
P

Passive Pumping

- ▶ [Capillary Flow](#)

Patterned Hydrogels

- ▶ [Nanoengineered Hydrogels for Cell Engineering](#)

Patterning

- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)

Peltier Coefficient

- ▶ [Thermoelectric Heat Convertors](#)

Peltier Cooler

- ▶ [Thermoelectric Heat Convertors](#)

Peltier Couple

- ▶ [Thermoelectric Heat Convertors](#)

Peltier Effect

- ▶ [Thermoelectric Heat Convertors](#)

Peltier Element

- ▶ [Thermoelectric Heat Convertors](#)

Peltier Module

- ▶ [Thermoelectric Heat Convertors](#)

Peltier-Seebeck Effect

- ▶ [Thermoelectric Heat Convertors](#)

Percutaneous Absorption

- ▶ [Dermal and Transdermal Delivery](#)

Perfluorocarbon Nanoparticles

Patrick M. Winter¹, Gregory M. Lanza² and Samuel A. Wickline²

¹Department of Radiology, Imaging Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

²C-TRAIN Labs, Washington University, St. Louis, MO, USA

Synonyms

PFC nanoparticles; PFOB nanoparticles

Definition

Perfluorocarbon (PFC) nanoparticles represent a multifunctional platform consisting of a PFC core encapsulated within a phospholipid monolayer. The particles have been developed for magnetic resonance (MR) molecular imaging to noninvasively map the expression of important biomarkers of disease.

Introduction

Perfluorocarbon (PFC) nanoparticles consist of a liquid PFC core encapsulated by a phospholipid monolayer. PFCs represent a family of modified hydrocarbon molecules in which the hydrogens are replaced with fluorine atoms [1–3]. The carbon–fluorine bond is very stable and biologically inert due to the dense electron cloud, which prevents chemical reactions with other compounds [2]. PFCs with molecular weights between 460 and 520 are biocompatible, showing no toxicity, carcinogenicity, mutagenicity, or teratogenic effects. PFCs have tissue half-life residencies ranging from 4 to 65 days. They are not metabolized. Instead, they are reintroduced into the blood circulation from clearance organs, such as the liver, by lipid carriers and are exhaled via the lungs [2]. PFC nanoparticles have been formulated for a wide range of molecular imaging and targeted drug delivery applications. The particles are typically 250 nm in diameter, which is large enough to constrain them to the intravascular space. As a result, PFC

particles are often targeted to intraluminal biomarkers expressed by endothelial cells, such as integrins, selectins, or adhesion molecules [2, 3].

Magnetic resonance (MR) molecular imaging aims to identify and map the expression of important biomarkers on a cellular scale. Typical MR imaging techniques lack sufficient resolution and sensitivity to directly detect these molecular signatures due to their very low concentrations in vivo. To overcome this limitation, nanoparticle agents are often employed to specifically bind to the biomarker of interest, accumulate at the site of pathology, and generate sufficient image contrast. Nanoparticles offer a large surface area for the incorporation of multiple binding ligands (to improve targeting efficacy) and multiple paramagnetic chelates (to amplify the signal enhancement). In some cases, the nanoparticle may also serve as a drug delivery agent, providing diagnostic and therapeutic information via noninvasive MR imaging.

A number of other medical imaging modalities are capable of molecular imaging, including positron emission tomography (PET), single photon emission computed tomography (SPECT), and optical imaging [4]. Like MR imaging, these modalities typically rely on specifically designed contrast agents that bind to the molecular biomarker of interest. For PET and SPECT imaging, the contrast agents consist of a radioactive element for detection. In optical imaging, the contrast agent will utilize fluorescence or bioluminescence for mapping expression of cellular receptors. MR imaging, however, has certain important advantages over these other modalities for molecular imaging applications. MR imaging offers higher spatial and temporal resolution than PET or SPECT and provides greater tissue penetration than optical imaging. This allows MR imaging to combine anatomical and/or functional information with molecular imaging, much like the multimodality benefits of PET/CT. In addition, MR imaging may be better suited for serially tracking disease progression or therapeutic response compared to nuclear imaging due to concerns of cumulative radiation dose.

Molecular imaging with PFC nanoparticles has been applied to a wide array of disease biomarkers, therapeutic drugs, and imaging modalities [2, 3]. PFC nanoparticle formulations have been developed for ultrasound imaging [5], computed tomography [6], nuclear imaging [7], optical imaging [8–10], MR [3], as well as multi-modality applications. For example, combined ^{99m}Tc SPECT and MR

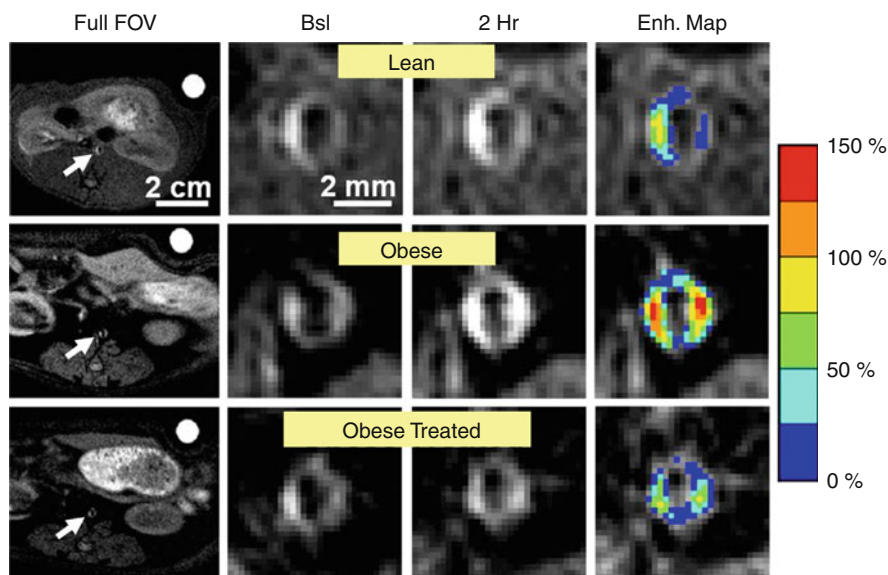
imaging of $\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles has been used to map tumor angiogenesis with high sensitivity and high resolution [7]. The $\alpha_v\beta_3$ -integrin is a heterodimeric transmembrane glycoprotein expressed by numerous cell types, including endothelial cells, macrophages, platelets, lymphocytes, smooth muscle cells, and tumor cells [2]. Endothelial expression of $\alpha_v\beta_3$ -integrin is required for the migration and invasion of vascular endothelial cells, and therefore, serves as a biomarker of angiogenesis [11]. Since PFC nanoparticles are constrained to the vasculature, they are not exposed to nonendothelial integrin-expressing cells, which improves binding specificity to the neovasculature [2]. Nuclear imaging offers very high sensitivity, which is preferred for the detection of occult tumors or metastases within large tissue regions. However, the limited spatial resolution of nuclear imaging precludes detailed characterization of the pathology. Incorporating a ^{99m}Tc nuclear imaging agent onto $\alpha_v\beta_3$ -integrin-targeted gadolinium nanoparticles yielded a high-sensitivity nuclear signal that was used to guide further high-resolution MR imaging and characterization of tumor angiogenesis [7]. MR molecular imaging demonstrated a coherent, asymmetric, patchy pattern of angiogenesis along the outer aspects of the tumor mass, consistent with an angiogenic phenotype.

In the case of ultrasound imaging, PFC nanoparticles represent a significant departure from conventional contrast agent design. Most ultrasound contrast agents consist of gas bubbles to maximize the backscatter of ultrasound waves and thereby increase the image contrast. However, these bubbles are quite large (more than 1 μm in diameter) and are quickly entrapped within the lungs. This severely limits the practical application of bubbles in molecular imaging because they do not circulate long enough to bind to biomarkers of disease and accumulate at the sites of pathology. PFC nanoparticles, however, have a long circulatory half-life and are inherently echogenic due the differences in the speed of sound between water and the PFC core. This ultrasound backscatter is typically not sufficient to visualize particles in the blood stream. Instead, the particles must accumulate at the target site, forming an “acoustic mirror” to generate significant image contrast. Targeted PFC nanoparticles have been used for molecular imaging of fibrin in acute arterial thrombus, high-frequency intravascular imaging, and imaging with novel ultrasound receivers, called information theoretical detectors [5].

Along with MR molecular imaging, PFC nanoparticles have been employed to label cells and noninvasively track implantation and migration of cell therapies [1, 3]. While MR molecular imaging with PFC nanoparticles typically utilizes paramagnetic chelates for detection with ^1H MR, tracking of cells labeled with PFC nanoparticles often relies on ^{19}F MR. ^{19}F has a relatively high sensitivity, 83% compared to ^1H , and there is virtually no native background signal. Thus, the PFC core can provide a definitive and quantitative MR signature [1]. PFC nanoparticles can be internalized by stem cells without the need of a transfection agent and ^{19}F MR imaging can quantitatively track the cells in vivo [1]. The high sensitivity of ^{19}F MR allows the in vivo detection of as few as 2,000 labeled cells. For some applications, modification of the surfactant surface of the particles improves cellular uptake. For example, PFC nanoparticles formulated with cationic phospholipids on the particle surface can be effectively internalized by neural stem cells [1].

Molecular Imaging of Therapeutic Response

One potential clinical application of MR molecular imaging with PFC nanoparticles is to serially map biomarkers of disease progression and/or therapeutic response. Noninvasive imaging could be used to monitor the biochemical signatures of disease, providing earlier and more specific detection of pathology compared to tracking symptomology or the occurrence of clinical events. For example, molecular imaging of angiogenesis can serve as a biomarker of atherosclerosis and can be used to monitor the physiological effects of a range of therapeutic treatments [3]. Paramagnetic $\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles have been utilized to characterize angiogenesis in the aortic wall of an atherosclerotic rat model and to monitor the effects of benfluorex (an appetite suppressant) on aortic angiogenesis [12]. The JCR:LA-cp rat is a model of metabolic syndrome characterized by obesity, insulin resistance, hyperlipidemia, and vasculopathy. The obese rats displayed higher food consumption and body weight, as well as increased insulin, leptin, cholesterol, and triglyceride levels compared to the lean controls. In conjunction with this typical metabolic syndrome phenotype, MR



Perfluorocarbon Nanoparticles, Fig. 1 Full field-of-view (“Full FOV”) images of JCR rats showing large abdominal fat deposits in the obese (*middle*) and obese-treated (*bottom*) animals that were absent in the lean (*top*) rat (*arrows* denote aortas). Magnified views of the abdominal aorta in lean, obese, and obese-treated rats at baseline (“Bsl”) and 2-h (“2 Hr”) post

injection of targeted nanoparticles displaying signal enhancement in the aorta wall. Percent signal enhancement maps (“Enh. Map”, false-colored from *blue* to *red*) demonstrate higher MR signal enhancement in the untreated obese animal, indicating active angiogenesis supporting development of atherosclerotic lesions (Reprinted with permission from Cai et al. [12])

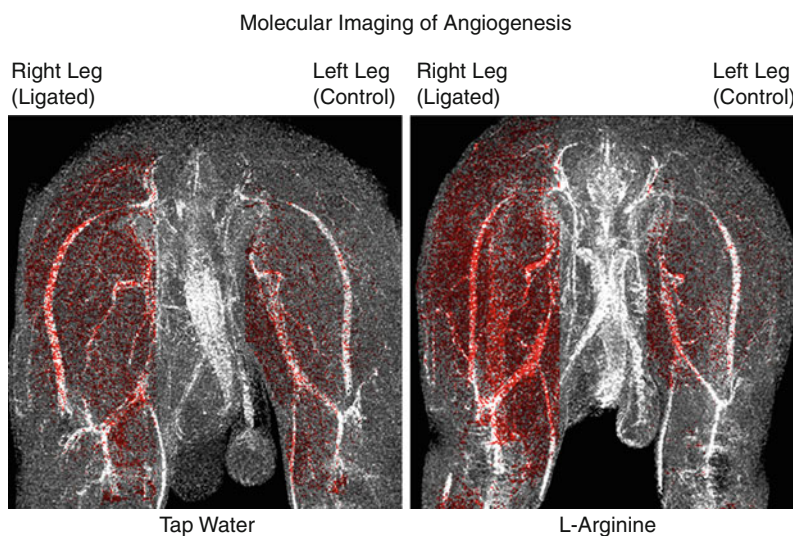
molecular imaging revealed increased aortic angiogenesis in the obese animals, indicative of actively developing atherosclerotic plaques in the arterial wall (Fig. 1). Benfluorex reduced food consumption, body weight, insulin, and leptin in obese-treated rats, but had no effect on serum cholesterol or triglyceride levels. In a similar fashion, benfluorex reduced aortic angiogenesis in the obese-treated animals to the same level as the control lean rats. These molecular imaging results were corroborated by histological measurements of microvascular density. Therefore, MR molecular imaging with $\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles can noninvasively monitor atherosclerosis in a rat model of obesity and can monitor the neovascular effects of an appetite suppression therapy.

However, atherosclerosis does not always lead to an increase in angiogenesis. In a rabbit model of peripheral vascular disease, hypercholesterolemia causes a reduction in the angiogenic response to chronic muscle ischemia, which can be normalized with a pro-angiogenic therapy, such as L-arginine supplementation [13]. MR molecular imaging was performed 10 days after femoral artery ligation with paramagnetic

$\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles. Molecular imaging revealed diffuse angiogenesis throughout the ischemic limb for both L-arginine-treated and untreated groups (Fig. 2). However, the angiogenic response was significantly increased in L-arginine-treated animals compared to the untreated controls. Furthermore, the integrin-targeted formulation produced two times higher MR enhancement compared to nontargeted particles, demonstrating improved identification of angiogenic vasculature with biomarker targeting. X-ray angiography was performed 40 days postligation to measure the development of collateral vessels feeding the ischemic muscle tissue. More collateral vessels were observed in L-arginine-treated rabbits compared to the untreated controls. Histologic staining of muscle capillaries revealed a denser pattern of microvasculature in L-arginine-treated animals, confirming the MR and X-ray imaging results. Noninvasive molecular imaging with PFC nanoparticles could be used for early detection of response to angiogenic therapies directed at ischemic diseases, such as peripheral vascular disease or myocardial infarction, to enable personalized optimization of treatment strategies.

Perfluorocarbon Nanoparticles,

Fig. 2 Molecular imaging of angiogenesis shows highly diffuse enhancement throughout the ligated (*right*) leg of cholesterol-fed rabbits with only slight enhancement of the control (*left*) leg. The L-arginine-treated rabbit (*right panel*) shows a more dense distribution of angiogenesis in the ligated limb compared with the control animal (*left panel*) (Reprinted with permission from Winter et al. [13])



Targeted Antiangiogenic Therapy

In the realm of targeted drug delivery, PFC nanoparticles have been utilized to carry a number of antiangiogenic drugs. In particular, the potent antiangiogenic drug, fumagillin suppresses angiogenesis by inhibition of methionine aminopeptidase 2 (MetAP2) [9]. Targeting the particles to a biomarker of angiogenesis, such as the $\alpha_v\beta_3$ -integrin, provides a means to specifically direct potent drugs to the sites of pathology, improving the therapeutic efficacy and decreasing the occurrence of unwanted side effects. Combining a specifically targeted molecular imaging agent with a therapeutic drug could benefit patients in a wide range of clinical scenarios. The imaging agent would allow confirmation and quantification of local drug delivery and enable personalization of treatment protocols based on the pharmacokinetics in the target tissue. In addition, specific targeting of the agent could improve uptake and retention of the drug at the sites of disease while lowering the exposure to other susceptible organs and reducing the occurrence of side effects.

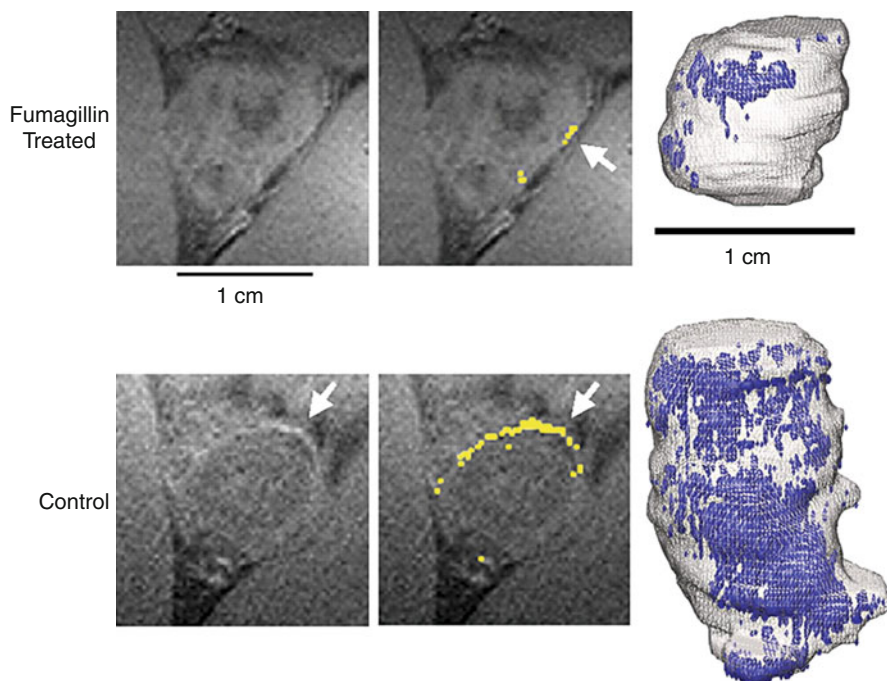
Paramagnetic PFC nanoparticles targeted to the $\alpha_v\beta_3$ -integrin have been used for sensitive detection of angiogenesis in Vx2 tumors implanted in New Zealand white rabbits [2]. In vivo competition studies demonstrated that the targeted particles specifically bind to the $\alpha_v\beta_3$ -integrin. Furthermore, MR molecular imaging with the targeted nanoparticles was able to differentiate between growing tumors and

inflammatory lesions resulting from rejected cancer cells. These tumor cell remnants appeared identical to the viable growing tumors on standard T2-weighted MR images, but molecular imaging of angiogenesis showed negligible neovasculature in these masses. Antiangiogenic therapy has also been demonstrated in this animal model using $\alpha_v\beta_3$ -integrin-targeted fumagillin nanoparticles [14]. Treatment with $\alpha_v\beta_3$ -integrin-targeted fumagillin nanoparticles reduced the tumor volume significantly compared with animals given nontargeted fumagillin nanoparticles, $\alpha_v\beta_3$ -integrin-targeted nanoparticles without drug or saline (Fig. 3). The MR imaging data was used to generate a 3D map of tumor angiogenesis, which showed a coherent asymmetric peripheral distribution characterized by dense neovessel regions interspersed with a finer, reticular pattern.

Another possible biomarker of angiogenesis is Robo4, an endothelial specific guidance receptor that is expressed on tumor neovasculature but not mature vessels [15]. Paramagnetic PFC nanoparticles were coupled to a Robo4 antibody and evaluated in a mouse melanoma tumor model [15]. Both Robo4-targeted and $\alpha_v\beta_3$ -integrin-targeted nanoparticle formulations generated greater MR image contrast than particles lacking a targeting ligand or coupled to scrambled targeting ligands. Microscopic analysis of particle binding was performed with Robo4-targeted rhodamine nanoparticles and $\alpha_v\beta_3$ -integrin-targeted AlexaFluor 488 nanoparticles. Binding of $\alpha_v\beta_3$ -integrin-targeted particles was generally more

Perfluorocarbon**Nanoparticles, Fig. 3**

Left: Molecular imaging of targeted antiangiogenic therapy (fumagillin) in rabbit Vx-2 tumors showing diminished MR contrast enhancement in the treated animal (*top*) versus the control animal (*bottom*). *Middle:* MR enhancement, color coded in *yellow* (*arrows*), demonstrating reduced tumor angiogenesis in the fumagillin-treated animal. *Right:* 3D maps of tumor angiogenesis showing an asymmetric distribution of neovasculature (*blue*) that is dramatically reduced in the fumagillin-treated rabbit (Reprinted with permission from Winter et al. [14])

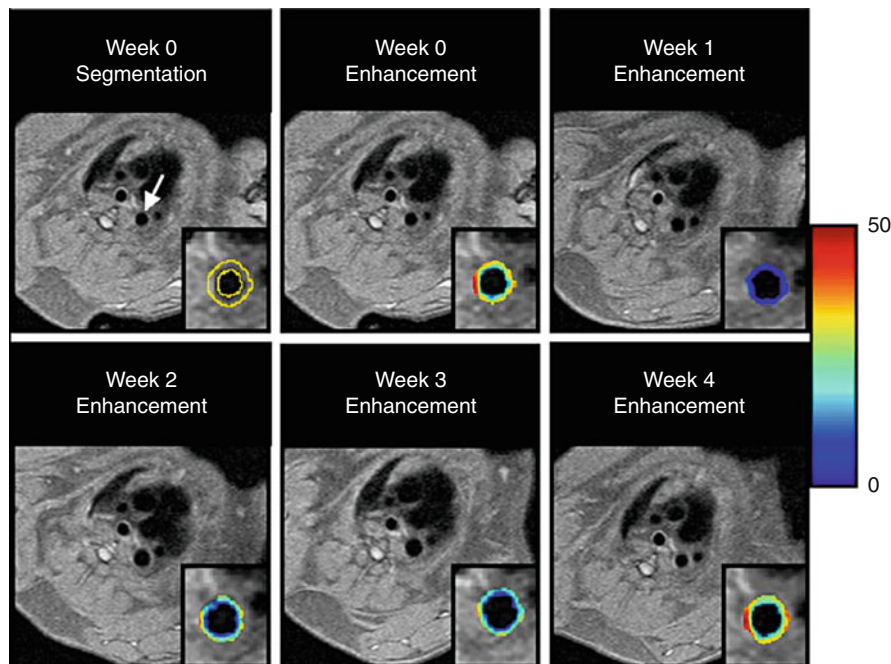


prevalent than the binding of Robo4-targeted nanoparticles, particularly in the tumor periphery. Thus, PFC nanoparticles can be used for both macroscopic MR molecular imaging as well as light microscopy to define cell binding of the particles and the coexpression of different cellular receptors.

MR molecular imaging of angiogenesis has also been demonstrated in human melanoma tumors (40 mm³) implanted in immunocompromised mice using $\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles [2]. Again, *in vivo* competition studies demonstrated the high specificity of ligand-directed targeting, which was further corroborated microscopically using $\alpha_v\beta_3$ -integrin-targeted fluorescent PFC nanoparticles. As an alternative to targeting the $\alpha_v\beta_3$ -integrin for mapping angiogenesis, the $\alpha_5\beta_1$ -integrin has been explored as a biomarker of tumor neovasculature [8]. The $\alpha_5\beta_1$ -integrin is an important adhesion molecule, which regulates endothelial cell migration and survival during angiogenesis. The $\alpha_5\beta_1$ -integrin is not expressed on normal quiescent blood vessels, but its expression is induced on tumor blood vessels and in response to angiogenic factors, including basic fibroblast growth factor, interleukin-8, tumor necrosis factor- α , and the angiogenic protein Del-1. The utility of a dual $\alpha_5\beta_1(\alpha_v\beta_3)$ -integrin targeting ligand for molecular imaging and targeted delivery of fumagillin was

compared with the use of a targeting ligand specific for only the $\alpha_v\beta_3$ -integrin. MR signal enhancement maps revealed a sparse network of neovessels predominantly located in the tumor periphery. Treatment with $\alpha_5\beta_1(\alpha_v\beta_3)$ -integrin-targeted nanoparticles containing fumagillin further decreased the already small amount of neovessel enhancement to an almost negligible level; however, the tumor volume was not decreased. Since this tumor model displayed a very sparse neovasculature and was relatively slow growing, it is not surprising that an antiangiogenic therapy did not significantly decrease the tumor size. Clinically, MR molecular imaging of angiogenesis could be used to characterize small tumors, such as breast, lung, or colon, and stratify patients into antiangiogenic therapies based on the abundance of neovascular biomarkers. Such a procedure for screening potential patients may yield improved therapeutic benefits with decreased risk of side effects.

Similar to the growth of tumors, the development of atherosclerotic plaques requires the recruitment of new blood vessels to meet the high metabolic demands of the abnormal tissue. In atherosclerosis, angiogenesis occurs in the vasa vasorum of large caliber arteries, such as the aorta, carotids, and coronaries. Since antiangiogenic therapies can slow the progression of cancer, it is reasonable to investigate if these



Perfluorocarbon Nanoparticles, Fig. 4 Serial imaging of aortic wall angiogenesis for monitoring targeted antiangiogenic therapy with fumagillin nanoparticles. *Top:* Black blood image of the thoracic aorta (*arrow*) prior to antiangiogenic treatment (Week 0) demonstrating segmentation of the vessel wall (*outlined in yellow*) and patchy areas of high angiogenesis within the aortic wall (color-coded as percent enhancement of MR

signal). At week 1, the MR enhancement is drastically reduced due to the antiangiogenic effect of fumagillin. *Bottom:* The level of signal enhancement gradually increases at Weeks 2 and 3 after fumagillin treatment. At Week 4, the level of enhancement is practically identical to the Week 0 image (Reprinted with permission from Winter et al. [9])

treatments could be employed to normalize atherosclerotic vessels. The acute antiangiogenic response to $\alpha_v\beta_3$ -integrin-targeted paramagnetic nanoparticles containing fumagillin has been studied in atherosclerotic rabbits [16]. MR imaging was used to estimate drug deposition in the aortic wall. One week later, the level of neovascular $\alpha_v\beta_3$ -integrin expression was assessed using $\alpha_v\beta_3$ -targeted paramagnetic nanoparticles without fumagillin. Aortic areas displaying high MR enhancement at the time of treatment had the largest subsequent reduction in $\alpha_v\beta_3$ -integrin-targeted MR signal 1 week later, suggesting that combining imaging with therapy may not only confirm and quantify the local delivery of chemotherapeutics but may also provide early predictions of the subsequent treatment effects.

In addition to the acute response to therapy, MR molecular imaging could provide serial monitoring of the end organ effects with the ultimate goal of optimizing drug treatment regimens and customizing patient protocols. The antiangiogenic effect of

atorvastatin with and without targeted delivery of fumagillin was serially monitored for 8 weeks by MR imaging in atherosclerotic rabbits [9]. Atorvastatin was dosed continuously via incorporation into the feed, while fumagillin treatment was provided once every 4 weeks. During 8 weeks of study, atorvastatin treatment did not reduce $\alpha_v\beta_3$ -integrin expression in the aortic wall. Fumagillin treatment, on the other hand, resulted in a transient (2–3 weeks) reduction in image enhancement, indicating successful antiangiogenic treatment (Fig. 4). Combining the fumagillin and atorvastatin treatments yielded a persistent decrease in the MR signal, suggesting that chronic statin treatment could be used to prolong the effects of discrete doses of an antiangiogenic agent. Histology revealed that the MR image enhancement was strongly correlated to the microvascular density in the aortic wall in a logarithmic fashion. An effective and sustained antiangiogenic treatment may halt plaque progression and stabilize atheromas by decreasing intramural hemorrhage and inflammation.

One of the benefits of using a molecular imaging contrast agent for targeted drug delivery is that imaging can provide a noninvasive measurement of drug deposition at the site of pathology and elimination from the body. Paramagnetic $\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles have been utilized to quantify the pharmacokinetics of the particles in both the blood pool as well as the target tissue in an atherosclerotic rabbit model [3]. Noninvasive monitoring of the particle biodistribution and elimination could be used as a surrogate marker for the tissue levels of lipophilic drugs incorporated in the nanoparticle formulation. Typically, pharmacodynamic measurements rely on blood sampling without consideration of the drug accumulation in the target tissue. Noninvasive mapping of drug concentrations in the end organ could provide significant clinical advantages in monitoring targeted therapeutic agents and determining the local drug concentrations beyond the blood pool. Atherosclerotic rabbits were injected with either $\alpha_v\beta_3$ -targeted or nontargeted paramagnetic nanoparticles at a dose of 1 mL/kg (0.0046 mmol Gd³⁺/kg), which is 20-fold lower than the standard dose of conventional gadolinium agents (0.1 mmol Gd³⁺/kg). Blood sampling was performed to determine the bulk pharmacokinetic behavior of the nanoparticles, and T1-weighted MR images of the descending aorta were collected over 24 h to ascertain binding to the angiogenic microvasculature in the vessel wall.

A three-compartment pharmacokinetic model was developed to describe the *in vivo* behavior of the nanoparticles in both the blood pool as well as the aortic wall. Compartments 1 and 2 represent the bulk distribution of nanoparticles throughout the blood stream. The third compartment consists of the vasa vasorum of the aortic wall where the nanoparticles can specifically bind to the $\alpha_v\beta_3$ -integrin. The constants that describe the transfer into and out of this compartment, k_{13} and k_{31} , are lumped parameters that are used to describe both passive transfer and active binding.

Blood sampling after IV injection of nanoparticles demonstrated the half-life for the distribution phase was 20.2 min, while the elimination half-life was 11.9 h. MR imaging of the aortic wall revealed that the targeted nanoparticle agent generated at least two-fold higher signal enhancement at 0.5 and 1 h post injection than the nontargeted nanoparticles. Fitting both the blood clearance and tissue uptake data with

a three-compartment model showed no significant differences in k_{12} , k_{21} , v_1 , or k_{31} between the targeted and nontargeted groups. However, k_{13} (representing transfer from the blood into the aortic wall) and k_e (elimination from the blood pool) were significantly higher for the targeted formulation. While k_{13} represents the passive transfer from the blood into the third compartment in a standard three-compartment model, the nanoparticles are not just passively transferred but also actively binding within the third compartment. These two effects cannot be differentiated with this model and combine together, resulting in an apparent increase in k_{13} for the targeted nanoparticles compared to the nontargeted formulation. The pharmacokinetic parameters were used to calculate the tissue exposure to the nanoparticles. The local tissue concentration of targeted nanoparticles was doubled compared to nontargeted particles (1.64 vs 0.84 nmol Gd³⁺/g tissue/h, respectively). The amount of gadolinium that reached the aortic vasa vasorum was 0.18% of the total dose for nontargeted particles and 0.38% for targeted nanoparticles. These results demonstrate that MR can be used to noninvasively track binding of a molecularly targeted contrast agent, which has previously only been accomplished with nuclear and optical imaging. The gadolinium chelate on the particle surface could provide a surrogate marker for tracking the biodistribution of a lipophilic drug payload [9, 16, 17]. Rather than relying on blood pool measurements of drug concentrations, noninvasive imaging could serially monitor the actual concentration of the drug in the end organ and lead to more accurate modeling of therapeutic response.

