric Materials

Menghan Zhou and Jian He
Department of Physics and Astronomy,
Clemson University, Clemson, SC, USA

Synonyms

Nanostructured thermoelectrics; Thermoelectric nanomaterials

Definition

Nanostructured thermoelectric materials are materials that contain nanoscale constituents and exhibit enhanced thermoelectric performance due to nanoscale phenomena.

Introduction

To a great extent, the development of civilization has evolved with the technological advent of energy production, storage, delivery, and use. As electricity will remain the most convenient form of energy in the foreseeable future, eco-friendly technology to generate electricity is of the utmost importance. Heat, however, has always been an abundant but low-quality form of energy as more than half of all the energy generated by mankind is lost as heat. As such, harvesting and converting the waste heat emanating from many different sources (e.g., solar, geothermal, and exhaust

from automobiles or other industrial processes) into electricity is highly desirable. One emerging market is to harvest the large amount of waste heat ($\approx 2/3$ of the generated power) from an automobile's engine and convert it into "onboard" electrical energy using thermoelectric (TE) devices. Meanwhile, there has always been a strong need for solid-state cooling (heat management) in microelectronics, which often requires responsive spot size cooling instead of traditional environmental cooling. This is best satisfied by TE refrigeration, the gain from temperature stabilization and device performance can be significant.

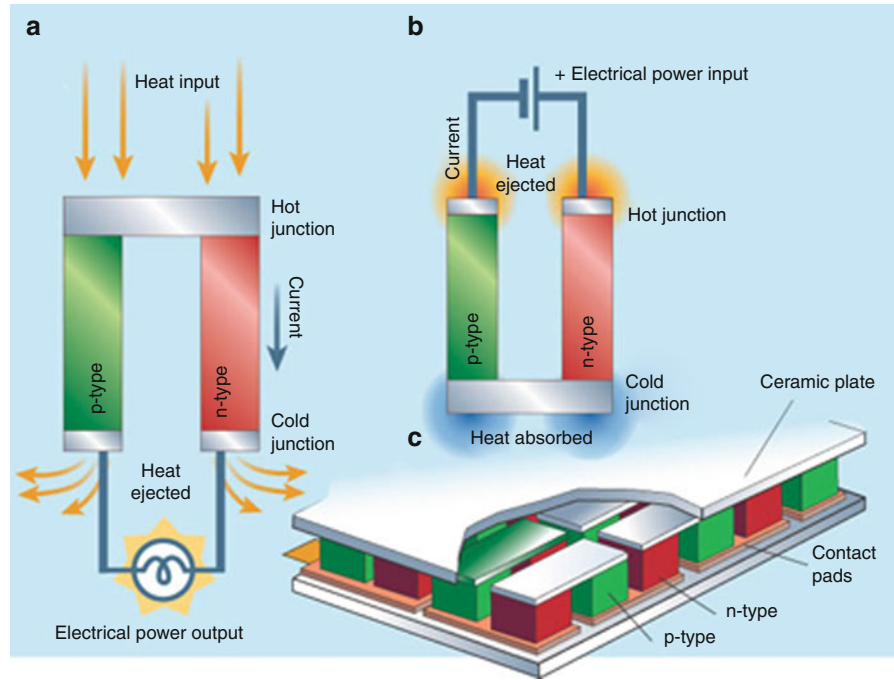
The simplest technology applicable for direct heat-electricity energy conversion is thermoelectricity [1]. Thermoelectricity is based on two basic effects, the Seebeck effect: generating electricity from a temperature gradient ("power generation mode"), and the Peltier effect: generating a temperature gradient when electrical current is applied ("refrigeration mode"). As shown in Fig. 1 [2], a basic thermoelectric couple consists of one leg made of n-type material and one leg made of p-type material in order to maximize the amount of power generated or the amount of heat absorbed, many of these couples are connected electrically in series and thermally in parallel to form a thermoelectric (TE) module or device.

The *working media* of TE materials are the electrons, rather than the atomic or molecular gases or liquids used in mechanical power generation and refrigeration systems. Accordingly, the TE device is a solid-state assembly without moving parts or green house emissions, lightweight and compact, responsive, and feasible for miniaturization, the TE device can readily work in tandem with other alternative energy technologies, such as photovoltaics. This is an important feature because no single technology can meet the world's energy needs in the twenty-first century; a combination of many technologies is necessary. Given the ubiquitous heat sources and the modular aspects of the TE device, thermoelectricity is guaranteed a position as an alternative and complementary energy technology of the twenty-first century.

At present, however, the conversion efficiency of thermoelectricity is inferior to that of equivalent mechanical systems by a factor of ~ 2 –4. For this reason, thermoelectricity has been long confined to *niche* applications, where the issue of efficiency is less of a concern than the issues of energy availability,

Nanostructured Thermoelectric Materials,

Fig. 1 The basic thermoelectric module, made of *n*-type and *p*-type materials, can work either in the (a) power generation mode or (b) refrigeration mode; and (c) a thermoelectric device is made of many basic modules [2]



reliability, compactness of the device, and quiet operation. For example, radioisotope TE power generators are the primary power source used in the *Voyager* and *Cassini* spaceships for NASA's deep-space missions beyond Mars, where fuel cells, solar cells, and nuclear power are not feasible.

The efficiency of a TE device can be calculated by regarding it as a heat engine working between a heat reservoir and a heat sink. The maximum efficiency of power generation, η , and that of TE refrigeration, ϕ , of a TE device [1] is determined by the Carnot efficiency, η_{Carnot} , and the dimensionless figure of merit, ZT , of the TE materials of which it is composed.

$$\eta = \eta_{Carnot} \eta_{TE} = \left(\frac{T_{hot} - T_{cold}}{T_{hot}} \right) \left[\frac{\sqrt{1 + ZT_m} - 1}{\left(\frac{T_{cold}}{T_{hot}} \right) + \sqrt{1 + ZT_m}} \right] \quad (1)$$

$$\phi = \left(\frac{T_{cold}}{T_{hot} - T_{cold}} \right) \left[\frac{\sqrt{1 + ZT_m} - \left(\frac{T_{hot}}{T_{cold}} \right)}{\sqrt{1 + ZT_m} + 1} \right] \quad (2)$$

The Carnot efficiency, η_{Carnot} , is the ratio of the temperature difference between the hot-end

temperature, T_{hot} , and the cold-end temperature, T_{cold} , to T_{hot} , and T_m is the mean temperature [1]. ZT is the material parameter of prime importance in thermoelectrics study:

$$ZT = \frac{\sigma \alpha^2 T}{\kappa} = \frac{PF}{\kappa} = \frac{\sigma \alpha^2 T}{(\kappa_{ph} + \kappa_e)} \quad (3)$$

where σ is the electrical conductivity, α the thermopower (i.e., Seebeck coefficient), κ the total thermal conductivity (often expressed as the sum of the lattice thermal conductivity, κ_{ph} , and the carrier thermal conductivity, κ_e), PF the power factor, and T the temperature in Kelvin. The Wiedemann–Franz relationship, $\kappa_e = L\sigma T$, is often used to estimate κ_e , where L is the Lorenz number. Although there is no theoretical upper limit for ZT values, the current state-of-the-art bulk TE materials have a maximum $ZT \sim 1$ in their respective temperature ranges of operation because the physical quantities (σ , α , and κ) that govern ZT are inherently interdependent: A helpful modification of any of these quantities often adversely affects the others. These materials ideally have conversion efficiencies of 7–15% depending on the specific materials and the temperature differences involved.

Nanostructured Thermoelectrics: Big Gains from Small Sizes

Equations (1)–(3) shift the fundamental issue of TE energy conversion to the materials' physical transport properties. As such, modern thermoelectrics studies are materials oriented; tremendous efforts have been contributed continuously to developing high ZT TE materials. A high ZT TE material is ideally a “phonon-glass electron-crystal” (PGEC) system that also possesses a large thermopower. A PGEC system is a system that simultaneously possesses a glass-like poor lattice thermal conductivity and a crystal-like good power factor. Over the past decade, the efforts of pursuing higher ZT materials have culminated into a two-pronged strategy, depending on the type of material studied. One approach is to develop novel bulk materials in line with the concepts of “PGEC,” while the other is to search for higher ZT in nanostructured thermoelectric materials (NTMs) that possess reduced dimensionality or nanometer characteristic lengths [3, 4].

The main themes of this entry are the strategies, implementations, and challenges of NTMs. Interested readers are referred to other reviews [5–7] and the references therein for more details about the topics discussed herein.

Classical and Quantum Size Effects

The paradigm of NTMs dates back to the early proposition by Hicks and Dresselhaus [3] that a low-dimensional material may have enhanced TE properties compared to its bulk counterpart due to quantum size effects. Low-dimensional material is the material whose system size in one or more directions is comparable with the wavelength or the mean free path of a quantum particle or an excitation. Accordingly, the charge-carrying electrons and heat-carrying phonons may be subject to different size effects, providing an extra layer of control of the charge and heat flows via altering the characteristic length scales of a low-dimensional (“confined”) system.

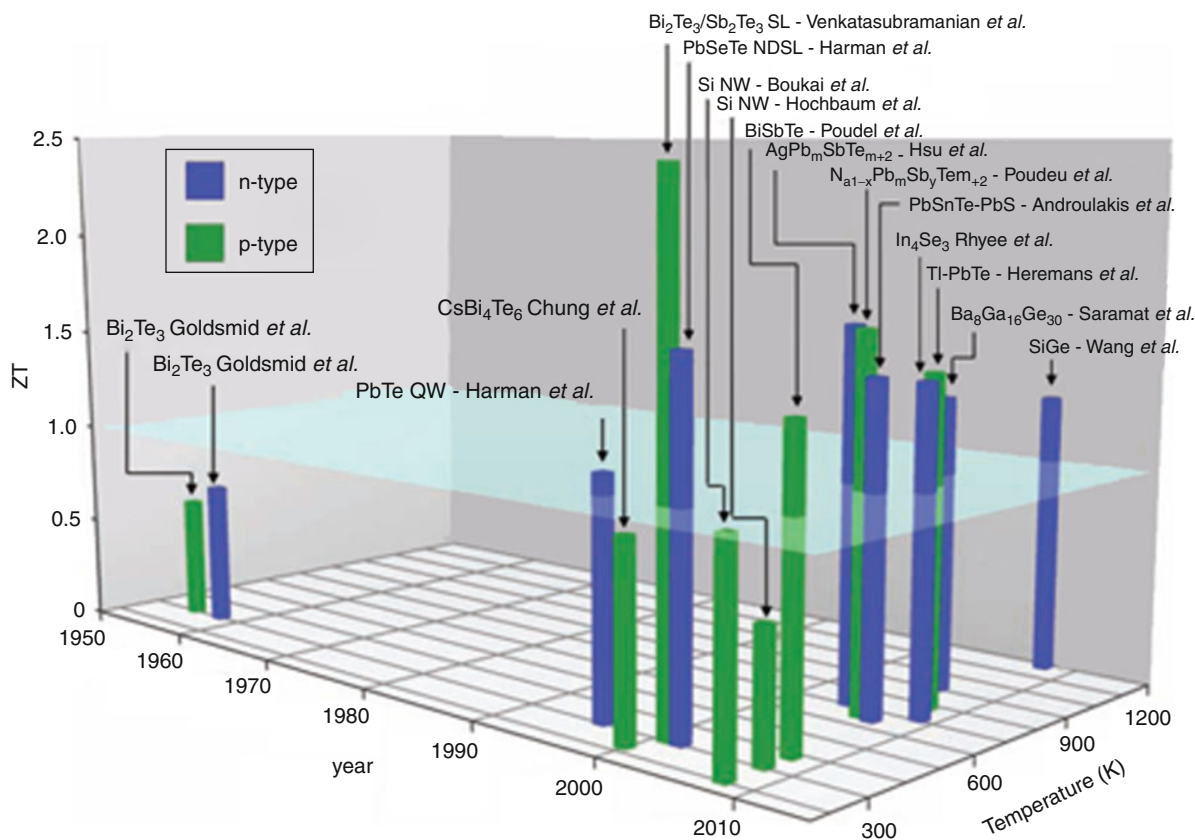
The experimental proof-of-principle studies in “custom nano-engineered” quantum dot and superlattice systems [4] show that (i) enhancement of σ and α can be achieved simultaneously with the reduction of κ in a nanostructured material, (ii) the degrading of the electrical transport property is

minimized by an optimal choice of band offsets in these heterostructures, and (iii) the primary gain in ZT is from the reduction of κ_{ph} , rather than the predicted enhancement of the PF , which has been attributed to the less-dispersed phonon modes or phonon scattering due to interfacial lattice mismatch. To date the record high ZT is still held by superlattice and quantum dot systems, but the present cost of the nanofabrication processes restricts the large-scale commercial use of these “custom nano-engineered” systems.

Theoretical studies of the mechanism of heat conduction in superlattice systems suggest that the thermal boundary resistance and a high density of interfaces, rather than the atomically perfect interfaces or the periodicity of nanostructures, accounts for such a significant reduction of κ_{ph} . It naturally leads to the idea of using bulk nanostructured materials (namely, nanocomposites) as a potential cost-effective and scalable alternative to these “custom nano-engineered” systems for high ZT materials [8]. Important advantages of nanocomposites include being made using a bulk synthesis procedure and in a form compatible with existing commercial TE devices, possessing better mechanical properties, and being more isotropic.

Despite the variety in micro-morphologies and TE properties of superlattice, quantum dot systems, and nanocomposites, the mechanisms underlying the enhancement of ZT are largely the same: While classical size effects limit the mean free paths of both electrons and phonons, quantum size effects may alter the electron band structures or phonon dispersion relations individually. As shown Fig. 2 [6], many advances in the enhancement of ZT are attained in these nanostructured thermoelectric materials (NTMs).

It is not always easy to distinguish classical and quantum effects. For example, the wave-particle dual nature of electrons may manifest itself in nanostructures, where the electron transport is between the coherent transport mode as a wave and the incoherent transport mode as a particle. Nonetheless, the combined and quantum size effects can make the physical properties of a NTM so different from its bulk counterpart(s) that the NTM is eligible to be treated as a *new* material, even they may have the same chemical composition. Si is a good example [5, 6].



Nanostructured Thermoelectric Materials, Fig. 2 ZT of the state-of-the-art TE materials, both n-type and p-type, as a function of temperature and year [6]. Many of recent advances

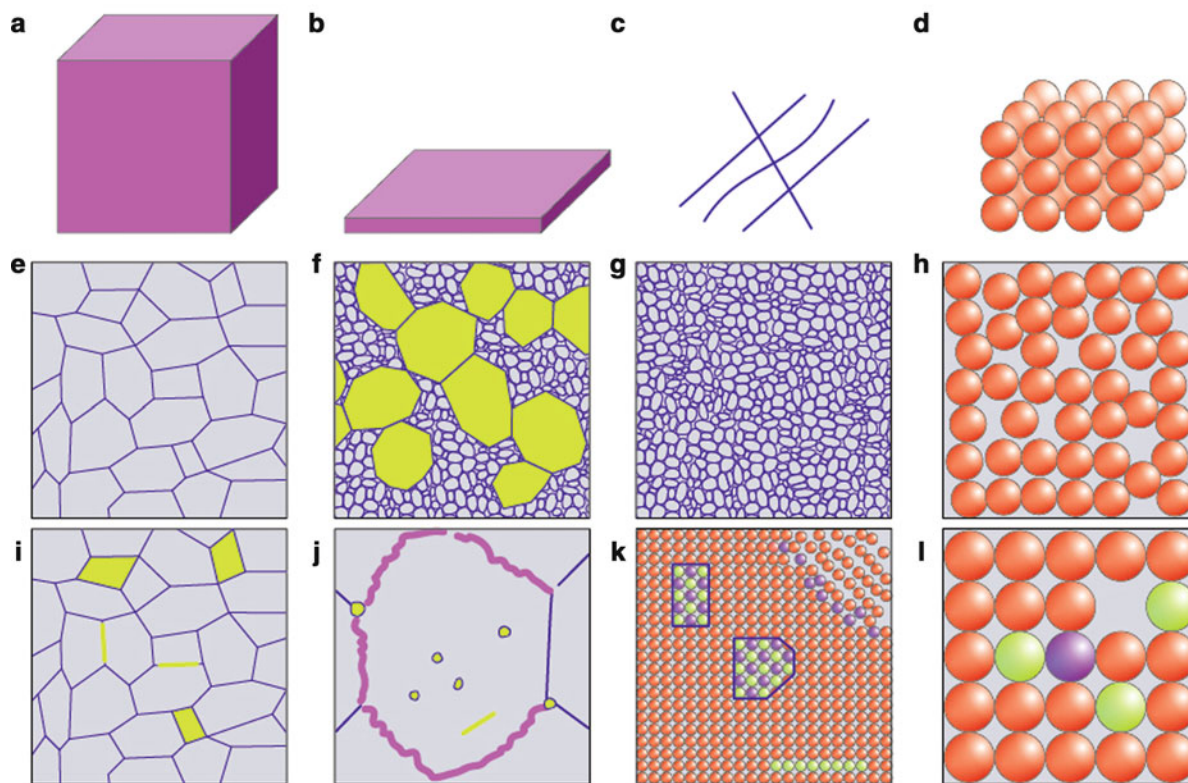
in the enhancement of ZT are achieved in NTMs, while it should be noted that some results have not yet been reproduced or verified

Micromorphology and Preparations of NTMs

The micromorphology of a NTM determines the material's TE properties. Hence creating a controlled micromorphology using a bulk synthesis procedure is crucial for the development of NTMs. As shown in Fig. 3 [9], the micromorphologies of NTM are comprised of three major aspects: (i) the composition, structure and characteristic dimension of nanostructures; (ii) the connectivity of nanostructures (i.e., the topology); and (iii) interfaces [10]. When the characteristic length scale becomes smaller, the aspects (ii) and (iii) often become more important than the aspect (i) in determining the TE properties of the NTMs. In certain cases, interfaces should be treated as an individual phase. Note that the effective medium theory suggests that the ZT of a composite cannot exceed the best performing constituent if there is no contribution from the interfaces.

The recent advent of nanomaterials and nanotechnology provides a variety of ways to prepare NTMs, to name a few: molecular epitaxial growth, laser ablation, chemical vapor transport, vapor-liquid-solid growth, etching, wet chemistry and hydrothermal synthesis, ball milling, mechanical alloying, melt spinning and phase separation. High-energy ball milling is an effective scalable method to produce large amount of nanoparticles of a variety of compositions. Another notable nanostructuring method is melt spinning, which can create rich nanostructures via highly dynamic quenching process. The current focus is mainly on nanostructuring bulk materials with well-known TE properties, such as Bi₂Te₃, PbTe, CoSb₃, and SiGe. The impacts of nanostructuring are thus amenable to a certain degree of analysis, prediction, and optimization.

In particular, nanocomposites can be prepared by means of ex situ or in situ methods. The ex situ approach



Nanostructured Thermoelectric Materials, Fig. 3 Typical micromorphologies of NTMs [9]

is typically performed in a two-step fashion: Nanomaterials are first prepared using the aforementioned methods, and then densified into bulk objects [11]. The two most widely used densification methods are hot pressing (HP) and spark plasma sintering (SPS). The SPS process can densify nanoparticles in a relatively short time, compared to HP process, to avoid the interdiffusion and growth of grains. The main drawback of the ex situ method is the lack of coherence in interfaces, which may severely degrade the carrier mobility. An alternative way is the in situ or self-assembled preparation [12], including precipitation nucleation and growth, spinodal decomposition, and matrix encapsulation (“arrested precipitation”), with the resulting interfaces being mostly coherent or semi-coherent.

Strategies and Implementation of Nanostructuring

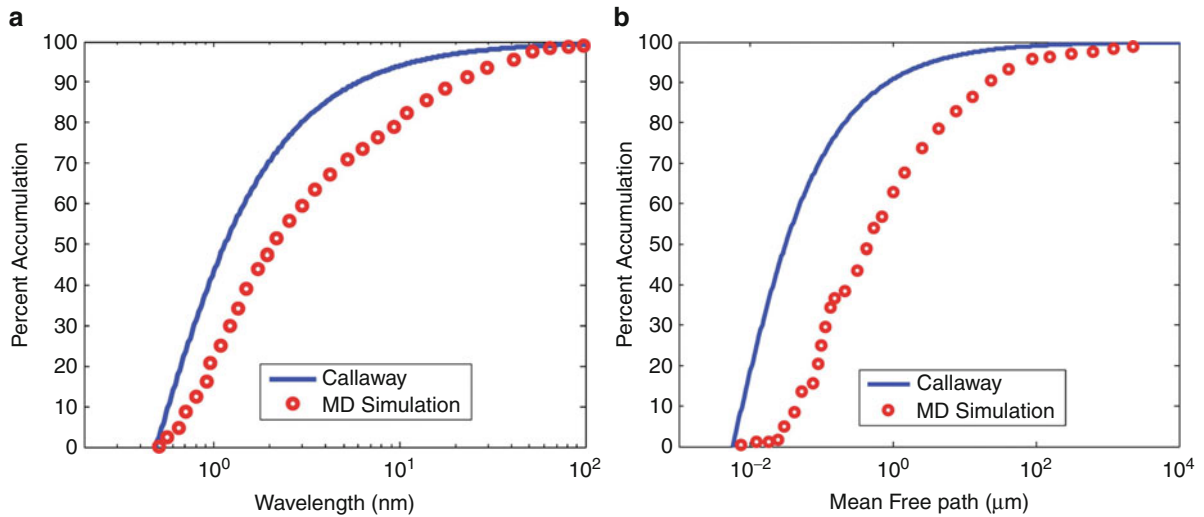
This section is devoted to several important strategies and implementation of nanostructuring, each of which

is directed toward one of those physical transport quantities that determine ZT (Eq. 3). The rationale behind these strategies and implementation is in line with the PGEC concept.

Reducing the Lattice Thermal Conductivity

Although the paradigm of NTMs started with the theoretical speculation that quantum confinement may induce enhanced electrical transport properties, to date almost all high-performance NTMs are a result of a very low κ_{ph} . An interesting question thus concerns the lower limit of κ_{ph} in NTMs.

In a bulk material, the lower limit of κ_{ph} was set by the so-called alloy limit obtained by amorphous solids, in which the phonon mean free path is close to half the phonon wavelength. Historically, the alloying approach was employed to reduce κ_{ph} as the atomic scale lattice imperfections can strongly scatter short-wavelength phonons, but alloying can often lead to deterioration in electrical conductivity. A somewhat oversimplified relationship, $\kappa_{ph} \approx (1/3)(v_s C_v L_{ph})$, is often employed to estimate the phonon mean free path, L_{ph} , where v_s is



Nanostructured Thermoelectric Materials, Fig. 4 Cumulative distribution of κ_{ph} of pure bulk Si with respect to (a) wavelength and (b) mean free path predicted by the Callaway model and a more exact molecular dynamics simulations [13]

the sound velocity and C_V is the heat capacity at a constant volume. This relationship is derived from the classical kinetic theory of heat conduction, regardless of the details of the scattering mechanisms. At temperatures comparable and above the Debye temperature, θ_D , both v_S and C_V are weakly dependent on temperature, κ_{ph} is thus governed by L_{ph} . At $T \ll \theta_D$, the dominant phonon wavelength, λ_d , is on the order of $(\theta_D/T)a$, where a is the interatomic spacing. At temperatures comparable and above θ_D , λ_d is on the order of $2a$. Experimentally, neutron scattering and Raman spectroscopy can provide a quantitative description of the phonon spectrum.

In several NTMs, the κ_{ph} is found to be lower than the “alloy limit.” Phonons in a material have temperature-dependent spectra of wavelengths and mean free paths, all of which contribute to κ_{ph} (Fig. 4, [13]). Hence it is the common belief that *multiscale structural complexity*, from the atomic scale alloying and strain fields to the interfacial boundary scattering peculiar to nanoscale constituents, is required to scatter phonons of a wide range of energies and wavelengths over a wide temperature range. As the mean free path and wavelength of phonons are typically on the nanometer scale, nanostructuring is likely the best option for lowering the κ_{ph} s, and probably the only option.

Reducing the Carrier Thermal Conductivity

Per Eqs. 1–3, the low TE conversion efficiency is due primarily to a low ZT , which stems from irreversible processes such as entropy generation and entropy flow

in the TE process. A so-called energy-specific equilibrium approach [14] has been proposed to minimize κ_e via limiting the irreversible processes. The basic idea is that the electronic efficiency (assuming $\kappa_{ph} = 0$) can reach the Carnot limit when the electron transport mechanism between the hot end and the cold end is reversible diffusion under a finite temperature gradient and a finite voltage drop. The reversibility arises from the situation where a system may have counter-propagating currents in both directions that may result in no net flow of charge while still transporting energy. As such, the Lorenz number in the Wiedemann–Franz relationship, which relates the thermal and electronic conductivities, is minimized. This proposal can be experimentally realized in a nanowire superlattice or quantum dot superlattice system with a carefully spatially graded chemical potential. In each quantum dot, the electrons are distributed in a very narrow energy band (a delta function like single level density of states); the value of energy can be tuned such that the Fermi-Dirac distribution is invariant between the hot end and the cold end.

Suppressing the Bipolar Effect

Many TE materials are degenerate semiconductors or semimetals, in which both electrons and holes contribute to the transport. The bipolar effect is detrimental to ZT for two reasons: First, the presence of Peltier heat flow can take place even when the total current is zero, leading to an extra term in κ_e ; second, the electronic and hole contributions to α are of opposite sign and

therefore cancel each other, often leading to a minimum in α . The bipolar effect can be a significant term at elevated temperatures. The key to suppressing the bipolar effect is to restrict the contribution from the minority carrier. For example, one way is to scatter the minority carrier more effectively than the majority carrier by the grain boundaries via preferential scattering (further discussed in Sect. “[Multifold Preferential Scattering at Interfacial Boundaries](#)”); as a result, the onset of the bipolar effect’s influence has been shifted to a higher T .

Increasing the Power Factor

From the practicality point of view, a $ZT \sim 3$ will make thermoelectricity an important contributor in the areas of waste heat recovery and heat management, a $ZT > 3$ is certainly desired but the return in the conversion efficiency becomes marginal [1]. At present, unless people can employ coherent or correlated phonon scattering effects to actively reduce the group velocity and the number of heat-carrying phonon modes [6], any further reduction of κ_{ph} below the “alloying limit” is not only difficult but also insufficient to reach a $ZT \sim 3$.

Due to the interdependence of σ and κ_e , the enhancement in the PF must come mainly from the enhancement in the thermopower. In general, the magnitude of thermopower is a measure of the entropy per unit charge carrier. For the purpose of illustration, one can estimate the minimum thermopower required for a viable TE material by assuming a hypothetical material in which κ_{ph} is zero and the Wiedemann–Franz relationship holds. As such, Eq. 3 is rewritten as $ZT = \alpha^2/L$, requiring that $\alpha \sim 157 \mu\text{V/K}$ for a $ZT = 1$, $225 \mu\text{V/K}$ for a $ZT = 2$, and $275 \mu\text{V/K}$ for a $ZT = 3$.

Multifold Preferential Scattering at Interfacial Boundaries

There are numerous results underscoring the important role of interfacial boundary scattering [10]. A well-known mechanism is the energy-dependent scattering of electrons, namely, the electron-filtering mechanism [4]. The introduction of appropriate energetic barriers at interfaces restricts the energy of carriers entering a material. At an interface, those carriers with a mean energy substantially above the Fermi level, E_F , will pass through the interface preferentially, thereby enhancing the thermopower that directly depends on the excess energy ($E - E_F$) of charge carriers [7].

This leads to an important strategy that implements multifold *preferential scattering* at interfacial boundaries: scattering phonons more effectively than electrons as for a “PGEC”; and scattering those carriers with a mean energy substantially lower than the Fermi level more effectively than those with higher mean energy to ensure a large thermopower. Certainly, the interfacial boundary may somewhat degrade the electrical conductivity, but the effect should not be significant because most TE materials are heavily doped and the electron mean free path is very short. Such multifold preferential scattering mechanisms can be implemented via controlling the physical parameters of the barrier at interfacial boundaries [10].

Electron Band Structure Engineering Based on semiclassical theory, Pichanusakorn and Bandaru recently proposed that [7] (i) the maximum PF of a bulk material is a function of the reduced Fermi level, an energy-dependent carrier scattering constant, and the dimensionality; (ii) for any bulk materials and at any temperatures, the PF maximizes when the magnitude of thermopower is in the range of 130–187 $\mu\text{V/K}$, while the magnitude of PF can be increased by tuning the detailed material parameters; (iii) in both bulk and nanostructured materials, the PF is eventually limited by the onset of the bipolar conduction; and (iv) the PF of a NTM can be enhanced beyond the bulk value, say, via quantum-confinement-induced enhancement of magnitude of the DOS, rather than the change in shape of the DOS.

Several Transformative Concepts There are several interesting concepts that are viable to implement in conjunction with the nanostructuring approach. In a recent work, Heremans et al. [15] showed that a 2% doping of Tl in bulk PbTe significantly increased ZT from 0.7 to 1.5, which has been attributed to the increase in $\text{DOS}(E_F)$ induced by resonant levels. Also in the PbTe system, Pb and Sb co-nanostructuring was found to significantly enhance the electron mobility and lead to a $ZT \sim 1.4$, while this phenomenon was not observed with Pb or Sb precipitates alone [8], the underlying mechanism are not clear. Another promising class of materials are the strongly correlated semiconductors, in which the virtual excitations from d - or f -levels into the conduction band give rise to a sharp feature in the $\text{DOS}(E_F)$, and a giant thermopower at low temperatures. For example, despite a very high κ_{ph} , FeSb_2 exhibits a colossal thermopower and

a record high PF at low temperatures [16], the PF is about 65 times higher than that of Bi_2Te_3 . The high PF at low temperatures makes these materials prime candidates for transient TE cooling (supercooling in a Peltier cooler). There are currently very few materials, such as BiSb and CsBi_2Te_3 , available for cooling in the temperature range below room temperature [1].

Challenges in Nanostructured Thermoelectrics

The effectiveness of “nanostructuring” as a means to improve ZT has been corroborated by numerous studies. While a systematic search and subsequent thorough investigation of promising materials may eventually yield the desired NTMs, the present trial-and-error approach is not sufficient, and the concerted efforts of theoreticians and experimentalists are needed to address the key question as to why certain nanostructures are thermoelectrically favorable while others are not. Some of the challenges will be briefly addressed in this section.

Understanding the Phonon and Electron Transport

Current understanding of phonon and electron transport in NTMs is fairly limited. This lack of understanding affects the ability to improve and design higher performance NTMs. For example, although it is experimentally shown that NTMs are able to exhibit κ_{ph} lower than the “alloy limit,” and the multiscale structural complexity can suppress κ_{ph} by strongly scattering a wide spectrum of phonons, the existing models largely fail to provide satisfactory predictions for the spectral dependence of κ_{ph} or the phonon transmissivity at interfacial boundaries. A pertinent question is what is the lower limit of κ_{ph} ? [6, 13]. There is also much work to be done on understanding electron transport in NTMs. The primary difficulties are (i) the standard analysis based on the Boltzmann transport equation loses its validity when the characteristic length scale of NTMs is comparable or less than the electron de Broglie wavelength; and (ii) modeling the electron transport across the interfacial boundaries. Models based on molecular dynamics and nonequilibrium Green’s functions are promising options.

Creation and Evolution of Nanostructures

At present, the factors that control the nanostructure formation and evolution in solids are not well

understood. More importantly, nanostructures are a metastable state of the material. As NTMs will work under high temperatures and temperature gradients, it is important to study and optimize their long-term stability, especially the impacts of interdiffusion and coarsening on the TE performance of the NTMs. It is noted that the nanoinclusions in the $\text{AgPb}_m\text{SbTe}_{m+2}$ (LAST-m) NTMs are thermodynamically and kinetically stable to at least 800 K over a long period of time [8].

Cost of Raw Materials

Even with the desired improvement in ZT , a TE device can only compete with the equivalent mechanical systems in some decentralized applications with a power scale from μW to kW . After performance, the cost of raw materials and fabrication is a major consideration for any large-scale industrial application of NTMs. For example, the price of tellurium (Te) has significantly increased in the past few years due largely to the use of Te in solar cells. This somewhat vitiates the enhancement of ZT enhancement in NTMs and especially limits the applications of Te-containing NTMs such as those based on Bi_2Te_3 and PbTe . High-performance NTMs based on abundantly available and nontoxic elements are crucial for scaled-up applications of NTMs in power generation and refrigeration. Figure 5 presents a histogram of the relative abundance of the elements [17].

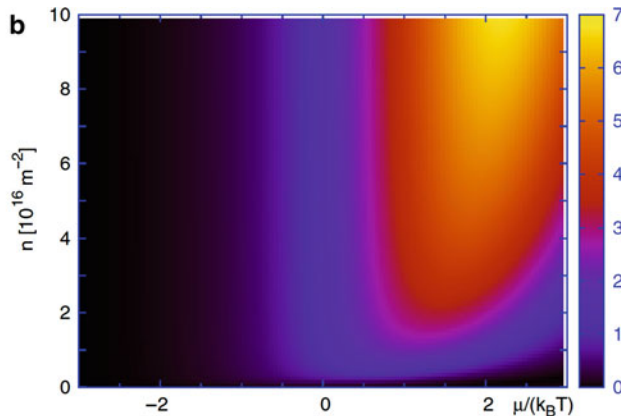
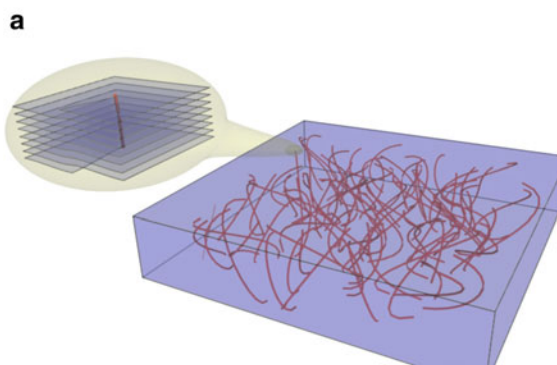
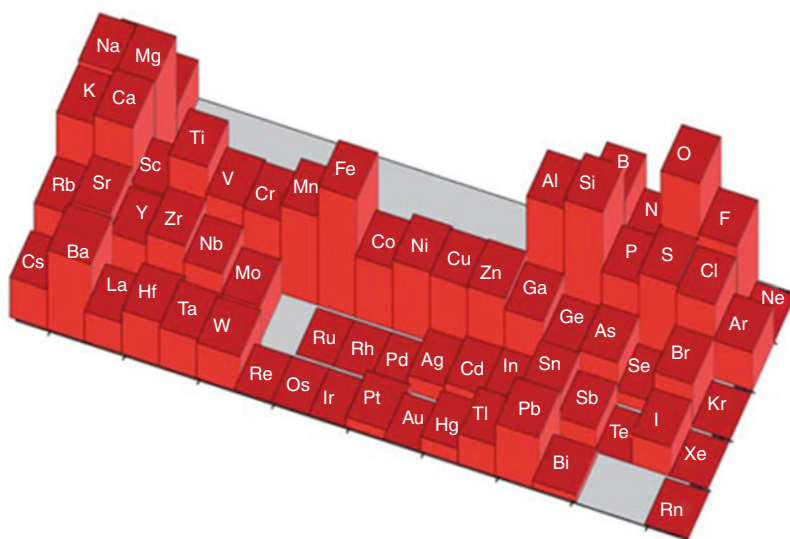
New Concepts

The advance in thermoelectrics studies hinges primarily upon the discovery of novel materials and novel concepts that implement decoupling the interrelated TE quantities (σ , α , and κ) in order to optimize the charge flow and entropy flow (the magnitude of α is a measure of the entropy per charge carrier) in the material.

Spin Caloritronics

The charge and lattice are the two degrees of freedom most relevant to the conventional thermoelectricity while the role of the spin degree of freedom has received much less attention with the exception of the layered cobaltates. A notable new area is spin caloritronics, basically a combination of spintronics and thermoelectrics. While the central theme of

Nanostructured Thermoelectric Materials, Fig. 5 Relative abundance of the elements in a logarithmic scale [17]



Nanostructured Thermoelectric Materials, Fig. 6 (a) Screw-type lattice dislocations (shown in the inset) form topologically protected 1-D conduction paths in a 3-D topological

insulator. (b) The contour plot of room temperature ZT as a function of chemical potential, μ , and the density of lattice dislocations, n , for a sample's length of 1 μm [20]

thermoelectrics is controlling charge flows by means of heat flows and vice versa, that of spintronics is the active manipulation of the population and the phase of spins, and that of spin caloritronics is the controlling of spin currents by means of heat flows and vice versa. The recently discovered spin Seebeck effect is conceptually different from the traditional charge Seebeck effect in that the spin Seebeck effect is not related to the charge-particle diffusion [18]. Uchida et al. [19] proposed an exciting possibility of using the spin Seebeck effect in magnetic insulators, in combination with the inverse spin Hall effect, to build TE devices “operating on an entirely new principle.”

Topological Insulators

Beyond the local order parameter and long-range correlation in the Landau symmetry-breaking description, topological order results from long-range quantum entanglement of quantum states and is described by a new set of quantum numbers (variables), such as ground state degeneracy, quasiparticle fractional statistics, edge states, and topological entropy. Topological insulators (TIs) have recently attracted great attention in condensed matter physics. Among the state-of-the-art bulk TE materials, Bi_2Te_3 , (Bi,Sb) and certain half-Heusler compounds have been recently identified as TIs. With the bulk insulating and topologically protected surface conducting nature, TIs have been suggested to be promising candidates for

high-performance TE materials. As shown in Fig. 6, a possible approach is to fabricate TI nanocomposites to combine enhanced electric conductivity and thermopower from the topologically protected 1-D channels (engineered line dislocations) and suppressed thermal conductivity (due to the vast phonon scattering at the interfaces and the insulating bulk states) [20].

Perspective Remarks

As a fundamental phenomenon discovered in the nineteenth century, thermoelectricity has been widely used in thermocouples for temperature measurements, but it has only found some *niche* applications in power generation and refrigeration in the twentieth century. The current state-of-the-art bulk TE materials such as Bi₂Te₃, PbTe, CoSb₃, and SiGe have been in the place for 20–30 years; expansion of thermoelectricity into broader applications is limited by the low *ZT* of materials. This could soon change due to the progress in the emerging field of NTMs in the past decade. Quantum and classical size effects in NTMs offer a means to tune the thermoelectric conversion efficiency beyond that of bulk materials. It is plausible to expect the recent advances in NTMs, especially the creative preparation methods, new materials, novel concepts, and better understanding, will lead to eco-friendly, cost-effective, higher performance thermoelectricity and a broader application of TE power generation and refrigeration.

Acknowledgments M.H. Z. and J. H. acknowledge Drs. Don Liebenberg and Malcolm Skove for useful discussions, the support of the DOE/EPSCoR Implementation Grant (#DE-FG02-04ER-46139), and the SC EPSCoR cost-sharing program.

Cross-References

- ▶ [Self-assembly of Nanostructures](#)
- ▶ [Thermal Conductivity and Phonon Transport](#)
- ▶ [Thermoelectric Heat Convertors](#)

References

1. Nolas, G.S., Sharp, J., Goldsmid, H.J.: Thermoelectrics basic principles and new materials developments. Springer, Berlin/Heidelberg (2001)
2. Vining, C.B.: Semiconductors are cool. *Nature* **413**, 577–578 (2001)
3. Hicks, L.D., Dresselhaus, M.S.: Effect of quantum-well structures on the thermoelectric figure of merit. *Phys. Rev. B* **47**, 12727–12731 (1993)
4. Dresselhaus, M.S., Chen, G., Tang, M.Y., Yang, R., Lee, H., Wang, D., Ren, Z.F., Fleurial, J.-P., Gogna, P.: New directions for low-dimensional thermoelectrics. *Adv. Mater.* **19**, 1043–1053 (2007)
5. He, J., Tritt, T.M.: Thermal-to-electrical energy conversion from the nanotechnology perspective. In: Garcia-Martinez, J. (ed.) *Nanotechnology for the energy challenge*, pp. 47–78. Wiley-VCH, Weinheim, Germany (2010)
6. Vineis, C.J., Shakouri, A., Majumdar, A., Kanatzidis, M.G.: Nanostructured thermoelectrics: big efficiency gains from small features. *Adv. Mater.* **22**, 3970–3980 (2010)
7. Pichanusakorn, P., Bandaru, P.: Nanostructured thermoelectrics. *Mater. Sci. Eng. Rep.* **67**, 19–93 (2010)
8. Kanatzidis, M.: Nanostructured thermoelectric: the new paradigm? *Chem. Mater.* **22**, 648–659 (2010)
9. Li, J., Liu, W., Zhao, L., Zhou, M.: High-performance nanostructured thermoelectric materials. *NPG Asia Mater.* **2**(4), 152–158 (2010)
10. Medlin, D.L., Snyder, G.J.: Interfaces in bulk thermoelectric materials: a review for current opinion in colloid and interface science. *Curr. Opin. Colloid Int. Sci.* **14**, 226–235 (2009)
11. Lan, Y., Minnich, A.J., Chen, G., Ren, Z.F.: Enhancement of thermoelectric figure-of-merit by a bulk nanostructuring approach. *Adv. Funct. Mater.* **19**, 1–20 (2009)
12. Sootsman, J.R., Chung, D.-Y., Kanatzidis, M.G.: New and old concepts in thermoelectric materials. *Angew. Chem. Int. Ed.* **48**, 8616–8639 (2009)
13. Minnich, A.J., Dresselhaus, M.S., Ren, Z.F., Chen, G.: Bulk nanostructured thermoelectric materials: current research and future prospects. *Energy Environ. Sci.* **2**, 466–479 (2009)
14. Humphery, T.E., Linke, K.: Reversible thermoelectric nanomaterials. *Phys. Rev. Lett.* **94**, 096601(1)–096601(4) (2005)
15. Heremans, J.P., Jovovic, V., Torberer, E.S., Saramat, A., Kurosaki, K., Chroenphaskdee, A., Yamanaka, S., Snyder, G.J.: Enhancement of thermoelectric efficiency in PbTe by distortion of the electronic density of states. *Science* **321**, 554–557 (2008)
16. Bienten, A., Johnsen, S., Madsen, G.K.H., Iversen, B.B., Steglich, F.: Colossal Seebeck coefficient in strongly correlated semiconductor FeSb₂. *Euro. Phys. Lett* **80**, 17008(1)–17008(5) (2007)
17. Vaquero, P., Powell, A.V.: Recent developments in nanostructured materials for high-performance thermoelectrics. *J. Mater. Chem.* **20**, 9577–9584 (2010)
18. Jaworski, C.M., Yang, J., Mack, S., Awschalom, D.D., Heremans, J.P., Meyers, R.C.: Measurement of the spin-Seebeck effect in a ferromagnetic semiconductor. *Nat. Mater.* **9**, 898–903 (2010)
19. Uchida, K., Xiao, J., Adachi, H., Ohe, J., Takahashi, S., Leda, J., Ota, T., Kajiwara, Y., Umezawa, U., Kawai, H.,

- Bauer, G.E.W., Maekawa, S., Saitoh, E.: Spin Seebeck insulator. *Nat. Mater.* **9**, 894–897 (2010)
20. Tretiakov, O.A., Abanov, Ar, Murakami, S., Sinova, J.: Large thermoelectric figure of merit for three-dimensional topological Anderson insulators via line dislocation engineering. *Appl. Phys. Lett.* **97**, 073108(1)–073108(3) (2010)

Nanostructured Thermoelectrics

► Nanostructured Thermoelectric Materials

Nanostructures Based on Porous Silicon

Luca Boarino and Giampiero Amato
Quantum Research Laboratory, Electromagnetism
Division, Istituto Nazionale di Ricerca Metrologica,
Torino, Italy

Definition

The term Porous Silicon (or, the more general, Porous Semiconductor) has been and is largely used to define a class of semiconducting material in which pores are opened, basically by means of electrochemical or chemical processes. According to the present definition (and to IUPAC), three different material types are identified, say, nanoporous, mesoporous, or macroporous. The important point is that these types are defined according to the dimension of the pores, instead of the residual structures.

In practice, a correlation between pore diameters and structure sizes is found, e.g., ensembles of macropores and nanostructures are found very seldom, and are the product of more complicated technological steps, not the simple (electro)chemical etching.

The term Porous Silicon has been extended in the past to define some kinds of disordered arrays of Silicon nanowires. This extended definition does not keep into account for some regular macroporous structures obtained by means of electrochemical etching coupled with photolithography. Moreover, disorder is a typical consequence for the (electro) chemical processes employed, and cannot be considered as distinctive for the material. In other words, nanostructured arrays can

be produced in many different ways, not only (electro) chemical ones, and ordered (electro) chemical etching is in principle possible.

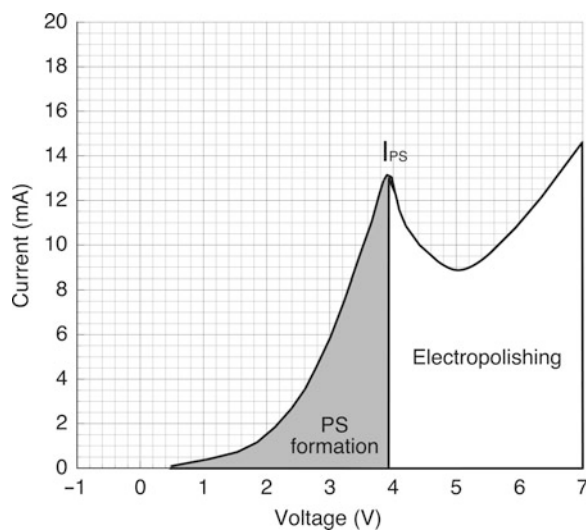
Overview

Historical Notes

Porous silicon (PS) was accidentally discovered by Arthur and Ingenborg Uhlir at the Bell Labs, in 1956, in the framework of a research work on silicon and germanium electropolishing. If the electropolishing current was not reached for the electrochemical etching of the silicon wafer, instead of a clean and polished surface, a brown or black layer appeared, due to a nonuniform dissolution of silicon atom (see Fig. 1). The material is porous, full of fine holes, and the electropolishing regime is reached only at the pore tips in these conditions. This former evidence of PS was published in the Bell Lab's technical notes [1] and the study of this system abandoned until the 1970s–1980s, when the huge specific surface of the material, of the order of hundreds of square meters per square centimeter gave origin to spectroscopic studies, as a precursor for thick oxide growth on silicon and to the first applications to gas sensing. In 1977 the first patent of a humidity sensor based on porous silicon was deposited [2]. PS was also applied to Silicon On Insulator (SOI) technology [3].

At the end of the 1980s, Leigh T. Canham working at the British Defense Research Agency obtained the experimental confirmation of his intuition that the nanofilaments of porous could exhibit good luminescence properties, due to quantum confinement effects. Also other authors were investigating on the phenomenon, V. Lehmann and U. Gösele, but Canham was the first to publish, in 1990, the experimental evidence of PS room temperature photoluminescence [4, 5]. This experiment stimulated the work of many research groups all over the world in the nonlinear optical and electrical properties, and the number of published papers arose exponentially from 1991 to 1995. The premises of a brilliant future for silicon-based lasers were not maintained, due to the low efficiency of electroluminescence of this material, slow decay times and to the chemical and mechanical stability.

In the mid 1990s, the main applications of porous silicon were micromachining, gas sensing and with



Nanostructures Based on Porous Silicon, Fig. 1 Porous Silicon formation voltammogram. Over the I_{PS} , electropolishing regime occurs. At lower currents, porous silicon is formed

1995 the medical and biological applications received a big boost, due to the demonstration of hydroxyapatite growth on porous silicon and to the proven biocompatibility and hemocompatibility of nanostructure silicon, again by L. T. Canham [6].

Other studies showed that B50 rat hippocampal cells have a better adhesion onto porous silicon respect to untreated surfaces, so indicating suitability of porous silicon for cell culturing purposes.

In 2001 a group of researchers at the Technical University of Munich [7] announced that hydrogenated porous silicon, in contact with liquid oxygen, reacts explosively, releasing more energy than an equivalent amount of TNT with higher speed, and with no necessity of a detonating material also for small quantities. The application of this discovery, substituting the liquid oxygen requiring very low temperatures with salts rich in oxygen, is in the air bag igniters in the automotive field, in the satellite propulsion, and in atomic emission spectroscopy [8].

Fabrication of Porous Silicon

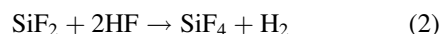
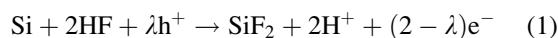
Anodization

Porous silicon is realized by electrochemical anodization of a silicon substrate, in a solution of HF and Ethanol in different ratio, depending on the silicon wafer resistivity on the desired etch rate and thickness and characteristics of the porous layer. The cathode is a platinum wire or a grid; the beaker can be Teflon or

PVC (Fig. 2). To obtain homogeneous and thick layers, a pulsed bias is preferable, so giving time to the solution at the pore tips to recover its title by Fluorine ions diffusion from the bath.

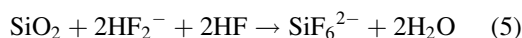
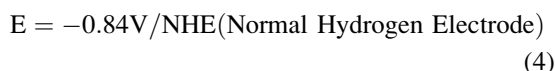
The ethanol is added to the hydrofluoric acid to favor the hydrogen bubbles removal from the porous silicon layer during formation.

Formation mechanisms The selective removal of the silicon atoms (local electropolishing at the pore tips) or complete electropolishing is performed by the Fluorine ions in presence of two or four holes in the bulk of the semiconductor, allowing the breaking of the Si-Si backbonding (Fig. 3). After the initial work of Turner, Memming, and Schwandt proposed the overall dissolution reaction:



Where h^+ is the hole, e^- the electron, and λ is the number of charges exchanged.

Similar mechanisms have been proposed in literature, including silicon oxide formation in presence of H_2O and dissolution in presence of HF, like in the case of the electropolishing regime:



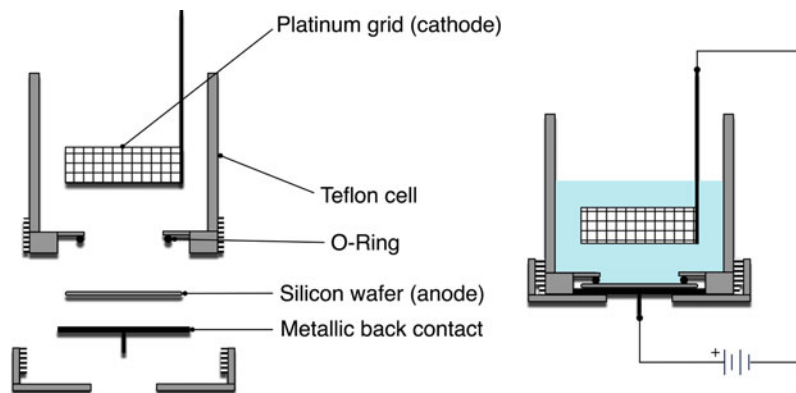
Stain Etching

Porous silicon formation occurs also in absence of an external biasing, by stain etching with hydrofluoric acid, nitric acid, and water.

The vacancies necessary to silicon dissolution in this case are provided by HNO_3 , the etchant concentrations are generally higher, and the etch rates are lower than in the electrochemical process. The first studies were carried on during the 1960s to probe p-n junction, then the discovery of room temperature photoluminescence of porous silicon obtained electrochemically renewed the interest for stain etching, and

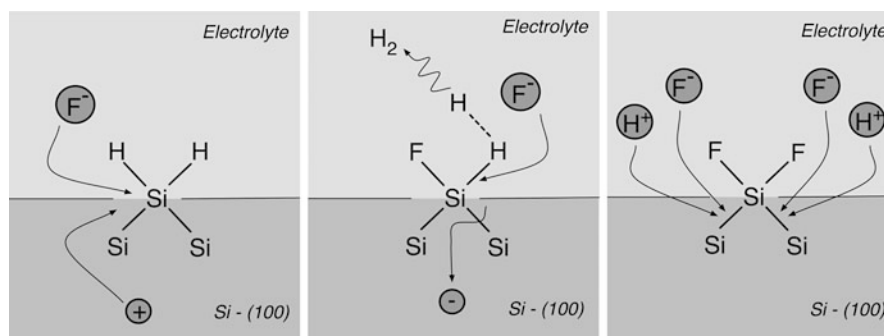
Nanostructures Based on Porous Silicon,

Fig. 2 Electrochemical cell scheme for Porous Silicon fabrication, (exploded, *left*, and assembled, *right*)



Nanostructures Based on Porous Silicon,

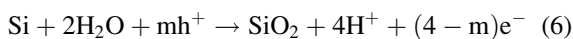
Fig. 3 Dissolution sequence of silicon in presence of HF and holes



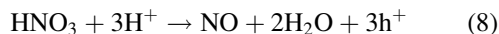
a systematic study regarding the structural and optical properties of porous films have been exploited [9].

The reactions are the following:

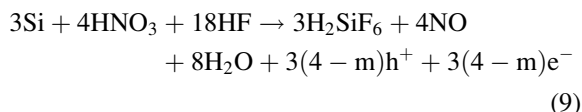
Anode:



Cathode:



Overall:



Recently, K. Kolasinsky [10] developed a new type of PS stain etching oxidants like, Fe(III), Mn(VII), and Cr(VI) drastically reducing hydrogen bubbling so obtaining very homogeneous and highly luminescent materials.

Metal-Assisted Etching

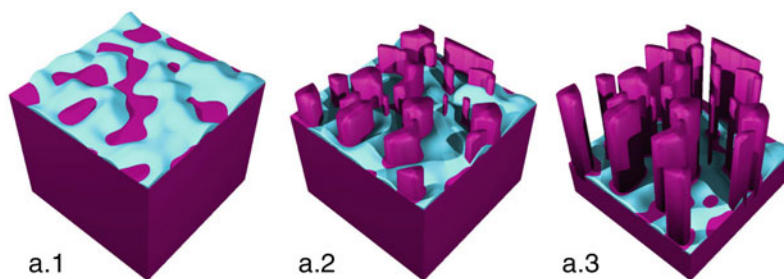
Metal-assisted etching (MAE) has been recently reported, by Li, Bohn and coworkers [11], it is a sort of self-catalytic etching of silicon, with no need of polarization, since the formation mechanism is ruled by local redox potential of the metal respect to silicon. This drives the holes necessary to dissolution at the metal-semiconductor interface. Noble metals like gold, silver, or platinum can be used and obtained by precipitation starting from liquid metallic salts like AgNO_3 and others, but also deposited in form of thin solid films by sputtering or evaporation.

The metal clusters at the surface of the silicon wafer, rapidly sink following the crystallographic directions, and define silicon walls or wires (Fig. 4), depending on the metal particle distribution at surface, on the temperature and concentration of the etchants. When the metal is patterned on the surface, MAE leads to regular morphologies.

The dissolution of silicon in HF and an oxidizing agent is a combination of chemical and electroless process, having strong similarities with porous silicon formation or electropolishing by stain etching in HF-HNO₃ [12].

Nanostructures Based on Porous Silicon,

Fig. 4 Dissolution sequence of silicon metal assisted etching. In *light blue* the silver clusters are depicted, for different etching times (from a.1 to a.3)



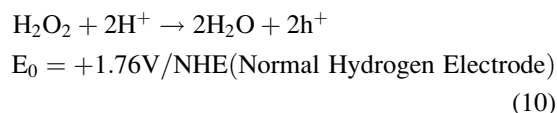
In electroless metal-assisted etching, holes are injected in silicon by H_2O_2 , independently from the original substrate doping. In this case, the oxidizing agent concentration can be considered equivalent to current density in electrochemical process. HF concentration rules dissolution and surface chemistry at the same way for both the processes. The ratio between oxidizing agent and HF deeply affects the etching regimes and the morphologies.

The reaction mechanisms involved in the electroless dissolution of silicon in presence of metal particles are rather complex and will be given in summary following the work of Chartier and coworkers [13] for the case of Ag on Si.

The Ag metallic nuclei at the silicon surface are charged of electrons, and for charge compensation, localize holes at silicon-silver interface, in majority coming from H_2O_2 dissociation at cathode in presence of catalyzing Ag, as shown in Eq. 10 and forming SiO_2 or removing Si at anode, as in Eq. 11.

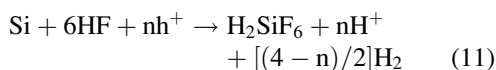
Cathodic reaction

The Ag particles at cathode, catalyze the reduction of H_2O_2 as in Eq. 3 taking advantage from a favorable redox potential, and enhancing the etch rates by supplying free holes in proximity of Si-Ag interface.



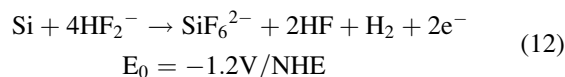
Anodic reaction

In analogy with the well-known chemical dissolution of Si in HF– HNO_3 [14], summarizing the different mechanisms proposed in literature and following the work of Chartier, the following anodic reaction is proposed:



The general form is valid for the two already cited electrochemical regimes, porous silicon formation and electropolishing, where n , ranging from 2 to 4, is the number of holes per dissolved Si atom.

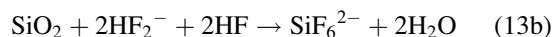
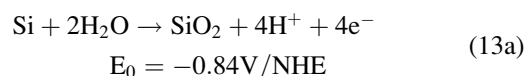
In the case of $n = 2$, porous silicon formation occurs, with the following overall dissolution anodic reaction:



also valid for HF/ HNO_3 stain etching well below the critical current density of electropolishing (J_{PS}).

In the case of $n = 3$, $\frac{1}{2} \text{H}_2$ per dissolved silicon atom, as in the case of 6 M HF – 6 M HNO_3 solutions or in stable electrochemical dissolution at J_{PS} [15].

When $n = 4$, the electropolishing regime occurs, as in the case of low HF/ HNO_3 ratio or for current densities higher than J_{PS} , with a different mechanism involving the formation and dissolution of intermediate species like SiO_2 and without H_2 formation (Eqs. 13a and 13b).



The local mechanism of silicon dissolution in proximity of Ag particle is still unclear, but the pore morphology indicates hole consumption close to the silicon/metal clusters interface, also thanks to the catalytic reduction of H_2O_2 .

This phenomenon can be ascribed to a faster surface reaction with respect to holes diffusion due to coupling between silicon valence bonds and energy levels of silver, or a depletion layer around Ag nanoparticles allowing an easier access of holes in proximity of Ag/Si interface with respect to HF/Si interface to be produced.

Classification of Porous Silicon

Porosity

Porosity is defined as the volume fraction between silicon structures and voids in the porous layer, and is determined by gravimetry, weighting the silicon wafer before the etching (m_1), after the etching and drying (m_2), and after PS removal by alkaline solution (m_3) as in Fig. 5. Porosity is a number between 0 and 1, and is obtained by the formula:

$$P(\%) = (m_1 - m_2)/(m_1 - m_3)$$

The most critical parameters affecting PS structure and porosity are current density and HF concentration. Porosity can range from 4% of macroporous silicon to 95% for mesoporous layers. Porosity is only an indicative parameter and it does not yield information regarding the microstructure of the PS layer, structure size, and relative size distributions are the much more significant for the prediction and interpretation of physical phenomena like optical, thermal, and electrical response of the system.

From gravimetry measurements, also the thickness of the porous layer can be determined using the following formula:

$$W = (m_1 - m_3)/S \times d$$

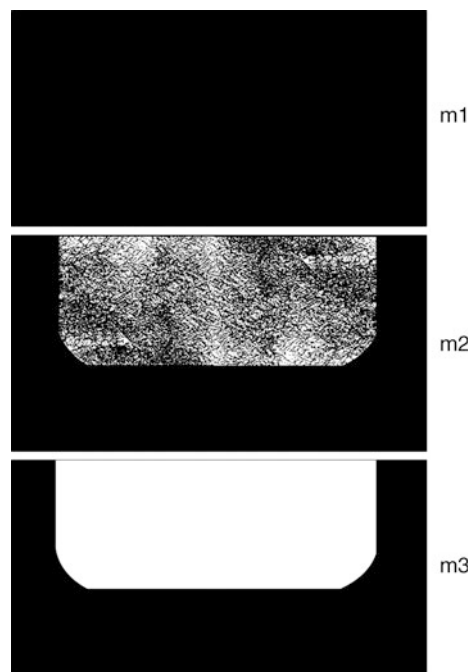
Where d is the silicon density and S the silicon surface exposed during etching. Other destructive methods to evaluate thickness are SEM sections, profiling after the porous layer removal, and among the nondestructive methods is ellipsometry, even if the final result is dependent on the adopted model.

Pore Size, Morphology, and Orientation

According to an IUPAC convention [16], Porous silicon has been divided into three categories based on the size of its pores; macroporous, mesoporous, and microporous.

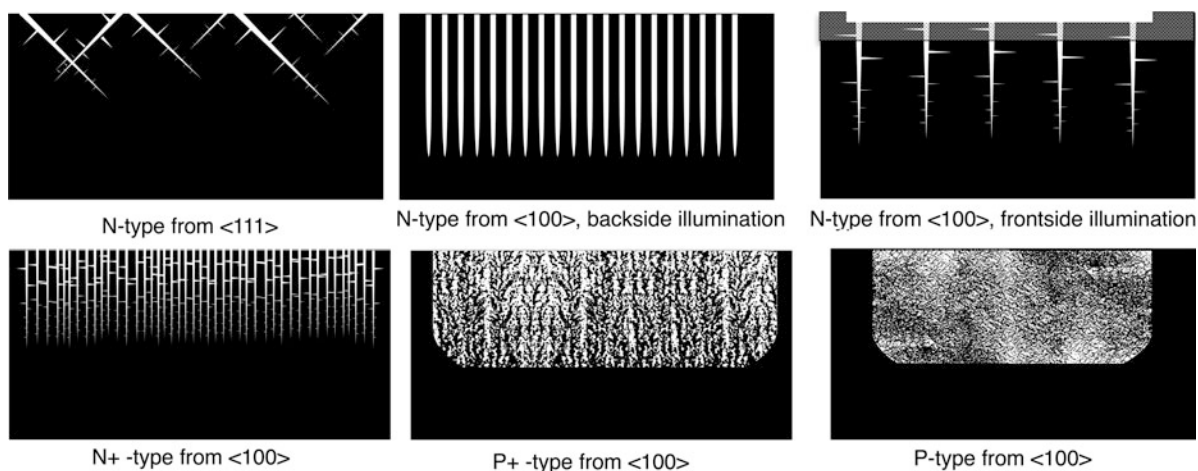
Type	Microporous	Mesoporous	Macroporous
Pore width (nm)	Less than 2	Between 2 and 50	Larger than 50

The pores are strongly dependent from the substrate characteristics, like crystalline orientation and impurity concentration. In general, pores initiate at the wafer surface, and tend to follow the crystalline



Nanostructures Based on Porous Silicon, Fig. 5 The three weights used in gravimetry for porosity evaluation. The first, m_1 , is the weight of the silicon sample before anodization, the second, m_2 , is the weight of the sample after PS formation and at least 20 min of drying, the third, m_3 , is the weight of the sample after alkaline removal of the PS layer

directions, an example for different orientations is shown in Fig. 6. Since the electrochemical etching is dependent on the number of holes available in the substrate, PS morphology is deeply affected by the original impurity content of donors or acceptors in the substrate. In low resistivity N-type silicon, no etching occurs in absence of light or high bias. In presence of light, if the energy $h\nu$ of the radiation is greater or equal the silicon band gap, 1.12 eV at room temperature, the energy excess is transferred to the crystal lattice by rapid thermal processes on the timescale of picoseconds, and the amount equal to the silicon band gap creates an electron-hole couple in Conduction Band and Valence Band, respectively. The holes can so contribute to silicon dissolution, in presence of Fluoride ions of the solution, removing the silicon-silicon bonding. If the illumination is from the top of the silicon surface, form the same side of the etching solution, the pores will initiate from the top and will follow the 100 direction, forming regular or branched macropores, depending on parameters like electric field and free carriers concentration available, but also



Nanostructures Based on Porous Silicon, Fig. 6 Some PS morphologies: from left to right, first row: N-type PS silicon from (111) substrate orientation, N-type PS silicon from (100) substrate orientation and backside illumination, N-type PS silicon from (100) substrate orientation and front illumination. Second row: N + type PS silicon from (100), P + type PS silicon from (100), P-type PS silicon from (100)

a microporous phase, due to the quantity of holes generated by light. The depth of this layer depends from the spectral component of the lamp. If the light reaches the substrate from the bottom, the holes will drift through the silicon substrates, reaching the pore pits at the upper interface with HF and ethanol. There is a high probability that the holes will reach the pore tips, so contributing to maintain a stable and regular pore growth.

In N-type macroporous silicon pore etching, there are several critical parameters to be monitored for a regular and sufficiently fast etching, temperature, current density, electrode potential, and illumination intensity.

For the current density, galvanostatic or potentiostatic conditions can be used. Typically constant current conditions (i.e., galvanostatic) are adopted, so allowing the formation of pores under constant dissolution regime.

So regular macropores on n-type silicon with backside illumination can be controlled in two ways: (1) By using galvanostatic/potentiostatic control maintaining the illumination intensity constant, (2) Fixing current and potential and adjusting the illumination intensity via feedback loop.

The highly doped N-type and P-type silicon give origin to mesoporous structures. In the case of P-type the great availability of free holes due to doping and the low resistivity of the material give origin to high etch rates, of the order of hundreds of nanometers per second, depending on the current density. A typical

value can be 300 nm/s with 300 mA/cm² in HF:Eth 1:1 starting from HF 49%. Mesoporous silicon has an intrinsic anisotropy. The material shows an interconnected columnar structure, with structure size distribution extending from 1.5 to 15 nm [17]. This material is scarcely luminescent, it is mechanically robust and its morphology is determined by the presence of impurities, since the etching does not remove selectively them (e.g., Boron) but the anodization deactivates the trivalent atoms leaving them in subsurface position [18].

For the highly doped N-type silicon the mechanism is rather different. The right branch of the curve is determined by the applied current density, and the regime is electrochemical, the left branch of the curve, is more a chemical dissolution, with slow etch rates, and so reaching high porosities.

Low resistivity P-type silicon gives origin to a tiny coral structure, with typical sizes around 1–2 nm. These PS layers are fragile and show room temperature photoluminescence.

Optical Properties and Photoluminescence

Generally, nano-PS emits bright visible light at room temperature, in spite of the fact that, the pristine c-Si is a well-known indirect-gap semiconductor, with a fair emission in the near-IR, at cryogenic T.

The application of Photoluminescence (PL) technique in PS involves several different physical processes, which make the final product (the PL spectrum) of difficult interpretation. Absorption of the incident photon creates an electron-hole pair. If nanocrystallites are somehow interconnected, or imbedded in a medium, these particles can escape through different mechanisms. If not, they can decay to lower states *in the same nanocrystal*, or could recombine radiatively, directly from the excited state. If localized states are present, trapping into these shallow states can occur. On the contrary, deep states, e.g., unpassivated dangling bonds, can enable the competing non-radiative recombination of electron-hole pairs.

After several years of investigation, there is a general consensus about the origin of such strong light emission. Resonant PL technique proves that emission arises in the crystalline core of the material (instead of other Si-O, Si-H, or siloxene compounds) [19] where the energy gap is broadened thanks to Quantum Confinement (QC) effects. On the other hand, the efficient passivation of the non-radiative pathways (due to H-, or O- termination of the Si nanostructures) plays an important role.

PL decay times are in the order from μs to ms, which makes practical application of PS in integrated optoelectronics unlikely. In particular, the red PL shows non-exponential decay, with typical times ranging in the μs and ms range at room T and at low T, respectively. The competition between radiative, and non-radiative recombination processes depends on T. Generally, the PL intensity reaches a maximum around 50–200 K (depending upon the sample), the simple meaning being that one of the competing processes, say, the radiative one, prevails at those T values. The T dependence of the PL lifetime is the consequence of exchange splitting of excitons. Indeed, the electron and hole, with spin 1/2, can be combined into a singlet ($S = 0$) and a triplet ($S = 1$) state, the exchange interaction stabilizing the triplet. While at high T both of them are populated, the PL arising from the “fast” decay of the singlet state, at low T the PL becomes much slower because the triplet state is virtually stable. In the T range at which the PL emission is maximum, the radiative process is still related to population of singlet states, whereas the non-radiative one is quenched. When T is further decreased, radiative

processes are now related to the slow decay of triplet states, and non-radiative ones become competing again.

PS Structures

PS has been fruitfully employed as a sacrificial layer since it is a strongly reactive material, and easy to etch [20].

Generally, a wide class of materials, from metals to semiconductors and insulators can be deposited onto PS. Due to the intrinsic reactivity of PS, it can be easily etched, giving rise to large undercutting of the layers deposited on it. This method is proved to be very efficient for realizing suspended membranes, bridges, or cantilevers

PS has been also employed for the epitaxial deposition of c-Si, because it conserves the lattice of the pristine material and can be used as a seed for further growth. Thanks to its reactivity and fragility, it allows transfer of the epitaxial film onto foreign substrates.

Macroporous Silicon

The former studies of n-type silicon anodization in the dark started in the early 1970s. Etch pits and silicon dissolution in dark is due to and breakdown between surface states tunneling in conduction band. From these studies, some evidences were clear, the channels formation is not due to crystal defects, the pores direction follow the orientation of the substrate, the growth rate is in general constant with anodization time, so the diffusion within the pores is not a rate-limiting process, and the pores' walls are passivated.

Beale proposed a depletion of porous layer, explaining the high resistivity of the material after anodization, and localizing carriers at the pore tips. Several concepts introduced by this physical model, were then adopted in successive works, but no chemical mechanisms were included, so providing poor comprehension of pore size variations and morphologies. Smith and coworkers, at the end of 1908s, postulated that the rate of pore growth is limited by the diffusion of the holes randomly walking in the bulk of silicon. This model was mainly ruled by the diffusion length L, a quantity that is function of dopant concentration and of potential. A sticking factor varying from point to point of the partially oxidized pore tip was

necessary to explain different morphologies and pore diameter variations. This model was too general to explain conduction mechanisms and morphologies in different type of substrate and like the previous was not taking in account the electrochemical reactions at the interface. Zhang et al., in the same period, thanks to a systematic study of the *i*-*V* characteristics, clarified that the PS formation or electropolishing is not dependent by electronic properties of the substrate, but as the result of the two competing reactions occurring at the pore tips, direct dissolution by fluoride species and indirect removal of silicon through silicon dioxide formation and dissolution in HF. In the following years, Zhang also proposed that surface curvature of the pore tips could reduce the effective width of the space charge layer greatly increasing the interface tunneling current and the dissolution rates of silicon dissolution and oxide formation and removal. This model is also applicable to other substrates and PS morphologies.

Lehmann and Föll in 1990 described the formation of macropores from low-doped n-type silicon in aqueous HF-electrolyte, backside illumination and pore nucleation predefined lithographically. They experimentally verified the hypothesis that the space charge region (SCR) around the pores can focus the holes diffusing in the silicon bulk on the pore tips, so enabling the growth of regular macropores. In further studies, Lehmann demonstrated that the equilibrium of the carrier transport in the semiconductor and the mass transfer in the electrolyte determines the etch rates and the morphologies in n-type silicon. At low current densities, there is no accumulation of holes and dissolution takes place only at the pore tips, leaving the walls unaffected thanks to carrier depletion. For high current densities, the reaction is mass transport limited and holes accumulate at the pore tips, dissolving walls and enlarging the pores.

Lehmann tried to unify the models describing the formation mechanisms of all types of PS, using the assumption that all the porous structure except the pore tips are passivated by carrier depletion. Depletion is due to quantum confinement for microporous silicon and to space charge layer for porous structures larger than few nanometers.

Föll et al. developed the current burst theory, one of the most recent and comprehensive models of PS formation, it is so far the only model considering the

electrochemical reactions under formation conditions in the different PS morphologies. The main assumptions are that the electrochemical reaction of silicon dissolution operates in microscopic units with a definite spatial and temporal distribution. Other important points are the following:

Current flow is inhomogeneous in space and time, from here the local current burst.

Current flow induces direct silicon dissolution and oxidation and removal following current burst.

As a result of hydrogen termination and desorption, the position of Fermi level and so of the space charge layer oscillates with the time

The rate of hydrogen termination is fastest on (111), determining the probability of current burst on surfaces of different orientations.

This model is rather powerful in explaining many aspects involved in the PS formation but it is extremely difficult to be expressed in mathematical form [21].

Porous Silicon Based Photonics

In nano- and mesoporous cases, application in 1D photonics structures deserves a mention. From an optical point of view, PS is seen from visible light as a homogeneous material. In fact, the sizes of the pores and the Si nanocrystals are well below the light wavelength. When porosity is varied, the macroscopic dielectric constant (and the refractive index) is varied, too. In principle, it is possible to “choose” the refractive index of the layer between the limits of $n = 3.4$ (porosity = 0) to $n = 1$ (porosity = 1). The range of variation for n is smaller, but n can be varied in a continuous way in this wide range. Being the dependence of n from porosity relatively simple and considering that porosity monotonically increases with the anodization current, by changing the current during anodization in the right way, one is able to obtain optical devices like Bragg Reflectors or Fabry-Perot microcavities done by PS. The advantage of this technology is evident: it allows obtaining optical devices directly integrated on Si [22].

A Bragg reflector is a periodical structure of layers with high and low refractive index. The optical thickness of both is chosen to be $\lambda/4$, λ being the light wavelength at which maximum rejection is wanted. The Bragg reflector is then a sort of 1D photonic crystal, there the rejection band stands for photons like energy gap stands for electrons in semiconductors.

Correspondingly, one can choose to insert defects states within such gap: this can be accomplished by means of a Fabry-Perot microcavity, where a layer with optical thickness equal to $\lambda/2$ is sandwiched between two Bragg reflectors. The PS anodization technology allows for realizing Bragg reflectors with >98% reflectivity at the rejection band and Fabry-Perot microcavities with peak FWHM <10 nm. This method could in principle be used to confine light arising from PL emission, but, unfortunately, the best optical 1D structures are obtained in highly doped material, in which the PL is generally poor. The roughness between layer of different porosity is dramatic in lightly doped PS, and greatly affects the optical performances of the 1D devices. A possible solution has been found in freezing the anodization etchant: roughness is greatly reduced in this way, but the quality is still far from devices realized by standard vacuum deposition methods.

Cross-References

- ▶ [Dry Etching](#)
- ▶ [Microfabricated Probe Technology](#)
- ▶ [Nanostructured Functionalized Surfaces](#)
- ▶ [Nanostructured Materials for Sensing](#)
- ▶ [Nanostructures for Photonics](#)
- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)
- ▶ [Wet Etching](#)

References

1. Uhler, A.: Electrolytic shaping of germanium and silicon. *Bell Syst. Technol. J.* **35**, 333–347 (1956)
2. Johannes, P., (Poughkeepsie Burkhardt, NY), Poponiak, M. R. (Newburgh, NY): Porous silicon dioxide moisture sensor and method for manufacture of a moisture sensor United States Patent 4057823 (1977)
3. Yonehara, T.: BESOI with porous silicon: ELTRAN[®]. In: Canham, L.T. (ed.) *Properties of Porous Silicon*, pp. 23–28. Institution of Engineering and Technology, London (1997). ISBN 0852969325
4. Canham, L.T.: Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers. *Appl. Phys. Lett.* **57**(10), 1046–1048 (1990)
5. Lehmann, V., Gosele, U.: Porous silicon formation: a quantum wire effect. *Appl. Phys. Lett.* **58**(8), 856–858 (1991)
6. Canham, L.T.: Bioactive silicon structure fabrication through nanoetching techniques. *Adv. Mater.* **7**, 1033–1037 (1995)
7. Kovalev, D., Timoshenko, V.Y., Künzner, N., Gross, E., Koch, F.: Strong explosive interaction of hydrogenated porous silicon with oxygen at cryogenic temperatures. *Phys. Rev. Lett.* **87**, 068301–068303 (2001)
8. Mikulec, F.V., Kirtland, J.D., Sailor, M.J.: Explosive nanocrystalline porous silicon and its use in atomic emission spectroscopy. *Adv. Mater.* **14**, 38–41 (2002)
9. Coffey, J.L.: Porous silicon formation by stain etching. In: Canham, L.T. (ed.) *Properties of Porous Silicon*, pp. 23–28. Institution of Engineering and Technology, London (1997). ISBN 0852969325
10. Kolasinski, K.W.: New approaches to the production of porous silicon by stain etching. In: Granitzer, P., Rumpf, K. (eds.) *Nanostructured Semiconductors: From Basic Research to Applications*. Pan Stanford, Singapore (2011)
11. Li, X., Bohn, P.W.: Metal-assisted chemical etching in HF/H₂O₂ produces porous silicon. *Appl. Phys. Lett.* **77**(16), 2572 (2000)
12. Amato, G.: Optical and morphological properties of light-emitting porous silicon prepared by chemical dissolution of silicon wafers. *Jpn. J. Appl. Phys.* **34**, 1716 (1995)
13. Chartier, C., Bastide, S., Levy-Clement, C.: Metal-assisted chemical etching of silicon in HF–H₂O₂. *Electrochim. Acta* **53**(17), 5509–5516 (2008)
14. Turner, D.R.: On the mechanism of chemically etching germanium and silicon. *J. Electrochem. Soc.* **107**, 810–816 (1960)
15. Lehmann, V.: *Electrochemistry of Silicon*, p. 58. Wiley, Weinheim (2002)
16. Rouquerol, J., Avnir, D., Fairbridge, C.W., Everett, D.H., Haynes, J.H., Pernicone, N., Ramsay, J.D.F., Sing, K.S.W., Unger, K.K.: Recommendations for the characterization of porous solids. *Pure Appl. Chem.* **66**(8), 1739 (1994)
17. Canham, L.T.: Skeleton size distribution in porous silicon. In: Canham, L.T. (ed.) *Properties of Porous Silicon*, pp. 23–28. Institution of Engineering and Technology, London (1997). ISBN 0852969325
18. Garrone, E., Geobaldo, F., Rivolo, P., Amato, G., Boarino, L., Chiesa, M., Giamello, E., Gobetto, R., Ugliengo, P., Viale, A.: A nanostructured porous silicon near insulator becomes either a p- or an n-type semiconductor upon gas adsorption. *Adv. Mater.* **17**(5), 528–531 (2005)
19. Amato, G., Rosenbauer, M.: Absorption and photoluminescence in porous silicon. In: Amato, G., Delerue, C., von Bardeleben, H. (eds.) *Optical and Structural Properties of Porous Silicon Nanostructures*, pp. 3–52. Gordon and Breach, Amsterdam (1997)
20. Lang, W.: Micromachining applications of porous silicon. In: Amato, G., Delerue, C., von Bardeleben, H.J. (eds.) *Optical and Structural Properties of Porous Silicon Nanostructures*, pp. 597–620. Gordon and Breach, Amsterdam (1997)
21. Zhang, X.G.: Macropores in n-type silicon and breakdown model. In: Zhang, X.G. (ed.) *Electrochemistry of Silicon and its Oxide*, pp. 410–418. Kluwer/Plenum, New York (2001)
22. Berger, M.G., Thönissen, M., Theiß, W., Münder, H.: Microoptical devices based on porous silicon. In: Amato, G., Delerue, C., von Bardeleben, H.J. (eds.) *Optical and Structural Properties of Porous Silicon Nanostructures*, pp. 557–595. Gordon and Breach, Amsterdam (1997)

Nanostructures for Coloration (Organisms other than Animals)

Ille C. Gebeshuber^{1,2} and David W. Lee³

¹Institute of Microengineering and Nanoelectronics (IMEN), Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

²Institute of Applied Physics, Vienna University of Technology, Vienna, Austria

³Department of Biological Sciences, Florida International University Modesto Maidique Campus, Miami, FL, USA

Synonyms

Iridescent colors (organisms other than animals);
Physical colors; Structural colors

Definition

Structural colors refer to colors generated by minuscule structures, with the characteristic dimension of the structures on the order of the wavelength of the visible light (i.e., some tens up to hundreds of nanometers). Examples for structural colors are the colors of CDs and DVDs, the colors of soap bubbles, or oil films on water (thin films), or the colors of certain butterfly wings (e.g., photonic crystals), and even plants. Tiny wax crystals in the blue spruce scatter the light (Tyndall scattering), resulting in the blue hue. Thin films in tropical understory plants and diffraction gratings in hibiscus and tulip flowers are just some more examples of the amazing variety of natural nanostructures that are the basis for coloration in some plants. This entry reviews the physics behind structural colors, lists plants, microorganisms, and virus species with nanostructures responsible for their coloration, and also touches upon the multifunctionality of materials in organisms, nanobioconvergence as an emergent science, and biomimetics as a promising field for knowledge transfer from biology to engineering and the arts.

Introduction

Iridescent, metallic, or greyish coloration in some plants is not caused by pigments, but rather by physical structures with characteristic sizes on the order of some tens up to several hundreds of nanometers [12]. The name-giver for the term iridescent, that is, color change with the viewing angle, is Iris (Ἴρις), the Greek goddess and personification of the rainbow.

According to previous studies about various kinds of colors and their origins, colors are divided to two different categories: chemical colors and structural colors. Chemical colors are caused by pigments, whereas structural colors are caused by structures. The interaction of visible light with structures causes shiny, bright colors that might also show iridescence and/or metallic appearance. Vertical spatial structuring may limit the change in color bandwidth with angle – in this way structural colors of specific hues such as, for example, the brilliant blue of the Morpho butterfly can be generated [1, 10]. In some of these structural colors with small bandwidth at ambient conditions, viewing at small angles and high light intensities reveals a wide spectrum of colors (e.g., in Raja Brooke's Birdwing).

In chemical colors, light is selectively reflected, absorbed, and transmitted. Pigments reflect the wavelengths of light that produce a color, and absorb the other wavelengths. Ultramarine blue, for example, reflects blue light and absorbs other colors. Light absorbed by the pigment yields altered chemical bonds of conjugated systems and other components of the pigment, and is finally released as heat, as reemission of light via fluorescence or by passing the energy on to another molecule. Plant pigments are largely soluble in aqueous solutions or lipids, and they actually work more like dyes in the sense that they do not actually directly reflect but allow the passage of wavelengths that are then reflected by structures with different refractive indices, such as cell walls. Plant pigments are therefore more like water colors on white paper than the colors used in an oil painting.

In structural colors, on the other hand, the incident light is reflected, scattered, and deflected on structures, with negligible energy exchange between the material and the light, resulting in strong, shiny coloration.

The objective of this entry is to give a review of the physical mechanisms leading to structural colors, introduce plant species as well as microorganisms and viruses with the respective nanostructures thought to be responsible for or contribute to the coloration, give some examples for structural colors in current technology, and introduce nanobioconvergence as an emergent science and biomimetics as a promising field for knowledge transfer from biology to engineering and the arts.

The Physics of Structural Colors

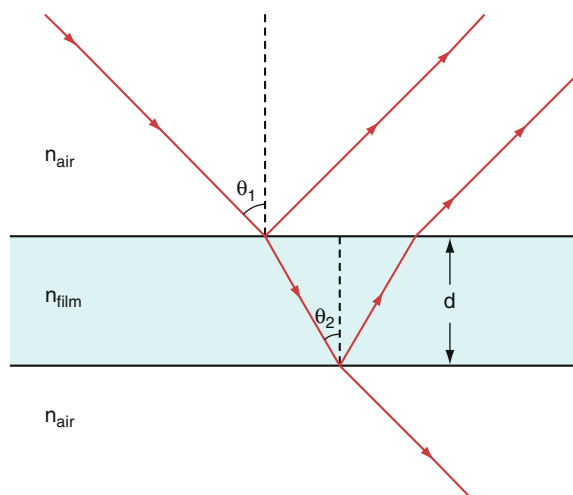
Five physical phenomena lead to structural coloration: thin film interference, multilayer interference, scattering, diffraction, and photonic crystals.

Thin Film Interference

Thin film interference occurs when an incident light wave is partly reflected by the upper and lower boundaries of a thin film (Fig. 1). It is one of the simplest phenomena in structural coloration and can be found in various instances in nature, for example, in the wings of houseflies, in oil films on water, or in soap bubbles.

Light that strikes a thin film surface can be either transmitted or reflected from upper or lower boundary, respectively. This can be described by the Fresnel equation. Interference between two reflected light waves can be constructive or destructive, depending on their phase difference. Of relevance are the thickness of the film layer (d), the refractive index of the film (n), and the angle of incidence of the original wave on the film (θ_1). In addition, if the refractive index n_2 of the thin film is larger than the refractive index of the outside medium n_1 , there is a 180° phase shift in the reflected wave. This is for example the case in a soap bubble, with $n_1 = n_{air} = 1$ and $n_2 = n_{film}$, with $n_{air} < n_{film}$.

The reflection on the upper boundary of the film (the air–film boundary) causes a 180° phase shift in the reflected wave (since $n_{air} < n_{film}$). Light that is transmitted at this first boundary continues to the lower film–air interface, where it is either reflected or transmitted. The reflection at this boundary has no phase change because $n_{film} > n_{air}$.



Nanostructures for Coloration (Organisms other than Animals), Fig. 1 Thin film interference in the case of, for example, soap bubbles

The condition for constructive interference for a soap bubble is given in (Eq. 1), the condition for destructive interference is given in (Eq. 2):

$$2n_{film}d \cos(\theta_2) = (m - 1/2)\lambda \quad (1)$$

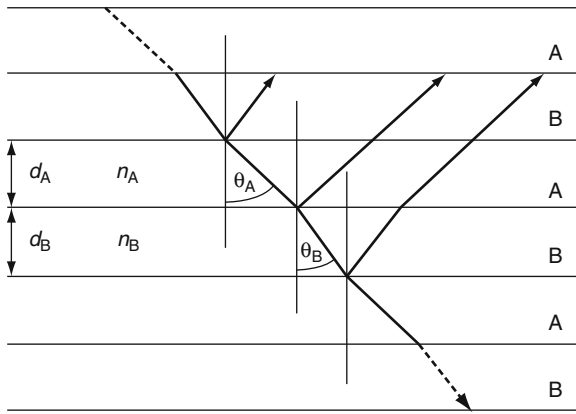
$$2n_{film}d \cos(\theta_2) = m\lambda \quad (2)$$

where d equals the film thickness, n_{film} equals the refractive index of the film, θ_2 equals the angle of incidence of the wave on the lower film–air boundary, m is an integer, and λ is the wavelength of the light.

Summing up, there are two conditions that should be satisfied for constructive interference: firstly, the thin film should be thin enough to crest the reflected waves and secondly, the two reflected waves should be in one phase.

Multilayer Interference

As described above, a light wave can be reflected from both boundaries of a thin film layer and dependent on the phase of the two reflected light waves either destructive or constructive interference occurs. The same phenomenon occurs in a series of thin films, a multilayer. A schematic of a multilayer is given in Fig. 2. With $n_B > n_A$, between each A-B interface a 180° change in phase takes place, while at the B-A



Nanostructures for Coloration (Organisms other than Animals), Fig. 2 Multilayer interference. $n_B > n_A$. A change in viewing angle corresponds to a change in the perceived color, due to changes in the phases of the interfering waves

interface there is no phase change. Reflection or refraction at the surface of a medium with a lower refractive index causes no phase shift. Reflection at the surface of a medium with a higher refractive index causes a phase shift of half a wavelength.

Constructive interference:

$$2(n_A d_A \cos \theta_A + n_B d_B \cos \theta_B) = m\lambda \quad (3)$$

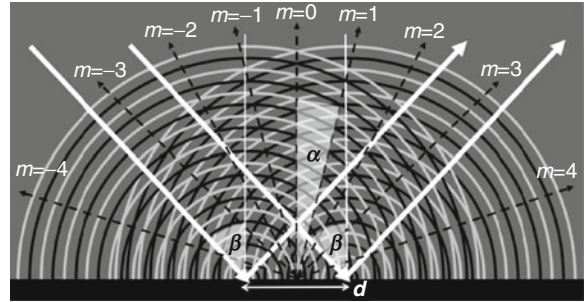
Maximum reflection:

$$2n_A d_A \cos \theta_A = (m - 1/2)\lambda \quad (4)$$

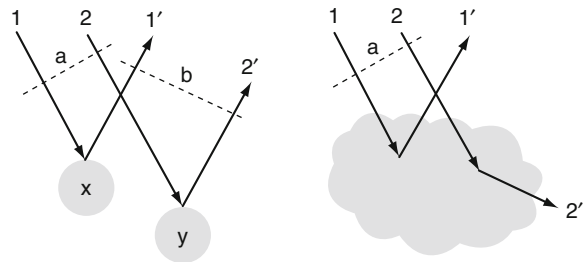
In ideal multilayers all waves interfere constructively, resulting in colorful reflections, whereas in non-ideal multilayers some waves interfere destructively and the reflection is less colorful [17].

Diffraction Gratings

The grating equation (Eq. 5) correlates the spacing of the grating (d) with the angles of the diffracted and incident beams (α, β) and the wavelength of illuminating light (λ) (Fig. 3). The coloration in a particular direction (viewing angle) is generated by the interfering components from each slit of the grating. The diffracted light has maxima at angles where there is no phase difference between the waves. The 0th order reflection corresponds to direct transmission and is denoted $m = 0$. Other maxima occur at angles with $m = \pm 1, \pm 2, \pm 3, \dots$



Nanostructures for Coloration (Organisms other than Animals), Fig. 3 Schematic of the color generating mechanism by the diffraction of light by a diffraction grating (black bar on bottom)



Nanostructures for Coloration (Organisms other than Animals), Fig. 4 Coherent (left) and incoherent scattering (right)

$$d(\sin \alpha - \sin \beta) = m\lambda \quad (5)$$

CDs and DVDs are examples for this phenomenon. Very recently, diffraction gratings that produce colors have been described in flowers [18].

Scattering

The term scattering is a more general denomination for the interference of light waves with different wavelengths reflected from scattering objects either in a constructive or destructive way (Fig. 4). In terms of coloration, scattering can yield either strong or weak colors. Examples of scattering, and more precisely, coherent scattering are the thin film interference and the multilayer film interference described above.

In coherent scattering there is a definite phase relationship between incoming and scattered waves, whereas in incoherent scattering this is not the case.

The two main types of scattering are termed Rayleigh scattering and Tyndall scattering. In both types of scattering, the intensity of the scattered light depends on the fourth power of the frequency.

In Rayleigh scattering, the responsible particles are much smaller than the wavelength of the incident light, whereas in Tyndall scattering, the particles are macroscopic (e.g., dust particles in air or small fat particles in water, as in milk). Since blue light (i.e., shorter wavelengths) is scattered more than red light (i.e., long wavelengths), milk, tobacco smoke, and the sky are blue or have a blue hue. In plants, for example, the blue of the blue spruce can be explained by scattering [14]. Multiple scattering within regular structures produces highly reflective bands in the reflection spectrum, and leads to the brilliant colors observed in, for example, iridescent fruits ([12] and references therein).

Photonic Crystals

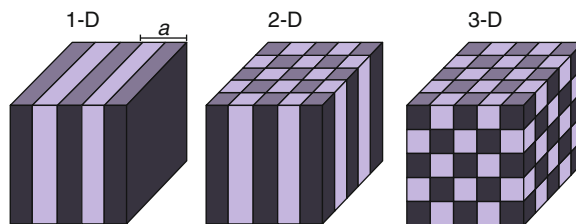
Photonic crystals exhibit a periodicity a in the refractive index (Fig. 5). There are one-, two-, and three-dimensional photonic crystals. Depending on their wavelength, photons either can be transmitted through the crystal or not (allowed and forbidden energy bands). For effects in the visible range, the periodicity of the photonic crystal has to be between about 200 nm (blue) and 350 nm (red).

Vegetable opal is an example for a photonic crystal produced by plants. Opal is the only gemstone that is known to be formed also as a plant-produced product. There are three different types of opal produced by photosynthetic organisms: minuscule iridescent diatoms, phytoliths, and tabasheer. Tabasheer (bamboo opal, plant opal) is a hard, translucent to opaque whitish substance extracted from the joints of bamboo [5]. It can have beautiful blue hue, and was – cut as cabochons – used in China for fashion jewelry.

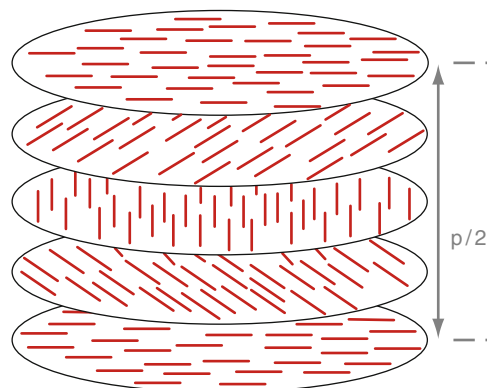
Cholesteric Liquid Crystals

Cholesteric liquid crystals are layered helicoidal structures, also known as chiral nematic liquid crystals, within which molecules take a preferred direction that gradually changes from layer to layer (Fig. 6). The variation of the director axis tends to be periodic in nature. The period of this variation (the distance over which a full rotation of 360° is completed) is known as the pitch, p .

Certain iridescent tropical understory ferns contain a helicoidal layering of cellulose microfibrils that produces multiple interference layers ([12] and references therein), similar to the helicoidal microstructure



Nanostructures for Coloration (Organisms other than Animals), Fig. 5 Schematics of one-, two-, and three-dimensional photonic crystals (1-D, 2-D, 3-D). The colors represent materials with different refractive indices. The spatial period of the material is called the lattice constant, a



Nanostructures for Coloration (Organisms other than Animals), Fig. 6 The chiral nematic phase in a cholesteric liquid crystal. p refers to the chiral pitch. Image reproduced with permission under the terms of the GNU Free Documentation License

in a cholesteric liquid crystal and the characteristic liquid crystals in beetles [1].

The peak wavelength of the reflected light, λ , is determined by the helicoidal pitch, p , where

$$\lambda = pn \tag{6}$$

and n is the average refractive index of the anisotropic material [3].

Plants (and Other Nonanimals) with Structural Colors

Structural colors in nature are found in various animals, plants, microorganisms, and even viruses [1, 10, 12, 19]. Studies of structural colors date back to the seventeenth century when the earliest scientific



description of structural colors appeared in “Micrographia,” written by Robert Hooke in 1665. In his book “Opticks” Sir Isaac Newton already in 1704 related iridescence to optical interference:

The finely colour'd feathers of some birds, and particularly those of the peacocks' tail, do in the very same part of the feather appear of several colours in several positions of the eye, after the very same manner that thin plates were found to do.

Since then, various organisms with structural colors (almost all in animals) have been reported. This chapter focuses on organisms outside of the animal kingdom, primarily plants, but also macroalgae, slime molds, diatoms, and viral particles. For plants, the discussion is organized along the physical phenomena as well as the various organs where such color is found, that is, leaves, flowers, fruits, and seeds. Other organisms with simpler organization are classified by the major groups, for example, the red algae (Rhodophyta).

The first report on structural colors in plants was by Lee and Lowry [13], who described the physical basis and ecological significance of iridescence in blue plants. The most common mechanisms in plants leading to structural coloration are multilayer interference and diffraction gratings. Multilayer interference is found predominantly in shade-plant leaves, suggesting a role either in photoprotection or in optimizing capture of photosynthetically active light ([12] and references therein). Surprisingly, diffraction gratings may be a common feature of petals, and recent work has shown that bees use them as cues to identify rewarding flowers. Structural colors may be surprisingly frequent in the plant kingdom, and still much remains to be discovered about their distribution, development, and function [7].

Nonanimals with Coloration Caused by Thin Film Interference

True Slime Molds

Slime molds are organisms that normally live as single cells and that can aggregate to a multicelled organism and form fruiting bodies that produce spores. They have the characteristics of single-celled microorganisms as well as of fungi and occur worldwide on decaying plant material. Inchaussandague and coworkers [9] described structural colors produced by a single thin layer in the slime mold *Diachea*

leucopoda. Interference within the peridium, a 200 nm transparent layer invariant along the different cells of the “organism,” is the basis for this color.

Plants with Coloration Caused by Multilayer Interference

Tropical Understory Ferns

The “peacock fern,” *Selaginella willdenowii* (actually a lycophyte, or fern ally), is a common plant in the Malaysian rainforests. It exhibits blue iridescence; investigation with TEM reveals various layers (with less than 100 nm thickness each) on the surface. The photo shown in Fig. 7 was taken in the Bukit Wang Recreational Forest in Malaysia, with very long exposure time as to better reveal the blue coloration of the leaves. Seen with the naked eye, the fern is bluish green and seems to glow in semidarkness. The blue coloration is iridescent and changes with the viewing angle. Holding and tilting the leaves results in color change from green to nearly total blue, although not as strong as seen in the long exposure time photograph shown in Fig. 7.

Also another species of *Selaginella* shows iridescence: *Selaginella uncinata*. As in *S. willdenowii*, a multilayer system where each layer is thinner than 100 nm is responsible for the structural coloration. When the leaves of these two *Selaginella* species are dried, the coloration disappears, but comes back again when the plants are hydrated again. When a droplet of water is put on a *Selaginella* leaf, the blue coloration disappears. Other tropical understory ferns where coloration is caused by multilayer interference are *Diplazium crenatoserratum*, *Lindsaea lucida*, *Danaea nodosa*, and *Trichomanes elegans* (the neotropical bristle fern, also called the “dime store plant” because it looks like it is made from plastic and is shade tolerant, is native to the American tropics) ([12] and references therein).

Tropical Understory Begonias

Gould and Lee (1996) reported structurally modified chloroplasts (termed “iridoplasts”) in the Malaysian tropical understory plant *Begonia pavonina* (the peacock begonia) ([12] and references therein). The structural coloration is generated by many layers, each less than 100 nm thick. In addition to *B. pavonina*, the authors have observed several other iridescent blue species on the Malaysian peninsula. Modified plastids produce blue colors in *Begonia*, *Phyllagathis* (see



Nanostructures for Coloration (Organisms other than Animals), Fig. 7 The “peacock” fern *Selaginella willdenowii*, a common plant in the Malaysian rainforest. © Mr. Foozi Saad, IPGM, Malaysia. Image reproduced with permission

section “[Other Tropical Understory Plants](#)”), and *Trichomanes* (see section “[Tropical Understory Ferns](#)”). Species with modified chloroplasts are not helicoidal.

Other Tropical Understory Plants

A traditional Malaysian medicinal plant, the melastome *Phyllagathis rotundifolia* (in Malay known as *tapak sulaiman*) and also the Malaysian understory species *P. griffithii* exhibit iridoplasts, similar to the ones in *Begonia pavonina* described above.

Iridescent Macroalgae

Among the red algae (Rhodophyta) several species of *Mazzaella* (= *Iridaea*) are strikingly iridescent [6]. Green and brown algae have been recorded as iridescent, but the phenomenon is most commonly found in red algae, where blue and green colors appear on the surface at certain stages of the lifecycle. Iridescent red algae (e.g., *Mazzaella flaccida*, *Mazzaella cordata*) exhibit shiny blue, emerald green, and deep red colors. The blades are relatively thin and vary in shape and size (up to 3 ft long and 10 in. wide) depending on the habitat and the amount of wave exposure. *Mazzaella flaccida* is iridescent yellow-green with purple or brown near the base of the blade, and the iridescent blade of *Mazzaella splendens*, is dark purple with a blue iridescent sheen. It is also known as the rainbow-leaf seaweed. The coloration disappears when the algae are dried, and is restored when the algae become wet again.



Nanostructures for Coloration (Organisms other than Animals), Fig. 8 *Mazzaella* sp., a highly iridescent red alga. © 2003 by J. Harvey, <http://www.JohnHarveyPhoto.com>. Image reproduced with permission

Wet *Iridaea* look like they have been coated with oil (Fig. 8), because of their rainbow sheen; iridescent spots, shaped like a teardrop, with a multilayer system of 17 electron opaque and translucent layers, some tens to a few hundred of nanometers thick, are responsible for the coloration. The iridescence might be a byproduct of a wear-protection mechanism of the algae.

The edible seaweed *Chondrus crispus* (carrageen, Irish moss) is a red algal species that shows blue iridescence due to multilayer interference effects on the tips [15].

The function of algal iridescence is still unclear. Suggestions include a role in camouflage or a role in optimizing photosynthesis by enhancing the absorption of useful wavelengths of light at the expense of increased reflection of other wavelengths.

Plants with Coloration Caused by Diffraction Gratings

Glover and Whitney [7] reviewed the original research on iridescence in plants and presented very recent results on floral iridescence produced by diffractive optics. They identified iridescence in flowers of *Hibiscus trionum* (the inner part of its petals has an oily iridescence overlying red pigmentation) and tulip species (e.g., *Tulipa kolpakowskiana*) and demonstrated that iridescence was generated through diffraction gratings. The iridescence in the hibiscus is obvious to human eyes (appearing blue, green, and yellow depending on the angle from which it is viewed), whereas the iridescence in the tulip is in the ultraviolet spectrum, which bees can see, but

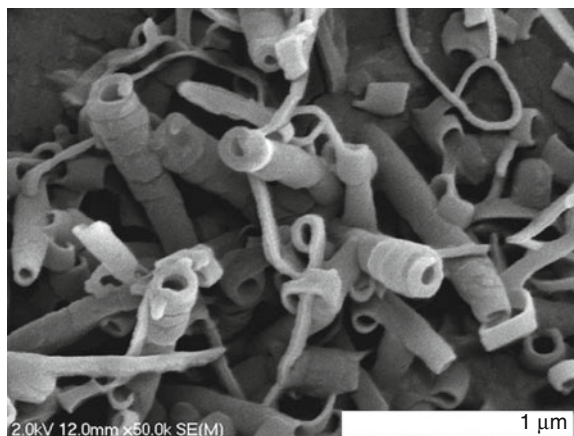
people cannot. The surface striations in the diffraction gratings observed in the tulips are about one micrometer apart. Whitney and her coworkers investigated 22 tulip species from around the world, and found ordered striations in 18 of them. They report iridescence generated by diffraction gratings in 10 families of angiosperms, tulips and hibiscus being just two of them: From the mallow family the species example they give is *Hibiscus trionum*, from the lily family, *Tulipa* sp.; from the aster, daisy or sunflower family, *Gazania klebsiana*; from the pea family, *Ulex europaeus*; from the Loasaceae family (of bristly hairy and often climbing plants), *Mentzelia lindleyi*; from the buttercup or crowfoot family, *Adonis aestivalis*; from the willowherb or evening primrose family, *Oenothera biennis*; from the nightshades, *Nolana paradoxa*; from the iris family, *Ixia viridiflora*; and from the peony family, *Paeonia lactiflora*. According to this extensive list, it might well be that many more flowers are iridescent than previously thought – since they are iridescent in parts of the spectrum that people cannot see. The principal reason why not more species with diffraction-based coloration are known might simply be that until now nobody has looked.

Plants with Coloration Caused by Scattering

Some green leaves look white or silvery because microscopic air spaces in surface hairs reflect the light. Other leaves (such as the wax palm) and some fruits (such as plums and grapes) have a white “bloom” which is a surface deposit of wax.

Whitish, silvery, and other metallic finishes of leaves are generally produced by reflection of the light off microstructures, nonliving plant hairs called trichomes. In deserts this increased reflection, especially of infrared radiation, off the microstructures is important since it reduces heating of the leaf. However, a reflecting leaf has a much lower photosynthetic capacity than would an equal leaf without the reflectant trichomes. The desert brittlebush *Encelia farinosa* produces nonreflectant green leaves during the cooler springtime, when it is beneficial to have warmer leaves.

Silvery, grey, white, or blue coloration in plants might arise due to light scattering on three-dimensional epicuticular wax structures (structural wax). Leaves with a bluish gray waxy surface are called glaucous.



Nanostructures for Coloration (Organisms other than Animals), Fig. 9 Monohelical tubular epicuticular wax structures in *Wollemia nobilis*. Scale bar 1 μm . From [4], reproduced with permission

Generally, radiation is scattered across the spectrum, with increased reflectance in the ultraviolet, visible, and infrared windows. The size, distribution, and orientation of wax crystals and other surface features determine the extent to which light is scattered at the plant surface. In some plant species such as the blue spruce (*Picea pungens*) preferential scattering of shorter wavelengths and enhanced reflectance of UV radiation occurs.

Epicuticular wax is a complex mixture of long-chain aliphatic and cyclic compounds. The bluish leaf hues of the evergreen desert mahonia *Berberis trifoliolata*, for example, arises from wax structures on nipple-like projections from the epidermal cells, called epidermal papillae. In some cases, for example in the Blue Finger, *Kleinia mandraliscae*, and in cabbage, *Brassica oleracea*, powdery wax can be rubbed from the surface, causing the grey or bluish hue to disappear, and revealing the green leaf color beneath. The Wollemia pine *Wollemia nobilis*, an evergreen tree with bluish green mature foliage, has monohelical tubular epicuticular wax structures with a diameter of about 100 nm (Fig. 9) [4]. Depending on the wax compound present, epicuticular wax crystals can have various shapes [11]: The glaucous eucalyptus tree (*Eucalyptus gunnii*) has nanoscale β -diketone epicuticular wax tubules; the blue-grey foliage of the dusty meadow rue (*Thalictrum flavum glaucum*) has nanoscale nonacosanol tubules. The Black Locust tree (*Robinia pseudoacacia*) has dark, dull, blue-green

summer foliage with nanoscale epicuticular wax platelets that are arranged in rosettes. The leaves of wild cabbage (*Brassica oleracea*) show simple nanoscale rodlets, whereas the nanoscale rodlets on the Silky Sassafras (*Sassafras albidum*) leaves are transversely ridged.

The epicuticular wax in the skin of plums (*Prunus domestica* L.) consists of an underlying amorphous layer adjacent to the cuticle proper together with crystalline granules of wax protruding from the surface. The epicuticular waxes in olive leaves serve as radiation inceptors.

Coloration Caused by Photonic Crystals: Plants, Diatoms, and Viruses

Iridescent Blue Fruits

The fruits of the blue quandong (*Elaeocarpus angustifolius*, syn. *E. grandis*) and the blue nun (*Delarbrea michiana*) have brilliant blue coloration ([12] and references therein). The iridescent blue color of the quandong fruit is actually enhanced by wetting, not like in *Selaginella*, mentioned above, where it disappears when wet. Iridosomes, multilayered systems arranged in 3-D structures, similar to the structure yielding coloration in Morpho butterflies, are responsible for the coloration of these fruits. Iridosomes are different from the iridoplasts described above for leaves: They are secreted by the epidermal cells of the fruit, and are located outside the cell membrane but inside the cell wall. The structures of the iridosomes are at least partly cellulosic. The stone of the silver quandong fruit is termed *rudraksha*. It is hard and highly ornamented, and is used throughout India and Southeast Asia in religious jewelry.

Edelweiss

Edelweiss (*Leontopodium nivale* subsp. *alpinum*) is a European alpine flower that grows at high elevations up to 3,400 m. Its body is covered with white hairs. Vigneron and coworkers showed in 2005 [16] that the internal structure of these hairs acts as a two-dimensional photonic crystal. The hairs are hollow tubes, about 10 μm in diameter, with an array of parallel striations with a lateral diameter of 176 nm around the external surface. Through diffraction effects determined by the optical fiber with photonic crystal cladding, the hairs absorb the majority of the UV light, thereby acting as an efficient sun-block for the structures beneath.

Spores

Some ferns produce two types of spores: microspores and megaspores (size: some hundreds of micrometers). Hemsley et al. [8] reported iridescent megaspores in *Selaginella* (fossil and modern). They report that spore iridescence is not generally visible until the outer layers of the spore wall (with different organization to the iridescent layer) are removed. The three-dimensional pattern yielding the iridescence, for example, in megaspores of *Selaginella exaltata*, is still unresolved. According to SEM images, an ordered structure of 240 nm diameter particles, perhaps an opal-like colloidal crystal that is generated via self-assembly, is the reason for the coloration.

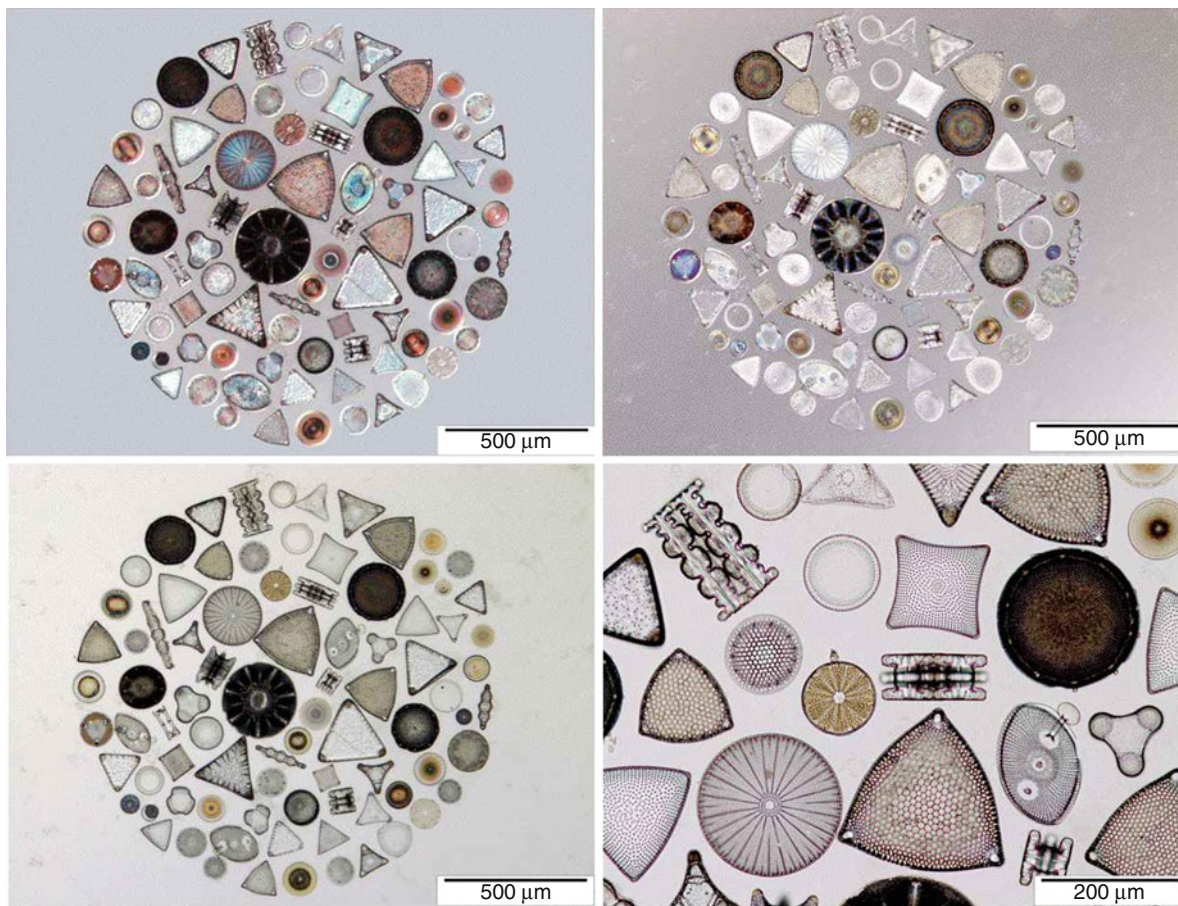
Tabasheer

Opal is the only gemstone that can be produced in biotic and abiotic ways. Tabasheer (vegetable opal, pearl opal) consists of pure silica and is produced when bamboo is hurt; the plant sap comes to the surface and dries in small nodules. These little nodules are opaque or translucent white, sometimes with a faint hint of blue. They can be polished as cabochons and are often used in the orient as jewelry. Tabasheer is highly porous and hygroscopic, and has been used in ancient times to remove snake poison from the body – therefore it is also known as snake stone [5].

Diatoms

Diatoms (Bacillariophyta) are found in both freshwater and marine environments, as well as in damp soils and on moist surfaces. They are unicellular microalgae with a cell wall consisting of a siliceous skeleton enveloped by a thin organic case. The cell walls of each diatom form a pillbox-like shell consisting of two parts that fit within each other.

Diatoms exhibit an amazing diversity of nanostructured frameworks (Fig. 10), including two-dimensional inverse photonic crystals. The cell wall exhibits periodic arrangement of pores in the micrometer to nanometer range – each diatom species has its own specific morphology. Individual diatoms range in size from 2 μm up to several millimeters, although only a few species are larger than 200 μm . A patch of diatoms can look like iridescent scum to the naked eye. Under the microscope, beautiful iridescence can be seen on the single cell level, and a microscope slide covered with a monolayer of diatoms also exhibits iridescence to the naked eye



Nanostructures for Coloration (Organisms other than Animals), Fig. 10 Sample with various diatoms from New Zealand, laid by hand by Elger in 1925. The slide was photographed by Friedel Hinz under the optical microscope with slightly varying angles of view, revealing color change in

some of the diatoms (*top images and bottom left*). *Bottom right*: Zoom. Marine fossil diatoms from Gave Valley, New Zealand, Sample T1/21 prepared by Elger, 1925. Image (c) 2011, F. Hinz, AWI Bremerhaven, Germany. Image reproduced with permission

(with good illumination or in sunlight). Iridescent effects are used widely in color cosmetic products and personal care packaging, and there is great potential for using diatoms in this industry.

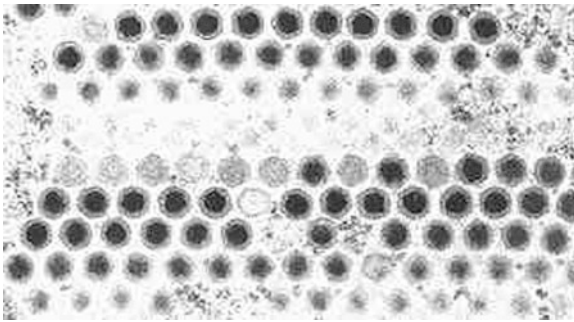
Viruses

Iridescent viruses (Iridoviridae) are between 120 and 140 nm in diameter and infect insects, fishes, and frogs. Iridescent viruses produce systemic infections, with the highest concentrations in the outer layer of the skin (the epidermis) and the fat bodies. Insect larvae infected with this virus develop iridescent lavender-blue, blue, and blue-green coloration. In older infections, the iridescent color is generally more vivid, and frequently there are also small,

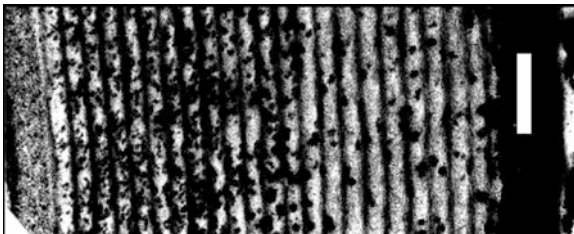
extremely brilliant islands. Electron microscope examination reveals numerous viruses in paracrystalline arrays (see Fig. 11) [20].

Plants with Coloration Caused by Cholesteric Liquid Crystals

In three iridescent tropical understory fern species, structurally modified chloroplasts with helicoidal structures, similar to the characteristic liquid crystals in beetles [1], might be the reason for the blue coloration. These ferns are *Danaea nodosa* (with many multilayers, each less than 100 nm thin), the necklace fern *Lindsea lucida* (17 layers, 192 nm spacing) and *Diplazium tormentosum* (20–23 layers, 141 nm spacing, see Fig. 12) [12].



Nanostructures for Coloration (Organisms other than Animals), Fig. 11 Paracrystalline array of virus particles within an infected cell. This array gives rise to the iridescence phenomenon. Image Source: <http://www.microbiologybytes.com/virology/kalmakoff/Iridoviruses.html>. Permission pending



Nanostructures for Coloration (Organisms other than Animals), Fig. 12 Ultrastructure of a *Diplazium tomentosum* leaf. Scale bar 0.5 μm . Image © with one of the authors (DWL)

Structural Coloration Caused by Not Yet Described or Not Yet Identified Mechanisms

Macroalgae

Various red algae (Rhodophyta) and brown algae (Phaeophyta) exhibit iridescence, and are completely unstudied. For example, of yet unknown origin is the bright blue, purple, or red iridescence in the red alga *Ochtodes secundiramea*.

Dictyota are widespread brown algae (Phaeophyta) along the Atlantic coast and very well known. *Dictyota mertensii* has blue/green iridescence; *Dictyota dichotoma* is also known as the purple peacock algae; *Dictyota barayresii* shows a translucent, iridescent blue; *Dictyota sandvicensis* exhibits a yellow green iridescence; and *Dictyota humifusa* is light brown, often with brilliant blue iridescence. Iridescent bodies in the brown algae *Cystoseira stricta* are specialized vacuoles (vesicles) that contain numerous dense globules inside. The physics of their iridescence has not yet been clarified. The Hawaiian brown alga *Styopodium hawaiiensis* has a beautiful blue or green color, and *Dictyopteris zonarioides* sometimes is iridescent blue [6].

Additional to the iridescence in red algae with multilayers responsible for the coloration (see section “Iridescent Macroalgae” above), the following red algae exhibit structural coloration: The blue branching seaweed *Faucheia laciniata* shows deep red with iridescent blue; thin layer interference (whether from one layer or a multilayer system still needs to be clarified) is the reason for its coloration. *Fryella gardneri* exhibits iridescent purple; *Cryptopleura ruprechtiana* is mildly iridescent. In *Chondracanthus corymbiferus* the younger blades are smooth and iridescent. Other iridescent red algae are *Chondria coerulescens* and *Drachiella spectabilis*. Also *Maripelta rotata* and *Botryoglossum farlowianum* are iridescent.

Slime Molds

As described above in section “True Slime Molds,” the physical reason for the iridescent color in at least one type of slime mold is a single transparent thin layer. Slime mold iridescence is also described for some other species; however, its physical basis still needs to be determined. Spores of *Diachea subsessilis* exhibit dull greenish grey varying to blue iridescence; the peridium in *Diachea deviata* is iridescent and persistent. Beautiful bronze iridescent colors, sometimes tinged with blue, can be observed in the peridium of *D. subsessilis* (with the membranous peridium being colorless in water mounts, suggesting that pigments are not involved in the color production).

Plants

Bryophytes Bryophytes include the mosses, liverworts, and hornworts. The moss *Schistostegia pennata* exhibits golden-green iridescence when it is grown in caves where the light always comes from only one direction. This moss (also known as dragon’s gold) shines like emerald jewels from the darkness of a rock crevice or cave. This unusual property is the result of lens-shaped cells with curved upper surface that focus the light on one point in the interior of the cell, where the chloroplasts aggregate. It remains to be determined if nanostructures are responsible for the coloration of this moss (Lee [12] and references therein).

Ferns *Elaphoglossum herminieri*, a strap fern that is native to tropical rainforests of the neotropics shows iridescent blue color (Fig. 13) that is not removed by wetting. The reason of the coloration is unknown, as it is



Nanostructures for Coloration (Organisms other than Animals), Fig. 13 The iridescent blue strap fern *Elaphoglossum herminieri*, at Heredia, Sarapiquí near Puerto Viejo, La Selva Biological Station, Costa Rica. © 2008 R.C. Moran, The New York Botanical Garden (rmoran@nybg.org) [ref. DOL23374]. Image reproduced with permission

in *Elaphoglossum wurdackii* and *E. metallicum*. The oil fern *Microsorium thailandicum* (*Microsorium steerei*) from Southeast Asia has blue-green iridescent leaves. Also ferns of the family Vittariaceae can have iridescent scales. In this family, especially ferns of the genus *Antrophyum* (e.g., *Antrophyum formosanum*) and *Haplopteris* can be densely covered with iridescent scales. *Antrophyum mannianum*, a mountain fern liking full shade, has more or less iridescent leaves. There are also reports on an *Antrophyum* species found in Borneo (Mt. Mulu, or Gunong Mulu) with an iridescence that makes the plant look like it is covered with metallic eye shadow. *Haplopteris* ferns live in tropical and subtropical Asia. Obscure iridescence is exhibited by *Haplopteris amboinensis*, *H. doniana*, *H. taeniophylla*, *H. himalayensis*, *H. mediosora*, *H. fudzinoi*, *H. linearifolia*, *H. hainanensis*, *H. sikkimensis*, *H. elongata* and *H. anguste-elongata* show bright iridescence. Also *H. flexuosa* is iridescent, but not as bright. Malaysian individuals of the fern *Didymochlaena truncatula* are reported to be iridescent. The Common maidenhair *Adiantum capillus-veneris* has iridescent stems, and the Western maidenhair *Adiantum aleuticum* has iridescent blackish foliage.

Grasses (Poaceae) Iridescence can be observed in the rosy inflorescences of the ornamental grass *Miscanthus sinensis*. Further grasses with startling

iridescent inflorescences are the small grass *Pennisetum alopecuroides*, little bunny, and the even smaller variety little honey. In late summer, these grasses are covered with blooms that look like hairy caterpillars.

Sedges (Cyperaceae) A sedge is a grass-like or rush-like plant growing in wet places having solid stems, narrow grass-like leaves, and spikelets of inconspicuous flowers. Leaves and seedlings of sedges can exhibit iridescence. The understory sedge *Mapania caudata* from Malaysia and also *M. graminea* exhibit iridescence in their leaves. The smooth black sedge grass *Carex nigra* shows whitish iridescence and the green sedge grass *Carex dipsacea* has iridescent olive green leaves. *Carex flagellifera* “toffee twist” has iridescent, slender leaves.

The seedlings of the pond flat sedge *Cyperus ochraceus* are dark iridescent gray because of an outer one-cell thick covering of translucent cells, whereas the seedlings of the marsh flat sedge *Cyperus pseudovegetus* are brown with a very thin translucent-iridescent layer of cells. Also *Carex squarrosa* has blackish seedlings with iridescent superficial cells (when fully mature). In *Fimbristylis annua* and *F. vahlii* the seedlings are often iridescent. The Harper’s fimbry *F. perpusilla* has pale brown seedlings with iridescent tints, and the southern fimbry *F. decipiens* has seeds of whitened-iridescent to brown.

Hypoxidaceae: Hypoxis The stargrass *Hypoxis sessilis* has seedlings that are black but with iridescent membranous coats.

Orchids (Orchidaceae) Iridescent foliage in orchids occurs, for example, in *Macodes petola*, *Masdevallia (Byrsella) caesia*, and *Aulosepalum pyramidale*, where it is bright green with an iridescent shine. Flowers of many species of orchid (such as *Ophrys apifera* and the mirror bee orchid *Ophrys speculum*) have been reported to have an iridescent patch, the speculum. The shape and iridescence of this iridescent patch is thought to mimic the closed wings of female wasps or bees, and sexually deceive the respective males, thereby pollinating the orchid (see [7] and references therein).

Others There are various further plant species that exhibit iridescence. Some of them are mentioned

below. The main point is that there is a lot to study and potentially several novel mechanisms of color production in organisms are yet to be identified.

Purslanes (Portulacaceae): *Portulaca*. Seeds mostly glossy black or iridescent gray have been reported for various species of *Portulaca*. The reason of their iridescence is still unknown. In the moss-rose purslane *Portulaca grandiflora* Hooker the mature seeds are steely gray and often iridescent; in the pink purslane *Portulaca pilosa* the seeds are black, with very slight purplish iridescence when mature; in *Portulaca psammotropha*, the seeds are black, turning iridescent gray when fully mature, and in the shrubby purslane (copper purslane) *Portulaca suffrutescens* Engelman and *P. halimoides*, the seeds are leaden and slightly iridescent.

Grapes (Vitaceae): *Ampelopsis*. The porcelain berry vine, *Ampelopsis brevipedunculata*, has slightly iridescent fruits showing white, green, turquoise, blue, brown, and violet coloration.

Lardizabalaceae (no common name): *Decaisnia*. The blue beans (*Decaisnea fargesii*) from Bhutan have iridescent, 4-in. large blue fruits.

Pinks (Caryophyllaceae): *Spergularia*. Other plants with iridescent seeds are the blackseed sandspurry, *Spergularia atrosperma*, whose black seeds are often iridescent.

Aizoaceae (ice plants): *Sesuvium*. Another plant with iridescent seeds is the slender sea purslane, *Sesuvium maritimum*, with brownish-black, smooth, and somewhat iridescent seeds.

Plantaginaceae (no common name): *Bacopa*. The blue hyssop, *bacopa caroliniana*, has iridescent seeds.

Goosefoot (Chenopodiaceae): *Chenopodium*. The purple goosefoot (tree spinach), *Chenopodium giganteum*, is an annual plant in which the young shoots are covered with a fine iridescent magenta powder.

Rapataceae: *Stegolepis*. The leaves of *Stegolepis ligulata* have blue-green iridescence; even more impressive is the larger *S. hitchcockii*.

Mallows (Malvaceae): *Hibiscus*. The red leaf hibiscus (*H. acetosella*) produces iridescent lobed maroon leaves on woody stalks.

Palms (Arecaceae): *Chamaedorea*. Very impressive are the leaves of the metallic palm *Chamaedorea metallica*.

Hydrangea (Hydrangeaceae): The United States patent PP18294 refers to the invention of

a *Hydrangea macrophylla* plant named “HYMMAD I” whose inflorescences mature to an iridescent lime green.

Star Flowers (Hydroxylidaceae): The flowering plant *Spiloxene capensis* can have white, cream yellow, or pink flowers with a dark center. In the pink form the center is iridescent blue-green, whereas in the white form the center is iridescent bronze-green.

Ice Plants (Aizoaceae): The inflorescences of the hardy ice plant *Delosperma cooperi* “Mesa Verde™” have iridescent hues of salmon-pink.

Gesneriads (Gesneriaceae): *Rhynchoglossum obliquum* is a highly shade-resistant flowering plant with deep blue iridescent inflorescences.

Sea hollies (Apiaceae): *Eryngium maritimum*, sea holly and *E. x tripartitum*, big blue, have surprisingly iridescent blue flowers.

Spikemoss (Selaginellaceae): *Selaginella kraussiana* “gold tips,” also known as golden jade, with the common names Krauss’s spikemoss and African clubmoss, is a spikemoss found naturally in the Canary Islands, the Azores, and parts of mainland Africa. It has dark green foliage with golden tips.

Begoniaceae: *Begonia*. Many begonia species have iridescent leaves; however, the coloration of the leaves is strongly dependent on the environment the plants are grown in. In *Begonia Rex* the leaves can be of shiny metallic red. A begonia plant named “lady rose” with iridescent foliage, flower bud, petals, and tepals was patented by Mr. Terry McCullough in December 1998 (patent number plant 10,736). There are no reproductive organs formed in this invention. Other iridescent begonias are *B. burkillii*, *B. limprichtii*, *B. congesta*, *B. hahiepiana*, *B. sizemoreae* and the hybrids *Begonia* bandit, Bethlehem star begonia, *B. comedian*, *B. Palomar prince*, *B. wild pony*, *B. his majesty*, *B. merry Christmas*, *B. little darling*, *B. max gold*, *B. stained glass*, *B. regal minuet*, *B. shirtsleeves*, *B. Venetian red*, and *B. hocking tutu terror*. The authors have observed several iridescent members of this genus in Malaysia, besides the *B. pavonina* described previously.

Conclusion and Outlook

In this entry, the authors give a review of the physical mechanisms leading to structural colors, introduce

plant species as well as microorganisms and viruses with the respective nanostructures thought to be responsible for or contribute to the coloration and present quite a number of organisms where the physical reason for the structural coloration has not yet been determined. Many animals have senses that either work in different bandwidths compared to the ones in humans, or they sense completely different properties (such as with echolocation or the magnetic sense). Since plants and animals interact on various levels and for various reasons, plants have evolved various properties tuned to animals.

Structural colors are increasingly being used in current and emerging technology as well as in the arts. Some examples: The Austrian company Attophotonics develops structural colors as sensors for lab-on-a-chip applications. The Japanese company Teijin Fibers Limited produced one blue dress made from Morphotex[®] structural colored fiber, a flattened polyester fiber mimicking Morpho butterfly structures. Qualcomm's mirasol[™] display technology was inspired by butterfly wings. ChromaFlair[®] and SpectraFlair[®] color-shifting paints from the company JDSU create color-shifting effects in cars. The emerging contemporary design practice Biornametics explores a new methodology to interconnect scientific evidence with creative design in the field of architecture, with role models from nature (e.g., structural colors in organisms) being investigated and the findings applied to design strategies.

More often than not, natural structural colors are the inspiration for such technical products and applications. The field of biomimetics is dealing with the identification of deep principles in living nature, and their transfer to humankind [2]. General biological principles identified by the German biologist Werner Nachtigall that can be applied by engineers who are not at all involved in biology are, for example, integration instead of additive construction, optimization of the whole instead of maximization of a single component feature, multifunctionality instead of mono-functionality, energy efficiency and development via trial-and-error processes. We are hopeful that biomimetic colors will not only show the bright and shiny properties as the natural examples do, but will also

exhibit one of the most important properties ensuring continuity of the biosphere: sustainability. Currently, with nanotechnology as booming new promising field, the sciences have started to converge, with nanobioconvergence as one of the most promising examples. New materials, structures, and processes can arise from this novel approach, perhaps exhibiting some of the multifunctional properties of the inspiring organisms, where various levels of hierarchy (with functionality on each and every level) are so refined. The convergence of the ways of thinking and novel approaches promises an interesting future.

Acknowledgments The National University of Malaysia funded part of this work with its leading-edge research project scheme "Arus Perdana," and the Austrian Society for the Advancement of Plant Sciences funded part of this work via the Biomimetics Pilot Project "BioScreen." Living in the tropics and exposure to high species diversity at frequent excursions to the tropical rainforests is a highly inspirational way to do biomimetics. Profs. F. Aumayr, H. Störi, and G. Badurek from the Vienna University of Technology are acknowledged for enabling ICG extensive research in the inspiring environment of Malaysia; Bettina G. Neunteufl is acknowledged for carefully reading the manuscript. Mrs. Friedel Hinz and Mr. Foozi Saad are acknowledged for providing Figs. 10 and 7.

Cross-References

- ▶ [Bioinspired Synthesis of Nanomaterials](#)
- ▶ [Biomimetics](#)
- ▶ [Biornametics – Architecture Defined by Natural Patterns](#)
- ▶ [Lotus Effect](#)
- ▶ [Nanoimprinting](#)
- ▶ [Nanomechanical Properties of Nanostructures](#)
- ▶ [Nanostructured Functionalized Surfaces](#)
- ▶ [Nanostructures for Photonics](#)
- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)
- ▶ [Nanotechnology](#)
- ▶ [Optical and Electronic Properties](#)
- ▶ [Rose Petal Effect](#)
- ▶ [Scanning Electron Microscopy](#)
- ▶ [Self-assembly of Nanostructures](#)
- ▶ [Self-repairing Materials](#)
- ▶ [Transmission Electron Microscopy](#)

References

- Berthier, S.: *Iridescences: the Physical Colors of Insects*. Springer, New York (2007)
- Bruckner, D., Gruber, P., Hellmich, C., Schmiedmayer, H.-B., Stachelberger, H., Gebeshuber, I.C. (eds.): *Biomimetics – Materials, Structures and Processes. Examples, Ideas and Case Studies*. Springer, Berlin (2011)
- de Gennes, P.G., Prost, J.: *The Physics of Liquid Crystals. The International Series of Monographs on Physics*. Oxford University Press, Oxford (1995)
- Dragota, S., Riederer, M.: Epicuticular wax crystals of *Wollemia nobilis*: morphology and chemical composition. *Ann. Bot.* **100**, 225–231 (2007)
- Eckert, A.W.: *The World of Opals*, p. 16 and 183. Wiley, New York (1997)
- Gerwick, W.H., Lang, N.J.: Structural, chemical and ecological studies on iridescence in *Iridaea* (Rhodophyta). *J. Phycol.* **13**(2), 121–127 (1977)
- Glover, B.J., Whitney, H.M.: Structural colour and iridescence in plants: the poorly studied relations of pigment colour. *Ann. Bot.* **105**(4), 505–511 (2010)
- Hemsley, A.R., Collinson, M.E., Kovach, W.L., Vincent, B., Williams, T.: The role of self-assembly in biological systems: evidence from iridescent colloidal sporopollenin in *Selaginella* megaspore walls. *Phil. Trans. R. Soc. Lond. B* **345**, 163–173 (1994)
- Inchaussandague, M., Skigin, D., Carmaran, C., Rosenfeldt, S.: Structural color in myxomycetes. *Opt. Express* **18**(15), 16055–16063 (2010)
- Kinoshita, S.: *Structural Colors in the Realm of Nature*. World Scientific Publishing Company, Singapore (2008)
- Koch, K., Barthlott, W.: Superhydrophobic and superhydrophilic plant surfaces: an inspiration for biomimetic materials. *Phil. Trans. R. Soc. A* **367**(1893), 1487–1509 (2009)
- Lee, D.: *Nature's Palette: the Science of Plant Color*. University of Chicago Press, Chicago (2007)
- Lee, D.W., Lowry, J.B.: Physical basis and ecological significance of iridescence in blue plants. *Nature* **254**, 50–51 (1975)
- Martin, J.T., Juniper, B.E.: *The Cuticles of Plants*, p. 109. Edward Arnold, Edinburgh (1970)
- Pedersén, M., Roomans, G.M., Hosten, A.V.: Blue iridescence and bromine in the cuticle of the red alga *Chondrus crispus* Stackh. *Bot. Mar.* **23**, 193–196 (1980)
- Vigneron, J.P., Rassart, M., Vértesy, Z., Kertész, K., Sarrazin, M., Biró, L.P., Ertz, D., Lousse, V.: Optical structure and function of the white filamentary hair covering the edelweiss bracts. *Phys. Rev. E* **71**(8p), 011906 (2005)
- Vukusic, P., Sambles, J.R.: Photonic structures in biology. *Nature* **424**, 852–855 (2003)
- Whitney, H.M., Kolle, M., Andrew, P., Chittka, L., Steiner, U., Glover, B.J.: Floral iridescence, produced by diffractive optics, acts as a cue for animal pollinators. *Science* **323**(1), 130–133 (2009)
- Williams, T.: The Iridoviruses. *Adv. Vir. Res.* **46**, 345–412 (1996)
- Willis, D.B. (ed.): *Iridoviridae. Current Topics in Microbiology and Immunology*, vol. 116. Springer, Berlin/Heidelberg/New York/Tokyo (1985)

Nanostructures for Energy

Stefano Bianco, Angelica Chiodoni, Claudio Gerbaldi and Marzia Quaglio
Center for Space Human Robotics, Fondazione Istituto Italiano di Tecnologia, Torino, Italy

Synonyms

[Nanodevices](#); [Nanomaterials](#); [Nanoparticles](#);
[Nanorods](#); [Nanotubes](#); [Nanowires](#)

Definition

Objects whose dimensions range between 100 and a few nanometers can be classified as nanostructures. Since the physical-chemical properties of matter change dramatically at the nanometer-scale, the introduction of this classification results particularly important to open a new field of investigation for next-generation materials and devices.

Introduction

The worldwide need for power is mainly satisfied by the use of well-known energy resources as oil, coal, hydropower, natural gas and nuclear power. The environmental impact of most of these energy sources is tremendous, mainly because of the emission of greenhouse gases that causes the increase in the globally averaged temperature. In the next decades a rise of the global energy demand is expected, mainly because of the growth of world population especially in the developing countries. It is interesting to notice how the preindustrial world's

energy consumption grew about 0.3 terawatts (TW) year per year, and this value dramatically increased to 14 TW in the years 2000s and it is expected to double to 28 TW by year 2050. In the same period the world population increased only five times, clearly showing how the energy consumption is not uniformly distributed. An improvement in life quality in developing countries will cause a further tremendous increase in energy demand in the next future [1, 2]. A fast growing country needs far more energy than a developed country, essentially because a mature industrialized economy has learned how to produce and use energy more efficiently [2].

In this scenario, exploring renewable, efficient and green energy sources is the greatest challenge to be won for sustainable human progress. Innovation is one of the key strategies to address this challenge, and some of the most promising solutions are related to nanotechnologies. Materials at the nanometer-scale are able to show completely different properties than those people are used to in their everyday life. Understanding the huge potentialities of materials at the nanoscale and learning how to drive them in real devices opens new extraordinary opportunities. The field of alternative energy is certainly the most challenging platform for some of nanotechnology's most exciting contributions [3]. New nanomaterials and nanostructures showed recently their potentiality in the development of devices for both energy production and storage, like third-generation photovoltaic cells, fuel cells and lithium-based batteries.

Solar energy conversion systems are limited mainly by the poor properties of current materials, showing limited charge-carrier mobility and significantly high charge recombination rates for materials with sustainable costs. Moreover, theoretical limitations (Shockley/Queisser limit) are well established and fix a superior insurmountable efficiency threshold. Nanotechnology could help in the design of innovative architectures for a new-generation solar cells, actually opening to a possible real alternative to traditional silicon solar cells. Valuable examples are dye-sensitized solar cells (DSCs). In the last few years, significant efforts were spent in designing new dye molecules with improved light harvesting properties as well as optimized mesoporous semiconducting layers with

lower charge recombination rates. Devices based on photosensitized nanostructured TiO₂ films and on liquid electrolytes have demonstrated light-to-current conversion efficiency greater than 10% [4]. Other approaches, like hybrid organic solar cells based on conjugated polymers or quantum dots solar cells with a light harvesting performed by semiconductor nanocrystals, are showing promising laboratory results. However, the achievement of satisfactory efficiency values still remains a challenge, mainly because of consistent losses during photoinduced charge separation and charge transport across the nanoassemblies.

Fuel cells represent another technological field in which new concepts from nanotechnology can be beneficial. The development of this kind of devices for large-scale use is still limited by the lack of adequate systems for hydrogen storage. At the same time nanotechnologies offer new interesting opportunities, directly storing hydrogen into the fuel cells by nanostructures of carbon, zeolites, stacked clays, or light metals hydrides [5]. The introduction of nanoengineering also opens to the design of new electrodes. Nanotechnologies can lead to a more reduced use of precious metals (e.g., platinum nanoparticles for high surface area and low volume) along with improved membrane function and durability [6].

Used batteries and accumulators represent a tremendous environmental problem in current days. Nanomaterials could help not only in developing batteries showing higher energy densities but also devices that will reduce the use of disposal energy-storage systems [7]. For example, it was demonstrated that silicon-based nanowires could significantly improve the specific capacity of Li-ion accumulators, currently powering-up electronic devices such as laptops, video cameras, cell phones, and many others [8]. In fact, such a new technology can grant a ten times increase of the capacity of traditional Li-ion batteries. Also, nanotechnology can reduce the possibility of batteries catching fire by providing less flammable electrode materials.

New applications of nanotechnology are underway in the energy field. Their use will undoubtedly impact on modern society, hopefully changing and improving the way people are currently thinking to energy production/storage, leading to a more sustainable human growth.

Energy Supply Systems

The traditional approach to energy supply is based on the production of huge amounts (from MW to GW) of energy in localized production sites and its following delivery for minute use, causing significant losses across the distribution net. Petroleum and coal industry, nuclear power stations, hydroelectric dams, wind farms, and geothermal power installations are common examples of this approach (concerning both nonrenewable and renewable energy sources).

Nanotechnologies are expected to dramatically increase the efficiency of alternative energy devices in all the different fields [3], creating the starting point for a new approach to a more sustainable energy management. In fact, it is reasonable to believe that the current scenario will be slightly modified in the next future. The centralized production of high quantities of energy will be more intensively joined by a distributed system of energetic hot spots, able to produce very low amounts of energy, with high spatial density. By this way, the production of energy occurs in direct connection or in close proximity with its use in devices (often miniaturized or microdevices), thus avoiding losses. In this framework, the efficiency of the single energy harvester could be no more a crucial point, while other aspects (like flexibility, integration in complex systems, and even aesthetics) gain importance.

These kinds of approaches cover a broad interest in different fields, starting from new-generation buildings to new smart portable devices. The firsts are designed and fabricated according to a sustainable construction approach and are able not only to reduce the total energy consumption, but also to produce a large part of the energy required inside the building [9]. The next future devices (as household electrical appliances, cellular phones, computers, and biomedical devices) will be able to manage energy more efficiently, thanks to the introduction of new power supply systems. These systems will present an energy source, hopefully a renewable one as solar, mechanical harvesting, or thermoelectric, connected to a storage system as lithium-ion batteries, both managed by proper electronic power management circuits [10]. In this new futuristic scenario the role of nanomaterials and nanostructures in devices for energy production and storage will become more and more crucial.

Two complementary strategies can be considered to manage energy in energy supply systems. A first class of devices, as batteries, fuel cells, and supercapacitors, need a finite quantity or a continuous supply of an active material or fuel to properly work. Devices of this kind can be classified as non-regenerative power supplies. A second class of tools, as photovoltaic cells, mechanical harvesters, and thermoelectric devices, do not need any material or fuel to work and can be classified as regenerative devices. In the next sections some valuable examples of both categories will be briefly introduced.

Nanostructures for Energy Production

Dye-Sensitized Solar Cells (DSCs)

Dye-sensitized solar cells (DSCs) were invented in the early 1990s by Michael Grätzel and Brian O'Regan at the École Polytechnique Fédérale de Lausanne (CH) [11], following the extensive research in semiconductor photo-electrochemistry started in the 1970s. These cells have immediately drawn a huge interest in the scientific community, because of their potentiality for the fabrication of devices with relatively high efficiency and potential low cost. In fact, these kinds of solar harvesters can be produced with abundantly available organic materials and widely diffused inexpensive semiconductors, using simple and scalable technologies and avoiding costly and energy-intensive high-vacuum treatments and materials purification steps.

The working principle of the cell [4] is based on the presence of a molecular light absorber, which is able to harvest photons from sunlight and convert them to electrons, pumped up from a lower to a higher energetic level, thus creating an electric potential that can be used to produce electric work. The light-absorbing dye is attached as a monolayer to the surface of a nanocrystalline film of a wide band gap semiconductor (TiO_2 in a pure anatase form in the most common DSC architecture). The semiconductor works as electron collector, since the photogenerated charges are fast injected in the conduction band of the nanostructured solid. A careful design of the molecular energy levels of the dye makes the charge transfer toward the oxide strongly favored from an energetic point of view, thus drastically reducing the electron-hole recombination process. The dye is electrically

regenerated by charge donations from an electrolyte, which penetrates the pores of the nanofilm to have a close proximity with the molecules of the sensitizer, creating a 3D interpenetrating network junction. The electrolyte is an ionic conductor able to accept hole from (or donate electrons to) the oxidized dye, and the best performance are noticed for liquid electrolytes containing a I^-/I_3^- redox system. The electrons percolate through the nanoparticles, which are interconnected because of a short sintering process, and are collected by an external circuit, thus generating electric power without causing any permanent chemical transformation inside the cell.

The role of the nanostructures in designing suitable photoelectrodes for high-performance cells is crucial. In fact, as reported from the first experiments of Grätzel, it emerges that it is possible to move from a sunlight-to-current conversion efficiency of 1.2% using a polycrystalline anatase film [12] to the 7.9% using a nanostructured mesoporous TiO_2 layer [11], just varying the active exposed surface area of the electrode. In fact, on a flat surface a monolayer of dye is able to absorb only a little percentage of the impinging photons, since its optical cross section is sensibly smaller with respect to the geometrical area it occupies. Thus, for good energy harvesting efficiency it is mandatory that a big area is exposed to the incident flux of photons, and this is possible only with a nanostructuring of the photoanode. By carefully controlling the porosity of the film, it is possible to find the optimal equilibrium between the thickness of the film and the path that the photogenerated charges are able to cover (e.g., the diffusion length) before recombination.

Harvesting

Energy harvesting is the direct extraction of background energy from the environment and its conversion in electrical power [13]. Harvestable ambient energies can be classified in three main classes, namely radiations (see Sect. 3.1), kinetic/vibration energy, and thermal energy. Energy recovery is possible through properly designed devices. The heart of energy harvesting devices is an active material that is actually able to change/modify at least one of its properties as a response to the stimulus of ambient energy. The change of the material behavior, induced by environmental energy, is usually turned into electrical power. The diffusion of harvesting devices

has been limited till now by their low efficiency in electrical energy production if compared to traditional energy sources, mainly due to the lack of well-performing materials. The development of new nanostructures and nanomaterials is promising to fasten up the development of a new generation of highly efficient devices for energy harvesting [3].

Energy from Vibrations

Several sources of environmental vibrations can be used to harvest energy, such as vibrations from household appliances, moving equipments, trains, helicopters flying overhead, or human movement. To extract that energy from the ambient a piezoelectric material is needed as the active material. Piezoelectric materials are able to directly convert the mechanical strain induced by vibrations into an electrical energy. The conversion is possible through the structural deformation of the material that induces a local spatial charge separation through an electric dipole. Piezoelectricity can occur in polycrystalline ferroelectric ceramic materials such as barium titanate ($BaTiO_3$), lead zirconate titanate (PZT) and polymeric materials as poly(vinylidene fluoride), PVDF.

Traditional piezoelectric ceramics are very brittle, needing to be handled carefully to avoid breakage. The usability is also affected by their hard and fragile behavior since they have limited ability in following complex curved surfaces. This aspect is particularly limiting when sensors and harvesters are considered, since the capability of devices as harvesters to capture and extract energy is also related to the ability of the material (and, consequently, of the whole device) to well couple to the surface of the object generating the vibrations. Traditional polymeric piezoelectric materials can overcome this limitation being flexible and compliant, but unfortunately their electromechanical coupling coefficient is not comparable to traditional ceramics. These limitations have increased the attention to other approaches, and hybrid nanocomposite materials are one of the most promising alternatives. The final properties of nanocomposite materials are related to the connectivity between phases [10]. Nanocomposite materials are so promising for energy harvesting devices because of high coupling factors, to which they add low acoustic impedance, mechanical flexibility, and broad bandwidth coupled with low mechanical quality factor that make them interesting candidates also for transducer and acoustic applications.

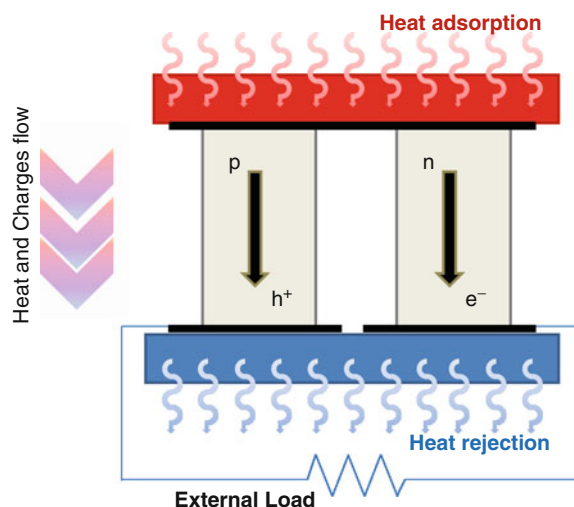
Energy from Temperature Gradients

A thermoelectric material is able to generate a voltage when a temperature gradient is applied to it (Seebeck effect) and vice versa (Peltier effect). The thermoelectric effect is related to the ability of charge carriers to move freely in metals and semiconductors (like molecules in a gas); while moving they can conduct charge and heat too. If a temperature gradient is applied to the material, charge carriers tend to move from the hot part to the cold region of the material through a diffusion process. While moving to the cold side charges are accumulated, creating a voltage that can be externally exploited [14]. The conversion of thermal energy in electrical energy is possible through p-n junctions made of thermoelectric materials as described in Fig. 1.

Thermoelectric generators are characterized by no moving parts that make them silent. They are reliable and scalable, and so they are considered quite interesting as power generators. Efficiency for thermoelectric materials and devices can be described by a nondimensional figure of merit that can be evaluated as $ZT = \sigma S^2 T / \kappa$, where S is the Seebeck coefficient, σ is the electrical conductivity, k is the thermal conductivity, and T is the absolute temperature.

For a long time, thermoelectric materials have been considered too inefficient to be cost effective for applications. The current state-of-the-art material in this field is a Bi_2T_3 -based alloy showing a ZT value close to one. New interest in thermoelectrics began in the middle of the 1990s, since theory suggested that thermoelectric efficiency could be significantly enhanced by the use of nanostructures. A well-performing thermoelectric material needs a variety of conflicting properties to be well optimized in order to work properly. It needs a large thermopower, which means a large value of the Seebeck coefficient; it must be a good electrical conductor, showing at the same time a very low thermal conductivity. In traditional materials, these transport characteristics are dependent the on one another, being deeply linked to interrelated material properties.

Based on both theoretical and experimental predictions, the current approach to increase ZT is to use nanostructured materials, both in form of films than quasi-1D structures, like nanowires, nanobelts, and nanotubes. Nanostructured films are quite interesting because they can use the same fabrication routes of traditional materials, while showing improved thermal behavior thanks to phonon scattering at the interfaces



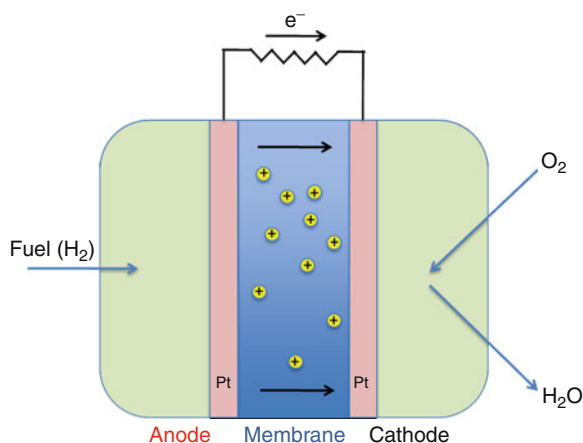
Nanostructures for Energy, Fig. 1 Schematic of a thermoelectric device based on a p-n junction between semi-conducting thermoelectric materials

among nanograins. The main issue in nanostructured films is to control electron scattering at grain boundaries between randomly oriented grains that could decrease also electrical conductivity [14]. In 1D materials as nanowires, the thermoelectric parameters (S , σ and k) can be optimized independently. A high value of ZT is predicted for 1D nanostructures, as a result of an increase of charge mobility due to the quantum confinement of the density of the states, together with a reduction in thermal diffusivity because of scattering at the grain boundaries [3].

Fuel Cells

A fuel cell (FC) is an electrochemical cell that converts directly chemical energy into electricity. This characteristic makes it an appealing system to provide a relatively cheap power source, by using hydrogen or ethanol as fuels, and by producing water as waste. They are constituted by an anode (hydrogen or ethanol fuel source) and a cathode (usually oxygen from air) separated by an electrolyte. The basic operation is depicted in Fig. 2.