Targeted Antiproliferative Therapy

In addition to antiangiogenic drugs, PFC nanoparticles have been formulated to deliver drugs that inhibit the pathological proliferation of various cell types. The progression of both cancer and atherosclerosis typically involve uncontrolled cell division and a wide range of antiproliferative drugs have been developed to fight these diseases. For example, smooth muscle cells line the luminal aspect of arteries and are involved in the progression of atherosclerosis. Tissue factor, a transmembrane glycoprotein with a prominent role in angiogenesis, thrombosis, cell signaling, hemostasis, and mitogenesis, is expressed on

smooth muscle cells in response to vascular injury. MR molecular imaging of tissue factor targeted PFC nanoparticles showed that the targeted particles bind to cultured smooth muscle cells at a concentration of 468 pM, while nontargeted particles only reach a concentration of 88 pM [3]. The MR measurements were validated by gas chromatography of the samples, revealing a particle concentration of 530 pM for the targeted nanoparticles and 111 pM for the nontargeted nanoparticles. Smooth muscle cell proliferation portends restenosis following coronary angioplasty. Therefore, the utility of PFC nanoparticles for delivery of antiproliferative drugs, including doxorubicin and paclitaxel, to smooth cells has been studied [2]. Exposure of the cells to targeted doxorubicin or paclitaxel nanoparticles significantly inhibited proliferation. Including the targeting ligand on the particle surface significantly increased the antiproliferative effect, particularly for paclitaxel. In vitro dissolution studies revealed that nanoparticle drug release persisted over 1 week. Targeted antiproliferative results were dependent on the hydrophobic nature of the drug and noncovalent interactions with other surfactant components. MR molecular imaging of the bound nanoparticles was demonstrated with both ^1H and ^{19}F MR imaging. A targeted paramagnetic nanoparticle carrying an antiproliferative payload to specific cell types could provide a novel, MR-visualizable, and quantifiable drug delivery system for the prevention of restenosis after angioplasty.

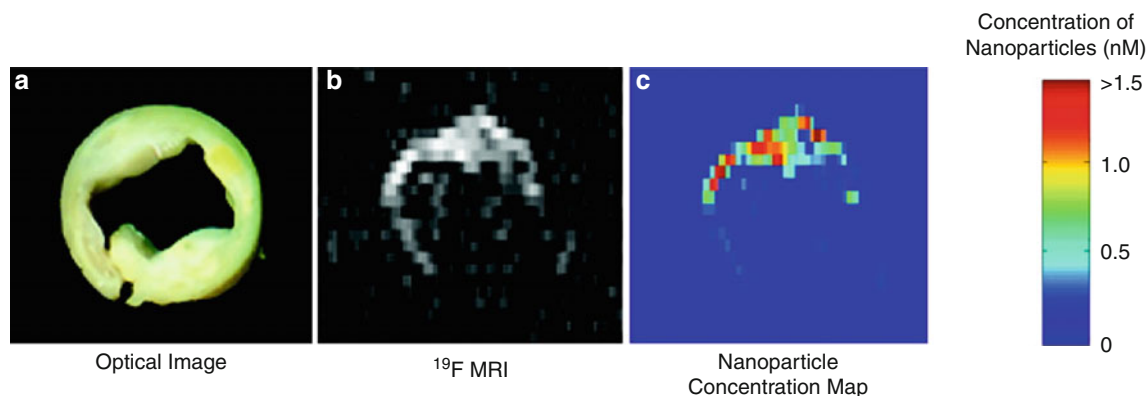
Clinically, drug-eluting stents are often used to locally release antiproliferative drugs for inhibiting restenosis following angioplasty. However, in-stent thrombosis can occur with these devices due to endothelial damage from the antiproliferative drugs. An alternative approach has been pursued with PFC nanoparticles loaded with antiproliferative agents and targeted to cellular biomarkers or the extracellular matrix [18]. Furthermore, these nanoparticle agents can be used to visualize ruptures in the artery wall resulting from the angioplasty balloon. Paramagnetic $\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles containing rapamycin were used to locally deliver a drug payload through microfractures created by balloon overstretch injury. The local delivery of rapamycin inhibited arterial stenosis without impeding endothelial healing [18]. MR imaging demonstrated that the integrin-targeted nanoparticles remained bound to the injured arterial wall 30–40 min after balloon overstretch and

reestablishment of blood flow. Two weeks after injury, MR imaging showed significant luminal stenosis in vessels treated with control formulations: $\alpha_v\beta_3$ -integrin-targeted nanoparticles without drug, nontargeted nanoparticles with rapamycin or saline. In vessels treated with $\alpha_v\beta_3$ -integrin-targeted rapamycin nanoparticles, no lumen irregularities were observed on MR angiograms. Histology showed $\alpha_v\beta_3$ -integrin-targeted rapamycin nanoparticles decreased neointimal proliferation by 40% compared to controls. Endothelial healing was evaluated 1, 7, 14, and 28 days after balloon injury and showed no differences between $\alpha_v\beta_3$ -integrin-targeted rapamycin nanoparticles and the controls.

Another therapeutic approach to inhibit cellular proliferation uses host defense peptides that embed in the phospholipid cell membrane to create pores and disrupt normal cellular physiology. Melittin is a cytolytic peptide derived from honeybee venom that ruptures lipid membranes. PFC nanoparticles conjugated with melittin have been used as a therapeutic delivery vehicle in cancer models [19]. Melittin PFC nanoparticles were injected intravenously in mice bearing MDA-435 tumors [19]. Ultrasound imaging showed that treated tumors grew 25% slower than controls. In a mouse melanoma model, four intravenous injections of melittin nanoparticles reduced the tumor size by 80% compared to control injections of saline or nanoparticles lacking melittin. In a genetically engineered model of squamous carcinoma, treatment with integrin-targeted melittin nanoparticles resulted in an 80% reduction in dermal papillae, the site of dysplasia in this model. In cultured cells, free melittin causes necrosis resulting in the release of lactate dehydrogenase (LDH) due to the disruption of cell membranes and cytochrome c indicating mitochondria damage. Cultured cells exposed to targeted melittin nanoparticles showed cytochrome c release, but only minimal LDH, suggesting a shift to apoptotic cell death rather than necrotic cell death. Thus, drug delivery with PFC nanoparticles could be utilized to modulate the therapeutic mechanism of action.

^{19}F MR Molecular Imaging

MR detection of paramagnetic PFC nanoparticles typically relies on the effect of gadolinium on the ^1H MR



Perfluorocarbon Nanoparticles, Fig. 5 Direct detection of fibrin-targeted PFC nanoparticles utilizing ^{19}F MR imaging. (a) Optical image of a cross section of a human carotid endarterectomy sample showing moderate luminal narrowing as well as several atherosclerotic lesions. (b) A ^{19}F projection image of the

carotid artery sample showing high signal along the lumen due to nanoparticles bound to fibrin. (c) Concentration map of bound nanoparticles on the carotid artery (Reprinted with permission from Morawski et al. [20])

signal. The paramagnetic nature of gadolinium increases the water relaxation rate and generates signal enhancement on T1-weighted MR images. An alternative approach achieves direct detection of the PFC core of the nanoparticles via ^{19}F MR imaging [1, 3]. ^{19}F has a relatively high sensitivity, 83% compared to ^1H , and there is virtually no native background signal. Thus, the PFC core can provide a definitive and quantitative MR signature [1]. For example, ^{19}F MR of fibrin-targeted PFC nanoparticles could be used to map the formation of thrombi on ruptured atherosclerotic plaques and quantify the extent of ruptures in the fibrous cap. Human carotid endarterectomy samples were treated with fibrin-targeted paramagnetic nanoparticles and imaged at 4.7 T [20]. ^1H MR imaging showed high levels of signal enhancement along the luminal surface due to nanoparticle binding to fibrin deposits (Fig. 5). A ^{19}F projection image of the artery showed an asymmetric distribution of nanoparticles around the vessel wall corroborating the ^1H signal enhancement. Co-registration of the quantitative nanoparticle map with the ^1H image permitted visualization of both anatomical and pathological information in a single image. Combining information from ^1H and ^{19}F MR imaging could allow prediction of subsequent occlusion or distal embolization from unstable or disrupted plaques, and aid clinical decision-making for acute invasive intervention versus pharmaceutical therapies.

PFC nanoparticles can be formulated with different perfluorochemicals to yield particles that can be

differentiated by ^{19}F MR. Fibrin-targeted nanoparticles consisting of perfluorooctylbromide (PFOB) or perfluoro-15-crown-5-ether (CE) have been distinguished based on the unique chemical shifts of each compound [1]. Both ^{19}F imaging and spectroscopy could identify PFOB and CE nanoparticles. The signal to noise for PFOB was lower than CE (10 vs 25, respectively), presumably due to the single CE peak (20 equivalent fluorine atoms) compared to the multiple PFOB peaks (17 fluorine atoms distributed over 5 peaks). A clear linear relationship between the ^{19}F signal intensity and nanoparticle concentration was demonstrated for both PFOB and CE using both imaging and spectroscopy. Multiple PFC nanoparticle agents could be used to target different epitopes and achieve a noninvasive analogy to immunohistochemistry. For example, simultaneous quantification of angiogenesis in the vessel wall and fibrin deposition on the plaque cap could be used to evaluate the pathophysiological stage of an obstructive atherosclerotic lesion.

Tissue inflammation is another biological marker of unstable atherosclerotic plaques, inducing the release of cytokines and upregulating the production of proteins like vascular cell adhesion molecule-1 (VCAM-1). PFC nanoparticles targeted to VCAM-1 were injected into genetically engineered ApoE $^{-/-}$ mice, which mimic the clinical progression of atherosclerosis, to map inflammation [3]. These mice display focal inflammation and macrophage infiltration in the kidneys. To definitively identify nanoparticle binding in the kidneys, ^{19}F MR imaging was performed on a 11.7 T

research scanner 2 h after injection of VCAM-1 targeted PFC nanoparticles. The ^{19}F signal arising from the PFC core provided an unambiguous marker of particle accumulation, without the innate signal variations that can confound typical ^1H signal enhancement with traditional paramagnetic or superparamagnetic MR contrast agents. VCAM-1-targeted nanoparticles accumulated in ApoE $^{-/-}$ kidneys to a greater extent than nontargeted nanoparticles (3.7 billion particles per gram of tissue vs 0.9 billion particles/gram). The uptake of targeted nanoparticles was also higher in the kidneys of ApoE $^{-/-}$ mice compared to non-ApoE $^{-/-}$ controls (3.7 vs 1.6 billion particles/gram). Control animals displayed no significant difference in the uptake of targeted versus nontargeted nanoparticles (1.6 vs 1.5 billion particles/gram).

Typically, targeted MR contrast agents generate image enhancement as a result of both nonspecific blood pool signal as well as the specific binding of the agent to the biomarker of interest. Separating out these two contributions can be very difficult to achieve in vivo. One method to suppress the nonspecific signal utilizes diffusion-weighted ^{19}F spectroscopy to null the signal arising from moving particles, representing the unbound fraction in the blood pool, while retaining the signal from stationary particles that are specifically bound to the target epitope [1]. A genetically engineered mouse model of squamous cell cancer, derived by incorporating human papilloma virus into the mouse genome, was studied with a 11.7 T research scanner. Transgenic and age-matched control mice were scanned 90 min after injection of $\alpha_v\beta_3$ -integrin-targeted PFC nanoparticles to assess angiogenesis in these precancerous lesions. Both the transgenic and control mice displayed decreased ^{19}F signal with increasing b-values. However, 60–100% of the ^{19}F signal remained at b-values near 60,000 s/mm 2 in the transgenic animals, while no detectable ^{19}F signal was observed in control mice at b-values of 1,500 s/mm 2 . The calculated ADC of PFC nanoparticles was 33.1 m 2 /s in the transgenic mice, significantly lower than the 19,563 m 2 /s ADC in the controls.

Angiogenesis is a prominent feature in the progression of aortic valve stenosis. Cholesterol-fed rabbits develop aortic valve sclerosis, characterized by gross thickening, macrophage infiltration, calcification, and eventual bone formation that mimics the clinical presentation of the disease. Cholesterol-fed rabbits underwent ^{19}F MR imaging after injection of

$\alpha_v\beta_3$ -integrin-targeted nanoparticles to quantify angiogenesis in the aortic valve leaflets [1]. The cholesterol feeding caused gross thickening of the aortic valves accompanied by extensive foam cell infiltration, non-calcified bone deposition, activation of myofibroblasts, abnormal microvascular proliferation, and upregulation of $\alpha_v\beta_3$ -integrin expression. None of these abnormalities were observed in the normal valve tissue from control animals.

The crown ether (CE) signal arising from the nanoparticle contrast agent was readily detected and distinguished from a perfluorooctylbromide (PFOB) quantification reference based on the chemical shifts of these perfluorocarbon species, allowing quantification of the total volume of bound nanoparticles. The volume of targeted nanoparticles bound to the valves was 19.5 nL, which was more than three times higher than the amount of nontargeted nanoparticles (5.6 nL). Competitive inhibition of $\alpha_v\beta_3$ -integrin binding reduced the amount of nanoparticles in the valves by about half (10.3 nL). Valves from healthy rabbits treated with targeted nanoparticles contained almost nine times fewer nanoparticles (2.3 nL) than the valves from cholesterol-fed rabbits. These techniques may be useful for assessing atherosclerotic components of pre-clinical aortic valve disease in patients and could assist in defining efficacy of medical therapies. The sensitivity of this approach for molecular detection of sparse quantities of inflammatory epitopes in very thin structures at high field strengths establishes a basis for future efforts to develop localized spectroscopic methods at clinical field strengths that could be useful for detecting disease and monitoring therapies.

Conclusions

PFC nanoparticles have been utilized for a wide range of molecular imaging, cell tracking, and targeted drug delivery applications. The relatively large size of the particles constrains them to the intravascular space, allowing specific targeting to endothelial expression of integrins, selectins, and adhesion molecules. PFC nanoparticles have a large surface area, which provides a means to incorporate multiple binding ligands and multiple paramagnetic chelates to the particle surface. The particle surface can also serve as a drug reservoir that selectively delivers a therapeutic payload to the targeted cell population. The dual role of PFC

nanoparticles as both a molecular imaging contrast agent and a drug delivery vehicle can be exploited to provide diagnostic and therapeutic information via noninvasive MR imaging, or most of the other common clinical imaging modalities, including ultrasound, nuclear imaging, and computed tomography. Although PFC nanoparticles not currently approved for clinical use, the future applications of these agents in the management of patients with chronic diseases could aid in diagnostics, therapeutics, and/or follow-up of treatment efficacy or disease recurrence, by noninvasively mapping biomarkers associated with cancer or cardiovascular disease and specifically directing highly potent drugs directly to the sites of pathology.

Cross-References

- ▶ [Angiogenesis](#)
- ▶ [Imaging Human Body Down to Molecular Level](#)
- ▶ [Intravital Microscopy Analysis](#)
- ▶ [Liposomes](#)
- ▶ [Microfabricated Probe Technology](#)
- ▶ [Nanomedicine](#)
- ▶ [Nanotechnology in Cardiovascular Diseases](#)

References

1. Chen, J., Lanza, G.M., Wickline, S.A.: Quantitative magnetic resonance fluorine imaging: today and tomorrow. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2**, 431–440 (2010)
2. Pan, D., Caruthers, S.D., Chen, J., Winter, P.M., Senpan, A., Schmieder, A.H., Wickline, S.A., Lanza, G.M.: Nanomedicine strategies for molecular targets with MRI and optical imaging. *Future Med. Chem.* **2**, 471–490 (2010)
3. Winter, P.M., Caruthers, S.D., Lanza, G.M., Wickline, S.A.: Quantitative cardiovascular magnetic resonance for molecular imaging. *J. Cardiovasc. Magn. Reson.* **12**, 62 (2010)
4. Culver, J., Akers, W., Achilefu, S.: Multimodality molecular imaging with combined optical and SPECT/PET modalities. *J. Nucl. Med.* **49**, 169–172 (2008)
5. Lanza, G.M., Winter, P.M., Caruthers, S.D., Hughes, M.S., Hu, G., Schmieder, A.H., Wickline, S.A.: Theragnostics for tumor and plaque angiogenesis with perfluorocarbon nanoemulsions. *Angiogenesis* **13**, 189–202 (2010)
6. Winter, P.M., Shukla, H.P., Caruthers, S.D., Scott, M.J., Fuhrhop, R.W., Robertson, J.D., Gaffney, P.J., Wickline, S.A., Lanza, G.M.: Molecular imaging of human thrombus with computed tomography. *Acad. Radiol.* **12**(Suppl 1), S9–13 (2005)
7. Lijowski, M., Caruthers, S., Hu, G., Zhang, H., Scott, M.J., Williams, T., Erpelding, T., Schmieder, A.H., Kiefer, G., Gulyas, G., et al.: High sensitivity: high-resolution SPECT-CT/MR molecular imaging of angiogenesis in the Vx2 model. *Invest Radiol.* **44**, 15–22 (2009)
8. Schmieder, A.H., Caruthers, S.D., Zhang, H., Williams, T.A., Robertson, J.D., Wickline, S.A., Lanza, G.M.: Three-dimensional MR mapping of angiogenesis with alpha5beta1(alpha nu beta3)-targeted theranostic nanoparticles in the MDA-MB-435 xenograft mouse model. *FASEB J.* **22**, 4179–4189 (2008)
9. Winter, P.M., Caruthers, S.D., Zhang, H., Williams, T.A., Wickline, S.A., Lanza, G.M.: Antiangiogenic synergism of integrin-targeted fumagillin nanoparticles and atorvastatin in atherosclerosis. *JACC Cardiovasc. Imaging* **1**, 624–634 (2008)
10. Zhou, H.F., Chan, H.W., Wickline, S.A., Lanza, G.M., Pham, C.T.: Alphanubeta3-targeted nanotherapy suppresses inflammatory arthritis in mice. *FASEB J.* **23**, 2978–2985 (2009)
11. Horton, M.A.: The alpha v beta 3 integrin “vitronectin receptor”. *Int. J. Biochem. Cell Biol.* **29**, 721–725 (1997)
12. Cai, K., Caruthers, S.D., Huang, W., Williams, T.A., Zhang, H., Wickline, S.A., Lanza, G.M., Winter, P.M.: MR molecular imaging of aortic angiogenesis. *JACC Cardiovasc. Imaging* **3**, 824–832 (2010)
13. Winter, P.M., Caruthers, S.D., Allen, J.S., Cai, K., Williams, T.A., Lanza, G.M., Wickline, S.A.: Molecular imaging of angiogenic therapy in peripheral vascular disease with alphanubeta3-integrin-targeted nanoparticles. *Magn. Reson. Med.* **64**, 369–376 (2010)
14. Winter, P.M., Schmieder, A.H., Caruthers, S.D., Keene, J.L., Zhang, H., Wickline, S.A., Lanza, G.M.: Minute dosages of alpha(nu)beta3-targeted fumagillin nanoparticles impair Vx-2 tumor angiogenesis and development in rabbits. *FASEB J.* **22**, 2758–2767 (2008)
15. Boles, K.S., Schmieder, A.H., Koch, A.W., Carano, R.A., Wu, Y., Caruthers, S.D., Tong, R.K., Stawicki, S., Hu, G., Scott, M.J., et al.: MR angiogenesis imaging with Robo4-vs. alphaVbeta3-targeted nanoparticles in a B16/F10 mouse melanoma model. *FASEB J.* **24**, 4262–4270 (2010)
16. Winter, P.M., Neubauer, A.M., Caruthers, S.D., Harris, T.D., Robertson, J.D., Williams, T.A., Schmieder, A.H., Hu, G., Allen, J.S., Lacy, E.K., et al.: Endothelial alpha(v)beta3 integrin-targeted fumagillin nanoparticles inhibit angiogenesis in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **26**, 2103–2109 (2006)
17. Lanza, G.M., Yu, X., Winter, P.M., Abendschein, D.R., Karukstis, K.K., Scott, M.J., Chinen, L.K., Fuhrhop, R.W., Scherrer, D.E., Wickline, S.A.: Targeted antiproliferative drug delivery to vascular smooth muscle cells with a magnetic resonance imaging nanoparticle contrast agent: implications for rational therapy of restenosis. *Circulation* **106**, 2842–2847 (2002)
18. Cyrus, T., Zhang, H., Allen, J.S., Williams, T.A., Hu, G., Caruthers, S.D., Wickline, S.A., Lanza, G.M.: Intramural delivery of rapamycin with alphanubeta3-targeted paramagnetic nanoparticles inhibits stenosis after balloon injury. *Arterioscler. Thromb. Vasc. Biol.* **28**, 820–826 (2008)
19. Soman, N.R., Baldwin, S.L., Hu, G., Marsh, J.N., Lanza, G.M., Heuser, J.E., Arbeit, J.M., Wickline, S.A., Schlesinger,

P.H.: Molecularly targeted nanocarriers deliver the cytolytic peptide melittin specifically to tumor cells in mice, reducing tumor growth. *J. Clin. Invest.* **119**, 2830–2842 (2009)

20. Morawski, A.M., Winter, P.M., Yu, X., Fuhrhop, R.W., Scott, M.J., Hockett, F., Robertson, J.D., Gaffney, P.J., Lanza, G.M., Wickline, S.A.: Quantitative “magnetic resonance immuno-histochemistry” with ligand-targeted (19)F nanoparticles. *Magn. Reson. Med.* **52**, 1255–1262 (2004)

Perfluoropolyethers

- ▶ [Boundary Lubrication](#)

Petal Effect

- ▶ [Rose Petal Effect](#)

Petroleum Lubricants

- ▶ [Boundary Lubrication](#)

PFC Nanoparticles

- ▶ [Perfluorocarbon Nanoparticles](#)

PFOB Nanoparticles

- ▶ [Perfluorocarbon Nanoparticles](#)

Photoetching

- ▶ [Chemical Milling and Photochemical Milling](#)

Photofabrication

- ▶ [Chemical Milling and Photochemical Milling](#)

Photolithography

Marc Madou¹ and Chunlei Wang²

¹Department of Mechanical and Aerospace Engineering & Biomedical Engineering, University of California at Irvine, Irvine, CA, USA

²Department of Mechanical and Materials Engineering, Florida International University, Miami, FL, USA

Synonyms

[Optical lithography](#)

Definition

Photolithography is a process used in microfabrication to selectively pattern a thin film substrate using light to transfer a geometric pattern from a photomask to a light-sensitive photoresist.

Introduction

Lithography is a technique used to transfer copies of a master pattern onto the surface of a solid material such as a silicon wafer. The word *lithography* (Greek for the words *stone* [*lithos*] and to *write* [*gráphein*]) refers to the process invented in 1796 by Aloys Senefelder, who inked Bavarian limestone and transferred a carved image from stone onto paper. The most widely used form of lithography is photolithography. In the IC industry, pattern transfer from masks onto thin films is accomplished almost exclusively via photolithography. Photomasking, followed by chemical processing, led to the photolithography now used in fabricating ICs (integrated circuits) and in miniaturization science. The first applications of the printed circuit board (PCB) (also printed wiring board or PWB) were invented in 1943 by the German Paul Eisler. By 1961, Fairchild Semiconductor had introduced the first commercial integrated circuit in which photo-etching processes produce large numbers of transistors on a thin slice of silicon (Si). In 2008, integrated circuits were ubiquitous and they could be found in microprocessors, audio and video equipment,

dishwashers, garage openers, security systems, automobiles, etc. With the first ICs, patterns had a resolution not better than 5 μm . Today, photolithography, x-ray lithography, and charged particle lithography all achieve submicron-printing accuracy. Intel introduced the Prescott Pentium IV chip in the fourth quarter of 2003, an IC chip incorporating 90 nm sized features, using UV photolithography. Chips based on a 65 nm technology became available in 2006 and the 32 nm node was achieved in 2009.

The combination of accurate registration and exposing a series of successive patterns leads to complex multilayered ICs. This essentially two-dimensional (2D) process has a limited tolerance for nonplanar topography, creating a major constraint for building non-IC miniaturized systems, which often exhibit extreme topographies. Photolithography has matured rapidly and its ability to resolve ever-smaller features is constantly improving. For the IC industry, this continued improvement in photolithography resolution has impeded the adaptation of alternative, higher-resolution lithography techniques, such as x-ray lithography. Research over the last 10 years, in high-aspect-ratio resists features to satisfy the needs of both IC and non-IC miniaturization, is also finally improving dramatically photolithography's capacity to cover wide ranges of topography. Performance of a photolithographic process is determined by its resolution, the minimum feature size that can be transferred with high fidelity, the registration, how accurately patterns on successive masks can be aligned and throughput, the number of wafers that can be transferred per hour. Photolithography and pattern transfer involve a set of process including photoresist coating, UV exposure, baking, developing, etc.

Wafer Cleaning and Clean Room

An important step, even before lithography proper, is wafer cleaning. Contaminants include solvent stains (methyl alcohol, acetone, trichloroethylene, isopropyl alcohol, xylene, etc.), dust from operators and equipment, smoke particles, etc. Stains or films may lead to adverse effects during oxidation and evaporation processes. Particulates, chunks of granular matter, may cause undesirable masking effects and scratches on the photomask during contact printing. All lithography processes take place inside a semiconductor clean room, which is a specially constructed enclosed area

environmentally controlled with respect to airborne particulates, temperature ($\pm 0.1^\circ\text{F}$), air pressure, humidity (from 0.5% to 5% RH), vibration, and lighting. In a Class 1 clean room, the particle count does not exceed 1 particle per cubic foot with particles of a size of 0.5 μm and larger, and in a Class 100 clean room, the particle count does not exceed 100 particles per cubic foot with particles of a size of 0.5 μm and larger. The allowable contamination particle size in IC manufacture has been decreasing hand in hand with the ever-decreasing minimum feature size.

Many different dry and wet methods for wafer cleaning currently in use, such as RCA1 and RCA2, developed by W. Kern, use mixtures of hydrogen peroxide and various acids or base followed by deionized (DI) water rinses. Others include vapor cleaning; thermal treatment, for example, baking at 1,000°C in vacuum or in oxygen; and plasma or glow discharge techniques, for example, in freons with or without oxygen. Mechanical methods include ultrasonic agitation, polishing with abrasive compounds, and supercritical cleaning. Ultrasonic cleaning, which is excellent for removing particulate matter from the substrate, is unfortunately prone to contamination and mechanical failure of deposited films.

Masks

The stencil used to repeatedly generate a desired pattern on resist-coated wafers is called a *mask*. In typical use, a photomask – a nearly optically flat glass (transparent to near ultraviolet [UV]) or quartz plate (transparent to deep UV) with an absorber pattern metal (e.g., an 800 Å thick chromium layer) – is placed above the photoresist-coated surface, and the mask/wafer system is exposed to UV radiation. The absorber pattern on the mask can be generated by e-beam lithography which yields higher resolution than photolithography. A *positive* or *dark field* mask is a mask on which the pattern is clear with the background dark. A *negative* or *clear field* mask is a mask on which the pattern is dark with the background clear. A light field or dark field image, known as *mask polarity*, is then transferred to the semiconductor surface. This procedure results in a 1:1 image of the entire mask onto the silicon wafer.

Masks, making direct physical contact with the substrate, are called *contact masks*. Unfortunately, these masks degrade faster due to wear than

noncontact, proximity masks (also referred to as *soft contact masks*), which are slightly raised, say 10–20 μm , above the wafer. Contact mask and proximity mask printing are collectively known as *shadow printing*. A more reliable method of masking is *projection printing* where, rather than placing a mask in direct contact with (or in proximity of) a wafer, the photomask is imaged by a high-resolution lens system onto the resist-coated wafer. The imaging lens can reduce the mask pattern by 1:5 or 1:10, making mask fabrication less challenging. Lithography is still mostly carried out using masks, but due to problems caused by masks, such as expense and time in fabricating them, contamination introduced by them, their disposal, and the difficulties in their alignment, research into maskless optical projection lithography (MOPL) is growing rapidly and broadly. One MOPL, approach already on the market (<http://www.intelligentmp.com>), is based on the Digital Micromirror Device (DMD) chip from Texas Instruments Inc. (TI), and relies on the same spatial light modulation technology used in DLP (digital light processing) projectors and HDTVs (high definition television) arrays in a DMD chip to project images on the photoresist. The resolution of DMD-based maskless photolithography (about 5 μm) is significantly less than with e-beam lithography (0.25 μm) or laser writers (<1.0 μm), but it is a parallel technique and for many applications, that is, in microfluidics, the lower resolution might not be an obstacle. Maskless zone-plate-array lithography (ZPAL) is another optical maskless technology which operates on the principle of diffraction rather than refraction.

Process

Photoresist Deposition

Lithography is the most expensive step in microelectronics technology, representing up to 35% of the wafer manufacturing cost. Within lithography, photoresist coating is one of the more expensive steps. For silicon integrated circuits, the resist thickness after a prebake (see below for what a prebake does) typically ranges between 0.5 and 2 μm . For miniaturized three-dimensional (3D) structures, much greater resist thicknesses are often required and complex topography might also call for a conformal resist coat over very high-aspect-ratio features. For thick

resist coats, techniques such as casting and the use of thick sheets of dry photoresists replace the ineffective resist spinners.

As the first step in the lithography process itself, a thin layer of an organic polymer, a photoresist sensitive to ultraviolet radiation, is deposited on the oxide surface. The photoresist is dispensed onto the wafer which is held by a vacuum chuck. A speed of about 500 rpm is commonly used during the dispensing step enabling the spread of the fluid over the substrate. After the dispense step it is common to accelerate to a relatively high speed to thin the fluid to near its final desired thickness. Typical spin speeds for this step range from 1,500 to 6,000 rpm, depending on the properties of the fluid (mostly its viscosity) as well as the substrate. This step can take from 10 s to several minutes. The combination of spin speed and time selected for this step will generally define the final film thickness. At these speeds, centrifugal force causes the solution to flow to the edges, where it builds up until expelled when surface tension is exceeded. The resulting polymer thickness, T , is a function of spin speed, solution concentration, and molecular weight (measured by intrinsic viscosity).

As alternatives to spin coating, several other methods can be used for photoresist application, including spray coating, electrostatic spraying or electrostatic deposition (ED), meniscus coating, silk screening, plasma-deposition, electrophoretic photoresist deposition, dip coating, roller coating, curtain coating, extrusion coating, etc.