Over the past few years, fuel cells have demonstrated increased reliability and lower costs because of the use of nanomaterials. The use of nanotechnology in the manufacturing process results in increased surface area, which originates large power and energy densities, long shelf life, and ease of miniaturization. These features, combined with MEMS technology



Nanostructures for Energy, Fig. 2 Basic operation of a fuel cell. Hydrogen enters at the anode, where it is catalytically dissociated and ionized. Protons migrate by ionic conduction to the cathode through the polymer electrolyte membrane. Electrons travel through the external circuit where they do electrical work before arriving at the cathode. At the cathode, oxygen molecules react with protons and electrons to form water molecules

[10], are important for the development of a more powerful fuel cell for portable electronic devices. Nanomaterials are also increasingly used in the production, purification and storage of hydrogen. In fact, at present, the ability to store hydrogen at high volumetric and gravimetric densities, as well as the ability to extract/insert it at sufficiently rapid rates, is still a hot topic that could condition further improvements of the fuel cell technology. In particular, a big effort has been done by the research community in the past decades to identify candidates that could fit such strict characteristics. Examples of investigated nanostructures are CNTs, considered a promising solution, even if recently partially abandoned, light metal hydrides [5], and graphene.

There are different types of fuel cells, depending on the type of electrolyte: the most diffused are proton exchange membrane fuel cells (PEMFCs) and high temperature fuel cells such as solid oxide fuel cells (SOFCs) and molten carbonate fuel cells (MCFCs). Among these PEMFC technology has become the dominant choice for the portable electronics market. Nowadays, nanotechnology is deeply embedded in the design of fuel cell components and, therefore, the development of nanostructured catalysts at the electrodes or optimized proton exchange membranes (PEM) has become an active area of research for fuel

cell applications. At present, many efforts are devoted to an optimization of the PEM characteristics, but despite great advancements in the development of nanostructured membranes, some restrictions and limitations remain. One of the challenges in current PEM fuel cell development is the lack of appropriate materials for the fabrication of proton-conductive membranes with improved performances. The high cost of the commercially available Nafion[®] membrane, the most diffused polymeric electrolyte in FC, has induced the development of new proton-conducting materials, including polybenzimidazoles, polyamides, polyether imides, polysulfones, polyphenylene sulfides, polyetheretherketones, and polyphenylquinoxalines. However, the Nafion[®] membranes show superior performance because of their high ionic conductivity compared to other non-fluorinated membranes. Proton conductivity is the most important property of membrane efficiency in fuel cells.

Membranes with acceptable conductivity at high temperatures are desirable for fuel cell applications, but there is always the issue of decreased performance at high temperatures due to dehydration of the membrane and a resulting reduction in ionic conductivity. To overcome the problems faced by proton exchange membranes, numerous studies were conducted to develop alternative materials for membrane preparation. One of the approaches to improve water retention in the membrane is to incorporate hydrophilic, inorganic materials [6]. However, with this method it is difficult to avoid nanoparticle self-agglomeration in the polymeric matrix, and the method therefore induces non-homogeneous dispersions and poor electrochemical performance. In order to achieve a homogeneous distribution, addition of nanosized inorganic particles is carried out by various methods, such as the solgel method, the solution casting method, and PECVD to obtain high-performance nanocomposite membranes. The use of surfactants can assist in creating a uniform distribution of nanoparticles as well.

A second challenge for the ultimate commercialization of fuel cells is the preparation of active, robust, and low-cost catalysts to replace platinum. Emphasis is placed on nanoengineering-based fabrication, processing, and characterization of multimetallic nanoparticles (transition metals) with controllable size (1–10 nm), shape, composition, and morphology.

Nanostructures for Energy Storage Systems

Many of the sustainable energy alternatives described produce or require electricity. Therefore, novel more efficient ways to store electricity are very much needed in the way to a more sustainable production, transformation, and use of energy. In this respect, electrochemical systems, such as batteries and supercapacitors, which can efficiently store and deliver energy on demand in stand-alone power plants, as well as provide power quality and load-leveling of the electrical grid in integrated systems, are playing a crucial role in the present energy economy. In addition, they are also seen as the power sources of choice for the progressive diffusion of sustainable hybrid/electric vehicles at high levels.

Lithium-Ion Secondary Batteries

Lithium-ion batteries (LIB) are one of the greatest successes of modern electrochemistry of materials. Their science and technology have been extensively reported in reviews [7] and dedicated books [8] to which the reader is referred for more details. A commercial LIB does not contain lithium metal; it comprises a negative electrode (generally, graphitic carbon), and a positive electrode (generally, layered lithium metal oxides such as LiCoO_2), both capable of reversibly intercalate Li^+ ions, these being separated by a nonaqueous lithium-ion conducting electrolyte (generally, a separator diaphragm soaked in a mixed ethylene – diethyl carbonate solution of LiPF_6). During discharge, Li^+ ions carry the current from the negative to the positive electrode, through the nonaqueous electrolyte [8]. During charge, an external electrical power source (the charging circuit) applies a higher voltage (but of the same polarity) than that produced by the battery, forcing the current to pass in the reverse direction. Lithium ions then migrate from the positive to the negative electrode, where they become embedded in the porous electrode material in a process known as “intercalation.”

The current-generation LIBs are increasingly being used by virtue of their high (volumetric and gravimetric) energy density; but, although commercially successful, some disadvantages arise, related to relatively low power density, large volume change on reaction, safety, and costs. In fact, the limits in performance using the current electrodes/electrolytes are being reached and further breakthroughs in materials

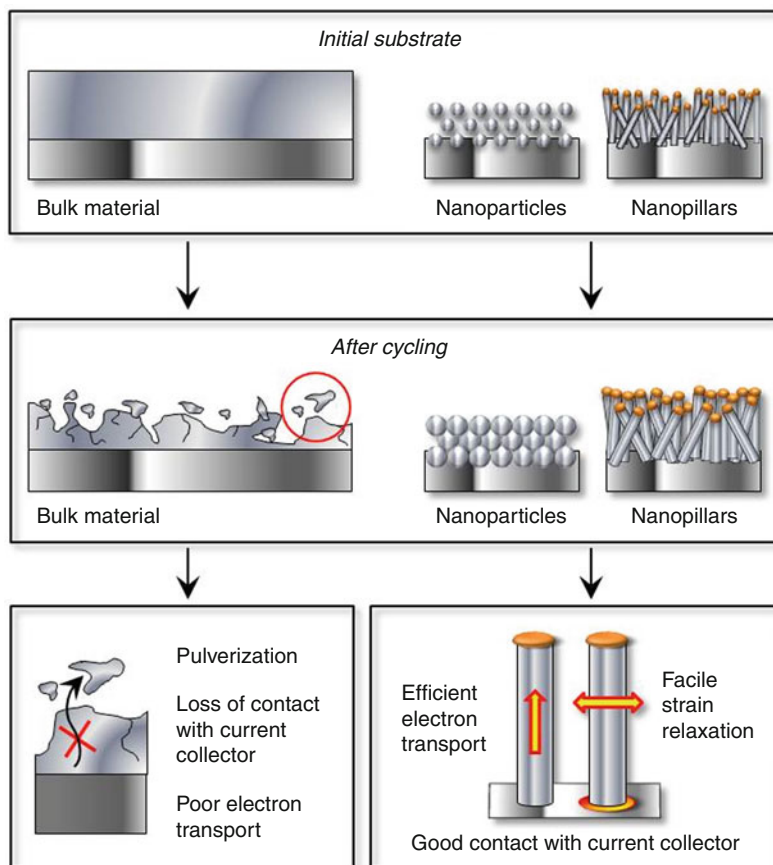
chemistry are essential, especially modifying and improving already known materials. The aforesaid shortcomings can be reduced by the application of nanotechnology. It is important to design and fabricate nanostructured electrode materials that provide high surface area and short diffusion paths for ionic transport and electronic conduction. This gives rise to a larger electrode/electrolyte surface and enables Li-ion diffusion through the electrode material to be limited as much as possible. Moreover, the volume changes and structural evolution that occur as the electrodes accept and release ions can be quite different for nanoscale materials, thus greatly affecting the cycle/service life.

The potential advantages and disadvantages associated with the development of nanomaterials applied to LIBs can be summarized as follows [7]. Advantages include (a) the strain associated with lithium ions insertion/removal is better accommodated (along with a wider solid-solution region), thus improving service life as shown in Fig. 3; (b) new electrode reactions are enabled to occur that are not possible for materials composed of micrometer-sized particles; (c) a high surface area permits a more extensive electrode/electrolyte contact area, leading to higher charge/discharge rates; and (d) the reduced dimensions significantly reduce the path lengths for electronic and Li^+ transport, thus permitting operation with low electronic conductivity and/or at higher power. As for disadvantages, the following have been identified: (a) an increase in undesirable (side) reactions with the electrolyte due to the high electrode/electrolyte contact area, possibly leading to self-discharge, poor cycling, and calendar life; (b) the volume of the electrode increases because of poor particle packing, leading to lower volumetric energy densities unless special compaction methods are developed; and (c) potentially, nanoparticles may be more complex to synthesize, especially their dimensions may be difficult to control.

There are numerous examples in which nanostructured materials have been used in Li-ion batteries as both electrodes and electrolytes. In meso- and/or nanoporous architectures, for instance, the nanoscale domains of the material are accessible to molecular reactants because of rapid mass transport via diffusion through the continuous volume of mesopores. Very recently, B. Scrosati, Y.-K. Sun, and colleagues developed a high-capacity, nanostructured tin-carbon anode, and a high-voltage, high-rate

Nanostructures for Energy,

Fig. 3 Example of how nanostructuring can help in improving the performances of Li-ion battery electrodes by reducing the strain associated with lithium ions insertion/removal (Readapted from Chan et al. [15])



$\text{Li}[\text{Ni}_{0.45}\text{Co}_{0.1}\text{Mn}_{1.45}]\text{O}_4$ spinel cathode. When the two parts are put together, the result is a high-performance battery with a high energy density and rate capacity, highly suitable for powering low or zero emission HEV or EV vehicles [15]. The ultimate expression of the nanoscale in rechargeable lithium batteries is the formation of 3D nanoarchitected cells, in which pillared anodes and cathodes are interdigitated [16]. The prospect of having active and passive multifunctional components interconnected within a 3D architecture is envisioned to lead to superior energy-storage capacity, high-volume synthesis, and improved safety. A nanowire battery has been invented by a team led by Dr. Y. Cui at Stanford University in 2007 [17]; it produces ten times the amount of electricity of existing LIBs. The team's invention consisted of a stainless steel anode covered with silicon nanowires, to replace the traditional graphite anode. Silicon, which stores ten times more lithium than graphite, allows a far greater energy density on the

anode, thus reducing the mass of the battery. The large surface area further allows for fast charging and discharging and efficient electron transport down the length of each nanowire.

Electrochemical Capacitors

Electrochemical capacitors (ECs), also named "supercapacitors," are capacitors with relatively high energy density (typically on the order of thousands of times greater with respect to conventional electrolytic capacitors) [18]. The term "supercapacitor" is usually used to describe an energy-storage device based on the charge storage in the electrical double layer (EDL) of a high-surface-area carbon in aqueous electrolytes. In fact, they store electrical energy, like batteries, but using a different mechanism: while batteries do it chemically, supercapacitors store electricity physically, by separating the positive and negative charges. Their lifetime is in principle infinite, as they operate solely on electrostatic surface charge

accumulation. In comparison with conventional batteries or fuel cells, ECs also have a much higher power density. They can be charged or discharged at a rate that is typically limited by current heating of the electrodes. So, while existing ECs have energy densities that are perhaps one tenth that of a conventional battery, their power density is generally 10–100 times greater. Supercapacitors can be used as uninterruptable power sources (UPSs), can be coupled with batteries to provide peak power, and can replace batteries for memory backup. The supercapacitor is ideal for energy storage that undergoes frequent charge and discharge cycles at high current and short duration. They have applications as energy-storage devices used in vehicles and for smaller applications like home solar energy systems where extremely fast charging is a valuable feature.

Since their discovery, these devices have attracted considerably less attention with respect to batteries as energy-storage devices. Nevertheless, thanks to both the contribution of nanotechnology and the better understanding of the charge storage mechanisms (ion behavior in small pores) the interest on ECs has noticeably increased very recently [18]. Nanotechnology is important because it enables large surface areas to be obtained to optimize performance in terms of capacitance, ionic and electronic resistance. There are numerous examples in the literature showing that nanosized crystallites and/or mesoporous materials exhibit significantly higher specific capacitance as compared to nonporous materials or micron-sized materials. The assembly of nanoscale materials is also important.

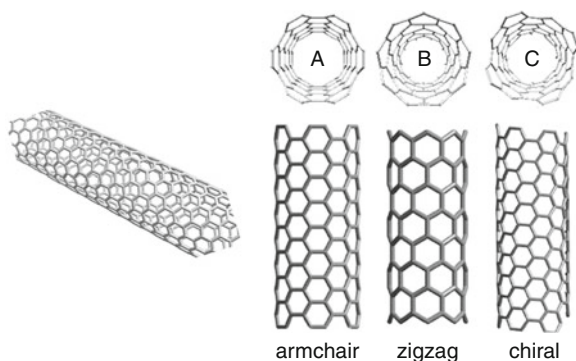
Recent trends in supercapacitors involve the development of high-surface-area activated carbon electrodes to optimize the performance in terms of capacitance and overall conductivity. Attention has been focused on nanostructured carbons, such as aerogels, nanotubes, and nanotemplates [19]. One structure envisioned to be of interest is an array of vertically aligned carbon nanotubes where the spacing between the tubes is matched to the diameters of the solvated electrolyte ions. The main focus has been the optimization of interparticle contact resistance and electrolyte wettability of the pores. The capacitance increase is almost 50% compared with the best performing commercially available activated carbons. Nanostructured metal oxide materials have also been investigated for supercapacitor applications, especially RuO_2 and $\text{RuO}_2 \cdot x\text{H}_2\text{O}$ nanotubes composites [20].

The Impact of Nanostructures on Traditional Energy Sources: An Overview

As the worldwide demand for energy rises at an ever-increasing rate, there is a new urgency to improve the efficiency and sustainability of actual power generation technologies. One of the keys to address this challenge is innovation, and some of the most promising solutions are occurring at the nanoscale. Perhaps the most mainstream acceptance within renewable technology has come in the form of wind energy, which is approaching cost competitiveness with traditional energy sources. Countries such as Germany, Spain, and Denmark are already beginning to utilize substantial amounts of wind energy to meet their growing electricity needs. Nanotechnology impacts the wind industry by improving turbine performance and reliability to allow for longer lifetime, less fatigue failure, and lower costs of generation. New lubricants containing nanoparticles act like mini ball bearings, thus reducing the friction generated from turbine rotation, consequently improving its life cycle. Advancements in nanocoatings, such as deicing and self-cleaning technologies, also help improving efficiencies, thus rendering ice and dirt buildup on the turbines virtually nonexistent.

In spite of the optimistic predictions of its proponents, alternative energy sources such as wind, hydro, and biomass are unlikely to seriously challenge coal, oil, natural gas, or nuclear power for a meaningful share of the worldwide energy needs in the near future. However, how these traditional energy sources are produced will rapidly change. In fact, nanotechnology is already at work modifying how coal, oil, and gas are being produced. One of the more prominent examples can be found in the application of new nanoscale catalysts to the production and refining processes of fossil-fuel resources. Moreover, modern fossil-fuel power plants, including coal-fired systems, require components that are able to withstand higher temperatures and pressures to improve efficiency, lower maintenance costs and decrease emissions levels. In this field, nanotechnology-based innovation is being credited with improvements in efficiencies and life spans of existing technologies, as well as the introduction of novel and disruptive power-generating components.

Nanotechnology has the potential to revolutionize the electric utility industry as well. Many of the



Nanostructures for Energy, Fig. 4 Structure of single-walled carbon nanotubes (SWCNTs); carbon nanotubes can be metallic or semiconducting, depending on how their chickenwire-like lattice rolls into a tube

components used in the electric utility industry can be modified to yield incremental improvements in the overall product performance. The end-benefit is that electric utility providers can improve their operating margins by making existing equipment both last longer and more efficient. Nanotechnology can provide immediate improvements to the electrical utility system of moving power by efficiently improving the transmission of electrons, as well as create newer and even better conducting materials. The real payoff may be carbon nanotubes (CNTs), which are unbelievably thin, possess 100 times the strength of steel, have only one sixth the density of aluminum, and efficiently conduct electricity. Most promising are the single-walled carbon nanotubes (SWCNTs). The broad range of uses for SWNTs is connected to the different nanotube properties according to how the hexagonal lattice is oriented (see Fig. 4), which together with the nanotube diameter determines the chirality. This combination of strength and conductivity suggests that in the future, energy providers should not have to dig up city streets to lay new wires. The existing lines could be simply replaced with super strong, highly conductive carbon nanotube wires.

Reduction of energy consumption is another active area in nanotechnology, targets being the development of more efficient lighting, better combustion systems, and/or use of lighter and stronger materials in the transportation sector. Light-emitting diodes (LEDs) are interesting examples of how energy consumption can be reduced; or “smart” windows, which can change reflective properties to attract or deflect sunlight depending on the customer’s needs. Similarly,

improved insulation materials could have a noticeable impact on overall energy demand. All of these products are made possible by advancements in nanotechnology. Nevertheless, the greatest long-term cost savings will probably come from the creation of “intelligent energy networks,” that is, integrated systems of more energy-efficient computers and sensors that gauge and regulate customer energy use. Such networks already exist today, but as nanotechnology enables computers and sensors to become smaller, but correspondingly more powerful, the cost effectiveness of deploying such devices will increase, along with the number of possible applications.

Conclusions

New materials hold the key to fundamental advances in energy production and storage, both of which are vital in order to meet the challenge of global warming and the finite nature of fossil fuels. Nanomaterials in particular offer unique properties or combinations of properties as electrodes and electrolytes in a range of energy devices. Through innovation, the efficiencies of present technologies can be improved and discover new ways by which mankind can prosper. Nanotechnology provides researchers with the opportunity to attain sustainable development, thus overcoming one of the greatest challenges of modern times, by using some of the simplest and smallest available means.

In conclusion, nanotechnology, known as paying attention to structures technology of the twentieth century, owing to benefits such as being potentially inexpensive, effective, and promising for to come. Hence, many companies and researchers are getting curious day by day in order to explore especially its energy application. Indeed, to minimize energy waste, it should become common and applicable as soon as possible in spite of the fact that it is quite a new technology.

Cross-References

- ▶ [Carbon-Nanotubes](#)
- ▶ [Hybrid Solar Cells](#)
- ▶ [Nanomaterials for Electrical Energy Storage Devices](#)
- ▶ [Nanomaterials for Excitonic Solar Cells](#)

- ▶ Nanostructured Materials for Sensing
- ▶ Piezoresistivity
- ▶ Thermoelectric Heat Convertors

References

1. US DOE Office of Basic Energy Sciences, Alivisatos, P., Cummings, P., De Yoreo, J., Fichthorn, K., Gates, B., Hwang, R., Lowndes, D., Majumdar, A., Makowski, L., Michalske, T., Misewich, J., Murray, C., Sibener, S., Teague, S., Williams, E.: Nanoscience research for energy needs. Report of the Department of Energy NanoSummit: Nanoscale Science and Our Energy Future, Wardman Park Marriott, NW, Washington, DC, 23–24 June 2004. This document is a work of the U.S. Government and is in the public domain. Printed in the United States of America, 2nd edn (2005)
2. Arunachalam, V.S., Fleisher, E.L.: Harnessing materials for energy. *MRS Bull.* **33**, 261 (2008)
3. Liu, J., Cao, G., Yang, Z., Wang, D., Dubois, D., Zhou, X., Graff, G.L., Pederson, L.R., Zhang, J.G.: Oriented nanostructures for energy conversion and storage. *ChemSusChem* **1**, 676–697 (2008)
4. Grätzel, M.: Recent advances in sensitized mesoscopic solar cells. *Accounts Chem. Res.* **42**, 1788–1798 (2009)
5. Crabtree, G.W., Dresselhaus, M.S.: Harnessing of materials for energy – the hydrogen fuel alternative. *MRS Bull.* **33**, 421–428 (2008)
6. Thiam, H.S., Daud, W.R.W., Kamarudin, S.K., Mohammad, A.B., Kadhum, A.A.H., Loh, K.S., Majlan, E.H.: Overview on nanostructured membrane in fuel cell applications. *Int. J. Hydrogen Energ.* **36**, 3187–3205 (2011)
7. Aricò, A.S., Bruce, P., Scrosati, B., Tarascon, J.M., van Schalkwijk, W.: Nanostructured materials for advanced energy conversion and storage devices. *Nat. Mater.* **4**, 366–377 (2005)
8. van Schalkwijk, W., Scrosati, B. (eds.): *Advances in Lithium-Ion Batteries*. Kluwer/Plenum, New York (2002)
9. Sartori, I., Hestnes, A.G.: Energy use in the life cycle of conventional and low-energy buildings: a review article. *Energ. Buildings* **39**, 249–257 (2007)
10. Cook-Chennault, K.A., Thambi, N., Sastry, A.M.: Powering MEMS portable devices – a review of non-regenerative and regenerative power supply systems with special emphasis on piezoelectric energy harvesting systems. *Smart Mater. Struct.* **17**, 1–33 (2008)
11. O'Regan, B., Grätzel, M.: A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO₂ films. *Nature* **353**, 737–740 (1991)
12. Vlachopoulos, N., Liska, P., Augustynski, J., Grätzel, M.: Very efficient visible light energy harvesting and conversion by spectral sensitization of high surface area polycrystalline titanium dioxide films. *J. Am. Chem. Soc.* **110**, 1216–1220 (1988)
13. Angrist, S.: *Direct Energy Conversion*. Allyn & Bacon, Boston (1982)
14. Snyder, G.J., Toberer, E.S.: Complex thermoelectric materials. *Nat. Mater.* **7**, 105–114 (2008)
15. Hassoun, J., Lee, K.S., Sun, Y.K., Scrosati, B.: An advanced lithium ion battery based on high performance electrode materials. *J. Am. Chem. Soc.* **133**, 3139–3143 (2011)
16. Long, J., Dunn, B., Rolison, D., White, H.: Three-dimensional battery architectures. *Chem. Rev.* **104**, 4463–4492 (2004)
17. Chan, C.K., Peng, H., Liu, G., McIlwrath, K., Zhang, X.F., Huggins, R.A., Cui, Y.: High-performance lithium battery anodes using silicon nanowires. *Nat. Nanotechnol.* **3**, 31–35 (2008)
18. Conway, B.E.: *Electrochemical Supercapacitors*. Kluwer/Plenum, New York (1999)
19. Frackowiak, E., Béguin, F.: Carbon materials for the electrochemical storage of energy in capacitors. *Carbon* **39**, 937–950 (2001)
20. Kim, I.H., Kim, J.H., Kim, K.B.: Electrochemical characterization of electrochemically prepared ruthenium oxide/carbon nanotube electrode for supercapacitor application. *Electrochem. Solid St. E.* **8**, A369–A372 (2005)

Nanostructures for Photonics

Andrea Toma¹, Remo Proietti Zaccaria¹, Roman Krahne¹, Alessandro Alabastri¹, Maria Laura Coluccio^{1,2}, Gobind Das¹, Carlo Liberale¹, Francesco De Angelis¹, Marco Francardi^{2,3}, Federico Mecarini¹, Francesco Gentile^{1,2}, Angelo Accardo^{2,4}, Liberato Manna¹ and Enzo Di Fabrizio^{1,2}

¹Nanobiotech Facility, Istituto Italiano di Tecnologia, Genoa, Italy

²Lab. BIONEM, Dipartimento di Medicina Sperimentale e Clinica, Università “Magna Grecia” di Catanzaro, Catanzaro, Italy

³International School for Advanced Studies (SISSA), Edificio Q1 Trieste, Italy

⁴Soft Matter Structures Group ID13 – MICROFOCUS Beamline, European Synchrotron Radiation Facility, Grenoble Cedex, France

Definition

“Nanostructures for Photonics” represents a collection of devices whose working principles are based on photons at the nanoscale level.

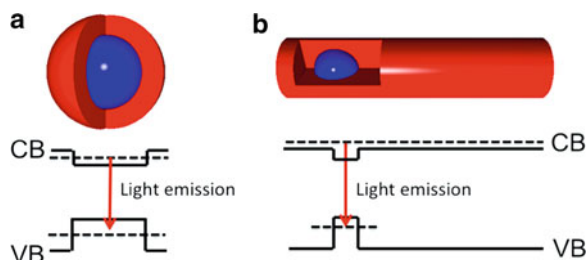
Overview

In the last decade, new disciplines emerged where the terms *nano* and *photon* happen to play a fundamental

descriptive role. Their origin dates back to the ancient Greece, with *nano* ($=\nu\alpha\nu\omicron\zeta$) indicating something of small dimensions, and *photon* ($=\phi\omega\psi\rho\zeta$) indicating the light. In Physics, but also in Chemistry, Biology, Material Science, and many more fields these terms have been used for indicating the studies of extremely small objects by means of light. This new technological frontier has already produced some amazing results: LCD based mobile phones, nanoparticles for optical tagging of biological molecules and hyperthermia, or devices for single-molecule detection are just few of many examples where nanophotonics plays a crucial role. In terms of optical devices, their functionality can be divided in three general key parts: *light generation*, *light guiding* and *light harvesting*. Here will be introduced some approaches describing these aspects, starting from colloidal nanocrystals for photonics to end, through the introduction of photonic crystals and metamaterials, with plasmonic devices for Raman spectroscopy. The aim is to prove that photonics at nanoscale level is not just an exotic research field but a mature discipline capable of offering devices for a multitude of uses.

Light Generation: Semiconductor Nanocrystals for Lasing Applications

Colloidal semiconductor nanocrystals are a very attractive alternative as light emitting materials to more conventional nanostructures, because they can be fabricated in a cost-effective bottom-up approach by wet chemical synthesis. In the last decade, great progress has been achieved regarding the control on nanocrystal shape and composition, leading to size dispersions of less than 5% in spherical particle samples, and to rod- and multibranching shapes. One particular property of colloidal nanocrystals is their exposed surface that needs to be stabilized via surfactant molecules. The related obstacle that has to be overcome toward highly efficient light emission is the nonradiative recombination of the electrons and holes at surface states. The most promising solution to this problem are core-shell architectures, where a low band gap material nanocrystal is capped by a second material with higher band gap. In such structures, the photoexcited charges are confined by the nearly defect-free internal interface between the two materials (see band alignment illustration in Fig. 1).



Nanostructures for Photonics, Fig. 1 Schemes illustrating the core-shell architectures and the related band structure of spherical (a) and dot/rod nanocrystal structures (b)

Here, the emission wavelength is determined on one hand by the band gap of the core material, and on the other hand by the quantum confinement effects, since typical diameters of nanocrystals are in the range of 1–8 nm. The most popular systems for light emission in the visible consist of CdSe, CdS, ZnSe, and ZnS materials. For example, the emission of CdSe based systems can be tuned by the nanocrystal size in the range from 1.8 to 2.5 eV, as demonstrated in Fig. 2a.

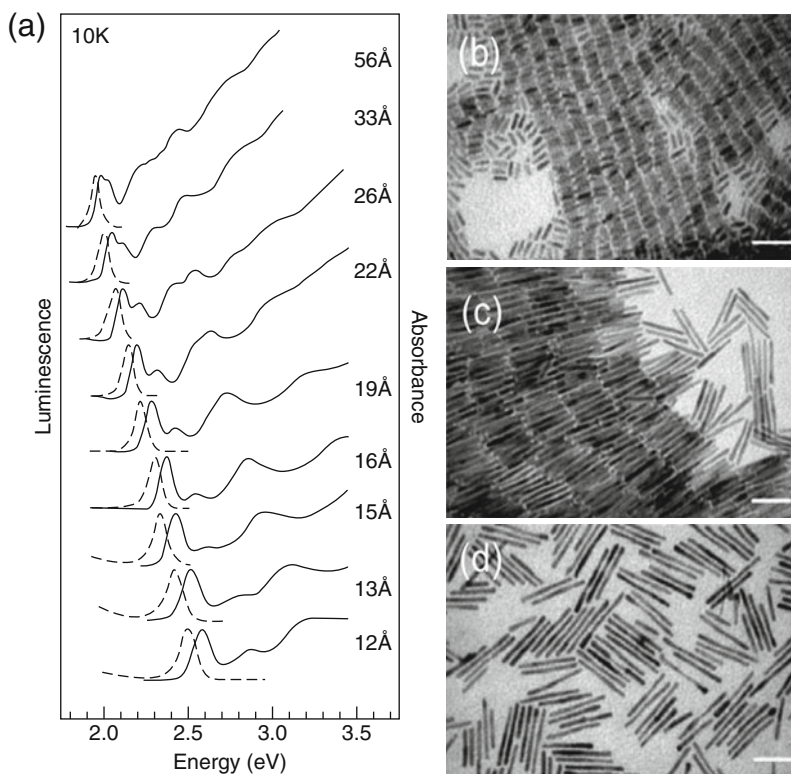
Two very successful core-shell architectures are sketched in Fig. 1, the spherically symmetric core-shell structure, and the dot-in-a-rod (dot/rod) structure that can be fabricated by the seeded growth technique [1]. The dot/rod architecture in Fig. 1b combines the advantageous optical properties of rod-shaped nanocrystals, such as strongly linearly polarized emission and oriented assembly, with a high quantum yield (ratio photons in/photons out) comparable to that of spherical nanocrystals (Fig. 2).

For lasing devices two factors are important: the optical gain of the light emitting material and the quality of the resonant cavity that provides the optical feedback. The optical gain can be improved by the composition and architecture of the nanocrystals, and by the homogeneity and density of their assembly. The optical feedback can be obtained from semitransparent mirrors or distributed feedback resonators that consist of Bragg reflectors. The optical gain of the material is typically characterized by transient (time-resolved) optical absorption spectroscopy, which reveals the population dynamics of the energetic levels of the nanocrystals [3]. If the emitting ground state population maintains an occupation larger than unity, optical gain is achieved. One peculiar property of the dot/rod core-shell nanocrystals is that optical gain can be obtained from the optical transitions related to the core and from the transitions in the shell material.

Nanostructures for Photonics, Fig. 2

(a) Emission and absorption spectra of spherical CdSe nanocrystals (With permission from Ref. [2]). (b)

Transmission electron microscopy images of dot-in-a-rod core-shell nanorods with similar core size and different rod length. The scale bar corresponds to 50 nm (With permission from Ref. [1])

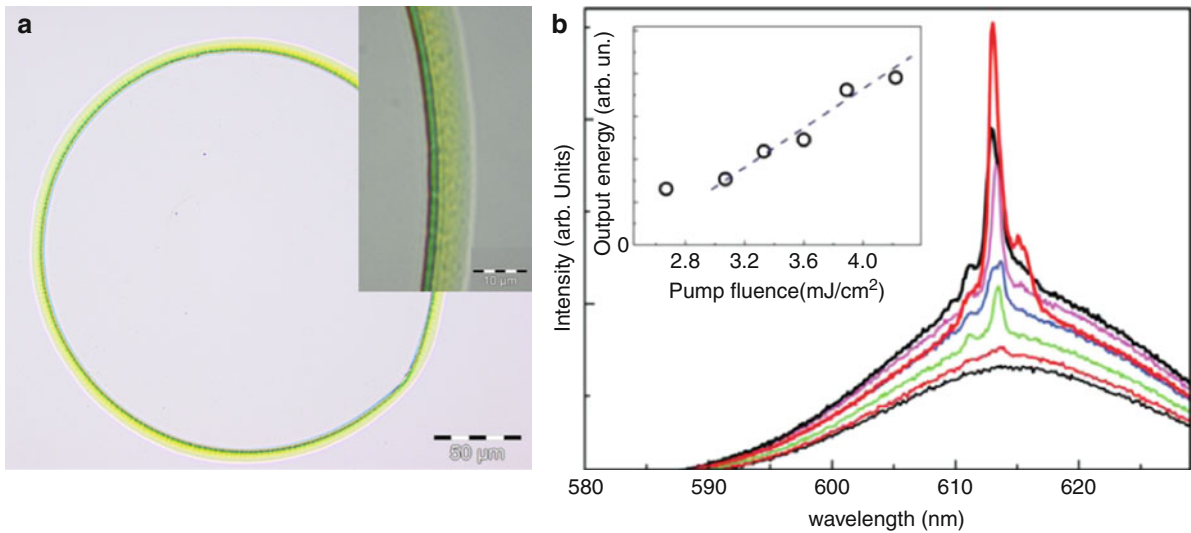


This effect extends the spectral range significantly, for example in the CdSe/CdS system from 570 to 650 nm (core emission) to 490 nm emission wavelength of the shell [4]. So far, the fabrication of efficient optical cavities for nanocrystal assemblies remain challenging since, for example, conventional sequential evaporation of Bragg reflectors does not work due to the incorporation of the inhomogeneous nanocrystal layer. Hence, novel approaches to obtain cavities have been sought, like the loading of cylindrical microcavities with nanocrystal solutions via the capillary effect [5], or the hot embossing of two pre-fabricated Bragg reflectors which embed the nanocrystal layer in a sandwich structure [6]. Another innovative approach consists in the laser emission from core-shell nanorod assemblies where the layer of nanorods itself formed the resonant cavity [3]. Here, circular rings with micron scale width and sharply defined lateral edges were formed via the coffee stain effect by simple deposition of nanocrystals dissolved in toluene (see Fig. 3a). The lateral edges of the rings themselves functioned as optical interfaces and were suitable for achieving feedback in a Fabry–Perot resonator scheme (Fig. 3b). Such microlasers

could be very interesting for a new generation of photonic devices integrated onto substrates for lab-on-a-chip or for other compact, integrated applications.

Light Guiding: From Photonic Crystals to Metamaterials

Transportation of light to a detector is one of the fundamental aspects of any optical device. The light, during its path, might be affected by the surrounding world thus undergoing a modification of its properties. This change can be temporary, namely existing only during the path toward the detector or even permanent. An example of the first kind is given by the effective wavelength of surface plasmon polaritons on a metallic interface whereas permanent modification can be well described by white light entering a fiber. In fact, in this case the light will lose most of its spectral components once the detector will be reached. A further example is given by metamaterials, usually defined as exotic devices capable of strongly modifying the standard properties of the light. These different situations indicate a strong relation between the intended



Nanostructures for Photonics, Fig. 3 (a) Optical microscope images of a self-assembled *circular* ring formed via coffee stain evaporation dynamics. The ring consists of densely packed core-shell nanorods and acts as a microcavity for laser feedback.

(b) Emission spectra from a Fabry–Perot self-assembled microcavity made of nanorods with 25 nm length recorded at different pump fluences. Inset: emission intensity as a function of the pump fluence (With permission from Ref [3])

application and the method of transportation of light which must be properly chosen.

Periodic and Quasiperiodic Fibers

A standard and well-established method for the transportation of light is by means of optical fibers. These devices are nowadays commonly used for applications such as Internet, allowing a high-speed data transfer (Fig. 4).

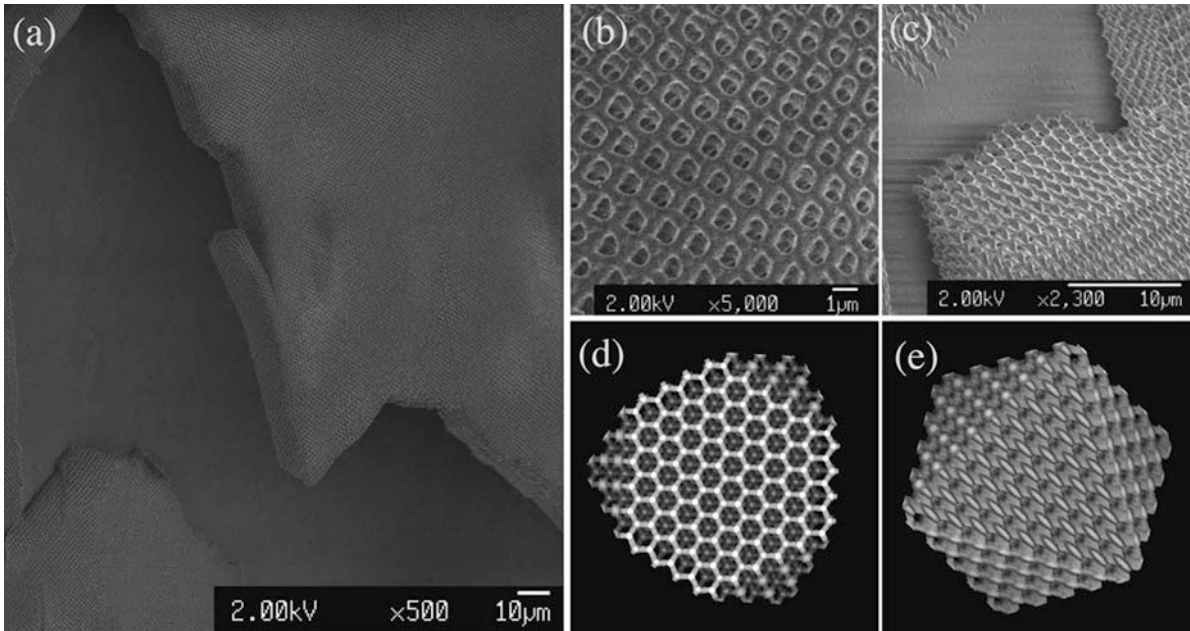
A different kind of fiber known as Photonic Crystal Fiber (PCF) was introduced in [8] with the idea of improving the classical fibers in many aspects, such as single-mode operation regime in a wide range of wavelengths, air-core guidance or reduced cross-sectional dimensions. In fact, while standard fibers have diameters of some hundreds of micrometers, PCFs can be reduced down to few microns [9–11].

The physics behind the working principle of PCFs fibers lies in their geometrical characteristics. Photonic crystals (PCs) are structures showing a *translational* periodicity of the medium permittivity [12, 13]. A three dimensional example of a face centered cubic PC is given in Fig. 4. The main characteristic of a PC is the capability of controlling the interaction between radiation and matter in order to obtain new physics. For example, they can induce a local modification of the field, which in turn will affect the optical

characteristics of samples embedded in the crystal. For the present purposes, these devices will be seen either as filters or waveguides. In fact, according to their geometrical and material properties, PCs can realize optical gaps in specific spectral regions. Consequences are the impossibility of the light to propagate inside the crystal. Figure 5a is an example of a square-like photonic crystal waveguide where light at 2.85 μm is obliged in following the free path of the crystal. The structure is formed by Silicon cylinders of 200 nm radius in air. The periodicity (lattice constant) of the crystal is 1 μm . When the optical gap concept is associated to the fiber world, Photonic Crystal Fibers are obtained. In this case, light propagates along the axis of the cylinders as in Fig. 5b.

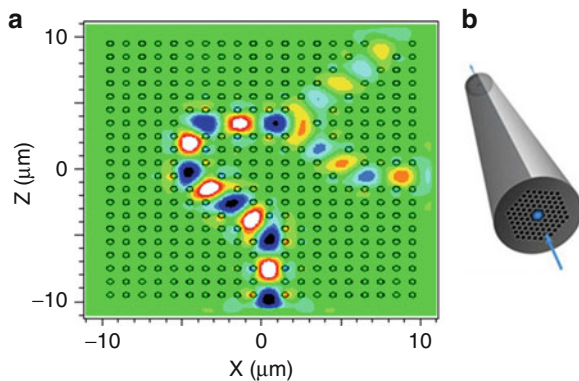
Among the PCFs special attention deserves the Photonic Quasicrystal Fibers (PQFs). Their geometry is similar to photonic crystals with the fundamental characteristic of loosing the translational periodicity. Only rotational and mirror symmetry are proper of a quasicrystal [14, 15]. Figure 6 gives some examples of quasicrystal. One of the advantages of considering PQFs versus PCFs is in a wider single-mode operation regime than the PCFs, maintaining similar cross-sectional dimensions.

The typical approach for determining the optical characteristic of a fiber is the calculation of its



Nanostructures for Photonics, Fig. 4 The scanning electron microscope image of the scaffold type a fcc optical lattice. The lattice constant is about 1.5 μm . (a) A bird's eye view. (b, c) Magnified images of (b) *top* surface of the structure, and (c)

cross section of a crack in the rim of the structure. (d, e) Simulated cross section of fcc lattice in (a) $\langle 1,1,1 \rangle$ and (b) $\langle 1,1,-1 \rangle$ (With permission from Ref. [7])



Nanostructures for Photonics, Fig. 5 (a) Two-dimensional photonic crystal waveguide. Transverse magnetic polarized light at 2.85 μm coming from the *bottom* of the figure travels along the free path. The period (distance between two Silicon cylinders) is 1 μm , the radius is 200 nm. (b) Photonic crystal fiber. The *blue arrow* indicates the direction of propagation of the light

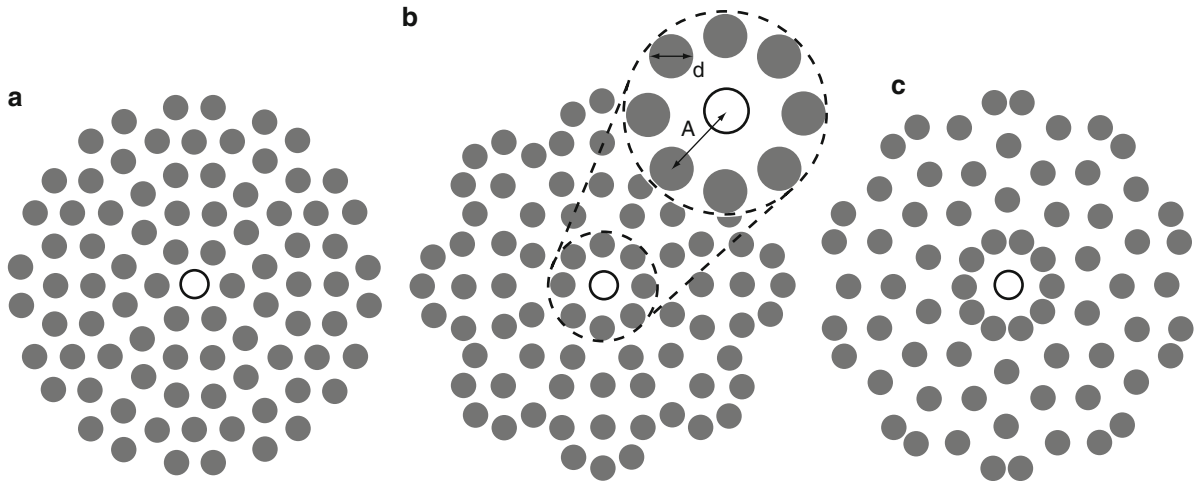
V parameter. This value is usually considered as the quantity capable of identifying the single-mode operation regime of a fiber. In fact, in the case of single-mode fiber, for a given wavelength and polarization, only a particular mode can propagate. This avoids the problem of the destructive interference that occurs in

a multimode fiber, which makes single-mode fibers the best choice for applications related to long distance communications. This explains why fibers working as single mode in a wide range of frequencies are very desirable for a number of applications. However, in case of PQFs the V parameter method has been demonstrated not to be working in many cases [15]. For this reason, a different approach consisting in the direct calculation of the effective area of the second-order mode is preferable. Figure 7 shows a comparison between these two approaches for three kinds of fibers.

Besides a disagreement between the V parameter and the effective area approach as in Fig. 7b, c, it is possible to notice that the single-mode operation regime for both the eightfold and tenfold fibers is much more extended than the sixfold one. This makes these kinds of fiber more robust for transporting signal at long distances.

Fishnet Structures: A First Step Toward Metamaterials

Metamaterials are artificial man-made materials exhibiting particular properties such as negative refraction [16–19], which can lead to several

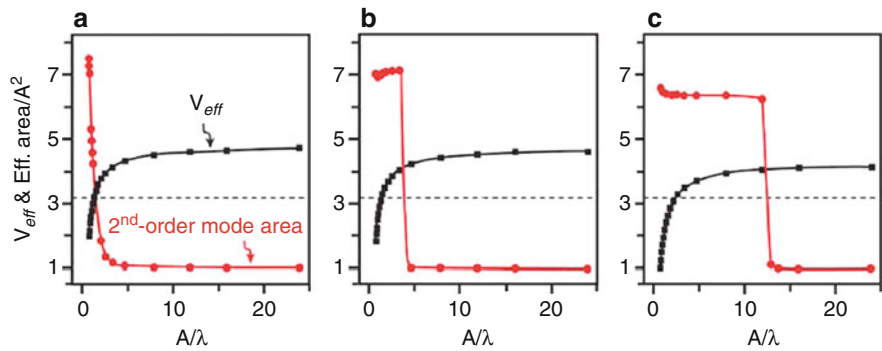


Nanostructures for Photonics, Fig. 6 Cross section of three kinds of PQFs: (a) sixfold, (b) eightfold, and (c) tenfold. Gray filled circles denote air holes drilled in silica fiber. The central

hole (empty circle) is absent in the fibers. The quantities A and d represent the lattice constant and the hole diameter, respectively (With permission from Ref. [15])

Nanostructures for Photonics, Fig. 7

Effective V parameter (black squares and curves) and the second-order mode normalized area (red circles and curves) as a function of A/λ for (a) sixfold, (b) eightfold, and (c) tenfold quasicrystal fibers. The dashed lines are drawn at π . A is the lattice constant (With permission from Ref. [15])

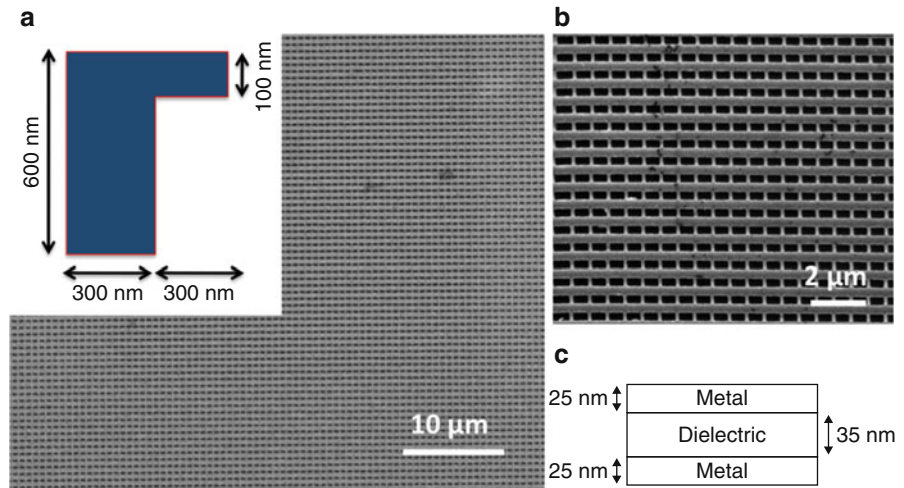


applications like the realization of superlenses or invisibility cloak. In order to reach a negative refractive index, simultaneously negative electric permittivity and magnetic permeability need to be reached as pointed out by Veselago in 1968 [16]. A negative permittivity is a relatively easy task to obtain, since all metals below their plasma frequency exhibit it. On the other hand, a negative permeability is more difficult to achieve; the lack of magnetic charges and the weak interaction between matter and the magnetic part of light, make scientists capable to attain it only in certain frequency regions. Combining thin metallic wires and rods in a three layers Metal/Dielectric/Metal structure (as shown in Fig. 8), constitutes the so called *fishnet structure* which has been proved to exhibit peculiar characteristics both in metamaterials and plasmonics fields.

Fishnet structures have been widely investigated in metamaterials research field both from an experimental and a theoretical point of view [20]. The reason is that their geometry is simple and its fabrication is relatively easy. However, the mechanism behind the negative refractive index behavior is nowadays not completely understood. The main idea is that properly polarized light passing through the structure can experience at the same time both negative permittivity and permeability. Considering, for instance, an impinging electromagnetic field polarized as shown in Fig. 9b: the electric field oscillates along thin wires allowing a negative electric permittivity and the magnetic field can induce a resonance through the aligned sets of rods, just like in a LC circuit. In fact, whereas the capacitor-like response is directly given by the vertical structure of the fishnet (two metallic plates separated

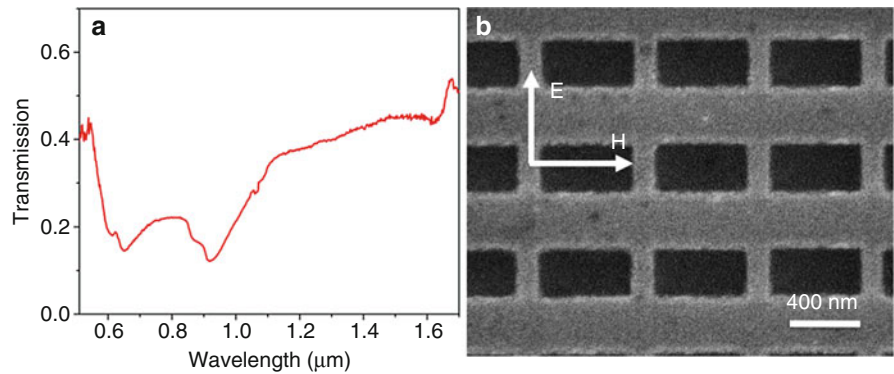
Nanostructures for Photonics, Fig. 8

(a, b) SEM images of a fabricated fishnet structure. (c) Sketch of the vertical structure. The inset in (a) shows the cell unit of the fishnet structure



Nanostructures for Photonics, Fig. 9

(a) Transmissivity spectrum for the presented fishnet structure. (b) SEM micrograph of measured fishnet structure and fields direction of the incident light



by a dielectric, see Fig. 8c), the inductive response is expressed by the electric currents, generated by the oscillating magnetic field, that circularly flow in the plates (as regular currents) and across the dielectric (as displacement current). The resonant behavior can be observed, for example, in the transmission spectrum (Fig. 9a): if the magnetic resonance occurs, a dip in the transmission coefficient can be noticed.

Before performing the experiments, also according to previous studies where similar structure were investigated, a magnetic resonance between 1,300 nm and 1,500 nm was expected [20]. In Fig. 9b, this kind of resonance is not clearly visible, probably because of high losses in that spectral range which damp transmission or reflection features. Nevertheless, it is possible to notice a resonant behavior at shorter wavelengths. Indeed current loops, which are supposed to induce the magnetic resonances, are generated by internal gap surface plasmons polaritons (SPPs). This

means that, more than one resonance exist, actually several of them may occur as far as the matching condition subsists:

$$k_{SPP} = k_{wave} + mG_x + nG_y \text{ with : } G_x = \frac{2\pi}{a_x} \text{ and } G_y = \frac{2\pi}{a_y}$$

In this relation, k_{SPP} and k_{wave} are, respectively, the wave vector of the excited SPP and the one of the incident light. a_x, a_y are the lattice periodicities along the x and y fishnets axis. Therefore, higher frequency resonances can be reached as far as the lattice provides the missing momentum through G_x and/or G_y . However, a negative refractive index is not guaranteed in the entire spectrum since the equivalent plasma frequency of the medium is dominated by the cutoff wavelength of the holes of the fishnet.

To conclude, fishnet structures represent a good example on the concept of metamaterials and how

their fascinating properties are related to plasmonic effects. A deep study of these structures can be considered a first step to understand the underlying physics of negative refraction enabling, eventually, the realization of more complex devices exploiting the effects of negative index materials.

Light Harvesting: Supported Metal Nanoarrays for Plasmonic Devices

Light harvesting plays an important role for the operation of many kinds of devices, such as biosensors or photovoltaic cells. In fact, besides the generation and transportation of light, an efficient coupling mechanism between the guiding device and the investigated bio-sample in order to produce a readable signal is a fundamental step of a generic photonic device. For example, metallic nanostructures have the property of realizing a high localized electric field which can be efficiently coupled to a bio-sample. This, in turn, means high signal-to-noise ratio from the sample, namely a good signal for the detector can be produced.

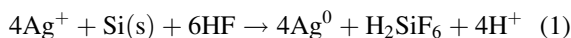
Metallic nanoparticles exhibit specific optical properties in the visible/near-infrared range due to the collective excitation of the conduction electrons (Localized Surface Plasmon Resonances – LSPRs). The field enhancement provided by metallic nanostructures has been used to manipulate light-matter interactions and enhance nonlinear phenomena. The strong effects associated to the LSPR excitation on metal nanoparticles are the core of many interesting perspective applications such as single-molecule detection and photovoltaic conversion. Here two specific examples are introduced, that is, nanoaggregate and nanoantenna arrays where light harvesting and field enhancement will be exploited with respect to high sensitivity Raman spectroscopy.

The use of advanced lithographic techniques allows precise control of particle shape and size, as well as interparticle separation [21, 22]. Electron beam lithography (EBL) has become the method of choice for fabricating single or periodic arrays of metal nanostructures endowed with the desired morphological features [23]. It is versatile and provides high resolution and precise control over the geometry and separation of nanostructures, guaranteeing a fabrication reproducibility, and precision down to the nanometer scale.

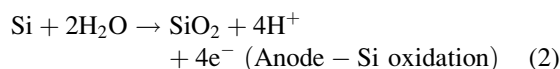
Nanoaggregate Array

The fabrication of an active SERS device was achieved by means of site selective deposition of Ag nanoparticles, exploiting electroless technique on Si pre-patterned substrates [23]. A large array of nanostructures with high degree of reproducibility at the nanometer scale was realized using electron beam lithography onto a Si substrate spin-coated with 50 nm thick ZEP layer.

The Ag deposition was realized by means of a redox reaction between the metal and the Si surface remained uncovered after the lithographic process. The reaction occurs in a dilute (0.1–0.5 M) HF solution in which is dissolved a metal salt, that is, AgNO_3 , at a concentration of about 1 mM. The global chemical reaction is:



which in its turn can be separated into two half-cell reactions:

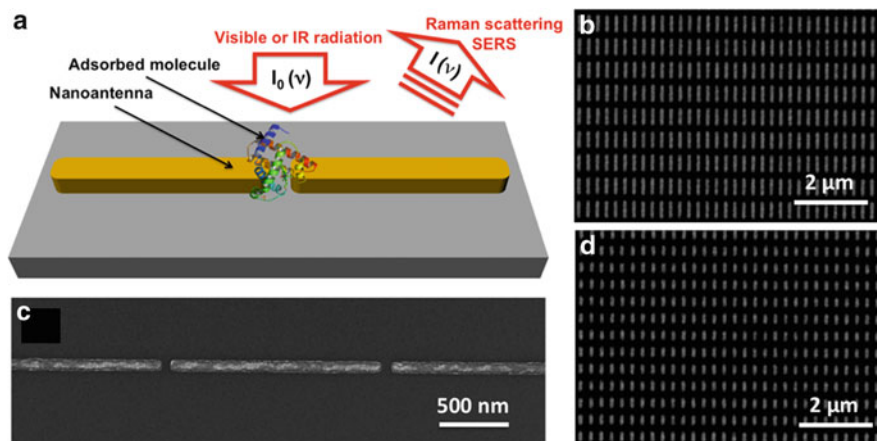
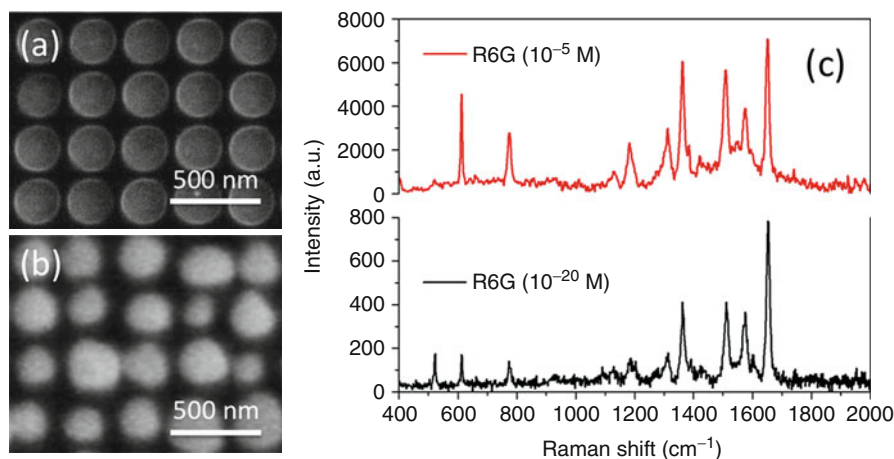


The driving force of the whole process is coming from the potential difference in the two half-reactions, thus fostering the growth of Ag grains and the formation of well-organized metallic nanostructures (see Fig. 10a, b).

In order to probe the enhancement behavior in Raman spectroscopy related to the plasmon excitation, the sample was immersed in Rhodamine 6G (R6G) solutions with different concentration (10^{-5} M and 10^{-20} M in H_2O) for 30 min. The sample was gently rinsed in DI-water to remove molecules in excess not chemisorbed on the metal surface and then dried in N_2 flow.

In Fig. 10c, Raman spectra of R6G molecules, adsorbed on a Si substrate with a 60 s Ag deposition, are shown. There are various bands for R6G compound at around 1,650, 1,509, and 1,361 cm^{-1} , attributed to the xanthenes ring stretching of C-C vibrations, 1,575 cm^{-1} attributed to C-O stretching; while the band at 1,183 cm^{-1} is related to the C-H bending and N-H bending vibration of xanthenes ring. As can be clearly observed in Fig. 10c, the Raman intensity

Nanostructures for Photonics, Fig. 10 SEM micrographs of (a) array of empty holes fabricated by EBL technique and (b) Ag nanoaggregates selectively grown inside the nanohole pattern. (c) SERS spectra acquired from 10^{-5} to 10^{-20} M R6G, adsorbed on the silver structures



Nanostructures for Photonics, Fig. 11 (a) The nanoantenna concept: fabrication of a highly sensitive and specific nanobiosensor based on extraordinary optical signal enhancement. (b, d) Representative SEM images of two different

nanoantenna arrays fabricated by EBL technique. The dimensions of the NAs are of length 410 nm (b), 200 nm (d) while width and height both are 60 nm for all NA. (c) SEM of coupled nanoantenna array

decreases with the decreasing concentration of R6G, as expected. Though, the concentration is very low, various bands are clearly visible even for R6G molecule with concentration of 10^{-20} M, keeping accumulation time 60 s and laser power 0.0018 mW.

Nanoantenna Array

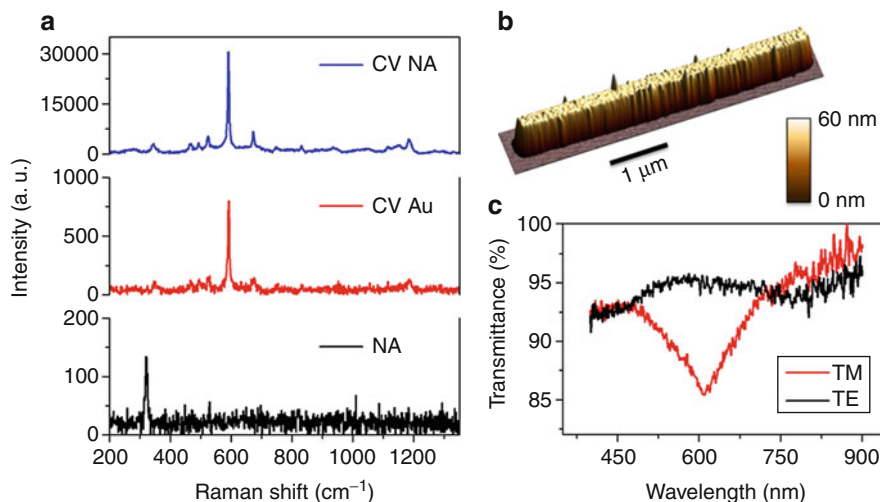
The ability to tailor the optical properties of structured plasmonic materials has proven to be crucial for a wide variety of applications, using them for surface-enhanced spectroscopy and biosensing. While a symmetrical geometry allows polarization independent plasmon excitation, a dichroic absorption is well expected for nanoantenna (NA) fabrication. Large scan

overview of different nanoantenna patterns are reported in Fig. 11. The sample topography has been characterized recurring to scanning electron microscopy (SEM) and atomic force microscopy (AFM) operated in tapping mode.

The optical properties of the nanoantenna array, shown in Fig. 12c, were investigated by means of spectroscopic transmission of polarized light in the range between 400 and 900 nm. The polarization of the incident light was varied from transverse magnetic TM (electric field perpendicular to the short NA axis) to transverse electric TE (electric field parallel to the long NA axis). The optical transmittance spectra present evidence of a clear anisotropic behavior; for TM

Nanostructures for Photonics, Fig. 12

(a) SERS spectra of cresyl violet molecules deposited on NA sample (blue line) and flat Au film (red line). Background measurement performed on bare NA structure is also reported (black line). (b) AFM 3D image. (c) Transmission optical spectra of the NA array of Fig. 1.11b. Red/black lines are for TM/TE polarization respectively



polarization a localized minimum around 620 nm is found. This is the typical behavior exhibited by subwavelength metal nanoparticles sustaining a localized surface plasmon resonance.

Evidence of plasmon-induced SERS effects is given in Fig. 12a in which the different Raman emission lines from cresyl violet (CV) molecules (with a concentration of 3.46 μM) deposited on the gold nanoantenna arrays are shown. The sample was excited by 633 nm laser wavelength (laser power = 0.14 mW and accumulation time = 50 s) through a 150X objective. The measurements were performed on bare nanoantenna arrays (sample NA), on cresyl violet molecules chemisorbed on flat Au film (sample CV Au) and on CV deposited on the nanoantenna arrays (sample CV NA). To ensure the quality of the sample surface, the background measurements were performed at different positions of the nanoantenna SERS device. As shown in Fig. 12a (NA line) no characteristic Raman band, except at around 320 cm⁻¹ related to the Ca-F vibration from the substrate, is observed. The characteristic vibrational bands of CV are observed in the SERS spectrum (CV NA trace). Intense Raman bands centered at around 591, 882, 927, and 1,189 cm⁻¹ can be attributed to the N-H₂ rocking vibration, two benzene group bending, out-of phase N-H₂ rocking vibration, and combination of N-H₂ rocking and C-H_x rocking, respectively.

CV molecules chemisorbed over flat Au film are also shown in Fig. 12a. A giant enhancement in Raman signal for CV NA sample with respect to the CV Au

one can be clearly observed. The maximum enhancement of the Raman emission is obtained when the excitation is polarized along to the nanoantenna short axis, in correspondence to the resonant excitation of the LSP resonance. The evaluated SERS enhancement for the fabricated nanoantenna device is 10⁷.

Novel Nanophotonic Devices

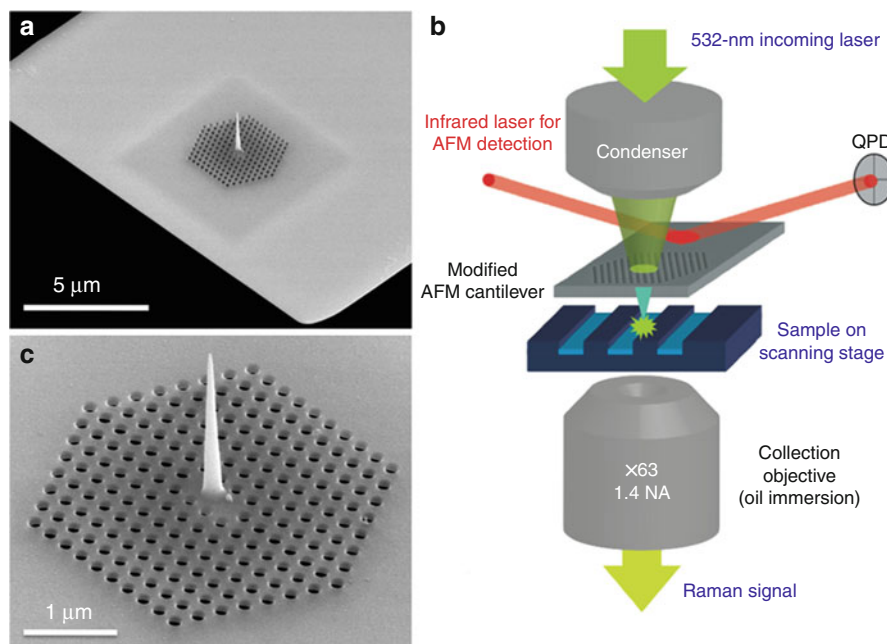
The three standard aspects of a photonic device, namely light generation, guiding, and harvesting can be combined in a number of ways to obtain many kinds of functionalities. For example, metallic-nanorods can be used for shrinking the wavelength of the light to improve the coupling with detectors such as nanoantenna arrays, or optical fiber can be properly modified for three-dimensional optical trapping [24]. Here will be introduced two specific examples: photonic crystals combined to plasmonic devices for the realization of background-free Raman probes, and superhydrophobic devices for single-molecule detection.

Integration of Photonic Crystals and Plasmonic Devices for Label-free Chemical and Structural Detection

A combination of different nanostructures for photonics give rise to a novel photonic-plasmonic device that allows label-free detection, through Raman spectroscopy, of single inorganic nanoparticles and of monolayers of organic compounds [25] and that allows,

Nanostructures for Photonics, Fig. 13

(a) SEM image of the device fabricated on a silicon nitride AFM cantilever. (c) Zoom on the photonic crystal cavity and the tapered plasmonic waveguide with radius of curvature at the apex of 5 nm. (b) Sketch of the experimental setup, showing the integration of the photonic–plasmonic device into an AFM–Raman microscope



integrated on an atomic force microscope (AFM) cantilever, to obtain topographic, chemical, and structural information at an unprecedentedly high spatial resolution (below 10 nm) [26].

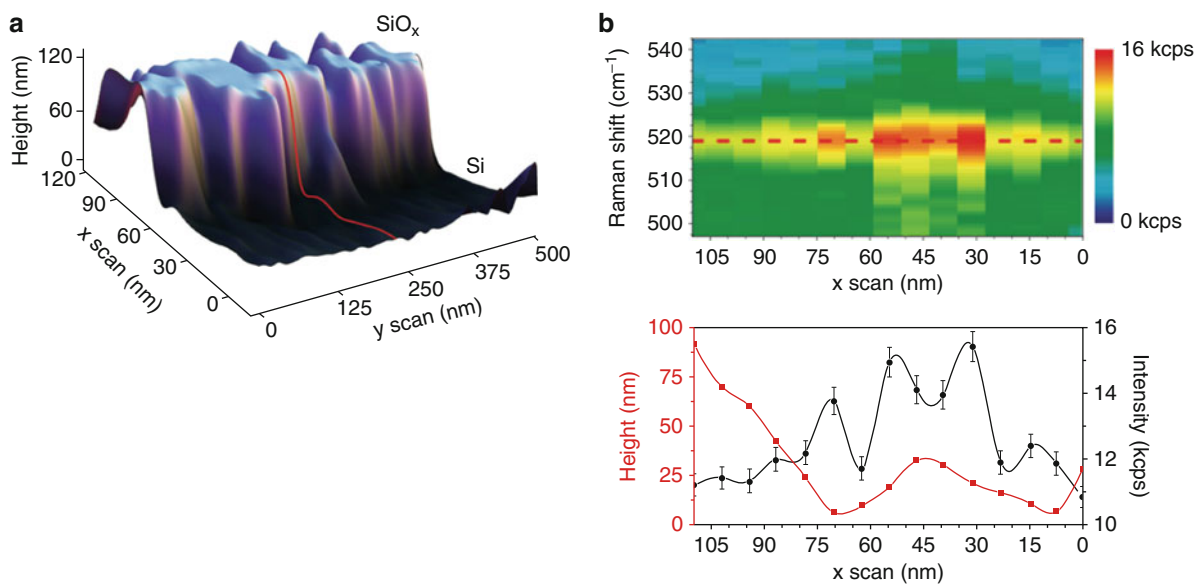
The device consists of a two-dimensional dielectric photonic crystal cavity, which can be patterned on a Si_3N_4 AFM cantilever, together with a tapered silver plasmonic waveguide placed at the center of the cavity (Fig. 13a). The photonic crystal consisted of a triangular lattice of air holes (lattice constant = 250 nm, hole diameter = 160 nm) patterned on a 100-nm-thick Si_3N_4 membrane. Three missing holes in the center generated a photonic crystal cavity, termed L3, tuned at $\lambda = 532$ nm. The plasmonic waveguide is a 2.5- μm -tall metallic cone, with a base diameter of 300 nm, and the minimum radius of curvature at the apex is controllable in the range 2.5–5 nm (Fig. 1c). The entire device is fabricated by means of Focused Ion Beam milling and electron-beam induced deposition [27]. Ion beam milling was used to define the photonic crystal whereas the silver tapered waveguide was grown in the center of the cavity using electron-beam induced deposition from a gas precursor containing a platinum–carbon polymer $(\text{CH}_3)_3\text{Pt}(\text{C}_p\text{CH}_3)$.

The photonic crystal cavity enables an efficient coupling between the external laser source (used for Raman excitation) and the tapered waveguide, together with optimal sample illumination. In fact,

the focal plane from which the SPPs are launched is 2.5 μm away from the cone apex where the Raman excitation is generated. This drastically lowers the background intensity at the sample plane. Thus, the spatial separation between SPP generation and Raman excitation efficiently increases the signal-to-noise ratio in the Raman spectrum, giving rise to a new optical configuration, namely Surface Plasmon Polariton Enhanced Spectroscopy (SPPERS). The focusing of SPPs follows an adiabatic compression mechanism, which causes a strong localization of the electrical field in a region with a size comparable to the apex radius of curvature [28]. This device can be fully integrated with an inverted Raman microscope combined with an AFM stage, as shown in Fig. 13b.

The SPP spatial confinement in the near field is comparable to the radius of curvature of the tapered waveguide apex, which lies between 2.5 and 5 nm. The expected electric field enhancement E due to adiabatic compression is defined as $E = E_{\text{tip}}/E_{\text{base}}$, where E_{tip} and E_{base} are the maximum electrical fields at the apex and bottom of the waveguide, respectively, and reaches values of practical interest in the range of 100 when the radius of curvature of the tip is below 5 nm. For successful spectroscopy applications, this necessitates state-of-the-art fabrication quality.

The architecture of the device and the mechanical stiffness of the tapered waveguide allow using it both



Nanostructures for Photonics, Fig. 14 (a) SEM image of the device fabricated on a silicon nitride AFM cantilever. (c) Zoom on the photonic crystal cavity and the tapered plasmonic waveguide with radius of curvature at the apex of 5 nm. (b) Sketch of

the experimental setup, showing the integration of the photonic–plasmonic device into an AFM–Raman microscope (With permission from Ref. [26])

as an AFM tip and as a nanometer light source for near-field Raman excitation. In other words, it is possible to obtain topographic and chemical mapping to yield structural and stress information with the same spatial resolution as is imposed by the geometry of the tip, that is, limited only by its radius of curvature.

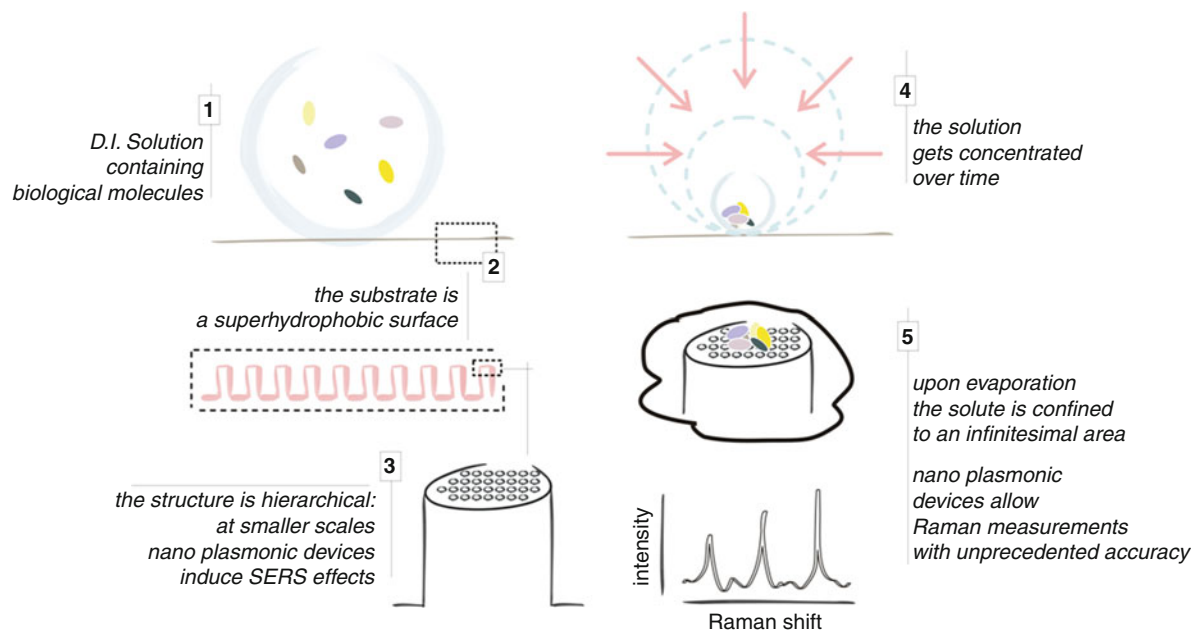
Device performances have been assessed by measuring a test sample for which the chemical composition had periodic spatial variations in the range of few nanometers. The test sample consisted of patches and gratings of silicon nanocrystals having different pitches (from several micrometers to a submicrometer scale) obtained by laser melting of a silica substrate previously fabricated by plasma-enhanced chemical vapor deposition.

The AFM–Raman setup was operated in tapping mode and wet conditions (in distilled water) to measure simultaneously the topography and Raman intensity map by scanning the sample and acquiring Raman data across the lithographic structures, point by point.

A severe test on the described sample was performed by scanning the sample in a patterned region in which the crystalline and amorphous topography forms a grating with submicrometer pitch. From the AFM measurements, it could be seen that silicon,

upon laser-induced crystallization, has its height reduced by 100 nm, which results in the formation of a trench on the sample surface, as shown in the three-dimensional AFM map in Fig. 14a. The AFM scan step size was 7 nm. The red line in Fig. 14a is a representative AFM line scan which provides simultaneously the Raman intensity map shown in Fig. 2b (upper panel), in which the Raman spectrum extends from 504 to 540 cm⁻¹. The acquisition time of each Raman map was 5 s. Fig. 14b (lower panel) shows the topography and Raman intensity for the same line scan. The Raman intensity is taken at the maximum of the silicon phonon peak at 520 cm⁻¹. The Raman intensity and relative peak amplitudes change from point to point, indicating an optical resolution of better than 7 nm.

In conclusion, the development of a device based on the combination of photonic–plasmonic nanostructures that is fully compatible with AFM and Raman spectroscopy opens up a number of significant opportunities in the nanoscale chemical mapping of materials. The approach could be of particular benefit in the nanoscale analysis of biological matter, such as in the investigation of cell membrane protein spatial distributions, in which chemical/physical bond strengths and amino acid composition *in vivo*



Nanostructures for Photonics, Fig. 15 Biophotonic nanostructures integrated with superhydrophobic surfaces can lead to the fabrication of novel devices with advanced sensing capabilities

conditions could be revealed. Device architectures based on the adiabatic generation of SPPs have considerable potential for applications in which a reliable nanometer-sized light source is of utmost importance.

Integration of Biophotonic and Superhydrophobic Devices for the Detection of Low-concentrated Biomolecules

Here, will be investigated how biophotonic nanostructures can be integrated to superhydrophobic surfaces (SHSs) to obtain devices with advanced sensing capabilities, thus offering new possibilities in the field of clinical medicine, especially toward the early diagnosis.

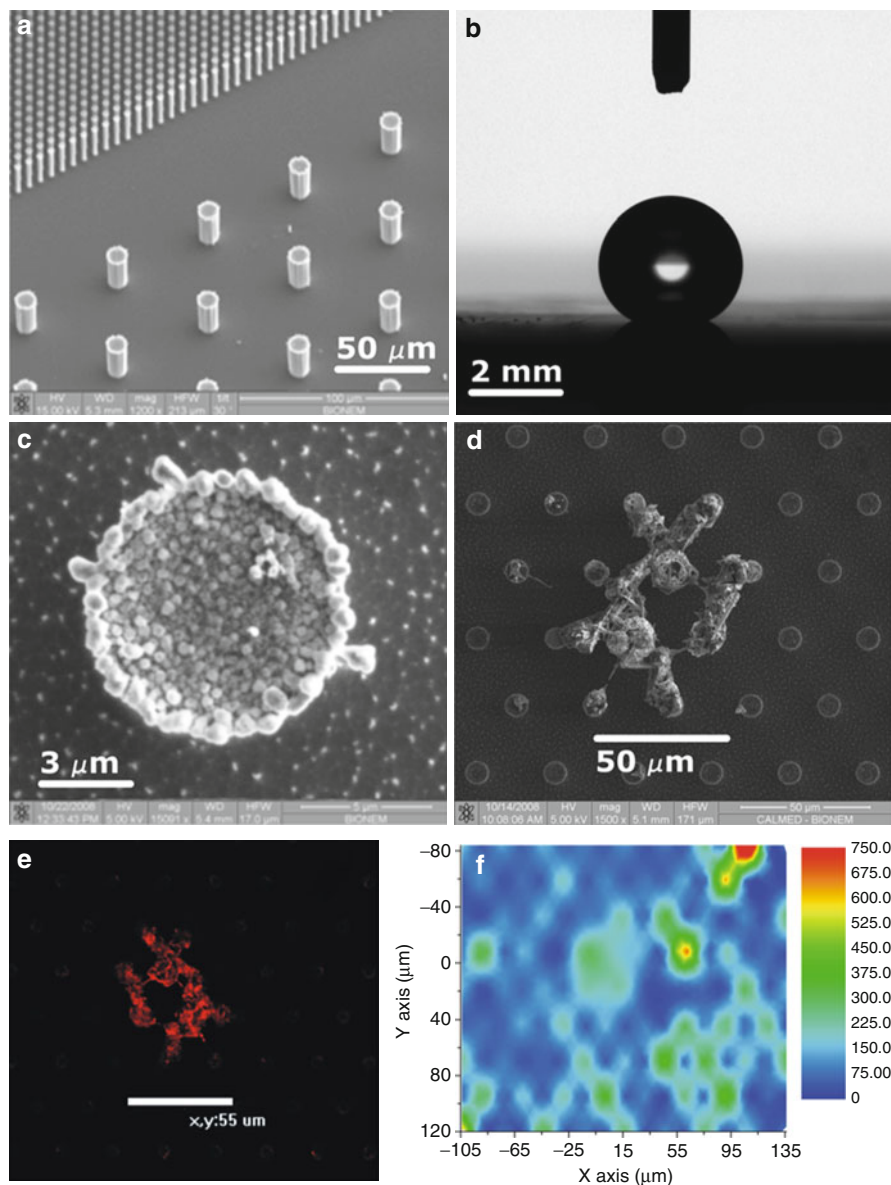
Superhydrophobicity is a phenomenon whereby a drop post upon a surface would preserve its original spherical shape rather than spreading or wetting indefinitely the plane of contact (Fig. 15). In the celebrated model of Cassie [29], a surface would be superhydrophobic on account of the pockets of air that remain trapped between the liquid and the substrate, and the smaller the fraction of solid in contact with the drop the larger the apparent contact angle. SHSs retain unique properties in terms of wettability that can be reviewed as follows: (a) SHSs have superior adhesive properties, in the sense that they exhibit

vanishing friction coefficients; (b) a droplet, deposited upon these surfaces, would accordingly preserve a quasi-spherical shape while evaporates, and the contact area at the interface would thus progressively reduce; (c) SHSs can be artificially reproduced using micro and nanofabrication techniques. Using the properties above, micro textured surfaces may be successfully exploited to concentrate tiny amounts of moieties over micrometric areas, and consequently measure these moieties with unprecedented accuracy [30, 31].

SHSs typically comprise a periodic hexagonal lattice of cylindrical Si micro pillars with a certain diameter and pitch (Fig. 16a). The micro pillars are realized combining optical lithography, electroless growth, and Bosch Reactive Ion Etching (RIE) techniques. Nonconventional biophotonic nanostructures such as arrays of silver nanograin aggregates, conveniently positioned upon the pillars, would complete a hierarchical structure thus permitting the identification of substances in the single-molecule regime. These “two-stages” micro and nanostructures act as superhydrophobic surfaces with an increased contact angle ranging from 155° to about 175° .

A particular kind of nanogeometry based plasmonic device is here reported as example, which is an electroless grown random assembly of silver nanograins

Nanostructures for Photonics, Fig. 16 (a) SEM micrograph of periodic hexagonal lattice of cylindrical Si micro pillars used as SHS. (b) The *spherical* shape of a drop positioned upon the SHS is preserved. (c) Arrays of Ag nanograin aggregates positioned upon the pillars. (d) SEM image of the residual solute of R6G at the end of process of evaporation. The intensity of fluorescence is shown in (e), while Micro-Raman mapping measurements are reported in (f)



[23]. Nevertheless, the method can be extended to a number of different plasmonic nanostructures, as those described in the sections above.

Imagine to deposit a drop of an extremely diluted solution upon a textured, superhydrophobic substrate. The drop would evaporate over time and thus the solution would get more and more concentrated (Fig. 15). At the end of the process, the residual solute would be confined within an exceptionally small region of the plane. With a suitable design, few molecules may be conveniently enforced to confine into the smallest area conceivable, at the limit upon a sole

pillar. Nanogeometry based biophotonic devices, conveniently tiling these surfaces, would probe/detect the moieties with extremely small resolution limits. Here, the beneficial effects of superhydrophobicity and nanogeometry based spectroscopy are combined and conveyed into a unique platform, and from the combination of the two novel properties arise permitting the identification of proteins or analytes with unprecedented accuracy.

In order to clarify the concept, an experiment performed using R6G molecules, is reported. Evaporation processes of drops of D.I. water were followed

over time. Few molecules were conveniently enforced to confine into a small area. A solution was investigated with initial concentration as small as 10^{-18} M. [Figure 16d](#) shows a SEM image of the residual solute of R6G at the end of process of evaporation. The intensity of fluorescence as in [Fig. 16e](#) is correctly proportional to the quantity of substance deposited upon the pillars, and thus the intensity signal of fluorescence follows the solute distribution upon the substrate. Micro-Raman mapping measurements were also performed ([Fig. 16f](#)). The mapping analysis was performed by referring the band centered at $1,650\text{ cm}^{-1}$. The results demonstrate that the method is effective in measuring extremely small diluted solutions.

Cross-References

- ▶ [Biosensors](#)
- ▶ [Electron Beam Lithography \(EBL\)](#)
- ▶ [Light Localization for Nano-Optical Devices](#)
- ▶ [Nanoparticles](#)
- ▶ [Nanophotonic Structures for Biosensing](#)
- ▶ [Optical Properties of Metal Nanoparticles](#)
- ▶ [Surface Plasmon-Polariton-Based Detectors](#)

References

1. Carbone, L., Nobile, C., De Giorgi, M., Sala, F.D., Morello, G., Pompa, P., Hytch, M., Snoeck, E., Fiore, A., Franchini, I.R., Nadasan, M., Silvestre, A.F., Chiodo, L., Kuder, S., Cingolani, R., Krahne, R., Manna, L.: Synthesis and micrometer-scale assembly of colloidal CdSe/Cds nanorods prepared by a seeded growth approach. *Nano Lett.* **7**, 2942 (2007)
2. Efros, A.L., Rosen, M., Kuno, M., Nirmal, M., Norris, D.J., Bawendi, M.: Band-edge exciton in quantum dots of semiconductors with a degenerate valence band: dark and bright exciton states. *Phys. Rev. B-Condens. Mat.* **54**, 4843 (1996)
3. Zavelani-Rossi, M., Lupo, M.G., Krahne, R., Manna, L., Lanzani, G.: Lasing in self-assembled microcavities of Cdse/Cds Core/Shell colloidal quantum rods. *Nanoscale* **2**, 931 (2010)
4. Krahne, R., Zavelani-Rossi, M., Lupo, M.G., Manna, L., Lanzani, G.: Amplified spontaneous emission from core and shell transitions in Cdse/Cds nanorods fabricated by seeded growth. *Appl. Phys. Lett.* **98**, 063105 (2011)
5. Kazes, M., Lewis, D.Y., Ebenstein, Y., Mokari, T., Banin, U.: Lasing from semiconductor quantum rods in a cylindrical microcavity. *Adv. Mater.* **14**, 317 (2002)
6. Martiradonna, L., Carbone, L., De Giorgi, M., Manna, L., Gigli, G., Cingolani, R., De Vittorio, M.: High Q-factor colloidal nanocrystal-based vertical microcavity by hot embossing technology. *Appl. Phys. Lett.* **88**, 181108 (2006)
7. Shoji, S., Proietti Zaccaria, R., Sun, H., Kawata, S.: Multi-step multi-beam laser interference patterning of three-dimensional photonic lattices. *Opt. Express* **14**, 2039 (2006)
8. Knight, J.C., Birks, T.A., Russell, P.St.J., Atkin, D.M.: All-silica single-mode optical fiber with photonic crystal cladding. *Opt. Lett.* **21**, 1547 (1996)
9. Galli, M., Agio, M., Andreani, L.C., Atzeni, L., Bajoni, D., Guizzetti, G., Businaro, L., Di Fabrizio, E., Romanato, F., Passaseo, A.: Optical properties and photonic bands of GaAs photonic crystal waveguides with tilted square lattice. *Eur. Phys. J. B* **27**, 79 (2002)
10. Malvezzi, A.M., Vecchi, G., Patrini, M., Guizzetti, G., Andreani, L.C., Romanato, F., Businaro, L., Di Fabrizio, E., Passaseo, A., De Vittorio, M.: Resonant second-harmonic generation in a GaAs photonic crystal waveguide. *Phys. Rev. B* **68**, 1613061 (2003)
11. De Vittorio, M., Todaro, M.T., Stomeo, T., Cingolani, R., Cojoc, D., Di Fabrizio, E.: Two-dimensional photonic crystal waveguide obtained by e-beam direct writing of SU8-2000 photoresist. *Microelectron. Eng.* **73–74**, 388 (2004)
12. Proietti Zaccaria, R., Verma, P., Kawaguchi, S., Shoji, S., Kawata, S.: Manipulating full photonic band gaps in two dimensional birefringent photonic crystals. *Opt. Express* **16**, 14812 (2008)
13. Zhao, H., Proietti Zaccaria, R., Verma, P., Song, J., Sun, H.: Single-mode operation regime for 12-fold index-guiding quasicrystal optical fibers. *Appl. Phys. B* **100**, 499 (2010)
14. Zhao, H., Proietti Zaccaria, R., Song, J., Kawata, S., Sun, H.: Photonic quasicrystals exhibit zero-transmission regions due to translational arrangement of constituent parts. *Phys. Rev. B* **79**, 115118 (2009)
15. Zhao, H., Proietti Zaccaria, R., Verma, P., Song, J., Sun, H.: Validity of the V parameter for photonic quasi-crystal fibers. *Opt. Lett.* **35**, 071064 (2010)
16. Veselago, V.: The electrodynamics of substances with simultaneously negative values of ϵ and μ . *Sov. Phys. Uspekhi* **10**, 509 (1968)
17. Papanikolaou, N., Luo, Z., Shen, Z.X., De Angelis, F., Di Fabrizio, E., Nikolaenko, A.E., Zheludev, N.I.: Graphene in a photonic metamaterial. *Opt. Express* **18**, 8353 (2010)
18. Nikolaenko, A.E., De Angelis, F., Boden, S.A., Papanikolaou, N., Ashburn, P., Di Fabrizio, E., Zheludev, N.I.: Carbon nanotubes in a photonic metamaterial. *Phys. Rev. Lett.* **104**, 153902 (2010)
19. Sámson, Z.L., MacDonald, K.F., De Angelis, F., Gholipour, B., Knight, K., Huang, C.C., Di Fabrizio, E., Hewak, D.W., Zheludev, N.I.: Metamaterial electro-optic switch of nanoscale thickness. *Appl. Phys. Lett.* **96**, 143105 (2010)
20. Mary, A., Rodrigo, S.G., Garcia-Vidal, F.J., Martin-Moreno, L.: Theory of negative-refractive-index response of double-fishnet structures. *Phys. Rev. Lett.* **101**, 103902 (2008)
21. Pérennès, F., Marmiroli, B., Matteucci, M., Tormen, M., Vaccari, L., Di Fabrizio, E.: Sharp beveled tip hollow microneedle arrays fabricated by LIGA and 3D soft lithography with polyvinyl alcohol. *J. Micromech. Microeng.* **16**, 473 (2006)
22. Tormen, M., Businaro, L., Altissimo, M., Romanato, F., Cabrini, S., Perennes, F., Proietti, R., Sun, H.-B., Kawata, S., Di Fabrizio, E.: 3D patterning by means of nanoimprinting.

X-ray and two-photon lithography. *Microelectron. Eng.* **73–74**, 535 (2004)

23. Das, G., Mecarini, F., Gentile, F., De Angelis, F., Kumar, M.-H.G., Candeloro, P., Liberale, C., Cuda, G., Di Fabrizio, E.: Nano-patterned SERS substrate: application for protein analysis vs. temperature. *Biosen. Bioelect.* **24**, 1693 (2009)
24. Liberale, C., Minzioni, P., Bragheri, F., De Angelis, F., Di Fabrizio, E., Cristiani, I.: Miniaturized all-fibre probe for three dimensional optical trapping and manipulation. *Nat. Photonics* **1**, 723 (2007)
25. De Angelis, F., Patrini, M., Das, G., Maksymov, I., Galli, M., Businaro, L., Andreani, L.C., Di Fabrizio, E.: A hybrid plasmonic – photonic nanodevice for label-free detection of a few molecules. *Nano Lett.* **8**, 2321 (2008)
26. De Angelis, F., Das, G., Candeloro, P., Patrini, M., Galli, M., Bek, A., Lazzarino, M., Maksymov, I., Liberale, C., Andreani, L.C., Di Fabrizio, E.: Nanoscale chemical mapping using three-dimensional adiabatic compression of surface plasmon polaritons. *Nat. Nanotech.* **5**, 67 (2010)
27. Cabrini, S., Carpentiero, A., Kumar, R., Businaro, L., Candeloro, P., Prasciolu, M., Gosparini, A., Andreani, C., De Vittorio, M., Stomeo, T., Di Fabrizio, E.: Focused ion beam lithography for two dimensional array structures for photonic applications. *Microelectron. Eng.* **2005**(78–79), 11 (2005)
28. Stockman, M.I.: Nanofocusing of optical energy in tapered plasmonic waveguides. *Phys. Rev. Lett.* **93**, 137404 (2004)
29. Lafuma, A., Quéré, D.: Superhydrophobic states. *Nat. Mater.* **2**, 457 (2003)
30. Accardo, A., Gentile, F., Mecarini, F., De Angelis, F., Burghammer, M., Di Fabrizio, E., Riekkel, C.: In situ x-ray scattering studies of protein solution droplets drying on micro- and nanopatterned superhydrophobic PMMA surfaces. *Langmuir* **26**, 15057 (2010)
31. De Angelis, F., Gentile, F., Mecarini, F., Das, G., Moretti, M., Candeloro, P., Coluccio, M.L., Cojoc, G., Accardo, A., Liberale, C., Zaccaria, R.P., Perozziello, G., Tirinato, L., Toma, A., Cuda, G., Cingolani, R., Di Fabrizio, E.: Breaking the diffusion limit with super hydrophobic delivery of few molecules to plasmonic nanofocusing structures. *Nat. Photon.* **5**, 682 (2011)

Nanostructures for Surface Functionalization and Surface Properties

Paola Rivolo

Dipartimento di Scienza dei Materiali e Ingegneria
Chimica – Politecnico di Torino, Torino, Italy

Synonyms

Chemical modification; Functionalization; Organosilanes; Patterning; Physical modification; Plasma grafting; Plasma polymerization; Self-assembled monolayers; Surface properties; Thiols

Definition

Modification and functionalization of substrates can lead to the presence at materials surfaces of molecular nanostructures. Their 2D packing and chemical functionalities may impart to surfaces particular properties that can be very different from the pristine ones.

Modification is a very general term concerning every kind of treatment leading to change the physical morphology or surface chemistry of a given material in order to obtain tailored properties with dependence on the material application field.

Functionalization is mainly related to the introduction at a surface of chemical groups with a variable surface density which are requested to subsequently react with other species.

When a process or a group of processes are applied to the surface of an inorganic or organic material, affecting both topography/morphology and surface chemistry at the micro- and/or nanoscale level, the material surface results physically and chemically patterned.

In the following the main modification, functionalization, and patterning surface techniques will be described according to process methods, substrate chemistry, and nanotechnological application fields.

Surface Modification and Functionalization

In general, some materials which have excellent bulk physical and chemical properties often do not possess suitable surface properties required for specific applications.

For this reason, surface-modification techniques, that can transform these materials into valuable finished products, become an important part of surface science and technology.

Processes, developed in the last decade, are able to simply modify surface morphology of materials or give rise to nanostructured coatings aimed to simply affect the adhesion or anti-adhesion properties, mainly hydrophilicity and hydrophobicity, or to promote specific bonding with other species. These features can be a crucial advantage in several applications in the fields of protective, anti-adhesion, adhesion coatings, friction and wear, composites, microelectronics and thin-film technology, molding and lithography, technical fabrics, biomaterials, sensing and in particular biosensing.

Generally, these surface-modification methods can be divided into two classes: physical modifications and chemical modifications [1].

In this entry will be reported the main topics regarding surface chemical modifications, both the surface derivatization or grafting method which depend on the chemistry of the substrate surface and the functional organic coating deposition which can be performed on the surface of very different materials without the need of a specific chemical binding with the surface (no-dependence on chemistry).

Finally, some surface patterning procedures devoted to get functional structures will be described as following process of the previously reported methods or based on other modification techniques including physical processes.

Chemical Derivatization or Grafting

The key advantage of these techniques is that the surface of the materials can be modified or tailored to acquire very distinctive properties through the choice of different grafting precursors, while maintaining the substrate properties.

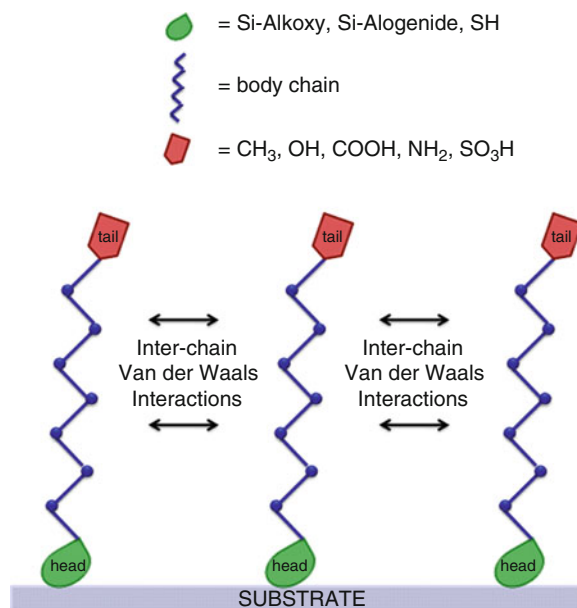
The crucial point is to single out the suitable molecule to be grafted according to the chemistry of the material substrate, that is, the kind of atoms/groups with which the substrate is truncated with.

Before the grafting step, commonly a surface activation is needed in order to create reactive sites on the substrate surface or enhance the amount of the ones naturally already present.

Practically, one can generate reactive groups through UV, high-energy electrons, γ irradiation, plasma treatment, ozone exposure [1] and chemical reactions performed in liquid and vapor phase. The kind of activation treatment heavily depends on materials.

For sake of simplicity, the different types of derivatizing treatments and the related materials activation procedures will be considered according to substrate chemistry.

Generally, among the surface-modification techniques, the formation of a Self-Assembled Monolayer (SAM) has been intensively studied. SAM is a kind of “bottom-up” technique which provides advantages regarding the surface modification. It leads to a structure which is at, or close to, thermodynamic equilibrium, and thus tends to self-healing/defect rejection and leads to a closely packed, well-ordered, and stable configuration on the surface [1]. SAMs offer



Nanostuctures for Surface Functionalization and Surface Properties, Fig. 1 Scheme of a general SAM formed on the surface of a substrate

unique opportunities to increase fundamental understanding of self-organization, structure-property relationships, and interfacial phenomena. The ability to tailor both head and tail groups of the constituent molecules makes SAMs excellent systems for a more fundamental understanding of phenomena (e.g., ordering and growth, wetting, adhesion, lubrication, corrosion, etc.) affected by competing intermolecular, molecule-substrate and molecule-solvent interactions.

SAMs provide the needed design flexibility, both at the individual molecular and at the material levels, and offer a vehicle for investigation of specific interactions at interfaces and of the effect of increasing molecular complexity on the structure and stability of two-dimensional assemblies.

The key factors characterizing the SAM precursor molecules consist in (a) the “head,” (b) the “tail,” and (c) the body chain length of the molecule (Fig. 1). The “head” is one of the terminal group of these normally linear species and has the role to anchor the surface by a favored and fast reaction producing the formation of a usually stable (covalent) bond linking the surface. The choice of this termination is driven by the chemistry of the material substrate.

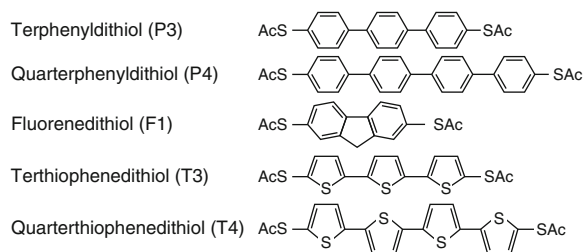
The “tail” of the molecule, that is, the termination opposite to the “head,” is deputed to perform the

desired effect at the SAM surface. So as a rule of thumb, it can be a methyl termination ($-\text{CH}_3$), a perfluoroalkyl termination that generate at the most a hydrophobic or super-hydrophobic and repulsive behavior or a functional group such as $-\text{SH}$, $-\text{OH}$, $-\text{NH}_2$, $-\text{COOH}$ able to carry out a subsequent reaction with other molecule (for example, biomolecules, polymers, suitable linker molecules) or simply express hydrophilic and adhesion properties.

The molecule length (the hydrocarbons chains can range, for instance, from C_3 to C_{24}) will strongly depend on the application for which SAM structures are needed.

Some reviews on the self-assembled monolayers have covered the SAM formation and structure their characterization by electrochemical and scanning tunneling microscopy [1], by PM-IRRAS vibrational spectroscopy techniques and XPS analysis [2], their application to biosensors [1], anti-stiction in MEMS devices and in nano-imprinting lithography techniques [3], and their advantages in providing controlled surface properties and unique reactions [1].

In particular, such monolayers have been generated mostly on metal surfaces, such as Au, Ag, and Cu, using the thiol chemistry [1]. Nuzzo and Allara, reported in 1983, [4] the procedure of bifunctional organic sulfides species self-assembly on the surface of Au (111) freshly evaporated substrates in the liquid phase by simply putting in contact the samples with a diluted (milliMolar range) solution of the reagent, for some hours. Successively many authors observed with other n-alkyl thiolates that in the first few seconds the species are able to bind the gold surface and in more than 1 h, they self-assemble into a compact, well-ordered flat intercrossed molecular monolayer. The $-\text{SH}$ head is able to establish a strong Au-S bond and the chain-to-chain interaction performed by the adjacent grafted molecules is dominated by weak dispersion forces if nonfunctional interacting groups are present in the middle of the aliphatic chain. In this case, H-bond or other stronger van der Waals forces can be exerted. The interchain forces are able to produce, as calculated by Godin [5] an equilibrium separation between two alkyl chains of 0.44 nm. The alkanethiol SAM is formed as a densely packed structure showing $(\sqrt{3} \times \sqrt{3})\text{R}30^\circ$ overlayer on gold surface where the molecule is tilted about 30° from the surface normal with the sulfur atom chemically bonded to gold atoms. This implies that a compressive surface stress has to be



Nanostructures for Surface Functionalization and Surface Properties, Fig. 2 Schematic structures of conjugated molecules studied in [7]

generated during the self-assembly, and this stress will increase with increasing alkyl chain length.

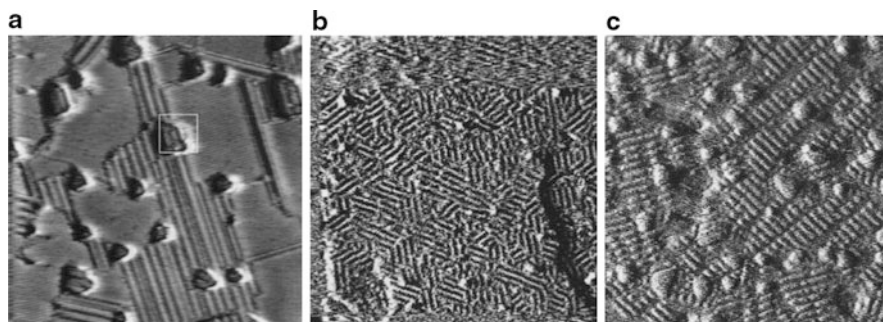
Patel et al. [6] studied the effect of chain length on the reactivity of SAMs toward protein immobilization generating SAMs from short- and long-chain carboxylic acid terminating alkyl-thiols, 11-MercaptoUndecanoic Acid (11-MUA), 3-MercaptoPropanoic Acid (3-MPA) and from the mixture, 1:10, of the both. The accessibility of the carboxylic groups by the amino group of a lysine termination exposed by the protein, *catalase*, to be anchored, decreases in the order Mixed > 11-MUA > 3-MPA. The experiment was performed by previously activating the carboxylic groups by a water soluble 1-ethyl-3-[3-(dimethylamino) propyl]carbodiimide hydrochloride (EDC) and *N*-hydroxysuccinimide (NHS). Successful formation of the NHS ester intermediate is reliant on the accessibility of the terminal carboxylate groups. It was assessed that a steric packing of these acid groups can limit the rate of intermediate formation with full conversion of accessible acid groups only occurring after several repeated reaction cycles. The co-adsorption of different alkyl-thiols provides the system a degree of disorder which is strategic for an efficient protein immobilization.

SAMs of conjugate alkyl-dithiols are one of the more interesting solutions focused by research in molecular electronics which is open to the possibility of substituting inorganic materials with organic molecules in the core regions of devices. Jiang [7] reported STM characterization of the electronic transport properties of SAM of containing double bonds (insaturation) in the body chain such as the aromatic dithiols (Fig. 2).

Co-assembly of conjugated dithiols and alkanemiothiols is suggested as a route to obtain dithiols that are denser packed and more vertical (with respect to what is obtained from pure dithiol routes). This surface

Nanostructures for Surface Functionalization and Surface Properties,

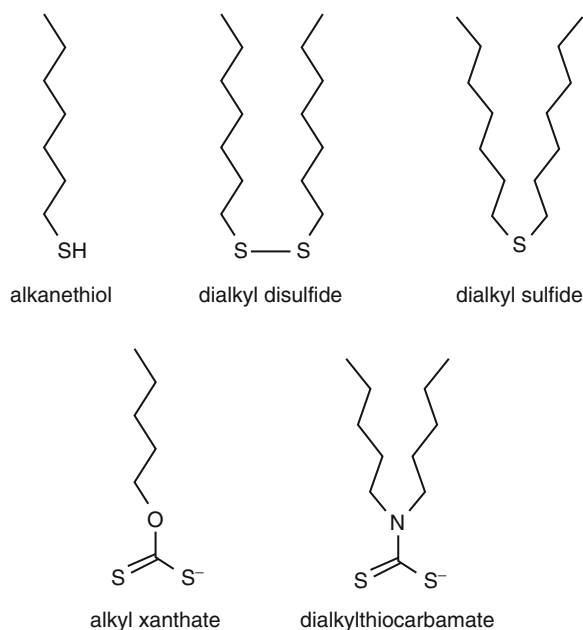
Fig. 3 (a) STM image (95×95 nm) of octanethiol SAMs on Au(111); (b) STM image (70×70 nm) of decanedithiol SAMs on Au(111); (c) STM image (75×75 nm) of the mixture SAMs on Au(111)



condition is suitable to the evaporation on the top formed SAM of an Au layer and to the formation of other Au–S bonds with the free thiol end of the dithiol molecules.

Kobayashi [8] reported the study by Ultra High Vacuum Scanning Tunneling Microscopy (UHV-STM) of alkyl-thiol and alkyl dithiol SAMs on Au surfaces, characterizing the morphology of resulting nanostructures and the strength of molecules interaction with the surface. Figure 3 shows the different appearance of the surface according to the nature mono- or dithiolic of the molecule used to form the SAMs both by a liquid process in ethanolic solution (24 h) and by a vapor functionalization process (from a 10 mM ethanolic solution, for 2 h in a closed vessel). In (a) the 1,8-octanethiol SAM results in nanometric islands where molecules assume a surface-normal arrangements. In (b) the nanometric stripe structures are evident due to the molecules of dithiols (1,6-hexanedithiol, 1,10-decanedithiol) lying with their axes parallel to the surface (the bright spots are assigned as sulfur atoms). The different arrangement between (a) and (b) is probably ascribed to the stronger interaction between both thiol groups at the molecular ends with the gold atoms with respect to the case of alkanedithiol SAMs. The parallel arrangement is more preferred than the normal arrangement in the case of alkanedithiol SAMs. In (c) the SAM is obtained by a mixture of the vapor of the ethanol solution which contained octanethiol and octanedithiol molecules at 1 mM each. The UHV-STM image shows stripe phase regions corresponding to the areas where alkanedithiol SAMs are formed. On the other hand, the island regions may be attributed to alkanethiol SAMs.

Ag (Silver) and other materials such as GaAs (Gallium Arsenide) can be functionalized by alkyl-thiols. Dubowski et al. [9] recently studied the formation mechanism of T-Alkanethiols (where $T = \text{NH}_2$,



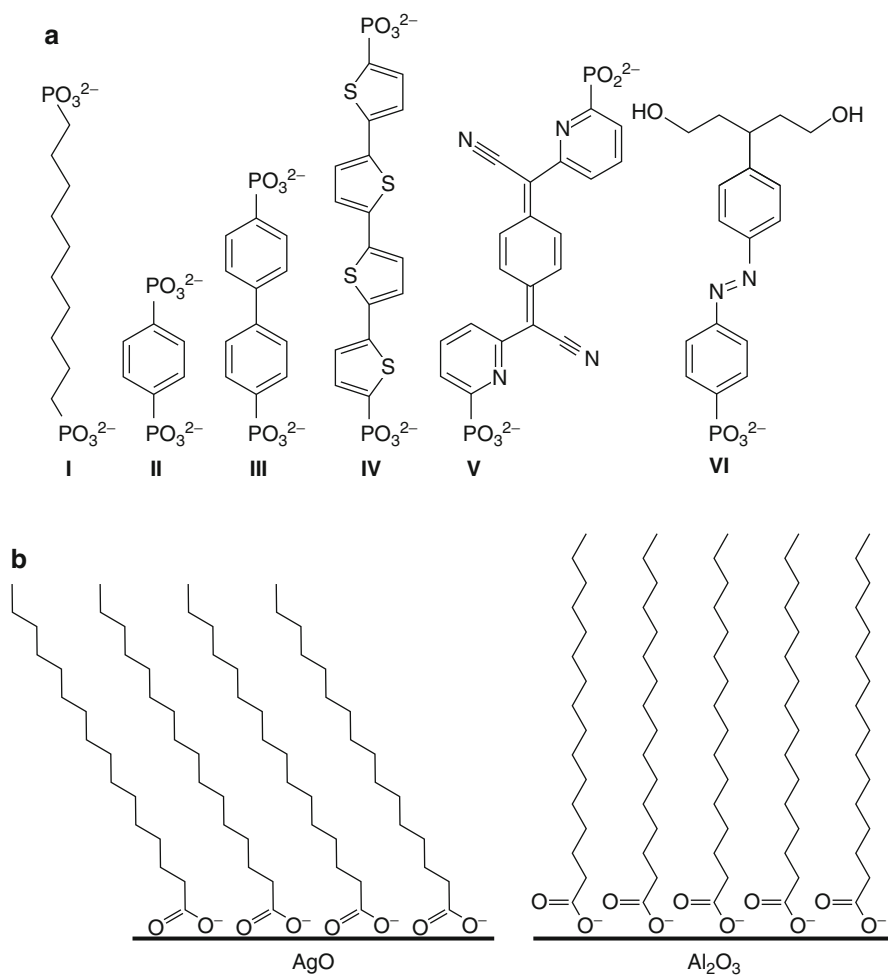
Nanostructures for Surface Functionalization and Surface Properties, Fig. 4 Surface-active organosulfur compounds that form monolayers on metals and semiconductors [10]

COOH , CH_3) SAMs on III-V semiconductors surfaces, in particular GaAs (001). They proposed the formation at the interface of a S–As or S–Ga bond according to which of the two elements is the richer in the exposed surface. They demonstrated the dependence of PhotoLuminescence signal on the terminal T group and the possibility to use these functionalized semiconducting materials in biosensing by PL technique.

In 1996 Ulman [10] reported other organosulfur molecules (Fig. 4) in addition to alkane thiols able to strongly graft and self-assemble on semiconductors surface, such as GaAs and InP, and other metals such as copper, platinum, mercury, iron, nanosize $\gamma\text{-Fe}_2\text{O}_3$ particles.

Nanostructures for Surface Functionalization and Surface Properties,

Fig. 5 (a) Different diphosphonic acids used in self-assembled multilayer preparation. (b) A schematic description of fatty acid monolayers on AgO and on Al₂O₃



In Fig. 5a are displayed the diphosphates molecules used for Zirconium (Zr^{4+}) ions exploiting their ability to form highly insoluble salts with tetravalent transition metal ions. At last, fatty acids, which are long-chain n-alkanoic acids ($\text{C}_n\text{H}_{2n+1}\text{COOH}$), are known, by an acid-base reaction, to form a surface ordered salt between the carboxylate anion and a surface metal cation, for example, with AgO and Al₂O₃ (Fig. 5b).

For inorganic oxides, the typical molecules used to form SAMs monolayers are organosilanes. They are characterized by a variability of possible structures and by the easiness of reaction performed both in the liquid and in the vapor phase.

In particular, Silicon/Silica (Si/SiO_2), glass and quartz surfaces, Aluminum/Alumina ($\text{Al}/\text{Al}_2\text{O}_3$), phyllosilicates such as mica, magnetic materials such as maghemite ($\gamma\text{-Fe}_2\text{O}_3$) nanoparticles, germanium oxide [1], and Zinc oxide (ZnO) [11] exposing

naturally hydroxyl-terminated ($-\text{OH}$) surfaces or zinc selenide [1] and Gallium Nitride (GaN) Aluminum Nitride (AlN) and Silicon Nitride (SiN) [12–14], where the $-\text{OH}$ species can be easily created by a pretreatment, react with alkylchlorosilanes and alkylalkoxysilanes. Stable siloxy linkages [2, 3] are formed between hydroxylated Silica and 3-aminopropyltriethoxysilane (3-APTES). The reaction proceeds through the hydrolysis of silane alkoxy terminations and the condensation with surface hydroxyls and elimination of an alcohol molecule. Further condensation between adjacent molecules already grafted to the surfaces take part stabilizing the SAM through the formation of siloxane bridges ($\text{Si}-\text{O}-\text{Si}$).

Glass and Si/SiO_2 are the most common inorganic materials used in many technological applications as optical and electrical materials and as MEMS (MicroElectroMechanical Systems). One major factor

that limits the widespread use and reliability of MEMS is adhesion. Adhesion is a result of the dominance of surface forces, such as capillary, hydrogen bonding, electrostatic, and van der Waals forces, over body forces at the microscale. When internal restoring forces of microstructures cannot overcome surface adhesive forces, the devices are said to suffer from stiction. Many procedures based on organosilanes SAMs formation are reported to impart to MEMS anti-stiction properties. Comparing the SAM formation from dimethyldichlorosilane ($(\text{CH}_3)_2\text{SiCl}_2$, DDMS) and tridecafluoro-1,1,2,2-tetrahydrooctyltrichlorosilane ($\text{CF}_3(\text{CF}_2)_5(\text{CH}_2)_2\text{SiCl}_3$, FOTS) [3], by a liquid (thermal annealing about 65°C in anhydrous toluene solution) and a vapor procedure (dosage of organosilane vapor pressure on an Oxygen plasma pre-activated surface in a low pressure CVD style reactor), it results that the vapor phase method avoid vertical polymerization of precursor molecule producing a flatter SAM with the same efficiency in anti-stiction properties.

3-APTES has been shown to be an efficient SAM precursor [2, 15] for biosensing of protein tumoral markers based on micro-cantilever devices.

Both the functionalization routes described above are efficient in reducing the mass effect of the functional coating which is detrimental for the device sensitivity and in providing a maximum coverage in the order of 10^{14} $-\text{NH}_2$ groups/ cm^2 [2].

In the field of magneto-chemotherapeutics, where spatial control of conjugated active biomolecules under magnetic field and Magnetic Resonance Imaging (MRI) contrast can be achieved simultaneously are becoming always more relevant functionalized magnetic nanoparticles (NP) able to conjugate specific biomolecules, for instance, Matrix MetalloProteinase (MMP), zinc-containing endopeptidases, which are responsible for the regulated degradation and processing of extracellular matrices (ECM), with physiological functions such as wound healing, tissue remodeling, tumor growth, and metastasis [16]. The silanization of the NP was conducted under anoxic (N_2 purging) suspension consisting of glycerol, aqueous methanol, and acetic acid. Mercaptopropyltrimethoxysilane (MPTMS) was added slowly to the suspension to ensure homogeneous hydrolysis of the silane agent, followed by reflux heating for 16 h. It is demonstrated that by this procedure it is possible not only to get at the NP surface a $-\text{SH}$ terminated SAM able to react

with MMP by the cross-linker Sulfosuccinimidyl-4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (Sulfo-SMCC), but also to reduce the iron oxide phase from $\gamma\text{-Fe}_2\text{O}_3$ to the higher magnetic moment phase of Fe_3O_4 .

Other methods to functionalized Iron Oxide NP are reported by Wu et al. in [17]. For instance, alkyl phosphonates and phosphates are used which bind efficiently to iron oxide particle surfaces, and favor, due to their good biocompatibility, the utilization of encapsulated magnetic NPs in nanomedicines.

The no-oxide surfaces, such as nitrides, can be easily hydroxylated in order to exploit organosilanes chemistry in order to achieve surface functionalization. The well-assessed procedure consists in the substrate immersion in $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2$ (3:1) solution, also called Piranha solution, for 20 min, rinsed with deionized water and dried under Nitrogen flux prior to silanization. The highly oxidizing reagents produce, for GaN (deposited by MetalOrganic Chemical Vapor Deposition – MOCVD), an oxide layer exposing hydroxyl species. For AlN (also obtained by MOCVD) which is already covered by a native oxide layer, the only effect is surface hydroxylation [12].

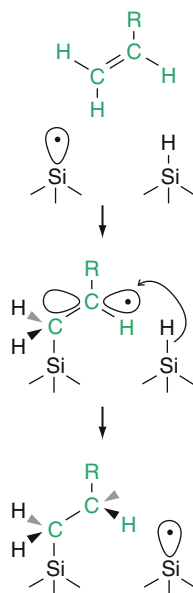
The Piranha solution is used also on crystalline Silicon and on SiO_2 thermally grown from crystalline Silicon. On SiN (deposited by Plasma Enhanced Chemical Vapor Deposition, PECVD), a solution with 10% nitric acid (HNO_3) at 80°C for 20 min [13] or a mixture of CH_3COOH and H_2O_2 for 3 h [14] is used to activate the silanols ($\text{Si}-\text{OH}$); for the both water is used for rinsing and N_2 for drying.

Also polymers can be activated in order to react with organosilane “head” (trialkoxo- or trichloro-terminations) as suggested by Kühn et al. in [18]. The double procedure consists in an Oxygen plasma step in order to produce mixed oxygen containing species OH, C=O, C–O–C, CHO, COOH, C–O–OH, CO_3 , C=C, etc., on a polyolefin polymer such as PolyEthylene (PE) or PolyPropylene (PP) and then a reduction step in diborane (gaseous procedure) or LiAlH_4 solution in order to obtain a monotype homofunctional surface exposing hydroxyls.

Since long time, Silicon surface reconstruction has been studied in order to understand the most suitable Silicon face (100) and (111) to get an organic functionalization avoiding detrimental side oxidation effects. Bent [19] reported some example of Silicon hydrosilylation by alkenes or alkynes, molecules

Nanostructures for Surface Functionalization and Surface Properties,

Fig. 6 Proposed mechanism for surface hydrosilylation. Initial loss of silicon hydride generates a silicon dangling bond. Reaction between the silicon and an alkene molecule leads to an attached alkyl radical, which may abstract a hydrogen atom from a neighboring silicon



bearing C=C terminal groups that react with the Si-H of the Si surface under different activation methods.

A radical initiator, a species used to create free radicals in reactions, starts the reaction by abstracting H from a surface Si-H bond, producing a silicon dangling bond (which is also referred to as a silicon radical). The mechanism is illustrated in Fig. 6 on a (111) silicon surface.

The dangling bond is represented as an orbital with a dot, where the dot represents a single, unpaired electron. The silicon dangling bond then reacts with the alkene molecule, forming a new Si-C bond at the surface and leading to the attachment of the organic molecule, which itself is left with a carbon radical. This surface bound organic radical is thought to abstract a hydrogen atom either from an unreacted alkene molecule or from a neighboring Si-H group on the surface. The net reaction sequence produces a stable, closed-shell organic group bound directly to the Si (111) surface. Hydrosilylation can be achieved also under UV irradiation in the range 200–400 nm, thermal activation at $T > 150^{\circ}\text{C}$ (but a low quality of resulting organic monolayers has been noticed), introducing in the reactant mixture acidic molecules (called Lewis acids) to catalyze the reaction.

The main applications of organic layers on Silicon are related to protection of the underlying substrate from attack by etchants, attribution of chiral properties to the surface, addition of new optical properties to

silicon and recently the substitution of organosilanes in sensing technology when the presence of oxide has to be avoided, also if the organic monolayers lacks in terms of order and cohesion, typical of SAMs.

Silicon/Silica and Diamond surface can be also directly derivatized through polymers grafting. On their own, polymers expose functionalities which are able to give further reactions.

Surface graft polymerization is a chemical modification method. In surface graft polymerization, the modification is achieved by grafting suitable macromolecular chains on the surface of materials through covalent bonding. The key advantage of these techniques is that the surface of the materials can be modified or tailored to acquire very distinctive properties through the choice of different grafting monomers, while maintaining the substrate properties. It also ensures an easy and controllable introduction of graft chains with a high density and exact localization onto the surface. Compared with the physically coated polymer chains, the covalent attachment of the grafted chains onto a material surface avoids their desorption and maintains a long-term chemical stability of the introduced chains.

Chiari et al. [20] used a *ter*-polymer of *N,N*-dimethylacrylamide, *N*-acryloyloxysuccinimide, and 3-(trimethoxysilyl)propyl methacrylate. The copolymer (DMA-NAS-MAPS) graft the SiO₂ surface of oxidized silicon or glass by a simple and robust procedure consisting in the dip coating in an aqueous solution, of substrates of different shapes in a short time span (less than 1 h) and in mild conditions without harmful treatments with harsh washing solutions. The film formed on the surface is thin, on the order of a few nanometers. For fluorescent microarray chips, this is a suitable coating as it does not affect the fluorescence intensification provided by the silicon/silicon oxide substrate. The trimethoxysilyl terminations react with surface silanols, as well as an organosilane, the acryloyloxysuccinimidyl ends give rise to covalent bonds with amino groups of biomolecules and the DMA units form the structural backbone of the polymer.

Recently, Oxidized Ultrananocrystalline diamond (UNCD) has been successfully coated with poly(styrene) (PS) using the self-initiated photografting and photopolymerization (SIPGP) approach with styrene as the monomer, within 16 h, under constant irradiation with UV light ($\lambda_{\text{max}} = 350 \text{ nm}$) [21]. The

obtained coating stable grafted to surface is then easily derivatized by sulphonation, nitration, and amido-alkylation carried out according to well-known liquid chemistry routes and so allowing the incorporation of additional functional molecules, such as fluorescent labels. This will open the way for designing more advanced biosensing schemes, incorporating multifunctional elements and with a higher loading capacity for biomolecules.

Ruckenstein et al. reported the graft polymerization reaction on already SAMs coated surfaces of glass and silicon wafers in order to introduce additional reactive sites. The method involves the aniline substitution at Br-terminated SAM through the aniline amino termination and further the surface oxidative graft polymerization of aniline on the modified glass surface via the covalently immobilized aniline sites. A PolyAniline (PANI) coating is so obtained, performing the reaction in the polymerization solution containing aniline, and due to its well-known conductivity (conductive polymer) this can be a suitable substrate to be used as biosensor and biomaterial.

Another route to further derivatize an organosilane surface for biosensing purposes consists in the grafting, from a 8.5 pH solution, of homofunctional linkers such as Glutaraldehyde ($\text{OHC-CH}_2\text{-CH}_2\text{-CH}_2\text{-CHO}$) which is able through an aldehyde end to react with NH_2 exposed at amino-SAM modified surface and, then, through the second aldehyde termination to immobilize biomolecules exposing amino terminations.

This method has been successfully applied by Ricciardi et al. [15] for functionalization of cantilever based biosensors aimed to the detection of tumoral markers (Angiopoietin).

Surface Modification: No-Dependence on Materials Chemistry

Among the surface-modification techniques, deposition of coating obtained by plasma polymerization presents, respect to surface derivatization discussed in the previous paragraph, the advantage to allow the tailoring of properties of any kind of surface without involving the substrate surface chemistry in the functionalization mechanism.

Macromolecular plasma chemistry [22] has developed in the last two decades in the following directions: plasma-enhanced synthesis (deposition and/or grafting) of thin layer macromolecular structures;

surface functionalization of polymeric materials; and etching of inorganic or polymeric substrate surfaces. Plasma-enhanced synthesis involves the dissociation of starting materials and reorganization of the resulting neutral and charged molecular fragments into macromolecular structures on the surfaces located inside or outside of the plasma zone. Outside the plasma zone, recombination mechanisms usually result in the incorporation of molecular fragments with lower degrees of dissociation into the nascent macromolecular structures. When the starting materials are common monomers the recombination processes are more complex due to the development of simultaneous conventional polymerization reactions along with the fragment-recombination mechanisms initiated by the plasma-created and surface-attached active species (e.g., ions, free radicals). In these cases, complex structures result and structural identification is difficult.

However, an advantage of conventional monomer-based processes is that specific functionalities can be retained in the polymer structures and consequently some characteristics can be more easily predicted and designed into the final product.

Macromolecular thin layers can also be synthesized by the plasma generation of active sites (e.g., ions, free radicals) or reactive functionalities (primary amine groups, etc.) on substrate surfaces, followed by the development of graft-polymerization reactions in the absence of the plasma in the presence of “true monomers.” Remote (the recombination processes are developed outside the plasma zone) and pulsed (plasma sustained according to preselected duty cycles) plasma processes, allow recombination mechanisms in the absence of the plasma state in time or in space, and are characterized by significantly minimized dissociation processes. These approaches do not change dramatically the nature of molecular fragmentation processes (the nature and relative ratios of charged and neutral active species); however, they limit significantly the intensities of plasma-induced fragmentation mechanisms, and allow the development of recombination processes in the partial or total absence of plasma state. As a result, molecular precursors can be physically or through covalent bonding incorporated through surface-located charged and neutral (e.g., free radicals) active species, into the nascent macromolecular layers.

Owing to the high reactivity of plasma species, surface functionalization reactions of even the most

inert polymeric substrates can be conveniently achieved. These mechanisms usually involve non-polymerizing gas plasmas, which generate active molecular fragments that covalently attach during the plasma reactions to the activated substrate surfaces. Recently, it has been demonstrated that even neutral molecules can be connected to active sites with covalent bonds on plasma-exposed substrates under in situ conditions [22]. At high particle energies, etching processes dominate both the deposition and functionalization mechanisms.

Depending on the nature of the plasma gases, inert (sputtering) or reactive etching mechanisms can be developed. It should, however, be noted that these mechanisms can have practical importance; they allow the creation of specific surface morphologies by selective ablation, for instance, of amorphous and crystalline surface regions.

The nature of plasmas, the modalities of transferring electric or electromagnetic field intensities to the reaction systems, the geometry of reactor (shape and volumes of the reaction vessel, geometrical location of electrodes and substrates, etc.) and the selected experimental conditions (pressure, power, flow rates of gases, temperature of the substrates, etc.) crucially influence the gas-phase and surface-related plasma chemistry. Consequently, comparing experimental results from a large variety of plasma systems, even for identical starting components, is extremely difficult.

Cold-plasma-mediated processes involve both gas-phase and surface reaction mechanisms. The gas-phase reaction mechanisms involve the interaction of *ab initio* existing and nascent neutral and charged plasma-created species, including atoms, molecules, free radicals, ions of either polarity, excited species, electrons, and photons. Besides the recombination mechanisms developed on the surfaces which confine the plasma, the active species of the discharge interact and continuously tailor the artificially exposed (reactor walls, various substrates, etc.) and self-generated (e.g., plasma-synthesized macromolecular structures) surface layers. The competition between the recombination-deposition (deposition, grafting, functionalization) processes and “destructive interaction” of plasma species (etching) with the nascent macromolecular structures will control the intensities and the predominance of ablation, surface functionalization and macromolecular-film-formation reactions.

The mechanisms of surface functionalization of polymeric substrates are different from the gas-phase processes. While electrons play the most important role in the plasma state, positive ions also play a significant role in the surface chemistry during the interactions of plasma species with polymers. The resulting valence-ionized polymer chains undergo neutralization reactions leading to sufficiently intense localized internal energy concentrations (electronically excited states), which can induce homolytic bond cleavages. Free-radical and unsaturated-bond development can result in cross-linking of polymeric layers. Free radicals can further induce chemical reactions involving plasma species in situ and affect reactions controlled by specific chemical environments (gas-phase or condensed-phase compounds, including, oxygen, monomer molecules, etc.) in the absence of plasma.

Qualitative and quantitative evaluation of charged and neutral species of the gas phase and identification of charged and neutral surface functionalities on substrate surfaces is essential for understanding the plasma-induced reaction mechanisms and physical processes. Adaptation of in situ diagnostic techniques (including real time measurements) for relatively high and low pressure (100–1,000 mTorr) plasma environments is a research area where many questions still remain unanswered. Continued research will promote the development of novel and/or improved processes for tailoring materials surface characteristics and the development of advanced laboratory plasma installations, required for scaling up plasma processes to industrial levels.

Recently, surface treatments and modification based on vapor phase plasma-assisted techniques have been widely applied to several biomedical fields. These processes are able to provide specific monotype chemical functionalities, exposed at materials surfaces. Respect to traditional liquid phase procedures, advantages such as the high process control and the use of small quantities of reagents, are crucial for the enhancement of efficiency of the interactions that materials must perform according to the application.

In particular, procedures for in situ plasma polymerization, starting from vapor released by liquid monomers, are presently used to synthesize innovative polymeric thin films applicable as substrates for cell culture [23], as adhesion promoter in prosthetic implants and for molecular recognition in sensors and biosensors.

The low pressure plasma polymerization, starting from the same reagent and only varying the process parameters (modulation of the plasma discharge, monomer or mixture of monomers/process gas, reactant vapor partial pressure, etc.) leads to the deposition of polymeric materials with different features in terms of density of functional groups, wettability (hydrophilicity/hydrophobicity), and chemical stability [24].

The most studied polymers in biomedical and sensing fields, display as functional groups mainly carboxyls ($-\text{COOH}$) and amines ($-\text{NH}_2$), hydroxyls ($-\text{OH}$), and aldehydes ($-\text{CHO}$). Fewer reports exist on hydroxy and aldehyde surfaces prepared by plasma methods. Hydroxy surfaces can be prepared by water plasma treatment or the plasma polymerization of alkyl alcohol vapors. Water plasma treatment on many polymer substrates suffers from aging, with surface adaptation leading to the movement of surface-modification effects into the polymer. Both hydroxy and aldehyde surfaces have been used for the covalent immobilization of biologically active molecules. Aging effects are less well documented than for amine surfaces.

Aminated surfaces have been fabricated using various plasma vapors or mixtures (allylamines, ethylenediamine, diaminopropane, *n*-heptylamine [24]) and have found wide use for biointerface applications. However, in many cases the amine surfaces have a rather limited shelf life, with post-plasma oxidation reactions and surface adaptation leading to the disappearance of amine groups from the surface. Aging is a widespread phenomenon that often has not been recognized, particularly in some of the earlier studies on the use of plasma-fabricated surfaces for biointerfacial applications, and can markedly alter the surface chemistry [23].

Plasma-fabricated surfaces that contain carboxyl groups have also been well documented and are obtained, for example, by acrylic acid [23] and vinyl acetic acid or applying CO_2 plasma on olefin polymers [24].

In order to better tune the distribution and density of functional groups, are also widely known processes of plasma copolymerization, where the monomer containing functional groups of interest is mixed with a precursor which does not contain specific functionalities (e.g., styrene or octadiene [25]) and which has the function to dilute the reactive groups at the

surface. For example, carboxylate copolymer surfaces have shown excellent ability to support the colonization of some human cell lines of clinical interest. Immobilization of proteins onto plasma-carboxylated surfaces is also well established. Generally, these surface chemistries show good ability to support cell colonization, though the effectiveness seems to depend on the process vapor and the plasma conditions.

An example in which process parameters affect the final properties of the plasma-polymer film is reported by Detomaso et al. [23]. Starting from Acrylic Acid vapor, processes obtained by Continuous Wave (CW) plasma discharge and by Modulated Wave (MW), that is pulsed, plasma discharge are compared respect to the efficiency of the final polymer coating toward cells growth. The $-\text{COOH}$ terminated polymer obtained by MW, with a Duty Cycle ($\text{D.C.} = (T_{\text{on}} / T_{\text{on}} + T_{\text{off}})$) of 5, where $T_{\text{on}} = 5$ ms and $T_{\text{off}} = 95$ ms, is less useful to murine fibroblast growth than the CW plasma film, because more soluble in water due to the higher density of $-\text{COOH}$ groups and the higher retention of the monomer (acrylic acid) structure (low cross-linking and backbone chain propagation).

Materials Surface Modification: Morphology Effects

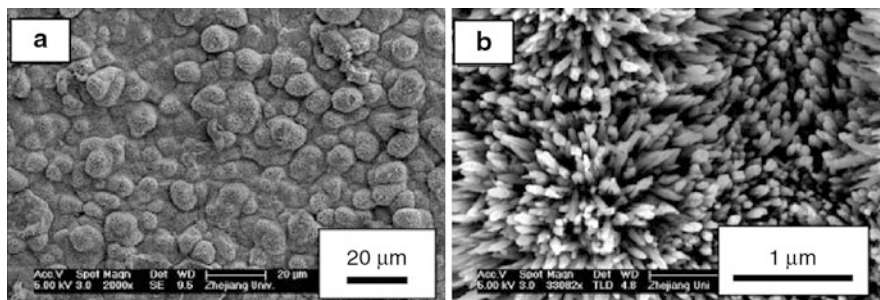
Fabrication of micro- and nano-patterned surfaces is a fundamental step for the development of engineered materials.

For biomaterials, it is well known that a series of interactions occur between their surface and the biological environment after they have been implanted into the human body. Therefore, the biomaterials surface plays an extremely important role in the response of artificial medical devices to the biological environment.

The efficacy of artificial implants is determined mainly by their surface characteristics such as surface morphology, microstructure, composition, and properties. Liu et al. [26] reported many surface-modification methods mainly based on inorganic nano-coatings deposition and functionalization at the nanoscale of obtained structures. Typical coating and film deposition techniques are plasma spraying, plasma immersion ion implantation and deposition (PIII&D), solgel, chemical vapor deposition (CVD), physical vapor deposition (PVD), cold spraying, self-assembly, and so on, whereas in situ surface-modification techniques

Nanostructures for Surface Functionalization and Surface Properties,

Fig. 7 SEM images of a titania nanorod formed on titanium substrate: (a) low magnification, (b) high magnification

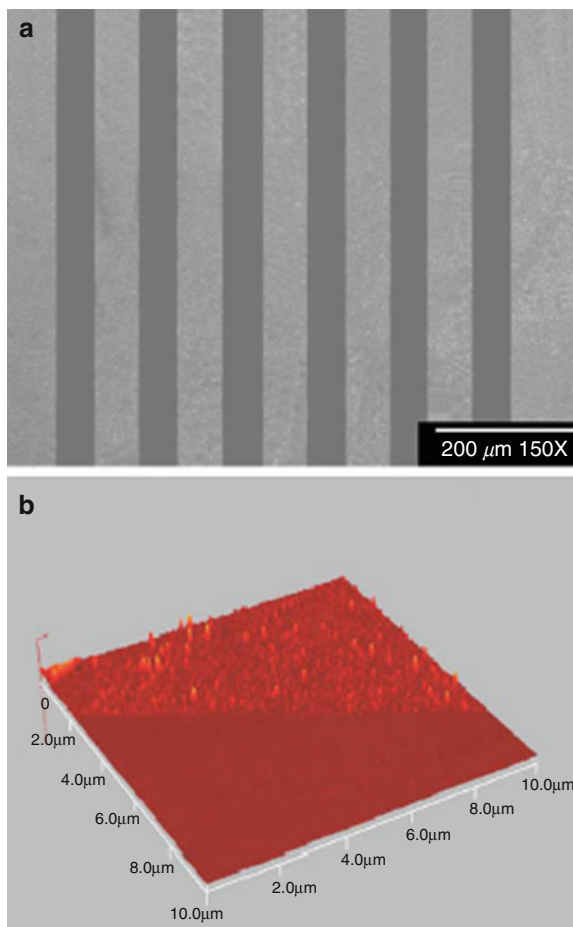


include laser etching, shot blasting, acid and alkali treatments, anodic oxidation, micro-arc oxidation, ion implantation, etc.

In bioceramics field, for example, a critical grain size for osteoblast adhesion between 49 and 67 nm for alumina and 32 and 56 nm for titania, provides evidence about the ability of nanophase alumina and titania to simulate the desirable materials characteristics of bones in a physiological environment, particularly those associated with the enhancement of protein interactions such as adsorption, configuration, and bioactivity as well as subsequent osteoblast adhesion.

Metal surfaces with low micrometer to nanophase topography enhance adhesion of osteoblasts. A surface with various nanostructures can be obtained by chemically treating titanium and its alloys. Chemical treatments are suitable for the creation of micrometer-scale and nanometer-scale textures on large-area surfaces and multifaceted devices with complex 3D shape due to the non-line-of-sight nature. Titania nanorods can be produced at low temperature on a large scale via direct oxidation of titanium with hydrogen peroxide. The ordered nanorod layer with a thickness of approximately 1 μm is deposited on a 2 mm thick condensed anatase layer as illustrated by Fig. 7.

Concerning chemical patterning, microstructured surfaces containing opposite bio-properties such as cell and/or protein adhesive features in a non-bio-adhesive matrix provide valuable tools in the field of biotechnology and numerous methods have been successfully developed for the production of such surfaces. Besides, the development of chemical nanopatterning technology is great of interest due to its ability to provide powerful tools for studying fundamental aspects of cell adhesion at the level of single protein/receptor molecules or for the implementation



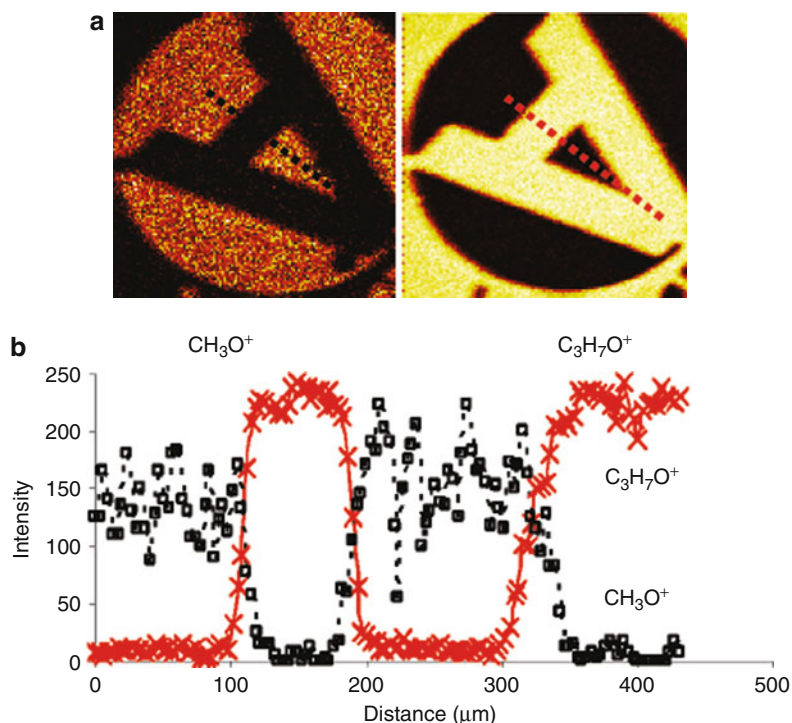
Nanostructures for Surface Functionalization and Surface Properties, Fig. 8 (a) SEM image of the patterned PANI film on a silicon substrate; and (b) selected edge area of (a) under AFM

of new generations of miniaturized biochips requiring smaller reagent quantities with lower detection limits [27].

In general [28], the chemical patterning techniques commonly used in bioengineering can be divided into

Nanostructures for Surface Functionalization and Surface Properties,

Fig. 9 Positive single ion ToF-SSIMS images of ppMA patterned photolithographically on Si wafer precoated with ppTTg ($500 \times 500 \mu\text{m}^2$). Unique secondary molecular fragments are presented in inset (a). The pixel intensity plot for secondary ions $\text{C}_3\text{H}_7\text{O}^+$ (ppTTg) and CH_3O^+ (ppMA) shown in inset (b). The image resolution was calculated to be $\sim 10 \mu\text{m}$, based on 80:20 definition



four categories: (a) lithographic patterning (photo-assisted and electron-beam assisted), (b) scanning probe microscopies (SPM); (c) transfer via polymeric stamps (soft lithography), and (d) physical patterning. The challenge with plasma polymerization lies in developing patterning techniques by which, reproducible spatial resolution and distribution of chemistries is achieved and the specific functional properties of the thin films are retained.

The simplest patterning approach reported by literature is the one based on the use of physical mask (solid materials deposited on defined areas of the substrate), e.g., polymeric masks and TEM grids.

Sometimes, the substrates are previously coated with another plasma polymer obtained starting from a precursor which does not contain specific functionalities (e.g., styrene or octadiene [25]) with the role of non-bio-adhesive matrix.

Ruckenstein et al. [1] reported a method for graft polymerization patterning of polyaniline on Si surface previously coated by SAM monolayers from phenyltrichlorosilanes and further subjected to a laser photolithographic process. The subsequent treatment with trifluoromethanesulfonic acid (triflic acid; HOTf) results in the generation of reactive silyl cation sites

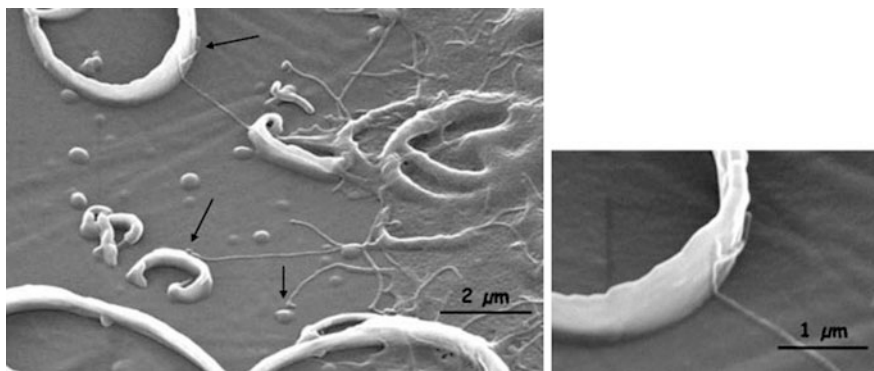
on the backbones, which can react with suitable monomers to form functionalized polysilanes.

The aniline monomer undergoes to nucleophilic substitution reaction followed by the surface oxidative graft polymerization of aniline on the covalently immobilized aniline sites. The grafted PANI thus formed provided a negative of the lithographic mask. Figure 8a presents the scanning electron microscopy (SEM) image of the patterned PANI after the graft polymerization. The white strips represent the grafted PANI areas, whereas the black strips, the nongrafted ones. A relatively clear boundary between the patterned and nonpatterned areas could be observed in the SEM image. A closer look at the boundaries between the two can be achieved under a higher magnification by AFM (Fig. 8b), which demonstrated that this method provides a high line edge acuity of the patterned PANI.

Mishra et al. [28] studied, by using a TEM grid placed onto an already plasma polymer coated silicon wafer, a wide range of standard size patterns in order to vary the resultant pattern features from approximately 5 to $200 \mu\text{m}$. After the plasma polymerization process starting from, for example, glycidyl methacrylate (ppGMA), maleic anhydride (ppMA), acrylic acid

Nanostructures for Surface Functionalization and Surface Properties,

Fig. 10 SEM images of MG 63 osteoblasts on hydrophobic structured samples showing in detail how filopodia could interact with nanostructures on the surface



(ppAAc), allylamine (ppAAm), tetraglyme (ppTG), and grid removal, AFM characterization highlighted the presence, at the physical boundaries, of a diffused interface rather than a sharp step (Fig. 9) and Time-of-Flight Secondary Ion Mass Spectrometry (ToF-SIMS) analysis showed that there was diffusion of the plasma deposit below the physical mask during the second polymer layer deposition.

On the other side, ion beam etching and plasma etching through masks as well as electron beam lithography (EBL) [27], have been shown to be able to produce chemically resolved microscale and in some cases nanoscale plasma polymer patterns.

Bretagnol et al. [27] showed that chemically active features of polyacrylic acid (pAA), with lateral size down to 300 nm, exposing $-\text{COOH}$ groups, obtained by EBL and the use of a photoresist mask, were able to graft a model protein (bovine serum albumin), even if a wet chemical step in acetone was performed.

Photolithographic patterning for plasma polymers is a particularly attractive approach for the fabrication of complex chemical arrays and the opportunity presented by plasma polymerization to produce thin films from a very wide range of different monomers with several properties such as surface charge, hydrophilicity, biospecificity, and adhesion properties alternated to the opposite such as hydrophobicity and, generally, non-fouling characteristics. Recently, it has been demonstrated that specific properties including protein resistance are maintained after performing the photolithographic patterning technique, with both spatial and chemical resolution down to 1 μm . [27].

Finally, an exotic method for surface nanopatterning, based on plasma polymerization has been reported by Gristina et al. [29]. Teflon ribbon-like nanostructures (Fig. 10) have been obtained by C_2F_4

MW (low D.C.) plasma polymerization and it has been demonstrated that these structures promote fibroblast cell adhesion through proteins interaction of cell filopodia with the nanostructures surface notwithstanding the well-known teflon super-hydrophobicity.

Cross-References

- ▶ [Applications of Nanofluidics](#)
- ▶ [Bioadhesion](#)
- ▶ [Biosensors](#)
- ▶ [BioPatterning](#)
- ▶ [Charge Transport in Self-assembled Monolayers](#)
- ▶ [Microfabricated Probe Technology](#)
- ▶ [Nanophotonic Structures for Biosensing](#)
- ▶ [Nanostructured Functionalized Surfaces](#)
- ▶ [Self-assembled Monolayers](#)

References

1. Ruckenstein, E., Li, Z.F.: Surface modification and functionalization through the self-assembled monolayer and graft polymerization. *Adv. Colloid Interface Sci.* **113**, 43–63 (2005)
2. Fiorilli, S., Rivolo, P., Descrovi, E., Ricciardi, C., Pasquardini, L., Lunelli, L., et al.: Vapor-phase self-assembled monolayers of aminosilane on plasma-activated silicon substrates. *J. Colloid Interface Sci.* **321**, 235–241 (2008)
3. Ashurst, W.A., Carraro, C., Maboudian, R., Frey, W.: Wafer level anti-stiction coatings for MEMS. *Sens Actuator A-Phys.* **104**, 213–21 (2003)
4. Nuzzo, R.G., Allara, D.L.: Adsorption of bifunctional organic disulfides on gold surfaces. *J. Am Chem Soc.* **105**, 4481–3 (1983)
5. Godin, M.: Surface stress, kinetics and structure of alkanethiol self-assembled monolayers. PhD Dissertation, McGill University, Canada (2004)

6. Patel, N., Davies, M.C., Heaton, R.J., Roberts, C.J., Tendler S.J.B., Williams, P.M.: A scanning probe microscopy study of the physisorption and chemisorption of protein molecules onto carboxylate terminated self-assembled monolayers. *Appl Phys A*. **66**, S569–74 (1998)
7. Jiang, W., Zhitenev, N., Bao, Z., Meng, H., Abusch-Magder, D., Tennant, D., et al.: Structure and bonding issues at the interface between gold and self-assembled conjugated dithiol monolayers. *Langmuir*. **21**, 8751–7 (2005)
8. Kobayashi, K., Horiuchi, T., Yamada, H., Matsushige, K.: STM studies on nanoscopic structures and electric characteristics of alkanethiol and alkanedithiol self-assembled monolayers. *Thin Solid Films*. **331**, 210–5 (1998)
9. Dubowski, J.J., Voznyy, O., Marshall, G.M.: Molecular self-assembly and passivation of GaAs (0 0 1) with alkanethiol monolayers: a view towards bio-functionalization. *Appl Surf Sci*. **256**, 5714–21 (2010)
10. Ulman, A.: Formation and structure of self-assembled monolayers. *Chem. Rev.* **96**, 1533–54 (1996)
11. Selegård, L., Khranovskyy, V., Söderlind, F., Vahlberg, C., Åhrén, M., Käll, P.-O., Yakimova, R., Uvdal, K.: Biotinylation of ZnO nanoparticles and thin films: a two-step surface functionalization study. *ACS Appl. Mater. Interfaces* **2**, 2128–2135 (2010)
12. Baur, B., Steinhoff, G., Hernando, J., Purrucker, O., Tanaka, M., Nickel, B., Stutzmann, M., Eickhoff, M.: Chemical functionalization of GaN and AlN surfaces. *Appl. Phys. Lett.* **87**, 263901-1–263901-3 (2005)
13. Tlili, A., Jarboui, M.A., Abdelghani, A., Fathallah, D.M., Maaref, M.A.: A novel silicon nitride biosensor for specific antibody–antigen interaction. *Nat Sci Eng C-Biomimetic Supramol Syst.* **25**, 490–5 (2005)
14. Lee, S.-H., Lee, C.-S., Shin, D.-S., Kim, B.-G., Lee, Y.-S., Kim, Y.-K.: Micro protein patterning using a lift-off process with fluorocarbon thin film. *Sens Actuator B-Chem.* **99**, 623–32 (2004)
15. Ricciardi, C., Fiorilli, S., Bianco, S., Canavese, G., Castagna, R., Ferrante, I., et al.: Development of microcantilever-based biosensor array to detect Angiopoietin-1, a marker of tumor angiogenesis. *Biosens Bioelectron.* **25**, 1193–8 (2010)
16. Li, D., Teoh, W.Y., Djunaedi, D., Gooding, J.J., Selomulya, C., Amal, R.: Facile functionalization and phase reduction route of magnetic iron oxide nanoparticles for conjugation of matrix metalloproteinase. *Adv. Eng. Mater.* **12**, B210–4 (2010)
17. Wu, W., He, Q., Jiang, C.: Magnetic iron oxide nanoparticles: synthesis and surface functionalization strategies. *Nanoscale Res. Lett.* **3**, 397–415 (2008)
18. Kühn, G., Weidner, S., Decker, R., Ghode, A., Friedrich, J.: Selective surface functionalization of polyolefins by plasma treatment followed by chemical reduction. *Surf Coat Technol.* 116–119:796–801 (1999)
19. Bent, S.F.: Organic functionalization of group IV semiconductor surfaces: principles, examples, applications, and prospects. *Surf. Sci.* **500**, 879–903 (2002)
20. Cretich, M., Di Carlo, G., Longhi, R., Gotti, C., Spinella, N., Coffa, S., et al.: High sensitivity protein assays on microarray silicon slides. *Anal. Chem.* **81**, 5197–203 (2009)
21. Steenackers, M., Lud, S.Q., Niedermeier, M., Bruno, P., Gruen, D.M., Feulner, P., et al.: Structured polymer grafts on diamond. *J. Am Chem Soc.* **129**, 15655–61 (2007)
22. Denes, F.S., Manolache, S.: Macromolecular plasma-chemistry: an emerging field of polymer science. *Prog Polym Sci.* **29**, 815–85 (2004)
23. Detomaso, L., Gristina, R., D’Agostino, R., Senesi, G.S., Favia, P.: Plasma deposited acrylic acid coatings: surface characterization and attachment of 3 T3 murine fibroblast cell lines. *Surf Coat Technol.* **200**, 1022–5 (2005)
24. Siow, K.S., Britcher, L., Kumar, S., Griesser, H.J.: Plasma methods for the generation of chemically reactive surfaces for biomolecule immobilization and cell colonization – a review. *Plasma Process Polym.* **3**, 392–418 (2006)
25. Swaraj, S., Oran, U., Friedrich, J.F., Lippitz, A., Unger, W.E.S.: Surface chemical analysis of plasma-deposited copolymer films prepared from feed gas mixtures of ethylene or styrene with allyl alcohol. *Plasma Process Polym.* **4**, 376–89 (2007)
26. Liu, X., Chu, P.K., Ding, C.: Surface nano-functionalization of biomaterials. *Mater Sci Eng R-Rep.* **70**, 275–302 (2010)
27. Brétagnol, F., Ceriotti, L., Valsesia, A., Sasaki, T., Ceccone, G., Gilliland, D., et al.: Fabrication of functional nanopatterned surfaces by a combination of plasma processes and electron-beam lithography. *Nanotechnology.* **18**, 135303–7 (2007)
28. Mishra, G., Easton, C.D., McArthur, S.L.: Physical vs photolithographic patterning of plasma polymers: an investigation by ToF-SSIMS and multivariate analysis. *Langmuir.* **26**, 3720–30 (2010)
29. Gristina, R., D’Aloia, E., Senesi, G.S., Milella, A., Nardulli, M., Sardella, E., et al.: Increasing cell adhesion on plasma deposited fluorocarbon coatings by changing the surface topography. *J. Biomed Mater Res Part B: Appl Biomater.* **88B**, 139–49 (2009)

Nanotechnology

Bharat Bhushan

Nanoprobe Laboratory for Bio- & Nanotechnology and Biomimetics, The Ohio State University, Columbus, OH, USA

Synonyms

[BioMEMS/NEMS](#); [MEMS](#); [Micro/Nanodevices](#); [Micro/Nanomachines](#); [Micro/Nanosystems](#); [MOEMS](#); [NOEMS](#); [NEMS](#); [RF-MEMS/NEMS](#)

Definition

Nanotechnology literally means any technology done on a nanoscale that has applications in the real world. Nanotechnology encompasses the production and

application of physical, chemical, and biological systems at scales ranging from individual atoms or molecules to submicron dimensions, as well as the integration of the resulting nanostructures into larger systems.