Soft Baking or Prebaking

After resist coating, the resist still contains up to 15% solvent and may contain built-in stresses. The wafers are therefore soft baked (also pre-exposure baked or prebaked) at 90–100°C for about 20 min in a convection oven, or at 75–85°C for 1–3 min with a vacuum hot plate, to remove solvents and stress and to promote adhesion of the resist layer to the substrate. This is a critical step in that failure to sufficiently remove the solvent will affect the resist profile. Excessive baking destroys the photoactive compound and reduces sensitivity. Thick resists may benefit from a longer bake time. The resist thickness, for both negative and positive resists, is typically reduced by 10–25% during soft baking. Hot plating the resist is faster, more controllable, and does not trap solvent like convection oven baking. In convection ovens, the solvent at the surface

of the resist is evaporated first, and this can cause an impermeable resist skin, trapping the remaining solvent inside. So heating in a convection oven must proceed slowly to avoid solvent bursts. Commercially, microwave heating or IR lamps are also used in production lines. The optimization of the prebaking step may substantially increase device yield.

Exposure and Postexposure Treatment

After soft baking, the resist-coated wafers are transferred to an illumination or exposure system where they are aligned with the features on the mask. For any lithographic technique to be of value, it must provide an alignment technique capable of a superposition precision of mask and wafer that is a small fraction of the minimum feature size of the devices under construction. In the simplest case, an exposure system consists of a UV lamp illuminating the resist-coated wafer through a mask without any lenses between the two. The purpose of the illumination is to deliver light with the proper intensity, directionality, spectral characteristics, and uniformity across the wafer, allowing a nearly perfect transfer or printing of the mask image onto the resist in the form of a latent image.

In photolithography, wavelengths of the light source used for exposure of the resist-coated wafer range from the very short wavelengths of extreme ultraviolet (EUV) (10–14 nm) to deep ultraviolet (DUV) (150–300 nm) to near ultraviolet (UV) (350–500 nm). In near UV, one typically uses the g-line (435 nm) or i-line (365 nm) of a mercury lamp. The brightness of shorter-wavelength sources is severely reduced compared to that of longer-wavelength sources, and the addition of lenses further reduces the efficiency of the exposure system. In general, the smallest feature that can be printed using projection lithography is roughly equal to the wavelength of the exposure source, in this example, 248 nm would be expected. The same laser, in combination with sophisticated resolution-enhancing techniques (RETs), may be used to produce more advanced circuits with transistor gate features of 160 nm and below. RET methods (see below) enables one to go quite a bit beyond the conventional Rayleigh diffraction limit. Other exposure systems now available include two DUV excimer lasers, the ArF at 193 nm and the F₂ at 157 nm, and extreme UV lithography (EUV) at 13.4 nm. The current generation of lithography is

using 193 nm light from ArF lasers. In the case of EUV, a plasma or synchrotron source and all reflective reduction optics (4×) (at this wavelength all materials absorb!) are used.

Postexposure Treatment

A postexposure treatment of the exposed photoresist is often desired, because the reactions initiated during exposure might not have run to completion. To halt the reactions or to induce new ones, several postexposure treatments are in use: postexposure baking (PEB), flood exposure with other types of radiation, treatment with a reactive gas, and vacuum treatment. Postexposure baking (sometimes in a vacuum) and treatment with reactive gas are used in image reversal and dry resist development. In the case of a chemically amplified resist, the postexposure bake is most critical. Although reactions induced by the catalyst that forms during exposure take place at room temperature, their rate is greatly increased by baking at 100°C. The precise control of this type of postexposure bake critically determines the subsequent development itself.

Development

Development is the dissolution of unpolymers resist that transforms the latent resist image, formed during exposure, into a relief image that will serve as a mask for further subtractive and additive steps. During the development of an exposed resist, selective dissolving takes place. Two main technologies are available for development: wet development, which is widely used in circuit and miniaturization manufacture in general, and dry development, which is starting to replace wet development for some of the ultimate line-width resolution applications.

During batch spray development, fan-type sprayers direct fresh developing solution across wafer surfaces. Positive resists are typically developed in aqueous alkaline solutions, and negative resists in organic ones. For alkaline developers, simple solutions of NaOH (Shipley 351), or KOH (AZ/Clariant 400 K) may be used, but because of the possibility of mobile ion contamination in MOS devices, developers free of metal ions are preferred. The latter are usually tetramethyl ammonium hydroxide (TMAH) based (Shipley CD-26, MF-321, AZ/Clariant 300-MIF). Each developer has a different dilution, and some require longer development times than others. Developers are

generally matched to a type of photoresist. Though they may be interchangeable to some extent, changing the type of developer used in a process will usually change the exposure time necessary to resolve the pattern. It should be noted that all of the above mentioned alkaline developers etch aluminum. If the micromachine or microchip, under construction, contains aluminum features, Shipley Microposit Developer Concentrate (MDC), a mixture of proprietary alkaline salts (mostly phosphates) with the slowest aluminum etch rate, is a better developer choice.

The use of organic solvents leads to some swelling of the resist (especially for negative resists; see below) and loss of adhesion of the resist to the substrate. Dry development overcomes these problems, as it is based either on a vapor phase process or a plasma. In the latter, oxygen-reactive ion etching (O_2 -RIE) is used to develop the latent image. The image formed during exposure exhibits a differential etch rate to O_2 -RIE rather than differential solubility to a solvent.

De-scumming and Post-baking

A mild oxygen plasma treatment, so-called de-scumming, removes unwanted resist left behind after development. Negative, and to a lesser degree positive, resists leave a thin polymer film at the resist/substrate interface. The problem is most severe in small ($< 1 \mu\text{m}$) high-aspect-ratio structures where the mass transfer of a wet developer is poor. Patterned resist areas are also thinned in the de-scumming process, but this is usually of little consequence. Before etching the substrate or adding a material, the wafer must be post-baked. Post-baking or hard baking removes residual coating solvent and developer, and anneals the film to promote interfacial adhesion of the resist that has been weakened either by developer penetration along the resist/substrate interface or by swelling of the resist (mainly for negative resists). Hard baking also improves the hardness of the film and avoids solvent bursts during vacuum processing. Improved hardness increases the resistance of the resist to subsequent etching steps. Post-baking frequently occurs at higher temperatures (120°C) and for longer times (say 20 min) than soft or prebaking. The major limitation for heat application is excessive flow or melt, which degrades wall profile angles and makes it more difficult to remove the resist. Post-bake does induce some stress and resist shrinkage. Special care needs to be taken when the baking temperature is

above the glass transition temperature, T_g , when impurities are easily incorporated into the resist due to the plastic flow of the resist.

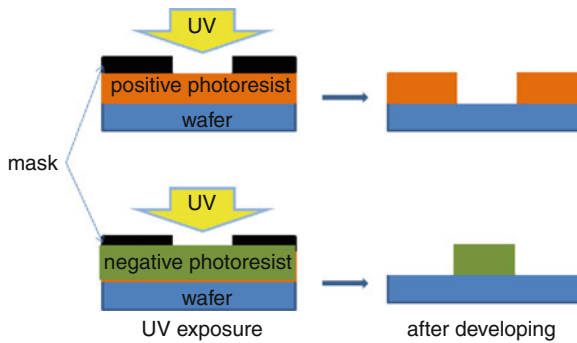
Resists

The principal components of photoresists are a polymer (base resin), a sensitizer, and a casting solvent. The polymer changes structure when exposed to radiation; the solvent allows spin application and formation of thin layers on the wafer surface; sensitizers control the chemical reactions in the polymeric phase. Resists without sensitizers are single-component or one-component systems, whereas sensitizer-based resists are two-component systems. Solvent and other potential additives do not directly relate to the photoactivity of the resist.

If the photoresist is of the type called *positive* (also *positive tone*), the photochemical reaction during exposure of a resist weakens the polymer by rupture or scission of the main and side polymer chains, and the exposed resist becomes more soluble in developing solutions (say ten times more soluble). In other words, the development rate, R , for the exposed resist is about ten times faster than the development rate, R_0 , for the unexposed resist. Two well-known families of positive photoresists are the single component poly (methylmethacrylate) (PMMA) resists and the two-component DNQ resists comprised of a photoactive component (PAC) such as diazonaphthoquinone ester (DNQ) (20–50 wt%) and a phenolic novolak resin.

If the photoresist is of the type called *negative* (also *negative tone*), the reaction strengthens the polymer by random cross-linkage of main chains or pendant side chains, becoming less soluble (slower dissolving). Exposure, development, and pattern-transfer sequences for negative and positive resists are shown in Fig. 1.

The first negative photoresists were based on free-radical-initiated photo-cross-linking processes of main or pendant polymer side chains, rendering the exposed parts insoluble. Commonly used negative-acting, two-component resists are bis(aryl)azide rubber resists, whose matrix resin is cyclized poly(cis-isoprene), a synthetic rubber. An example of a commercial, two-component negative photoresist is the Kodak KTRF [an azide-sensitized poly(isoprene) rubber] with a lithographic sensitivity (also photospeed) of $75\text{--}125 \text{ mJ/cm}^2$. Negative photoresists, in general,



Photolithography, Fig. 1 Positive and negative resist: exposure, development, and pattern transfer

adhere very well to the substrate, and a vast amount of compositions are available (stemming from R&D work in paints, UV curing inks, and adhesives all based on polymerization hardening). Negative resists are highly resistant to acid and alkaline aqueous solutions as well as to oxidizing agents. As a consequence, a given thickness of negative resist is more chemically resistant than a corresponding thickness of positive resist. A comparison of negative and positive photoresist features is presented in Table 1. This table is not exhaustive and is meant only as a practical guide for selection of a resist tone.

Resists typically are removed (stripped) once they have served their function as temporary stencils. Some negative resists, hardened through UV exposure, are used as permanent components of miniature devices. Two prominent examples in this category are polyimide and SU-8-based resists. Polyimides are a class of polymers synthesized from two monomers: a dianhydride and a diamine. Commercial products are supplied as soluble polyamic acid (PAA) intermediates, which undergo a thermal imidization with the evolution of water to form an insoluble polyimide. Polyimide is used as an interlayer dielectric because of its desirable characteristics such as a high thermal stability with a glass transition temperature, $T_g > 300^\circ\text{C}$, a low dielectric constant ϵ , good chemical stability, and ease of processing.

Most resists in IC and MEMS fabrication are deposited as liquids, whereas resists used in printed wiring board (PWB) manufacture are usually dry film resists that come in rolls (ranging from 2 to 60 in. wide and 125 to 1,000 ft. long) and are laminated onto the substrate instead of being spin coated on it. Dry film

resist formulations are sandwiched between a polyolefin release sheet and a polyester base, rolled up on a support core. Dry film resists offer advantages such as excellent adhesion on most substrates, no liquid handling since there is no solvent, high process speed, excellent thickness uniformity over a whole substrate, simple handling, no formation of edge beads, low exposure energy, low cost, short processing time, and near vertical sidewalls. There are a variety of dry film photoresists widely used and commercially available, examples include Riston[®], Ordy BF 410, Etertec[®] 5600, DF 4615, and DFR-15. All are used in the manufacture of circuit boards and can be made quite thick and all are candidates for broad use in MEMS as well.

Photoresist stripping, in slightly oversimplified terms, is organic polymer etching. The primary consideration is complete removal of the photoresist without damaging the device under construction. There are many commercial strippers available; some are specific for positive resists (e.g., ACT-690 C from Ashland), others for negative resists, and still others are universal strippers (e.g., ACT-140 from Ashland). Other popular commercial strippers are Piranha and RCA clean. Wet stripping solutions lose potency in use, causing stripping rates to change with time. Accumulated contamination in solutions can be a source of particles, and liquid phase surface tension and mass transport tend to make photoresist removal difficult and uneven. Dry stripping or oxygen plasma stripping, also known as ashing, has become more popular as it poses fewer disposal problems with toxic, flammable, and dangerous chemicals. Dry stripping is more controllable than liquid stripping, less corrosive with respect to metal features on the wafer, and, more importantly, it leaves a cleaner surface under the right conditions. Finally, it does not cause the undercutting and broadening of photoresist features that can be caused by wet strippers.

Critical Dimension and Overall Resolution

The absolute size of a minimum feature in an IC or a miniature device, whether it involves a line-width, spacing, or contact dimension, is called the *critical dimension* (CD). The overall resolution of a process describes the consistent ability to print a minimum size image, a critical dimension, under conditions of

Photolithography, Table 1 Comparison of traditional negative and positive photoresists

Characteristic	Resist type	
	Positive resist	Negative resist
Adhesion to Si	Fair (priming required)	Excellent (priming not required)
Available compositions	Many	Vast
Baking	In air (+)	In Nitrogen (–)
Contrast γ	Higher, e.g., 2.2	Lower, e.g., 1.5
Cost	More expensive	Less expensive
Developer	Temperature sensitive (–) and aqueous based (Ecologically sound)	Temperature insensitive (+) and organic solvent (–)
Developer process window	Small	Very wide, insensitive to overdeveloping
Influence of oxygen	No (+)	Yes (–)
Lift-off	Yes	Yes
Mask type	Dark-field: lower-defects	Clear-field: higher-defects
Opaque dirt on clear portion of mask	Not very sensitive to it	Causes printing of pinholes
Photospeed	Slower	Faster
Pinhole count	Higher	Lower
Pinholes in mask	Prints mask pinholes	Not so sensitive to mask pinholes
Plasma etch resistance	Not very good	Very good
Proximity effect	Prints isolated holes or trenches better	Prints isolated lines better
Residue after development	Mostly at $<1 \mu\text{m}$ and high aspect ratio	Often a problem
Resolution	High	Low ($> 1 \mu\text{m}$)
Sensitizer quantum yield Φ	0.2–0.3	0.5–1
Step coverage	Better	Lower
Strippers of resist over		
Oxide steps	Acid	Acid
Metal steps	Simple solvents	Chlorinated solvent compounds
Swelling in developer	No	Yes
Thermal stability	Good	Fair
Wet chemical resistance	Fair	Excellent

reasonable manufacturing variation. Many aspects of the process, including hardware, materials, and processing considerations can limit the resolution of lithography. Hardware limitations include diffraction of light or scattering of charged particles (in the case of charged-particle lithography or hard x-rays), lens aberrations, mechanical stability of the system, etc. The resist material properties that impact resolution are contrast, swelling behavior, thermal flow, and chemical etch resistance, etc. The most important process-related resist variables include swelling (during development) and stability (during etching and baking steps). Resolution frequently is measured by line-width measurements using either transmitted or reflected light or other metrology techniques. Optical techniques perform satisfactorily for features of $1 \mu\text{m}$ and larger, providing a precision of

$\pm 0.1 \mu\text{m}$. The successful performance of devices depends upon the control of the size of critical structures across the entire wafer and from one wafer to another, referred to as line-width control. A rule of thumb is that the dimensions must be controlled to tolerances of at least $\pm 1/5$ of the minimum feature size.

Correct feature size must be maintained within a wafer and from wafer to wafer, as device performance depends on the absolute size of the patterned structures. In the shadow printing mode, including contact and proximity arrangements of mask and wafer, optical lithography has a resolution with limits set by a variety of factors. These include diffraction of light at the edge of an opaque feature in the mask as the light passes through an adjacent clear area, alignment of wafer to mask, nonuniformities in wafer

flatness, and debris between mask and wafer. Diffraction causes the image of a perfectly delineated edge to become blurred or diffused. The theoretical resolution, R , that is, the minimum resolved dimension in a grating mask (b_{\min} for a line or a space) is given by:

$$R = b_{\min} = k\sqrt{\lambda\left(s + \frac{z}{2}\right)} \quad (1)$$

where

b_{\min} = half the grating period and the minimum feature size transferable

s = the gap between the mask and the photoresist surface

λ = wavelength of the exposing radiation

z = the photoresist thickness

k = a constant which theoretically is ~ 1.5

Depending on the relative magnitude of s and z , one distinguishes between contact and proximity printing. The square root relation in contact and proximity printing is a consequence of the near field or Fresnel diffraction theory valid in the near field region just below the mask openings, this contrasts with the far field behavior (Fraunhofer diffraction) of projection lithography (see below). Self-aligned printing can be regarded as an extreme form of contact printing where there is no gap between the mask and the substrate. In contact printing, a photomask is pressed against the resist-covered wafer with pressures in the range of 0.05–0.3 atm., and the s , in Eq. 8.13 is zero. The theoretical maximum resolution is seldom achieved, however, as only diffraction effects were taken into account to derive Eq. 1. The other factors mentioned above (wafer flatness, mask alignment, etc.) and shadowing effect (penumbral blur) usually conspire to make the resolution worse.

In proximity printing, spacing of the mask removed from the substrate (by at least 10 μm) minimizes defects that result from contact. On the other hand, diffraction of the transmitted light reduces the resolution. The degree of reduction in resolution and image distortion depends on the wafer-to-substrate distance, which may vary across the wafer.

In projection printing, wafer contact is completely avoided; a high-resolution lens projects an image of the photomask onto the photoresist-covered wafer. Imaging in this case is far-field or Fraunhofer diffraction dominated and the Rayleigh criterion for far-field

resolution can be considered. A dimensionless constant k_1 can be used to define the practical limiting resolution R in projection printing as:

$$R = k_1 \frac{\lambda}{\text{NA}} \quad (2)$$

where k_1 = experimentally determined dimensionless parameter that depends on resist parameters, process conditions, mask aligner optics, etc.

To achieve higher resolutions, instrumentation designers want to work at shorter wavelengths λ , higher NA, and smaller k_1 .

Since the mid-1980s, the demise of optical lithography has been predicted as being only a few years away, but each time a new resolution limit approaches, some new method extends the useful life of the technology. The three main technologies involved in printing ICs and other miniaturized devices are resist technology, mask technology, and the exposure tool, all of which need to be addressed to optimize lithography resolution. Strategies for improved resolution can be achieved through using chemically amplified resists, antireflection coatings, improved mask technology (such as Optical Proximity Correction), improved Exposure Equipment (such as off-axis illumination), etc. DUV photolithography is expected to continue until the 45 nm node through these resolution enhancing techniques. The 32 nm node is viewed as beyond the scope of DUV lithography.

Next-Generation Lithographies (NGLs) and Lithography Research

In the IC industry, continuous improvements to DUV photolithography have postponed the industrial adoption of alternative next-generation lithographies (NGLs), such as extreme ultraviolet lithography (EUVL), x-ray lithography, charged particle beam lithography based on electrons and ions (such as electron and ion projection techniques), and imprint lithography (Nanoimprint lithography [NIL] and Step-and-Flash Imprint Lithography [SFIL]), because of the huge financial investment in existing photolithography equipment. These NGLs are regarded today as important technologies for the beyond-the-DUV-lithography-era, either for mask making or for actual IC production. In addition, many lithography

approaches in the R&D stage, including lithography based on very thin resist layers and block copolymers, zone plate array lithography (ZPAL), quantum lithography (two-photon lithography) and proximal probe-based techniques such as atomic force microscopy (AFM), scanning tunneling microscopy (STM), dip-pen lithography (DPL) and near-field scanning optical microscopy (NSOM), and apertureless near-field scanning optical microscopy (ANSOM), are attracting more and more research interest. It is quite possible that some of the alternative lithography tools in the R&D category will emerge as serious next-generation lithographies (NGL) in the coming years.

Extreme Ultraviolet Lithography (EUVL)

Extreme ultraviolet lithography (EUVL), using wavelengths in the 10–14 nm range to carry out projection imaging, is perhaps the most natural extension of optical projection lithography as, in principle, it only differs in terms of the wavelength. This type of radiation is also referred to as *soft-x-ray radiation* and *vacuum UV*. Sources for this type of radiation are laser-produced plasmas and synchrotrons. EUV is strongly absorbed in virtually all materials and, consequently, imaging must be carried out in vacuum; also, all camera optics as well as masks used must be reflective rather than refractive. New resists and processing techniques must be developed as well.

X-ray Lithography

X-ray lithography employs a shadow printing method similar to optical proximity printing. The x-ray wavelength (4–50 Å) is much shorter than that of UV light (Table 2). Hence, diffraction effects are reduced and higher resolution can be attained. For instance, for an x-ray wavelength of 5 Å and a gap of 40 μm, resolution can be around 0.2 μm. In contrast, with electron lithography and ion-beam lithography, no charged particles are directly involved in x-ray exposures, which makes the need for vacuum less stringent.

X-ray lithography is superior to optical lithography due to its use of shorter wavelengths and its very large DOF, and because exposure time and development conditions are not as stringent. Reproducibility is high since results are independent of substrate type, surface reflections, and wafer topography. Another important benefit is that x-ray lithography is immune to low-atomic-number (Z) particle contamination

Photolithography, Table 2 Light sources for various types of lithography

Wavelength [nm]	Source	Range
436	Hg arc lamp	G-line
405	Hg arc lamp	H-line
365	Hg arc lamp	I-line
248	Hg/Xe arc lamp, KrF excimer laser	Deep UV (DUV)
193	ArF excimer laser	DUV
157	F2 laser	Vacuum UV (VUV)
~10	Laser-produced plasma sources	Extreme UV (EUV)
~1	X-ray tube, synchrotron	X-ray

(dust). With an x-ray wavelength on the order of 10 Å or less, diffraction effects generally are negligible, and proximity masking can be used, increasing the lifetime of the mask. In x-ray lithography, there are essentially no optics involved and, although this sounds like an advantage, it also presents one major disadvantage since one can only work with 1:1 shadow printing. No image reduction is possible, so the mask fabrication process is very complicated. In the United States, IBM remains the only major champion of x-ray lithography for next-generation lithography. At the end of 1999, IBM fabricated several PowerPC 604e microprocessor batches to demonstrate the viability of the method. In Japan, there are still many players involved such as NTT, Toshiba, Mitsubishi, and NEC.

LIGA

The LIGA technique (a German acronym for Lithographie, Galvanoformung, Abformung) was invented about 20 years ago. LIGA exploits all of the advantages of x-ray lithography listed above. The LIGA process involves a thick layer of resist (from micrometers to centimeters), high-energy x-ray radiation, and resist development to make a resist mold. By applying galvanizing techniques, the mold is filled with a metal. The resist structure is removed, and metal products result. Alternatively, the metal part can serve as a mold itself for precision plastic injection molding. Several types of plastic molding processes have been tested, including reaction injection molding, thermoplastic injection molding, and hot embossing. LIGA enables new building materials and a wider dynamic range of dimensions and possible shapes.

Cross-References

- ▶ [BioPatterning](#)
- ▶ [Biosensors](#)
- ▶ [CMOS MEMS Biosensors](#)
- ▶ [CMOS-CNT Integration](#)
- ▶ [Dip-Pen Nanolithography](#)
- ▶ [Dry Etching](#)
- ▶ [DUV Photolithography and Materials](#)
- ▶ [Electron Beam Lithography \(EBL\)](#)
- ▶ [Focused-Ion-Beam Chemical-Vapor-Deposition \(FIB-CVD\)](#)
- ▶ [MEMS Packaging](#)
- ▶ [Microcontact Printing](#)
- ▶ [Nanoimprint Lithography](#)
- ▶ [NEMS Piezoelectric Switches](#)
- ▶ [NEMS Resonant Chemical Sensors](#)
- ▶ [SU-8 Photoresist](#)
- ▶ [Wet Etching](#)

References

1. Madou, M.: Fundamentals of Microfabrication, 2nd edn. CRC Press, Boca Raton (2002)

Photomilling

- ▶ [Chemical Milling and Photochemical Milling](#)

Photonic Crystal Nanobeam Cavities

Parag B. Deotare and Marko Loncar
Electrical Engineering, Harvard School of Engineering
and Applied Sciences, Cambridge, MA, USA

Definition

Photonic crystal nanobeam cavities are high quality factor, small mode volume optical resonators formed by using photonic crystal Bragg mirrors in the longitudinal direction and total internal reflection in the others.

Introduction

On-chip manipulation of light is of great importance for applications ranging from optoelectronics and nonlinear optics to biochemical sensing and cavity quantum electrodynamics. Ultra-high quality factor (Q) optical resonators are indispensable for these applications since they enhance the probability of light-matter interaction due to their ability to store photons for many optical cycles [1]. Traveling-wave-based resonators, and micro-toroid resonators in particular, have been used to achieve some of the largest Q factors to date [1]. For many applications, however, the figure of merit is proportional to the ratio between resonator's Q and mode volume (V_{mode}), which can be quite large in traveling-wave geometries. Therefore, standing-wave optical resonators, capable of having very small mode volume, have been explored as an attractive alternative [1]. These structures, based on a Fabry-Perot geometry, typically consist of the region in which light can propagate (e.g., an optical waveguide), terminated at its both ends with optical mirrors. The mirrors are often in the form of a 1-D Bragg mirror, as in the case of vertical cavity surface emitting laser (VCSEL), for example. In 1987, the Bragg mirror idea was extended to 3-D by Yablonovitch [2] and John [3] who proposed 3-D photonic crystals as a platform suitable for realization of ultimate cavities with wavelength-scale mode volumes and ultra-high Q s. Unfortunately, fabrication of these 3-D periodic optical structures has proven to be challenging and in spite of great effort no ultra-high Q small mode volume cavity based on 3-D photonic crystals has been realized to date. This prompted researchers to consider alternative approaches, based on lower-dimensionality photonic crystals. In particular, planar photonics crystals, based on 2-D periodic lattices realized in thin semiconductor membranes [4] received considerable interest due to the ease of fabrication and compatibility with standard planar micro-fabrication technology. State-of-the-art fabrication techniques, combined with creative cavity designs, have resulted in wavelength-scale optical cavities with $Q > 10^6$ (experimental) and Q/V_{mode} values larger than traveling wave resonators. Recently, however, it has been demonstrated that photonic crystals with even lower dimensionality, resembling some of the early distributed feedback (DFB) cavity designs, can be used to realize cavities with performance rivaling and even outperforming those

based on 2-D planar photonic crystals. These photonic crystal nanobeam cavities (PCNCs), based on chirped 1-D optical lattices, also have a physical footprint which is exactly the same as that of an optical waveguide and therefore, represent the smallest cavity geometry that can be made using dielectric materials only. Furthermore, nanobeams enable realization of high-Q small- V_{mode} cavities even in low-index materials (e.g., SiO_2 , polymers) [5, 6], thus vastly increasing the library of photonic materials suitable for realization of ultra-high Q/V_{mode} cavities. The near-field of PCNCs is highly accessible, which allows for easy light-material interaction and which is useful for sensing [6], optical trapping [7], and optomechanics [8, 9]. PCNCs can be easily and naturally integrated with optical waveguides for efficient delivery of light [10, 11]. Cavities that support ultra-high Q modes with both TE and TM polarization [12] can be easily realized. All of these features, many of which are hard to achieve using 2-D photonic crystals, make photonic crystal nanobeam cavities an ideal platform for practical applications.

Design Evolution

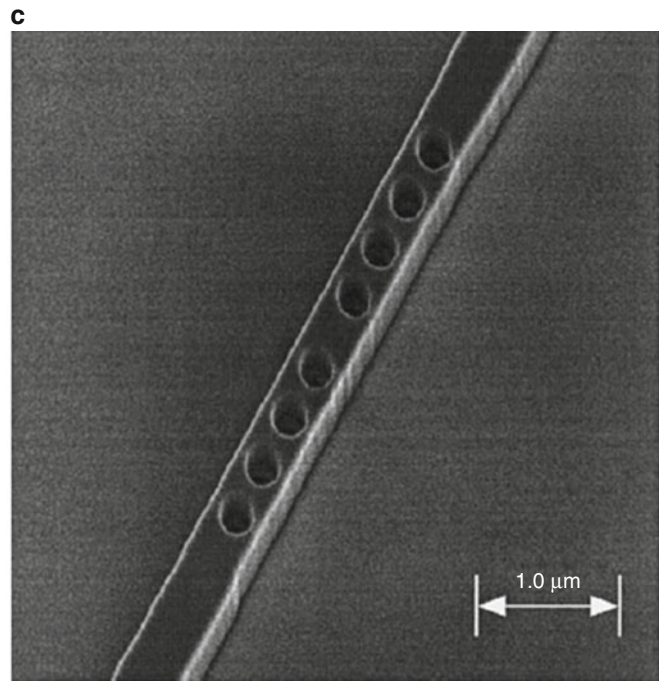
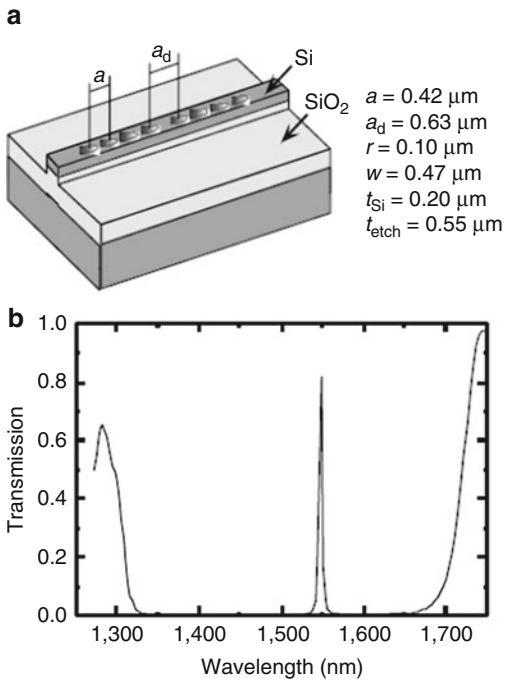
The idea of waveguide-based micro-cavities has been around for more than 15 years. It was first proposed at MIT [13] in 1995 and then experimentally realized 2 years later [14]. Figure 1a shows the schematic of the device which consisted of a waveguide segment surrounded with two lattices of holes that play the role of Bragg mirrors. Light trapping is achieved in this system via a combination of Bragg scattering (in longitudinal direction) and total internal reflection (in two transverse directions), resulting in an optical resonance as shown in Fig. 1b. The theoretical value of Q was calculated to be $Q = 280$ whereas the experimentally measured value was 265 (Fig. 1c). Such a low value of Q factor was due to the suboptimal design. In particular, light scattering at the interface between the Bragg mirror and the central cavity region was responsible for the low Q factor (Fig. 2a) due to the mismatch in the effective mode indices of the two regions.

In order to increase the Q of the cavity it is necessary to *taper* propagating mode of the cavity region into the exponentially decaying mode of the Bragg mirror [15, 16]. This mode matching, analogous to impedance matching problem in electronics, can be achieved using a multi-hole taper toward the cavity

[15]. By doing so, the waveguide mode index can slowly be reduced to match the Bragg index which significantly reduces the scattering and provides exponential increase in cavity Q, at an expense of slight, linear increase in the cavity mode volume. By increasing the number of holes used in the taper, near adiabatic conversion between the two modes could be achieved. By putting two of such engineered tapered (chirped) mirrors back to back, ultra high-Q cavity could be realized. The electric field of a cavity mode formed by using a 7-hole taper is shown in Fig. 2a. The effect of tapered holes on the Q and mode volume was first studied in silicon nitride material system for operation in visible [16] and the results are shown in Fig. 2b. The figure shows Q as a function of cavity length, defined as distance between two central holes in the structure, for two different numbers of tapered holes used in the mirror. It can be seen that the drastic increase in Q can be achieved for an optimal value of the cavity length. Also, it is apparent that the Q values depend strongly on the precise cavity length which puts stringent requirement on the fabrication quality needed for the realization of ultra-high Q cavities.

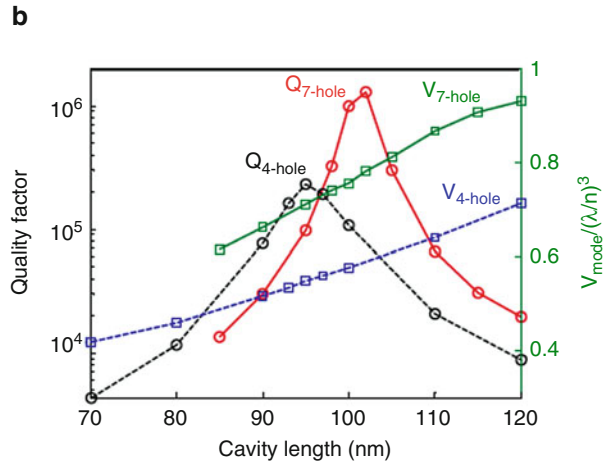
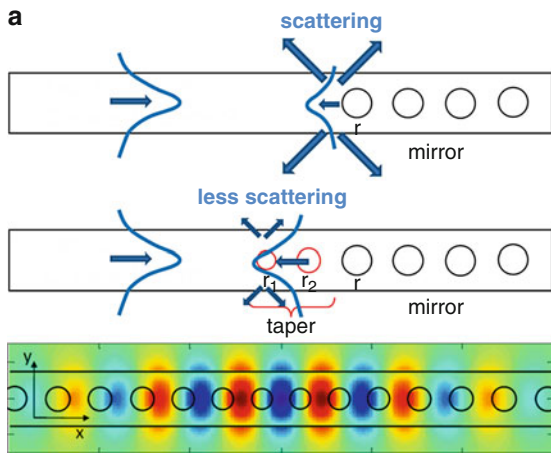
Ultra-High Q Nanobeam Cavities

In order to validate the designs, devices were fabricated in silicon-on-insulator (SOI) material platform (Fig. 3) using state-of-the art nanofabrication technology. The highest Q measured in such cavities that operated in telecom wavelength range was 750,000 [17]. The cavities were probed using a free-space cross-polarization technique popularly known as resonance scattering. This technique is relatively straightforward to implement but suffers from poor in- and out-coupling of light, in particular in the case of high-Q cavity modes. Other characterization techniques that have been used to characterize nanobeam cavities include evanescent coupling using a tapered fiber [8] and butt coupling [10] (Fig. 4a). The latter technique, particularly suitable for realization of integrated on-chip networks, takes advantage of spot-size converters to couple light from tapered optical fiber into the on-chip optical waveguide. In this case, the converter is in the form of polymer (SU-8 photoresist) waveguides (ref. Fig. 4b) whose cross-section and resulting mode profile is comparable to the spot emerging from the tapered optical fiber. Light is then adiabatically



Photonic Crystal Nanobeam Cavities, Fig. 1 (a) Schematic of photonic crystal nanobeam cavity (PCNC) (b) simulated transmission of the device showing the bandgap and the cavity

resonance [13] (c) SEM image of the fabricated device. The measure Q of the device was 265 [14]



Photonic Crystal Nanobeam Cavities, Fig. 2 (a) Schematic showing the effect of taper on scattering loss due to mismatch of the waveguide mode and the Bloch mode. Cavity mode electric field (E_y) for a 7-hole taper (b) simulated results from [16]

showing the variation in Q and mode volume for different taper and cavity lengths. Cavity length was defined as the distance between two innermost holes (indicated in Part a)

coupled from polymer waveguide into the silicon waveguide using the inverse taper geometry, and then from the silicon waveguide into the nanobeam cavity using adiabatically tapered photonic crystal [10, 11].

Efficient coupling of light into the cavity is achieved by using hole taper similar to the one described earlier [11]. Light transmitted through the system is collected by a second spot-size converter and

tapered fiber. **Figure 4b** shows an optical micrograph of such a fabricated structure with the input coupling pads, silicon waveguides, cavities, and output coupling pads. Inset shows the zoomed-in image of the polymer coupling pad. Coupling efficiencies as high as 80% over more than 150 nm could be achieved using this technique. The transmission spectrum shown in **Fig. 4c** features the excellent quality of signal to noise achieved using the above technique. The inset shows the zoomed region around the cavity resonance.

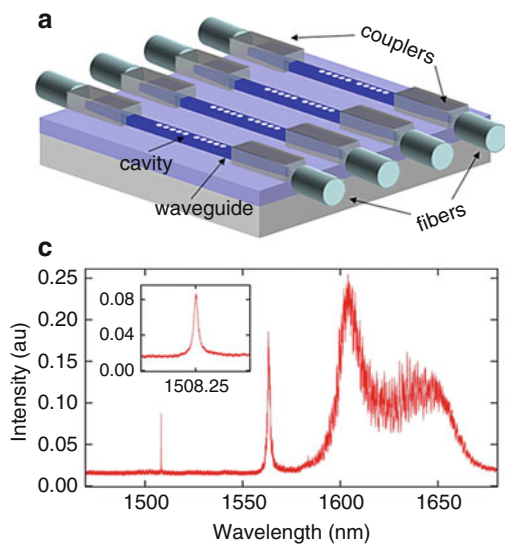
There have been quite a few variations in the design of the periodic structure of the ultra-high Q nanobeam cavity, including ladder cavities [18] with rectangular



Photonic Crystal Nanobeam Cavities, Fig. 3 SEM image of the fabricated structure with a measured Q of 7.5×10^5 [17]. The inset shows the zoomed cavity region

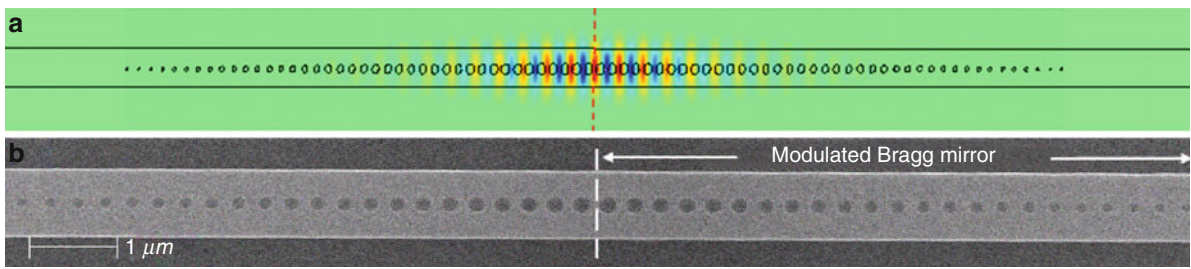
holes, bookshelf/domino structures [18], and parabolic width-taper cavity [19]. However, the working principle of all of these approaches is the same and is based on efficient effective index matching between cavity and mirror regions. Another common feature of all the different design strategies is that they are all intuition-based and involve extensive parameter search and trial-and-error-based optimization in order to find the optimal taper profile (the radius and spacing of the holes in the taper section). To overcome these limitations and streamline the design process, thus significantly reducing the time needed for the design of ultra-high Q structures, a deterministic method to design ultra-high Q cavities has recently been proposed and experimentally demonstrated [10]. This approach not only eliminates trial-and-error-based shifting and resizing of holes, but also results in cavities that feature high transmission at the cavity resonance: $Q > 10^7$ with transmission of 97% (**Fig. 5**).

As mentioned in the introduction, nanobeam cavities enable realization of ultra-high Q cavities, with small mode volumes, even in materials with low refractive index. This is typically very hard to achieve using 2-D planar photonic crystals since the bandgap is typically closed due to low-index contrast. Nanobeam cavities do not suffer from this problem since 1-D bandgap is always open even for the smallest index



Photonic Crystal Nanobeam Cavities, Fig. 4 (a) Schematic showing butt coupling technique using spot-size converter couplers based on polymer waveguides and inverse-tapered silicon waveguides. (b) Optical micrograph showing the fabricated

device with the in- and out-coupling pads, silicon waveguide, and PCNC. Inset shows an SEM image of the polymer coupling pads. (c) Transmission from one of the devices shown in (b)



Photonic Crystal Nanobeam Cavities, Fig. 5 Ultra-high Q cavity designed using deterministic approach [10] (a) Electric field profile (b) SEM image of structure fabricated in SOI

modulation. Recently, high-Q all-polymer cavities with index contrast between core (cavity material) and cladding as low as 1.15 have been experimentally demonstrated [6].

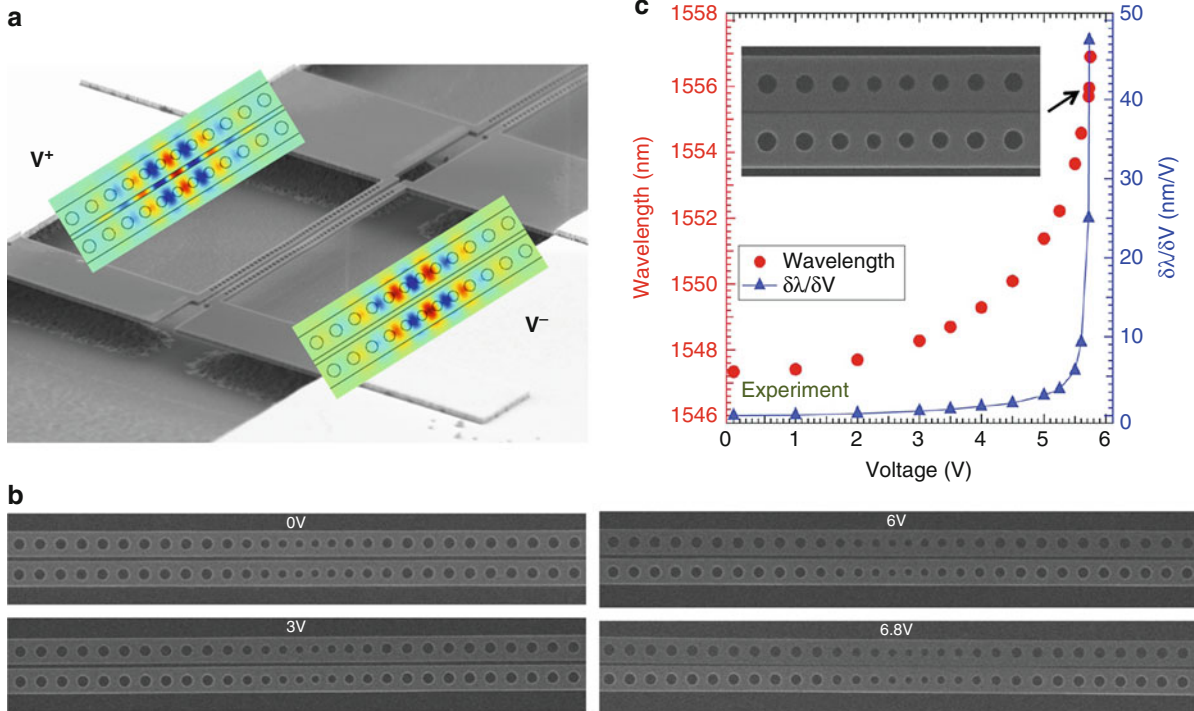
Finally, it has been demonstrated that photonic crystal nanobeam cavities also play a role of *phononic crystals* and are capable of localizing phonons [9]. Furthermore, simultaneous confinement of photons and phonons allows for strong interaction between mechanical and optical excitations, thus enabling both exploration of exciting physics at nanoscale and realization of novel devices. For example, electromagnetically induced transparency (EIT) in photonic/phononic crystal nanobeam cavity was demonstrated and successfully used to alter the group velocity of light thus realizing optical delay lines of interest for optical buffering, filtering, and signal processing [20].

Applications

One promising application of nanobeam cavities is realization of compact optical filters and spectrometers. However, due to fabrication limitations, some degree of post-fabrication wavelength trimming is needed in order to bring the resonance of the filter (cavity) to the desired wavelength. Moreover, optical devices that could be dynamically and reversibly configured to operate at vastly different wavelengths (optical analog to FPGAs) would be of great interest for practical applications. Conventional techniques, based on thermal effects and free carrier injection, suffer from high steady-state power consumption and small tuning range, respectively. These limitations can be overcome by taking advantage of mechanical degrees of freedom of suspended photonic crystal nanobeam resonators.

Using coupled photonic crystal nanobeam cavities [21], a reconfigurable photonic crystal filter [22] that can be dynamically, continuously, and reversibly tuned over a wide wavelength range was proposed and demonstrated. These devices, which combine the fields of NEMS and nanophotonics, consist of two doubly clamped photonic crystal nanobeam cavities separated by a gap smaller than 100 nm. Coupling between cavity modes results in two super modes with even and odd symmetry [8, 21] – the situation analogous to coupled quantum wells.

The symmetric mode of such coupled system is extremely dispersive to the separation between the nanobeams which allows for the tuning of this mode by controlling the separation. On the other hand, the asymmetric mode of the system is virtually insensitive to the changes in the separation, which can be attributed to higher-order effects such as a coupling-induced frequency shift [21]. The separation between nanobeams can be controlled using capacitive force, for example: An external bias voltage can be applied to two gold electrodes (Fig. 6a) that are in contact with Si with which the cavities were made of. In this way, the two cavities form the plates of a parallel plate capacitor. The capacitive force is strong enough to bend the beams toward each other as shown in Fig. 6b, which in turn has a strong effect on the resonant wavelength of the structure. Using this approach, tunable filters with a tuning range of ~ 10 nm for less than 6 V bias and negligible steady-state power consumption were demonstrated. The latter is a very important feature, because once reconfigured, *they do not dissipate any power*. In addition, due to their low mass and flexibility, the energy required to deflect nanobeams is extremely low, and is estimated to be well below the femtojoule level. Therefore, nanoscale optomechanical devices can be used to realize programmable optical systems that



Photonic Crystal Nanobeam Cavities, Fig. 6 (a) SEM image of the fabricated structure with gold contact pads. The insets show the field profiles for even and odd super modes. (b) SEM images showing the reduction of gaps for different applied

potentials between the two nanobeams; (c) red graph shows the tuning of the even mode for various applied voltages while the blue graph plots the sensitivity at the operating points [22]

can be pre-configured (and reconfigured) using electrical signals, and will likely find applications in low-power optical circuit switching and programmable photonics. It should also be noted that when properly biased, only a few mV of additional electrical signal is enough to move the nanobeams and induce the shift in cavity resonance larger than one linewidth, thus completely detuning the cavity from the incoming optical signal (originally in resonance with the structure) [22].

Significant improvements in the tuning range can be obtained by using more sophisticated actuation methods. For example, Fig. 7 shows a tuning technique that uses a mechanical comb drive to move one of the cavities thus changing the gap between the cavities and resonance of the system [23]. This technique can achieve extremely small gaps as it does not suffer from the pull-in effects which can be detrimental for capacitive actuation approach discussed above.

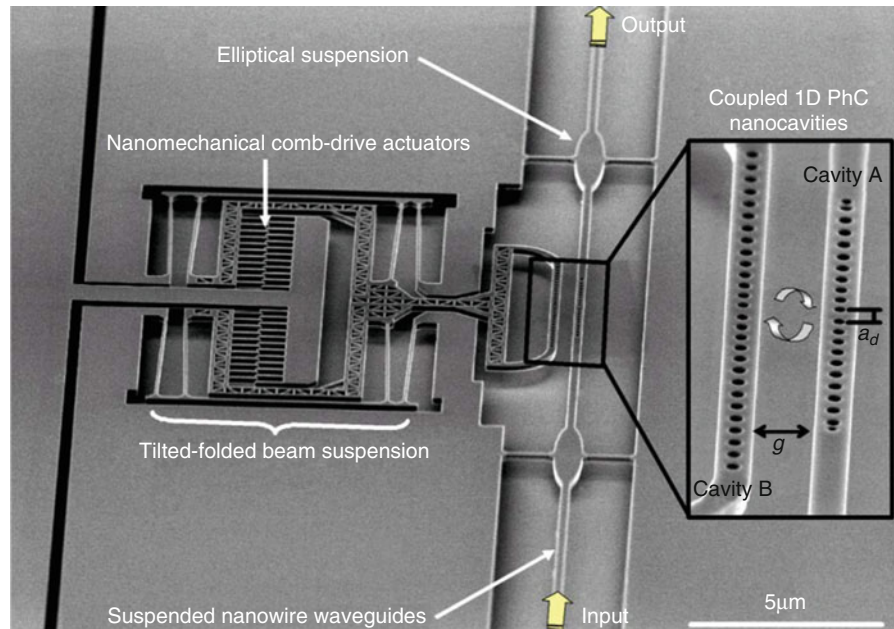
In addition to electrical actuation, nanoscale optical devices can be actuated using light [8, 24]. Figure 8a shows a structure consisting of freestanding coupled

nanobeam cavities interfaced with optical waveguides and fiber couplers (for efficient in- and out-coupling of light). Optical transmission through the system (experimental) is shown in Fig. 8b, whereas the tuning of cavity resonance for 6 mW of pump power at 1,583 nm is shown in Fig. 8c. The cavity tuning was due to contributions from thermo-optical, nonlinear optical, and optomechanical effects [24].

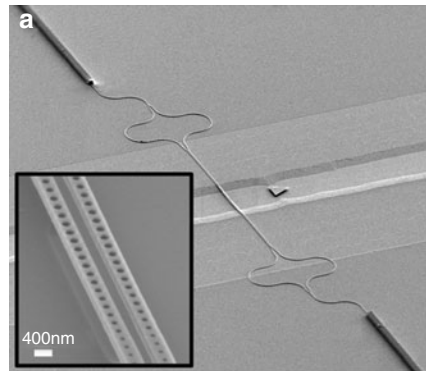
Alternatively, optomechanical systems can be used to read out mechanical motion optically. Figure 9 shows the mechanical spectrum of the structure in air (mechanical modes excited by Brownian motion), detected by monitoring optical transmission through the device. However, it is important to note that this measurement also affects the mechanical properties of the devices, as seen in Fig. 9: Optical signal can stiffen or soften mechanical resonances (induce an additional effective spring constant) via optical gradient force [8].

Owing to their small mode volume and footprint, nanobeam cavities are also excellent candidates for the realization of low-threshold nanolasers with a

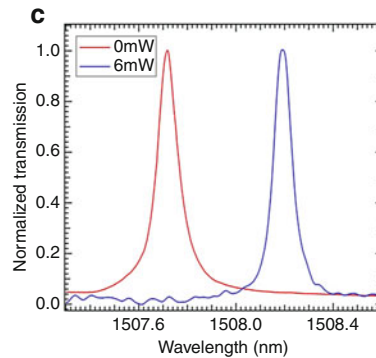
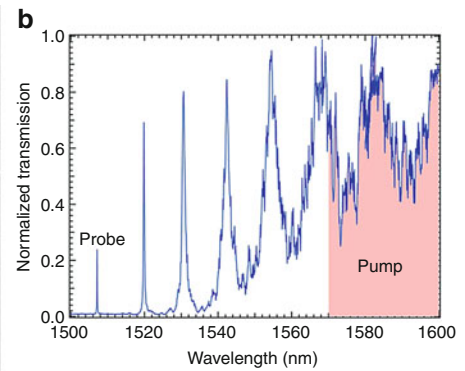
Photonic Crystal Nanobeam Cavities,
Fig. 7 SEM image showing the PCNCs and comb drive used to achieve tuning of the resonant wavelength [23]



Photonic Crystal Nanobeam Cavities,
Fig. 8 Integrated optomechanical system based on coupled nanobeam cavities (a) SEM image showing the complete device with the fiber-coupling pads, interferometer, waveguides, and the suspended nanobeam cavity region (inset). (b)

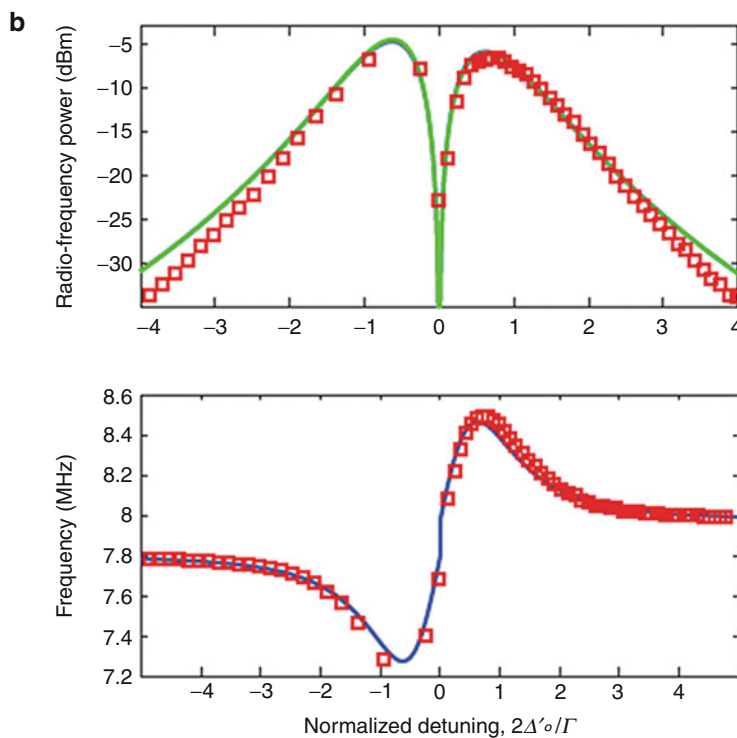
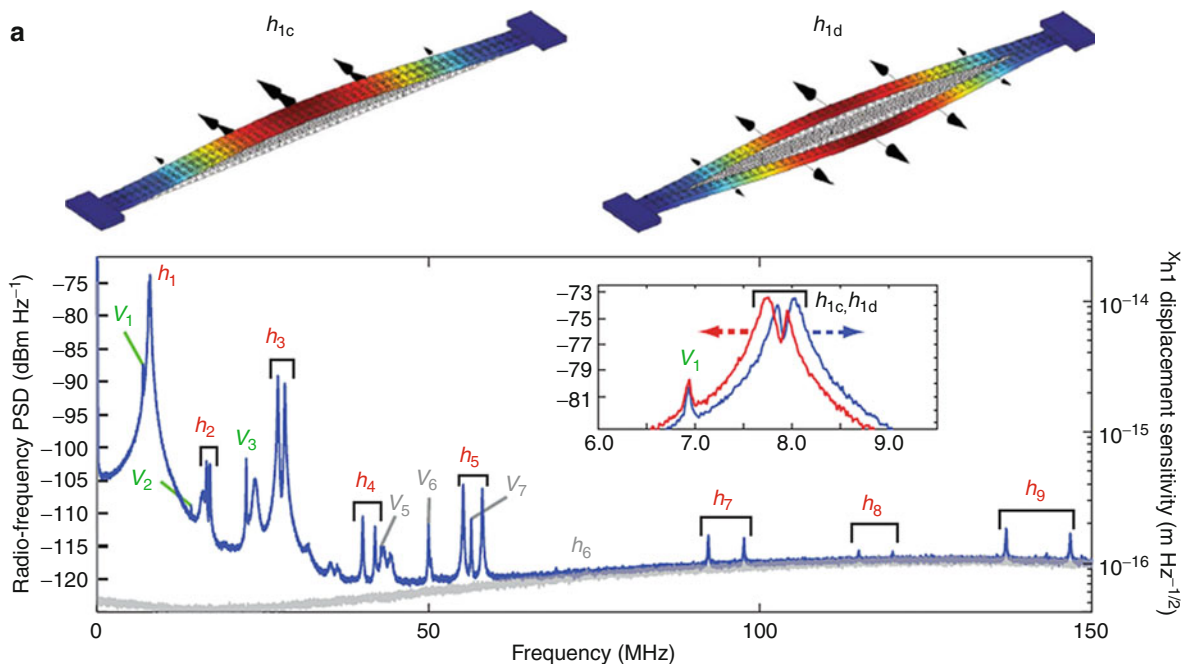


(b) A transmission spectrum of the fabricated cavity. The even cavity mode – *probe* – is at $\sim 1,507$ nm ($Q \sim 15,000$) and the *shaded* region can be used to pump outside the bandgap. (c) Cavity resonance was tuned from 1,507.72 to 1,508.19 nm when 6 mW of *pump* power at 1,583 nm was used. Cavity tuning was due to contribution from thermo-optical, nonlinear, optical, and optomechanical effects



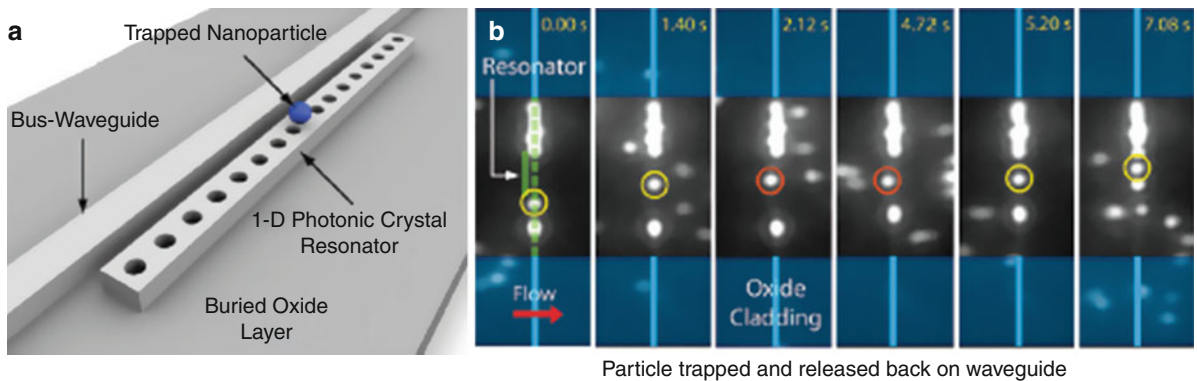
large spontaneous emission (β) factor [19, 25]. Furthermore, it is possible to realize tunable/programmable nanolasers by taking advantage of mechanical degrees of freedom of coupled nanobeam cavities, as

discussed above [26]. PCNCs can also be designed to support TE and TM modes simultaneously, which is of interest for nonlinear frequency conversions [12]. Scaling properties of photonic crystals can be used in



Photonic Crystal Nanobeam Cavities, Fig. 9 Zipper cavities (a) FEM simulation showing the lowest-order common and differential mechanical modes. Detected radio frequency spectrum showing various in-plane and out-of-plane mechanical

modes of zipper structure. (b) Power and resonance frequency of fundamental mechanical mode (common) for various optical probe detunings



Photonic Crystal Nanobeam Cavities, Fig. 10 (a) Schematic of side-coupled PCNC used for trapping polystyrene particles; (b) snapshots showing particle being trapped by the cavity field and propelled by the guided modes of the waveguide [26]

PCNCs to design cavities in microwave regime allowing for integrating of RF with photonics. Finally, PCNCs are excellent candidates for applications in sensing [6, 7] and for manipulation of particles suspended in fluid (optical trapping). Figure 10a shows the schematic of the device while Fig. 10b reports snapshots from a movie showing particle manipulation. Polystyrene particles were propelled along the waveguides by the guided modes and trapped by the intense field in the cavity region when the laser was tuned to the cavity resonance.

Outlook

Nanobeam photonic crystal cavities have developed significantly since they were first proposed more than a decade ago. Greater understanding and improvements in the design methods along with the advancement in nanofabrication processes has significantly helped research outgrow the challenges in making ultra-high Q wavelength-scale optical resonators. The last few years of research has seen nanobeam photonic crystal cavities mature as a photonic device and allowed researchers to not only solve current technological problems but also explore new research areas involving enhanced light-matter interaction. Their use in tunable filters, switches, delay lines, sensing, particle manipulation, etc., has already been demonstrated. Their wavelength-scale dimensions, flexible structure, and small footprint area along with the scalable nature make them ideal to be used for applications in the entire electromagnetic spectrum. Emerging areas of research include RF, Terahertz, and Mid-IR photonics.