Overview

Nanotechnology is likely to have a profound impact on world economy and society in the early twenty-first century, comparable to that of semiconductor technology, information technology, or cellular and molecular biology. Science and technology research in nanotechnology promises breakthroughs in areas such as materials and manufacturing, nanoelectronics, medicine and healthcare, energy, biotechnology, information technology, and national security. It is widely felt that nanotechnology will be the next industrial revolution.

Nanometer-scale features are mainly built up from their elemental constituents. Examples include chemical synthesis, the spontaneous self-assembly of molecular clusters (molecular self-assembly) from simple reagents in solution, biological molecules (e.g., DNA) used as building blocks for the production of three-dimensional nanostructures, or quantum dots (nanocrystals) of an arbitrary diameter (about 10–105 atoms). The definition of a nanoparticle is an aggregate of atoms bonded together with a radius between 1 and 100 nm. It typically consists of 10–105 atoms. A variety of vacuum deposition and nonequilibrium plasma chemistry techniques are used to produce layered nanocomposites and nanotubes. Atomically controlled structures are produced using molecular beam epitaxy and organo-metallic vapor phase epitaxy. Micro- and nanosystem components are fabricated using top-down lithographic and nonlithographic fabrication techniques and range in size from micro- to nanometers. Continued improvements in lithography for use in the production of nanocomponents have resulted in line widths as small as 10 nm in experimental prototypes. The nanotechnology field, in addition to the fabrication of nanosystems, provides impetus to the development of experimental and computational tools.

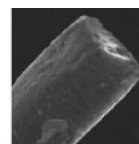
The discovery of novel materials, processes, and phenomena at the nanoscale and the development of new experimental and theoretical techniques for research provide fresh opportunities for the

development of innovative nanosystems and nanostructured materials. The properties of materials at the nanoscale can be very different from those at a larger scale. When the dimension of a material is reduced from a large size, the properties remain the same at first; then small changes occur, until finally when the size drops below 100 nm, and dramatic changes in properties can occur. If only one length of a three-dimensional nanostructure is of the nanodimension, the structure is referred to as a quantum well; if two sides are of nanometer length, the structure is referred to as a quantum wire. A quantum dot has all three dimensions in the nanorange. The word quantum is associated with these three types of nanostructures because the changes in properties arise from the quantum-mechanical nature of physics in the domain of the ultra-small. Materials can be nanostructured for new properties and novel performance. This field is opening new venues in science and technology.

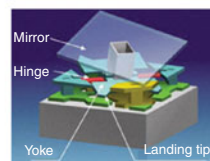
Micro- and nanosystems include Micro/NanoElectroMechanical Systems. MEMS refers to microscopic devices that have a characteristic length of less than 1 mm but more than 100 nm and combine electrical and mechanical components. NEMS refers to nanoscopic devices that have a characteristic length of less than 100 nm and combine electrical and mechanical components. In mesoscale devices, if the functional components are on the micro- or nanoscale, they may be referred to as MEMS or NEMS, respectively. These are referred to as an intelligent miniaturized system comprising of sensing, processing, and/or actuating functions and combine electrical and mechanical components. The acronym MEMS originated in the USA. The term commonly used in Europe is microsystem technology (MST), and in Japan it is micromachines. Another term generally used is micro/nanodevices. MEMS/NEMS terms are also now used in a broad sense and include electrical, mechanical, fluidic, optical, and/or biological function. MEMS/NEMS for optical applications are referred to as micro/nanooptoelectromechanical systems (MOEMS/NOEMS). MEMS/NEMS for electronic applications are referred to as radio-frequency-MEMS/NEMS or RF-MEMS/RF-NEMS. MEMS/NEMS for biological applications are referred to as BioMEMS/BioNEMS.

To put the dimensions of MEMS/NEMS and BioNEMS in perspective, see [Fig. 1](#) and [Table 1](#) [4]. Individual atoms are typically a fraction of

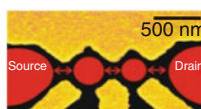
MEMS - characteristic length less than 1 mm,
larger than 100 nm
NEMS - less than 100 nm



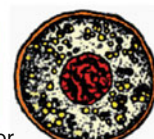
Human hair
50-100 μm



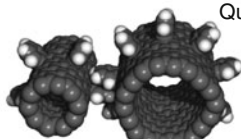
DMD 12 μm



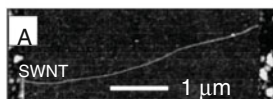
Quantum-dots transistor
300 nm



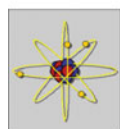
Red blood cell
8 μm



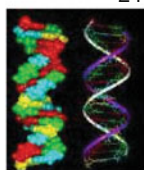
Molecular gear
10 nm-100 nm



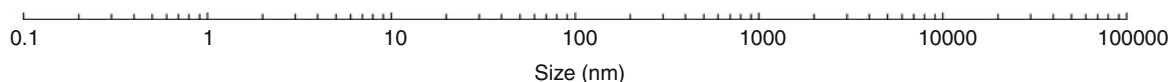
SWNT chemical sensor
2 nm



C atom
0.16 nm



DNA
2.5 nm



Nanotechnology, Fig. 1 Dimensions of MEMS/NEMS and BioNEMS in perspective. Examples shown are of a single-walled carbon nanotube (SWNT) chemical sensor [1], of molecular dynamic simulations of carbon-nanotube-based gears [2],

quantum-dot transistor obtained from van der Wiel et al. [3], and DMD obtained from www.dlp.com. For comparison, dimensions and weights of various biological objects found in nature are also presented

a nanometer in diameter, DNA molecules are about 2.5 nm wide, biological cells are in the range of thousands of nm in diameter, and human hair is about 75 μm in diameter. Smaller length of BioNEMS shown in the figure is about 2 nm, NEMS ranges in size from 10 to 300 nm, and the size of MEMS is 12,000 nm. The mass of a micromachined silicon structure can be as low as 1 nN, and NEMS can be built with mass as low as 10–20 nN with cross-sections of about 10 nm. In comparison, the mass of a drop of water is about 10 μN , and the mass of an eyelash is about 100 nN.

MEMS/NEMS and BioMEMS/BioNEMS are expected to have a major impact on people’s lives, comparable to that of semiconductor technology, information technology, or cellular and molecular biology [5, 6]. MEMS/NEMS and BioMEMS/BioNEMS are used in electromechanical, electronics, information/communication, chemical, and biological applications. The MEMS industry in 2004 was worth about \$4.5 billion with a projected annual growth rate of 17%, Fig. 2 [7]. The NEMS industry was worth about \$10 billion in 2004, mostly in nanomaterials,

Nanotechnology,**Table 1** Characteristic dimensions and weights in perspective

(a) Characteristic dimensions in perspective	
NEMS characteristic length	< 100 nm
MEMS characteristic length	< 1 mm and > 100 nm
SWCNT Chemical sensor	~2 nm
Molecular gear	~10 nm
Quantum-dots transistor	300 nm
Digital micromirror	12,000 nm
Individual atoms	typically fraction of a nm in diameter
DNA molecules	~2.5 nm wide
Biological cells	In the range of thousands of nm in diameter
Human hair	~75,000 nm in diameter
(b) Weight in perspective	
NEMS built with cross-sections of about 10 nm	As low as 10–20 N
Micromachines silicon structure	As low as 1 nN
Eyelash	~100 nN
Water droplet	~10 μ N

Fig. 2 [8]. Growth of Si-based MEMS/NEMS may slow down, and nonsilicon MEMS may pick up during the next decade. It is expected to expand in this decade, in nanomaterials and biomedical applications as well as in nanoelectronics or molecular electronics. For example, miniaturized diagnostics could be implanted for the early diagnosis of illness. Targeted drug delivery devices are under development. Due to the enabling nature of these systems and because of the significant impact they can have on both commercial and defense applications, industry as well as federal governments have taken special interest in seeing growth nurtured in this field. MEMS/NEMS and BioMEMS/BioNEMS are the next logical step in the “silicon revolution.”

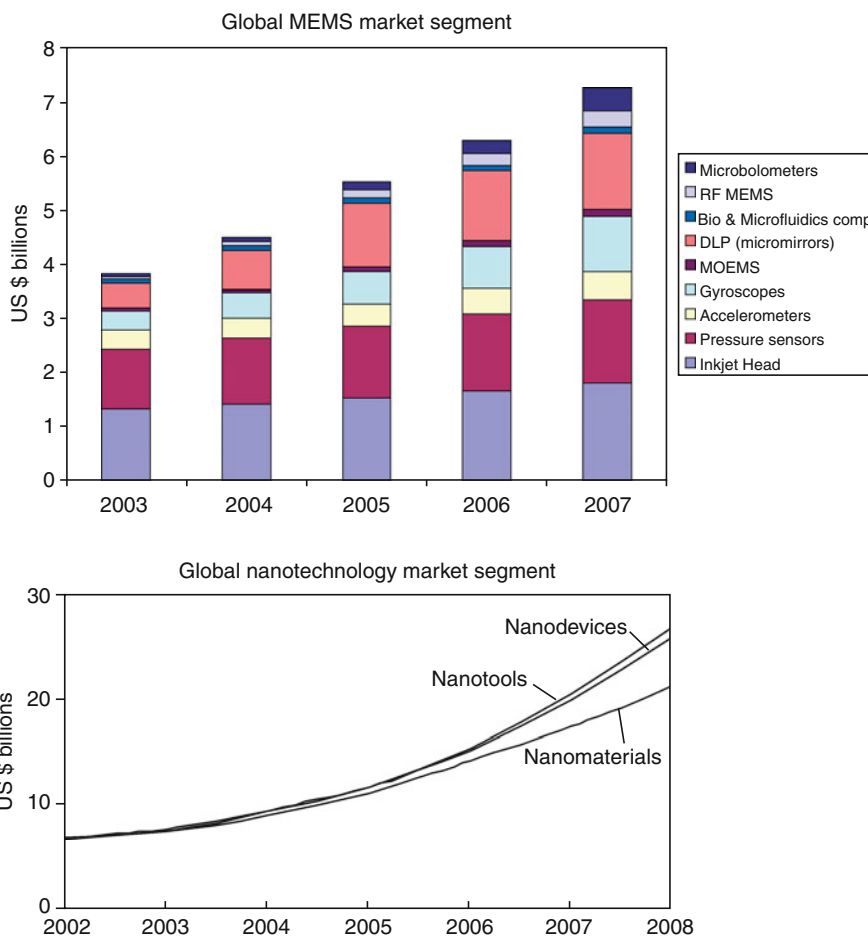
Background and Research Expenditures

On December 29, 1959 at the California Institute of Technology, Nobel Laureate Richard P. Feynman gave a talk at the Annual Meeting of the American Physical Society that has become one of the twentieth century classic science lectures, titled “There’s Plenty of Room at the Bottom” [9]. He presented a technological vision of extreme miniaturization in 1959, several years before the word “chip” became part of the lexicon. He talked about the problem of manipulating and controlling things on a small scale. Extrapolating from known physical laws, Feynman envisioned a technology using the ultimate toolbox of nature,

building nanoobjects atom by atom or molecule by molecule. Since the 1980s, many inventions and discoveries in the fabrication of nanoobjects have been testaments to his vision. In recognition of this reality, the National Science and Technology Council (NSTC) of the White House created the Interagency Working Group on Nanoscience, Engineering and Technology (IWGN) in 1998. In a January 2000 speech at the same institute, former President W. J. Clinton talked about the exciting promise of “nanotechnology” and the importance of expanding research in nanoscale science and technology more broadly. Later that month, he announced in his State of the Union Address an ambitious \$497 million federal, multi-agency national nanotechnology initiative (NNI) in the fiscal year 2001 budget, and made the NNI a top science and technology priority [10, 11]. The objective of this initiative was to form a broad-based coalition in which the academe, the private sector, and local, state, and federal governments work together to push the envelope of nanoscience and nanoengineering to reap nanotechnology’s potential social and economic benefits.

The funding in the USA has continued to increase. In January 2003, the US senate introduced a bill to establish a National Nanotechnology Program. On December 3, 2003, former President George W. Bush signed into law the twenty-first century Nanotechnology Research and Development Act. The legislation put into law programs and activities supported by the National Nanotechnology Initiative. The bill gave

Nanotechnology,
Fig. 2 Global MEMS and
 nanotechnology market
 segments



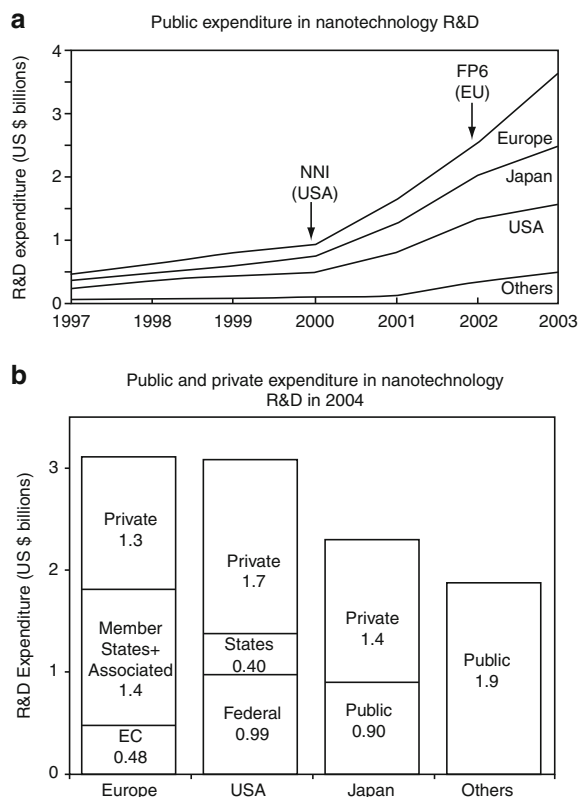
nanotechnology a permanent home in the federal government and authorized \$3.7 billion to be spent in the 4-year period beginning in October 2005 for nanotechnology initiatives at five federal agencies. The funds would provide grants to researchers, coordinate R&D across five federal agencies (National Science Foundation (NSF), Department of Energy (DOE), NASA, National Institute of Standards and Technology (NIST), and Environmental Protection Agency (EPA)), establish interdisciplinary research centers, and accelerate technology transfer into the private sector. In addition, the Departments of Defense (DOD), Homeland Security, Agriculture, and Justice as well as the National Institutes of Health (NIH) also fund large R&D activities. They currently account for more than one-third of the federal budget for nanotechnology.

The European Union (EU) made nanosciences and nanotechnologies a priority in the Sixth Framework

Program (FP6) in 2002 for a period of 2003–2006. They had dedicated small funds in FP4 and FP5 before. FP6 was tailored to help better structure European research and to cope with the strategic objectives set out in Lisbon in 2000. Japan identified nanotechnology as one of its main research priorities in 2001. The funding levels increased sharply from \$400 million in 2001 to around \$950 million in 2004. In 2003, South Korea embarked upon a 10-year program with around \$2 billion of public funding, and Taiwan has committed around \$600 million of public funding over 6 years. Singapore and China are also investing on a large scale. Russia is well funded as well.

Figure 3a shows the public expenditure breakdown in nanotechnology R&D around the world, with about \$5 billion in 2004, being about equal by USA, Japan, and Europe. Next public expenditure on a per capita basis is compared. The average expenditures per





Nanotechnology, Fig. 3 Breakdown of public expenditure in nanotechnology R&D (a) around the world (European Commission, 2003), and (b) by public and private resources in 2004 (European Commission 2005; figures from private sources based upon Lux Research)

capita for the USA, EU-25, and Japan are about \$3.7 billion, \$2.4 billion, and \$6.2 billion, respectively [12]. Figure 3b shows the breakdown of expenditures in 2004 by public and private sources with more than \$10 billion spent in nanotechnology research. Two thirds of this came from corporate and private funding. The private expenditure in USA and Japan was slightly larger than that of the public, whereas in Europe it was about one-third. Figure 4 shows the public and private expenditure breakdown in 2004 in various countries. Japan and USA had the largest expenditure, followed by Germany, Taiwan, South Korea, UK, Australia, China, France, and Italy. Figure 5 shows a breakdown of worldwide publications and patents. USA and Canada led, followed by Europe and Asia. Figure 6 shows the breakdown in start-up companies around the world (1997–2002). Entrepreneurship in USA is clearly evident followed by Europe.

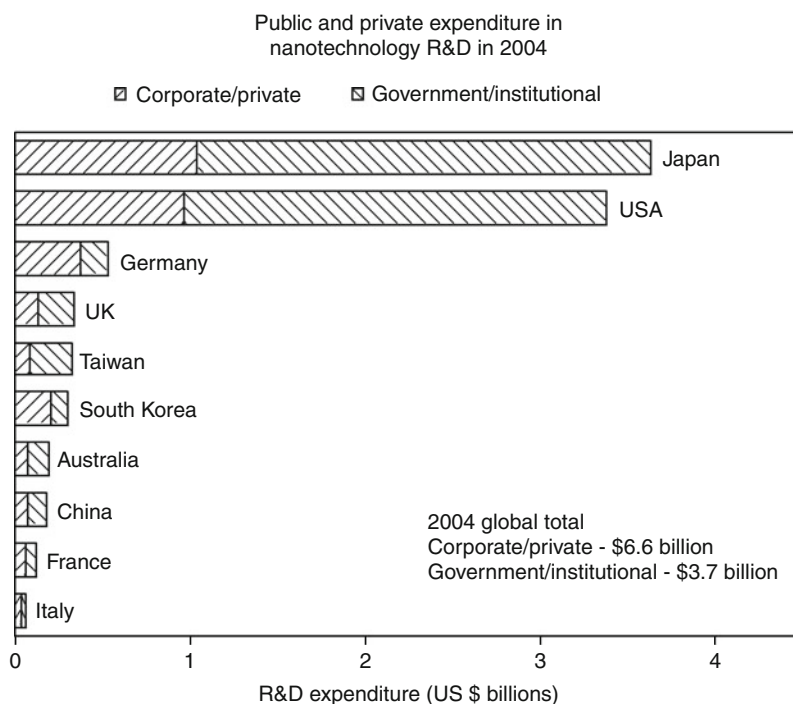
Applications in Different Fields

Science and technology continue to move forward in masking the fabrication of micro/nanodevices and systems possible for a variety of industrial, consumer, and biomedical applications (e.g., [4, 13, 14]). A variety of MEMS devices have been produced, and some are commercially used [4, 15–23, 45]. A variety of sensors are used in industrial, consumer, defense, and biomedical applications. Various micro/nanostructures or micro/nanocomponents are used in micro-instruments and other industrial applications such as micromirror arrays. The largest “killer” MEMS applications include accelerometers (some 90 million units installed in vehicles in 2004), silicon-based piezoresistive pressure sensors for manifold absolute pressure sensing for engines and for disposable blood pressure sensors (about 30 million units and about 25 million units, respectively), capacitive pressure sensors for tire pressure measurements (about 37 million units in 2005), thermal inkjet printheads (about 500 million units in 2004), and digital micromirror arrays for digital projection display (about \$700 million revenue in 2004), and optical cross-connections in telecommunications. Other applications of MEMS devices include chemical/biosensors and gas sensors, microresonators, infrared detectors and focal plane arrays for earth observations, space science and missile defense applications, pico-satellites for space applications, fuel cells, and many hydraulic, pneumatic, and other consumer products. MEMS devices are also being pursued in magnetic storage systems [24], where they are being developed for super compact and ultrahigh-recording-density magnetic disk drives.

NEMS are produced by nanomachining in a typical top-down approach and bottom-up approach, largely relying on nanochemistry [4, 25–30, 46]. Examples of NEMS include microcantilevers with integrated sharp nanotips for scanning tunneling microscopy (STM) and atomic force microscopy (AFM), quantum corral formed using STM by placing atoms one by one, AFM cantilever array for data storage, AFM tips for nanolithography, dip-pen lithography for printing molecules, nanowires, carbon nanotubes, quantum wires (QWRs), quantum boxes (QBs), quantum-dot transistors, nanotube-based sensors, biological (DNA) motors, molecular gears by attaching benzene molecules to the outer walls of carbon nanotubes, devices incorporating nm-thick films (e.g., in giant

Nanotechnology,

Fig. 4 Breakdown of public and private expenditures in nanotechnology R&D in 2004 in various countries [8]



magneto-resistive or GMR read/write magnetic heads and magnetic media) for magnetic rigid disk drives and magnetic tape drives, nanopatterned magnetic rigid disks, and nanoparticles (e.g., nanoparticles in magnetic tape substrates and magnetic particles in magnetic tape coatings).

Nanoelectronics can be used to build computer memory using individual molecules or nanotubes to store bits of information, molecular switches, molecular or nanotube transistors, nanotube flat-panel displays, nanotube integrated circuits, fast logic gates, switches, nanoscopic lasers, and nanotubes as electrodes in fuel cells.

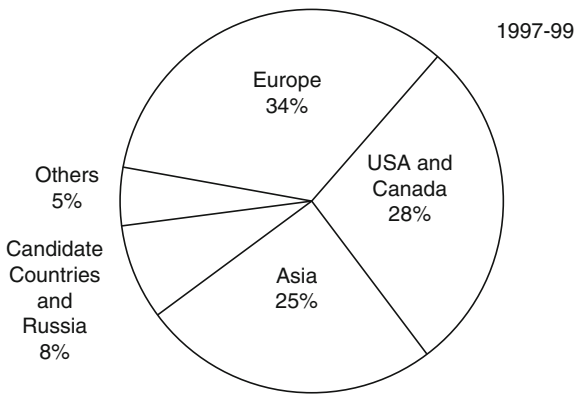
BioMEMS/BioNEMS are increasingly used in commercial and defense applications (see, e.g., [4, 31–37]). They are used for chemical and biochemical analyses (biosensors) in medical diagnostics (e.g., DNA, RNA, proteins, cells, blood pressure and assays, and toxin identification) [37], tissue engineering [38], and implantable pharmaceutical drug delivery [39, 40]. Biosensors, also referred to as biochips, deal with liquids and gases. There are two types of biosensors. A large variety of biosensors are based on micro/nanofluidics. Micro/nanofluidic devices offer the ability to work with smaller reagent volumes and shorter reaction times, and perform analyses multiple

times at once. The second type of biosensors includes micro/nanoarrays which perform one type of analysis thousands of times. Micro/nanoarrays are a tool used in biotechnology research to analyze DNA or proteins to diagnose diseases or discover new drugs. Also called DNA arrays, they can identify thousands of genes simultaneously [33]. They include a microarray of silicon nanowires, roughly a few nm in size, to selectively bind and detect even a single biological molecule, such as DNA or protein, by using nanoelectronics to detect the slight electrical charge caused by such binding, or a microarray of carbon nanotubes to electrically detect glucose.

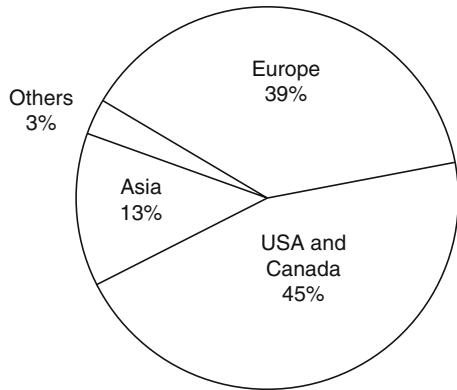
After the tragedy of September 11, 2001, concern of biological and chemical warfare has led to the development of handheld units with bio- and chemical sensors for detection of biological germs, chemical or nerve agents, and mustard agents, and chemical precursors to protect subways, airports, water supply, and population at large.

BioMEMS/BioNEMS are also being developed for minimal invasive surgery including endoscopic surgery, laser angioplasty, and microscopic surgery. Other applications include implantable drug-delivery devices – micro/nanoparticles with drug molecules encapsulated in functionalized shells for site-specific

a Worldwide publications in nanotechnology



b Worldwide patents in nanotechnology



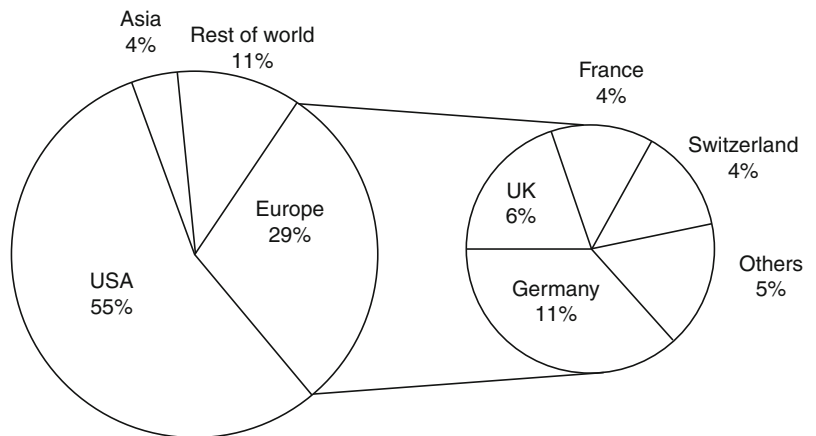
Nanotechnology, Fig. 5 Breakdown of (a) worldwide publications and (b) worldwide patents (European Commission 2003)

targeting applications, and a silicon capsule with a nanoporous membrane filled with drugs for long-term delivery.

Various Issues

There is an increasing need for a multidisciplinary, system-oriented approach to the manufacturing of micro/nanodevices which function reliably. This can only be achieved through the cross-fertilization of ideas from different disciplines and the systematic flow of information and people among research groups. Common potential failure mechanisms for MEMS/NEMS requiring relative motion that need to be addressed in order to increase their reliability are: adhesion, friction, wear, fracture, fatigue, and contamination [4, 41–44]. Surface micro/nanomachined structures often include smooth and chemically active surfaces. Due to large surface area to volume ratio in MEMS/NEMS, they are particularly prone to stiction (high static friction) as part of normal operation. Fracture occurs when the load on a microdevice is greater than the strength of the material. Fracture is a serious reliability concern, particularly for brittle materials used in the construction of these components, since it can immediately or would eventually lead to catastrophic failures. Additionally, debris can be formed from the fracturing of microstructures, leading to other failure processes. For less brittle materials, repeated

Start-up companies in nanotechnology (1997-2002)



Nanotechnology, Fig. 6 Breakdown of start-up companies around the world (1997–2002) (CEA, Bureau d’Etude Marketing)

loading over a long period of time causes fatigue that would also lead to the breaking and fracturing of the device. In principle, this failure mode is relatively easy to observe and simple to predict. However, the material properties of thin films are often not known, making fatigue predictions error-prone.

Many MEMS/NEMS devices operate near their thermal dissipation limit. They may encounter hot spots that may cause failures, particularly in weak structures such as diaphragms or cantilevers. Thermal stressing and relaxation caused by thermal variations can create material delamination and fatigue in cantilevers. In large temperature changes, as experienced in the space environment, bimetallic beams will also experience warping due to mismatched coefficients of thermal expansion. Packaging has been a big problem. The contamination that probably happens in packaging and during storage also can strongly influence the reliability of MEMS/NEMS. For example, a particulate dust landed on one of the electrodes of a comb drive can cause catastrophic failure. There are no MEMS/NEMS fabrications standards, which make it difficult to transfer fabrication steps in MEMS/NEMS between boundaries.

Obviously, studies of determination and suppression of active failure mechanisms affecting this new and promising technology are critical to high reliability of MEMS/NEMS and are determining factors in successful practical application.

Adhesion between a biological molecular layer and the substrate, referred to as “bioadhesion,” and reduction of friction and wear of biological layers, biocompatibility, and biofouling for BioMEMS/BioNEMS are important.

Mechanical properties are known to exhibit a dependence on specimen size. Mechanical property evaluation of nanometer-scaled structures is carried out to help design reliable systems since good mechanical properties are of critical importance in such applications. Some of the properties of interest are: Young’s modulus of elasticity, hardness, bending strength, fracture toughness, and fatigue life. Finite element modeling is carried out to study the effects of surface roughness and scratches on stresses in nanostructures. When nanostructures are smaller than a fundamental physical length scale, conventional theory may no longer apply, and new phenomena

emerges. Molecular mechanics is used to simulate the behavior of a nano-object.

The societal, ethical, political, and health/safety implications are getting major attention [12]. One of the prime reasons is to avoid some of the public skepticism that surrounded the debate over biotech advances such as genetically modified foods, while at the same time dispelling some of the misconceptions the public may already have about nanotechnology. Health/safety issues need to be addressed as well. For example, one key question is what happens to nanoparticles (such as buckyballs or nanotubes) in the environment and are they toxic in the human body, if digested.

Cross-References

- ▶ [Biomimetics](#)
- ▶ [MEMS](#)
- ▶ [Micromachines](#)
- ▶ [Nanocalorimetry](#)
- ▶ [Nanofabrication](#)
- ▶ [NEMS](#)

References

1. Chen, R.J., Choi, H.C., Bangsaruntip, S., Yenilmex, E., Tang, X., Wang, Q., Chang, Y.L., Dai, H.: An investigation of the mechanisms of electrode sensing of protein adsorption on carbon nanotube devices. *J. Am. Chem. Soc.* **126**, 1563–1568 (2004)
2. Srivastava, D.: Computational nanotechnology of carbon nanotubes. In: Meyyappan, M. (ed.) *Carbon Nanotubes: Science and Applications*, pp. 25–36. CRC Press, Boca Raton (2004)
3. van der Wiel, W.G., De Franceschi, S., Elzerman, J.M., Fujisawa, T., Tarucha, S., Kouwenhoven, L.P.: Electron transport through double quantum dots. *Rev. Mod. Phys.* **75**, 1–22 (2003)
4. Bhushan, B.: *Springer Handbook of Nanotechnology*, 3rd edn. Springer, Heidelberg (2010)
5. Anonymous.: *Microelectromechanical systems: advanced materials and fabrication methods*, NMAB-483, National Academy, Washington, DC (1997)
6. Roukes, M.: Nanoelectromechanical systems face the future. *Phys. World* **14**, 25–31 (2001)
7. Eloy, J.C.: Status of the MEMS industry 2005. Report Yole Developpement, France. Presented at SPIE Photonics West, San Jose, California (2005)

8. Lawrence, S.: Nanotech grows up. *Technol. Rev.* **108**(6), 31 (2005)
9. Feynman, R.P.: There's plenty of room at the bottom. *Eng. Sci.* **23**, 22–36 (1960). www.zyvex.com/nanotech/feynman.html
10. Amato, I.: Nanotechnology, www.ostp.gov/nstc/html/iwgn/iwgn.public.brochure/welcome.htm or www.nsf.gov/home/crssprgm/nano/nsfnireports.htm (2000)
11. Anonymous.: National nanotechnology initiative, www.ostp.gov/nstc/html/iwgn.fy01budsuppl/nni.pdf or www.nsf.gov/home/crssprgm/nano/nsfnireports.htm (2000)
12. Anonymous.: Towards a European Strategy for Nanotechnology, European Commission, Research Directorate, Brussels (2004)
13. Anonymous.: Small Tech 101 – An Introduction to micro and nanotechnology, Small Times (2003)
14. Schulenburg, M.: Nanotechnology – Innovation for Tomorrow's World. European Commission Research DG, Brussels (2004)
15. Bhushan, B.: Tribology Issues and Opportunities in MEMS. Kluwer, Dordrecht (1998)
16. Elwenspoek, M., Wiegerink, R.: Mechanical Microsensors. Springer, Berlin (2001)
17. Fujimasa, I.: Micromachines: A New Era in Mechanical Engineering. Oxford University Press, Oxford (1996)
18. Hierlemann, A.: Integrated Chemical Microsensor Systems in CMOS Technology. Springer, Berlin (2005)
19. Hsu, T.R.: MEMS and Microsystems: Design and Manufacture. McGraw-Hill, Boston (2002)
20. Kovacs, G.T.A.: Micromachined Transducers Sourcebook. WCB McGraw-Hill, Boston (1998)
21. Madou, M.: Fundamentals of Microfabrication: The Science of Miniaturization, 2nd edn. CRC Press, Boca Raton (2002)
22. Senturia, S.D.: Microsystem Design. Kluwer, Boston (2000)
23. Trimmer, W.S. (ed.): Micromachines and MEMS, Classical and Seminal Papers to 1990. IEEE Press, New York (1997)
24. Bhushan, B.: Tribology and Mechanics of Magnetic Storage Devices, 2nd edn. Springer, New York (1996)
25. Dresselhaus, M.S., Dresselhaus, G., Avouris, Ph.: Carbon Nanotubes – Synthesis, Structure, Properties, and Applications. Springer, Berlin (2001)
26. Drexler, K.E.: Nanosystems: Molecular Machinery, Manufacturing and Computation. Wiley, New York (1992)
27. Timp, G. (ed.): Nanotechnology. Springer, New York (1999)
28. Goddard, W.A., Brenner, D.W., Lyshevski, S.E., Iafate, G.J. (eds.): Handbook of Nanoscience, Engineering, and Technology. CRC Press, Boca Raton (2002)
29. Nalwa, H.S. (ed.): Nanostructured Materials and Nanotechnology. Academic, San Diego (2002)
30. Poole, C.P., Owens, F.J.: Introduction to Nanotechnology. Wiley, New York (2003)
31. Cheng, J., Kricka, L.J. (eds.): Biochip Technology. Harwood Academic, Philadelphia (2001)
32. Heller, M.J., Guttman, A. (eds.): Integrated Microfabricated Biodevices. Marcel Dekker, New York (2001)
33. Lai Poh San, C., Yap, E.P.H. (eds.): Frontiers in Human Genetics. World Scientific, Singapore (2001)
34. Manz, A., Becker, H. (eds.): Microsystem Technology in Chemistry and Life Sciences, Topics in Current Chemistry 194. Springer, Heidelberg (1998)
35. Mastrangelo, C.H., Becker, H. (ed.): Microfluidics and BioMEMS. In: Proceedings of SPIE, vol. 4560, SPIE, Bellingham, Washington (2001)
36. Becker, H., Locascio, L.E.: Polymer microfluidic devices. *Talanta* **56**, 267–287 (2002)
37. van der Berg, A. (ed.): Lab-on-a-Chip: Chemistry in Miniaturized Synthesis and Analysis Systems. Elsevier, Amsterdam (2003)
38. Lanza, R.P., Langer, R., Vacanti, J. (eds.): Principles of Tissue Engineering, 2nd edn. Academic, San Diego (2000)
39. Oeberg, P.A., Togawa, T., Spelman, F.A.: Sensors in Medicine and Health Care. Wiley, New York (2004)
40. Park, K. (ed.): Controlled Drug Delivery: Challenges and Strategies. American Chemical Society, Washington, DC (1997)
41. Bhushan, B.: Principles and Applications of Tribology. Wiley, New York (1999)
42. Bhushan, B.: Handbook of Micro/Nanotribology. CRC Press, Boca Raton (1999)
43. Bhushan, B.: Introduction to Tribology. Wiley, New York (2002)
44. Bhushan, B.: Nanotribology and Nanomechanics I – Measurement Techniques and Nanomechanics, II – Nanotribology, Biomimetics, and Industrial Applications, 3rd edn. Springer, Heidelberg (2011)
45. Muller, R.S., Howe, R.T., Senturia, S.D., Smith, R.L., White, R.M.: Microsensors. IEEE Press, New York (1991)
46. Rietman, E.A.: Molecular Engineering of Nanosystems. Springer, New York (2001)

Nanotechnology Applications in Polymerase Chain Reaction (PCR)

Kuo-Sheng Ma

IsaCal Technology, Inc., Riverside, CA, USA

Synonyms

Nucleic acid amplification; Oligonucleotide amplification

Definition

Polymerase chain reaction (PCR) is the most prevalent technology used in modern molecular biology research. Generally, it is a technique to amplify the amount of a piece of DNA with a specific sequence. Although the PCR technique is mature, improvement of its efficiency is still an emerging area of research. Recently, nanotechnology is getting more attention in

this application. Several nanometer-sized materials such as carbon nanotubes, gold nanoparticles, quantum dots, and metal oxide nanoparticles have been employed.

Introduction

The polymerase chain reaction (PCR) process is a technique to amplify the number of copies of a specific DNA sequence, producing hundreds of thousands of copies of DNA. This revolutionary technique was first developed by Kary Mullis [1] in 1983. Since then, it has been broadly used in medical and biological research laboratories for a variety of applications [2, 3]. These include DNA sequencing, DNA cloning, identification of functional gene, the disease diagnosis, and the identification of genetic fingerprints.

A basic PCR reaction needs several components in the mixture. These include: (a) A small amount of original DNA (mRNA in some applications) as a template which contains the oligonucleotides segment to be amplified. (b) Primers. These are short sequences of oligonucleotides in which their sequences are designed complementary to the specific site of the template DNA. Two primers are used for one amplification process and each primer binds to the different end of the segment on one of the paired DNA strands. (c) DNA polymerases. These are the enzymes that perform DNA synthesis in the amplification process. The Taq polymerase is most prevalently used in the conventional PCR process. (d) Deoxynucleotide triphosphates (dNTPs). These are the building blocks of DNA which are polymerized by the enzyme through the sequence of DNA templates. (e) Buffer solution. This is the salt solution to provide optimum condition for DNA synthesis. Divalent cations such as Mg^{2+} or Mn^{2+} are normally used in buffer solution to mediate DNA mutagenesis or decrease the error rate of DNA polymerase.

The technique is based on the cycles of repeated temperature changes of the reactant in reaction tubes. The temperature changes are normally carried out in a thermal cycler which performs heating and cooling of the tubes. A typical PCR process consists of 20–40 cycles, with each cycle consisting of three basic heating and cooling steps on discrete temperatures. The temperatures employed and the duration of specific temperature in a cycle depends on the reactant

used in the mixture for each specific process. In a conventional PCR process, as the first step (or called as denature step), the mixture is heated up to 94–98°C for 20–30 s to break the hydrogen bonds between complementary double-stranded DNA (templates) and yield single-stranded DNA molecules. In the second step (or called as annealing step), the temperature of the solution is lowered to 50–65°C for 20–40 s allowing the primers bind to the separated strands of DNA template. Sequentially, the enzyme (DNA polymerase) binds to the hybridization sites of primer and template to start DNA amplification. In the next step (also called the extension/elongation step), the temperature of solution is risen to a temperature at which the enzyme has its optimum activity. A temperature of 72°C is used for the commonly used Taq polymerase. At this step, the amount of DNA is doubled due to the enzyme synthesizing a new complementary strand to each DNA template by polymerizing dNTPs in 5'–3' direction. These three basic steps form a cycle and repetitions of this cycle lead to exponential amplification of the amount of DNA fragments. Figure 1 schematically shows a conventional PCR process.

Variations on the Basic PCR Technique

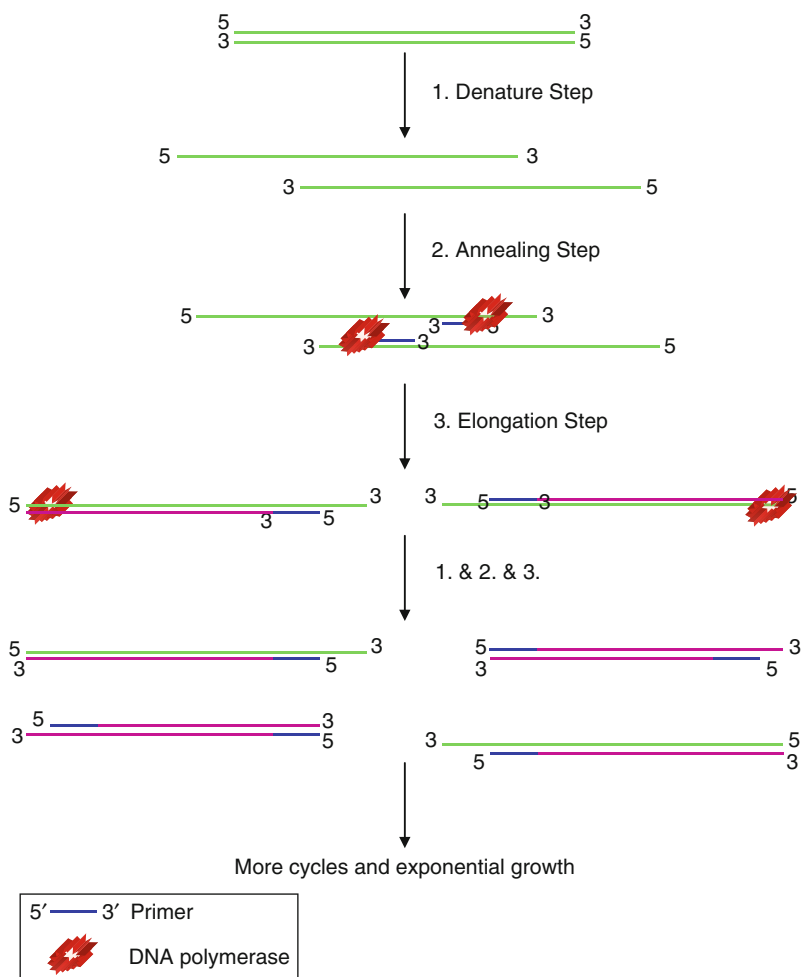
PCR is such a powerful technique to provide a relatively easy, efficient method to amplify the amount of desired DNA sequence. Since it was invented, many modifications to tailor the basic PCR thermal cycle approach to fit specific applications have been made. The features of these modified techniques are summarized in Table 1.

Nanotechnology in PCR

Although the PCR technique is mature and has been broadly employed in various applications, there is still an effort to improve the specificity and efficiency of it by using different advanced technologies. In the nanotechnology, the development and discovery of more applications of nanoparticles is an emerging area of research. Nanoparticle-assisted PCR has attracted more attentions and many nanomaterials have been studied for their potential in developing better PCR technique. Various types of nanoparticles such as carbon nanotubes, metal nanoparticles (especially

Nanotechnology Applications in Polymerase Chain Reaction (PCR),

Fig. 1 Schematic illustration of PCR processes



gold nanoparticles), semiconductor quantum dots, and metal oxide nanoparticles have been introduced into PCR and their effect on reaction efficiency studied. Some studies showed these nanoparticles dramatically improve PCR in efficiency and specificity but some studies showed contrary results. In this entry, the development of the technique which employed the nanoparticles in PCR would be introduced. The different types of nanoparticles utilized in the studies and their effect in assisting PCR as well as mechanisms will be discussed.

Carbon Nanotubes

Gao et al. [14] investigated the effects of single-walled carbon nanotubes (SWCNTs) on PCR. Their results showed the amount of PCR product increased by adding single-walled carbon nanotubes into the PCR

reactant mixture. They further discovered that adding higher concentrations of SWCNTs (> 3 ug/ul) reversed the effect (amount of PCR product decreased). Using SEM and HRTEM, they showed that DNA templates and Taq enzymes were attached to bundles of SWCNTs. XPS results further suggested that there might be a chemical reaction between SWCNTs and components in PCR solution. The authors concluded that the SWCNTs have the potential to act as catalysts in the solution, which increased the yield of PCR products.

Zhang et al. [15] had done the similar study and reported the beneficial effect of using single-walled carbon nanotubes (SWCNTs) and multi-walled carbon nanotubes (MWCNTs) to enhance the amplification efficiency of long PCR up to 14 kb. Rather than the pristine carbon nanotubes, the authors employed

Nanotechnology Applications in Polymerase Chain Reaction (PCR), Table 1 Variation techniques based on PCR process

Name	Functions	References
1. Quantitative PCR (qPCR)	A technique used to quantify the PCR products during the amplification cycles. Because the measurement is in “real time” of amplification, it is also called quantitative real-time PCR (qRT-PCR). The technique employed the fluorescent dyes in the PCR solution to measure the amount of amplified products in the log phase of the reaction before the plateau. The commonly used dyes are Syber Green or TaqMan which is a DNA probe containing a fluorophore	[4]
2. Reverse transcription PCR (RT-PCR)	A preferred technique used to detect and quantify mRNA. In mRNA detection applications, the target RNA needs to be amplified by PCR. Thus, RNA samples need to be first reverse transcribed into cDNA by reverse transcriptase and cDNA sequences will be amplified by PCR. This technique is widely used in profiling of gene expressions	[5]
3. Polymerase cycling assembly (PCA)	A technique to synthesize long DNA sequences by a PCR. It employed a pool of oligonucleotides with designed short overlapping sequences in each piece of oligonucleotides. Rather than using two primers, this technique uses multiple short oligonucleotides which are part of the single strand of the target sequences. During the elongation cycle, the rest of sequences will be filled in by polymerase. The technique is popular for the production of synthetic genes and even synthetic genomes	[6]
4. Helicase-dependent amplification	This technique is similar to traditional PCR but uses a constant temperature instead of temperature cycling in denaturation and annealing steps. Rather than thermal denaturation, a DNA helicase is used to separate the double strand of DNA which provides single-stranded DNA template for primers and polymerase binding	[7]
5. Hot start PCR	A modified process to reduce nonspecific amplification of DNA by inactivating the Taq polymerase at a lower temperature in the initial stages of the conventional PCR. It can be performed manually by adding a primer to start amplification after heating the reaction components to the melting temperature. Common methods for this technique include chemical modifications, wax-barrier methods, inhibition by a Taq-directed antibody, and by the presence of covalently bound inhibitors. Due to lack of nonspecific hybridization of primers to template, the amplified DNA bands tend to be cleaner	[8, 9]
6. Methylation-specific PCR (MSP)	A technique is used to detect methylation of CpG islands in the genomic DNA. DNA is first treated with sodium bisulfite, which converts unmethylated cytosine bases to uracil. A special primer is designed to distinguish methylated from unmethylated DNA. The process takes advantage of the sequence differences resulting from bisulfite modification. MSP using qPCR can also be performed to obtain quantitative rather than qualitative information about methylation	[10]
7. Allele-specific PCR	It is a variation of the PCR which is used as a diagnostic or cloning technique to identify single-nucleotide polymorphism (SNPs) (single-base differences in DNA). It requires prior knowledge of a DNA sequence and differences between alleles. The allele-specific PCR uses primers whose 3' ends encompass the SNP	[11]
8 Solid phase PCR	In a solid phase PCR (SP-PCR), the amplification takes place directly on a solid substrate. The technique encompasses multiple modifications, including polony amplification where PCR colonies are derived in a gel matrix; bridge PCR in which primers are covalently linked to a solid-support surface; conventional solid phase PCR in which an asymmetric PCR is employed. A primer is immobilized on a solid support which has the sequence matching one of the aqueous primers	[12, 13]

the CNTs with surface functionalization. Both hydroxylic and carboxylic functionalized CNTs had been studied and had similar enhancing effects of long PCR.

The opposite effect on CNTs enhancing PCR had also been reported. Yang et al. [16] studied the effects of multi-walled carbon nanotubes and single-walled carbon nanotubes functionalized with and without carboxylic groups on PCR. Their

results showed that CNTs reduced the efficiency of PCR. From their report, the inhibition effect on PCR was increased in the order of CNT-COOH > pristine CNT and SWCNT > MWCNT. The authors attributed the inhibition effect of nanomaterials on PCR to the decrease of the DNA polymerase activity, which they believed to be a significant effect of how nanomaterials impact biological systems.

Gold Nanoparticles

Small gold nanoparticles (Au-NPs) or nanogold had also been studied of their influences on PCR. Hu et al. [17] reported nanogold significantly improved the specificity and yield of PCR. The author attributed the enhancement to the greater affinity of Au-NPs to single-stranded DNA (ssDNA) than to double-stranded DNA (dsDNA) which reduced the mispairing of nucleotides to the template sequence in the elongation step. In their another work [18], the authors did the multiple rounds of PCR with 10 nm Au-NPs. The result showed the Au-NPs effectively enhanced the specificity and yield of error-prone amplification. As the result showed, after six rounds of amplification, the target band was still manifested by a single bright band in gel electrophoresis. However, the target band disappeared after four rounds of amplification of the sample without adding Au-NPs. This study has shown with Au-NPs enhancement, a large amount of target DNA can be produced even with very low copies of original DNA samples. Liu et al. [19] had done the similar study using 13 nm Au-NPs in the PCR reagent and discovered the efficiency was increased. In this study, the authors also compared different PCR systems by using different DNA polymerases, various DNA sizes for amplification, and complex samples. Their results demonstrated that Au-NPs increased the efficiency by fivefold to tenfold in a conventional PCR and by 10^4 -fold in a real-time PCR. The authors hypothesized that the enhancement is due to the excellent heat transfer property of the Au-NPs in PCR solution.

Although Au-NPs enhanced PCR had been investigated, different results had been reported. Emslie et al. [20] reported the contrary results from their study. The authors investigated the efficiency of real-time quantitative polymerase chain reaction (qPCR) by adding Au-NPs into the PCR solution. It was found that Au-NPs affect the qPCR amplification due to fluorescence quenching by Au-NPs. Rather than the quenching effect, it was discovered that Au-NPs destabilized the double-stranded PCR products which further reduced the efficiency of qPCR. In contrast to those previous reports, Willson et al. [21] discovered that Au-NPs suppressed the amplification of longer products and rather enhanced the amplification of shorter products. In their discovery, when Au-NPs were added in PCR solution, some of the DNA polymerases were nonspecifically absorbed on nanoparticles, and thus reduced the amount of polymerase and the efficiency

of amplification. In their recent report, Hu et al. [22] also discovered that the Au-NPs can enhance the PCR efficiency only by employing native Taq polymerase. No improvement was found in a PCR based on the recombinant Taq polymerase. However, the underlying mechanism was still unclear.

Yeow et al. [23] had done the study of the proposed mechanism for Au-NPs affecting the PCR. The authors studied the PCR efficiency with different Au-NPs size (5 nm, 10 nm, and 20 nm) and different concentrations. Evidence from their study showed that the possible mechanism for Au-NPs affecting the PCR is through DNA polymerase-AuNPs binding which modulated the enzyme activity for the amplification. Besides, the binding strongly related to the total surface area of the nanoparticles in the reaction solution.

Semiconductor Quantum Dots

Another type of nanoparticles such as semiconductor quantum dots (QDs) has been studied in affecting the PCR efficiency. Li et al. [24] reported that the QDs dramatically improved the yield and specificity of PCR. In their study, by employing mercaptoacetic acid (MAA)-coated QDs, yield and specificity of PCR enhanced even at lower annealing temperature. Rather than MAA surface coated QDs, streptavidin surface-modified QDs also effectively improved the specificity of PCR which demonstrated that the effect was not due to the molecules of surface modification but instead due to the QDs itself. Further study is needed to discover the underlying mechanism. Another study using QDs in PCR was done by Xu et al. [25]. In this study, the authors also discovered that the QDs could increase the specificity of the PCR at different annealing temperatures. Thus, as the authors concluded, this method could make the PCR widely applicable, especially in the multi-round PCR with different annealing temperatures.

Other Nanoparticles

Several different types of nanoparticles have also been studied of their affects in PCR. Mahapatra et al. [26] recently reported that the TiO₂ nanoparticles with the size of 25 nm diameter significantly enhanced the PCR efficiency even with various types of templates (i.e., plasmid DNA, genomic DNA, and complementary DNA). Besides, the yield was also enhanced by utilizing TiO₂ nanoparticles in the PCR. Investigations of the mechanism by simulations showed

evidences of fast and efficient heat transfer in the PCR solution in the presence of TiO₂ nanoparticles to enhance the PCR efficiency. Consistent with the simulation results, experimentally, the authors observed the increase of denaturation of genomic DNA in the presence of TiO₂ nanoparticles in PCR solution which indicated higher thermal conductivity through the reaction buffer. This technique may be useful for reduction of the overall PCR reaction time and further enhanced the efficiency and specificity of PCR in a variety of samples.

The effect of C-60 on PCR was studied by Jiang et al. [27] using a qPCR system. The authors reported a dramatical enhancement of the qPCR at the beginning of the exponential phase but a significant inhibition at the plateau region. In addition, the resulting melting curve demonstrated that C-60 decreased the melting temperatures (T_m) of the template DNA fragments. The authors hypothesized this is due to the enhancement of qPCR efficiency by C-60 in the initial phase of PCR. However, when there is more DNA amplicon in solution, C-60 can disrupt DNA replication by binding to DNA to decrease the access of templates and changing the conformation of DNA templates which reduced the amplification efficiency at the later cycles of the reaction.

Conclusion

Although the utilization of nanoparticles in conventional PCR is just emerging, the potentials to enhance the process have been demonstrated by some of the works. Due to its excellent thermal conductivity, high surface-to-volume ratio, and high affinity to bind molecules on surface, nanoparticle-assisted PCR can greatly reduce the total process time and specificity of the process. This is critical for the technique to be used in medical diagnosis and in-field detection applications. In addition, due to the high efficiency of process with NPs enhancement, less amount of expensive reagents is needed which significantly reduce the experimental costs. However, until now, the NP-assisted technology has not yet been revolutionized in practice. To enable users to embrace the technique, future development in identifying type of NPs, size of NPs, concentration of NPs in PCR reagent, and specific PCR components are all essential. In addition, optimization of process design to maximize the

enhancements and further reduce the time and cost of the process is also important to evolve this technology out of research and to be used routinely.

Cross-References

- ▶ [Carbon-Nanotubes](#)
- ▶ [Functionalization of Carbon Nanotubes](#)
- ▶ [Gold Nanorods](#)
- ▶ [Nanoparticles](#)

References

1. Mullis, K., et al.: Specific enzymatic amplification of DNA invitro – the polymerase chain-reaction. Cold Spring Harb. Symp. Quant. Biol. **51**, 263–273 (1986)
2. Saiki, R.K., et al.: Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle-cell anemia. *Science* **230**(4732), 1350–1354 (1985)
3. Saiki, R.K., et al.: Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. *Science* **239**(4839), 487–491 (1988)
4. Kellogg, D.E., Sninsky, J.J., Kwok, S.: Quantitation of HIV-1 proviral DNA relative to cellular DNA by the polymerase chain-reaction. *Anal. Biochem.* **189**(2), 202–208 (1990)
5. Mackay, I.M., Arden, K.E., Nitsche, A.: Real-time PCR in virology. *Nucleic Acids Res.* **30**(6), 1292–1305 (2002)
6. Stemmer, W.P.C., et al.: Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164**(1), 49–53 (1995)
7. Vincent, M., Xu, Y., Kong, H.M.: Helicase-dependent isothermal DNA amplification. *EMBO Rep.* **5**(8), 795–800 (2004)
8. Chou, Q., et al.: Prevention of pre-PCR mis-priming and primer dimerization improves Low-copy-number amplifications. *Nucleic Acids Res.* **20**(7), 1717–1723 (1992)
9. Kellogg, D.E., et al.: Taqstart antibody(T_m) – Hot start PCR facilitated by a neutralizing neutralizing monoclonal-antibody directed against Taq DNA-polymerase. *Biotechniques* **16**(6), 1134–1137 (1994)
10. Herman, J.G., et al.: Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci. U.S.A.* **93**(18), 9821–9826 (1996)
11. Newton, C.R., et al.: Analysis of any point mutation in DNA – the amplification refractory mutation system (arms). *Nucleic Acids Res.* **17**(7), 2503–2516 (1989)
12. Bing, D.H., Sawosik, T.M., Word, C.J.: Assay performance results with the AmpliType(R) PM PCR amplification and typing kit. *Crime Lab. Dig.* **23**(2), 27–45 (1996)
13. Khan, Z., Poetter, K., Park, D.J.: Enhanced solid phase PCR: mechanisms to increase priming by solid support primers. *Anal. Biochem.* **375**(2), 391–393 (2008)
14. Cui, D.X., et al.: Effects of single-walled carbon nanotubes on the polymerase chain reaction. *Nanotechnology* **15**(1), 154–157 (2004)

15. Zhang, Z.Z., et al.: Aqueous suspension of carbon nanotubes enhances the specificity of long PCR. *Biotechniques* **44**(4), 537–545 (2008)
16. Yi, C.Q., et al.: Interactions between carbon nanotubes and DNA polymerase and restriction endonucleases. *Nanotechnology* **18**(2), 6 (2007)
17. Li, H.K., et al.: Nanoparticle PCR: nanogold-assisted PCR with enhanced specificity. *Angew. Chem.Int. Ed.* **44**(32), 5100–5103 (2005)
18. Pan, J.K., et al.: Nanogold-assisted multi-round polymerase to chain reaction (PCR). *J. Nanosci. Nanotechnol.* **7**(12), 4428–4433 (2007)
19. Li, M., et al.: Enhancing the efficiency of a PCR using gold nanoparticles. *Nucleic Acids Res.* **33**(21), 10 (2005)
20. Haber, A.L., et al.: Addition of gold nanoparticles to real-time PCR: effect on PCR profile and SYBR Green I fluorescence. *Anal. Bioanal. Chem.* **392**(5), 887–896 (2008)
21. Vu, B.V., Litvinov, D., Willson, R.C.: Gold nanoparticle effects in polymerase chain reaction: favoring of smaller products by polymerase adsorption. *Anal. Chem.* **80**(14), 5462–5467 (2008)
22. Yang, W.C., et al.: Evaluation of gold nanoparticles as the additive in real-time polymerase chain reaction with SYBR Green I dye. *Nanotechnology* **19**(25), 9 (2008)
23. Wan, W.J., Yeow, J.T.W.: The effects of gold nanoparticles with different sizes on polymerase chain reaction efficiency. *Nanotechnology* **20**(32), 5 (2009)
24. Ma, L., et al.: Maximizing specificity and yield of PCR by the quantum dot itself rather than property of the quantum dot surface. *Biochimie* **91**(8), 969–973 (2009)
25. Wang, L.B., et al.: Effects of quantum dots in polymerase chain reaction. *J. Phys. Chem. B* **113**(21), 7637–7641 (2009)
26. Khaliq, A., et al.: Enhancement in the efficiency of polymerase chain reaction by TiO₂ nanoparticles: crucial role of enhanced thermal conductivity. *Nanotechnology* **21**(25), 11 (2010)
27. Liang, Y., et al.: C-60 affects DNA replication in vitro by decreasing the melting temperature of DNA templates. *Carbon* **47**(6), 1457–1465 (2009)

Nanotechnology in Cardiovascular Diseases

Lu Dai¹, Yongfen Qi¹, Lixin Jia¹, Wei Yu¹ and Jie Du²

¹The Key Laboratory of Remodeling-related Cardiovascular Diseases, Capital Medical University, Ministry of Education, Beijing, China

²Beijing Institute of Heart Lung and Blood Vessel Diseases, Beijing Anzhen Hospital, Beijing, China

Definition

Nanomolecular imaging is a new type of imaging with a combination of nanomaterial and molecular biology,

using nanoparticles to visualize, characterize, and measure biological processes at the molecular and cellular levels in living systems, including radiotracer imaging/nuclear medicine, magnetic resonance (MR) imaging, MR spectroscopy, optical imaging, ultrasound, and other nanomolecular imaging.

Overview

Recently, the emergence of nanomolecular imaging has set the stage for an evolutionary leap in the diagnosis and therapeutics of diseases such as cancer, and neurological and cardiovascular diseases (CVD) [1–3]. CVD is the leading cause of morbidity and mortality in the world. The pathology underpinning CVD is chronic inflammation involving the injury of arterial wall, infiltration leukocytes, and remodeling of vasculature. The main benefits of nanomolecular imaging are derived from combination of cell-type-specific molecular-based, molecular-molecular recognition with nanoparticle-based probes, leading to accurate and high resolution of diseased sites, and to the early diagnosis and treatment for CVDs, including atherosclerosis, thrombosis, and myocardial infarction [4–6]. The progress in nanomolecular imaging will play an important role not only in molecular imaging, but also in the development of biosensors, biomarkers, and new drug delivery system.

Nanoparticles are defined as ranging from 5 to 700 nm in diameter and can enter the cells and function inside the cells (such as generation of intracellular imaging and targeting-specific molecular pathways) [4] (see ► [Nanoparticles](#)). For example, new types of nanoparticles have been recently introduced to generate contrast enhancement and improved sensitivity, such as infrared dyes and quantum dot (QD) nanocrystals for optical imaging [7], gold colloids or radiotracer-based nanoassemblies for X-ray contrast [4], gas-filled microbubbles for contrast using ultrasound [8], and nanoparticles of gadolinium or iron oxide for magnetic resonance contrast [4]. These abilities open up a large number of exciting possibilities for early detection, the potential treatment of disease, and a novel type of drug delivery system. Nanomolecular imaging has also been applied to imaging techniques including luminescence imaging and spectroscopy. Indeed, current imaging device can be applied for nanomolecular imaging, such as positron emission tomography (quantitative-PET) and

magnetic resonance spectroscopy (MRS), magnetic resonance spectroscopic imaging (MRSI), diffusion spectroscopy (d-MRI) and functional magnetic resonance (f-MRI).

Key Research Findings

Nanomolecular Imaging and Potential Therapy in Cardiovascular Disease

CVD is the largest cause of mortality in the world, accounting for about 40% of deaths. CVD occurs at a higher rate in diabetes with mortality exceeding 50%. Risk factors, such as visceral obesity, hypertension, and dyslipidemia, are all associated with increased risk of CVD. The major underlying disease of most CVD is atherosclerosis, characterized by a complex multifactorial pathophysiology such as the slow accumulation of lipid and inflammatory products in the vessel wall that further leads to calcification and fibrosis in atherosclerotic plaques. Inflammation in the vessel wall is now considered to play an essential role in the initiation, progression, and the final steps of atherosclerosis, namely, plaque destabilization and eventually plaque rupture (Fig. 1) [9]. Atherosclerotic plaques can be divided into two types: stable plaque and unstable plaque (a vulnerable plaque, causing myocardial infarction and stroke). Unstable plaques are at higher risk for plaque disruption, characterized by fibrous cap rupture, ulceration following spontaneous denudation of the endothelium, or intraplaque hemorrhage. This plaque disruption leads to local thrombus formation and vessel occlusion, recognized as the proximate cause of the majority of clinical CVD events such as a heart attack or stroke. Although there have been significant advances in vascular imaging (such as intravascular ultrasound and MRI), current imaging methods are still insufficiently capable for elucidating plaque composition. Therefore, it is critical to develop novel techniques to identify vulnerable plaques. In recent years, nanomolecular imaging provides exciting opportunities to identify vulnerable plaque to reduce the risk of acute coronary syndromes.

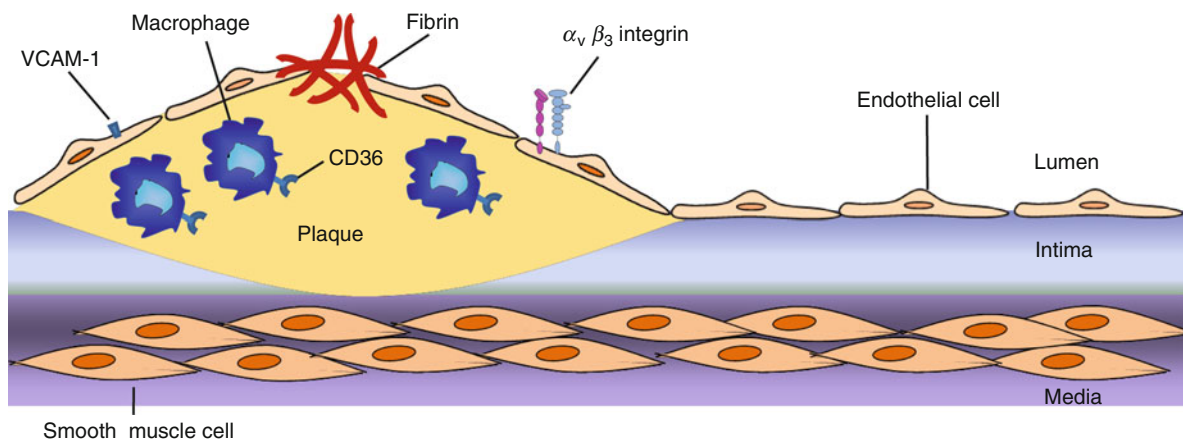
Endothelial Cells

The initiating event for the development of atherosclerosis is the endothelial layer inflammation or dysfunction. Increased adhesiveness of the endothelium leads to the infiltration of leukocytes which could engulf oxidized lipid particles and develop into inflammatory foam cells

[10]. These processes can be determined by measuring inflammatory markers, such as vascular cell adhesion molecule-1 (VCAM-1), a molecule found in blood vessel walls at the onset of inflammation – a first stage in the development of atherosclerotic plaques. Kelly et al. identified VCAM-1 internalizing peptide-28 nanoparticles (VINP-28), a novel VCAM-1 peptide affinity ligand from phage display that undergoes internalization and trapping in cells that express VCAM-1, resulting in higher signal amplification ion [11]. That offers an opportunity for detecting early atheromata (fatty streak-like lesions), and targeting VCAM-1-expressing cells of human carotid atherosclerotic plaques.

Macrophages

Macrophages have emerged as a key component of atherosclerosis, which accumulate cholesterol esters to form lipid-laden foam cells in the arterial intima and generate proteinases that lead to plaque destabilization and rupture. The accumulation of macrophage in the subendothelial space could be observed through all stages of vascular lesion development, and the increased macrophage density is associated with lesion progression and rupture [9]. Additionally, the macrophage may serve as an ideal target for the detection and potential therapy of vulnerable inflamed lesions. Recently, macrophage-specific lipid-based nanoparticles were created which improved cardiac magnetic resonance detection and characterization of human atherosclerosis [12]. These nanoparticles were made from phospholipids, a surfactant (Tween 80), and an aliphatic gadolinium complex, which targets CD36, a macrophage class B scavenger receptor. The macrophage-specific (CD36) nanoparticles bind human macrophages and improve cardiac magnetic resonance (CMR) detection. A light-activated theranostic nanoagent for targeted macrophage ablation was invented [13]. This nanoparticle was based upon cross-linked dextran-coated iron oxide (CLIO). It was modified with near-infrared fluorophores and light-activated therapeutic moieties. The therapeutic component was activated at 650 nm and could result in eradication of inflammatory macrophages. Moreover, other potential targets, such as scavenger receptor class B type 1 (SRB1), Lectin-like oxidized LDL receptor-1 (LOX-1), Receptor for Advanced Glycation Endproducts (RAGE), and urokinase-type plasminogen activator (uPA) receptor, could be more efficacious and provide an integrated imaging and therapeutic nanopatform for atherosclerosis [14–17].



Nanotechnology in Cardiovascular Diseases, Fig. 1 The structure of plaque in atherosclerosis. *VCAM-1* vascular cell adhesion molecule-1

Atherosclerotic Plaque

A major component of atherosclerosis plaques is fibrin, which forms the fibrils that entangle the blood platelets and seal up the rent in the artery wall. Fibrin deposition is one of the earliest signs of plaque rupture or erosion and intraplaque hemorrhage. Densely packed fibrin-targeted nanoparticles can be detected by scanning electron micrographs (SEM) [18]. These nanoparticles were formulated with gadolinium DTPA-bis-oleate (Gd-DTPA-BOA) and presented a highly detectable, homogeneous T1-weighted contrast enhancement that improved with increasing gadolinium level. Higher-resolution scans and scanning electron microscopy revealed that the nanoparticles were present as a thin layer over the clot surface. Then, the development of other matrix-targeted nanoparticles could provide accurate detection and localization of matrix deposits and may allow early, direct identification of vulnerable plaques.

Angiogenesis

In atherosclerosis, angiogenesis is not only related to plaque growth, but also related to plaque instability and rupture. Integrins, cell surface receptors involved in cell–cell and cell–matrix interaction, such as $\alpha_v\beta_3$, are associated with angiogenesis and have been extensively used as a target for nanomolecular imaging [19]. Patrick Winter et al. developed a paramagnetic nanoparticle contrast agent targeted specifically to $\alpha_v\beta_3$ integrins to permit noninvasive molecular imaging of plaque-associated angiogenesis. Moreover, $\alpha_v\beta_3$ integrins-targeted

paramagnetic nanoparticles can deliver fumagillin and elicit a marked antiangiogenic response with a minimal drug dosage [19].

Future Directions for Research

Nanomolecular imaging is a new evolving field, which represents the beginning of a revolutionary new age for the clinicians to enhance the diagnosis and treatment of the cardiovascular diseases relying on the powerful imaging probes and hardware. While it is still in the developing phase, the nanomolecular imaging has resulted in lots of nanoparticles that vary in size, shape, charge, chemistry, coating, and solubility. More importantly, nanomolecular imaging must utilize the progress of molecular and cellular biology in pathological conditions. At the same time, the potential toxicity of these nanoparticles should also be paid more attention to [20] (see ► [Cellular Mechanisms of Nanoparticle's Toxicity](#) and ► [Nanoparticle Cytotoxicity](#)). The nano-imaging agent synthesis requires the efforts on their subsequent translation to clinical practice. The rapid growth of nanomolecular imaging will change the field of imaging and imaging-guided interventions for cardiovascular diseases.

Cross-References

- [Cellular Mechanisms of Nanoparticle's Toxicity](#)
- [Nanoparticle Cytotoxicity](#)
- [Nanoparticles](#)

References

- Ferrari, M.: Cancer nanotechnology: opportunities and challenges. *Nat. Rev. Cancer* **5**, 161–171 (2005)
- Bhaskar, S., et al.: Multifunctional nanocarriers for diagnostics, drug delivery and targeted treatment across blood–brain barrier: perspectives on tracking and neuroimaging. *Part. Fibre Toxicol.* **7**, 3 (2010)
- Yang, X.: Nano- and microparticle-based imaging of cardiovascular interventions: overview. *Radiology* **243**, 340–347 (2007)
- Wickline, S.A., Neubauer, A.M., Winter, P., Caruthers, S., Lanza, G.: Applications of nanotechnology to atherosclerosis, thrombosis, and vascular biology. *Arterioscler. Thromb. Vasc. Biol.* **26**, 435–441 (2006)
- Godin, B., et al.: Emerging applications of nanomedicine for the diagnosis and treatment of cardiovascular diseases. *Trends Pharmacol. Sci.* **31**, 199–205 (2010)
- Douma, K., et al.: Nanoparticles for optical molecular imaging of atherosclerosis. *Small* **5**, 544–557 (2009)
- Stinaff, E.A., et al.: Optical signatures of coupled quantum dots. *Science* **311**, 636–639 (2006)
- Anderson, D.R., Tsutsui, J.M., Xie, F., Radio, S.J., Porter, T.R.: The role of complement in the adherence of microbubbles to dysfunctional arterial endothelium and atherosclerotic plaque. *Cardiovasc. Res.* **73**, 597–606 (2007)
- Tabas, I.: Macrophage death and defective inflammation resolution in atherosclerosis. *Nat. Rev. Immunol.* **10**, 36–46 (2010)
- Sitia, S., et al.: From endothelial dysfunction to atherosclerosis. *Autoimmun. Rev.* **9**, 830–834 (2010)
- Nahrendorf, M., et al.: Noninvasive vascular cell adhesion molecule-1 imaging identifies inflammatory activation of cells in atherosclerosis. *Circulation* **114**, 1504–1511 (2006)
- Lipinski, M.J., et al.: Macrophage-specific lipid-based nanoparticles improve cardiac magnetic resonance detection and characterization of human atherosclerosis. *JACC Cardiovasc. Imaging* **2**, 637–647 (2009)
- McCarthy, J.R., Korngold, E., Weissleder, R., Jaffer, F.A.: A light-activated theranostic nanoagent for targeted macrophage ablation in inflammatory atherosclerosis. *Small* **6**, 2041–2049 (2010)
- Stephen, S.L., et al.: Scavenger receptors and their potential as therapeutic targets in the treatment of cardiovascular disease. *Int. J. Hypertens.* **2010**, 646929 (2010)
- Farmer, D.G., Kennedy, S.: RAGE, vascular tone and vascular disease. *Pharmacol. Ther.* **124**, 185–194 (2009)
- Oka, H., et al.: Lysophosphatidylcholine induces urokinase-type plasminogen activator and its receptor in human macrophages partly through redox-sensitive pathway. *Arterioscler. Thromb. Vasc. Biol.* **20**, 244–250 (2000)
- White, S.J., Sala-Newby, G.B., Newby, A.C.: Overexpression of scavenger receptor LOX-1 in endothelial cells promotes atherogenesis in the ApoE(–/–) mouse model. *Cardiovasc. Pathol.* (2010)
- Flacke, S., et al.: Novel MRI contrast agent for molecular imaging of fibrin: implications for detecting vulnerable plaques. *Circulation* **104**, 1280–1285 (2001)
- Winter, P.M., et al.: Endothelial alpha(v)beta3 integrin-targeted fumagillin nanoparticles inhibit angiogenesis in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **26**, 2103–2109 (2006)
- Stone, V., Donaldson, K.: Nanotoxicology: signs of stress. *Nat. Nanotechnol.* **1**, 23–24 (2006)

Nanotechnology in Dental Medicine

► [Nanodentistry](#)

Nanothermodynamics

► [Surface Energy and Chemical Potential at Nanoscale](#)

Nanotoxicology

► [Physicochemical Properties of Nanoparticles in Relation with Toxicity](#)

Nanotransfer Printing

► [Nanoscale Printing](#)

Nanotribology

Bharat Bhushan

Nanoprobe Laboratory for Bio- & Nanotechnology and Biomimetics, The Ohio State University, Columbus, OH, USA

Synonyms

[Adhesion](#); [Friction](#); [Lubrication](#); [Wear](#)

Definition

The word “Tribology” was coined in 1966. It is derived from the Greek word *tribos*, meaning rubbing so the

literal translation would be the science of rubbing. However, it encompasses studies of surface characterization, adhesion, friction, wear, and lubrication. According to Webster dictionary, Tribology is defined as the science and technology of interacting surfaces in relative motion and of the related subjects and practices. Tribology is an interdisciplinary field. Rather complex surface interactions in a tribological interface require knowledge of various disciplines including physics, chemistry, mechanical engineering, solid mechanics, materials science, rheology, applied mathematics, and reliability. At most interfaces of technological relevance, contact occurs at numerous asperities. It is of importance to investigate a single asperity contact in the fundamental tribological studies and thus emphasizing the importance of nanotribology.

Introduction

The mechanisms and dynamics of the interactions of two contacting solids during relative motion, ranging from atomic- to microscale, need to be understood in order to develop fundamental understanding of adhesion, friction, wear, indentation, and lubrication processes. For most solid–solid interfaces of technological relevance, contact occurs at multiple asperities. Consequently, the importance of investigating single asperity contacts in studies of the fundamental micro/nanomechanical and micro/nanotribological properties of surfaces and interfaces has long been recognized. The recent emergence and proliferation of proximal probes, in particular scanning probe microscopes (the scanning tunneling microscope and the atomic force microscope), the surface force apparatus, and of computational techniques for simulating tip-surface interactions and interfacial properties, has allowed systematic investigations of interfacial problems with high resolution as well as ways and means for modifying and manipulating nanoscale structures. These advances have led to the appearance of the new field of nanotribology, which pertains to experimental and theoretical investigations of interfacial processes on scales ranging from the atomic- and molecular-scale to the microscale, occurring during adhesion, friction, scratching, wear, indentation, and thin-film lubrication at sliding surfaces [2, 4–8, 10–12, 15, 18, 22, 25]. Proximal probes have also been used for mechanical and electrical

characterization, in situ characterization of local deformation, and other nanomechanics studies.