Cross-References

- ▶ [Capacitive MEMS Switches](#)
- ▶ [Light Localization for Nano-optical Devices](#)
- ▶ [Nanophotonic Structures for Biosensing](#)
- ▶ [Nanostructures for Photonics](#)
- ▶ [Whispering Gallery Mode Resonator Biosensors](#)

References

1. Vahala, K.J.: Optical microcavities. *Nature* **424**(6950), 839–846 (2003)
2. Yablonovitch, E.: Inhibited spontaneous emission in 3 dimensionally modulated periodic dielectric structures. *Journal De Physique* **48**(C-5), 615–616 (1987)
3. John, S.: Strong localization of photons in certain disordered dielectric superlattices. *Phys. Rev. Lett.* **58**(23), 2486–2489 (1987)
4. Joannopoulos, J.D., et al.: *Photonic Crystals: Molding the Flow of Light*. Princeton University Press, Princeton (2008)
5. Gong, Y., Vuckovic, J.: Photonic crystal cavities in silicon dioxide. *Appl. Phys. Lett.* **96**(3), 031107 (2010)
6. Quan, Q.M., et al.: High-Q, low index-contrast polymeric photonic crystal nanobeam cavities. *Opt. Express* **19**(22), 22191–22197 (2011)
7. Mandal, S., Serey, X., Erickson, D.: Nanomanipulation using silicon photonic crystal resonators. *Nano Lett.* **10**(1), 99–104 (2010)
8. Eichenfield, M., et al.: A picogram- and nanometre-scale photonic-crystal optomechanical cavity. *Nature* **459**(7246), 550–U79 (2009)
9. Eichenfield, M., et al.: Optomechanical crystals. *Nature* **462**(7269), 78–82 (2009)
10. Quan, Q.M., Deotare, P.B., Loncar, M.: Photonic crystal nanobeam cavity strongly coupled to the feeding waveguide. *Appl. Phys. Lett.* **96**(20), 203102 (2010)

11. Zain, A.R.M., et al.: Ultra high quality factor one dimensional photonic crystal/photonic wire micro-cavities in silicon-on-insulator (SOI). *Opt. Express* **16**(16), 12084–12089 (2008)
12. McCutcheon, M.W., et al.: High-Q transverse-electric/transverse-magnetic photonic crystal nanobeam cavities. *Appl. Phys. Lett.* **98**, 111117 (2011)
13. Villeneuve, P.R., et al.: Air-Bridge Microcavities. *Appl. Phys. Lett.* **67**(2), 167–169 (1995)
14. Foresi, J.S., et al.: Photonic-bandgap microcavities in optical waveguides. *Nature* **390**(6656), 143–145 (1997)
15. Lalanne, P., Hugonin, J.P.: Bloch-wave engineering for high-Q, small-V microcavities. *IEEE J. Quantum. Electron.* **39**(11), 1430–1438 (2003)
16. McCutcheon, M.W., Loncar, M.: Design of a silicon nitride photonic crystal nanocavity with a quality factor of one million for coupling to a diamond nanocrystal. *Opt. Express* **16**(23), 19136–19145 (2008)
17. Deotare, P.B., et al.: High quality factor photonic crystal nanobeam cavities. *Appl. Phys. Lett.* **94**(12), 121106 (2009)
18. Notomi, M., Kuramochi, E., Taniyama, H.: Ultrahigh-Q nanocavity with 1D photonic gap. *Opt. Express* **16**(15), 11095–11102 (2008)
19. Ahn, B.H., et al.: One-dimensional parabolic-beam photonic crystal laser. *Opt. Express* **18**(6), 5654–5660 (2010)
20. Safavi-Naeini, A.H., et al.: Electromagnetically induced transparency and slow light with optomechanics. *Nature* **472**(7341), 69–73 (2011)
21. Deotare, P.B., et al.: Coupled photonic crystal nanobeam cavities. *Appl. Phys. Lett.* **95**(3), 31102 (2009)
22. Frank, I.W., et al.: Programmable photonic crystal nanobeam cavities. *Opt. Express* **18**(8), 8705–8712 (2010)
23. Chew, X.Y., et al.: Dynamic tuning of an optical resonator through MEMS-driven coupled photonic crystal nanocavities. *Opt. Lett.* **35**(15), 2517–2519 (2010)
24. Deotare, P.B., et al.: All optical reconfiguration of optomechanical filters, accepted *Nature Communications*. DOI: 10.1038/ncomms1830 (2012)
25. Zhang, Y., et al.: Photonic crystal nanobeam lasers. *Appl. Phys. Lett.* **97**(5), 051104 (2010)
26. Perahia, R., et al.: Electrostatically tunable optomechanical “zipper” cavity laser. *Appl. Phys. Lett.* **97**(19), 191112 (2010)

Photonics in Nature

- ▶ [Biomimetics of Optical Nanostructures](#)

Photovoltaic Devices

- ▶ [Self-repairing Photoelectrochemical Complexes Based on Nanoscale Synthetic and Biological Components](#)

Physical Colors

- ▶ [Nanostructures for Coloration \(Organisms other than Animals\)](#)

Physical Dry Etching

- ▶ [Dry Etching](#)

Physical Modification

- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)

Physical Vapor Deposition

Yoke Khin Yap

Department of Physics, Michigan Technological University, Houghton, MI, USA

Synonyms

[Anodic arc deposition](#); [Arc discharge](#); [Boron nitride nanotubes \(BNNTs\)](#); [Carbon nanotubes \(CNTs\)](#); [Catalyst](#); [Cathodic arc deposition](#); [Chemical beam epitaxial \(CBE\)](#); [Electron beam evaporation](#); [Electron beam physical vapor deposition \(EBPVD\)](#); [Gallium arsenide \(GaAs\)](#); [Gas-phase molecular beam epitaxy \(gas-phase MBE\)](#); [Graphene](#); [Ion beam](#); [Magnetron sputtering](#); [Metal-organic molecular beam epitaxy \(MOMBE\)](#); [Molecular beam epitaxy \(MBE\)](#); [Nanorods](#); [Nanotubes](#); [Nanowires](#); [Plasma](#); [Pulsed-laser deposition \(PLD\)](#); [Si nanotubes](#); [Sputtering](#); [Thermal evaporation](#)

Definition

Physical vapor deposition (PVD) is referred to deposition processes of thin films and nanostructures through the evaporation of solid precursors into their vapor phase by physical approaches followed by the

condensation of the vapor phase on substrates. The whole process consists of three stages: (1) evaporation of the solid source, (2) vapor phase transport from the source to the substrates, and (3) vapor condensation on the substrates. Since PVD technique can convert solid materials into vapor phase without chemical processes, they are very convenient and possible to deposit many types of materials into thin films and nanostructures.

Classification

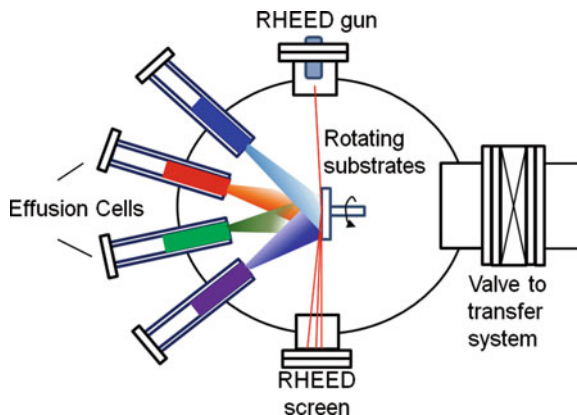
PVD techniques can be classified based on the techniques used to evaporate the solid source materials into their vapor phase [1, 2]. These techniques include (1) thermal evaporation, (2) ion sputtering, and (3) arc discharge. In the following subsections, examples of PVD based on these evaporation techniques will be described.

PVD by Thermal Evaporation

Molecular Beam Epitaxy

Molecular Beam Epitaxy (MBE) is a PVD technique conducted under ultrahigh vacuum (UHV, better than 10^{-9} mbar) where epitaxial deposition of thin film crystal on crystalline substrates can be obtained by using one or more vapor sources (*Effusion cells* or sometimes called *Knudsen cells*). An effusion cell consists of a crucible (made of graphite, pyrolytic boron nitride, quartz, or tungsten) where the solid source can be heated and evaporated by filaments (Fig. 1). For example, powders of gallium (Ga) and arsenic (As) can be evaporated in separated cells until they sublimate into vapors. When these separated beam of element vapors are condensed on a heated Si substrate, crystallized thin films of gallium arsenide (GaAs) will be formed. A typical MBE system has a reflection high-energy electron diffraction (RHEED) system to provide in situ monitoring of the crystal growth processes.

The major advantage of MBE comes from its controllable deposition rate at the atomic or molecular scale. This can be achieved due to the use of effusion cells that enable the generation of “molecular beam,” where the generated vapors have low interparticle collisions before they reach to the substrate surface. This means the evaporated atoms escape from the orifice of the cells through effusion and have long mean free path. They neither interact with each other nor with the gases in the vacuum chamber. Due to this capability, MBE was recently used for the growth of graphene



Physical Vapor Deposition, Fig. 1 Schematic drawing of a typical molecular beam epitaxy system

but the outcome is debatable [3, 4]. Finally, gases were also used as the precursors instead of solid sources and such configuration is called *gas-phase MBE* or more appropriately *chemical beam epitaxial (CBE)* [5]. In this case, the pyrolysis of either trimethylindium (TMIn) or triethylindium (TEIn), and trimethylgallium (TMGa) or triethylgallium (TEGa) at the heated substrates were used as the In and Ga source, while As_2 and P_2 were obtained by thermal decomposition of the trimethylarsine (TMAs) and triethylphosphine (TEP). Such approach is more like a combined MBE and *metal-organic chemical vapor deposition (MOCVD)* and is sometimes called *metal-organic molecular beam epitaxy (MOMBE)*.

In fact, *MBE* is an advanced form of thermal evaporation technique with highly controllable deposition thickness (and thus with the drawback of slow deposition rate) due to the use of effusion cells and high vacuum. For the deposition of amorphous thin films, much simple setup can be employed with less demanding vacuum level ($\sim 10^{-2}$ mbar in a glass bell jar chamber). This approach is simply called *direct resistive evaporation* and can be achieved by passing a large current through a resistive wire or foil containing the solid material to be deposited. The heating element is often referred to as an “evaporation source.” The evaporation source can appear in many forms including tungsten wires in the form of filaments or baskets, and thin tungsten, tantalum, or molybdenum foils in the form of boats.

PVD by Electron Beam (E-Beam) Evaporation

In contrast to *MBE*, *electron beam physical vapor deposition (EBPVD)* provides very high deposition

rate on large deposition areas. *EBPVD* is referring to thin film deposition initiated by electron beam evaporation. The deposition rate is typically ranging from tens of nanometer/minute to hundreds of micrometer/minute, although a rate as high as 50 $\mu\text{m/s}$ was reported [6]. *EBPVD* is based on electron beam evaporation of anode materials (usually in ingot form) in high vacuum (10^{-4} mbar). A typical *EBPVD* system can have multiple electron guns and sources. The electron beams can be generated by either thermionic emission or field emission of filaments. These beams are then accelerated to a high kinetic energy and focused on the ingot. The energy of these electrons is converted into thermal energy as the beam bombards the surface of the ingot. The high deposition rate of *EBPVD* without corrosive products meets the needs of industrial applications. A potential drawback of *EBPVD* is that filament degradation in the electron gun results in a nonuniform evaporation rate. In addition, *EBPVD* may cost contamination of the deposition chamber as the vapors generated from the source materials are coated everywhere in the vacuum chamber, the so-called line-of-sight process.

Pulsed-Laser Deposition

Pulsed-laser deposition (PLD) is a *PVD* technique where a high-power pulsed-laser is used to generate vapors (i.e., laser ablation) from a solid target followed by the condensation of the vapors on substrates. The experimental setup for *PLD* is relatively simpler than many *PVD* techniques but the laser ablation mechanism is quite complex [7]. As shown in Fig. 2a, a *PLD* system consists of a pulsed-laser, a vacuum chamber with a rotating target holder and a substrate holder with a heater. *PLD* can be performed in vacuum at various substrate temperatures. During film deposition, laser pulses are focused on the surface of a solid target through a viewport of the chamber by using a thin lens. This will generate a laser spot with very high irradiance (I , power per unit area) and local electric field, where $I = \frac{1}{2}\epsilon_0 c E_o^2$ [8]. Here, E_o is the amplitude of electric field of the laser, c is the speed of light in vacuum, and ϵ_0 is the vacuum permittivity. The local electric field is sufficient to cause atomic bond breaking and dielectric breakdown of the target materials to initiate vaporization of the target materials. This vapor will absorb the remaining portion of the laser pulse, and the electric field of the pulse will accelerate electrons and ions

inside the vapor to create rapid heating and to form laser plasma. Continued absorption of the laser pulse will cause a highly directional expansion of the laser plasma perpendicular to the target surface. This expanding vapor is called laser plume that propagates toward the substrate placed several centimeters opposite to the target.

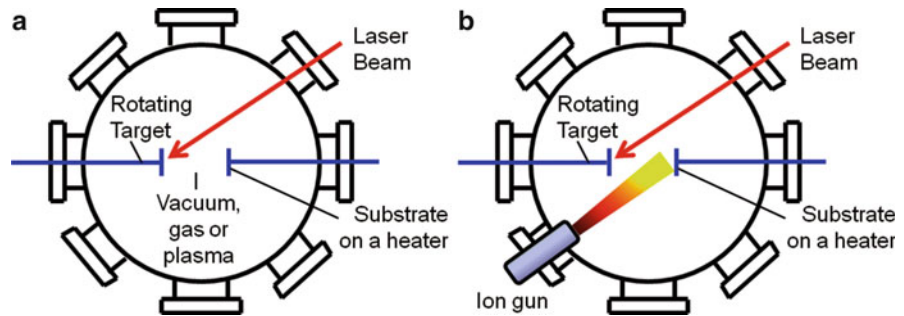
PLD can be conducted in vacuum or in a low-pressured gas ambient. When *PLD* is conducted in reactive gas ambient (say oxygen or nitrogen) the technique is called *reactive-PLD*. In this case, the expanding laser plume will form shock waves and cause chemical reaction between the target vapor and the gas molecules. *PLD* can also be modified into various configurations. For example, plasma can be applied to enhance the interaction of the vapors with the gas ambient and the technique is called *plasma-enhanced or plasma-assisted PLD*. For example, such an approach was shown to control the phases of nitride thin films [9, 10]. Sometimes, an auxiliary ion beam is irradiated on the surface of the substrates during thin film deposition as shown in Fig. 2b, and the technique is called *ion beam-assisted PLD*.

In fact, the mechanism of laser ablation and the quality of samples deposited by *PLD* depend strongly on the wavelength and energy of the pulsed-laser. This can be understood by the absorption of the laser light by the laser plasma when the plasma index of refraction, $n(\omega)$, becomes complex. According to the relation $n^2(\omega) = 1 - [\omega_p/\omega]^2$, a complex $n(\omega)$ will occur when $\omega_p > \omega$ [8]. This means, shorter laser wavelength (larger ω) will reduce the chances of such plasma heating effect. However, one should also consider that the laser plasma frequency, $\omega_p = [N e^2 m_e \epsilon_0]^{1/2}$, where m_e is the mass of electron and N is the electron density. In this case, to avoid plasma heating, N must be minimized, which can be obtained by keeping the laser energy or irradiance low. Plasma heating will lead to the so-called photothermal ablation. For example, photothermal ablation of h-BN target will transfer h-BN target to the substrates and prevent the formation of cubic phase (c-BN) [8]. Photothermal ablation of graphite target at longer laser wavelength will also prevent the formation of diamond-like carbon film [11] as the sp^2 species from the graphite are transferred to the substrates.

PVD by Sputtering

Sputtering is a process where atoms are released from a solid target material due to bombardment of the

Physical Vapor Deposition, Fig. 2 Schematic drawings of a typical pulsed-laser deposition (PLD) system operated (a) in vacuum or with auxiliary plasma, or (b) with auxiliary ion beam



target by energetic particles such as ions and atoms. The atoms released from the target can then be deposited as thin film on substrates and is commonly referred as PVD by sputtering. The physics behind the process is momentum transfer from the incident ions/atoms to the target materials through collisions. The average number of atoms ejected from the target per incident particle is called the sputter yield. The yield is depend on the ion incident angle, the energy of the ion, the masses of the ion and target atoms, the surface binding energy of atoms in the target, and the crystalline orientation of the target when crystals are used as the target. There are many approaches to generate sputter including the use of DC, AC, and RF plasmas or external ion beams.

Sputtering offers several advantages over other PVD techniques including: (1) large area deposition as compared to PLD; (2) convenient for thin film deposition of alloy/composite and materials with high melting points; and (3) avoid device damaging from X-rays generated by electron beam evaporation. However, sputtering also has the following disadvantages: (1) more complicated setup than PLD; (2) contamination is more likely to occur due to the use of plasma as well as relatively lesser vacuum level; and (3) the film morphology may be rougher and even damaged due to bombardment of energetic growth species and clustering of the growth species in the relatively high deposition pressures.

PVD by DC, AC, RF, and Magnetron Sputtering

PVD can be obtained by applying plasma in various configurations. The simplest one is to apply a DC potential across the substrate (anode) and target (cathode). In this case, plasma can be generated when gases are introduced (10s–100s of mTorr) at sufficient electric field strength (\sim a few kV/cm). For example, argon gas is often used as argon ion can lead to high sputter

yield and uniform glow discharge. These Ar ions will bombard the target surface to generate vapors that will be deposited to the substrates. Despite the simplicity, the DC configuration will have the following issues: (1) positive charges will build up on the target surface and cause sparking within the plasma and (2) not suitable for insulating target. To avoid these issues, AC and RF configurations are often used. In these cases, the potential between the target and substrate is alternating so that the target will be negative in potential more often (to generate sputtering) than positive potential. These configurations will also allow lower operation pressure (\sim 10–30 mTorr).

The use of magnetron can further enhance the deposition rate at even lower pressures by the use of DC, AC, or RF electric fields. RF potential is of advantage since it can be used for insulating targets and is applicable in an antenna configuration, that is, without positive bias voltages on the substrates (less electron bombardments and less substrate heating). In this case, the target is placed on the magnetron surrounded by strong electric and magnetic fields. Electrons in the RF plasma will follow helical paths around the magnetic field lines and generate more ionizing collisions near the target surface with the surrounding gas. Thus, more ions will be created and will lead to a higher sputtering and deposition rates. Because of higher ionization rates, plasma can be sustained at a lower pressure.

PVD by Ion-Beam Sputtering

Ion-beam sputtering (*IBS*) is a PVD technique in which the target is evaporated by an external ion source (ion beam). In a typical ion source, ions are generated by collisions of neutral gas atoms/molecules with electrons. They are then accelerated by the electric field by a grid electrode toward the target. Before these ions leave the source, they are neutralized by electrons from

a second external filament. The major advantage of PVD by *IBS* is that the energy and flux of ions can be controlled independently. In addition, the *IBS* approach is applicable to insulating and conducting targets since the flux of “ions” (which are neutralized as they are emitted from the ion beam) that strikes the target is composed of neutral atoms/molecules.

Reactive Sputtering

Like the PLD process, reactive sputtering can be achieved by mixing reactive gases (N_2 , O_2 , etc.) with the inert gas (Ar, Kr, etc.). While the inert gas is responsible for the sputtering of the target materials, the reactive gases are used to initiate chemical reactions with the target vapors to form oxide or nitride films. Like *PLD*, reactive sputtering can also be achieved by various configurations. For example, plasma can be applied to enhance the interaction between the target vapors and the reactive gas ambient and the technique is called *plasma-enhanced or plasma-assisted deposition*. Sometimes, an auxiliary ion beam is irradiated on the surface of the substrates during thin film deposition, and the technique is called *ion beam-assisted deposition (IBAD)*.

PVD by Arc Discharge

In addition to PVD by thermal evaporation and sputtering, the desired vapors for thin film deposition can be obtained by arc discharge of the target. This is sometimes called *arc vapor deposition* [1], where a high-current (10^4 – 10^6 A/cm²), low-voltage (tens of voltages) discharge was applied on the target. For example, *cathodic arc deposition* is referring arc vapor deposition using the target as the cathode. In this case, a high-voltage “trigger arc” will be ignited between the cathode and a sharp auxiliary anode. This will form “seed” electrons and ions to initiate a low-voltage, high-current discharge between the target cathode and the adjacent anode.

On the other hand, target materials can be evaporated by using the target as the anode of the arc discharge. This PVD process is referred as *anodic arc deposition*. In fact, this process is similar to *electron beam physical vapor deposition (EBPVD)* discussed earlier, where arc discharge is generated by irradiating an unfocused electron beam on the target anode. The electrons can be made to spiral in a magnetic field to further ionize the evaporated materials from the target prior to deposition. Like most other PVD processes,

reactive arc vapor deposition can be performed by both *anodic arc deposition* and *cathodic arc deposition* using reactive gases like nitrogen and oxygen. This will lead to the formation of nitride and oxide films.

The deposition rates of *arc vapor deposition* are usually higher than that of PVD by sputtering. This is one of the advantages of *anodic arc deposition* and *cathodic arc deposition* in addition to the low operation voltage. On the other hand, contamination is the potential drawback of most plasma-based PVD processes. In addition, particulates or the so-called macros are undesired for thin films deposited by *anodic arc deposition* and *cathodic arc deposition*. Macros are formed by ablation of molten or solid particles by thermal shock during arc discharges.

Examples of PVD Approaches for Nanotechnology

Physical evaporation techniques like arc discharge [12] and laser ablation [13] are important on the discovery of carbon nanotubes (CNTs). However, PVD techniques discussed here are not popular for the growth of nanomaterials in comparison to chemical vapor deposition (CVD). The major reason is the tendency of forming thin films and amorphous by-products on the surfaces of catalysts which cause “poisoning” effect and prohibit the formation of nanotubes or nanowires [14]. However, there have been some pioneering works reported on the growth of nanotubes and nanowires by PVD techniques such as sputtering, PLD, and MBE. For example, vertically aligned carbon nanotubes (CNTs) were first successfully deposited by RF plasma-assisted pulsed-laser deposition [15]. In this approach, two RF plasmas were employed. The first one was applied on the substrates to form a negative bias voltage during deposition. The second one was applied on a ring electrode located between the graphite target and the substrates. Hydrogen gas was used for the plasma generation and Fe powders (20 nm in diameter) were employed as the catalysts. Although the density of these CNTs is low, these nanotubes have high structural order with internal nanotube diameter as small as 0.4 nm. The formation of these smallest possible CNTs was confirmed by transmission electron microscopy (TEM). Later, CNTs were deposited by magnetron sputtering with a dc bias voltage on substrates [16]. Prior to the growth of CNTs,

Ni films (10 nm thick) were deposited by e-beam evaporation to form nanoparticles that later function as the catalyst for the vapor-liquid-solid (VLS) process. However, these CNTs are defective with bamboo-like structures. Later, RF plasma-assisted pulsed-laser deposition was used to achieve low-temperature growth of boron nitride nanotubes (BNNTs) [14, 17]. In this case, pyrolytic boron nitride pellets were used as the targets and a RF plasma was applied on the substrates to create bias voltages, which are important to eliminate the formation of boron nitride thin films. Again, these BNNTs have internal nanotubes of 1 nm in diameter with high structural order.

Various PVD approaches were also employed for the synthesis of nanowires/nanorods. For example, ZnO and $Zn_{1-x}Mg_xO$ nanorods were grown by MBE on oxidized Si substrate coated with Ag catalyst [18]. In both cases, an ozone/oxygen mixture was used as the oxidizing source. The Zn and Mg cation fluxes were provided by a Knudsen effusion cell using high purity Zn and Mg metals as the sources. The growth of vertically aligned ZnO [19] and MgO [20] nanowires can also be obtained by PLD using ZnO or MgO targets, respectively. Gold particles were used as the catalysts for both cases.

Cross-References

- ▶ [Atomic Layer Deposition](#)
- ▶ [Carbon Nanotube-Metal Contact](#)
- ▶ [Carbon Nanotubes for Chip Interconnections](#)
- ▶ [Carbon-Nanotubes](#)
- ▶ [Chemical Vapor Deposition \(CVD\)](#)
- ▶ [Electron-Beam-Induced Deposition](#)
- ▶ [Focused-Ion-Beam Chemical-Vapor-Deposition \(FIB-CVD\)](#)
- ▶ [Graphene](#)
- ▶ [Growth of Silica Nanowires](#)
- ▶ [Nanomaterials for Electrical Energy Storage Devices](#)
- ▶ [Nanomaterials for Excitonic Solar Cells](#)
- ▶ [Nanotechnology](#)
- ▶ [Synthesis of Carbon Nanotubes](#)

References

1. Mattox, D.M.: Handbook of Physical Vapor Deposition (PVD) Processing: Film Formation, Adhesion, Surface Preparation and Contamination Control (Hardcover). Noyes, New Jersey (1998)
2. Mahan, J.E.: Physical Vapor Deposition of Thin Films. Wiley-Interscience, New York (2000)
3. Park, J., et al.: Epitaxial graphene growth by carbon molecular beam epitaxial (CMBE). *Adv. Mater.* **22**, 4140–4145 (2010)
4. Hackley, J., et al.: Graphitic carbon growth on Si (111) using solid source molecular beam epitaxy. *Appl. Phys. Lett.* **95**, 133114 (2009)
5. Tsang, W.T., et al.: Chemical beam epitaxy of InP and GaAs. *Appl. Phys. Lett.* **45**, 1234–1236 (1984)
6. Schiller, S., Jäsch, G.: Deposition by electron beam evaporation with rates of up to $50 \mu\text{m s}^{-1}$. *Thin Solid Films* **54**, 9–21 (1978)
7. Chrisey, D.B., Hubler, G.H. (eds.): Pulsed Laser Deposition of Thin Films. Wiley-Interscience, New York (1994)
8. Wang, J., Yap, Y.K.: Growth of adhesive cubic phase boron nitride films without argon ion bombardment. *Diam. Relat. Mater.* **15**, 444–447 (2006)
9. Yap, Y.K., Kida, S., Aoyama, T., Mori, Y., Sasaki, T.: Influence of negative dc bias voltage on structural transformation of carbon nitride at 600°C . *Appl. Phys. Lett.* **73**, 915 (1998)
10. Yap, Y.K., Aoyama, T., Kida, S., Mori, Y., Sasaki, T.: Synthesis of adhesive c-BN films in pure nitrogen radio-frequency plasma. *Diam. Relat. Mater.* **8**, 382–385 (1999)
11. Yamamoto, K., Koga, Y., Fujiwara, S., Kokai, F., Heimann, B.: Dependence of the sp^3 bond fraction on the laser wavelength in thin carbon films prepared by pulsed laser deposition. *Appl. Phys A* **66**, 115–117 (1998)
12. Iijima, S.: Helical microtubules of graphitic carbon. *Nature* **354**, 56–58 (1991)
13. Iijima, S., Ichihashi, T.: Single-shell carbon nanotubes of 1-nm diameter. *Nature* **363**, 603–605 (1993)
14. Wang, J., et al.: Low temperature growth of boron nitride nanotubes on substrates. *Nano Lett.* **5**, 2528–2532 (2005)
15. Yap, Y.K., Yoshimura, M., Mori, Y., Sasaki, T., Hanada, T.: Formation of aligned-carbon nanotubes by RF-plasma-assisted pulsed-laser deposition. Special issue for Tsukuba symposium on carbon nanotubes in commemoration of the 10th anniversary of its discovery; S. Iijima et al. eds., *Physica B* **323**, 341–343 (2002)
16. Lee, K.-F., et al.: Synthesis of aligned bamboo-like carbon nanotubes using radio frequency magnetron sputtering. *J. Vac. Sci. Technol. B* **21**, 1437–1441 (2003)
17. Xie, M., Wang, J., Yap, Y.K.: Mechanism for low temperature growth of boron nitride nanotubes. *J. Phys. Chem. C* **114**, 16236–16241 (2010)
18. Heo, Y.W., Kaufman, M., Pruessner, K., Norton, D.P., Ren, F., Chisholm, M.F., Fleming, P.H.: Optical properties of $Zn_{1-x}Mg_xO$ nanorods using catalysis-driven molecular beam epitaxy. *Solid-State Electron.* **47**, 2269–2273 (2003); Heo, Y.W., Varadarajan, V., Kaufman, M., Kim, K., Norton, D.P., Ren, F., Fleming, P.H.: Site-specific growth of ZnO nanorods using catalysis-driven molecular beam epitaxy. *Appl. Phys. Lett.* **81**, 3046–3048 (2002)
19. Rahm, A., et al.: Pulsed-laser deposition and characterization of ZnO nanowires with regular lateral arrangement. *Appl. Phys. A* **88**, 31–34 (2007)

20. Nagashima, K., Yanagida, T., Tanaka, H., Kawai, T.: Epitaxial growth of MgO nanowires by pulsed laser deposition. *J. Appl. Phys.* **101**, 124304 (2007)

Physical-Chemical Etching

► Dry Etching

Physicochemical Properties of Nanoparticles in Relation with Toxicity

Jérôme Rose^{1,2}, Mélanie Auffan^{3,4}, Olivier Proux^{2,5}, Vincent Niviere^{2,6} and Jean-Yves Bottero^{3,4}

¹CEREGE UMR 6635– CNRS–Université Paul Cézanne Aix–Marseille III, Aix–Marseille Université, Europôle de l’Arbois BP 80, Aix–en–Provence Cedex 4, France

²GDRI ICEINT: International Center for the Environmental Implications of Nanotechnology, CNRS–CEA, Europôle de l’Arbois BP 80, Aix–en–Provence Cedex 4, France

³CEREGE, UMR 6635 CNRS/Aix–Marseille Université, Aix–en–Provence, France

⁴ICEINT, International Consortium for the Environmental Implications of Nanotechnology, Center for the Environmental Implications of NanoTechnology, Aix–En–Provence, Cedex 4, France

⁵OSUG, Université Joseph Fourier BP 53, Grenoble, France

⁶Laboratoire de Chimie et Biologie des Métaux, UMR 5249, iRTSV–CEA Bat. K’, 17 avenue des Martyrs, Grenoble Cedex 9, France

Synonyms

[Engineered nanoparticles](#); [Nano-effects](#); [Nanomaterials and nanoproducts](#); [Nanotoxicology](#)

Definition

Enhancement or modifications of properties in the nanometric range is the core of nanotechnologies. Nanoparticles exhibit size-related properties that differ significantly from those observed in fine

particles or bulk materials. Properties can be intrinsic (e.g., electrical, magnetic, optical, etc.) or extrinsic (e.g., aggregation state, surface area, etc.). Then what makes nanoparticles so attractive also raises questions of “new” toxic “nano-effects.” It is then worth trying to identify which specific properties can generate biological harmful effects.

More than these “new” nano-properties aspects, the second issue is to manufacture materials that incorporate those nanoparticles. Then exposure to nanoparticles will depend not only on intrinsic properties of nanoparticles, but also on extrinsic properties related to the stability and durability of their incorporation.

Overview

Introduction

The success story of nanoparticles in many scientific and technological fields is based on their unique physicochemical properties and reactivity. Indeed engineered nanoparticles are designed to exhibit specific optical, electronic, magnetic, or chemical properties that are strongly related to the exponential increase in the number of atoms localized at the surface as the size decreases [1]. However, the uniqueness of nanoparticles that makes them so attractive may also raise questions about the environmental and health risk they could generate. Indeed is there any relation between their new physicochemical properties and potential new biological hazards?

Moreover the tremendous development of nanotechnologies leads to a huge variety of manufactured nanoparticles with varying chemical composition, size, shape, crystal structure, surface area, surface chemistry and charge, solubility. . . etc., e.g., [1].

It is therefore challenging to determine which of those unique intrinsic or extrinsic properties can be at the origin of their biological danger. In the following, we will focus on the different properties, how to determine their evolution in biological media and how they can affect living organisms.

Physicochemical Properties of Nanoparticles

Even if there is no consensus concerning the definition of engineered nanoparticles, most of the definitions indicate that besides the size, “*Novel properties*

differentiate nanoparticles from the bulk material typically develop at a critical length scale of under 100 nm” or “*phenomena not observed in larger structures start to become apparent ...*”. But what are these novel properties or phenomena not observed for larger structures? What is the size at which the latter properties occur? One of the most emblematic example concerns gold. Indeed gold has been a topic of fascination since its discovery in almost all ancient and modern civilizations. Gold preserve its brightness for years which made his high value for art. Everybody has in mind that the magnificent death mask of Egyptian pharaoh Tutankhamun is made of gold. Moreover, gold chemical inertness is related to its high monetary value [2]. But its chemical inertness is a matter of size since at the nanoscale it can be exceptionally active as a catalyst for instance but only for particles below 2–3 nm [3].

Among the vast diversity of engineered nanoparticles, many other properties can be tailored in different scientific and technological fields. Indeed optical properties can be modified to use nanoparticles as biological dyes for instance. QDots are of good examples to trace, label, or image biological organisms or cells. In the case of QDots, the band gap energy (related to the fluorescence wavelength that is to say the color) is constant for large particles whatever their size but only if larger than 6–8 nm. The bandgap energy strongly increases when size of particles decreases below 6–8 nm. Photocatalytic properties may also be tailored not only as function of size, but also as function as the crystal structure. TiO₂ anatase phase exhibits optimum photocatalytic degradation of organic contaminants at various sizes that are much below 100 nm (e.g., ~7 nm for trichloroethylene or around 11 nm for chloroform). But in any case, anatase is much more efficient than rutile, indicating that besides size, crystal structure also strongly affects physicochemical properties.

Thermal properties may also vary in relation with size decrease. For instance, tin and indium melting point can be strongly reduced by 80°C and 120°C, respectively (e.g., in [1]) when their diameters decrease from 100 to 10 nm. More surprisingly, their normalized heat of fusion that is supposed to remain constant in classical thermodynamics exponentially decreases when size decreases below 20 nm in diameter. Solubility is also affected by size decrease. It is intuitive to predict that the dissolution kinetics of

particles is function of the specific area and aggregation state, with of course faster dissolution kinetic for the smallest particles, but what about the thermodynamic crystal solubility K_b ? K_b is supposed to be constant and generally similar to the solubility product K_{sp} .

$\ln K_b = \ln K_{sp} + f(\gamma/r)$ (with γ is the surface tension and r is the characteristic length of the crystal)

For particles larger than 25 nm, $f(\gamma/r)$ can be approximated as a constant function. But this simplification fails for smaller nanoparticles as the size dependence of the morphology and γ cannot be ignored.

Other examples of physicochemical properties modification with size could be given (magnetic, electrical, etc.). Besides these intrinsic properties, it is worth detailing nanoparticles when introduced in aquatic media. Indeed some confusion may exist between dissolution and dispersion. Due to their small size, when nanoparticles are homogeneously distributed in aquatic media, particles cannot be detected by eyes and appear as if particles may be totally dissolved to ionic species. However, (colloidal) dispersion is the term referred to suspended material in the 1–250 nm range. Colloidal dispersions are thermodynamically unstable even if the aggregation rate can be very slow.

Latter examples are good illustrations of the high variety of properties that can change or be enhanced as size of particles decreases in the nanometric size range. It is therefore challenging to prioritize which properties are of main concern in relation with toxicity.

Determination of Nanoparticle's Properties in Biological Media

Generalities

Even if challenging, determining the evolution of physicochemical properties of nanoparticles in relation to human and environmental exposure requires to answer to, at least, two questions, in order to define the appropriate useful techniques.

First of all, which structural and characterization information are considered to be important? As detailed before their aggregation in the media, their (mean) size and shape, dissolution or solubility, their surface area, redox stability, surface charge, and surface chemistry are of prime interest.

Secondly, in which conditions the nanoparticles have to be characterized? A three-step procedure have been formalized in which the characterizations need ideally to be performed (1) *on as-synthesized nanoparticles* in their native state (dry or hydrated state, depending on the synthesis), (2) *in situ on nanoparticles dispersed in the media* used for the environment and toxicity tests (ultrapure water, phosphate-buffered saline solution cell culture media...), and (3) *in vivo or in vitro on particles interacting with the living systems* tested (bacteria, cells, animals, plants, etc.).

Numerous characterization and detection techniques exist to determine these physicochemical properties of nanoparticles in a toxicological or biological context. There is no best technique, each of them has their specificities, advantages, and disadvantages. However, in order to limit at the maximum the discrepancies between the results, the same technique has to be employed for each step of the characterization. Several recent review papers summarize different and complementary analytical techniques (e.g., [4]...). We will here focus on recent techniques using X-ray. Their large penetration depth, especially in the hard X-rays regime, is interesting in a large range of applications because of its compatibility with *in situ* or *in vivo* conditions, in most of the cases without any special sample preparation, directly on the nanoparticles in suspension, for example. Moreover, specific interactions of X-rays with matter provide elemental and chemical sensitivity that have made X-ray techniques a very attractive tool.

X-Ray Imaging

Advances in X-ray spectromicroscopy techniques based on new third-generation synchrotron sources have opened unprecedented opportunities in terms of spatial resolution to analyze nano-sized compounds in biological systems [5]. Different possibilities exist to improve the contrast using X-ray microscopes with respect to optical or electron microscopes. They can be schematically described looking at the different interaction processes occurring between the incident photons (or the associated wave) and the sample. (1) The *photon absorption* which strongly depends on the chemical nature of the element (the higher the atomic number of an element, the stronger will be the absorption). (2) The *fluorescence emission* resulting from the decay process. (3) The *phase shift*

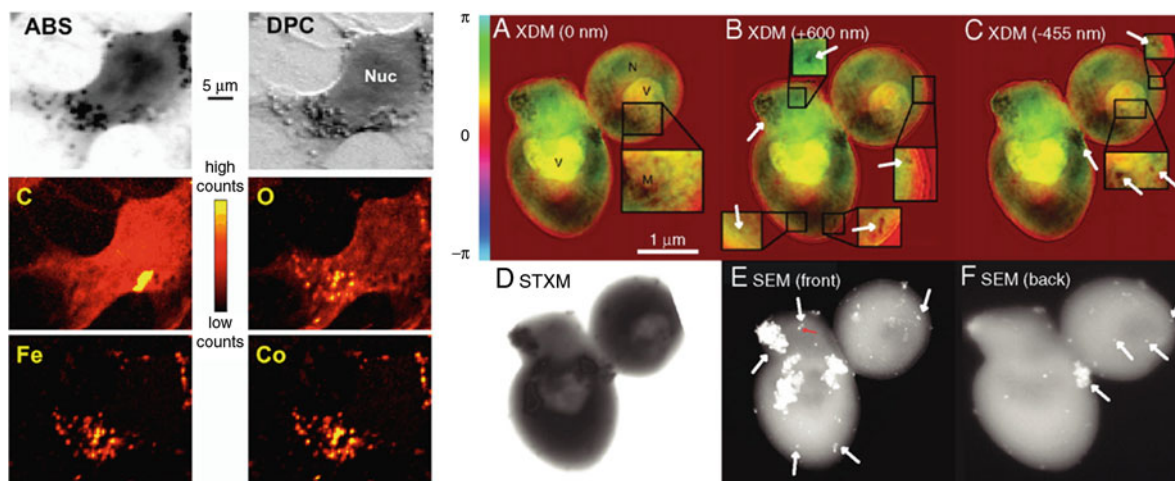
of the incoming wave induced by the matter/sample interaction. (4) The *photon diffraction* by the sample. These processes give rise to different kinds of X-ray microscopy, among them Transmission X-ray Microscopy (TXM), Scanning Transmission X-ray Microscopy (STXM), X-ray fluorescence with microbeam (μ XRF), and X-ray Diffraction Microscopy (XDM).

Examples of imaging performed with these techniques are shown on Fig. 1, on fibroblast cells exposed to nanoparticles using STXM in absorption (ABS) and differential phase contrast (DPC) modes combined with μ XRF (Fig. 1, left) [5], and on yeast using XDM compared with Scanning Electron Microscopy (SEM) and STXM (Fig. 1, right) [6].

First example deals with the toxicity of CoFe_2O_4 nanoparticles, nano-sized magnets that can be potentially used in combined cancer therapy if their biocompatibility is demonstrated. μ XRF maps indicate that, for the probed concentrations, these nanoparticles do not penetrate the nucleus. Precise localization of the elements with respect to the different constituting parts of the cell is possible by acquiring at the same time an image, the best contrast being obtained in DPC mode (Fig. 1, left, upper part). Spatial resolution in this soft-X-ray microscope is sub-50 nm. An improvement of the resolution can be achieved by shifting to DCM microscopy as shown by Nelson et al. [6] who obtained a "resolution between 11–13 nm, the highest resolution in X-ray imaging of a biological specimen." *Saccharomyces cerevisiae* yeasts were imaged after labeling using colloidal gold particles. These labels were positioned to correlate the images obtained by the three different ways, to have a probe to estimate the resolution and the probing depth. The 3D XDM image obtained after reconstruction allows to image the internal part of the cell (which can be compared to the image obtained in transmission using STXM) or the two opposite walls of the cell (which can be compared in this case to the two images obtained in SEM). This technique is really promising even if it still remains quite challenging, to image thick cells, too thick to be studied by transmission electron microscopy.

From Long-Range to Short-Range Orders in Nanomaterials

X-ray scattering and X-ray Absorption Spectroscopy techniques (XAS) techniques provide valuable information on long- and short-range orders existing inside



Physicochemical Properties of Nanoparticles in Relation with Toxicity, Fig. 1 *Left.* Uptake and interaction of CoFe_2O_4 nanoparticles with fibroblast cells imaged using absorption (ABS) and differential phase contrast (DPC) combined with

μXRF analysis of the C, O, Fe, and Co (From [5]). *Right.* *Saccharomyces cerevisiae* yeast imaged using XDM (three images corresponding to three different analyzing depth), STXM, and SEM ([6])

(X-Ray Diffraction, XRD) and between (Small-Angle X-ray Scattering, SAXS) nanoparticles, with a great sensitivity to their shape.

SAXS experiments have been performed to follow the kinetics of aggregation or dissolution of ZnS nanoparticles as a function of the mean size of the particles and of the pH of the solution [7]. Levard et al. [8] have combined SAXS and XAS in situ experiments to study the growth mechanisms of imogolite-like aluminogermanate nanotubes at various stages of their formation, from the precursor (roof-tile-shaped particles) to the nanotubular structure. 20% of Ge vacancies have been also determined in the Ge layer from XAS analysis. Finally, structure of iron oxide nanoparticles in aqueous suspension was probed by high-energy powder XRD on a large angular range coupled to atomic pair distribution function (PDF) analysis, well suited for determining the internal atomic ordering and the geometry of nanoparticles [9]. This PDF technique provided detailed structural information for $\gamma\text{-Fe}_2\text{O}_3$ spherical and tetrapod-shaped nanoparticles.

These techniques are performed using synchrotron radiation in order to perform kinetics experiments and XAS spectroscopy, but SAXS and XRD experiments can also be performed using laboratory sources, preferably at high energy (Mo or Ag $K\alpha$ anode).

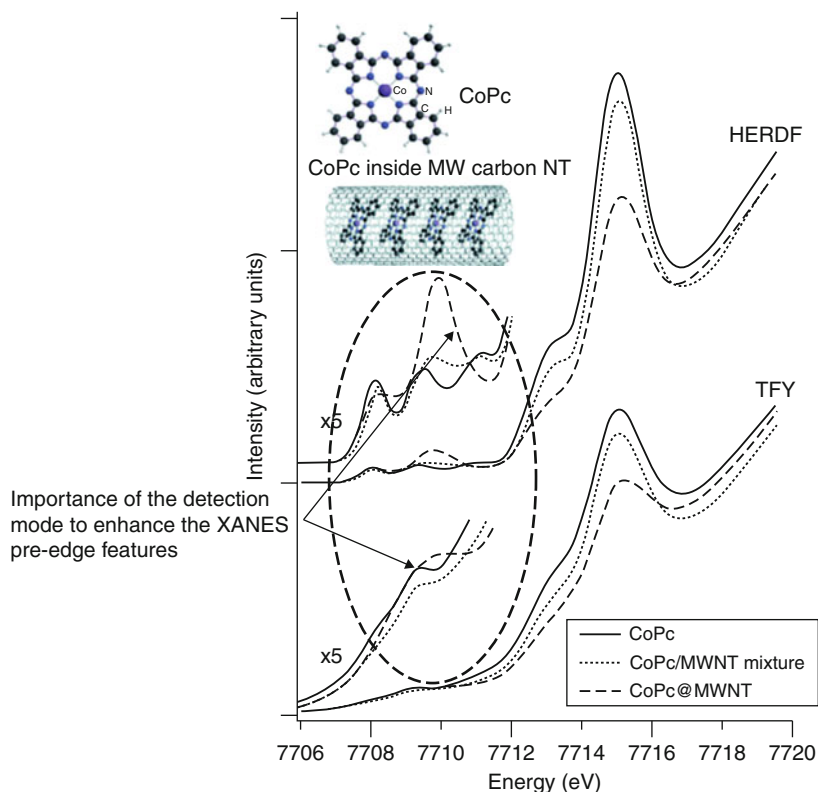
Electronic Structure and Short-Range Order in Nanomaterials

XAS is now a well-established structural technique that allows to determine the electronic structure (valence) and the local order (geometry of the site, chemical nature of the neighbors, etc.) of a selected element in a sample. Being only sensitive to the local order, XAS is well suited to study either amorphous, ill-ordered, or well-crystallized samples, in the solid or liquid states, and has been then quite naturally used for the nanoparticles studies [10]. Numerous examples of the application of this technique for nanotoxicology studies can be found (e.g., [11, 12]). In this particular field, two conditions are required. Acquisitions need to be performed in the fluorescence mode, to be sensitive to a small amount of nanoparticles, using, for example, of conventional solid state detector, with a medium energy resolution (ca. 200–250 eV). Moreover, a high resolution is necessary to detect accurately the subtle variations upon interaction with a solution or a biological media. This second requirement is especially true in the pre-edge and edge part of the signal (X-ray Absorption Near Edge Structure, XANES).

High-Energy Resolution Fluorescence Detection (HERFD) allows to further enhance XANES sensitivity. Even if not extensively used at present, this technique has gained a lot of interests and under development on many synchrotron beamlines. Using a detector with an

Physicochemical Properties of Nanoparticles in Relation with Toxicity,

Fig. 2 TFY XANES pre-edge and edge region compared to HERFD spectra for three CoPc samples. The pre-edge is magnified by 5 for clarity (From [13])



energy resolution better than the core hole lifetime broadening gives XANES spectra with sharper spectral features compared to conventional total fluorescence (TFY) spectra. Example of such sharpening is illustrated in Fig. 2. The structure of Cobalt phthalocyanine (CoPc), free or encapsulated inside multiwalled carbon nanotubes has been probed using either TFY or HERFD [13]. The differences in the pre-edge using conventional TFY spectra are really small, until those obtained by HERFD clearly put in evidence a modification of the Co conformation when CoP is free or encapsulated. Complete analysis shows that Co atoms move out of the molecular plane by a fraction of Angstrom.

Physicochemical Properties: Biological Effects Relationships

Some recent reviews inventoried NPs that are toxic to living organisms, e.g., [14–16]. There are now obvious examples of the toxic effects of C₆₀ fullerenes to bacteria, to daphnia, etc., Nano-TiO₂, nano-CeO₂, nano-Ag, etc., can also be toxic to various organisms ([14, 15] etc.), or even colloidal silica to rat and mice

[16]. But despite the growing body of literature, many conclusions remain contradictory. For instance, dose–response relationships are not always observed even for equivalent systems. Lovorn et al. [17] clearly observed such relationship in the case of daphnia mortality exposed to 25 nm TiO₂, while Hund-Rinke et al. [18] were unable to determine a LC₅₀ value for 30 nm TiO₂ in contact with *Daphnia* due to chaotic results. In many cases, the degree of physicochemical characterization of nanoparticles is not equivalent and one cannot identify why results are inconsistent. As previously discussed, nanoparticles can differ by many aspects due to variation in size, shape, redox properties, surface charge, crystal structure, chemical composition, adhesion to surfaces, etc., and beyond a list of results, it is worth addressing the effects of various physicochemical parameters of NPs on their toxic effects.

Effects of Aggregation State and Size

Most of pioneer published results dealing with NPs toxicity mainly focused on size effects. Can toxic effects only be due to the small size of particles? Size was mainly measured by TEM that implied dry state characterization. However, the size was not characterized in

nutritive or aquatic test media. Indeed the dispersion state of NPs (aggregation vs. dispersion) in contact with living organisms can change from one study to another. At this stage, it is worth detailing that here aggregation is referred to an assembly of particles which are loosely attached to each other. It means that aggregation is a reversible phenomenon and that NPs can aggregate in aquatic media but can disaggregate in presence of stabilizing agents like proteins. Then one has to keep in mind that aggregation can be reversible. There is a generally accepted idea assuming that aggregation of NPs impedes the determination of specific properties of NPs. It is obvious that the transport and mobility are strongly affected. Nevertheless, intrinsic or specific “nano” physicochemical properties of NPs are not necessarily (or even significantly) affected by aggregation processes. In the particular case of surface chemical reactivity, Auffan et al. [1, 19] have shown that surface normalized quantity of probe molecules sorbed to iron oxide particles is strongly enhanced for particles smaller than 12 nm. Using arsenite as surface reactivity molecular probe, the quantity of arsenite adsorbed to the maghemite ($\gamma\text{Fe}_2\text{O}_3$) surface per surface unit reached 3.6 As/nm² for 300 nm and 20 nm particle size while it increased up to 11 As/nm² for 12 nm particles. This surface reactivity was not affected by aggregation [19]. The example of Nano- Fe_2O_3 in contact with *E. coli* can illustrate that the aggregation state as well as the size may not be the most important parameters in terms of toxicity [20]. Indeed, [20] have shown that 8 nm maghemite (Fe_2O_3) had no toxic effect toward *E. coli* regardless of the aggregation state of the NPs.

Effects of the Redox State Evolution and Instability

As function of their chemical composition redox states of NPs can change in biological redox/pH conditions. Indeed most elements exhibit various possible redox states that are or are not stable as function of Eh and pH values. It is obvious for all metals that they can exist under metallic or oxidized forms. But even when oxidized, some elements can exist under various redox states. Metals like Fe, Mn, Cr, or rare earth like Ce illustrate such variability (Fe^{2+} , Fe^{3+} ; Mn^{2+} , Mn^{3+} , Mn^{4+} , Mn^{7+} ; Cr^{3+} , Cr^{6+} ; Ce^{3+} , Ce^{4+} , etc.). Living organisms and especially eukaryotic cells metabolism cannot support large pH variation. For most eukaryotic cells, pH is buffered in the 7–8 pH range. In the case of redox potential (noted Eh), oxidative phosphorylation in mitochondria is based on a series of redox reactions at near

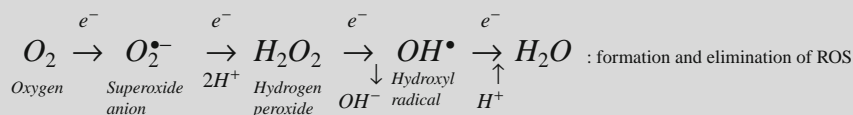
circumneutral pH for which potentials are from -0.32 (NAD^+/NADH) to 0.29 V (cytochromes). Besides oxidative phosphorylation (respiration), extracellular Eh is generally controlled by thiol/disulfide redox systems (mainly GSH/GSSH and Cys/CySS) for which Eh vary in the $-0.140/-0.08$ V range. In such conditions, many elements can be redox unstable leading to electron exchange between NP surface and surrounding media. This could then be the starting point of disequilibrium of the redox balance and then to oxidative stress (See Box 1). In the specific case of iron, Auffan et al. [21] have shown that nano-zero valent iron as well as nano-magnetite (Fe_3O_4) can be toxic to *E. coli*. The study reported that the oxidation of iron in contact to bacteria can be at the origin of the toxicity. The potential of the redox couple characteristic of Fe^0 nanoparticles at pH = 7 is in the $\text{Eh}_{\text{Fe}^{2+}/\text{Fe}(0)} = -0.40/-0.7$ V range as function of the Fe concentration [22] which is lower than the typical Eh in biological media. This allowed the oxidation of $\text{Fe}^0(\text{s})$ in contact with *E. coli*. This oxidation is directly responsible of toxic effects through the generation of an oxidative stress, as demonstrated by an enhanced toxicity of these Fe NPs toward a mutant strain of *E. coli* deprived of defense mechanism against oxidative stress. Other examples in the literature confirm the importance of the redox “instability” of nanoparticles. Using XANES at Ce L3-edge, Thill et al. [12] and Auffan et al. [11] have shown that the cytotoxicity/genotoxicity of CeO_2 was due to the reduction of surface Ce(IV) ions to Ce(III).

Effects of NPs Shape and Aspect Ratio

The toxicity of particles as function of their shape does not concern only nanoparticles. Indeed, many studies concerning asbestos toxicity have demonstrated that the toxicity can be partly related to the aspect ratio. Then comparison between asbestos and carbon nanotubes legitimated the worry concerning probable toxic effects of NPs. Poland and coworkers [23] have shown that “*long multiwalled carbon nanotubes result in asbestos-like, length dependent, pathogenic behavior.*” However, the mechanisms at the origin of the toxicity of long ($>15-20$ μm) nanotubes still remain unclear. In the case of nano-silver, it is possible to produce spherical, truncated triangular, rods nanoparticles. Pal et al. [24] have shown that the biocidal effects depend on silver NPs shape with the higher effect for truncated triangular NPs. It has been emphasized that the {111} lattice plane exhibit the

Box 1

Oxidative stress is a major problem that concerns all the cells living in the presence of dioxygen. It originates from the production of highly toxic reactive oxygen species (ROS), including superoxide radical $O_2^{\bullet-}$, hydrogen peroxide H_2O_2 and hydroxyl radical HO^{\bullet} . Although cells have developed efficient antioxidant systems to detoxify ROS, any deficiency of these systems or overload by overproduction of ROS can result in severe cellular damages such as proteins, lipids, and DNA oxidations, which can ultimately lead to cell death. It is now well documented that oxidative stress is involved in the pathogenesis of a variety of diseases such as cancers, asthma, pulmonary diseases, Alzheimer, Parkinson, and in ageing [1].



Several studies have related the toxicity of some NPs to an oxidative stress toward biological models [2]. In order to better understand the cellular oxidative stress pathways induced by NPs, identification and quantification of the specific ROS produced directly or indirectly by NPs represent an important issue.

In vivo detection and quantification of ROS are far from being trivial and require the use of specific techniques [3, 4]. This is mainly due to the fact that ROS are unstable species and are present at very low steady-state concentration in the cells. Various methods using synthetic probes have been developed to detect ROS in vivo. Some of these probes lead to a fluorescence signal upon oxidation by ROS, which is very well adapted for in cellulo detection by fluorescence microscopy. However, some of them lack of specificity and may lead to false interpretations for the presence of ROS [3, 4]. For example, the use of dichlorofluorescein (DCF) as a fluorescent peroxide probe is complex and crucially lacks of selectivity in vivo [5]. Hydroethidine (HE, MitoSOX™), a cell permeable compound which is specifically oxidized by $O_2^{\bullet-}$ into the fluorescent 2-hydroxyethidium (2-OH-E+) compound, has been described as an interesting $O_2^{\bullet-}$ probe [6]. However, in the cell, HE also generates other fluorescent compounds that are not related to the reaction with $O_2^{\bullet-}$. Consequently, direct quantification of 2-OH-E+ production in whole cells by fluorescence microscopy is not possible, and chromatography analysis of extracts obtained after cell lysis must be carried out in order to quantify 2-OH-E+ production [7].

In vitro, ROS detection is much less challenging since control experiments can be usually carried out in order to make corrections for nonspecific reactions. Probes like cytochrome c and tetrazolium salts XTT have been widely used to detect $O_2^{\bullet-}$ in vitro [3, 4]. Spin trap nitrones, like DMPO or DEPMPO, associated with electron paramagnetic resonance spectroscopy (EPR) are well adapted to detect the unstable radical species $O_2^{\bullet-}$ and HO^{\bullet} [3]. However, in vivo, their use is more difficult since the nitron adducts resulting from the reaction of the spin trap with $O_2^{\bullet-}$ or HO^{\bullet} are unstable and can be easily reduced by cellular reductants [3].

1. Winterbourn, C.: Reconciling the chemistry and biology of reactive oxygen species. *Nat. Chem. Biol.* **5**, 278–286 (2008)
2. Nel, A., Xia, T., Madler, L., Li, N.: Toxic Potential of Materials at the Nanolevel. *Science* **311**, 622–627 (2006)
3. Bartosz, G.: Use of spectroscopic probes for detection of reactive oxygen species. *Clin. Chim. Acta* **368**, 53–76 (2006)
4. Wardman, P.: Fluorescent and luminescent probes for measurement of oxidative and nitrosative species in cells and tissues: Progress, pitfalls, and prospects. *Free Rad. Biol. Med.* **43**, 995–1022 (2007)
5. Zielonka, J., Kalyanaraman, B.: ROS-generating mitochondrial DNA mutations can regulate tumor cell metastasis”— a critical commentary. *Free Rad. Biol. Med.* **45**, 1217–1219 (2008)

6. Robinson, K.M., Janes, M.S., Pehar, M., Monette, J.S., Ross, M.F., Hagen, T.M., Murphy, M.P., Beckman, J.S.: Selective fluorescent imaging of superoxide in vivo using ethidium-based probes. *Proc. Natl. Acad. Sci.* **103**, 15038–15043 (2006)
7. Zielonka, J., Vasquez-Vivar, J., Kalyanaram, B.: Detection of 2-hydroxyethidium in cellular systems: a unique marker product of superoxide and hydroethidine. *Nat. Protocol.* **3**, 8–21 (2008)

highest reactivity to sulfur- or phosphorus-containing soft bases, such as R-S-R, R-SH, RS, or PR₃. Then sulfur-containing proteins in the membrane interacting with Ag NPs (with high affinity {111} lattice plane) can lead to membrane disruption or strong damage.

Effects of NPs Solubility

As discussed in the “Physicochemical properties of Nanoparticles” part, NPs solubility may substantially vary from thermodynamic prediction deduced from larger particles. The small size and consequently large surface area is of high concern to address possible NPs dissolution. It has been mentioned that in the case of toxic metals the dissolution will increase the concentration of free metal ions in contact with living organism. This is particularly true for silver nanoparticles for which the dissolution is associated with a change of the redox state (oxidative dissolution). It has been reported that upon contact with a bacterial cell, biocide Ag⁺ is released from the surface of NPs, which then interacts with bacterial cell membranes and inhibits respiratory enzymes (e.g., [25]). In the case of ZnO NPs, dissolution and release of Zn²⁺ play an important role in the biological toxicity (e.g., [26]). Dissolution is also discussed in the case such as Quantum Dots, even if recent developments prevent or strongly decrease QDots solubility via specific surface coating.

Effects of NP Adhesion to Cells

NPs can be chemically unstable (dissolution, oxidation, reduction, etc.), inducing biological effects. However, the toxicity level may be enhanced or alleviated by additional parameters. For instance, it has been shown that nano-CeO₂ can be reduced when both *E. coli* and *Synechocystis* are exposed [27]. But the reduction of Ce⁴⁺ to Ce³⁺ induced a much more intense effect on *E. coli*. Indeed, at the highest nano-CeO₂ concentration, the percentage of survival cells was still 35% for *Synechocystis* while no cell was alive in

the case of *E. coli*. This difference may be explained by the production of extracellular polymeric substances (EPS) preventing direct contact between cell membranes and nano-CeO₂ due to strong adsorption and adhesion of NPs. Thus, in the event of redox phenomena, the added distance between NPs and the membranes and the EPS shield itself may provide at least partial protection against the toxicity.

Effects of Chemical Composition, Crystal Structure. . .

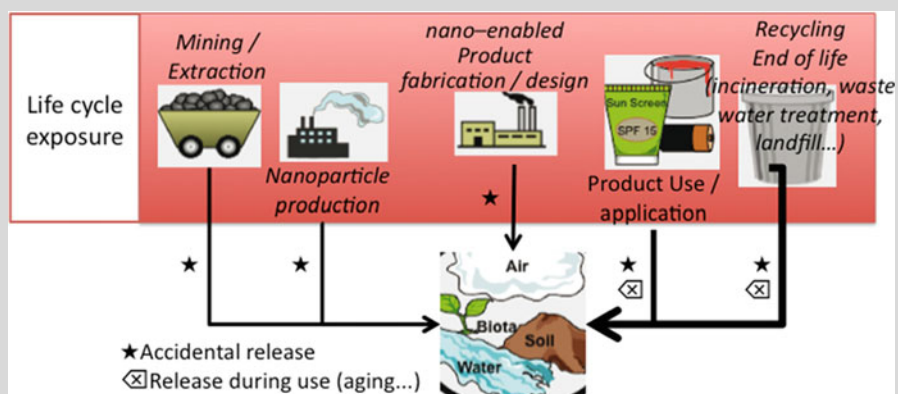
Besides NPs properties discussed above, other parameters can induce biological effects. The chemical composition of NPs is of utmost importance. Many metals and metalloids are known to be toxic like for instance Cd or Se. Their use by nanotechnology to elaborate nanoproducts or nanomaterials must be done with care. Any dissolution of NPs incorporating Cd (for instance) will lead to toxicity, and it is worth developing any specific design to enhance their chemical stability and prevent at best their direct contact with external media. Most importantly, one has to define whether or not other less toxic element can be used to reach the same level of performance.

From Nanoparticles to Nano-Residues of the Nanomaterials Life Cycle

Nanoparticles and nanomaterials may be released from point sources like factories or landfills that may be more related to accidental events or technical defects during the processes. But most of the release may be due to indirect exposure through the life cycle of the nanomaterials. Indeed, for most applications, nanoparticles are surface modified and generally are embedded in the final product. Therefore, they do not come into direct contact with consumers or the environment. Their surface modification can strongly affect their behavior and especially their toxicity. NPs in commercial sunscreens can illustrate the difference between bare nanoparticles and surface modified nanoproducts. Indeed, the nanoparticles of TiO₂ are

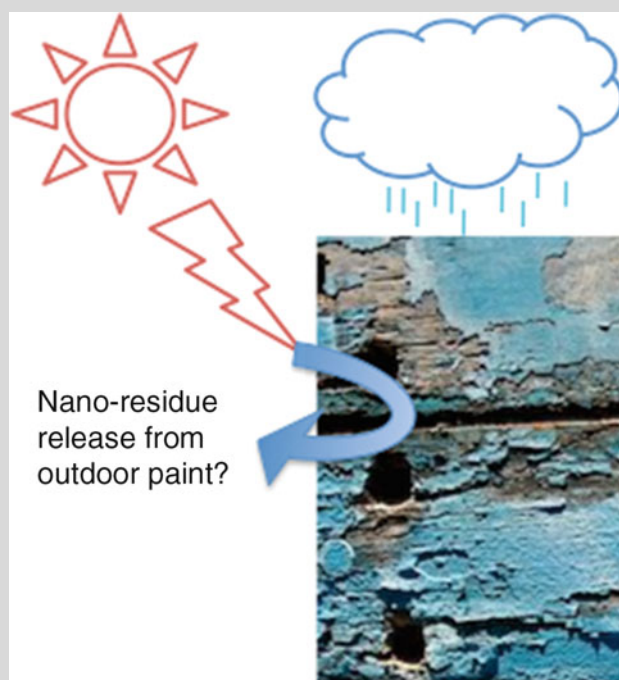
Box 2: Exposure During Life Cycle of Nanoparticles and Nanomaterials

Most of the time, nanoparticles are surface modified prior to their incorporation in final products. Therefore, it is very likely that nanoparticles will not be directly released in the environment.



The degradation of products due to external conditions (sun, rain, etc.) may generate the release of by-products containing nanoparticles.