Nanotribological and nanomechanics studies are needed to develop fundamental understanding of interfacial phenomena on a small scale and to study interfacial phenomena in micro/nanostructures used in magnetic storage devices, nanotechnology, and other applications [1–12]. Friction and wear of lightly loaded micro/nanocomponents are highly dependent on the surface interactions (few atomic layers). These structures are generally coated with molecularly thin films. Nanotribological and nanomechanics studies are also valuable in the fundamental understanding of interfacial phenomena in macrostructures, and provide a bridge between science and engineering.

The surface force apparatus (SFA), the scanning tunneling microscopes (STM), atomic force and friction force microscopes (AFM and FFM) are widely used in nanotribological and nanomechanics studies. Typical operating parameters are compared in Table 1. The SFA was developed in 1968 and is commonly employed to study both static and dynamic properties of molecularly thin films sandwiched between two molecularly smooth surfaces. The STM, developed in 1981, allows imaging of electrically conducting surfaces with atomic resolution, and has been used for imaging of clean surfaces as well as of lubricant molecules. The introduction of the AFM in 1985 provided a method for measuring ultra-small forces between a probe tip and an engineering (electrically conducting or insulating) surface, and has been used for morphological and surface roughness measurements of surfaces on the nanoscale, as well as for adhesion measurements. Subsequent modifications of the AFM led to the development of the FFM, designed for atomic-scale and microscale studies of friction. This instrument measures forces in the scanning direction. The AFM is also being used for various investigations including scratching, wear, indentation, detection of transfer of material, boundary lubrication, and fabrication and machining. Meanwhile, significant progress in understanding the fundamental nature of bonding and interactions in materials, combined with advances in computer-based modeling and simulation methods, has allowed theoretical studies of complex interfacial phenomena with high resolution in space and time. Such simulations provide insights into atomic-scale energetics, structure, dynamics, thermodynamics, transport, and rheological aspects of tribological processes.

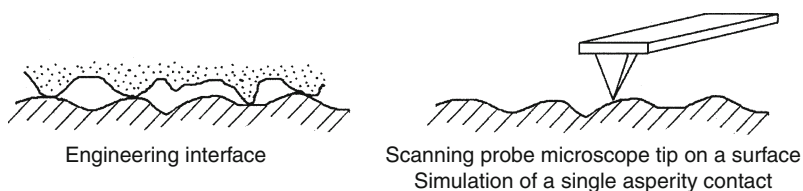
Nanotribology, Table 1 Comparison of typical operating parameters in SFA, STM and AFM/FFM used for micro/nanotribological studies

Operating parameter	SFA	STM ^a	AFM/FFM
Radius of mating surface/tip	~10 mm	5–100 nm	5–100 nm
Radius of contact area	10–40 μm	N/A	0.05–0.5 nm
Normal load	10–100 mN	N/A	<0.1–500 nN
Sliding velocity	0.001–100 $\mu\text{m/s}$	0.02–200 $\mu\text{m/s}$ (scan size ~1 nm \times 1 nm to 125 μm \times 125 μm ; scan rate <1–122 Hz)	0.02–200 $\mu\text{m/s}$ (scan size ~1 nm \times 1 nm to 125 μm \times 125 μm ; scan rate <1–122 Hz)
Sample limitations	Typically atomically-smooth, optically transparent mica; opaque ceramic, smooth surfaces can also be used	Electrically-conducting samples	None of the above

^aCan only be used for atomic-scale imaging

Nanotribology,

Fig. 1 Schematics of an engineering interface and scanning probe microscope tip in contact with an engineering interface



The nature of interactions between two surfaces brought close together, and those between two surfaces in contact as they are separated, have been studied experimentally with the surface force apparatus. This has led to a basic understanding of the normal forces between surfaces, and the way in which these are modified by the presence of a thin liquid or a polymer film. The frictional properties of such systems have been studied by moving the surfaces laterally, and such experiments have provided insights into the molecular-scale operation of lubricants such as thin liquid or polymer films. Complementary to these studies are those in which the AFM tip is used to simulate a single asperity contact with a solid or lubricated surface, Fig. 1. These experiments have demonstrated that the relationship between friction and surface roughness is not always simple or obvious. AFM studies have also revealed much about the nanoscale nature of intimate contact during wear, indentation, and lubrication.

Overview

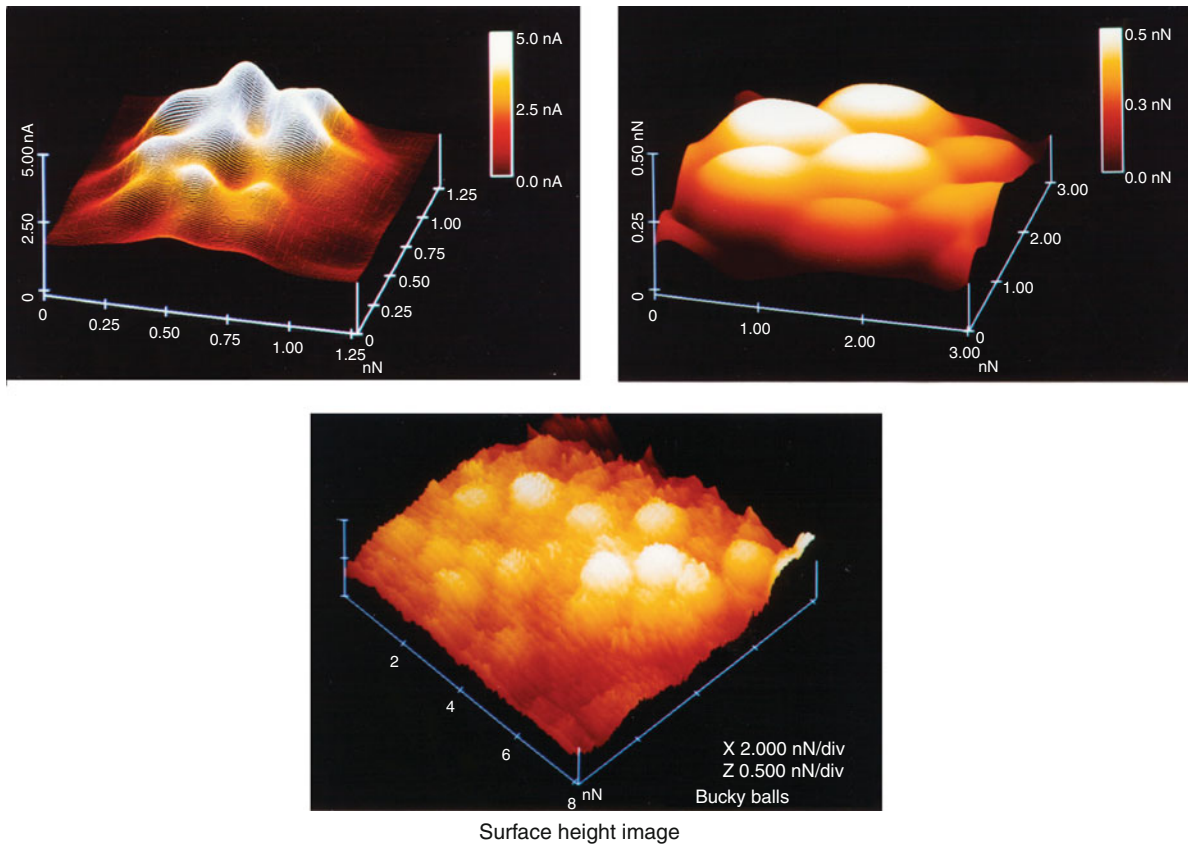
Surface Imaging, Friction, and Adhesion

Surface height imaging down to atomic resolution of electrically conducting surfaces can be carried out

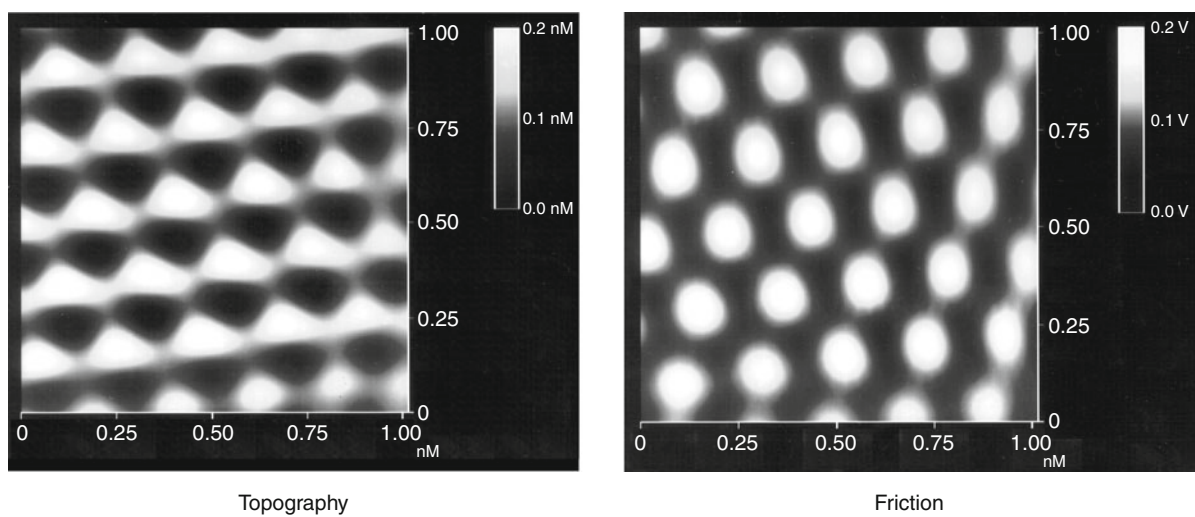
using an STM. An AFM can also be used for surface height imaging and roughness characterization down to nanoscale. Figure 2 shows a sequence of STM images at various scan sizes of solvent-deposited C60 film on a 200-nm thick gold-coated freshly cleaved mica [13]. The film consists of clusters of C60 molecules with 8 nm in diameter. The C60 molecules within a cluster appear to pack in a hexagonal array with a spacing of about 1 nm; however, they do not follow any long range order. The measured cage diameter of the C60 molecule is about 0.7 nm, very close to the projected diameter of 0.71 nm.

In an AFM measurement during surface imaging, the tip comes in intimate contact with the sample surface and leads to surface deformation with finite tip-sample contact area (typically few atoms). The finite size of contact area prevents the imaging of individual point defects, and only the periodicity of the atomic lattice can be imaged. Figure 3 shows the topography image of freshly cleaved surface of highly oriented pyrolytic graphite (HOPG) [24]. The periodicity of the graphite is clearly observed.

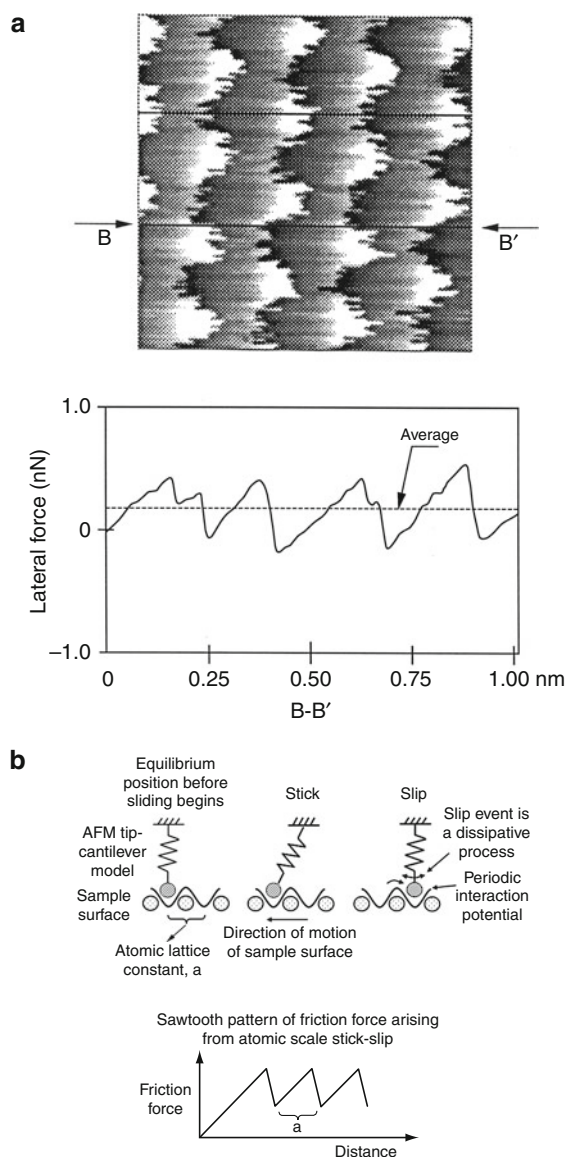
To study friction mechanisms on an atomic scale, a freshly cleaved HOPG has been studied by Mate et al. [21] and Ruan and Bhushan [24]. Figure 4a shows the atomic-scale friction force map



Nanotribology, Fig. 2 STM images of solvent deposited C_{60} film on a gold-coated freshly-cleaved mica at various scan sizes



Nanotribology, Fig. 3 Gray-scale plots of surface topography and friction force maps (2D spectrum filtered), measured simultaneously, of a $1 \text{ nm} \times 1 \text{ nm}$ area of freshly cleaved HOPG, showing the atomic-scale variation of topography and friction



Nanotribology, Fig. 4 (a) Gray scale plot of friction force map (raw data) of a $1 \text{ nm}^2 \times 1 \text{ nm}^2$ area of freshly cleaved HOPG, showing atomic-scale variation of friction force. High points are shown by lighter color. Also shown is line plot of friction force profile along the line indicated by arrows. The normal load was 25 nN and the cantilever normal stiffness was 0.4 N/m [23], and (b) schematic of a model for a tip atom sliding on an atomically flat periodic surface. The schematic shows the tip jumping from one potential minimum to another, resulting in stick-slip behavior

(raw data), and Fig. 3 shows the friction force maps (after 2D spectrum filtering with high frequency noise truncated) [24]. Figure 4a also shows line plot of friction force profiles along some crystallographic

direction. The actual shape of the friction profile depends upon the spatial location of axis of tip motion. Note that a portion of atomic-scale lateral force is conservative. Mate et al. [21] and Ruan and Bhushan [24] reported that the average friction force linearly increased with normal load and was reversible with load. Friction profiles were similar during sliding of the tip in either direction.

During scanning, the tip moves discontinuously over the sample surface and jumps with discrete steps from one potential minimum (well) to the next. This leads to a saw-tooth-like pattern for the lateral motion (force) with a periodicity of the lattice constant. This motion is called stick-slip movement of the tip [5, 8, 12, 21, 23]. The observed friction force includes two components – conservative and periodic, and nonconservative and constant. If the relative motion of the sample and tip were simply that of two rigid collections of atoms, the effective force would be a conservative force oscillating about zero. Slow reversible elastic deformation would also contribute to conservative force. The origin of the nonconservative direction-dependent force component would be phonon generation, viscous dissipation, or plastic deformation.

The stick-slip on the atomic scale discussed above is the result of the energy barrier required to be overcome for jumping over the atomic corrugations on the sample surface. It corresponds to the energy required for the jump of the tip from a stable equilibrium position on the surface into a neighboring position. The perfect atomic regularity of the surface guarantees the periodicity of the lateral force signal, independent of actual atomic structure of tip apex. Few atoms (based on the magnitude of the friction force, less than 10), on a tip sliding over an array of atoms on the sample, are expected to go through the stick-slip. For simplicity, Fig. 4b shows a simplified model for one atom on a tip with a one-dimensional spring mass system. As the sample surface slides against the AFM tip, the tip remains “stuck” initially until it can overcome the energy (potential) barrier, which is illustrated by a sinusoidal interaction potential as experienced by the tip. After some motion, there is enough energy stored in the spring which leads to “slip” into the neighboring stable equilibrium position. During the slip and before attaining stable equilibrium, stored energy is converted into vibrational energy of the surface atoms in the range of 1,013 Hz (phonon

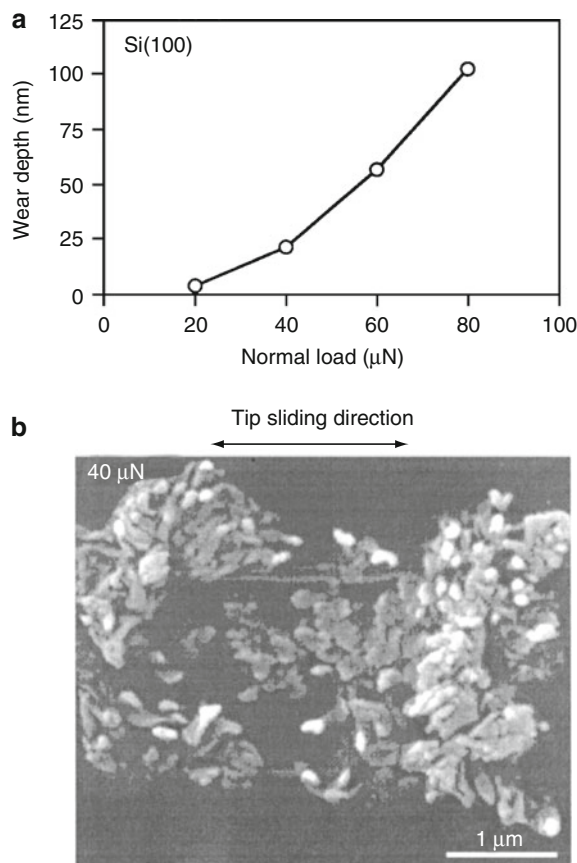
generation) and decays within the range of 10–11 s into heat. (A wave of atoms vibrating in concert is termed a phonon.) The stick–slip phenomenon, resulting from irreversible atomic jumps, can be theoretically modeled with classical mechanical models [27, 28]. The Tomanek-Zhong-Thomas model [27] is the starting point for determining friction force during atomic scale stick–slip. The AFM model describes the total potential as the sum of the potential acting on the tip due to interaction with the sample and the elastic energy stored in the cantilever. Thermally activated stick–slip behavior can explain the velocity effects on friction.

Wear

By scanning the sample in two dimensions with the AFM, wear scars are generated on the surface. Figure 5a shows the effect of normal load on wear depth on Si(100). Note that wear depth is very small below 20 μN of normal load [29]. A normal load of 20 μN corresponds to contact stresses comparable to the hardness of silicon. Primarily, elastic deformation at loads below 20 μN is responsible for low wear. Uniform material removal at the bottom of the wear mark has been reported. An AFM image of the wear mark shows small debris at the edges, probably swiped during AFM scanning. This indicates that the debris is loose (not sticky) and can be removed during the AFM scanning.

Next, the mechanisms of material removal were examined on a microscale in AFM wear experiments [29]. Figure 5b shows a secondary electron image of the wear mark and associated wear particles. The specimen used for the scanning electron microscope (SEM) was not scanned with the AFM after initial wear, in order to retain wear debris in the wear region. Wear debris is clearly observed. In the SEM micrographs, the wear debris appears agglomerated because of high surface energy of the fine particles. Particles appear to be a mixture of rounded and so-called cutting type (feather-like or ribbon-like material). Zhao and Bhushan [29] reported an increase in the number and size of cutting type particles with the normal load. The presence of cutting type particles indicates that the material is removed primarily by plastic deformation.

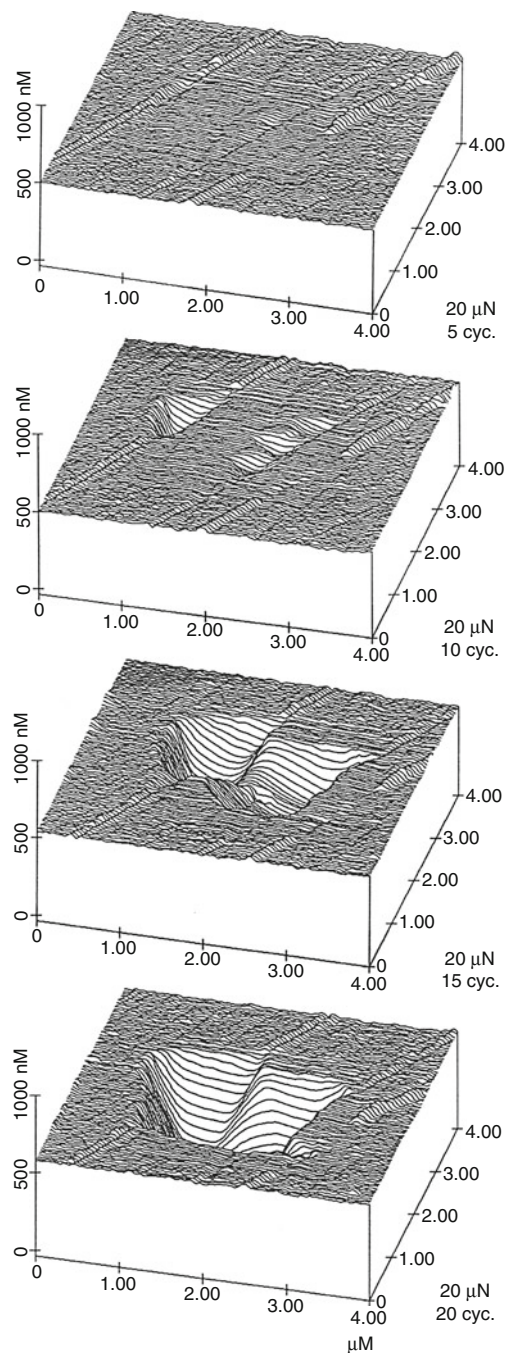
To better understand the material removal mechanisms, Zhao and Bhushan [29] used transmission electron microscopy (TEM). The existence of



Nanotribology, Fig. 5 (a) Wear depth as a function of normal load for Si(100) after one cycle [29], and (b) secondary electron image of wear mark and debris for Si(100) produced at a normal load of 40 μN and one scan cycle

dislocation arrays observed by them confirms that material removal occurs by plastic deformation. It is concluded that the material on microscale at high loads is removed by plastic deformation with a small contribution from elastic fracture.

To understand wear mechanisms, evolution of wear can be studied using AFM. Figure 6 shows evolution of wear marks of a DLC-coated disk sample. The data illustrate how the microwear profile for a load of 20 μN develops as a function of the number of scanning cycles [14]. Wear is not uniform, but is initiated at the nanoscratches. Surface defects (with high surface energy) present at nanoscratches act as initiation sites for wear. Coating deposition also may not be uniform on and near nanoscratches which may lead to coating delamination. Thus, scratch-free surfaces will be relatively resistant to wear.



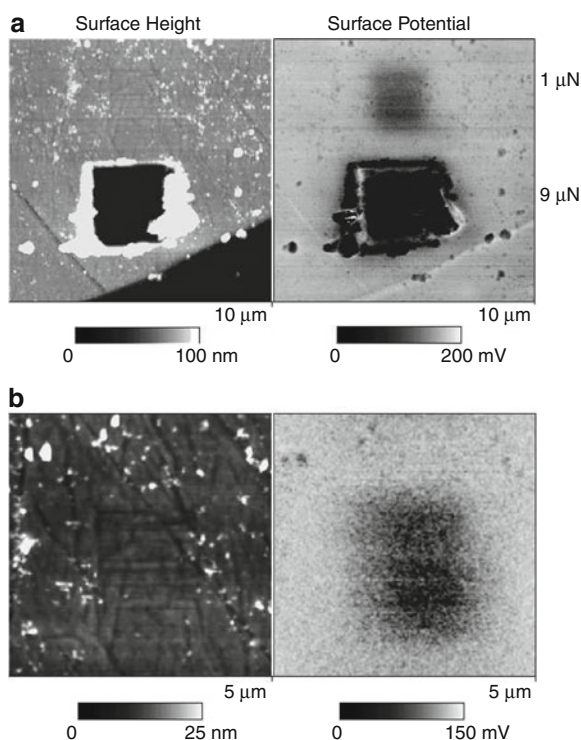
Nanotribology, Fig. 6 Surface plots of diamond-like carbon-coated thin-film disk showing the worn region; the normal load and number of test cycles are indicated [14]

Wear precursors (precursors to measurable wear) can be studied by making surface potential measurements [17]. The contact potential difference, or simply the surface potential between two surfaces depends

on a variety of parameters such as electronic work function, adsorption, and oxide layers. The surface potential map of an interface gives a measure of changes in the work function which is sensitive to both physical and chemical conditions of the surfaces including structural and chemical changes. Before material is actually removed in a wear process, the surface experiences stresses that result in surface and subsurface changes of structure and/or chemistry. These can cause changes in the measured potential of a surface. An AFM tip allows mapping of surface potential with nanoscale resolution. Surface height and change in surface potential maps of a polished single-crystal aluminum (100) sample, abraded using a diamond tip at loads of 1 and 9 μN , are shown in Fig. 7a (Note that the sign of the change in surface potential is reversed here from that in DeVecchio and Bhushan [17]). It is evident that both abraded regions show a large potential contrast ($\sim 0.17\text{ V}$), with respect to the non-abraded area. The black region in the lower right-hand part of the topography scan shows a step that was created during the polishing phase. There is no potential contrast between the high region and the low region of the sample, indicating that the technique is independent of surface height. Figure 7b shows a close up scan of the upper (low load) wear region in Fig. 7a. Notice that while there is no detectable change in the surface topography, there is nonetheless, a large change in the potential of the surface in the worn region. Indeed, the wear mark of Fig. 7b might not be visible at all in the topography map were it not for the noted absence of wear debris generated nearby and then swept off during the low load scan. Thus, even in the case of zero wear (no measurable deformation of the surface using AFM), there can be a significant change in the surface potential inside the wear mark which is useful for the study of wear precursors. It is believed that the removal of the thin contaminant layer including the natural oxide layer gives rise to the initial change in surface potential. The structural changes, which precede generation of wear debris and/or measurable wear scars, occur under ultra-low loads in the top few nanometers of the sample, and are primarily responsible for the subsequent changes in surface potential.

Boundary Lubrication

The classical approach to lubrication uses freely supported multimolecular layers of liquid lubricants



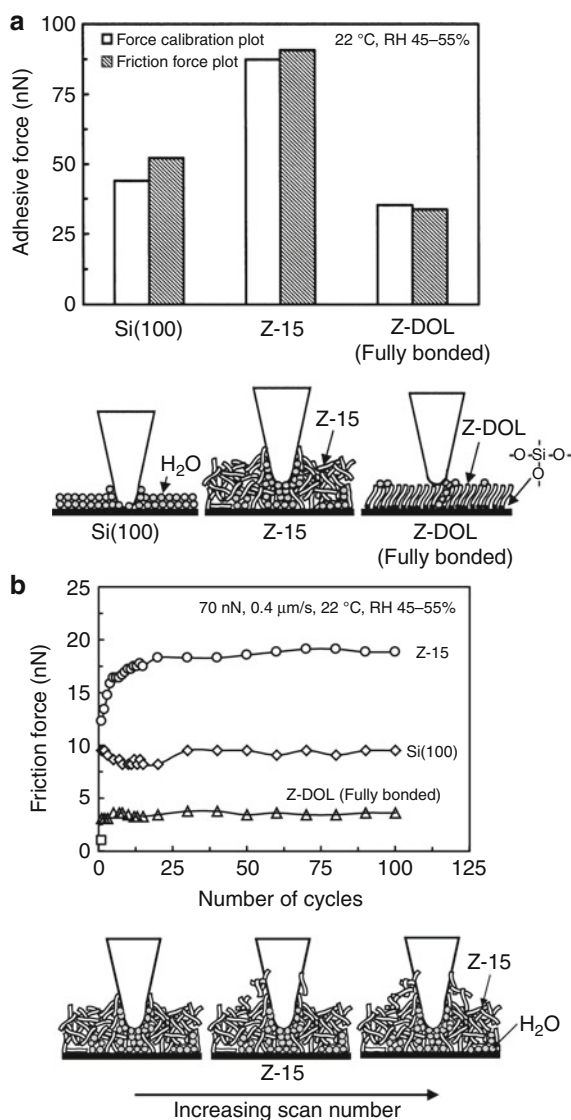
Nanotribology, Fig. 7 (a) Surface height and change in surface potential maps of wear regions generated at 1 μN (top) and 9 μN (bottom) on a single crystal aluminum sample showing bright contrast in the surface potential map on the worn regions. (b) Close up of upper (low load) wear region [17]

[1, 5, 8, 16]. The liquid lubricants are sometimes chemically bonded to improve their wear resistance [1, 5, 8]. Partially chemically-bonded, molecularly-thick perfluoropolyether (PFPE) films are used for lubrication of magnetic storage media because of their thermal stability and extremely low vapor pressure [1]. Chemically-bonded lubricants are considered as potential candidate lubricants for MEMS/NEMS. Molecularly-thick PFPEs are well suited because of the following properties: low surface tension and low contact angle which allow easy spreading on surfaces and provide hydrophobic properties; chemical and thermal stability which minimize degradation under use; low vapor pressure which provides low out-gassing; high adhesion to substrate via organic functional bonds; and good lubricity which reduces contact surface wear.

For boundary lubrication studies, friction, adhesion, and durability experiments have been performed on

virgin Si (100) surfaces and silicon surfaces lubricated with various PFPE lubricants [19, 26]. Results of two of the PFPE lubricants will be presented here, which are -Z-15 (with $-\text{CF}_3$ nonpolar end groups), $\text{CF}_3-\text{O}-(\text{CF}_2-\text{CF}_2-\text{O})_m-(\text{CF}_2-\text{O})_n-\text{CF}_3$ ($m/n \sim 2/3$) and Z-DOL (with $-\text{OH}$ polar end groups), $\text{HO}-\text{CH}_2-\text{CF}_2-\text{O}-(\text{CF}_2-\text{CF}_2-\text{O})_m-(\text{CF}_2-\text{O})_n-\text{CF}_2-\text{CH}_2-\text{OH}$ ($m/n \sim 2/3$). Z-DOL film was thermally bonded at 150°C for 30 min and unbonded fraction was removed by a solvent (BW) [1]. The thicknesses of Z-15 and Z-DOL films were 2.8 and 2.3 nm, respectively. Lubricant chain diameters of these molecules are about 0.6 nm and molecularly thick film generally lie flat on surfaces with high coverage.

The adhesive forces of Si(100), Z-15, and Z-DOL (BW) measured by force calibration plot and friction force versus normal load plot are summarized in Fig. 8a [20]. The data obtained by these two methods are in good agreement. Figure 8(a) shows that the presence of mobile Z-15 lubricant film increases the adhesive force as compared to that of Si(100) by meniscus formation. Whereas, the presence of solid-like phase of the Z-DOL (BW) film reduces the adhesive force as compared to that of Si(100) because of the absence of mobile liquid. The schematic (bottom) in Fig. 8a shows relative size and sources of meniscus. It is well known that the native oxide layer (SiO_2) on the top of Si(100) wafer exhibits hydrophilic properties, and some water molecules can be adsorbed on this surface. The condensed water will form meniscus as the tip approaches the sample surface. The larger adhesive force in Z-15 is not only caused by the Z-15 meniscus alone, the non-polarized Z-15 liquid does not have good wettability and strong bonding with Si(100). Consequently, in the ambient environment, the condensed water molecules from the environment will permeate through the liquid Z-15 lubricant film and compete with the lubricant molecules present on the substrate. The interaction of the liquid lubricant with the substrate is weakened, and a boundary layer of the liquid lubricant forms puddles. This dewetting allows water molecules to be adsorbed on the Si(100) surface along with Z-15 molecules. And both of them can form meniscus while the tip approaches to the surface. Thus the dewetting of liquid Z-15 film results in higher adhesive force and poorer lubrication performance. In addition, as the Z-15 film is pretty soft as compared to the solid Si(100) surface, and penetration of the tip in the film occurs while pushing tip down. This results in the large area of the tip



Nanotribology, Fig. 8 (a) Summary of the adhesive forces of Si(100) and Z-15 and Z-DOL (BW) films measured by force calibration plots and friction force versus normal load plots in ambient air. The schematic (*bottom*) showing the effect of meniscus, formed between AFM tip and the surface sample, on the adhesive and friction forces [19], and (b) Friction force versus number of sliding cycles for Si(100) and Z-15 and Z-DOL (BW) films at 70 nN, 0.4 $\mu\text{m/s}$, and in ambient air. Schematic (*bottom*) shows that some liquid Z-15 molecules can be attached onto the tip. The molecular interaction between the attached molecules onto the tip with the Z-15 molecules in the film results in an increase of the friction force with multi scanning [20]

wetted by the liquid to form the meniscus at the tip-liquid (mixture of Z-15 and water) interface. It should also be noted that Z-15 has a higher viscosity compared to water, therefore Z-15 film provides higher resistance to motion

and coefficient of friction. In the case of Z-DOL (BW) film, both of the active groups of Z-DOL molecules are mostly bonded on Si(100) substrate, thus the Z-DOL (BW) film has low free surface energy and cannot be displaced readily by water molecules or readily adsorb water molecules. Thus, the use of Z-DOL (BW) can reduce the adhesive force.

To study the durability of lubricant films at nanoscale, the friction of Si(100), Z-15, and Z-DOL (BW) as a function of the number of scanning cycles are shown in Fig. 8b [19]. As observed earlier for adhesive force, friction force of Z-15 is higher than that of Si(100) with the lowest values for Z-DOL(BW). During cycling, friction force of Si(100) shows a slight decrease during initial few cycles then remains constant. This is related to the removal of the native oxide layer. In the case of Z-15 film, the friction force shows an increase during the initial few cycles and then approaches to higher and stable values. This is believed to be caused by the attachment of the Z-15 molecules onto the tip. After several scans, the molecular interaction reaches an equilibrium and after that the friction force remains constant. In the case of Z-DOL (BW) film, the friction force starts out to be low and remains low during the entire test for 100 cycles. It suggests that Z-DOL (BW) molecules do not get attached or displaced as readily as Z-15.

Closure

For most solid-solid interfaces of technological relevance, contact occurs at multiple asperities. A sharp AFM/FFM tip sliding on a surface simulates just one such contact. However, asperities come in all shapes and sizes. AFM/FFM are used to study various tribological phenomena, which include surface roughness, adhesion, friction, scratching, wear, indentation, detection of material transfer, and boundary lubrication. Measurement of atomic-scale friction of a freshly-cleaved highly-oriented pyrolytic graphite exhibits the same periodicity as that of the corresponding topography. Relevant friction mechanism is atomic-scale stick-slip.

Wear rate on the microscale for single-crystal silicon is negligible below 20 μN and is much higher and remains approximately constant at higher loads. Elastic deformation at low loads is responsible for negligible wear. Most of the wear debris is loose. SEM and TEM studies of the wear region suggest that the material on the microscale is removed by plastic deformation with a small contribution from elastic fracture; this observation corroborates with the

scratch data. Evolution of wear has also been studied using AFM. Wear is found to be initiated at nanoscratches. For a sliding interface requiring near-zero friction and wear, contact stresses should be below the hardness of the softer material to minimize plastic deformation and surfaces should be free of nanoscratches. Further, wear precursors can be detected at early stages of wear by using surface potential measurements. It is found that even in the case of zero wear (no measurable deformation of the surface using AFM), there can be a significant change in the surface potential inside the wear mark which is useful for study of wear precursors.

Boundary lubrication studies and measurement of lubricant-film thickness with a lateral resolution on a nanoscale can be conducted using AFM. For chemically bonded lubricant films, the adsorption of water, the formation of meniscus and its change during sliding, and surface properties play an important role on the adhesion, friction, and durability of these films.

Investigations of adhesion, friction, wear, scratching and indentation on nanoscales using the AFM can provide insights into failure mechanisms of materials. Coefficients of friction, wear rates and mechanical properties such as hardness have been found to be different on the nanoscale than on the macroscale; generally, coefficients of friction and wear rates on micro- and nanoscales are smaller, whereas hardness is greater. Therefore, micro/nanotribological studies may help define the regimes for ultra-low friction and near zero wear. These studies also provide insight to atomic origins of adhesion, friction, wear, and lubrication mechanisms.

Cross-References

► Atomic Force Microscopy

References

- Bhushan, B.: Tribology and mechanics of magnetic storage devices, 2nd edn. Springer, New York (1996)
- Bhushan, B.: Micro/nanotribology and its applications, vol. E330. Kluwer, Dordrecht (1997)
- Bhushan, B.: Tribology issues and opportunities in MEMS. Kluwer, Dordrecht (1998)
- Bhushan, B.: Handbook of micro/nanotribology, 2nd edn. CRC Press, Boca Raton (1999)
- Bhushan, B.: Principles and applications of tribology. Wiley, New York (1999)
- Bhushan, B.: Modern tribology handbook. Principles of tribology, vol. 1. CRC Press, Boca Raton (2001)
- Bhushan, B.: Fundamentals of tribology and bridging the Gap Between the Macro- and Micro/Nanoscales. In: NATO science series II, vol. 10. Kluwer, Dordrecht (2001)
- Bhushan, B.: Introduction to tribology. Wiley, New York (2002)
- Bhushan, B.: Adhesion and stiction: mechanisms, measurement techniques, and methods for reduction. *J. Vac. Sci. Technol. B* **21**, 2262–2296 (2003). invited
- Bhushan, B.: Nanotribology, nanomechanics and nanomaterials characterization. *Philos. Tr. R. Soc. A* **366**, 1351–1381 (2008)
- Bhushan, B.: Springer handbook of nanotechnology, 3rd edn. Springer, Heidelberg (2010)
- Bhushan, B.: Nanotribology and nanomechanics I – measurement techniques and nanomechanics, II – nanotribology, biomimetics, and industrial applications, 3rd edn. Springer, Heidelberg (2011)
- Bhushan, B., Ruan, J., Gupta, B.K.: A scanning tunnelling microscopy study of fullerene films. *J. Phys. D Appl. Phys.* **26**, 1319–1322 (1993)
- Bhushan, B., Koinkar, V.N., Ruan, J.: Microtribology of magnetic media. *P. I. Mech. Eng. J-J. Eng. Tribol.* **208**, 17–29 (1994)
- Bhushan, B., Israelachvili, J.N., Landman, U.: Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* **374**, 607–616 (1995)
- Bowden, F.P., Tabor, D.: The friction and lubrication of solids, part I. Clarendon, Oxford (1950)
- DeVecchio, D., Bhushan, B.: Use of a nanoscale Kelvin probe for detecting wear precursors. *Rev. Sci. Instrum.* **69**, 3618–3624 (1998)
- Guntherodt, H.J., Anselmetti, D., Meyer, E.: Forces in scanning probe methods, vol. E286. Kluwer, Dordrecht (1995)
- Liu, H., Bhushan, B.: Adhesion and friction studies of microelectromechanical systems/nanoelectromechanical systems materials using a novel microtriboapparatus. *J. Vac. Sci. Technol. A* **21**, 1528–1538 (2003)
- Liu, H., Bhushan, B.: Nanotribological characterization of molecularly-thick lubricant films for applications to MEMS/NEMS by AFM. *Ultramicroscopy* **97**, 321–340 (2003)
- Mate, C.M., McClelland, G.M., Erlandsson, R., Chiang, S.: Atomic-scale friction of a tungsten tip on a graphite surface. *Phys. Rev. Lett.* **59**, 1942–1945 (1987)
- Persson, B.N.J., Tosatti, E.: Physics of sliding friction, vol. E311. Kluwer, Dordrecht (1996)
- Ruan, J., Bhushan, B.: Atomic-scale and microscale friction of graphite and diamond using friction force microscopy. *J. Appl. Phys.* **76**, 5022–5035 (1994)
- Ruan, J., Bhushan, B.: Frictional behavior of highly oriented pyrolytic graphite. *J. Appl. Phys.* **76**, 8117–8120 (1994)
- Singer, I.L., Pollock, H.M.: Fundamentals of friction: macroscopic and microscopic processes, vol. E220. Kluwer, Dordrecht (1992)
- Tao, Z., Bhushan, B.: Bonding, degradation, and environmental effects on novel perfluoropolyether lubricants. *Wear* **259**, 1352–1361 (2005)

27. Tomanek, D., Zhong, W., Thomas, H.: Calculation of an atomically modulated friction force in atomic force microscopy. *Europhys. Lett.* **15**, 887–892 (1991)
28. Tomlinson, G.A.: A molecular theory of friction. *Phil. Mag. Ser. 7*, 905–939 (1929)
29. Zhao, X., Bhushan, B.: Material removal mechanism of single-crystal silicon on nanoscale and at ultralow loads. *Wear* **223**, 66–78 (1998)

Nano-Tribology Applications in Microprojector Technology

Satish C. Chaparala and Vikram Bhatia
Science and Technology, Corning Incorporated
SP-PR-02-1, Corning, NY, USA

Definition

Microprojectors are being developed to project a larger image from a smaller display in handheld electronic devices such as cell phone, handheld PDA, etc. The technology associated with microprojectors has several applications of nano-tribology and those are described in this entry.

Introduction to Microprojector Technology

The capacity of high-speed wireless networks has enabled carriers to provide multimedia services to consumers such as web-browsing, live television and gaming, etc. At the same time, the electronics industry is driving toward miniaturization which in turn led to the smaller size of displays on the electronic devices. The smaller display limits the full utilization and enjoyment of these services. Microprojectors have been proposed to address this issue. These devices can either be as an accessory to any electronic device or an embedded component within a mobile device. These are expected to provide images of the size of A4 paper and larger in indoor lighting conditions. The challenge is to develop the correct microprojector configuration to balance size, cost, performance, and efficiency [1]. For instance, currently, the thickness of the projector could be in the range of 7–14 mm that results in overall projector volume between 10 and 20 cm³. Brightness is obviously a very important attribute of a

microprojector. While the current range of products offers 5–10 lm of brightness, 50 lm is required to provide enough brightness in all lighting conditions. Focus-free operation is another desirable attribute and this is possible with the projectors that have lasers as light sources. Resolution is expected to continue to grow as product technology improves. The wide-screen format is typically acceptable for viewing video content. Similar to brightness, contrast is also an important attribute that makes the difference between washed-out images and crisp, dramatic-looking images [2]. These trade-offs are primarily driven by the choices of light source and imaging mechanism.

The microprojector technology has two major components in the optical engine: (1) Light source and (2) Image producing technology. Optical engines are modules that combine the red-green-blue (RGB) light sources and an imaging element or elements. Optical engines accept an electronic input video signal and project an optical image. Optical engines do not contain power sources or any other functional elements such as wireless communication chips. Three primary colors, namely, red, blue, and green are required to create full color images. The light sources in the projection technology would be semiconductor devices that emit these colors. These devices could be either light emitting diodes (LEDs) or Lasers. Like a normal diode, when an LED is forward biased or in other words, switched on, electron combines with a hole, falls into a lower energy level, thereby releasing the energy in the form of photon. The color of the emitted light depends on the energy gap of the semiconductor. Red, blue, and green LEDs are commercially available. LEDs have several advantages that include low energy consumption, longer lifetime, smaller size, faster switching, and good reliability, typically lasting a few years. However, LEDs have higher etendue. Etendue is a property of an optical system that describes how “diverged” the light is in area and angle. Because of high etendue of LEDs, large optical elements (e.g., lenses, etc.) have to be used in the package to collect all the light that is emitted from LED, which in turn increases the size of the projector. If the projector volume is designed to be small using smaller optical elements, some of the optical power (larger angles) will be lost, which in turn decreases the conversion efficiency of the package. Lasers can be used to overcome this problem. A laser is a device that

emits light through a process of optical amplification based on the stimulated emission of photons. Laser light is spatially and temporally coherent. Red and blue lasers are commercially available for applications to the optical read/write products such as DVD players, whereas green has been a challenge to produce through currently available processes and materials. A green laser is often referred to as the “keystone” component to enable laser-based mobile projection. Extensive research has been conducted to develop native green lasers using either III–V or II–VI compound semiconductors. Recent technological advancements have led to creation of native green laser light by Sumitomo, Osram, and others. For instance, Sumitomo succeeded in producing 2.4 mW of optical power at 520 nm at room temperature continuous wave lasing of InGaN-based green LEDs on semipolar GaN substrates [3]. Native green lasers are easier to control, and could potentially demonstrate greater temperature stability, a smaller form factor and higher modulation capability at several 100 MHz. However, there has not been sufficient reliability data of these devices under operating conditions. While the native green lasers are under development, a synthetic green laser has already been developed based on frequency doubling of 1,060 nm IR light using second harmonic generation crystal by various companies such as Corning [1, 4–6].

The second critical component of a microprojector is image producing technology. There are three main image producing technologies: (1) Scanners, (2) Digital light processing (DLP), and (3) Liquid crystal on silicon (LCOS). Scanners are based on two-dimensional pixel-by-pixel scanning, typically using a single MEMS mirror, and uses modulated laser beams as the light source. The image remains in focus at any projection distance and the projector consumes relatively low power [2, 7]. DLP uses millions of tiny mirrors, one for each pixel, to reflect light. This technology involves modulating the image by tilting the mirrors either into or away from the lens path. It is a “reflective” technology and can use either lasers or LEDs as light sources. DLP projection technology using LED as light source is already commercialized. LCOS combines the idea of a DLP which is a reflective technology, but uses liquid crystals instead of individual mirrors. In LCOS, liquid crystals are applied to a reflective mirror substrate. As the liquid crystals are modulated based on the incoming video signal, the light is either reflected from the mirror

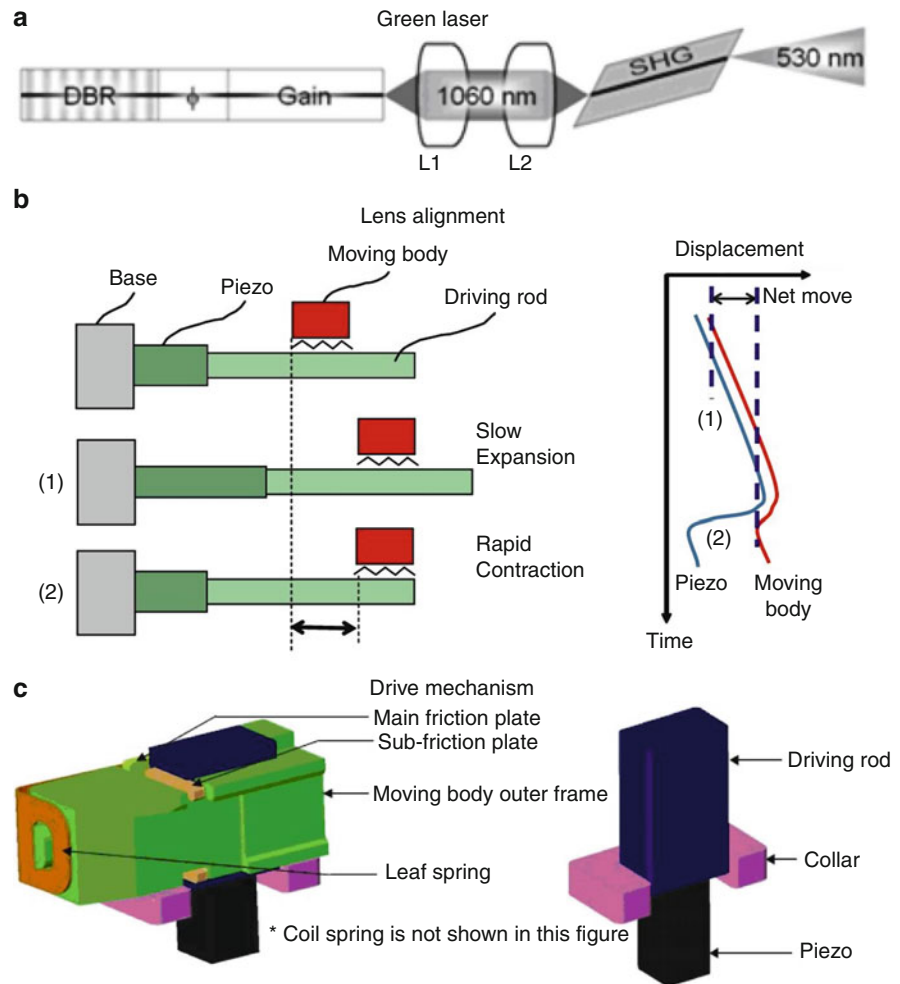
below, or blocked. This modulates the light from either laser or LED sources and creates the image. Any feasible combination of these three technologies with either lasers or LEDs could project an image. Finally, all these components can be either wrapped in a case (skin) and sold as a companion projector or embedded in a handset. Projector contains an optical engine, power source, and “industrial design or skin.”

Synthetic Green Laser Package

Until now, a detailed introduction to various projection technologies and components involved in the microprojector product has been provided. The next focus of this entry would be on the keystone component of a laser-based microprojector, which is a green laser optoelectronic package developed by Corning Incorporated and specifically, the role of nanotribology in its development. Several key tribology design studies will be presented. A frequency doubling technique is used to produce a compact green laser. A distributed bragg reflector (DBR) laser diode generated 1060 nm wavelength light is passed through second harmonic generating (SHG) crystal resulting in 530 nm green light. Corning’s green laser architecture is shown in the Fig. 1 [1, 4, 8]. As can be seen from the figure, there are two lenses in this configuration. The first lens that is in front of the IR laser diode is used to collimate the light beam that emits from the laser diode. The second lens converges the collimated beam onto the waveguide in SHG. Exposure to the operating environment, which can be at a temperature as high as 60°C and at 75% relative humidity (RH), can cause lens misalignment and affect optical output. To correct for this, the green laser module uses an adaptive optics component with a drive mechanism to align the optics and maintain constant power output with time and temperature. Lens alignment along the x- and y-axes is made by using adaptive optics, represented by L_1 and L_2 in the figure below. This adaptive optics component is referred to as Smooth Impact Drive Mechanism (SIDM) developed by Konica Minolta. Schematics of the operation and construction of this drive mechanism are shown in Fig. 1. The main components of the drive mechanism are the piezoelectric element, driving rod, moving body, and the friction plate. During operation, a rectangular waveform with a given

Nano-Tribology Applications in Microprojector Technology,

Fig. 1 Schematic of (a) the green laser architecture, (b) the operation, and (c) the construction of the drive mechanism used in the green laser



voltage, frequency, and number of cycles is applied to the piezo element. The rectangular waveform is then converted to a sawtooth motion by the mechanical transfer function of the piezo. On the gentle upward slope, the rod carries the moving body due to friction (“stick” operation), while on the rapid falling slope, the rod “slips” due to inertia. The net displacement is the minimum movement of the drive mechanism. This is repeated until the desired amount of displacement is achieved [4].

Carbon fiber reinforced plastic (CFRP) is used as a component for a wide variety of engineering applications, such as for aircraft structures, automobile, and solar cells, among many others [8–12]. In the drive mechanism, CFRP is the material used as the driving rod. For the friction plate, stainless steel is used. For

the moving body, Zn alloy is used as the die-cast part, which is commonly used in automotive applications. Adaptive optics sliding components in microprojectors should be stable at an operating environment up to 60°C and 75% relative humidity. To improve tribological performance of the sliding components, perfluoropolyether (PFPE) lubricants are applied to the surfaces of the components. The ideal lubricant should be molecularly thick to protect the surface from wear, easily applicable, able to chemically bond to the surface, and insensitive to environment. PFPE lubricants are known to be most desirable as they have low surface tension, high contact angle, and high adhesion to the substrate allowing easy application and spreading onto the surface as well as providing hydrophobicity [13–19]. Their chemical and thermal

stability and low vapor pressure provide low degradation and low outgassing. PFPEs have been extensively investigated, especially in the magnetic disk drive industry. It has been shown that they reduce friction and wear, resulting in lower disk drive failure. It is, therefore, very critical to conduct a comprehensive tribology study to understand the friction mechanisms and effect of temperature, humidity, plasma, etc., on the friction.

Applications of Nano-Tribology

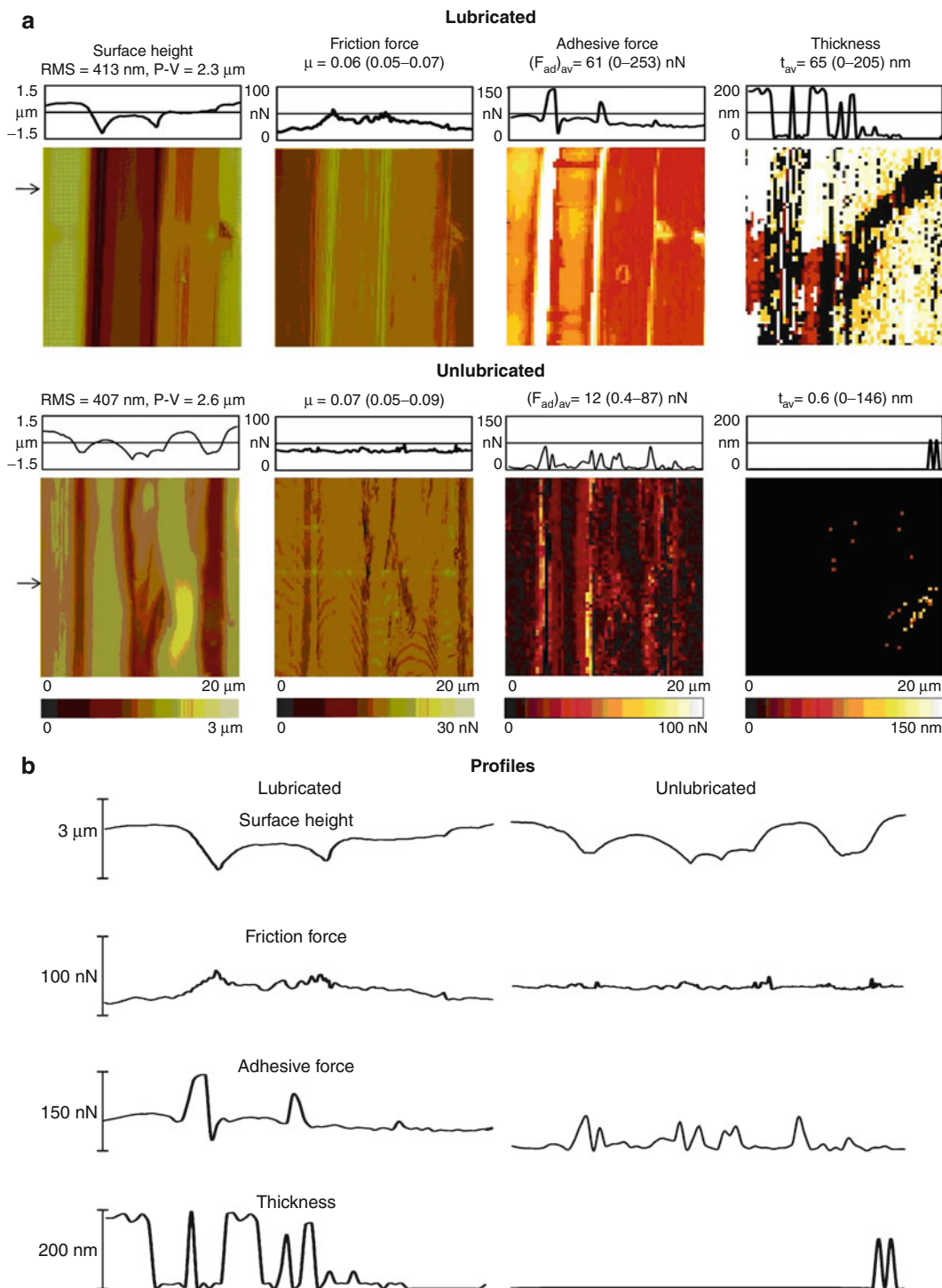
The following section presents several key findings from various nano-tribological studies made on the lubricant used in the adaptive optics component. A methodology to measure lubricant film thickness with a nanoscale resolution was developed. Atomic force microscopy (AFM) allows the simulation of a single asperity contact. The morphology measured by AFM is shown in Fig. 2. On the height images of the lubricated and unlubricated samples, the bright long columns are for the individual carbon fibers, and the dark ones for epoxy. The average surface roughness (RMS) of the lubricated CFRP is about 328 nm, and that of the unlubricated is 406 nm. Even though both lubricated and unlubricated CFRPs have high average surface roughness, the lubricated sample surface has lower roughness than that of the unlubricated surface. The roughness difference between the lubricated and the unlubricated samples can be attributed to lubricant thickness.

The degree of bonding was measured by measuring the film thickness before and after washing off the lubricant film. It was found that the lubricant was unbonded to the sample surfaces. Friction force, adhesive force, and lubricant thickness distribution correlate with surface morphology. Lubricant resides in the recessed epoxy regions which act as a lubricant reservoir. It is expected that lubricant will wick up during the operation and get transferred. Coefficient of friction on lubricated samples is lower than that on the unlubricated samples. The epoxy has a higher friction force than carbon fiber because of the presence of the lubricant. Wear depth of epoxy is found to be higher than on the carbon fiber. Effects of composite components, scan direction, length scale, temperature, and humidity were studied. Scanning in the transverse direction includes sliding on both the carbon fiber

and the epoxy, and the friction data is close to that of epoxy. Regarding length scale effects, the coefficient of friction on larger length scales is higher due to larger contact area. Figure 3 shows the effect of humidity and temperature. Humidity does not affect the coefficient of friction due to the condensed water layer from the environment which is expected to be thin compared to lubricant thickness. However, increasing temperature leads to viscosity drop of the lubricant, resulting in the decrease in the coefficient of friction.

Thermal, UV, and oxygen and argon plasma treatments in order to allow chemical bonding of lubricant were studied. It was found that the effect of thermal treatment on the lubricant chemical bonding was not obvious. However, the thermal treatment reduced adhesive force and coefficient of friction because of solvent evaporation. For UV treatment, it was found that coefficient of friction, adhesive force, and lubricant thickness are unchanged. UV treatment does not affect chemical bonding of lubricant. For plasma treatment, it is observed that lubricant on the lubricated and subsequently plasma-treated samples is bonded to the surfaces. The plasma treatment also improves coefficient of friction and adhesive force. Lubricant film thickness was reduced by the plasma treatment. However, no difference is found between oxygen and argon plasma treatments. There is no difference between various power levels in plasma treatments either. It is found that chemical bonding of lubricant to the sample surfaces due to plasma treatment helps to protect the carbon fiber areas of the CFRP rods from wear but not epoxy areas (Fig. 4).

Further studies revealed that the thermal and UV treatments do not improve wear for the CFRP rods. There is no measurable change in friction, adhesive force, and lubricant thickness after laser exposure to the samples. Therefore, it is believed that lubricant degradation by green and IR laser exposure does not occur, or is negligible. Finally, wettability by contact angle measurement for unlubricated and lubricated CFRP rods was studied to simulate the effect of atmospheric humidity. It is found that, in advancing mode, unlubricated CFRP rods are hydrophilic, but the lubricated rods are hydrophobic. Contact angles increase with lubricant thickness. Therefore, humidity should have less effect on coefficient of friction on the lubricated samples. However, contact angle hysteresis is observed due to the large difference of advancing and receding contact angles. It could be explained by

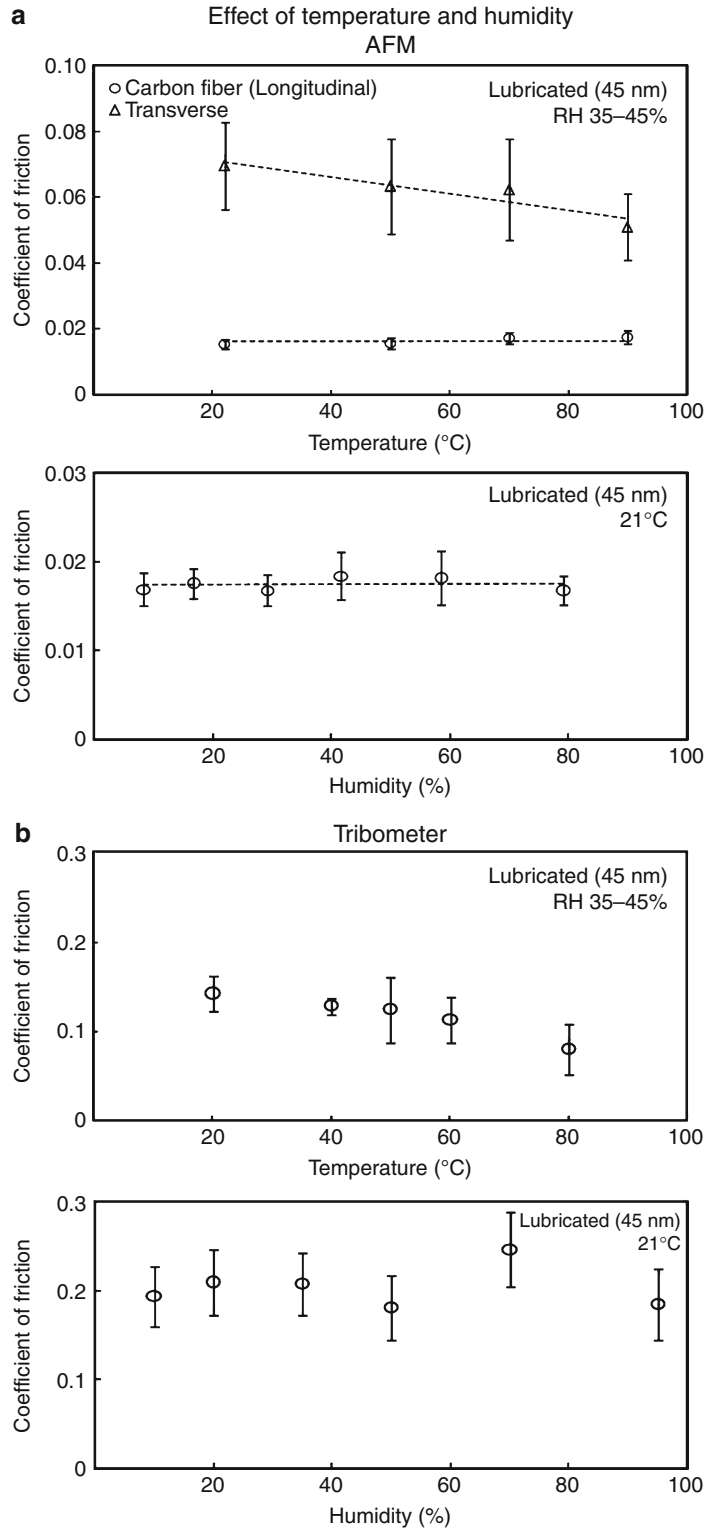


Nano-Tribology Applications in Microprojector Technology, Fig. 2 (a) AFM maps for surface height, friction force, adhesive force, and film thickness distribution for lubricated and unlubricated CFRPs. Shown above each image is a cross

section taken at a position denoted by the corresponding arrows. (b) 2-D cross-sectional profiles to correlate the surface morphology with friction force, adhesive force, and film thickness

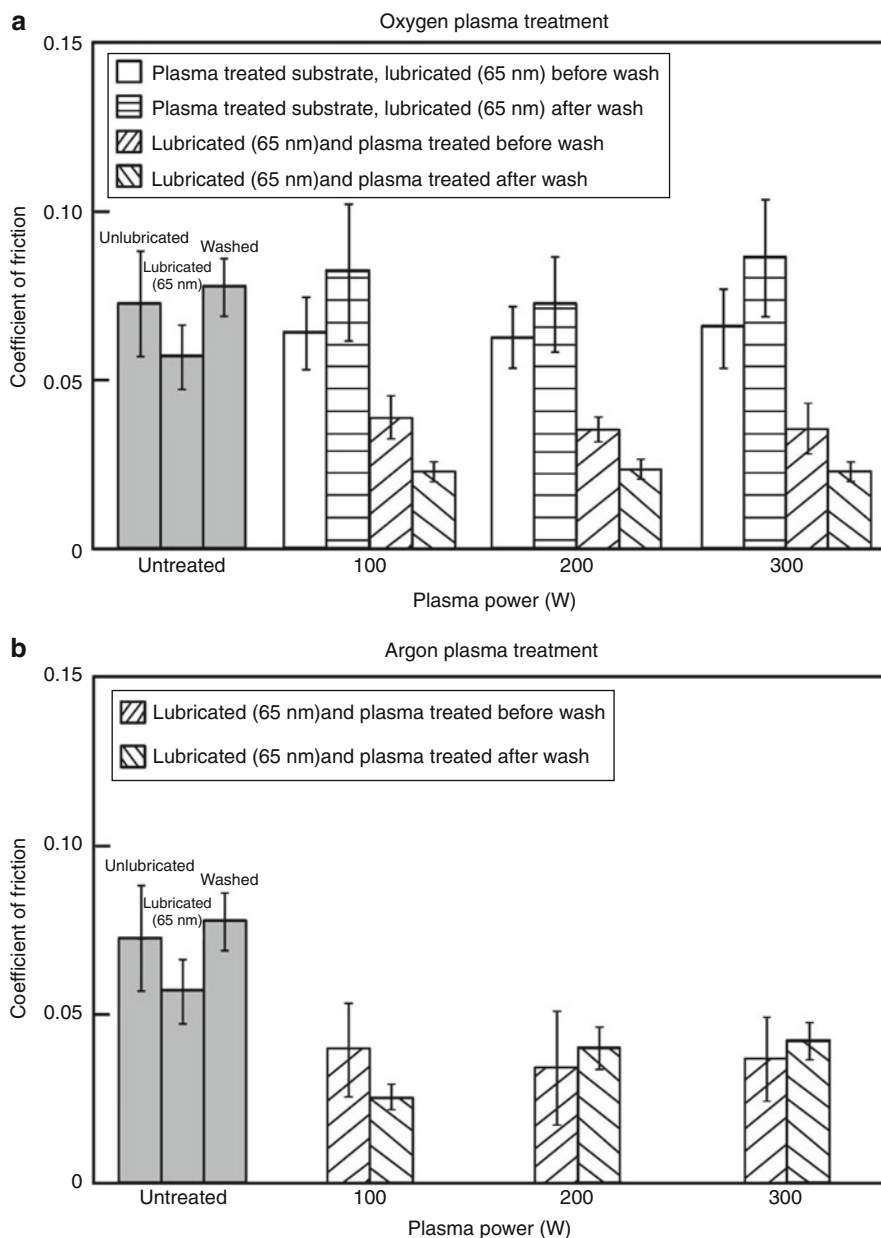
Nano-Tribology Applications in Microprojector Technology, Fig. 3

Effect of temperature and humidity on lubricated CFRP (average film thickness = 45 nm) using (a) AFM on carbon fiber in the longitudinal and transverse directions, and (b) tribometer



Nano-Tribology Applications in Microprojector Technology, Fig. 4

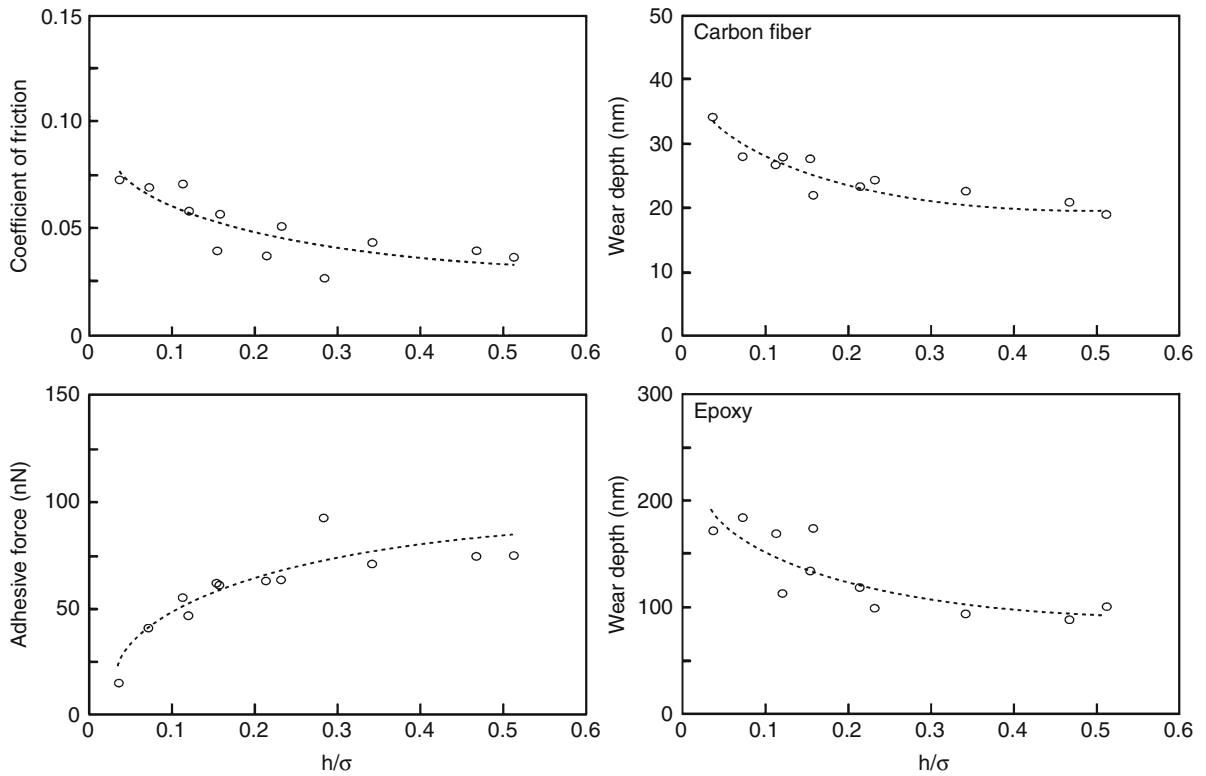
Effect of plasma power on coefficient of friction. (a) Oxygen and (b) argon plasma treatments were carried out. The first three *dark* columns are coefficient of friction on untreated samples. The next groups of columns are for plasma-treated samples. The *white blank* and *horizontally lined* columns in (a), which are not in (b), are for sample type #1. *Scatter bars* represent $\pm \sigma$ limits



chemical heterogeneity caused by the high roughness and the composite effect of the CFRP rods. Finally, the interplay of surface roughness and film thickness of lubricant is optimized for low adhesion, friction, and wear. The effect of h/σ on friction, adhesion, and durability is studied. With an increase in h/σ , the uniformity of the lubricant distribution on the CFRP rods is improved, resulting in stable and reliable operation of the actual devices. Coefficient of friction and wear depth decrease with a decrease in h/σ , but

adhesive force increases due to an increase in the amount of lubricant involved in surface interface and in the contact area between the two mating materials, and vice versa (Figs. 5 and 6).

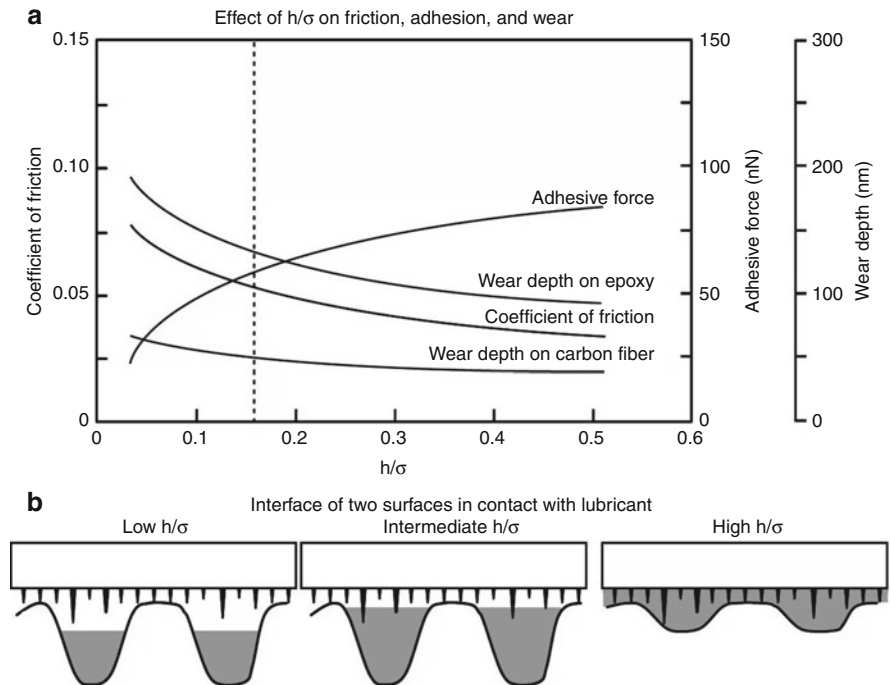
Therefore, there is a trade-off in lubricant film thickness and surface roughness for optimum values of coefficient of friction, adhesive force, and durability. The ratio of the mobile fraction of lubricant film thickness to sample roughness should be decided by the requirement and functionality of the actual devices.



Nano-Tribology Applications in Microprojector Technology, Fig. 5 The effect of h/σ on coefficient of friction, adhesive force, and wear depth on carbon fibers and epoxy regions at 30 μN of normal load. The *dotted lines* are the trend lines for the data

Nano-Tribology Applications in Microprojector Technology, Fig. 6

(a) The trend of the effect of h/σ on coefficient of friction, adhesive force, and wear depth. The dotted vertical line is the h/σ value (0.16) of the samples with 410 nm of RMS and 65 nm of lubricant thickness used in some commercial microprojectors. (b) A schematic showing the rough interface of CFRP rods and a moving part of microprojector. The gray colored area represents lubricant



Conclusion

An introduction to microprojection technology and the components involved is provided. Specifically, a keystone component in laser-based microprojector, green laser package, developed by Corning Incorporated is considered in this study. Applications of nano-tribology in the development of adaptive optics component used in this package are detailed. The adaptive optics component involves sliding mechanical elements and lubricant was used in between these elements. The tribological design has to be efficient in order for the adaptive optics component to function consistently. Therefore, a comprehensive study was done on tribological design of the lubricant and the surface characteristics of the sliding components. The design specifications obtained from this study were implemented in the final package design.