For instance:



coated with AlOOH and polydimethylsiloxane for their use as sunscreens. The release of nano-residues from commercial NPs-based sunscreens is no longer a hypothesis, but has become a fact [28–30]. The major effect of the remaining Al-based coating is a protection

against the generation of superoxide ions from the photoactive/phototoxic TiO₂ core. This surface modification of the nano-TiO₂ results in a substantially lower toxic potential of the residue as long as the protective layer remains stable (See Box 2).

Conclusion

This section discussed the relationships between the physicochemical properties of nanoparticles and their biological effects. As described above, nanoparticles differ from each other by many aspects and not only size and shape. Aggregation state, chemical stability, solubility, etc. can change from one to another NP leading to difficult comparison when toxicity occurs. Then, even if many issues remain unsolved, it clearly appears that the toxicity strongly depends on an accurate physicochemical characterization of nanoparticles and nanomaterials. However, while characterizing several properties like chemical composition, size, shape, crystal structure, surface area, surface chemistry, charge, and solubility of the initial NPs is now well accepted (but not always performed!), information is still lacking as, more than initial product characterization, it is worth determining the evolution of important parameters during the interaction between NPs and biological organisms. However, such accurate determination of nanoparticles in biological media remains highly challenging as discussed in this section.

Acknowledgements The authors would like to thank the Centre National de la Recherche Scientifique (CNRS) and the Commissariat à l'Énergie atomique (CEA) for funding the International Consortium for the Environmental Implications of NanoTechnology (GDRi iCEINT).

Cross-References

- ▶ [Cellular Mechanisms of Nanoparticle's Toxicity](#)
- ▶ [Effect of Surface Modification on Toxicity of Nanoparticles](#)
- ▶ [Electron Microscopy of Interactions Between Engineered Nanomaterials and Cells](#)
- ▶ [In Vitro and In Vivo Toxicity of Silver Nanoparticles](#)
- ▶ [In Vivo Toxicity of Titanium Dioxide and Gold Nanoparticles](#)
- ▶ [Nanoparticle Cytotoxicity](#)
- ▶ [Selected Synchrotron Radiation Techniques](#)

References

1. Auffan, M., Rose, J., Bottero, J.Y., Lowry, G.V., Jolivet, J.P., Wiesner, M.R.: Towards a definition of inorganic nanoparticles from an environmental, health and safety perspective. *Nat. Nanotechnol.* **4**, 634–641 (2009)
2. Hutchings, G.J., Brust, M., Schmidbauer, H.: Gold - an introductory perspective. *Chem. Soc. Rev.* **37**, 1759–1765 (2008)
3. Turner, M., Golovko, V.B., Vaughan, O.P.H., Abdulkin, P., Berenguer-Murcia, A., Tikhov, M.S., Johnson, B.F.G., Lambert, R.M.: Selective oxidation with dioxygen by gold nanoparticle catalysts derived from 55-atom clusters. *Nature* **454**, 981–U931 (2008)
4. Hassellöv, M., Readman, J.W., Ranville, J.F., Tiede, K.: Nanoparticle analysis and characterization methodologies in environmental risk assessment of engineered nanoparticles. *Ecotoxicology* **17**, 344–361 (2008)
5. Kaulich, B., Thibault, P., Gianoncelli, A., Kiskinova, M.: Transmission and emission x-ray microscopy: operation modes, contrast mechanisms and applications (Topical Review). *J. Phys. Condens. Matter* **23**(8), 083002 (2011)
6. Nelson, J., Huang, X., Steinbrener, J., Shapiro, D., Kirz, J., Marchesini, S., Neiman, A.M., Turner, J.J., Jacobsen, C.: High-resolution x-ray diffraction microscopy of specifically labeled yeast cells. *Proc Natl Acad Sci* **107**, 7235–7239 (2010)
7. Zhang, H., Chen, B., Banfield, J.F.: Particle size and pH effects on nanoparticle dissolution. *J. Phys. Chem. C* **114**, 14876–14884 (2010)
8. Levard, C., Rose, J., Thill, A., Masion, A., Doelsch, E., Maillat, P., Spalla, O., Olivi, L., Cognigni, A., Ziarelli, F., Bottero, J.Y.: Formation and growth mechanisms of imogolite-like aluminogermanate nanotubes. *Chem. Mater.* **22**, 2466–2473 (2010)
9. Petkov, V., Cozzoli, D., Buonsanti, R., Cingolani, R., Ren, Y.: Size, shape, and internal atomic ordering of nanocrystals by atomic pair distribution functions: a comparative study of γ -Fe₂O₃ nanosized spheres and tetrapods. *J. Am. Chem. Soc.* **131**, 14264–14266 (2009)
10. Modrow, H.: Tuning nanoparticle properties - the X-ray absorption spectroscopic point of view. *Appl. Spectrosc. Rev.* **39**, 183–290 (2004)
11. Auffan, M., Rose, J., Orsiere, T., De Meo, M., Thill, A., Zeyons, O., Proux, O., Masion, A., Chaurand, P., Spalla, O., Botta, A., Wiesner, M.R., Bottero, J.-Y.: CeO₂ nanoparticles induce DNA damage towards human dermal fibroblasts in vitro. *Nanotoxicology* **3**, 161–171 (2009)
12. Thill, A., Zeyons, O., Spalla, O., Chauvat, F., Rose, J., Auffan, M., Flank, A.M.: Cytotoxicity of CeO₂ nanoparticles for *Escherichia coli*. Physico-chemical insight of the cytotoxicity mechanism. *Environ. Sci. Technol.* **40**, 6151–6156 (2006)
13. Swarbrick, J.C., Weng, T.-C., Schulte, K., Khlobystov, A.N., Glatzel, P.: Electronic structure changes in cobalt phthalocyanine due to nanotube encapsulation probed using resonant inelastic X-ray scattering. *Phys. Chem. Chem. Phys.* **12**, 9693–9699 (2010)
14. Klaine, S.J., Alvarez, P.J.J., Batley, G.E., Fernandes, T.F., Handy, R.H., Lyon, D.Y., Mahendra, S., McLaughlin, M.J., Lead, J.R.: Nanomaterials in the environment: behavior fate, bioavailability, and effects. *Environ. Toxicol. Chem.* **27**, 1825–1851 (2008)
15. Handy, R.H., von der Kammer, F., Lead, J.R., Hassellöv, M., Owen, R., Crane, M.: The ecotoxicology and chemistry of manufactured nanoparticles. *Ecotoxicology* **17**, 287–314 (2008)

16. Napierska, D., Thomassen, L.C.J., Lison, D., Martens, J.A., Hoet, P.H.: The nanosilica hazard: another variable entity. *Part. Fibre Toxicol.* **7**(1), 39 (2010)
17. Lovern, S.B., Klaper, R.: *Daphnia magna* mortality when exposed to titanium dioxide and fullerene [C60] nanoparticles. *Environ. Toxicol. Chem.* **25**, 1132–1137 (2006)
18. Hund-Rinke, K., Simon, M.: Ecotoxic effect of photocatalytic active nanoparticles (TiO₂) on algae and daphnids. *Environ. Sci. Pollut. Res.* **13**, 225–232 (2006)
19. Auffan, M., Rose, J., Proux, O., Borschneck, D., Masion, A., Chaurand, P., Hazemann, J.L., Chaneac, C., Jolivet, J.P., Wiesner, M.R., Van Geen, A., Bottero, J.Y.: Enhanced adsorption of arsenic onto maghemite nanoparticles: As (III) as a probe of the surface structure and heterogeneity. *Langmuir* **24**, 3215–3222 (2008)
20. Auffan, M., Decome, L., Rose, J., Orsiere, T., De Meo, M., Briois, V., Chaneac, C., Olivi, L., Berge-LeFranc, J.L., Botta, A., Wiesner, M.R., Bottero, J.Y.: In vitro interactions between DMSA-coated maghemite nanoparticles and human fibroblasts: A physicochemical and cyto-genotoxicological study. *Environ. Sci. Technol.* **40**, 4367–4373 (2006)
21. Auffan, M., Achouak, W., Rose, J., Roncato, M.A., Chaneac, C., Waite, D.T., Masion, A., Woicik, J.C., Wiesner, M.R., Bottero, J.Y.: Relation between the redox state of iron-based nanoparticles and their cytotoxicity toward *Escherichia coli*. *Environ. Sci. Technol.* **42**, 6730–6735 (2008)
22. Pourbaix, M.: Atlas of electrochemical equilibria in aqueous solutions, 2nd edn. National Association of Corrosion Engineers, Houston (1974)
23. Poland, C.A., Duffin, R., Kinloch, I., Maynard, A., Wallace, W.A.H., Seaton, A., Stone, V., Brown, S., MacNee, W., Donaldson, K.: Carbon nanotubes introduced into the abdominal cavity of mice show asbestos-like pathogenicity in a pilot study. *Nat. Nanotechnol.* **3**, 423–428 (2008)
24. Pal, S., Tak, Y.K., Song, J.M.: Does the antibacterial activity of silver nanoparticles depend on the shape of the nanoparticle? A study of the gram-negative bacterium *Escherichia coli*. *Appl. Environ. Microbiol.* **73**, 1712–1720 (2007)
25. Choi, O., Hu, Z.: Size dependent and reactive oxygen species related nanosilver toxicity to nitrifying bacteria. *Environ. Sci. Technol.* **42**, 4583–4588 (2008)
26. Xia, T., Kovoichich, M., Liong, M., Madler, L., Gilbert, B., Shi, H.B., Yeh, J.I., Zink, J.I., Nel, A.E.: Comparison of the mechanism of toxicity of zinc oxide and cerium oxide nanoparticles based on dissolution and oxidative stress properties. *ACS Nano* **2**, 2121–2134 (2008)
27. Zeyons, O., Thill, A., Chauvat, F., Menguy, N., Cassier-Chauvat, C., Orear, C., Daraspe, J., Auffan, M., Rose, J., Spalla, O.: Direct and indirect CeO₂ nanoparticles toxicity for *E. coli* and *Synechocystis*. *Nanotoxicology* **3**, 284–295 (2009)
28. Auffan, M., Pedeutour, M., Rose, J., Masion, A., Ziarelli, F., Borschneck, D., Chaneac, C., Botta, C., Chaurand, P., Labille, J., Bottero, J.Y.: Structural degradation at the surface of a TiO₂-based nanomaterial used in cosmetics. *Environ. Sci. Technol.* **44**, 2689–2694 (2010)
29. Labille, J., Feng, J.H., Botta, C., Borschneck, D., Sammut, M., Cabie, M., Auffan, M., Rose, J., Bottero, J.Y.: Aging of TiO₂ nanocomposites used in sunscreen. Dispersion and fate of the degradation products in aqueous environment. *Environ. Pollut.* **158**, 3482–3489 (2010)
30. Botta, C., Labille, J., Auffan, M., Borschneck, D., Miche, H., Cabié, M., Masion, A., Rose, J., Bottero, J.-Y.: TiO₂-based nanoparticles released in water from commercialized sunscreens in a life-cycle perspective: structures and quantities. *Environ. Pollut.* **159**(6), 1543–1550 (2011)

Phytotoxicity

► Toxicology: Plants and Nanoparticles

Piezoelectric Effect at Nanoscale

Zhong Lin Wang and Ying Liu
School of Materials Science and Engineering,
Georgia Institute of Technology, Atlanta, GA, USA

Synonyms

Nano-piezoelectricity

Definition

Piezoelectricity is a phenomenon of strain induced electric polarization in certain crystals, which can be used to create a mechanical action by applying an external voltage for sensors and actuators, or an mechanical straining can produce a voltage for energy conversion.

Piezoelectricity

Piezoelectricity was discovered by Jacques and Pierre Curie during their systematic studies of the effect of pressure on the generation of electrical charge by crystals in 1880. It is a reversible physical phenomenon, as there is the direct piezoelectric effect and the reverse piezoelectric effect.

Direct piezoelectric effect is defined as: *electric polarization produced by mechanical strain in crystals belonging to certain classes, the polarization being proportional to the strain and changing sign with it* [1].

Reverse (or inverse) piezoelectric effect is defined as: *a piezoelectric crystal becomes strained, when electrically polarized by an amount proportional to the polarizing field.*

Structures of Piezoelectric Crystals

The nature of piezoelectricity comes from the lack of centrosymmetry/inversion symmetry in the crystal (Fig. 1). 21 out of all 32 crystal classes lack inversion symmetry, and 20 of them exhibit piezoelectric effect [2]. The two most important classes include perovskite structure, which occurs in ceramics such as Lead Zirconate Titanate (PZT) and Barium Titanate (BaTiO_3); and wurtzite structure, which occurs in semiconducting materials such as Zinc Oxide (ZnO), Gallium Nitride (GaN), Indium Nitride (InN), and Zinc Sulfide (ZnS).

Perovskite Structure

The perovskite structure gets its name from the mineral CaTiO_3 and is adopted by many oxides that have the chemical formula of ABO_3 [3]. The most widely used piezoelectric ceramics, including Barium Titanate (BaTiO_3) and Lead Zirconate Titanate (PZT) are formed in this structure.

As shown in Fig. 2a, in the unit cell of perovskite compound, type “A” atom is in the cube corner positions (0, 0, 0), type “B” atom is in the body center position (1/2, 1/2, 1/2) and oxygen atoms is in the face centered positions (1/2, 1/2, 0). However, in real cases, since ion sizes hardly meet the stringent requirement for stability, most perovskite structures are distorted and lose their cubic symmetry, and commonly shift into orthorhombic or tetragonal phases [4]. When “B” atoms are displaced in such distortion, as shown in Fig. 2b, the center for negative charge and positive charge in the crystal no longer coincide with each other, and results in an electric dipole. This is where the piezoelectricity comes from.

Since the piezoelectricity originates from rather randomly oriented distortion, it is much more complicated and is correlated with the concept of ferroelectricity. In most cases, perovskite ceramics forms polycrystalline structures with multiple electric “domains.” Each domain has a size in the order of 10–100 nm, and even though each individual domain is piezoelectric itself, the whole ceramic does not show macroscopic dipole

without poling, as they have differently oriented electric dipoles. Electrical alignment of the dipoles, or electric poling as stated in most cases, are required to activate the piezoelectric effect in the ceramics. Similar to ferromagnetic materials, a characteristic temperature, the Curie temperature (T_C) exists in perovskite ceramics, above which they lose their electric polarization induced in the poling process.

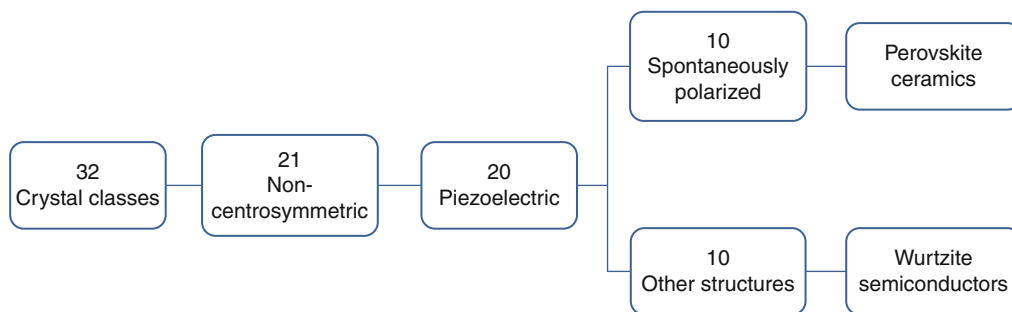
There are multiple methods for fabricating perovskite ceramic nanostructures [5]. The top-down methods include electron-beam assisted fabrication, nanoprinting, and self-assembling, and they provide high-precision positioning and size control. However, top-down methods are limited in resolution, and prone to processing damage. In comparison, the bottom-up approaches allow much smaller feature sizes but suffer from poor registration and thus require substrate patterning that in turn involves top-down processing. For one dimensional nanostructure, solution-phase decomposition and hydrothermal chemical synthesis [6] can result in single crystalline nanorods and nanowires, sol-gel method associated with nanochannel templates or electrospinning can produce polycrystalline nanowires, and wetting of the porous templates can produce nanotubes.

In all of the conventional piezoelectric materials, PZT has the best performance and the widest application. Its piezoelectric coefficient d_{33} can be as high as 400 pC/N, and can be higher than 600 pC/N when doped with lanthanum [2, 7]. However, since it is a lead containing chemical, it is not environmentally and biologically friendly without proper packaging, and its applications in the nanoscale region are thus restricted, as such applications are often related to biomedical or environmental issues.

Wurtzite Structure (ZnO, GaN)

The wurtzite crystal structure is named after the mineral wurtzite, and is the crystal structure for various binary compounds. Wurtzite structure occurs in most piezoelectric semiconductors, including Zinc Oxide (ZnO), Gallium Nitride (GaN), Cadmium Sulfide (CdS), Silicon Carbide (SiC) and so on, and among them, ZnO is most widely used in one dimensional nanostructure, and the key material for nanogenerators, piezotronics and piezophotonics, as will be discussed later.

As shown in Fig. 3, in a Wurtzite structure, each of the two individual atom types forms a sublattice which



Piezoelectric Effect at Nanoscale, Fig. 1 Scheme of piezoelectric crystal classes

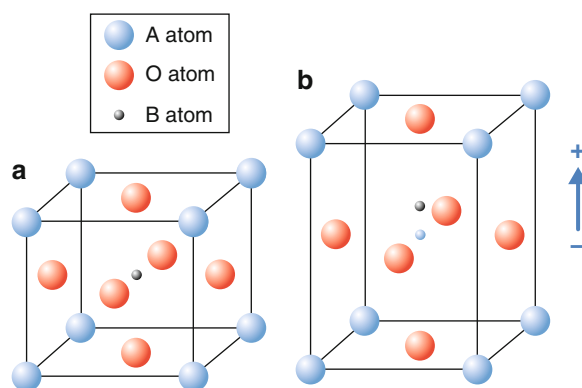
is hexagonal close pack (HCP) type, and they form tetrahedral coordination with each other [8]. Distinct from Perovskite structure, Wurtzite structure crystals do not need distortion or external poling to break the centrosymmetry. In single crystal Wurtzite materials, there are strong and uniform dipoles along the c axis, and the dipole direction is determined by the stacking order of the anions and cations, and more specifically, $+c$ for from cation to anion and $-c$ for from anion to cation. The orientation and polarity determines the direction of the piezoelectricity in the crystal. In ZnO, the interaction of the polar charges at the surface results in the growth of a wide range of unique nanostructures, such as nanobelts, nanosprings, nanorings, and nanohelices.

The synthesis of ZnO nanostructures is easy and convenient. To fabricate large scale nanowire arrays, chemical hydrothermal at 80°C can be carried out on any substrate, while physical vapor liquid solid (VLD) or physical or chemical vapor solid (PVD or CVD) growth can give results with higher aspect ratio and higher density on substrate with proper lattice match.

As will be discussed in latter sections, one dimensional wurtzite nanostructures, especially ZnO nanowires and nanobelts, promotes the new concept of “piezotronics,” which is the coupling between piezoelectricity and semiconductor transport property; and “piezo-phototronics,” which is the result of three-way coupling of piezoelectricity, photonic excitation, and semiconductor transport property.

Piezoelectric Polymer

Piezoelectricity also occurs in semicrystalline polymer materials. The most famous and widely used of them are polyvinylidene fluoride (PVDF) and its copolymers [9].

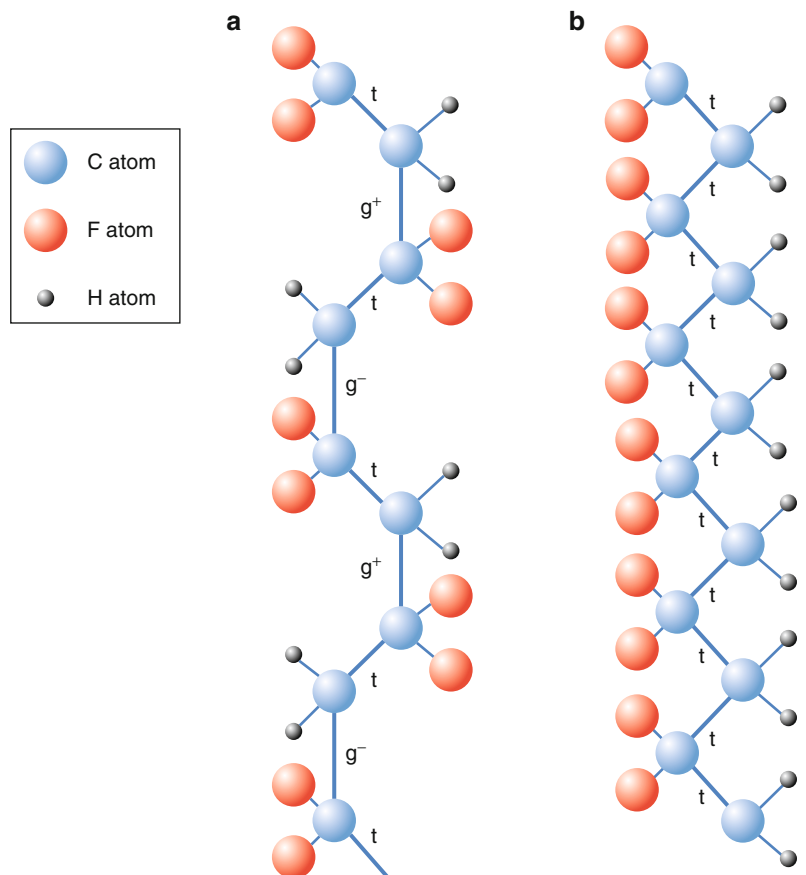
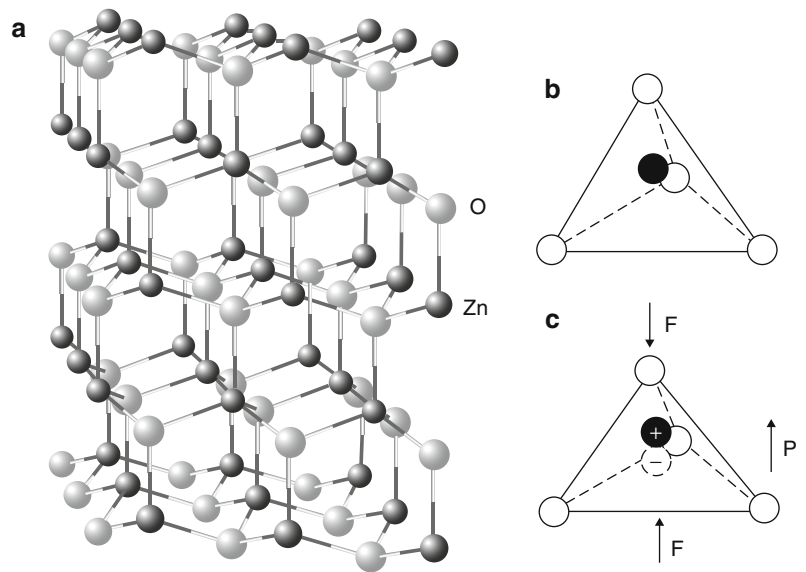


Piezoelectric Effect at Nanoscale, Fig. 2 (a) Ideal perovskite structure; (b) perovskite structure with distortion and polarization along c -axis

In PVDF, hydrogen atoms are positively charged and fluoride atoms are negatively charged with respect to the carbon atoms, so that each monomer has an inherent dipole moment. As shown in Fig. 4, PVDF has several conformations: α (tg^+tg^-), β (all trans), and γ ($\text{tttg}^+\text{tttg}^-$). In α conformation, the dipoles along the chain cancel each other, so that the overall piezoelectric effect is zero, and this conformation is the non-piezoelectric conformation. In the β conformation, the alignment of all its dipoles is in the same direction normal to the chain axis. In this way, once strain is applied on the β phased polymer, and the chain is either stretched or compressed, the dipole strength will change and extra net charge will be induced, so that this conformation is piezoelectric. Since the nature of PVDF is ferroelectric, electric poling is also needed for the dipoles on the polymer chains to align so that the polymer will have piezoelectric response.

The crystallinity of pure PVDF is usually 50~60%. In its copolymer with TrFE and TFE, the crystallinity can be increased to nearly 100% after annealing, and

Piezoelectric Effect at Nanoscale, Fig. 3 ZnO with wurtzite structure and illustration for the formation of dipoles ([8], Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission)



Piezoelectric Effect at Nanoscale, Fig. 4 Illustration of the two main conformation of PVDF. (a) α conformation; (b) β conformation

the portion of β conformation is also larger, and the piezoelectric response is thus greater. The piezoelectric coefficient in PVDF can be as high as a negative value of -33 pm/V, and that of PVDF-TrFE can be as high as -38 pm/V.

The synthesis of piezoelectric polymer nanofibers can be easily carried out through the process of electrospinning.

Theory of Piezoelectricity in One Dimensional Nanostructures

In theoretical concern, piezoelectricity refers to the linear coupling of electricity and mechanics [10]. Considering electrical behavior, we have Maxwell's equation: $D = \epsilon E$, where D is electric displacement, ϵ is permittivity, and E is electric field. Considering mechanical behavior, we have Hooke's law: $S = sT$, where S is strain, s is compliance, and T is stress. To describe piezoelectricity, the coupled equations are given as,

$$\begin{aligned} \{S\} &= [s^E]\{T\} + [d^T]\{E\} \\ \{D\} &= [d]\{T\} + [\epsilon^T]\{E\} \end{aligned}$$

Where $[d]$ is the direct piezoelectric effect matrix and its transpose $[d^T]$ is the inverse piezoelectric effect matrix, and their specific form is related to the point group to which the crystal belongs. The most frequently used parameter is the piezoelectric coefficient d_{33} , with expression $d_{33} = (\partial D_3 / \partial T_3)^E = (\partial S_3 / \partial E_3)^T$, so the unit of d_{33} is either C/N or m/V, and generally in materials d_{33} is often in the order of 10^{-10} to 10^{-12} C/N or m/V. When normal stress is applied along the polarization axis and on the surface charge is collected, d_{33} can be simplified as the volume change when subject to an electric field, or the electric polarization when subject to stress. A similar parameter d_{31} is applied when stress is applied at right angle with the polarization axis.

Theoretical approaches have been applied on calculating piezoelectric potential distribution in piezoelectric nanostructures. A continuum model was built for calculating the electrostatic potential in a laterally bent nanowire (NW) [11].

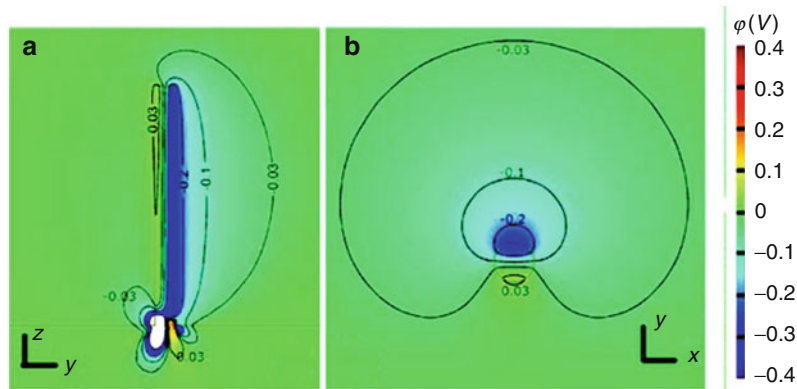
To derive the relationship between the potential distribution in a NW and the dimensionality of the

NW and magnitude of the force applied, four sets of equations for static piezoelectric material are first listed:

1. Mechanical equilibrium equation, $\nabla \cdot \sigma = \bar{f}_e^{(b)} = 0$, where σ is the stress tensor.
2. Constitutive equations,
$$\begin{cases} \sigma_p = c_{pq} \epsilon_q - e_{kp} E_k \\ D_i = e_{iq} \epsilon_q + \kappa_{ik} E_k \end{cases}$$
 where c_{pq} is the linear elastic constant, e_{kp} is the linear piezoelectric coefficient, and κ_{ik} is the dielectric constant.
3. Geometrical compatibility equation, $e_{ilm} e_{jpq} (\partial^2 \epsilon_{mp} / \partial x_l \partial x_q) = 0$, where e_{ilm} and e_{jpq} are Levi-Civita antisymmetric tensors. Small deformation of the NW is also assumed for simplicity of the derivation.
4. Gaussian equation of electric field, $\nabla \cdot \vec{D} = \rho_e^{(b)} = 0$, under the assumption that the piezoelectric NW has no conductivity.

For analytical analysis, a perturbation theory has been developed based on these equations. Under different orders of approximation, the four equations correspond to different extends of decoupling and coupling between the electric field and mechanical deformation: the zeroth order solution stands for purely mechanical deformation without piezoelectricity; the first order is the result of direct piezoelectric effect, where stress/strain generates an electric field in the NW; and the second order shows the first coupling of the piezoelectric field to the strain in the material. For the calculation of piezoelectric potential in an insulating NW, the first order approximation is sufficient, and it gives analytical results less than 5% away from finite element method (FEM) numerical calculation for potential distribution in a typical NW with diameter $d = 50$ nm and length $l = 600$ nm at a lateral bending force of $f = 80$ nN.

For further investigation on the interaction of piezoelectric potential with semiconductor transportation properties, charge carrier concentration can be substituted into the Gaussian equation of electric field, and the equation is modified as: $\nabla \cdot \vec{D} = (\partial / \partial x_i) (e_{iq} \epsilon_q + \kappa_{ik} E_k) = \rho_e^{(b)} = ep - en + eN_D^+ - eN_A^-$, where p is the hole concentration in the valance band, n is the electron concentration in the conduction band, N_D^+ is the ionized donor concentration, and N_A^- is the ionized acceptor concentration. For n-type nanowires, it can be assumed that $p = N_A^- = 0$, and vice versa for p-type. Considering band structure and charge carrier



Piezoelectric Effect at Nanoscale, Fig. 5 Plot of calculated piezoelectric potential φ . Regarding a ZnO NW with diameter $d = 50$ nm and length $l = 600$ nm at a lateral bending force of $f = 80$ nN for the NW are the same as in Fig. 4 except that it is n-type semiconducting with $N_D = 1 \times 10^{17} \text{ cm}^{-3}$ (Adapted with

permission from Ref. 36 of [12] “Gao, Y.F., Wang, Z.L.: Equilibrium potential of free charge carriers in a bent piezoelectric semiconductive nanowire. *Nano Lett.* **9**, 1103–1110 (2009).” Copyright 2009 American Chemical Society)

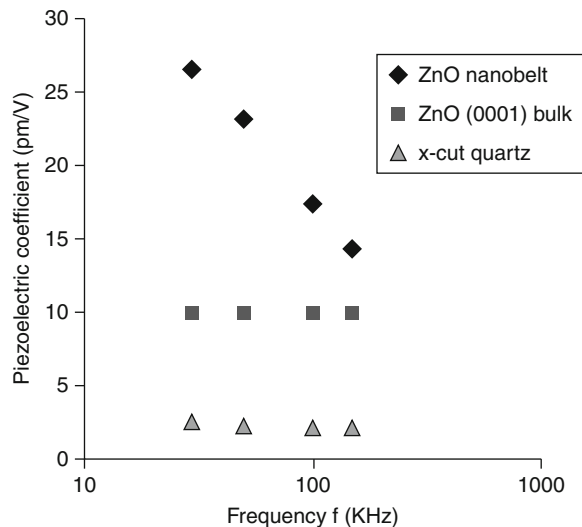
distribution equations, there can be numerical solution of the modified equations. As shown in Fig. 5, FEM solutions based on these equations is given for a typical ZnO NW with moderate doping [Ref. 36 of [12]].