Cross-References

- ▶ [AFM](#)
- ▶ [Nanotribology](#)

References

1. Bhatia, V., Gregorski, S.J., Pikula, D., Chaparala, S.C., Loeber, D.A.S., Gollier, J., Gregorski, J.D., Hempstead, M., Ozeki, Y., Hata, Y., Shibatani, K., Nagai, F., Mori, N., Nakabayashi, Y., Mitsugi, N., Nakano, S.: Efficient and compact green laser for micro-projector applications. *J. Soc. Inf. Disp.* **17/3**, 1–6 (2009)
2. Freeman, M., Champion, M., Madhavan, S.: Scanned laser pico-projectors. *Opt. Photonics News* **20**, 29–34 (2009)
3. Yoshizumi, Y., Adachi, M., Enya, Y., Kyono, T., Tokuyama, S., Sumitomo, T., Akita, K., Ikegami, T., Ueno, M., Katayama, K., Nakamura, T.: Continuous-wave operation of 520 nm green InGaN-based laser diodes on semi-polar {2021} GaN substrates. *Appl. Phys. Express* **2**, 092101-1–092101-3 (2009)
4. Bhatia, V., Gregorski, S.J., Pikula, D., Chaparala, S.C., Loeber, D.A.S., Gollier, J., Ozeki, Y., Hata, Y., Shibatani, K., Nagai, F., Mori, N., Nakabayashi, Y., Mitsugi, N., Nakano, S.: Efficient and compact green laser incorporating adaptive optics for wide operating temperature range. *SID Int. Symp. Digest Tech. Pap.* **39**, 962–965 (2008)
5. Bhushan, B., Zhao, Z.: Macroscale and microscale tribological studies of molecularly thick boundary layers of perfluoropolyether lubricants for magnetic thin-film rigid disks. *J. Info. Storage Proc. Syst.* **1**, 1–21 (1999)
6. Bhushan, B., Lee, H., Chaparala, S.C., Bhatia, V.: Nanolubrication of sliding components in adaptive optics used in microprojectors. *Appl. Surf. Sci.* **256**, 7545–7558 (2010)
7. Graham-Rowe, D.: Projectors get personal. *Nat. Photonics* **1**, 677–679 (2007)
8. Bhatia, V., Hempstead, M., Grochosinski, J., Sekiguchi, N., Okada, A., Loeber, D.: Compact and efficient green lasers for mobile projector applications. *J. Soc. Info. Disp.* **17/1**, 1–6 (2009)
9. Hancox, N.L.: The use of a torsion machine to measure the shear strength and modulus of unidirectional carbon fiber reinforced plastic composites. *J. Mater. Sci.* **7**, 1030–1036 (1972)
10. Ogisu, T., Shimanuki, M., Kiyoshima, S., Takeda, N.: A basic study of CFRP laminates with embedded prestrained SMA foils for aircraft structures. *J. Intel. Mater. Syst. Struct.* **16**, 175–185 (2005)
11. HeroldSchmidt, U., Floeth, E., Last, B., Schafer, W., Zaglauer, H.W.: Quantification of smart composites for the use in aerospace applications. *Proc. Soc. Photo-Opt. Instrum. Eng.* **3044**, 168–175 (1997)
12. Lee, C.K.: Structure, electrochemical and wear-corrosion properties of electroless nickel-phosphorus deposition on CFRP composites. *Mater. Chem. Phys.* **114**, 125–133 (2009)
13. Rion, J., Leterrier, Y., Manson, J.A.E., Blairon, J.M.: Ultra-light asymmetric photovoltaic sandwich structures. *Compos. A* **40**, 1167–1173 (2009)
14. Palacio, M., Bhushan, B.: Nanotribological properties of novel lubrications for magnetic tapes. *Ultramicroscopy* **109**, 980–990 (2009)
15. Tao, Z., Bhushan, B.: Bonding, degradation, and environmental effects on novel perfluoropolyether lubricants. *Wear* **259**, 1352–1361 (2005)
16. Liu, H., Bhushan, B.: Nanotribological characterization of molecularly thick lubricant films for applications to MEMS/NEMS by AFM. *Ultramicroscopy* **97**, 321–340 (2003)
17. Koinkar, V.N., Bhushan, B.: Microtribological studies of unlubricated and lubricated surfaces using atomic force/friction force microscopy. *J. Vac. Sci. Technol. A* **14**, 2378–2391 (1996)
18. Zhao, X., Bhushan, B.: Study on degradation mechanisms of lubricants for magnetic thin film rigid disks. *Proc. Instn. Mech. Engrs., Part J: J. Eng. Tribol.* **215**, 173–181 (2001)
19. Bhushan, B., Lee, H., Chaparala, S.C., Bhatia, V.: Chemical bonding and degradation of ultrathin liquid films for nanolubrication of sliding components in adaptive optics. *Coll. Surf. A* **369**, 39–52 (2010)

Nanotubes

- ▶ [Chemical Vapor Deposition \(CVD\)](#)
- ▶ [Nanostructures for Energy](#)
- ▶ [Physical Vapor Deposition](#)

Nanotweezers

- ▶ [Nanogrippers](#)

Nano-twinned Materials

- ▶ [Nano-sized Nanocrystalline and Nano-twinned Metals](#)

Nanowire FET Biosensor

- ▶ [Nanostructure Field Effect Transistor Biosensors](#)

Nanowires

- ▶ [Chemical Vapor Deposition \(CVD\)](#)
- ▶ [Nanostructures for Energy](#)
- ▶ [Physical Vapor Deposition](#)

Nature Inspired Microneedle

- ▶ [Biomimetic Mosquito-Like Microneedles](#)

Near Infrared

- ▶ [Thermal Cancer Ablation Therapies Using Nanoparticles](#)

Near-Field Molecular Optics

- ▶ [Nanopolaritonics](#)

Near-Field Optics

- ▶ [Light Localization for Nano-optical Devices](#)

Near-Field Scanning Optical Microscopy (NSOM)

- ▶ [Scanning Near-Field Optical Microscopy](#)

NEMS

- ▶ [Nanotechnology](#)

NEMS Gravimetric Chemical Sensors

- ▶ [NEMS Resonant Chemical Sensors](#)

NEMS Gravimetric Sensors

- ▶ [NEMS Resonant Chemical Sensors](#)

NEMS Mass Sensors

- ▶ [NEMS Resonant Chemical Sensors](#)

NEMS Piezoelectric Switches

Nipun Sinha
Department of Mechanical Engineering and Applied Mechanics, Penn Micro and Nano Systems (PMaNS) Lab, University of Pennsylvania, Philadelphia, PA, USA

Synonyms

[Dual-beam piezoelectric switches](#); [Piezoelectric switches](#)

Definition

NEMS Piezoelectric Switches are defined as mechanical switches that use piezoelectric actuation for switching and employ ultrathin or nanopiezoelectric films for realizing these switches.

Background and Motivation

MEMS switches were first demonstrated in 1979 [1]. Since then lot of research activities have focused on their development and optimization. MEMS switches have traditionally provided a multitude of advantages like higher isolation, lower insertion loss, very low power consumption, and better linearity when compared to traditional transistor/FET and pin diodes [2]. Most of the switch development till date has been done for applications in the Radio Frequency MicroElectroMechanical Systems (RF MEMS) domain. The main aim in this case is to use mechanical relays for switching on and off various components in RF architectures and also to facilitate the functioning of reconfigurable systems. More recently, there has been growing interest in the development of low voltage mechanical computing platforms using MEMS switches and reconfigurable mechanical RF front-ends. RF MEMS switches with their favorable on and off characteristics are ideal for such applications.

Various actuation methods can be employed for achieving switching, such as electrostatic [3–12], piezoelectric [13–23], electromagnetic [24], thermal [25], and electrothermal [26]. Out of these methods, electrostatic actuation has been the most commonly and widely used actuation method. However, it suffers from inherent disadvantages due to its nonlinear mechanism, which transduction methods like piezoelectric actuation have the ability to overcome.

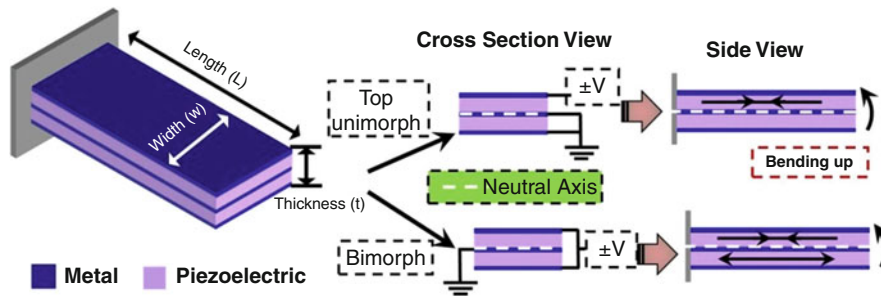
Piezoelectricity, as a phenomenon, was discovered in 1880 by Pierre Curie and Jacques Curie (the Curie brothers). It is simply characterized by the accumulation of charge in certain materials upon the application of mechanical stress. The converse piezoelectric effect also exists, which makes it a reversible process. In the converse piezoelectric effect, an applied electric field can lead to the generation of a mechanical force, thereby causing the film to strain. This converse effect was not predicted by the Curie brothers initially and was predicted by Gabriel Lippmann in 1881. It was later confirmed experimentally by the Curie brothers. The piezoelectric effect is a linear electromechanical interaction between the electrical and mechanical variables in crystalline and polycrystalline materials. For a material to exhibit this phenomenon its crystal structure needs to possess some asymmetry, so materials with inversion symmetry do not exhibit this property.

This linear interaction can be defined using a set of coupled equations between the strain (S) (6×1 matrix), compliance (s) (6×6 matrix), stress (T) (6×1 matrix), electric field (\vec{E}) (3×1 matrix), piezoelectric coefficients (d) (3×6 matrix), permittivity (ϵ) (3×3 matrix), and electric charge density displacement (D) (3×1 matrix).

$$[S] = [s^E][T] + [d^T][\vec{E}] \quad [D] = [d][T] + [\epsilon^T][\vec{E}] \quad (1)$$

Implementation of piezoelectric actuation using the converse piezoelectric effect necessitates the presence of two electrodes that sandwich the piezoelectric material to apply an electric field across it. For in plane stretching and consequently bending operation, the d_{31} piezo-coefficient is conventionally used in c -axis oriented, hexagonal materials such as AlN. In this case, an electric field applied across the film thickness generates an in-plane strain through the equivalent d_{31} piezo-coefficient. In a multilayered beam, this strain can be tailored to generate a bending moment about the neutral axis of the composite beam. This moment deflects the beam out-of-plane, i.e., vertically, and this motion is used to open and close an air gap in a mechanical switch. There are two commonly used configurations for implementing piezoelectric actuation, namely, unimorph and bimorph, shown in Fig. 1. In the unimorph configuration [27], there is only one operational piezoelectric layer. A unimorph actuator has both active and passive layers. The passive layers are used for offsetting the neutral axis so that the piezoelectric force can generate a moment about it. In the bimorph configuration [28] there are two active piezoelectric layers, which are simultaneously actuated. The electrodes sandwiching the actuator are properly designed and routed so that the two layers provide for additive bending moments and generate a greater deflection per volt than in the unimorph case, for a given piezoelectric layer thickness.

Majority of the piezoelectric switches made till date have employed either Lead Zirconate Titanate (PZT) or Aluminum Nitride (AlN) as their functional material. When compared to PZT, AlN has the advantage of being compatible with CMOS processes and foundries (by being lead free), exhibits a lower leakage current for comparable film thicknesses and has a lower dielectric constant, which ultimately translates to reduced power consumption for a given actuator dimension. PZT does exhibit a higher piezoelectric coefficient than AlN, but repeatability of high-quality PZT films has been known to be a challenge.



NEMS Piezoelectric Switches, Fig. 1 Schematic representation of the piezoelectric actuator (showing the length (L), width (w), and thickness (t) of the actuator) made of two piezoelectric layers and three metal layers illustrating the two common

topologies used for actuation, i.e., the unimorph (using one layer of piezoelectric material for actuation) and bimorph (using two layer of piezoelectric material for actuation)

This, along with the need to pole PZT post deposition (in the sol-gel process) has been the major reason why none of the prominent examples of PZT switches employ bimorph actuation, whereas this can be easily achieved when using AlN. Unlike PZT, AlN is easy to fabricate and has shown preservation of piezoelectric coefficients when scaled from thick ($1\ \mu\text{m}$) to thin ($100\ \text{nm}$) films [29]. Another unique advantage that AlN-based devices have over PZT-based devices is the ability of easy integration with the already well-established AlN-based contour-mode resonator technology [30–33]. With this integration new classes of transceivers that will use single-chip multifrequency and highly reconfigurable systems capable of low power and low loss operation can be realized.

AlN-Based Dual-Beam Piezoelectric RF MEMS Switch

RF MEMS Switch

The AlN-based MEMS switch described herein utilizes a novel dual-beam design, i.e., it consists of two symmetric, released beams that face each other and move in opposing directions (with respect to each other) with the same two inputs. Figure 2 shows an SEM of a dual-beam switch and the two inputs (i.e., the body and the gate) that are required for the switch to function. The reason for implementing such a design is the presence of residual stress in released multilayered structures. Residual stresses in MEMS and NEMS devices are unwarranted effects that result from the fabrication processes that occur at elevated temperatures. The residual stresses affect the structural reliability and the performance of the system.

The deflection induced by residual stresses can cause significant variation in threshold voltages and can lead to a normally closed switch, and, in extreme cases, the release can lead to large deflections rendering the switch inoperable. Therefore, it is essential to overcome this effect in MEMS switches. The dual-beam design helps reducing the impact of this and is introduced primarily because of the need to fabricate switches immune to these residual stresses.

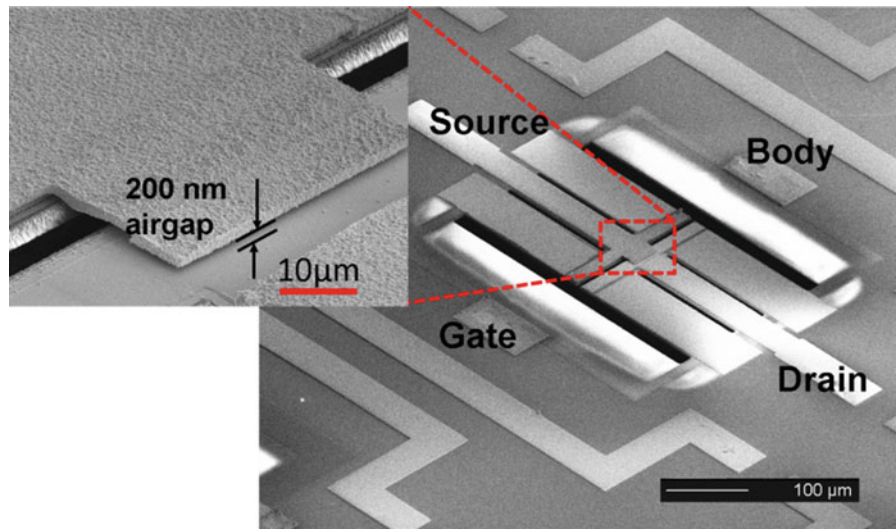
In the dual-beam mechanism employed, both beams can be actuated simultaneously, thus giving: (1) low actuation voltage (halved, as compared to just one beam moving the whole gap), (2) high contact forces (3) active pull-off to open the switch, and (4) faster switching times (as each beam has to move only half the gap). The nano-gap, shown in the zoomed-in view in Fig. 2, is what enables the switch to achieve both good contact forces and isolation.

The three-finger dual-beam switch design used for implementing piezoelectric devices was inherited from the initial RF switch development [21]. This topology comprises of using two symmetric released beams facing each other. Each of these two beams is made of three separate fingers (hence the three-finger design), the two outer fingers are used for actuation and the central finger is employed solely for carrying the signal from the source to the drain as to avoid any gate to source/drain overlap (Fig. 3). The actuators are made up of two layers of AlN, so this design helps in realizing bimorph actuation in piezoelectric switches.

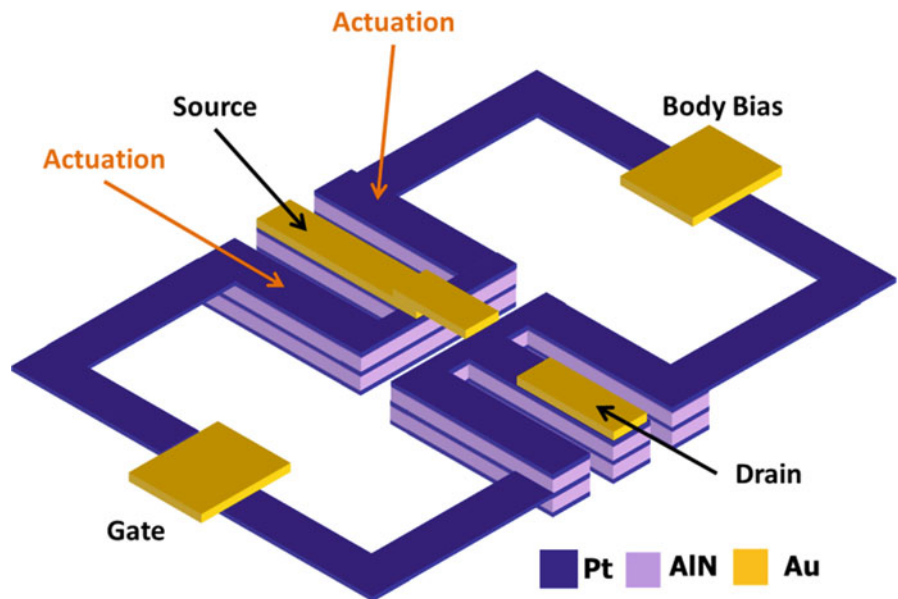
Integration with Contour-Mode Technology

As mentioned earlier, one of the biggest advantages of developing an AlN-based RF MEMS switch was the ability to integrate it with the already proven

NEMS Piezoelectric Switches, Fig. 2 A SEM of a fabricated three-finger AIN MEMS switch with a zoomed-in view of the planarized nano-air gap (~200 nm). Single gate and body terminals are created by appropriate electrode routing



NEMS Piezoelectric Switches, Fig. 3 Schematic representation of the piezoelectric switch, showing the three-finger dual-beam design. The two pads are used to apply the gate and body-bias voltages and are properly connected to the layers of either actuator so that they actuate both the beams simultaneously



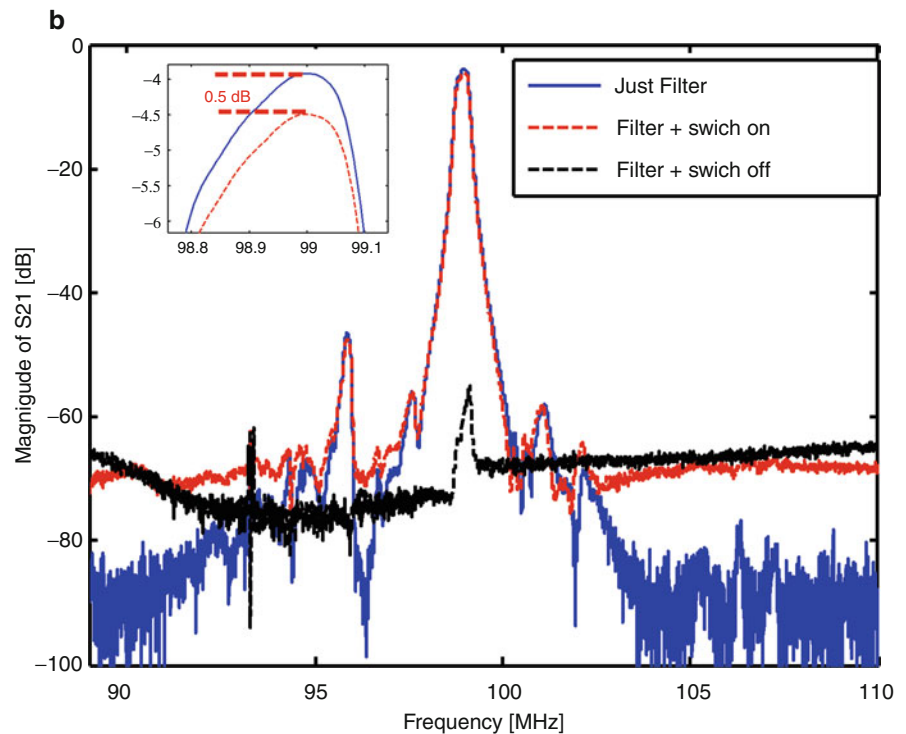
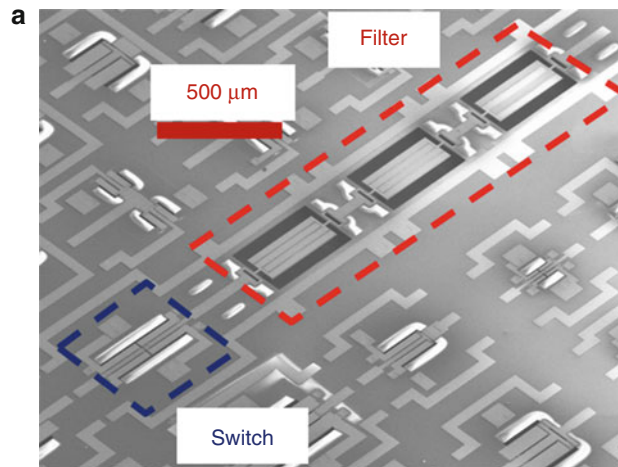
contour-mode technology to form switchable frequency control elements. This monolithic integration has the potential to tremendously impact the size and power efficiency of RF front-ends by drastically reducing parasitics and off-state leakage. Figure 4a shows an SEM of the monolithically integrated third order AIN contour-mode technology filter along with an AIN switch. Figure 4b shows the frequency response of the filter when it has been switched on and off. An attenuation of >50 dB was measured between the on

and off states of the filter. The measurements also show the performance of the filter without the switch. The inset in Fig. 4b shows a zoomed-in view of the difference in insertion loss of the filter with the switch and without the switch, at the filter's center frequency (~99 MHz). The presence of the switch in the measuring path introduces a low loss in the performance of the filter, by increasing the insertion loss by ~0.5 dB.

Such integration of switches and frequency control components has the potential to enable new

NEMS Piezoelectric

Switches, Fig. 4 (a) SEM of a monolithically integrated RF MEMS switch with a contour-mode filter. (b) Shows the experimental results of a switchable AIN filter effectively being turned on and off by an AIN RF MEMS switch. The switch introduces low loss and enables a single-chip switchable filter bank solution



reconfigurable RF architectures based on massive arrays of multifrequency AIN contour-mode resonators and filters.

MEMS/NEMS Switch-Based Logic Using Body-Biasing

MEMS switches can also be easily employed for computing applications. Scaling of CMOS technology

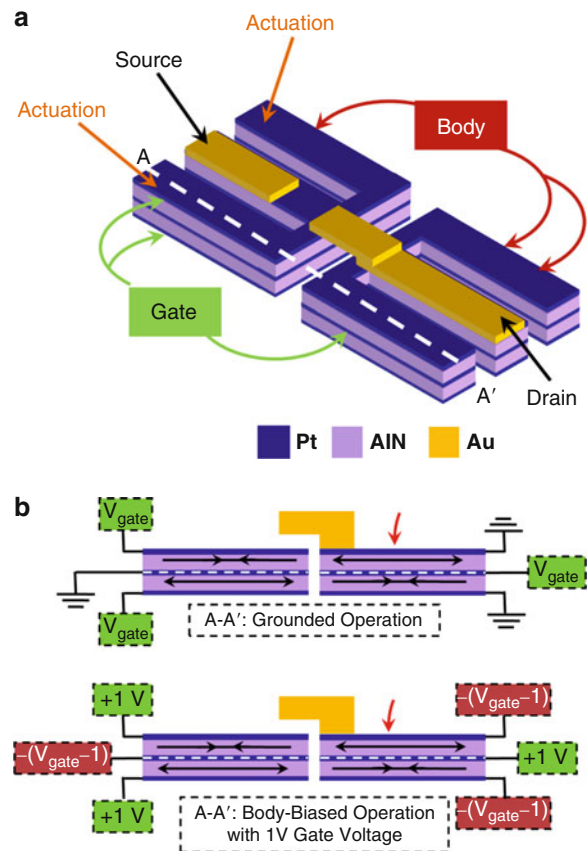
to the nano-regime is facing issues with power consumption and reduction of threshold voltages. For overcoming these challenges the International Technology Roadmap for Semiconductors (ITRS) has recognized the development of technologies, like nanomechanical computing based on NEMS switches, as a possible solution for future electronic transistor replacement. Unlike transistors, MEMS/NEMS switches do not suffer from the drawback of passive power loss as they usually have an air gap instead of

the physical channel between the source and the drain. Therefore, they can be used for developing efficient low power computing devices. AlN-based switches have been used to implement a unique MEMS switch actuation methodology, i.e., the body-biased actuation, which has enabled low voltage demonstration of basic logic-building blocks. These switches can be referred to as piezoelectric mechanical transistors (MTs) as they use mechanical transduction to achieve the function that is traditionally carried out by electronic transistors.

Body-Biased Actuation

As mentioned earlier, piezoelectric actuation necessitates the presence of two actuation electrodes that sandwich the piezoelectric material and create an electric field across it. It is this electric field across the piezoelectric film, which generates the strain that leads to actuation. A body bias, in the present context, stands for a constant and floating voltage applied to one of the actuation electrodes. Traditionally, the actuators are operated by applying a voltage (V_{gate}) to the gate terminal, which is greater in magnitude than the threshold voltage (V_{th}) of the device, i.e., $V_{gate} > V_{th}$. This mode of actuation is referred to as the “grounded” mode of actuation, since the body terminal of the device is grounded, shown in Fig. 5b. Instead, in the “body-biased” mode of operation, the previously grounded electrode is biased to a constant voltage (V_{bias}) close in magnitude (but opposite in polarity) to the threshold voltage so that a lower gate voltage can be used to open and close the switch. In Fig. 5, the body-biased mode of actuation is explained schematically using a cross-section of a three-finger dual-beam switch [19, 20]. Figure 5b presents a scenario where the gate voltage of the MT is reduced from V_{gate} to 1 V by using a body-bias voltage of $-(V_{gate}-1)$. This new and lower threshold voltage leads to a smaller swing voltage, which effectively lowers the energy required for actuation. From a mechanical perspective, body-biasing can be explained as preloading (prebending) of the beams that form the actuators. Effectively, body-biasing corresponds to the reduction of the air gap between the two contacting surfaces and translates into the ability of tuning the device threshold voltage by modulating the size of the initial contact gap.

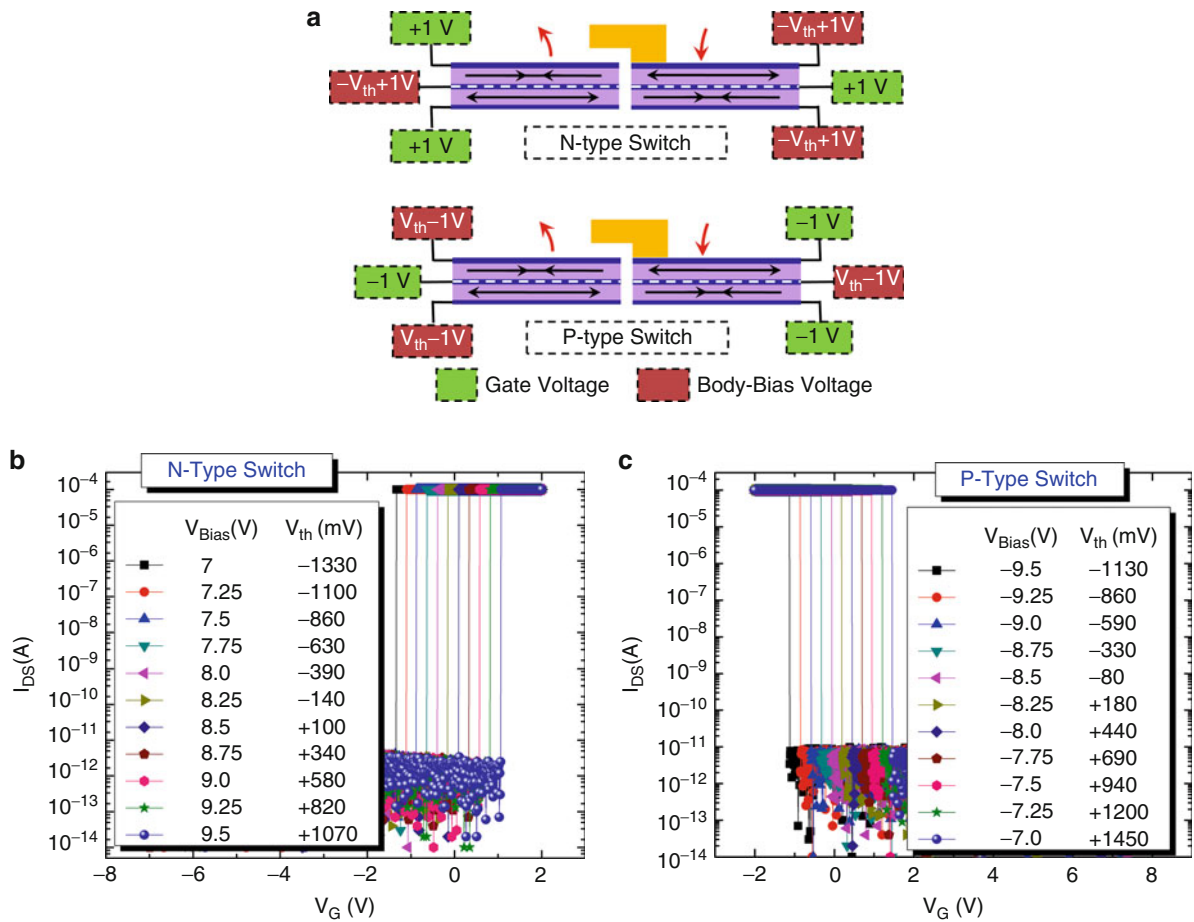
By using the body-biased mode of actuation, the threshold voltage of a piezoelectric switch can be



NEMS Piezoelectric Switches, Fig. 5 (a) 3D schematic of a three-finger dual-beam AlN switch highlighting the source, drain, gate, and body terminals. It also shows the A–A’ cross-section used to illustrate the body-biasing method. (b) Schematic representation of the principle of operation of an AlN MEMS switch in grounded and body-biased mode of operation. By applying a bias to the body terminal $[-(V_{gate}-1)]$ V, the same effective potential required to close the switch is applied across the AlN layer with a lowered gate voltage of 1 V (1 V is used here as an example, but, theoretically, any value is possible thanks to the extreme linearity of the body-bias technique in a piezoelectric AlN film). This effectively reduces the dynamic energy required for every switching operation. The gold tip is not present along the A–A’ cross-section, but is drawn here to aid the understanding of the mechanism of device operation

tailored to be any value. The same switch can have a negative or positive threshold voltage, as shown in Fig. 6. The major aim of this technique is to reduce the threshold voltage in magnitude (as low as 1–20 mV), which can be easily achieved by body-biasing.

The absence of a doped channel in the MT renders the classification of n- or p-type devices ambiguous. Moreover, this provides the MT with an unparalleled



NEMS Piezoelectric Switches, Fig. 6 (a) Schematic representation of the complementary operation made possible in piezoelectric switches by using the body-biasing method. By changing the polarity of the body-bias and the electrodes to which it is applied, the nature of the switch (p- or n-like) can be modified without any alteration to its layout design or

fabrication. (b, c) Experimental data verifying that the switch threshold voltage (V_{th}) can be tuned by varying the value of the body bias. This data show that the body-bias technique enables the implementation of variable threshold voltages on the same die and in the same process. The current measurement was limited to a maximum value of 100 μ A

ability to choose or vary the nature of a switch on the basis of need, i.e., the same switch can be made into an n- or p-like device (in analogy to p-type and n-type CMOS) without any change in design or fabrication by simply switching the polarity of the body-bias and applying it to the opposing actuating electrode set (as shown in Fig. 6a). For the switch to be actuated, the magnitude and direction of the electric field in the piezoelectric material has to be preserved. The polarity and value of the potential on any particular electrode is inconsequential.

The subthreshold slope has been measured to be <1 mV/dec as the actuation of these structures is not limited by thermal noise. This gives these switches

with body-bias the ability to lose nearly no power when they are in standby.

An Agilent B1500A Semiconductor Parameter Analyzer was used for measuring the switch threshold voltage as a function of body bias applied to the MT. Figure 6b, c illustrate the results of 11 successive scans, each, on a $200 \times 200 \mu\text{m}^2$ device, showing the change in the threshold voltage (V_{th}) as the body-bias (V_{bias}) is increased with a 250 mV step per scan, both for an n-like and a p-like MT. Equivalent to an n-type CMOS transistor, the MT can be made to operate as a depletion or enhancement mode device depending on the value of the body-bias. The absolute value of the body-bias voltages were the same for both the scan sets.

Good agreement between threshold values for corresponding positive and negative body-bias values was observed, but small differences in the magnitude of the threshold voltage (10–120 mV) between the n- and p-like cases were recorded and can be attributed to subtle changes in the adhesive forces due to the evolution of the contact. These data conclusively show that, by using body-biasing the same switch can be made to behave either as an n-like switch or as a p-like switch.

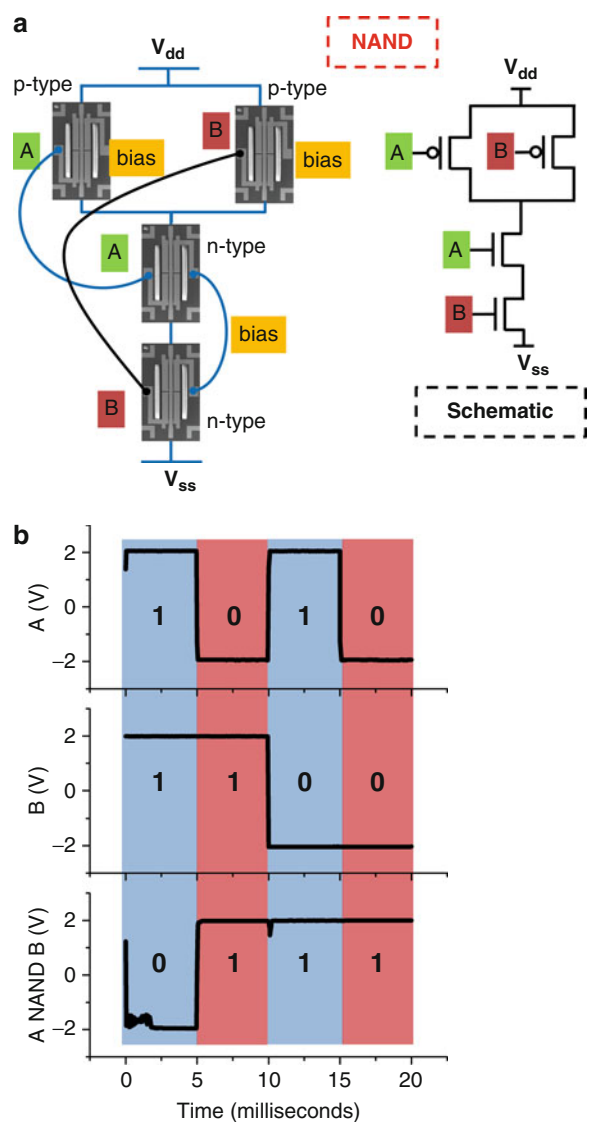
Logic-Building Blocks Using MEMS/NEMS Switches

For initial demonstration of mechanical logic achieved by AlN MTs the logic circuit layout was adopted directly from traditional CMOS design and mechanical transistors were used to simply replace the CMOS transistors. This approach was used to implement and validate the ability of realizing low swing voltage dynamically operated complementary piezoelectric logic circuits via body-biasing.

These three-finger dual-beam AlN MEMS switches have been able to demonstrate logic elements like inverters, NAND, and NOR gates. The key factors that have enabled this low voltage demonstration are the use of thin AlN films (250 nm) and of the body-biasing technique. Two AlN MEMS switches were used to demonstrate a low voltage digital inverter using body-biased AlN switches capable of fast switching (~220 ns). The mechanical inverter formed was tested at 100 Hz with a ± 1.5 V (square input wave) voltage swing. The NAND and NOR structures were formed by wire-bonding four separate switches in the desired configuration. The mechanical NAND was formed by using AlN mechanical switches with one input at 100 Hz and the other at 50 Hz with a ± 2 V voltage swing, as shown in Fig. 7a. The experimental results and superimposed equivalent truth table for the dynamic testing of a single cycle of a mechanical NAND gate is shown in Fig. 7b.

AlN Device Scaling and Nano-Actuator Development

The necessity to miniaturize the MT and improve device density, scaling of AlN MTs has to be investigated and projected. The results of an analysis of how dimensional scaling affects various important actuator parameters such as displacement, δ_E , resonance frequency, f_n ,

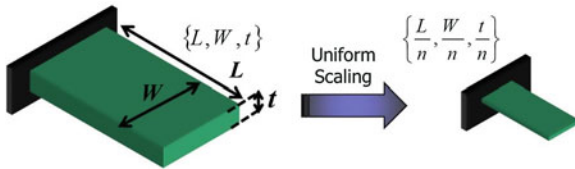


NEMS Piezoelectric Switches, Fig. 7 (a) Schematic representation of the switch layout showing the wiring that was required to form the NAND gate. The electronic schematic of a conventional NAND is also shown to illustrate the similarities between the two implementations. (b) Experimental data for one cycle of operation of a NAND gate made using mechanical switches and its corresponding truth table. This implementation was achieved with a ± 2 V swing

contact force, F_c , gap size, g , as well as the energy consumption, $(En)_E$, and power consumption, P_E , of a simple logic gate is show in Table 1. The analysis was carried out assuming uniform scaling of the actuator (Fig. 8), i.e., scaling of all the dimensions of the actuator together by the same factor. The other assumption of this

NEMS Piezoelectric Switches, Table 1 Scaling factors for constant-field scaling of piezoelectric MTs. Scaling factor: k

Property	Relation to independent variables (L, w, t, V)	Constant-field scaling
Displacement (δ_E)	$\propto \frac{L^2}{t}$	$\propto \frac{1}{k^1}$
Contact force (F_C)	$\propto \frac{wt^2}{L}$	$\propto \frac{1}{k^2}$
Resonance frequency (f_n)	$\propto \frac{t}{L^2}$	$\propto k^1$
Gap (g)	$\propto \frac{L^2}{t}$	$\propto \frac{1}{k^1}$
Body-bias voltage	$\propto \frac{L^2}{t}$	$\propto \frac{1}{k^1}$
Energy (En) _E	$\propto wLV$	$\propto \frac{1}{k^3}$
Power (P_E)	$\propto \frac{wV}{L}$	$\propto \frac{1}{k^2}$

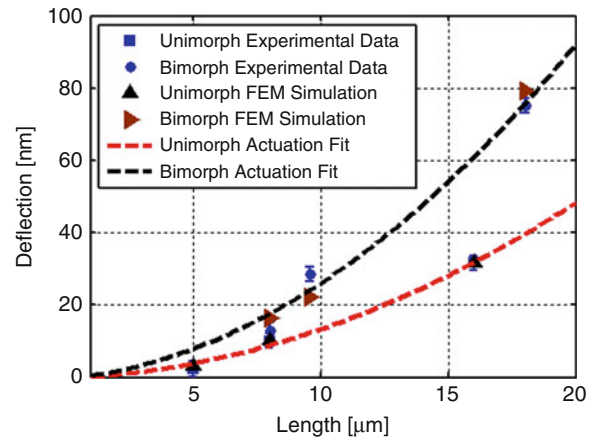


NEMS Piezoelectric Switches, Fig. 8 3D schematic showing uniform scaling as has been proposed for the mechanical transistors. All the dimensions of the actuator should be scaled by the same factor

study was the use of constant electric field, i.e., the magnitude of the electric field in the piezoelectric was kept constant throughout the scaling process. From the prospective of energy and power consumed the advantages of scaling are amply visible in the table. It is interesting to note the similarities to CMOS scaling. The scaling of the channel length and the thickness of the gate oxide are effectively analogous to the scaling of the actuator length and the piezoelectric film thickness. As for CMOS scaling, MT scaling also leads to a linear increase in speed and quadratic reduction in power.

In the past, the biggest hurdle in the scaling of piezoelectric transduction to the nanoscale has been the degradation of piezoelectric properties due to limited orientation in thin films and increases in internal stresses (cracking and excessive deformations in released structures). With better control over the deposition parameters and extensive process development ultrathin AIN films have been successfully developed and demonstrated [34–36].

Nanoactuators have been developed to study the effect of scaling (to the nano-regime) on the



NEMS Piezoelectric Switches, Fig. 9 Results of deflection testing of beams of different lengths with unimorph actuation (single layer) and bimorph actuation (double layer) at 4 V and their agreement with FEM analysis. The *dashed line* shows the square dependence that exists analytically between tip displacement and beam length

piezoelectric properties of AIN. The films were successfully scaled down to the thickness of 100 nm, and actuators as well as switches were fabricated using them. DC voltage-induced nano-beam deflections of six beams of different lengths (5–18 μm) under unimorph and bimorph actuation at 4 V are shown in Fig. 9. The measurements were made by a Zygo optical profilometer. Both layers of AIN have been individually actuated and the deflection measurement for each individual layer and both the layers actuated together closely match the FEM predictions (using COMSOL). This matching with FEM analysis verifies the dependence of vertical displacement on the square of the beam length as predicted by conventional Euler beam-bending equations. These deflection measurements along with analytical solutions for the deflection of piezoelectric unimorph and bimorph beams [37] were used for extracting the d_{31} piezoelectric coefficient of the ultrathin AIN film. The extracted experimental d_{31} is within 5% of the theoretical value of the piezoelectric coefficient.

Finally, the successful scaling of the AIN film led to the demonstration of an AIN piezoelectric nanoswitch that employed 100 nm thick AIN films. These switches validate the use of the body-biasing technique at the nanoscale. As the thickness of the films has been scaled to the nano-dimensions the body-bias voltages have been lowered to <6 V and

a better control over the actuator motion has been demonstrated. This enhanced control over switching translates into a threshold voltage of ~ 1 mV and the ability to potentially operate with supply voltages of few 10 s of mV.

These demonstrations, both in the RF domain as well as in the field of mechanical computing, clearly show the scalability and the potential that this piezoelectric mechanical switch/MT technology possesses. Given continued effort this can lead to development of low power highly integrated RF and computing systems.

Cross-References

- ▶ [Piezoelectric Effect at Nanoscale](#)
- ▶ [Piezoelectric MEMS Switch](#)

References

1. Petersen, K.E.: Micromechanical membrane switches on silicon. *IBM J. Res. Dev.* **23**, 376–385 (1979)
2. Rebeiz, G.M.: *RF MEMS Theory, Design and Technology*, 1st edn. Wiley, Hoboken (2003)
3. Zavracky, P.M., Majumder, S., McGruer, N.E.: Micromechanical switches fabricated using nickel surface micromachining. *J. MicroElectroMech. Syst.* **6**(1), 3–9 (1997)
4. Bozler C., Drangmeister, R., Duffy, S., Gouker, M., Knecht, J., Kushner, L., Parr, R., Rabe, S., Travis, L.: MEMS micro-switch arrays for reconfigurable distributed microwave components. In: *IEEE MTT-S International Microwave Symposium Digest*, Boston, pp. 153–156 (2000)
5. Weon Wi, J., JeongOen, L., Hyun-Ho, Y., Jun-Bo, Y.: Mechanically operated random access memory (MORAM) based on an electrostatic microswitch for nonvolatile memory applications. *IEEE Trans. Electron Dev.* **55**(10), 2785–2789 (2008)
6. Chong, S., Akarvardar, K., Parsa, R., Yoon, J.-B., Howe, R.T., Mitra, S., Wong, H.-S.P.: Nanoelectromechanical (NEM) relays integrated with SRAM for improved stability and low leakage. In: *IEEE/ACM International Conference on Computer-Aided Design*, San Jose, 2–5 Nov 2009
7. Parsa, R., Akarvardar, K., Provine, J., Lee, D., Elata, D., Mitra, S., Wong, H.-S.P., Howe, R.T.: Composite polysilicon-platinum lateral nanoelectromechanical relays. In: *14th Solid-State Sensors, Actuators, and Microsystems Workshop*, Hilton Head, 6–10 June 2010, pp. 7–10
8. Nathanael, R., Pott, V., Kam, H., Jeon, J., King Liu, T.-J.: 4-terminal relay technology for complementary logic. In: *IEDM Technical Digest*, pp. 223–226 (2009)
9. Spencer, M., Chen, F., Wang, C.C., Nathanael, R., Fariborzi, H., Gupta, A., Kam, H., Pott, V., Jeon, J., Liu, T.-J.K., Marković, D., Alon, E., Stojanović, V.: Demonstration of integrated micro-electro-mechanical relay circuits for VLSI applications. *IEEE J. Solid St. Circ* **46**(1), 308–320 (2011)
10. Kam, H., Pott, V., Nathanael, R., Jeon, J., Alon, E., King Liu, T.-J.: Design and reliability of a micro-relay technology for zero-standby-power digital logic applications. In: *IEDM Technical Digest*, pp. 809–812 (2009)
11. Abele, N., Fritschi, R., Boucart, K., Casset, F., Ancey, P., Ionescu, A.M.: Suspended-gate MOSFET: bringing new MEMS functionality into solid-state MOS transistor. In: *IEEE International Electron Devices Meeting. IEDM Technical Digest*, pp. 479–481, 5–7 Dec 2005
12. Duffy, S., Bozler, C., Rabe, S., Knecht, J., Travis, L., Wyatt, P., Keast, C., Gouker, M.: MEMS microswitches for reconfigurable microwave circuitry. *Microwave Wireless Comp. Lett.* **11**(3), 106–108 (2001)
13. Lee, H.C., Park, J.H., Park, J.Y., Nam, H.J., Bu, J.U.: Design, fabrication and RF performances of two different types of piezoelectrically actuated Ohmic MEMS switches. *J. Micromech. Microeng.* **15**, 2098–2104 (2004)
14. Park, J.H., Lee, H.C., Park, Y.H., Kim, Y.D., Ji, C.H., Bu, J., Nam, H.J.: A fully wafer-level packaged RF MEMS switch with low actuation voltage using a piezoelectric actuator. *J. Micromech. Microeng.* **16**, 2281–2286 (2006)
15. Gross, S.J., Tadigadapa, S., Jackson, T.N., Trolier-McKinstry, S., Zhang, Q.Q.: Lead-zirconate-titanate-based piezoelectric micromachined switch. *Appl. Phys. Lett.* **83**(1), 174–176 (2003)
16. Judy, D.C., Pulskamp, J.S., Polcawich, R.G., Currano, L.: Piezoelectric Nanoswitch. In: *22nd IEEE International Conference on Micro Electro Mechanical Systems*, pp. 591–594, 25–29 Jan 2009
17. Proie, R.M., Polcawich, R.G., Pulskamp, J.S., Ivanov, T., Zaghoul, M.E.: Development of a PZT MEMS switch architecture for low-power digital applications. *J. Microelectromech. Syst.* **20**(4), pp. 1–11 (2011)
18. Judy, D., Polcawich, R.G., Pulskamp, J.: Low voltage complementary MEMS logic using piezoelectric actuators. In: *2008 Solid State Sensor, Actuator and Microsystems Workshop*, Hilton Head Island, pp. 328–331 (2008)
19. Sinha, N., Jones, T.S., Guo, Z., Piazza, G.: Demonstration of low voltage and functionally complete logic operations using body-biased complementary and ultra-thin AlN piezoelectric mechanical switches. In: *The 23rd IEEE Conference on Micro Electro Mechanical Systems (MEMS 2010)*, Hong Kong, pp. 751–754 (2010)
20. Sinha, N., Jones, T.S., Guo, Z., Piazza, G.: Body-biased complementary logic implemented using AlN piezoelectric MEMS switches. In: *2009 IEEE International Electron Devices Meeting (IEDM 2009)*, Baltimore, pp. 813–816 (2009)
21. Mahameed, R., Sinha, N., Pisani, M.B., Piazza, G.: Dual beam actuation of piezoelectric AlN RF mems switches monolithically integrated with AlN contour-mode resonators. *J. Micromech. Microeng.* **18**, 105011 (2008)
22. Sinha, N., Mahameed, R., Zuo, C., Piazza, G.: Integration of AlN micromechanical contour-mode technology filters with three-finger dual beam AlN MEMS switches. In: *Frequency Control Symposium, 2009 Joint with the 22nd European Frequency and Time forum. IEEE International*, pp. 1–4, 20–24 Apr 2009

23. Polcawich, R.G., Judy, D., Pulskamp, J.S., Trolier-McKinstry, S., Dubey, M.: Advances in piezoelectrically actuated RF MEMS switches and phase shifters. In: Microwave Symposium, IEEE/MTT-S International, pp. 2083–2086 (2007)
24. Ruan, M., Shen, J., Wheeler, C.B.: Latching micromagnetic relays. *J. Microelectromech. Syst.* **10**, 511–517 (2001)
25. Streeter, R.D., Hall, C.A., Wood, R., Madadevan, R.: VHF high-power tunable RF bandpass filter using microelectromechanical (MEM) microrelays. *Int. J. RF Microwave CAE* **11**(5), 261–275 (2001)
26. Blondy, P., Cros, D., Rey, P., Charyet, P., Diem, B., Zanchi, C., Quorin, J.B.: Low voltage high isolation MEMS switches. In: Silicon Monolithic Integrated Circuits RF Systems, pp. 47–49 (2001)
27. Zuo, C., Sinha, N., Piazza, G.: Novel electrode configurations in dual-layer stacked and switchable AlN contour-mode resonators for low impedance filter termination and reduced insertion loss. In: 23rd IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2010), pp. 719–722 (2010)
28. Piazza, G., Stephanou, P.J., Pisano, A.P.: Single-chip multiple-frequency AlN MEMS filters based on contour-mode piezoelectric resonators. *J. Microelectromech. Syst.* **16**(2), 319–328 (2007)
29. Smits, J.G., Choi, W.-S.: The constituent equations of piezoelectric heterogeneous bimorphs. *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.* **38**(3), 256–270 (1991)
30. Smits, J.G., Dalke, S.L., Cooney, T.K.: The constituent equations of piezoelectric bimorphs. *Sensor. Actuator. A.* **28**(1), 41–61 (1991)
31. Sinha, N., Wabiszewski, G.E., Mahameed, R., Felmetsger, V.V., Tanner, S.M., Carpick, R.W., Piazza, G.: Piezoelectric aluminum nitride nanoelectromechanical actuators. *Appl. Phys. Lett.* **95**, 053106 (2009)
32. Zuo, C., Sinha, N., Pisani, M.B., Perez, C.R., Mahameed, R., Piazza, G.: Channel-select RF MEMS filters based on self-coupled AlN contour-mode piezoelectric resonators. In: IEEE Ultrasonics Symposium, pp. 1156–1159, 28–31 Oct 2007
33. Piazza, G., Stephanou, P.J., Pisano, A.P.: One and two port piezoelectric higher order contour-mode MEMS resonators for mechanical signal processing. *Solid State Electron.* **51** (11–12), 1596–1608 (2007)
34. Felmetsger, V., Laptev, P., Tanner, S.: Crystal orientation and stress in AC reactively sputtered AlN films on Mo electrodes for electro-acoustic devices. In: Proceedings IEEE IUS 2009, Beijing, pp. 2146–2149 (2008)
35. Felmetsger, V.V., Laptev, P.N., Graham, R.J.: Deposition of ultrathin AlN films for high frequency electroacoustic devices. *J. Vac. Sci. Tech. A.* **29**(2), 021014–021014-7 (2011)
36. Sinha, N., Guo, Z., Felmetsger, V.V., Piazza, G.: 100 nm thick aluminum nitride based piezoelectric nano switches exhibiting 1 mV threshold voltage via body-biasing. In: Proceedings of 2010 Solid-State Sensors, Actuators, and Microsystems Workshop (Hilton Head 2010), pp. 352–355 (2010)
37. DeVoe, D.L., Pisano, A.P.: Modeling and optimal design of piezoelectric cantilever microactuators. *J. MEMS* **6**(3), 266–270 (1997)

NEMS Resonant Chemical Sensors

Matteo Rinaldi

Department of Electrical and Computer Engineering,
Northeastern University, Boston, MA, USA

Synonyms

[NEMS gravimetric chemical sensors](#); [NEMS gravimetric sensors](#); [NEMS mass sensors](#); [NEMS resonant mass sensors](#)

Definition

Nanoelectromechanical System (NEMS) Chemical Sensors are devices that combine electrical and mechanical functionalities at the nanoscale for the detection of minute concentrations of target gaseous compounds in the environment. A NEMS Resonant Chemical Sensor is composed of three fundamental elements: an electrically driven nanomechanical resonator (NEMS resonator), a chemical interactive material (CIM) deposited on the surface of the NEMS resonator, and a readout system. The natural resonance frequency of the nanomechanical resonator is highly sensitive to variations in the device mass per unit area. Therefore, any small amount of mass per unit area loaded on the nanomechanical resonator will produce a relatively large shift in the natural resonance frequency of the device. The smaller is the mass per unit area of the mechanical resonator itself, the larger is the shift in resonance frequency for a given amount of mass per unit area loaded on the device surface. Therefore, an ultra-miniaturized (nano) mechanical resonator, with an extremely low mass per unit area, is ultrasensitive to mass loading per unit area. When the sensor is exposed to a certain concentration of a target gaseous compound, a number of gas molecules per unit area (mass per unit area), depending on the affinity between the CIM and the target gas, are absorbed by the CIM on the NEMS resonator surface causing a shift in its natural resonance frequency. Since a specific amount of mass per unit area is absorbed by the CIM for a given concentration of gas in the environment and a specific shift in resonance frequency is caused by a given amount of mass per unit area loaded on the

NEMS resonator, it is possible to quantitatively detect the concentration of the target gas by monitoring the natural resonance frequency of the NEMS resonator. The resonance frequency of the NEMS resonator is monitored over time and converted into an electrical signal by the readout system. The fundamental performances of a NEMS resonant chemical sensor are expressed in terms of limit of detection (LOD), response time, and selectivity. The LOD is the smallest amount of adsorbed mass per unit area (hence minimum gas concentration given a certain CIM) that can be resolved by the sensor. This quantity is inversely proportional to the resonator sensitivity to mass loading per unit area and directly proportional to the frequency fluctuations induced by the noise generated in the resonator and in the readout system. Therefore, a nanomechanical resonator characterized by a very large sensitivity to mass loading per unit area and a very low noise is the best candidate for the implementation of a high performance NEMS resonant chemical sensor. Response time and selectivity are instead determined by the chemical and physical properties of the CIM employed.

Main Text

In recent years, the demand for highly miniaturized sensor arrays capable of selectively detecting extremely small concentrations (part per trillion, ppt) of multiple gaseous analytes has steadily grown. The necessity to detect such small concentrations requires reliably measuring extremely small variations in the sensor output signal. In this perspective, gravimetric sensors, such as Quartz Crystal Microbalances (QCMs), have a significant advantage over other sensor technologies (conductance-based sensors, Chem-FET, or optical sensors) since they use frequency as the output variable, which is one of the physical quantities that can be monitored with the highest accuracy [1].

Nevertheless, smart, compact, portable, and low cost electronic noses for multiple gas detection cannot be implemented with bulky and unintegrable QCMs. Even if QCMs have been successfully employed as gravimetric sensors, thanks to their extremely high quality factors, which permit to obtain limits of mass detection in the order of few nanograms, their relatively large volume and their inability to be directly

integrated on silicon render them unattractive for the fabrication of sensor arrays composed of a large number of mechanical elements.

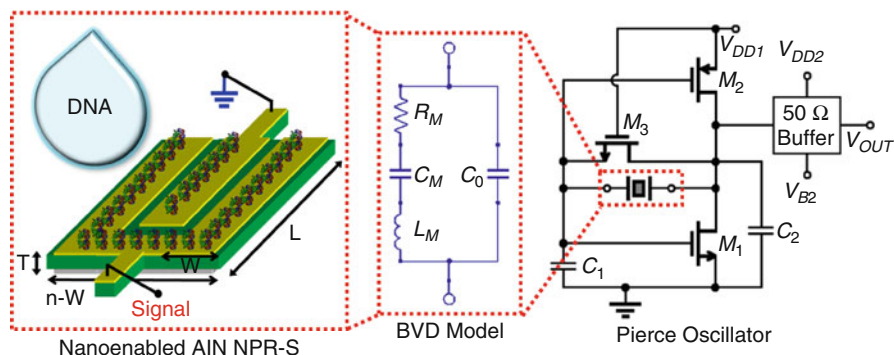
In response to this growing demand for miniaturization, the use of microelectromechanical (MEMS) and nanoelectromechanical (NEMS) resonators for sensing applications has been extensively explored. The scaling of the device dimensions enables the fabrication of ultrasensitive nanomechanical gravimetric sensors that can be arranged in single-chip arrays capable of selectively detecting extremely small concentrations (ppb-ppt) of multiple gaseous analytes.

Different MEMS/NEMS resonator technologies based on electrostatic or piezoelectric transduction [2, 3] or piezoresistive displacement transduction [4] have been proposed. In particular, NEMS resonators have been exploited as transducers for the realization of extremely sensitive gravimetric sensors. Sub-attogram mass resolution has been demonstrated in NEMS cantilevers [4], thanks to the minuscule mass and the relatively high quality factor (Q). Nevertheless the greatly reduced dimensions of these devices render their transduction extremely difficult and require the use of cumbersome, complex, and power-inefficient readout techniques.

In this perspective, optimal sensor performance is attained by synthesizing a transducer that occupies a large area (which facilitates efficient transduction) and is very thin (which allows fabricating low mass devices with ultrahigh sensitivity). Suspended membranes with thickness in the nanometer range are therefore desirable. Flexural plates (instead of beams) are a good compromise [5], but higher frequencies and power handling are achieved in bulk mode acoustic resonators [3, 6].

A particularly representative example of high performance bulk mode acoustic NEMS resonant sensors that will be described here involves exciting high frequencies bulk acoustic waves in a piezoelectric nanoplate (thickness 250 nm) made of Aluminum Nitride (AlN). Such AlN Piezoelectric Nano-Plate Resonant Sensor (NPR-S) technology is not only characterized by high values of sensitivity, due to the reduced mass and high frequency of operation of the nanomechanical resonant element, but it is also associated with low noise performance, due to the combination of high quality factor, Q , and power handling capability of the bulk acoustic wave NEMS resonators. In addition, the AlN nano-plate composing the NPR-S can be

NEMS Resonant Chemical Sensors, Fig. 1 Schematic representation of the fabricated nanoenabled AlN NPR-S connected in a self-sustained oscillator loop (Pierce oscillator). The butterworth van dyke (BVD) model is used to describe the resonator behavior



efficiently actuated and sensed piezoelectrically on-chip solving the fundamental transduction issues associated with electrostatically transduced NEMS resonators and enabling the use of compact and low power CMOS circuits for electronic readout (Fig. 1). The sensor affinity for the adsorption of volatile organic chemicals such as 2,6 dinitrotoluene is enhanced by functionalizing the top gold electrode of the device with a thiol-terminated single stranded DNA sequence (Thiol - 5' CTT CTG TCT TGA TGT TTG TCA AAC 3') and the ability to detect concentrations in the ppt range is demonstrated. Thanks to the unique combination of extremely important features such as high sensitivity, low noise, efficient transduction, small form factor, and CMOS integration capability, the NPR-S represents one of the most suitable technologies to devise high performance portable gas sensors arrays that can be easily incorporated in wireless devices and sensor networks for distributed environmental or personal health monitoring.

NEMS Resonant Chemical Sensor Design

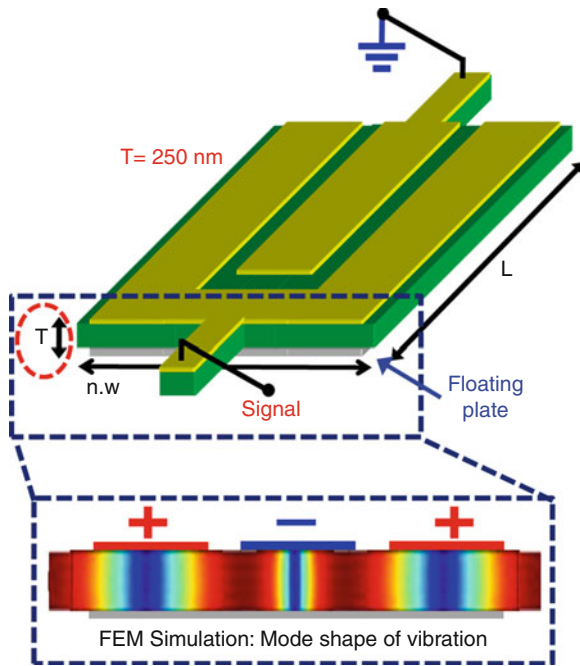
The fundamental metrics that need to be considered for the design of high performance NEMS resonant chemical sensors are expressed in terms of limit of detection (LOD), response time, and selectivity. The LOD is the smallest amount of adsorbed mass per unit area (hence minimum gas concentration given a certain chemical interactive material, CIM) that can be resolved by the sensor. This quantity is inversely proportional to the resonator sensitivity to mass loading per unit area and directly proportional to the frequency fluctuations induced by the noise generated in the resonator and in the readout system. Therefore, a nanomechanical resonator characterized

by a very large sensitivity to mass loading per unit area and a very low noise is the best candidate for the implementation of a high performance NEMS resonant chemical sensor. Response time and selectivity are instead determined by the chemical and physical properties of the CIM employed. A particularly representative example of high performance NEMS resonant chemical sensor design is given by the Aluminum Nitride (AlN) Nano-Plate Resonant Sensor (NPR-S) technology, which will be described in details herein.

The Aluminum Nitride (AlN) Nano-Plate Resonant Sensor (NPR-S) technology involves exciting a high frequency contour-extensional mode of vibration (Fig. 2) in piezoelectric nano-plates (thickness 250 nm) made out of AlN. The capability to excite higher order contour-extensional mode of vibration in micro-scale AlN plates has been demonstrated by the AlN Contour-Mode resonator (CMR) technology [3, 7, 8]. High performance AlN Contour-Mode resonators in the very high frequency (VHF), ultrahigh frequency (UHF), and superhigh frequency (SHF) range (20 MHz–10 GHz) and quality factors, Q_s , between 1,000 and 3,000 have been reported [7–9]. Piezoelectric sensing and actuation of these devices have been successfully employed in the MEMS domain. While the scaling into the nano-domain of the AlN CMR core technology has been exploited for the implementation of AlN Nano-Plate Resonators with unique performances in sensing applications [10–15].

The reduced mass of the mechanical resonator along with the high quality factor values and the power handling capabilities make the AlN NPR-S technology capable of achieving unprecedented values of limit of detection to mass per unit area in the order of $zg/\mu m^2$.

An AlN NPR-S is composed of an AlN nano-plate sandwiched between two metal electrodes (Fig. 2).



NEMS Resonant Chemical Sensors, Fig. 2 Schematic representation of an AlN Nano-Plate Resonator. The inset shows a 2-D FEM simulation of the device mode of vibration

When an AC voltage is applied across the thickness (T) of the device, a contour-extensional mode of vibration is excited through the equivalent d_{31} piezoelectric coefficient of AlN. Given the equivalent mass density, ρ_{eq} , and Young’s modulus, E_{eq} , of the material stack (AlN and metal electrodes) that forms the resonator, the center frequency, f_0 , of this laterally vibrating mechanical structure is univocally set by the period, W , of the metal electrode patterned on the AlN plate and can be approximately expressed as in Eq. 1.

$$f_0 = \frac{1}{2W} \sqrt{\frac{E_{eq}}{\rho_{eq}}} \quad (1)$$

The other two geometrical dimensions, thickness, T , and length, L , set the equivalent electrical impedance of the resonator [3] and can be designed independently of the desired resonance frequency.

The sensitivity to mass per unit area of an AlN NPR-S loaded on its top surface can be expressed as in Eq. 2 [10, 16].

$$S = -\frac{f_0}{2\rho_{eq}T} \quad (2)$$

It is worth noting that the resonance frequency, f_0 , and the thickness, T , of the AlN NPR-S are two independent variables while for QCMs they are intrinsically coupled and cannot be set independently from one another. This is an important and unique advantage of AlN NPR-S. In fact, it permits to set the frequency of operation of these devices according to the specifications of the desired application and independently obtain the required value of mass sensitivity by scaling the device thickness.

Nevertheless, the mass sensitivity of the device cannot be considered the only important parameter for the design of a high performance gravimetric sensor. In fact, the limit of detection, LOD , of the sensor (smallest amount of adsorbed mass per unit area that can be resolved) is defined as in Eq. 3:

$$LOD = \frac{\Delta f_{min}}{S} \quad (3)$$

where Δf_{min} indicates the minimum frequency shift detectable by the sensor readout. Thus, the performance of the sensor can be improved by reducing the thickness, T , of the AlN layer up to the limit at which the smallest value of the ratio, $\Delta f_{min}/f_0$, is preserved. For example, $\Delta f_{min}/f_0$ comparable to thicker films and as low as $9 \cdot 10^{-8}$ has been attained for a nanoscaled device.

Another crucial aspect directly related to the scaling of the device is the complexity, size, and power efficiency of the required sensor readout. In particular, the ability to connect the nanoelectromechanical resonator to a self-sustained oscillator loop for direct frequency readout greatly simplifies the design of multiplexed sensor platforms involving large arrays of resonant devices. Self-sustained, high performance, Pierce-like oscillator circuits based on thick film (2–3 μm) AlN Contour-Mode MEMS Resonator (CMR) have been demonstrated [17]. The primary power loss in these oscillator circuits is due to the motional resistance, R_m , of the resonator [17], whose value is inversely proportional to the device figure of merit, $k_t^2 \cdot Q$ [17], where k_t^2 represents the device electromechanical coupling and Q its quality factor. Therefore, the $k_t^2 \cdot Q$ product of the AlN NPR-S needs to be maintained sufficiently large (>10) to be able to connect it to a compact and low power self-sustained oscillator circuit for direct frequency readout.

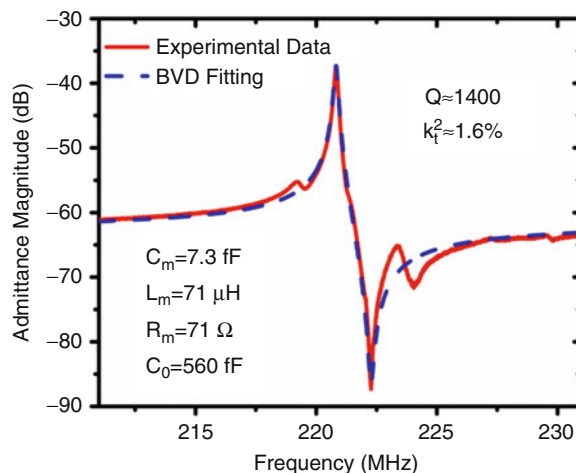
A particular device design that proves the scaling advantages of the AlN NPR-S technology consists of a 220 MHz ($W = 15 \mu\text{m}$) AlN resonant nano-plate whose thickness was set to be, $T = 250 \text{ nm}$, in order to improve the sensitivity of the resonant sensor.

The oscillator circuit topology used as readout for this 250-nm-thick AlN NPR-S is described in more detail in [17]. Briefly, the circuit (Fig. 1) consists of a Pierce oscillator implemented by means of a CMOS inverter biased in its active region. Transistors M1 and M2 form the CMOS inverter while transistor M3 acts as a large resistor to provide biasing in the active region of M1 and M2.

Being functionalization another important aspect in the design of a resonant sensor, an interface layer between the CIM and the VOC has to be selected. In the case of this 250-nm-thick ALN NPR-s, the top gold (Au) electrode was chosen and used to attach thiolated single stranded DNA (HS-ss-DNA) sequences [18]. The capability of single stranded DNA to selectively enhance the adsorption of volatile organic compounds such as dimethyl-methylphosphonate (DMMP, a stimulant for nerve agent sarin) and 2,6 dinitrotoluene (DNT, a stimulant for explosive vapors) has been recently demonstrated [16]. Therefore, in order to enhance the sensor affinity for the adsorption of volatile organic chemicals, the top gold electrode of the AlN NPR-S device was directly functionalized with a thiol-terminated ss-DNA sequence. The capability of employing CIMs with increased effective adsorbing area based on single wall carbon nanotubes (SWNTs) directly grown on the top surface of a micro-scale CMR and functionalized by DNA strands has also been demonstrated [16].

NEMS Resonant Chemical Sensor Performance

The performances of NEMS resonant chemicals sensor can be fully characterized by performing three sets of experiments: (1) Open loop measurements (i.e., frequency response of the nanomechanical resonator in Fig. 3) are performed to extract the device sensitivity to mass loading and extract the electromechanical parameters of the NEMS resonator [10]. (2) The NEMS resonator is then connected to a self-sustained

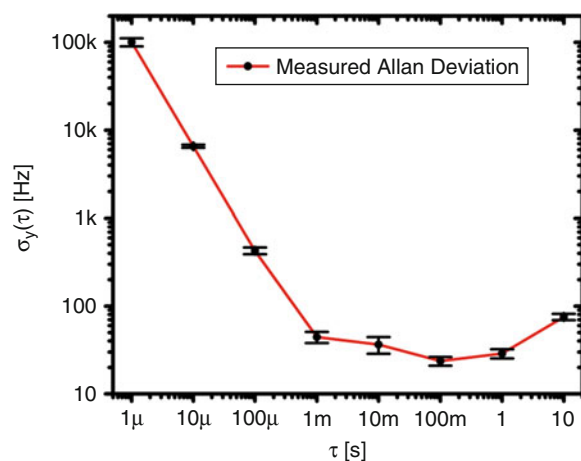
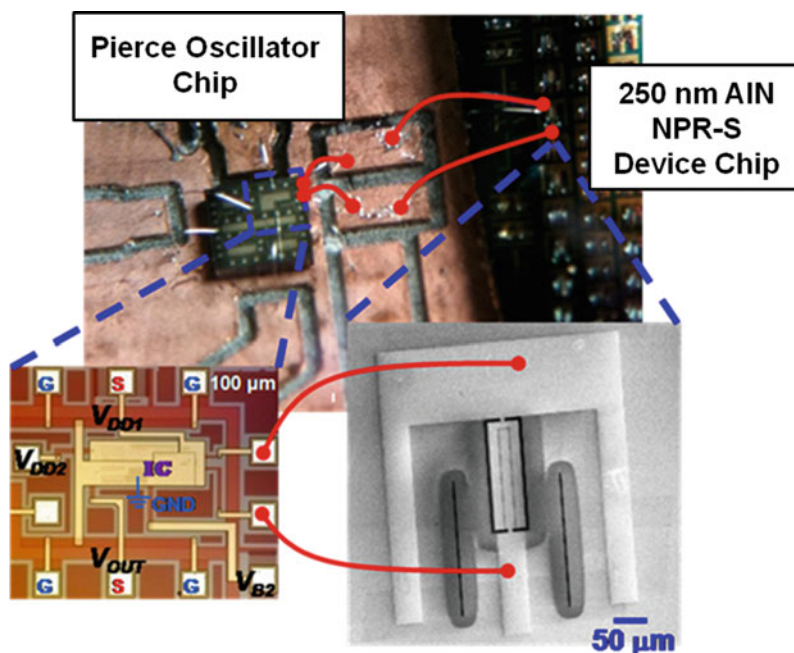


NEMS Resonant Chemical Sensors, Fig. 3 Experimental and fitted (BVD model) admittance curves of the fabricated AlN NPR-S

oscillator circuit (Fig. 4) for direct frequency readout and closed-loop characterization of the resonant sensor performance in terms of noise and power consumption is performed. (3) Finally the sensor is exposed to different concentrations of the target gaseous compounds to verify detection capability. For the ss-DNA-functionalized 250-nm AlN NPR-S described herein, an open loop measurement (frequency response) was first performed and the sensitivity to mass loading was extracted to be approximately $41.6 \text{ KHz} \cdot \mu\text{m}^2/\text{fg}$ [10]. Despite the device volume reduction, high Q ($\sim 1,400$ in air) and high k_f^2 ($\sim 1.6\%$) were attained (Fig. 3) enabling the direct wire bonding of the AlN NPR-S to an oscillator circuit chip fabricated in the ON Semiconductor $0.5 \mu\text{m}$ CMOS process. The 220 MHz oscillator used a 1.67 V supply voltage and $300 \mu\text{A}$ of biasing current, which give a total power consumption of $501 \mu\text{W}$.

In order to estimate the noise limited minimum frequency shift (Δf_{min}) detectable by the readout of a NEMS resonant chemical sensor, the short-term frequency stability of the oscillator output signal is characterized by measuring its Allan deviation, $\sigma_y(\tau)$, [19]. The measured values of Allan deviation for the 220 MHz AlN NPR-S described here are reported in Fig. 5. The minimum value of Allan deviation recorded is approximately 20 Hz and it is achieved

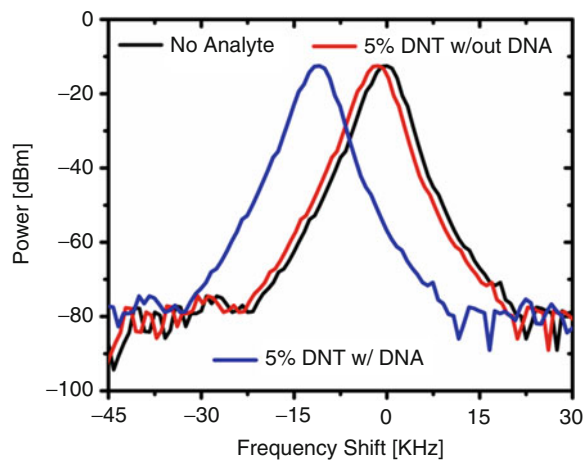
NEMS Resonant Chemical Sensors, Fig. 4 Picture of the AlN NPR-S wire-bonded to the Pierce oscillator chip fabricated in the ON Semiconductor 0.5 μm CMOS process



NEMS Resonant Chemical Sensors, Fig. 5 Measured Allan deviation of the 220 MHz oscillator

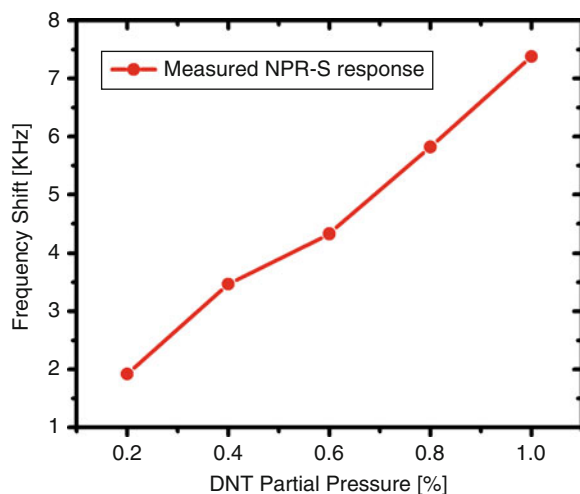
for a gate time, τ , of 100 ms. Therefore, considering a minimum detectable frequency shift, Δf_{min} , of 20 Hz, a limit of detection of mass per unit area of about $481 \text{ zg}/\mu\text{m}^2$ can be estimated. This translates to an extrapolated limit of detection for DNT concentrations of about 15 part per trillion (ppt) [20].

The detection capability of a NEMS resonant chemical sensor is finally tested by exposing the device to



NEMS Resonant Chemical Sensors, Fig. 6 Sensor response to 5% DNT vapor pressure before and after DNA functionalization

known concentrations of a target analyte and monitoring the respective frequency shifts of the oscillator output signal. The ss-DNA-functionalized AlN NPR-S in example was exposed to DNT vapor (generated by bubbling method [10]). The improved affinity of the AlN NPR-S to the DNT vapor, due to ss-DNA functionalization (Fig. 6), enabled the detection of



NEMS Resonant Chemical Sensors, Fig. 7 Response of the nanoenabled AlN NPR-S for different concentrations of DNT ranging between 0.2% and 1% of the saturated vapor pressure

very minute analyte concentrations. The measured sensor response for different concentrations of DNT, ranging between 0.2% and 1% of the saturated vapor pressure, is reported in Fig. 7. Considering that the DNT saturated vapor pressure at room temperature is equal to $5.67 \cdot 10^{-4}$ Torr [20], a concentration of DNT as low as 1.5 ppb was detected by the fabricated nanoenabled AlN NPR-S prototype.

Cross-References

- ▶ [Laterally Vibrating Piezoelectric Resonators](#)
- ▶ [Nanomechanical Properties of Nanostructures](#)
- ▶ [Piezoelectric Effect at Nanoscale](#)
- ▶ [Self-Assembled Monolayers](#)

References

1. Vig, J.R., Filler, R.L., Kim, Y.: Uncooled IR imaging array based on quartz microresonators. *J. Microelectromech. Syst.* **5**, 131–137 (1996)
2. Wang, K., Wong, A.-C., Nguyen, C.T.C.: VHF free-free beam high-Q micromechanical resonators. *J. Microelectromech. Syst.* **9**, 347–360 (2000)
3. Piazza, G., Stephanou, P.J., Pisano, A.P.: Piezoelectric aluminum nitride vibrating contour-mode MEMS resonators. *J. Microelectromech. Syst.* **15**, 1406–1418 (2006)
4. Li, M., Tang, H.X., Roukes, M.L.: Ultra-sensitive NEMS-based cantilevers for sensing, scanned probe and

- very high-frequency applications. *Nat. Nanotechnol.* **2**, 114–120 (2007)
5. Lee, H.J., Park, K.K., Oralkan, O., Kupnik, M., Khuri-Yakub, B.T.: CMUT as chemical sensor for DMMP detection. In: 2008 IEEE International Frequency Control Symposium, pp. 434–439 (2008)
6. Ruby, R., Merchant, P.: Micromachined thin film bulk acoustic resonators. In: Proceedings of the 48th IEEE International Frequency Control Symposium, pp. 135–138 (1994)
7. Rinaldi, M., Zuniga, C., Chengjie, Z., Piazza, G.: Super-high-frequency two-port AlN contour-mode resonators for RF applications. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **57**, 38–45 (2010)
8. Rinaldi, M., Zuo, C., Van der Spiegel, J., Piazza, G.: Reconfigurable CMOS oscillator based on multifrequency AlN contour-mode MEMS resonators. *IEEE Trans. Electron. Dev.* **58**, 1281–1286 (2011)
9. Rinaldi, M., Zuniga, C., Piazza, G.: 5–10 GHz AlN contour-mode nanoelectromechanical resonators. In: IEEE 22nd International Conference on Micro Electro Mechanical Systems, 2009. MEMS 2009, Sorrento, pp. 916–919 (2009)
10. Rinaldi, M., Zuniga, C., Piazza, G.: Ultra-thin-film AlN contour-mode resonators for sensing applications. In: Ultrasonics Symposium (IUS), 2009 IEEE International, pp. 714–717 (2009)
11. Rinaldi, M., Zuniga, C., Duick, B., Piazza, G.: Use of a single multiplexed CMOS oscillator as direct frequency read-out for an array of eight AlN contour-mode NEMS resonant sensors. In: Proceedings of IEEE Sensors 2010, pp. 2666–2670 (2010)
12. Rinaldi, M., Piazza, G.: Effects of volume and frequency scaling in AlN contour mode NEMS resonators on oscillator phase noise. In: 2011 Joint Conference of the IEEE International Frequency Control and the European Frequency and Time Forum (FCS), pp. 1–5 (2011)
13. Rinaldi, M., Zuniga, C., Piazza, G.: ss-DNA functionalized array of ALN contour-mode NEMS resonant sensors with single CMOS multiplexed oscillator for sub-ppb detection of volatile organic chemicals. In: IEEE 24th International Conference on Micro Electro Mechanical Systems (MEMS 2011), pp. 976–979 (2011)
14. Zuniga, C., Rinaldi, M., Piazza, G.: High frequency piezoelectric resonant nanochannel for bio-sensing applications in liquid environment. In: 2010 IEEE Sensors, pp. 52–55 (2010)
15. Zuniga, C., Rinaldi, M., Piazza, G.: Reduced viscous damping in high frequency Piezoelectric resonant nanochannels for sensing in fluids. In: 2011 IEEE 24th International Conference on Micro Electro Mechanical Systems (MEMS), pp. 960–963 (2011)
16. Zuniga, C., Rinaldi, M., Khamis, S.M., Johnson, A.T., Piazza, G.: Nanoenabled microelectromechanical sensor for volatile organic chemical detection. *Appl. Phys. Lett.* **94**, 223122–223122-3 (2009)
17. Zuo, C., Van der Spiegel, J., Piazza, G.: 1.05 GHz MEMS oscillator based on lateral-field-excited piezoelectric AlN resonators. In: Frequency Control Symposium, 2009 Joint with the 22nd European Frequency and Time forum. IEEE International, pp. 381–384 (2009)

18. Herne, T.M., Tarlov, M.J.: Characterization of DNA probes immobilized on gold surfaces. *J. Am. Chem. Soc.* **119**, 8916–8920 (1997)
19. Allan, D.W.: Statistics of atomic frequency standards. In: *Proceedings of the IEEE*, vol. 54, pp. 221–230 (1966)
20. Regulatory Determinations Support Document for Selected Contaminants from the Second Drinking Water Contaminant Candidate List (CCL 2), Chapter 7 (2008)

NEMS Resonant Mass Sensors

- ▶ [NEMS Resonant Chemical Sensors](#)

Neural Activity-Dependent Closed-Loop

- ▶ [Dynamic Clamp](#)

Neural Interfaces

- ▶ [MEMS Neural Probes](#)

N-Methyl-2-pyrrolidone (NMP)

- ▶ [SU-8 Photoresist](#)

Non-conventional Machining

- ▶ [Ultrasonic Machining](#)

Nonlinear Electrokinetic Transport

- ▶ [Concentration Polarization at Micro/Nanofluidic Interfaces](#)

Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures

V. Sai Muthukumar¹, Ramakrishna Podila², Benoy Anand¹, S. Siva Sankara Sai¹, K. Venkataramaniah¹, Reji Philip³ and Apparao M. Rao^{2,4}

¹Department of Physics, Sri Sathya Sai Institute of Higher Learning, Prashanthinilayam, Andhra Pradesh, India

²Department of Physics and Astronomy, Clemson University, Clemson, SC, USA

³Light and Matter Physics Group, Raman Research Institute, Sadashivanagar, Bangalore, India

⁴Center for Optical Materials Science & Engineering Technologies, Clemson University, Clemson, SC, USA

Definition

Nonlinear optics (NLO) studies the behavior of light in *nonlinear media*, whose dielectric polarization (P) responds nonlinearly to the electric field (E) of the light. In general, such nonlinear behavior is observed under highly intense light beams. Several fundamental properties of materials such as nonlinear absorption, nonlinear refraction, harmonic generation, etc., can be probed using NLO spectroscopy.

The field of nonlinear optics (NLO) emerged into prominence more than three decades ago with the development of the laser. These milestone discoveries have not only generated much interest in laser science but also set the stage for future work in nonlinear optics. In recent years, the discovery of various nano-scale materials such as fullerenes, carbon nanotubes, and metallic and other semiconducting nanostructures has allowed nonlinear spectroscopy to progress towards technological applications, and the focus has been on the discovery of materials with higher photonic functionalities and better performance [1, 2]. Knowledge of electronic transitions is essential in understanding the linear and nonlinear optical properties of nanomaterials. Many spectroscopic techniques such as optical absorption, photoluminescence, photoacoustics, and x-ray photoelectron spectroscopy are often employed for the investigation of energy levels

in nanostructures. More significantly, estimation of the nonlinear interaction of light with nanostructures helps one to probe the higher electronic levels, which are not easily accessible otherwise.

In nanostructures, quantum confinement effects and surface plasmon resonance (SPR), in tandem with other crucial factors like morphology, particle size, and chemical environment, can lead to interesting changes in nonlinear properties with many promising applications such as ultrafast optical switching, optical limiting, passive mode locking, frequency up-conversion lasing, optical data storage, multiphoton-based laser microscopy, etc. Interestingly, various nanomaterials exhibit distinct intensity-dependent nonlinear responses for laser excitation that depends on parameters such as laser wavelength, pulse duration, and repetition rate [2]. The physical basis of observed nonlinear optical responses in nanostructures is briefly elucidated in the following sections along with an outline for various experimental techniques used for probing their photophysical properties with special reference to the optical limiting application. Optical limiters are materials which are transparent at lower incident fluences, but attenuate light considerably at higher fluences; thus protecting eye and photosensors from damage due to intense optical sources. Various nanostructures are promising as efficient and broadband optical limiters when compared to conventional materials like organic chromophores and dyes [1, 2]

Nonlinear Optical Processes

In general, the response of a material medium (which is lossless and dispersionless) to input light can be described in terms of the electric polarization \mathbf{P} [1], which is usually expanded in a Taylor series in the electric field amplitude \mathbf{E} as

$$\mathbf{P} = \chi^{(1)}\mathbf{E} + \chi^{(2)}\mathbf{E}\mathbf{E} + \chi^{(3)}\mathbf{E}\mathbf{E}\mathbf{E} \quad (1)$$

where $\chi^{(n)}$ are the susceptibility tensors of order n . Far away from a material resonance, the susceptibilities are predominantly real and the material response is instantaneous, but as resonances are approached, the susceptibilities will become complex, with relatively slower responses. The first term, $\chi^{(1)}$, is responsible for

linear absorption and refraction, while the remaining terms are associated with different light-induced nonlinear effects. Second-order nonlinear effects due to $\chi^{(2)}$ are present only in noncentrosymmetric materials, and these include second harmonic generation and other three-wave mixing processes, optical rectification, and the electro-optic Pockels effect. Nonlinear optical effects involving the third-order susceptibility $\chi^{(3)}$ include phenomena such as third harmonic generation and four-wave mixing processes, the electro-optic Kerr effect, two-photon absorption, and Raman amplification.

For a nonlinear optical medium, the refractive index and absorption coefficient will be functions of the incident light intensity. For an isotropic centrosymmetric material (where the lowest order nonlinearity is cubic), the intensity-dependent nonlinear refractive index is given by [1–4]

$$n = n_0 + n_2 I \quad (2)$$

and the intensity-dependent nonlinear absorption coefficient is given by

$$\alpha = \alpha_0 + \beta I \quad (3)$$

where n_2 is the nonlinear refractive index coefficient, and β is the nonlinear absorption coefficient. n_0 and α_0 are the linear refractive index and linear absorption coefficient, respectively. I is the intensity of the incident light which is proportional to \mathbf{E}^2 . The third-order nonlinear susceptibility can be very generally written as

$$\chi^{(3)} = \chi_R^{(3)} + \chi_I^{(3)} \quad (4)$$

where the real part $\chi_R^{(3)}$ accounts for effects that are transient and consume no energy (like the optical Kerr effect, molecular orientation, electrostriction, etc.). The imaginary part $\chi_I^{(3)}$ accounts for processes that involve losses such as two-photon absorption. Thus, $\chi^{(3)}$ (both the real and imaginary parts) serves as a figure of merit for investigating potential novel materials for photonic applications.

It may be noted that ultrafast laser pulse excitation (typically in the femtoseconds regime) provides a way to measure pure-electronic nonlinearities exclusively. To observe longer timescale processes such as excited

state absorption and thermal lensing, longer laser pulses (typically in the nanosecond regime) are used.

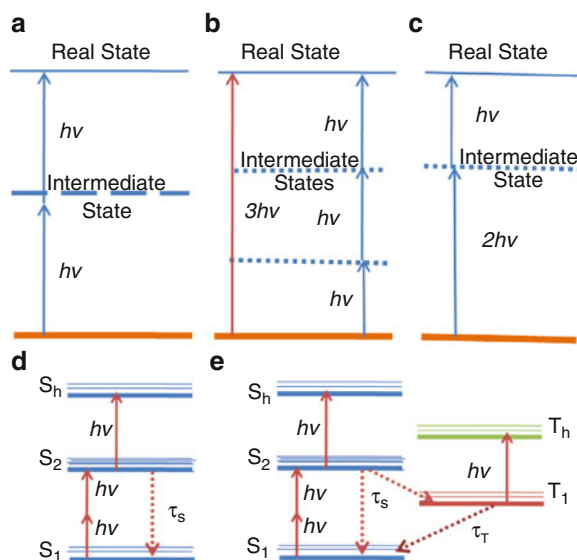
Nonlinear Refraction (NLR)

The dominant physical mechanism driving nonlinear refraction in a given medium can vary depending on the material, optical wavelength, irradiance, timescale of the light-matter interaction, and/or geometry. In addition to bound electronic nonlinearities, which are instantaneous, effects like intramolecular motion, molecular reorientation, electrostriction, induced population change, and photothermal effects may become responsible for nonlinear refractive index in a medium. Depending on the sign (positive or negative) and magnitude of the value of “ n_2 ” in Eq. 2, a laser beam passing through the nonlinear medium may be self-focused or defocused. Such focusing (or defocusing) arising from the real part of the third-order nonlinear susceptibility ($\chi_R^{(3)}$) is known as the optical Kerr effect (Kerr lens effect). The instantaneous nature of Kerr lensing is used in several applications including optical switching and laser mode-locking, especially in the ultrafast regime. At very high intensities, the optical Kerr effect may tend to saturate. For relatively shorter optical pulses with very high intensity, electronic nonlinearities may, on the other hand, lead to self-steepening and supercontinuum generation effects.

Nonlinear refraction can also occur due to thermal effects (usually referred to as thermal lensing), which typically build up with a relatively long response time of a few milliseconds. Thermal nonlinearity can be dominant for continuous-wave (CW) laser excitation and significant for nanosecond laser pulse excitation. The variation of the refractive index with temperature can be described mathematically as $\tilde{n} = n_o + (dn/dT)\tilde{T}_0$, where n_o is the linear refractive index, \tilde{T}_0 is the laser-induced change in temperature, and dn/dT , the temperature-dependent refractive index. The thermal nonlinear index of refraction can be determined using a closed aperture Z-scan measurement, employing either the thermal lens model (TLM) or Sheik-Bahae’s formalism [3] for analysis.

Nonlinear Absorption

Nonlinear absorption refers to the change in the transmittance of a material as a function of input light fluence (given in units of J/m^2) [1]. As given above, the nonlinear absorption coefficient of a third-order nonlinear medium can be written as $\alpha = \alpha_0 + \beta I$,



Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 1 Various nonlinear absorption schemes involving 2PA, 3PA and ESA processes. (a) depicts an ESA process while (b) shows 3PA and (c) shows 2PA. The dotted lines represent virtual state and solid lines represent real state. (d) and (e) show three- and five- level ESA models respectively, which have been extensively used in literature (see ref. [1] for details). (a) 2PA virtual intermediate state (dashed line). (b) 3PA virtual intermediate state (dotted line). (c) ESA real intermediate state (dotted line)

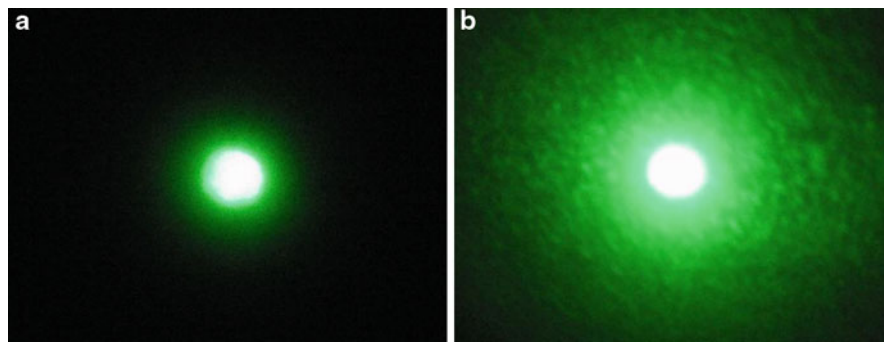
where β is the nonlinear coefficient of absorption. Far away from a material resonance, β usually reduces to the two-photon absorption coefficient. However, near a material resonance, two-step phenomena like excited state absorption (ESA) and free carrier absorption (FCA) may also contribute to nonlinear absorption. Such phenomena are collectively referred to as reverse saturable absorption (RSA). An overview of these phenomena is given below.

Multiphoton Absorption

Many semiconducting and metallic nanomaterials exhibit multiphoton absorption (MPA). In this discussion, MPA is limited to two- and three-photon absorption (2PA and 3PA, respectively) processes. 2PA involves the simultaneous absorption of two photons to promote an electron from an initial state to a final state through a virtual intermediate level (Fig. 1). 3PA has a lower transition probability and involves three photons and two intermediate states. Since the intermediate level is virtual, energy need be conserved only

Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures,

Fig. 2 Panel (a) shows the transverse profile of the incident laser beam, while panel (b) shows the beam profile of the scattered beam after traversing through the sample [6]



in the final state. In contrast to reverse saturable absorption, MPA is an instantaneous process occurring in femtosecond timescales. The mechanism of 2PA can be viewed as analogous to a three-level RSA process where the lifetime of the intermediate state approaches zero and the ground state absorption is extremely low (high linear transparency). Two- and three-photon-absorption-based nanomaterials can be used for applications like frequency conversion lasing, optical limiting, pulse stabilization, and reshaping [1].

Excited State Absorption

In systems like C_{60} , the excited state absorption (ESA) cross section is greater than the ground state absorption cross section, which can lead to an enhanced absorption at higher input intensities. Typically, the ESA process is much stronger than 2PA, and to indicate the collective contribution of 2PA and ESA, the term “effective 2PA” is sometimes used in literature. When excited with nanosecond laser pulses, which are relatively long compared to typical excited state lifetimes of nonfluorescent materials, ESA will contribute significantly to nonlinear absorption. For instance, the effective 2PA cross section of a medium measured under nanosecond (ns) pulse excitation has been found to be about two orders of magnitude larger than the intrinsic 2PA cross section measured under femtosecond (fs) excitation [5]. In the case of intrinsic 2PA, the nonlinear absorption coefficient (β) is independent of the input intensity, but in the case of effective 2PA, the nonlinear absorption coefficient (β_{eff}) varies with input intensity. Relevant examples are discussed in section “Experimental Methods”. Another interesting case is that of TPA-induced ESA, where a strong 2PA process considerably increases molecular

populations in an excited state, which then absorbs another photon by an ESA process.

Free Carrier Absorption

In semiconducting materials, charge carriers are generated by one-photon as well as two-photon excitations. These electrons or holes can be excited to still higher or lower states in the conduction or valence bands by absorbing additional photons. Such carrier-assisted nonlinear absorption processes are known as “free carrier absorption,” which is analogous to excited state absorption in molecular systems [2]. One difference is that free carrier absorption processes are usually independent of the input laser pulse width (assuming shorter timescales for carrier diffusion and recombination). In addition, they are insensitive to particle size and geometry and can be found in both solids and nanoparticle dispersions covering a broad range of temporal and wavelength ranges.

Induced Thermal Scattering

An important mechanism by which laser light transmitted through a medium suffers losses, particularly in the nanosecond excitation regime, is induced thermal scattering (ITS). Processes including the formation of microplasma and solvent microbubbles, and thermally induced local refractive index gradients lead to ITS [4]. Since the formation of thermal scattering centers takes several nanoseconds, these phenomena generally do not affect the transmission of ultrafast laser pulses through the medium. Mechanistically, ITS leads to a spatial redistribution of the incident beam (as shown in Fig. 2), resulting in spatial intensity modulation, which can be efficiently used for optical limiting if appropriate geometries are employed [2, 4].

Experimental Methods

In order to evaluate the contribution of various intensity-dependent nonlinear processes in any material, different methods such as second harmonic generation, four-wave mixing, ellipse rotation, nonlinear interferometry, and Z-scan are often employed [1, 2]. In recent years, Z-scan and four-wave mixing methods have been extensively adopted and applied by various research groups across the world to evaluate the nonlinear optical properties of various bulk and nanostructured materials.

Four-Wave Mixing

Four-wave mixing (FWM) is a phase-sensitive nonlinear effect arising from third-order optical nonlinearity. The concept of three optical fields interacting to produce a fourth optical field is central to the description of all four-wave mixing processes. Along with Z-scan, this is one of the most widely used techniques to measure the third-order susceptibility $\chi^{(3)}$ of materials [2, 3]. In a dielectric medium, when two light beams are present, each beam independently polarizes the dielectric. The interference between polarized fields then results in harmonics at the sum and difference frequencies. When a third field is incident on the dielectric, it too drives the polarization and will beat with both the other input fields as well as the sum and difference frequencies. This beating with the sum and difference frequencies gives rise to the fourth field in four-wave mixing.

The FWM process is essentially of two types: nondegenerate and degenerate (DFWM). In the former, the optical excitation sources are of different wavelengths, while in the latter, their wavelengths are identical. The strength of the fourth beam is dependent on a coupling constant that is proportional to the effective $\chi^{(3)}$, and hence, measurement of the observed signal will yield information about the $\chi^{(3)}$ tensor components of the medium. Furthermore, DFWM can be employed in the backward (phase conjugate) and BOXCARS configurations, with the choice being made through consideration of the experimental conditions and requirements. In the backward geometry, two waves travel in the backward direction and two travel in the forward direction, whereas in the BOXCARS geometry, all the four waves travel in the forward direction. Using different polarizations for the three beams, all the components of the $\chi^{(3)}$ tensor can

be measured. In terms of the intensity-dependent refractive index, the DFWM process is considered to be the interference of two input beams leading to a spatial periodicity resulting in the formation of a grating from which the third beam scatters, generating the fourth wave.

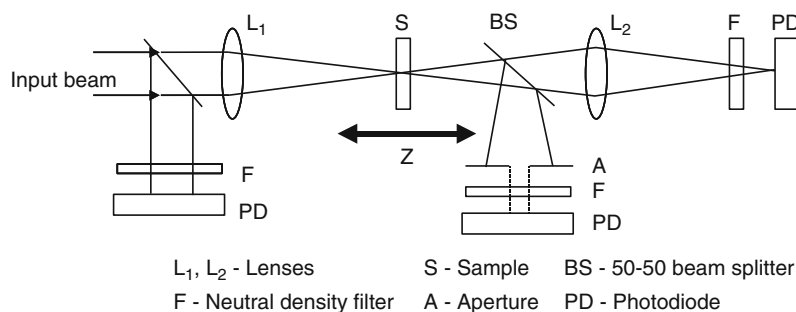
FWM can be used to probe either one-photon resonances or two-photon resonances in a material by measuring the resonant enhancement as one or more of the frequencies are tuned. By tuning the frequencies to multiple resonances in the material, excited state cross sections, lifetimes, and line widths can be measured. One of the main disadvantages of FWM is its sensitivity to alignment, requiring all three coherent beams to simultaneously overlap within the sample in space and time.

Z-Scan

The Z-scan technique is a sensitive and simple method employed to measure nonlinear absorption and nonlinear refraction. It has been used extensively to study the optical nonlinearity of various materials like semiconductors, nanocrystals, semiconductor-doped glasses, liquid crystals, organic materials, and biomaterials. Introduced by Sheik-Bahae et al. [3] in the early 1990s, this widely used technique has many advantages over other nonlinear spectroscopic methods. In a typical Z-scan, the light-induced change in transmittance (ΔT) of a medium due to optical nonlinearity is measured as a function of input light energy density (fluence) or intensity. A continuous variation of the input fluence is achieved by translating the sample under study through the focal region of a focused laser beam (the sample position is taken as z with $z = 0$ being the focal point for the incident laser beam, hence the name “Z-scan”). Consequent increase and decrease of the intensity incident at the sample position lead to wave front distortions (created by nonlinear optical effects in the sample). There are two types of Z-scan techniques, namely, the “closed aperture” Z-scan and the “open aperture” Z-scan.

In the closed aperture Z-scan, which is used for studying nonlinear refraction (NLR), the transmitted beam is passed through an aperture placed in the far field and then measured by a detector, for different values of sample position z . When the medium is far before the focal plane, no self-lensing occurs. As the medium approaches the focal plane, the high intensity begins to induce a lensing effect in the medium. For

Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 3 A conventional Z-scan layout for simultaneous open aperture and closed aperture measurement



a negative nonlinearity, this lens tends to collimate the beam, thereby increasing the transmittance through the aperture. At the focal plane, even though the intensity is highest, the influence of the induced lens is minimized, resulting in a transmittance comparable to the linear transmittance. This is similar to placing a thin lens at the focus of a beam; this results in a minimal effect on the far field beam pattern. As the sample is translated beyond the focal plane, the negative lens tends to increase the beam divergence, resulting in a decrease in the aperture transmittance. As the medium is translated still farther from focus, the intensity again becomes weak enough that the induced lensing is negligible. This sequence results in a change in transmittance with a characteristic peak followed by a valley, which is symmetric about the focal plane ($z = 0$). For a positive nonlinearity, the pattern consists of a valley followed by a peak. Thus, a simple Z-scan experiment can give both the sign and magnitude of the optical nonlinearity in a material. In both these cases, a purely refractive nonlinearity is considered, assuming that no absorptive nonlinearities are present. The sensitivity to nonlinear refraction results from the vital role played by aperture A (Fig. 3). If aperture A is removed, then the Z-scan is sensitive only to nonlinear absorption (NLA), and this configuration is commonly referred to as the open aperture Z-scan. The absorptive nonlinearity will be a maximum at the focal plane, where the intensity is highest. Thus, the open aperture scheme is employed to characterize materials that exhibit NLA, such as excited state absorption (ESA), two-photon absorption (TPA), saturable absorption (SA), etc.

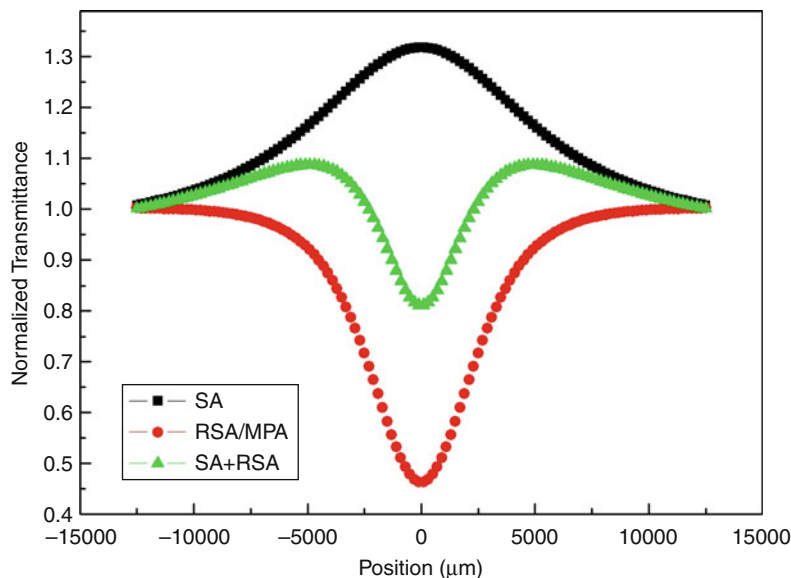
Even though closed aperture Z-scan is used to study only NLR, in practice, the closed aperture Z-scan curve may contain a contribution from NLA also, depending on the sample. To obtain the NLR coefficient in such a case, it is sufficient to normalize the

closed aperture data using the open aperture data before analysis. Although the open aperture Z-scan technique is adequate to detect the presence of nonlinear absorptive effects, it cannot singularly identify the nature of an absorptive nonlinearity. For example, at the outset, it cannot differentiate a 2PA signature from a two-step excitation signature. Figure 4 shows typical Z-scan curves simulated for the MPA, SA, and RSA processes. The Z-scan curves for SA (peak) and RSA (valley) have shapes, which are complementary to each other. It should be noted that the Z-scan curves for RSA and MPA look very similar to each other. To confirm the exact nature of the process, it is often necessary to know the linear absorption properties of the sample from a linear absorption spectrum taken using a spectrophotometer. Z-scans performed using ultrafast laser pulses (where RSA effects are a minimum) will also be instructive in this regard.

A recent variation in the Z-scan technique is the white-light continuum Z-scan, where a supercontinuum is used as the light source for doing the Z-scan experiment. The “supercontinuum” is an intense, ultrafast, broadband light pulse, with wavelengths spanning the UV to the near IR region, generated when an ultrafast laser pulse is passed through certain transparent NLO media like heavy water. Conversion of the spectrally narrow pump pulse into the resultant broad spectrum involves the interplay of self-phase modulation and self-focusing of the pulse, due to an intensity-dependent refractive index of the medium. Since the typical supercontinuum spectrum covers the whole visible range, it appears white, and hence it is also known as a “white-light continuum.” In the conventional Z-scan technique, nonlinearity can be measured only for one wavelength in one measurement. To determine the spectral dispersion of the nonlinearity, a tunable laser source has to be used, and separate

Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures,

Fig. 4 Simulated Z-scan curves for the saturable absorption (SA), reverse saturable absorption (RSA), and multiphoton absorption (MPA) processes



Z-scans need to be done at all the required wavelengths. Obviously, this approach is time consuming and tedious. The idea of using a strong white-light continuum (WLC) as the light source to measure the dispersion of the nonlinearity in a single Z-scan experiment was therefore proposed by Van Stryland's group [7]. Their "WLC Z-scan" allows for rapid, broadband characterization of degenerate NLA and NLR.

Nonlinear Optical Properties of Nanomaterials

Nonlinear optical measurements in low-dimensional materials continue to draw considerable attention since the nonlinear behavior of the complex refractive indices lead to unique optical phenomena [1–3, 7]. Furthermore, the quantum confined nature of the charge carriers, surface effects, and sharp density of states in low-dimensional materials such as nanoparticles, nanowires, nanotubes, and nanorods introduces various new fundamental physical phenomena that can influence the NLO properties. For example, surface effects can lead to ferromagnetic behavior in nanostructured ZnO whereas bulk ZnO is diamagnetic [8], and a semimetal to semiconductor transition occurs in the low dimensional forms of bismuth due to quantum confinement effects. Hence, one may utilize these interesting properties of nanostructures to tailor new materials for NLO applications. For example,

optical switching, saturable absorption, and optical limiting properties of one-dimensional nanostructures have shown considerable promise compared to bulk materials. In this section, several interesting NLO phenomena which can be observed in various nanostructures using the Z-scan technique are discussed. Although the discussion is limited to the Z-scan technique, various experimental techniques used for probing the NLO properties of popular nanostructures are summarized in Table 1.

Carbon Nanotubes

One-dimensional carbon nanotubes (CNTs) have revolutionized various fields of science and technology since their optical and electrical properties allow one to selectively modify and tailor nanotube properties for specific applications. Initial reports on the nonlinear optical limiting performance of multiwalled carbon nanotubes (MWNTs) highlighted the superior performance of nanotubes compared to corresponding properties of C_{60} and carbon black dispersions. Subsequently, it spurred interest in investigating NLO properties of functionalized CNT-based novel nanomaterials. Moreover, the broadband optical absorption profile of carbon nanotubes provides a unique advantage over other optical materials including fullerenes. Theoretical predictions by Margulis et al claimed that the third-order nonlinear optical susceptibility for MWNTs exceeds that of fullerene molecules by two orders of magnitude [13]. Such

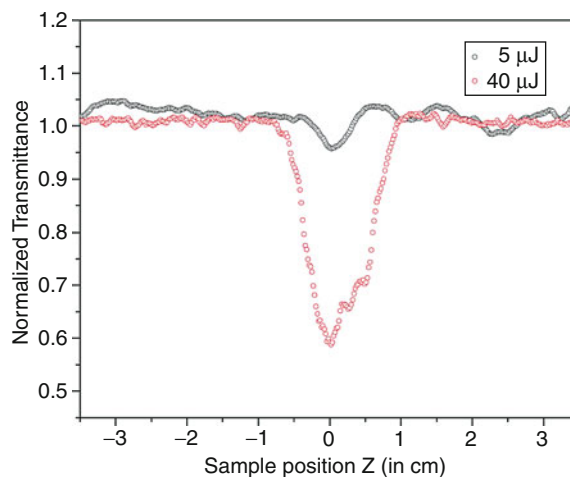
Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Table 1 Different nonlinear optical processes observed in different nanostructured materials (see text for details)

Type of nanostructures	Cause for NLO properties	Technique used for probing	Reference
<i>C₆₀ Dispersions</i>	NLA (2PA)	Pump-probe techniques	Tutt et al. [4]
	NLR, RSA, NLS, thermal NLO	Z-scan	
<i>C₆₀ Films</i>	NLR, NLA(SA)	Z-scan	Chin et al. [9]
<i>SWNT films</i>	NLR, NLA (FCA)	Z-scan, pump-probe techniques	
<i>SWNTs suspensions</i>	NLR, NLS	Z-scan	
<i>MWNT suspension</i>	NLS	Z-scan and optical limiting measurements	Pan et al. [10]
<i>Metal nanowires</i>	NLS	Optical limiting measurement	
<i>Gold nanoparticles dispersions</i>	NLA, NLS	Z-scan and optical limiting measurements	
<i>CdSe-based quantum dots suspensions</i>	NLR, NLA	Z-scan and optical limiting measurements	Ganeev et al. [12]

enhancement in third-order NLO values arises from the combined effect of π -electron transitions, which are dominant over interband/intraband transitions. Thus, the modification of the π -electron cloud via functionalization or doping provides an efficient means for engineering novel materials with high nonlinear optical susceptibility.

The optical limiting performance of MWNTs at 532-nm excitation wavelength with a 7-ns pulsewidth from a second harmonic generator of a high-power Nd:YAG system (quanta systems) is depicted in Fig. 5. The laser excitation was used in the conventional open aperture Z-scan setup (as described in section “Experimental Methods”). Optical limiting response was studied by varying parameters such as the solvent used for nanotube suspension and the incident energy. The MWNTs were suspended in solvents such as deionized water, 2-propanol, and ethanol using sonication (for about ~ 10 min).

The transmittance curves of MWNT suspensions, as shown in Fig. 5, at different energies indicated improved limiting at higher energies. As expected, in the case of higher energies, early onset of limiting was observed (greater width of transmittance curve at 40 μ J). The main mechanism responsible for the observed optical response is nonlinear-induced scattering, which was evident as the MWNTs suspension translated across the focal plane in the Z-scan experiment. The observed NLO response of MWNTs arises due to “microbubble” or “microplasma” formation that leads to the observed nonlinear scattering [4].

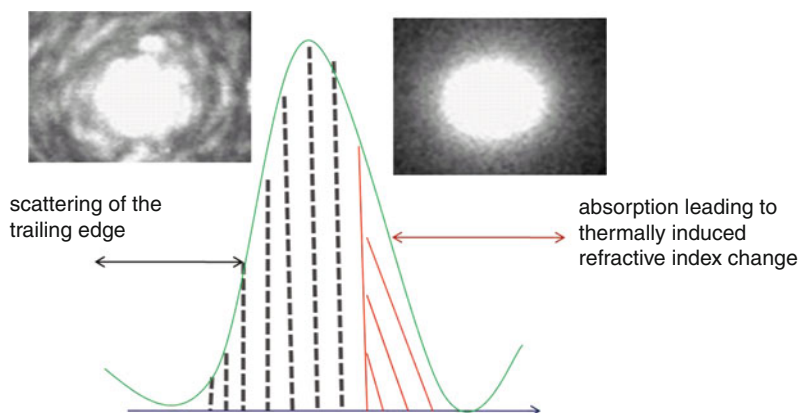


Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 5 Open aperture Z-scan signal for MWNTs suspensions dispersed in water and excited at two different incident energies, 5 μ J and 40 μ J; $\lambda_{\text{excitation}} = 532$ nm

The leading edge of the pulse is absorbed by the MWNTs and is nonradiatively transferred to the surrounding medium within a few nanoseconds, which causes thermal inhomogeneities in the suspension centered around the nanotubes. Depending on the fluence (F_0) and thermo-optic coefficient (dn/dT), a local change in the temperature results in a local refractive index gradient (Δn) since $\Delta n = (dn/dT)(F_0 \alpha / 2\rho C_v)$, where α is absorption coefficient, ρ is the density, and C_v is the specific heat of the sample [14]. The trailing edge of the optical pulse, therefore, “sees”

Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures,

Fig. 6 Scheme of nonlinear scattering or induced scattering leading to spatial redistribution of the beam as the intense laser pulse interacts with the MWNTs. The *right inset* shows the mm size laser spot before incidence, and the *left inset* shows the visible scattering after incidence

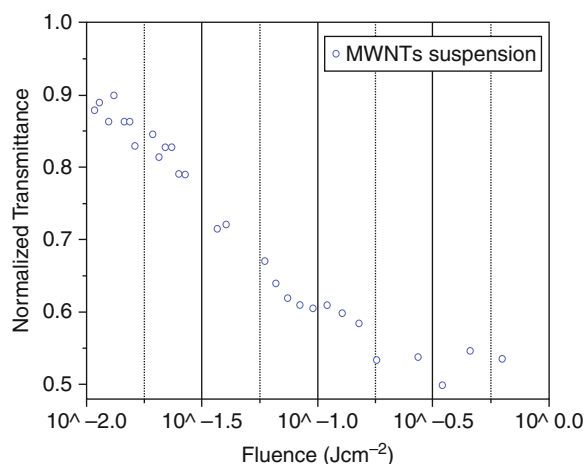


a nonuniform optical medium of scatterers created by the leading edge of the pulse. Thus, the trailing edge gets scattered leading to a spatial redistribution of pulse energy (as shown in Fig. 6)

Since the thermodynamical parameters of the solvent media surrounding the nanotubes play a crucial role in the formation of a scattering center, the optical limiting response was also found to be solvent dependent [9]. The threshold fluence for optical limiting (at which the transmittance falls to 50% its initial value) for MWNTs was determined to be $\sim 0.5 \text{ J cm}^{-2}$ (Fig. 7).

Bi Nanorods

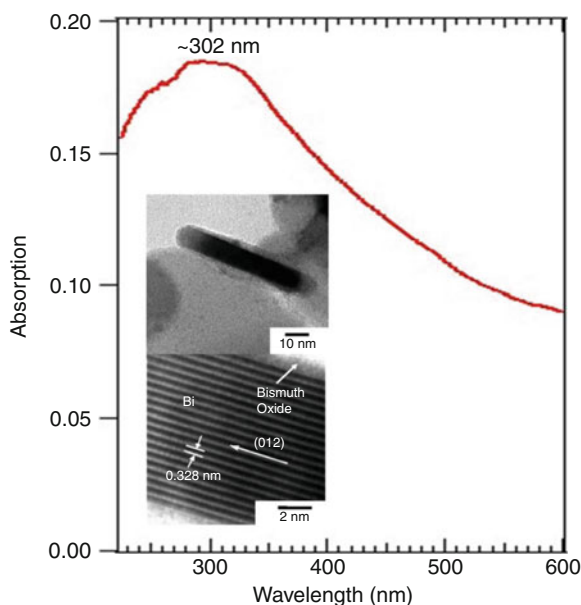
Bismuth is a semimetal in its bulk form and can be prepared in several different nanostructured forms, including thin films, nanowires, nanorods, and quantum dots [15]. Bi is an attractive material for studying how nanostructuring can alter the optical properties of metals. For example, because of the small effective mass components (in the anisotropic effective mass tensor) of Bi, quantum confinement effects are observed in nanowires with diameters as large as 40 nm at low temperature (77 K) [15]. In addition, since the mean free path is long in bismuth ($\sim 250 \text{ nm}$ at room temperature), the finite size effects in a Bi crystal are more pronounced. These effects, which are observed in the tens of nanometers range in Bi, occur in other metals only at much smaller length scales and, thus, are likely to be more difficult to probe experimentally. Therefore, understanding these effects in bismuth can assist in the prediction of the optical properties of other nanostructured materials.



Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 7 Nonlinear optical response of MWNTs suspension at 532-nm excitation

A semimetal to semiconductor transition was predicted in the low-dimensional forms of bismuth due to quantum confinement effects. When such a transition occurs, it is expected that several material properties undergo a radical change. For example, the thermoelectric figure of merit is poor in bulk Bi due to nearly equal and opposite contributions from electrons and holes, but is high, ~ 2 , in small-diameter Bi nanowires. Similar enhancements in NLO properties were also observed for Bi nanorods [16].

Figure 8 shows the transmission electron microscope (TEM) images for Bi NRs synthesized using pulsed vaporization techniques [16]. Unlike bulk Bi, bismuth nanoparticles and nanorods exhibit a broad surface plasmon (SP) peak at ~ 281 and $\sim 302 \text{ nm}$



Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 8 UV-Visible absorption of Bi nanorod suspension in the range of 200–600 nm. The broad peak at 302 nm is due to surface plasmons. TEM images of isolated Bi nanorods encased inside a thin oxide layer (*top inset*) and the crystalline planes (012) of trigonal Bi (*bottom inset*)

(4.4 and 4.1 eV), respectively. When irradiated by 532 nm (2.33 eV) laser pulses (pulse width ~ 7 ns), Bi NRs can absorb energy via interband and intraband transitions, free carrier absorption, and two-photon excitation. The presence of the SP peak makes the absorption a “near-resonant” process. When Bi NRs are suspended in any solvent, the absorbed energy is transferred to the surrounding medium through nonradiative de-excitation, resulting in localized heating and thus scattering centers. In the case of a Z-scan experiment, the leading edge of the pulse creates these scattering centers and the trailing edge of the pulse gets scattered (see Fig. 6). Such a mechanism has been observed for other nanostructured systems such as SWNTs and MWNTs (see section “Carbon Nanotubes”).

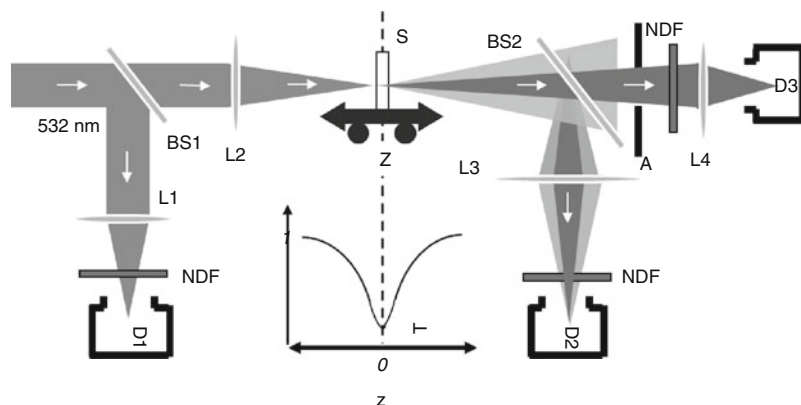
Previously, Sivaramkrishnan et al. used visible and IR laser pulses from a Nd:YAG laser (quanta system, 7 ns, 10 Hz), at both 532 and 1,064 nm, to measure the optical limiting performance of Bi nanorod suspensions by the Z-scan technique at room temperature [17]. In order to separate the contributions of NLA and NLS, a modified Z-scan setup (Fig. 9) proposed

by Jourdrier et al. was employed. In the modified setup, the beam is split 50:50 (using BS2) to reach two detectors D2 and D3 in order to measure the transmittance T of the sample with detector D1 as a reference. A “2” wide lens (L4) collects light from both the scattered and unscattered parts making D2 insensitive to transmittance changes due to spatial redistribution. This ensures that the collection efficiency is $\sim 90\%$. Thus, D2 records changes in transmittance solely due to absorptive effects, and the nonlinear extinction coefficient measured by the Z-scan curve is only due to contributions from NLA. On the other hand, D3 detects only the unscattered part since aperture (A) can be adjusted to prevent the scattered part from reaching D3. Therefore, D3 records a reduced transmittance due to NLS and NLA, unlike the case of D2. The Z-scan curves for the Bi nanorods suspended in chloroform (linear optical transmittance of 60%) in a 1-mm cuvette made of fused silica (under a 55- μJ , 532-nm, 7-ns excitation) are shown in Fig. 10.

Figure 10a shows the Z-scan curves recorded by detectors D2 (NLA alone) and D3 (NLA and NLS). The depth of the valley in the Z-scan curve is a direct measure of the extent of optical limiting and is different in the D2 and D3 curves with dips of 0.65 and 0.25, respectively. Since D2 records the contribution due to NLA alone and the D3 curve also has contributions from NLS, it is clear that NLS induced by NLA makes a significant contribution to the total extinction. In fact, the contribution of NLS to optical limiting is greater than that due to NLA. Interestingly, in the case of 1,064-nm excitation (55 μJ , 7 ns), the D2 and D3 curves overlap (Fig. 10b), and the minimum transmittance is ~ 0.75 in both the cases. Thus, NLS is absent at 1,064 nm, and only NLA makes a contribution to the total extinction observed at this wavelength. At 532 nm, however, both NLA and NLS are responsible for the observed optical limiting unlike the 1,064-nm case. This is very unlike the behavior of carbon black and SWNT/MWNT suspensions, where NLS dominates at both 532- and 1,064-nm excitations (see section “Carbon Nanotubes”).

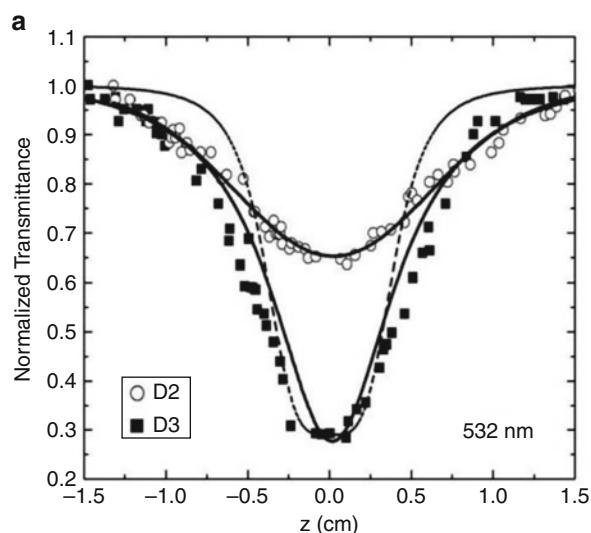
As shown in Fig. 8, Bi nanorods used in this study contain a native oxide layer. Z-scan experiments performed using 532-nm excitation on bulk Bi and bulk Bi_2O_3 did not display any nonlinearity (inset in Fig. 10b).

However, ball-milled Bi_2O_3 powder exhibited strong nonlinearity when excited at a higher fluence



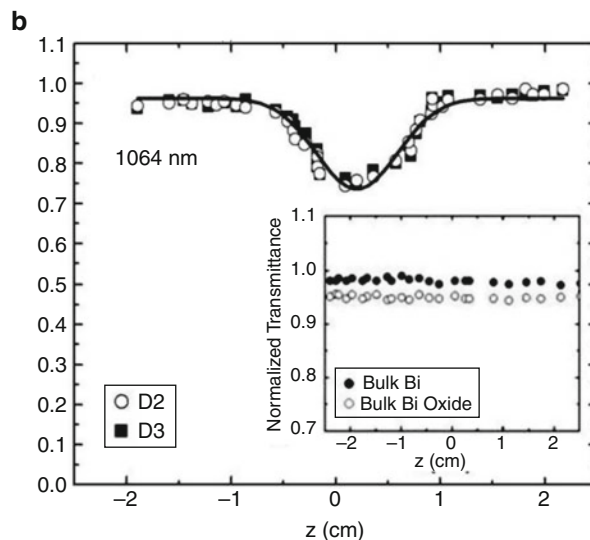
Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 9 Schematic of the experimental setup used for the NLO studies of Bi nanorods. A combination of lenses (L1–L4), beam splitters (BS1) and (BS2), neutral density filters (NDFs),

aperture A, and detectors (D1–D3) are used to measure contributions to optical limiting from the scattered and unscattered components. Bi nanorod suspension is positioned at S and translated through $z = 0$



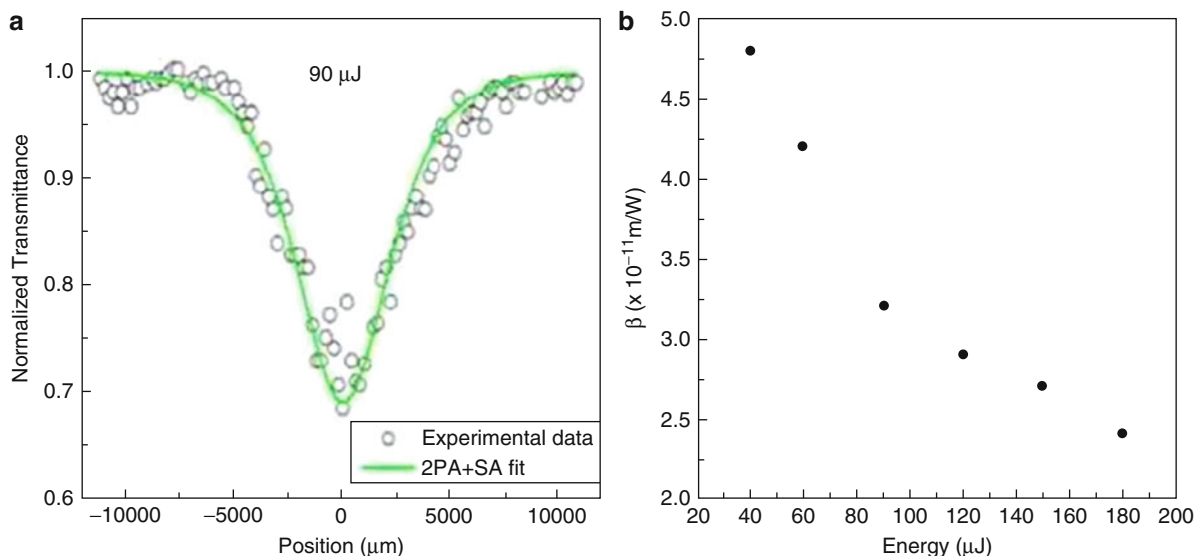
Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 10 (a) Normalized transmittance measured at 532-nm excitation by detectors D3 (nonlinear absorption and scattering) and D2 (nonlinear absorption only) for an input energy of 55 μJ . The minimum transmittance recorded by D3 is less than that of

(90 μJ ; Fig. 11a). Similar to the case of Bi nanorods, the predominant cause for such nonlinearity was found to be two-photon absorption (2PA). Additionally, the saturation of the ground state absorption was also observed for ball-milled Bi_2O_3 samples. Thus an effective nonlinear absorption coefficient $\alpha(I)$, given by $\alpha(I) = \frac{\alpha_0}{1+I/I_s} + \beta I$ was considered for the fitting of the



D2 as it has contributions due to both scattering and absorption. (b) Same as in (a) using the 1,064-nm excitation for an input energy of 55 μJ . The overlapping curves suggest that nonlinear scattering is absent when excited with 1,064 nm in contrast to the 532-nm case

data shown in Fig. 11a, where I_s is the saturation intensity (intensity at which the linear absorption drops to half its original value) and β is the 2PA coefficient. Since Bi nanorods possess a high surface area (hence more of a surface oxide layer), it is possible that the nonlinear response discussed in Fig. 10 arises from the native surface oxide layer rather than core Bi



Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 11 (a) Z-scan data from ball-milled Bi_2O_3 samples

obtained with 532-nm, 7-ns pulses. (b) The variation of 2 PA coefficients as a function of incident laser energy

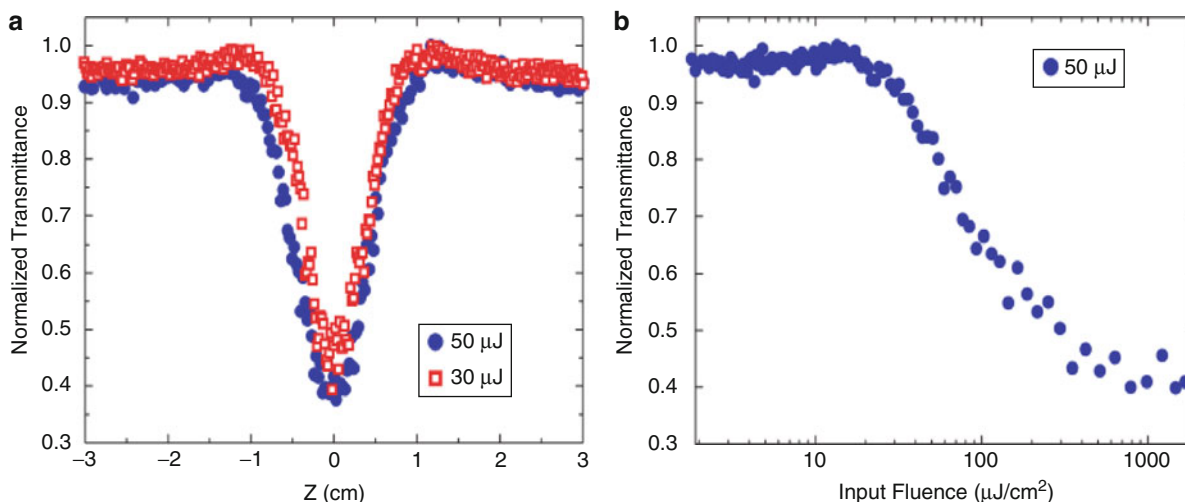
nanorods. Furthermore, in Bi_2O_3 , the numerically calculated β 's were found to be decreasing with increasing energy as shown in Fig. 11b. The falloff of β with increasing intensity is a consequence of reverse saturation absorption arising from a sequential 2PA [2]. Thus, the contribution of the native Bi_2O_3 layer to the nonlinear properties of Bi nanorods cannot be ruled out despite the absence of any nonlinearity in bulk forms of Bi and Bi_2O_3 .

Surface-Modified NiS_2 Nanoparticles

Photonic nanoparticle systems have attracted considerable attention due to a wide variety of NLO applications such as optical limiting, saturable absorption, second harmonic generation, two-photon absorption, all-optical control, and holographic storage control. Typically, such systems are designed with a view of enhancing one of the nonlinear optical properties over their bulk counterparts. Previously, it was shown that nanomaterials (such as surface-modified NiS_2) can also be engineered to deliver systems displaying more complex behavior involving the interplay of two or more nonlinear phenomena [17]. If in such systems, two competing mechanisms operate simultaneously, however, the stronger dominates the weaker. Decoupling them in a single experiment is not

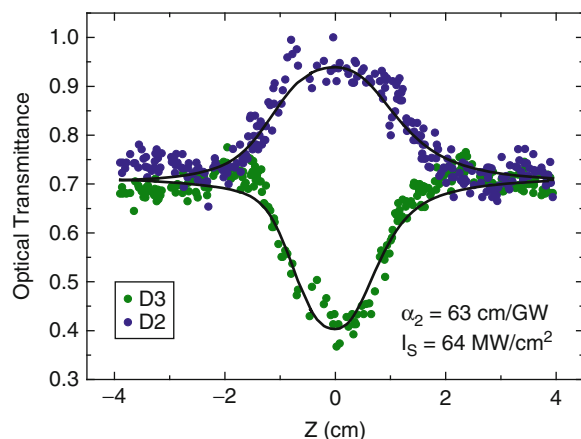
straightforward. This section discusses the case of complete quantitative characterization in a single experiment of two *simultaneous* but *contrary* optical transmission phenomena—optical limiting and saturable absorption in surface-modified NiS_2 nanoparticle suspensions by employing a modified Z-scan method as in the case of Bi nanorods (Fig. 9). Capping with dimethylglyoxime (DMG) results in the surface modification of the nanoparticles and exhibits a spectral feature essential for the saturable absorption.

NiS_2 nanoparticles were suspended in an appropriate amount of 2-propanol so that the sample held in a 1-mm-wide fused silica cuvette would have a linear transmittance of 80% at 532-nm wavelength. The second harmonic of a Q-switched Nd: YAG laser (532 nm) of pulse width 7 ns at 10-Hz repetition was used in the Z-scan studies employing a converging lens with a focal length of 20 cm. The Z-scan curve and transmittance as a function of input fluence are shown in Fig. 12a. It is clearly evident from Fig. 12b that the optical transmittance falls as intensity increases, exhibiting an *optical limiting* behavior. A transmission of 50% (which is the threshold fluence for optical limiting) is reached at $300 \mu\text{J}/\text{cm}^2$. Such optical limiting in NiS_2 occurs due to the intensity-dependent scattering that is also visible to the naked eye.



Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 12 (a) Z-scan curves for nickel sulfide nanoparticles in

2-propanol and (b) the corresponding plot of optical transmission as a function of input fluence



Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 13 Z-scan curves of detector D3 (●) and D2 (●). While D3 measures the contribution of both nonlinear scattering and nonlinear absorption, D2 gives a measure of the intensity-dependent absorption alone

Figure 13 shows the Z-scan curves (with the set up shown in Fig. 9) from detectors D2 (●) and D3 (●) for NiS₂ nanoparticles suspended in 2-propanol. The Z-scan curve due to detector D2 (●) shows an upward peak indicative of the saturation of absorption. The transmittance of the sample increases as it is brought closer to the focus of the Z-scan lens. This is in sharp contrast to the studies on Bi nanorods discussed earlier

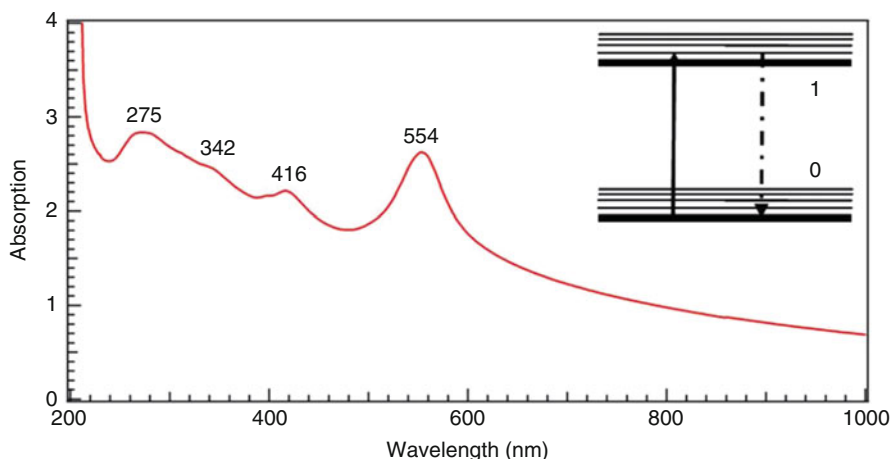
where the corresponding curve is a valley indicating that nonlinear absorption increases with intensity and transmission falls with intensity (Fig. 10). Saturable absorption observed in this case is similar to the behavior of rhodamine B under similar experimental conditions [6].

Furthermore, in Fig. 13, the Z-scan curve of detector D3 (●) which records intensity-dependent transmittance due to both absorption and scattering shows a typical optical limiting valley. Closer observation reveals that both the fall in transmittance due to optical limiting (●) and the rise in transmittance (●) due to saturable absorption are at a z-position of 1.75 cm with respect to the focal point ($z = 0$). Thus, the two effects are simultaneous and not sequential as has been observed in other experiments on nanoparticles systems [18]. The sequential occurrence of saturable absorption followed by an optical limiting behavior is indicative of excited state absorption.

The UV-Vis absorption spectrum for NiS₂ nanoparticles shows a strong absorption at ~ 554 nm (Fig. 14). Under 532-nm excitation, the nanoparticles are excited from the ground state (labeled “0”) to a higher state (labeled “1”) as shown in the inset in Fig. 14. As the intensity of the exciting laser pulse is increased to a maximum of 200 MW/cm², most of the nanoparticles occupy the excited state. As the ground state is bleached, the system becomes increasingly

Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures,

Fig. 14 UV-Visible spectrum of NiS₂ nanoparticles in 2-propanol. *Inset* is a schematic of the excited manifold for saturable absorption as explained in the text



transparent to the incident laser pulse at 532 nm, resulting in a saturation of absorption and the consequent Z-scan curve.

The intensity-dependent absorption coefficient for a two-level system is given by [2]:

$$\alpha_{SA}(I) = \frac{\alpha_0}{1 + \frac{I}{I_s}} \quad (5)$$

where α_0 is the absorption coefficient, I is the intensity of the laser beam, and I_s the saturation intensity. This explains the upper curve in Fig. 13. The bottom curve can be understood in terms of simultaneous nonlinear intensity-dependent scattering that occurs due to energy transfer from nonradiative decay of excited states, as explained earlier in discussions of MWNT and Bi nanorod systems.

A complete description of the Z-scan experiment requires that both the saturable absorption and optical extinction due to nonlinear scattering be taken into account. Thus, in the similar lines of Venkatram et al. [18], we use the following equation describing these two concomitant processes:

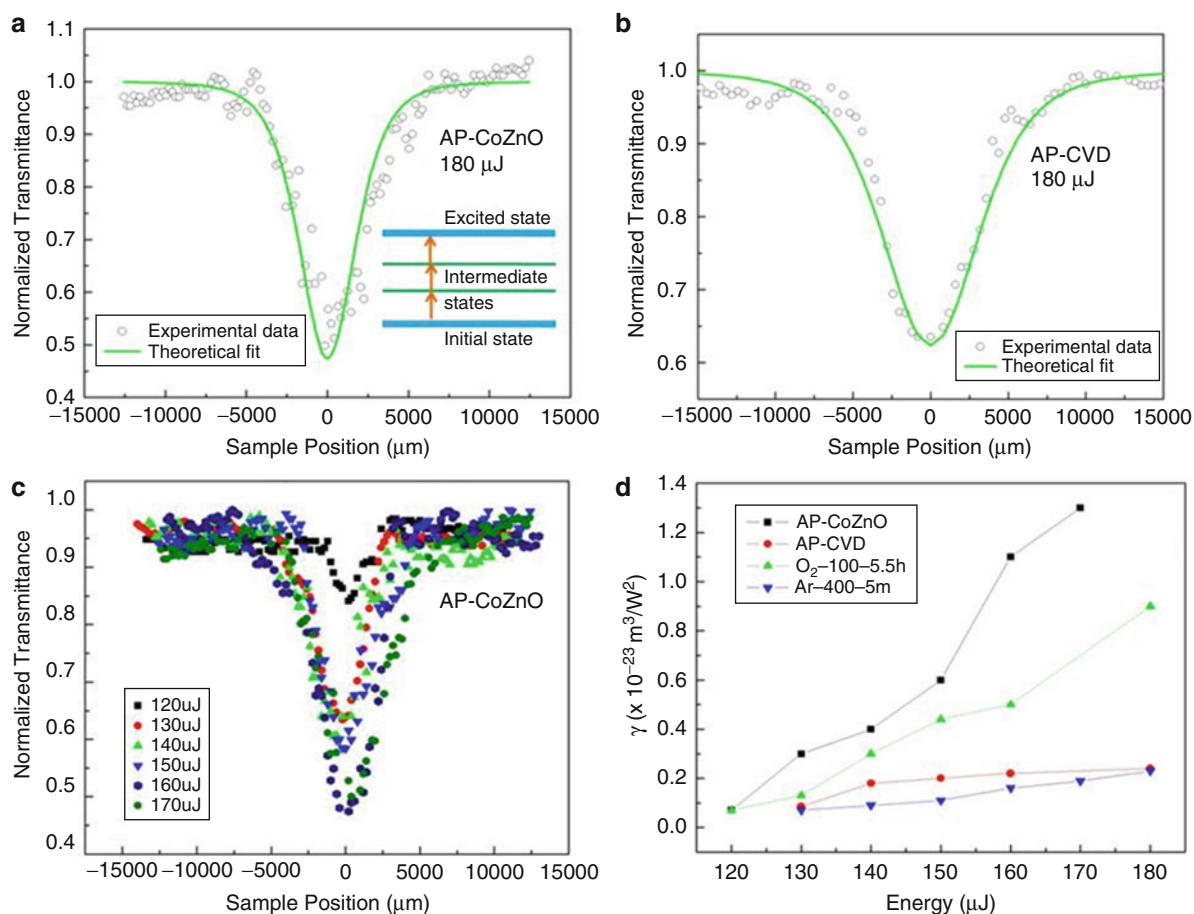
$$\frac{dI}{dz} = - \left[\alpha_0 + \alpha_2 I + \frac{\alpha_0}{1 + \frac{I}{I_s}} \right] I \quad (6)$$

In the above equation, α_2 is the nonlinear scattering coefficient. The value of I_s and α_2 can be extracted from numerical fitting of data to this propagation. Here, the fourth-order Runge-Kutta method is used to solve the equation while finding the best fit for both

parameters [6]. As mentioned earlier, the Z-scan curve of D2 is purely due to absorptive effects. The equation is first solved setting $\alpha_2 = 0$ so as to obtain a value for I_s which in this case is 64 MW/cm². This value of I_s is much smaller than the incident intensity at the focal point (at both 30 μ J and 50 μ J), eventually favoring the domination of nonlinear scattering following the absorption of a fraction of the leading part of the incident pulse. Thereafter, the value of α_2 was obtained by using this value of I_s in the complete equation and finding the best fit to the Z-scan curve from detector D3. The nonlinear scattering coefficient, α_2 , was found to be 63 cm/GW.

ZnO Nanostructures

Among the metal oxides, ZnO is a remarkable wide bandgap (3.37 eV in bulk) II–VI semiconductor. Its high exciton binding energy (60 meV) allows efficient exciton emission even at room temperature. As a consequence of its noncentrosymmetric crystal structure, ZnO is expected to have nonzero second-order susceptibility. Although many properties of bulk and nanostructured ZnO have been well studied, there is still a lack of understanding in the relationship between the density of defects, dopant level, synthesis conditions, and the observed optical and magnetic properties. In nanostructures, a fairly high surface state density significantly modifies or alters the chemical and physical properties, which can be reflected in optical measurements. NLO properties of different forms of pristine ZnO, such as single crystals, thin films, nanolayers, and nanowires, have been previously reported. It is also very important to understand



Nonlinear Optical Absorption and Induced Thermal Scattering Studies in Organic and Inorganic Nanostructures, Fig. 15 (a) Z-scan data for AP-CoZnO samples fitted to a three-photon absorption (3PA) equation. *Inset* is a schematic depicting effective 3PA. When the intermediate states are virtual, the effect is genuine 3PA, but when they are real, it is an effective

3PA arising mostly from excited state absorption. (b) Z-Scan data for pristine ZnO samples fitted to a 3PA equation. (c) Z-scan data for AP-CoZnO samples at different incident laser energies. (d) Variation of the 3PA coefficient (γ) as a function of the incident laser pulse energy for AP-CVD, AP-CoZnO, O₂-100-5.5 h and Ar-400-5 m samples

the NLO properties of transition metal (TM)-doped ZnO nanostructures due to their ferromagnetic ordering at room temperature and importance in spintronic applications. An extensive study of NLO properties of TM-doped ZnO nanostructures can lead to a better understanding of the impact of surface states.

In earlier studies of the origin of ferromagnetism (FM) in pristine micro- and nanostructured ZnO, Podila et al. observed the changes in magnetic properties with varying crystallinity. They attributed the FM to the presence of surface states [8]. The presence of such surface states in pristine and Co-doped ZnO nanostructures (denoted as AP-CVD and AP-CoZnO)

can be confirmed using the open aperture Z-scan. It is worth noting that, except when excited by ultrafast laser pulses, ΔT strongly depends on the excited state population density. In view of this, the nonlinear transmission of the ZnO samples was measured in the nanosecond excitation regime using linearly polarized 5-ns optical pulses from a Q-switched frequency-doubled Nd:YAG laser operating at 532 nm. The Z-scan data obtained for AP-CoZnO and AP-CVD samples (Fig. 15a, b, respectively) were found to be best-fit numerically by a three-photon absorption (3PA) process model. It is instructive to note that, compared to the 2PA process described in

Eq. 7, the nonlinear absorption coefficient α in the presence of 3PA is given by

$$\alpha(I) = \alpha_o + \gamma I^2 \quad (7)$$

In the above equation, I is the intensity of the laser, α_o the unsaturated linear absorption coefficient, and γ the 3PA coefficient. The value of γ was obtained by fitting the Z-scan curve to the 3PA propagation equation, given by

$$\frac{dI}{dz} = -\alpha_o I - \gamma I^3 \quad (8)$$

with z' being the propagation distance within the sample.

The values of γ numerically obtained from the fits indicated that nonlinear absorption in the present case arises not only from genuine 3PA but also from “effective” three-photon nonlinearity [1]. Such effective 3PA originates from sequential excited state absorption as shown in the inset of Fig. 15a. In the case of genuine three-photon absorption (3PA), where the transition states involved are virtual, the 3PA coefficient (γ) is a constant and is independent of the incident laser fluence. All genuine ground state absorption coefficients (such as 2PA or 3PA) are dependent on the ground state population N since the absorption coefficient is obtained by multiplying the corresponding absorption cross section (σ) by N . The absorption cross section is a microscopic parameter, independent of N that characterizes the average two-photon/three-photon absorbability per molecule. However, the change in γ will be evident only when there is a substantial change in the ground state population due to absorption. Therefore, if the absorptions are weak, then the coefficients can be considered to be almost constant, since there is a negligible change in N . Genuine 2PA and 3PA are usually very weak phenomena, and, hence, the corresponding coefficients can be considered as material constants at a given wavelength and concentration. But strong excited state absorptions (ESA) during an effective 3PA deplete the ground state population significantly so that the absorption coefficient is no longer a constant. Hence a change in γ with respect to the incident laser intensity can help confirm the presence of surface states in the forbidden gap of ZnO.

Both two-photon absorption and/or sequential absorption of two photons at 532 nm (2.33 eV; inset Fig. 15a) will result in a real terminal level which is above the lowest excitonic state that lies around 363 nm (3.4 eV) for ZnO, and another one-photon absorption to go from this level to the band edge. Figure 15c, d show the variation of γ for AP-CoZnO, AP-CVD, and AP-CVD samples annealed in O₂ (Ar) for 5.5 h (5 min) at 100°C (400°C)—O₂-100-5.5 h and Ar-400-5 m, respectively—as a function of incident laser energy. While γ remains almost constant with input energy in AP-CVD, there is a pronounced increase in γ with incident energy in AP-CoZnO and O₂-100-5.5 h samples. Moreover, the 3PA coefficient (and, hence, the optical limiting efficiency) for AP-CoZnO samples is obviously higher than that of pristine ZnO. These observations concur with the enhancement seen in the magnetic moment of AP-CoZnO and O₂-100-5.5 h samples and indicate the presence of enhanced surface state density due to Co atoms in the ZnO lattice. Thus, the increase of γ with the incident laser energy confirms effective 3 PA due to the existence of a wide range of surface states in AP-CoZnO and O₂-100-5.5 h samples. Furthermore, the lowest γ values and a constant trend with input fluence were observed for Ar-400-5 m samples indicating the absence of any surface states. An enhanced surface state density in AP-CoZnO due to Co doping results in stronger FM and better optical limiting compared to pristine ZnO nanostructures [19].

Conclusions

In the preceding sections, the reader is introduced to optical nonlinearities and given an overview of various methods used for nonlinear optical characterization with emphasis on the Z-scan technique. The chapter discusses in detail the NLO properties of various nanostructures, such as carbon nanotubes, bismuth nanorods, NiS₂ nanoparticles, and ZnO nanowires, studied using the Z-scan. In the case of MWNTs and Bi nanorods, nonlinear scattering is found to play a major role in determining the nonlinear transmission. The simultaneous onset of two different nonlinear processes (saturable absorption and nonlinear scattering) was observed in NiS₂ nanoparticles. In pristine and doped ZnO nanostructures, the intensity dependence of the 3PA coefficient could be used to confirm the

presence of intermediate states. These studies indicate that organic as well as inorganic nanostructures are excellent candidates for nonlinear optical transmission with potential strategic applications, including optical power limiting.

References

- Sutherland, R.L.: Handbook of Nonlinear Optics. Marcel Dekker, New York (1996)
- Nalwa, H.S., Miyata, S.: Nonlinear Optics of Organic Molecules and Polymers. CRC Press, BocaRato (1997). pp 841
- Sheik-Bahae, M., Said, A.A., Wei, T.H., Hagan, D.J., Van Stryland, E.W.: Sensitive measurement of optical nonlinearities using a single beam. *J. Quantum Electron* **26**, 760 (1990)
- Tutt, L.W., Boggess, T.F.: A review of optical limiting mechanism and devices using organics, fullerenes, semiconductors and other materials. *Prog. Quantum Electron* **17**, 299 (1993)
- Gu, B., Lou, K., Wang, H.T., Ji, W.: Dynamics of two-photon induces three-photon absorption in nanosecond, picosecond and femtosecond regimes. *Opt. Lett.* **35**, 3 (2010)
- Rao, S.V., Srinivas, N.K.M.N., Rao, D.N.: Nonlinear absorption and excited state dynamics in Rhodamine B studies using Z-scan and degenerate four wave mixing techniques. *Chem. Phys. Lett.* **361**, 439 (2002)
- Balu, M., Joel, H., David, H., Van Stryland, E.: White-light continuum Z-scan technique for nonlinear materials characterization. *Opt. Exp.* **12**(16), 3820–3826 (2004)
- Podila, R., Queen, W., Nath, A., Fazio, A., Schonalez, A., He, J., Dalpian, G., Skove, M.J., Hwu, S.J., Rao, A.M.: Origin of FM ordering in pristine micro- & nanostructured ZnO. *Nano Lett.* **10**, 1380 (2010)
- Chin, K.C., Gohel, A., Elim, H.I., Chen, W., Ji, W., Chong, G.L., Sow, C.H., Wee, A.T.S.: Modified carbon nanotubes as broadband optical limiting nanomaterials. *J. Mater. Res.* **21**, 2758 (2006)
- Pan, H., Chen, W., Feng, Y.P., Ji, W.: Optical limiting properties of metal nanowires. *Appl. Phys. Lett.* **88**, 223106 (2006)
- Philip, R., Kumar, G.R., Sandhyarani, N., Pradeep, T.: Picosecond optical nonlinearity in monolayer protected gold, silver, and gold-silver alloy nanoclusters. *Phys. Rev. B.* **62**, 13160 (2000)
- Ganeev, R.A., Rysanyansky, A.I., Redkorechev, V.I., Fostiropoulos, K., Prieb, G., Usmanov, T.: Variations of nonlinear optical characteristics of C60 thin films at 532 nm. **225**, 131 (2003)
- Margulis, A.: Theoretical estimations of third-order optical nonlinearities for semiconducting carbon nanotubes. *J. Phys. Condens. Matter* **11**, 3065 (1999)
- Joudrier, V., Bourdon, P., Hache, F., Flytzanis, C.: Nonlinear light scattering in a two-component medium: optical limiting application. *Appl. Phys. B: Lasers Opt.* **67**, 627 (1998).
- Black, M.R., Lin, Y.-M., Cronin, S.B., Rabin, O., Dresselhaus, M.S.: Infrared absorption in Bi nanowires resulting from quantum confinement. *Phys. Rev. B* **65**, 195417 (2002)
- Sivaramakrishnan, S., Muthukumar, V.S., Reppert, J., Anija, M., Sivasankara Sai, S., Philip, R., Venkataramaniah, K., Kuthirummal, N., Rao, A.M.: Nonlinear optical scattering and absorption in Bi nanorod suspensions. *Appl. Phys. Lett.* **91**, 093104 (2007)
- Muthukumar, V.S., Kiran, J.K., Reppert, J., Satyajit, R., Krishna, V., Nageshwar Rao, G., Siva Rama Krishnan, S., Siva Sankara Sai, S., Venkataramaniah, K., Rao, A.M.: Nonlinear optical transmission of surface modified NiS nanoparticles: Saturation of absorption and optical limiting. *NANO: Brief Rep. Rev.* **3**, 161 (2008)
- Venkatram, N., Sai Santosh Kumar, R., Narayana Rao, D.: Nonlinear absorption and scattering properties of CdS nanocrystals with its application as a potential optical limiter. *J. Appl. Phys.* **100**, 074309 (2006)
- Podila, R., Anand, B., West, J.P., Philip, R., He, J., Skove, M., Hwu, S.J., Rao, A.M.: Evidence for surface states in pristine and Co-doped ZnO nanostructures: magnetization and nonlinear optical studies. *Nanotechnology* **22**, 095703 (2011)

NS500

- [Nanoparticle Tracking Analysis](#)

NTA

- [Nanoparticle Tracking Analysis](#)

Nucleation of Silica Nanowires

- [Growth of Silica Nanowires](#)

Nucleic Acid Amplification

- [Nanotechnology Applications in Polymerase Chain Reaction \(PCR\)](#)