Further investigation is also carried out considering the contact of semiconducting NW with metals [13]. When a ZnO NW is under nonuniform deformation, the local piezoelectric potential distribution at different positions shows significant effect on and distinct trend of variation in the charge carrier transport characteristic. This can help understanding the characteristics of piezotronic devices based on NW of wurtzite materials by controlling local contact position and contact size.

Effect of Size on Piezoelectricity

Size Effect in Wurtzite Nanocrystals

At nanoscale, many interesting effects occur and this can affect physical properties of materials, including piezoelectric properties. In nanowires (NW) and nanobelts (NB), crystal defects such as dislocation and vacancies will concentrate and eliminate on the surface, and perfect single crystallinity is more easily achieved at this scale of dimension. As a result, during piezoelectric measurement of individual Zinc Oxide NB with dimension of tens of nanometers in thickness, hundreds of nanometers in width and tens of micrometers in length, the effective piezoelectric coefficient d_{33}



Piezoelectric Effect at Nanoscale, Fig. 6 Frequency dependence of piezoelectric coefficient of ZnO nanobelts, ZnO bulk material, and x-cut quartz (Adapted with permission from Ref. 41 of [12] “Zhao, M.H., Wang, Z.L., Mao, S. X.: Piezoelectric characterization of individual zinc oxide nanobelt probed by piezoresponse force microscope. *Nano Lett.* **4**, 587–590 (2004).” Copyright 2004 American Chemical Society)

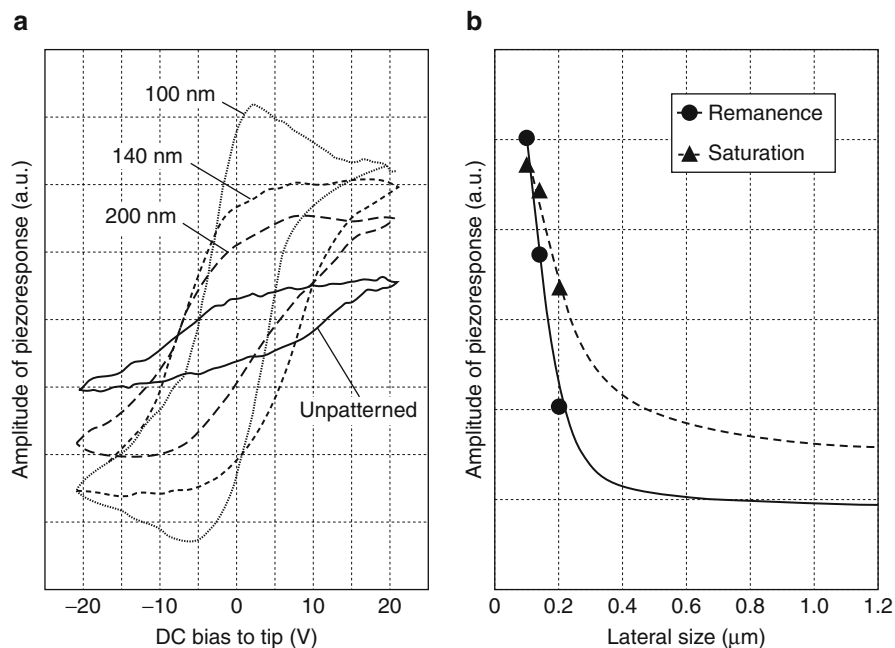
is revealed to be frequency dependent and varies from 14.3 to 26.7 pm/V, which is about one time larger than that in bulk ZnO, around 10 pm/V, as shown in Fig. 6.

Size Effect in Perovskite Nanostructures

Since in perovskite materials, piezoelectric performance is tightly related to the concept of domain, the

Piezoelectric Effect at Nanoscale,

Fig. 7 Ferroelectric loop showing increased piezoelectric response with shrinking size. (a) Loops on patterns with different lateral sizes and on the unpatterned film; and (b) remnant polarization and saturation as a function of lateral size [14]



case of size effect is much more complicated, there are multiple factors to take into consideration. Intrinsically, at nanoscale, competition exists between thermal vibrations and the correlation energy which aligns the dipoles, but greater surface area also helps eliminate defects. Extrinsically, grain boundaries and other microstructural factor can also have various influences.

Between 10 and 1,000 nm, which is bigger than or equal to the scale of a typical single domain and similar to the state of art piezoelectric nanodevices, controversial size effects can occur in differently processed materials regarding the sequence of patterning and crystallization of the nanostructure. This controversy is caused by extrinsic factors related with post annealing. In structures that were patterned with high-energy ion etching after crystallization, which leads to higher concentration of oxygen and lead vacancies, several reports reveal the inherent reduction in the piezoelectric coefficient in ultrathin PZT films. With post annealing, nonswitching of pinned domains help stabilize and enhance piezoelectric effect. In thin film studies, the observed increase of piezoelectric response amplitude can be as high as 300% when decreasing the feature size of patterns on epitaxial PZT thin film from 200 to 100 nm, as shown in Fig. 7. The crystalline nature of patterns might also have contributed to this difference: single crystalline tends to have better piezo performance

with smaller scale than polycrystalline. In template method, synthesized PZT NW enhanced piezoelectric response than the value in thin film was also discovered [5].

Under 10 nm, intrinsic factors have more influence on the piezoresponse. People have used phenomenological models and first principal/ab initial calculations to predict the behavior of such nanostructures, and their prediction is based on the competition and cancellation of thermal drift. In general, the results are shown as: (1) depolarizing fields exist near the surfaces to reduce polarization and (2) piezoelectric state becomes unstable below certain sizes, which is estimated to be 7–10 nm along each of the three directions of well-known materials such as BaTiO₃. Microscopy study in this scale has been done and verified these predictions to some extent. This gives us important information that fabricating piezoelectric devices under the scale of 10 nm using perovskite structure materials will hardly result in better performance.

Applications of Nanoscale Piezoelectric Properties

As mentioned in previous sections, for materials with perovskite structure, the enhancement of performance at the nanoscale is not always significant, and related

application is very limited. For materials with wurtzite structures however, especially for ZnO NW and NB, the interaction of piezoelectricity with semiconducting transportation process brings about new phenomena and draws great research interests, into the new concept of “piezotronics” and even “piezo-phototronics.” These applications includes piezoelectric field-effect transistor, piezoelectric diode, piezoelectricity based nanosensors, piezoelectricity enhanced photodetector, and most importantly, the piezoelectric nanogenerators.

Piezoelectric Nanogenerator

The future of nanotechnology research is likely to focus on the areas of integrating individual nanodevices into a nanosystem that acts like living specie with sensing, communicating, controlling, and responding. A nanosystem requires a nanopower source to make the entire package extremely small yet with high performance. The goal is to make self-powered nanosystem that can operate wirelessly, independently, and sustainably. Harvesting energy from the environment is a choice for powering nanosystems especially for biomedical applications. Since mechanical energy is a conventional form of energy that exists in our living environment, it is essential to explore innovative nanotechnologies for converting mechanical energy (such as body movement and muscle stretching), vibration energy (such as acoustic/ultrasonic wave), and hydraulic energy (such as body fluid and blood flow) into electric energy that will be used to power nanodevices without using battery. This is a key step toward *self-powered nanosystems* [15].

Dr. Z.L. Wang’s group in Georgia Institute of Technology is the first to invent an innovative approach for converting mechanical energy into electric energy by piezoelectric zinc oxide nanowire arrays [16]. The operation mechanism of the nanogenerator (NG) relies on the piezoelectric potential created by an external strain; a dynamic straining of the nanowire results in a transient flow of the electrons in the external load due to the driving force of the piezopotential. The nanogenerator technology has been developed from fundamental science, to engineering integration, and to technological scale-up. As today, a gentle straining can output 1–3 V at an instant output power of $\sim 2 \mu\text{W}$ from an integrated nanogenerator, using which a self-powered nanosensor has been demonstrated [6]. This

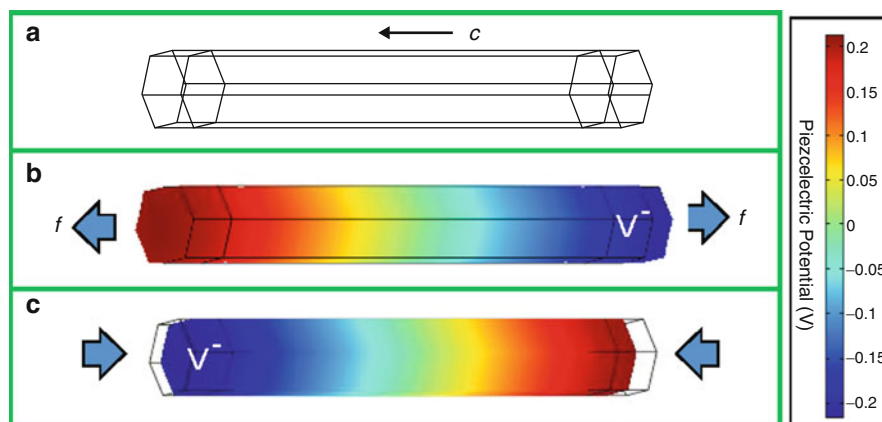
technology has the potential applications for powering MEMS/NEMS that requires a power of the microwatt to milliwatt range.

The fundamental of nanogenerator relies on the presence of the piezoelectric potential (piezopotential) generated in a nanowire/nanobelt via dynamic straining. As demonstrated in previous sections, ZnO has the wurtzite structure, in which the tetrahedral coordinated O^{2-} and Zn^{2+} are stacked layer by layer along the c -axis. The lack of central symmetry results in the piezoelectric property of ZnO, which is vital for the mechanic–electric energy conversion with the ZnO-based nanogenerator.

The cations and anions are tetrahedral-coordinated in the wurtzite-structured ZnO. At the strain free status, the charge-center of the cations and that of anions coincide with each other. When an external strain is applied, the structure is deformed so that the charge-centers for cations and anions separate and result in an electric dipole. Because the ionic charges are not free to move and the intrinsic free charge carriers can only partially screen them if the doping level is low, the piezoelectric field is preserved as long as the NW is strained. The potential created by the polar ions is called *piezoelectric potential*, or *piezopotential*.

The presence of piezopotential is the fundamental physical basis of the nanogenerators and piezotronics (Fig. 8). When a strained crystal is connected to an external load, the electrons in the circuit are driven to flow in to partially screen the piezopotential, which is the energy conversion process. Therefore, *the principle of the nanogenerator is the transient flow of electrons in external load as driven by the piezopotential created by dynamic straining*.

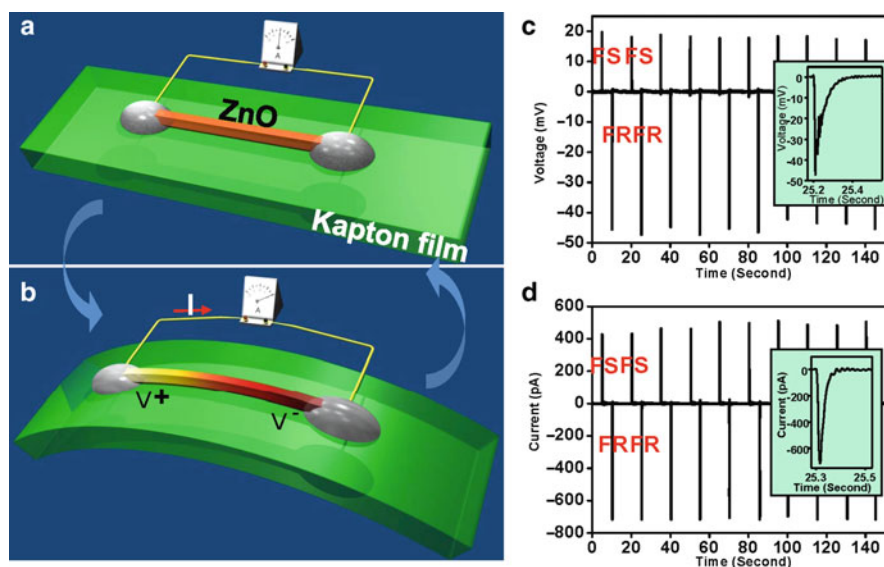
Piezopotential-driven transient flow of electrons in an external load is the principle of the nanogenerator. Figure 9 presents a single wire generator (SWG) using a laterally packaged ZnO wire. In brief, a single ZnO wire was placed laterally on the flexible polyimide film. Silver paste can be used to fix two ends of the ZnO wire to the substrate and link the wire to the external measuring instrument through metallic wires. Electric measurement shows that a functioning SWG usually has rectifying I – V characteristic, indicating the existence of the Schottky contact at least at one end. Before the SWG is deformed, there is no measurable potential drop from the ZnO wire. When the substrate is bent, both the substrate and the ZnO wire are under strain. Because the thickness of the substrate is much larger than the dimension of the ZnO wire, the ZnO wire is solely



Piezoelectric Effect at Nanoscale, Fig. 8 Numerical simulation of the piezopotential distribution in an insulating ZnO nanowire (a); The piezopotential distribution and the deformed shape of a ZnO nanowire grown along c -axis under a stretching force of 85 nN (b); or a compressing force of 85 nN (c) (Adapted

with permission from Ref. 13 of [12] “Gao, Z.Y., Zhou, J., Gu, Y.D., Fei, P., Hao, Y., Bao, G., Wang, Z.L.: Effects of piezoelectric potential on the transport characteristics of metal-ZnO nanowire-metal field effect transistor. *J. Appl. Phys.* **105**, 113707 (2009).” Copyright 2009, American Institute of Physics)

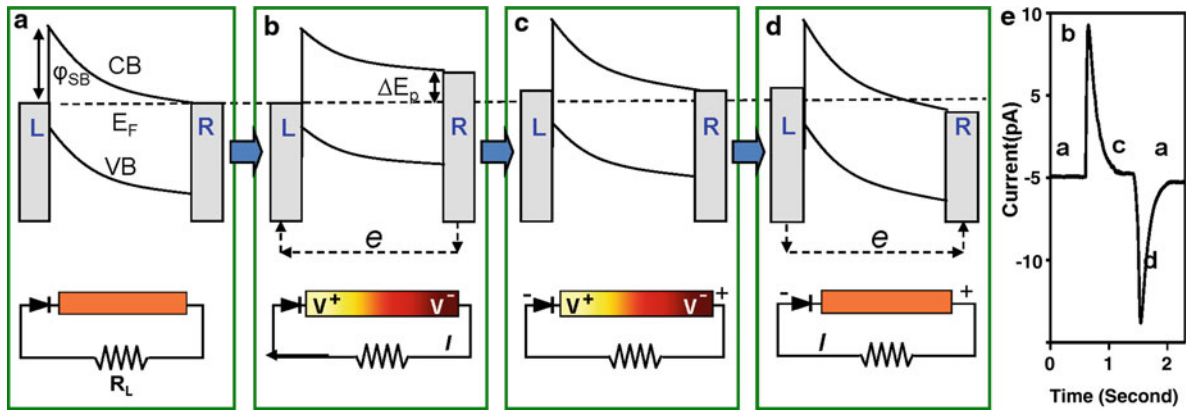
Piezoelectric Effect at Nanoscale, Fig. 9 Design and power output from a single wire generator. (a) and (b) schematic diagram of single wire generator before and after stretching. Open-circuit voltage (c) and short-circuit current (d) of the single wire generator, which was cyclically fast stretched and fast released (Adapted from Ref. 16 of [12] “Yang, R. S., Qin, Y., Dai, L.M., Wang, Z.L.: Power generation with laterally packaged piezoelectric fine wires. *Nat. Nanotechnol.* **4**, 34–39 (2009)”)



experiencing tensile strain when the bend is concave downward, as shown in Fig. 9b. As discussed earlier, tensile strain will result in piezopotential, such that the $+c$ -axis side gains positive potential (light yellow) and $-c$ -axis side gains negative potential (dark red). The potential difference between the two ends can then be measured either as open-circuit voltage (Fig. 9c) or short-circuit current (Fig. 9d).

The working mechanism of the SWG can be understood with the energy band diagram in Fig. 10. The

Schottky contact at least at one end is necessary for the energy generation. It is assumed the Schottky contact is at the $+c$ -axis side (left side in Fig. 10a), and the energy generation follows in a similar way when the Schottky contact is on the other side or both sides. The entire wire is in equilibrium state without any strain and power output (Fig. 10a, e). When the wire is stretched, the tensile strain induces polarization of atoms, as well as piezoelectric field, in the crystal. Consequently, the $+c$ -axis side holds positive potential



Piezoelectric Effect at Nanoscale, Fig. 10 Current generation mechanism explained with energy band diagram of the cyclically stretched NW. Energy band diagram of the ZnO NW in equilibrium and free condition (a); non-equilibrium and tensile strained condition (b); re-reached equilibrium and tensile strained condition (c); and nonequilibrium and free condition (d). After (d), the NW will reach equilibrium as in (a) after charges flow. E_F , CB, VB, Φ_{SB} , and ΔE_p indicate Fermi level

of the electrode, conduction band, and valence band of ZnO, Schottky barrier height, and piezopotential difference respectively. (e) Short-circuit current measurement in which the labels of a, b, c, and d indicate the corresponding process shown in subparts a–d (Adapted from Ref. 16 of [12] “Yang, R.S., Qin, Y., Dai, L.M., Wang, Z.L.: Power generation with laterally packaged piezoelectric fine wires. *Nat. Nanotechnol.* **4**, 34–39 (2009)”)

and the $-c$ -axis side holds negative potential. The potential difference is ΔE_p . The conduction band edge, as well as the Fermi level of the metal of the right side electrode, rises up for the same amount of ΔE_p . The electrons in the external circuit should flow from the right-hand side to the left-hand side to compensate the energy difference, which will generate the first output signal if the flow rate is sufficiently large, as indicated in Fig. 10b and e. The electrons cannot flow across the interface due to the presence of a Schottky barrier ϕ_{SB} . The accumulation of the electrons at the interface will partially screen the piezopotential built in the wire. When the electrons and the piezoelectric field reach equilibrium, the Fermi energy on both sides is on the same level and there is no more current flow (Fig. 10c, e). When the SWG is fast released, there is no strain inside the NW. As a result, the polarization and piezoelectric field vanish as well and the equilibrium with the accumulated electrons is broken. The accumulated charge carriers flow back from the left-hand side to the right-hand side through the external circuit, producing the second output signal in the opposite direction, as shown in Fig. 10d, e.

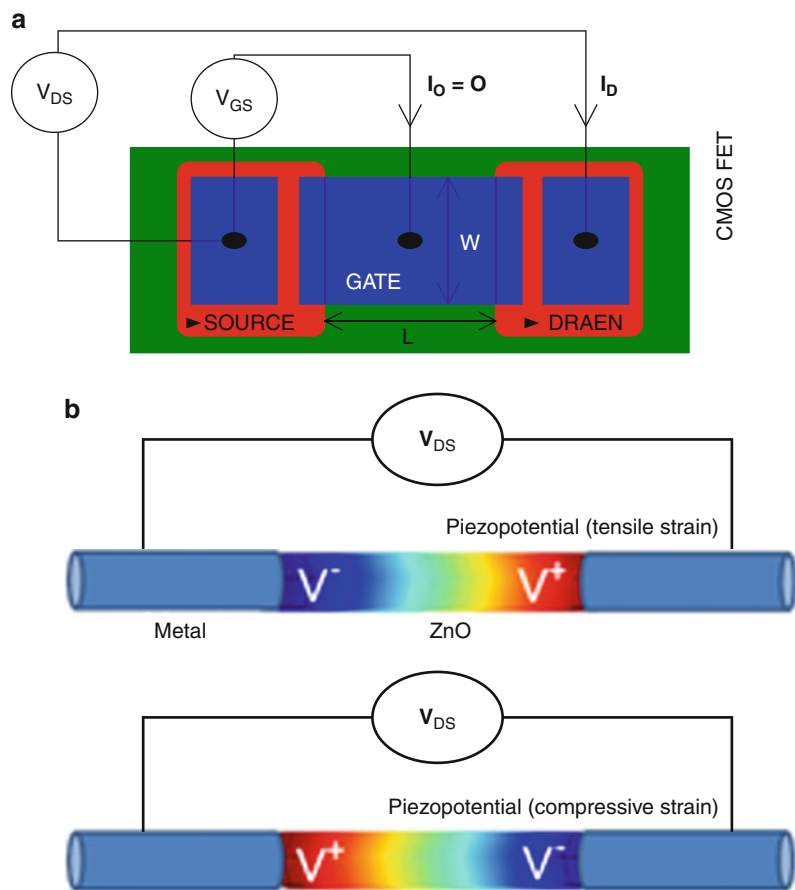
Integration of SWGs is a major step toward the practical applications. Taking advantage of the small size of the NW, multiple SWGs can be built on a single substrate such that the hamster can drive all SWGs simultaneously. Integration of up to four

SWGs in serial has been demonstrated with the open-circuit voltage of ~ 0.1 – 0.15 V. The output current can also be improved with multiple SWGs in parallel. The work on the human finger and living hamster clearly demonstrated the capability of the SWG to scavenge mechanical energy from the biosystem. It also confirms the feasibility of using the SWG for harvesting the energy created by regular and irregular motions.

Piezotronics

A most simple FET is a two ends bonded semiconductor wire, in which the two electric contacts at the ends are the source and drain, and the gate voltage can be applied either at the top of the wire through a gate electrode or at its bottom on the substrate [8, 13]. When a ZnO NW is strained axially along its length, the piezoelectric potential continuously drops from one side of the NW to the other, which means that the electron energy continuously increases from the one side to the other. Meanwhile, the Fermi level will be flat all over the NW when equilibrium is achieved, since there is no external electrical field. As a result, the effective barrier height and/or width of the electron energy barrier between ZnO and metal electrode will be raised at one side and lowered at the other side; thus, it has a nonsymmetric effect on the source and drain. This is the *piezotronic effect* [11].

Piezoelectric Effect at Nanoscale, Fig. 11 A comparison of a CMOS FET and a piezotronic FET. The key difference is that the externally applied gate voltage V_{GS} is replaced by the strain generated piezopotential, which tunes the charge transport insets are the enlarged view of the boxed area for one cycle of deformation. This is the core of piezotronics

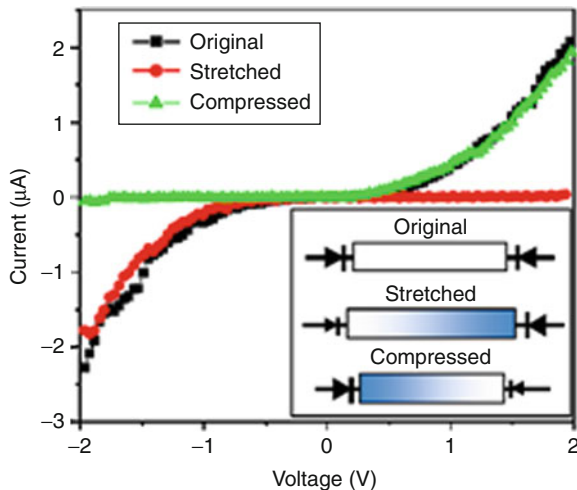


Once a strain is created in the semiconductor that also has piezoelectric property, a negative piezopotential at the semiconductor side effectively increases the local SB height, while a positive piezopotential reduces the barrier height. The polarity of the piezopotential is dictated by the direction of the c -axis for ZnO. The role played by the piezopotential is to effectively change the local contact characteristics through an internal field, thus, the charge carrier transport process is tuned/gated at the metal-semiconductor (M-S) contact. Figure 11 compares the piezopotential gated FET and the traditional CMOS FET, which are different only by replacing the gate voltage by the inner crystal piezopotential. With considering the change in piezopotential polarity by switching the strain from tensile to compressive, the local contact characteristics can be tuned and controlled by the magnitude of the strain and the sign of strain. This is the core of piezotronics.

The fundamental working principle of the p-n junction and the Schottky contact is that there is an effective

barrier that prevents the charge carriers of both sides from crossing. The height and width of the barrier are the characteristic of the device. In piezotronics, the role played by the piezopotential is to effectively change the width of p-n junction or height of Schottky barrier (SB) by piezoelectricity, which has applications in maximizing the performance of the solar cell, photon detector and LED by tuning the height of the local SB for achieving the optimum charge separation.

A simple piezotronic device is a polarity switchable diode that is made of a ZnO NW contacted with metal contacts at the two ends on an insulating polymer substrate. From the initial I-V curve measured from the device before applying a strain as shown in Fig. 12, the symmetric shape of the curve indicates that the SBs present at the two contacts are about equal heights. The equivalent circuit model of the device is a pair of back-to-back Schottky diodes, as illustrated in the inset in Fig. 12. Under tensile strain, the piezoelectric potential at the right-hand side of this NW was lower (denoted



Piezoelectric Effect at Nanoscale, Fig. 12 Piezotronic strain sensor/switch. Changes of transport characteristics of an Ag/ZnO-nanowire/Ag device from symmetric I - V characteristic (black) to asymmetric rectifying behavior when stretching (red) and compressing (green) the wire. The inset is the equivalent circuit models of the device in corresponding to the observed I - V curves, different sizes of diode symbol are used to illustrate the asymmetric Schottky contacts at the two ends of the nanowire. The blue side is the negative potential side, and the other side is positive side (Ref. 22 of [12], Adapted with permission from “Zhou, J., Gu, Y.D., Fei, P., Mai, W.J., Gao, Y.F., Yang, R.S., Bao, G., Wang, Z.L.: Flexible piezotronic strain sensor. *Nano Lett.* **8**, 3035–3040 (2008).” Copyright 2008 American Chemical Society)

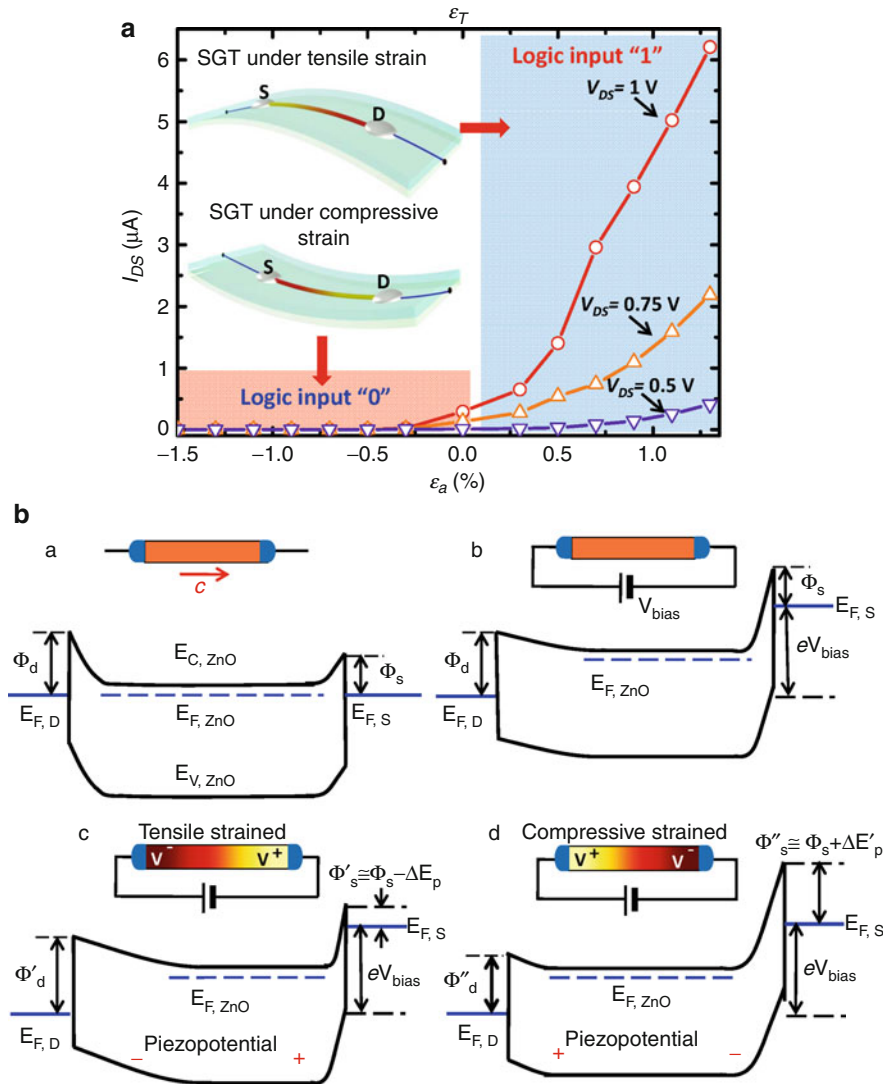
by blue color in the inset in Fig. 12), which raised the local barrier height (denoted by a large diode symbol in the inset). Since the positive piezoelectric potential was partially screened by free electrons, the SB height at the left-hand side remained almost unchanged. As a result, under positive bias voltage with the left-hand side positive, the current transport was determined by the reverse biased SB at the right-hand side. While under the reverse biased voltage with the right-hand side positive, the current transport depended on the reverse biased SB at the left-hand side, which had a much lower barrier height than the right-hand one. Experimentally, the device thus exhibited a rectifying behavior in the positive voltage region, and the I - V curve in the negative voltage region overlapped with that of the original curve without straining. By the same token, under compressive strain the device exhibited a rectifying behavior in the negative voltage region, and the I - V curve in the positive voltage region overlapped with that of the original curve without straining, as shown by the green line in Fig. 12.

Piezotronic Logic Operations

The existing semiconductor NW logic devices are based on electrically gated field effect transistors (FETs), which function as both the drivers and the active loads of the logic units by adjusting the conducting channel width [17]. Moreover, the currently existing logic units are “static” and are almost completely triggered or agitated by electric signals, while the “dynamic” movable mechanical actuation is carried out by another unit possibly made of different materials. It is highly desirable to integrate the static logic operations with the dynamic mechanical action into one unit utilizing a single material. The first piezoelectric triggered mechanical-electronic logic operation has been demonstrated using the piezotronic effect, through which the integrated mechanical actuation and electronic logic computation are achieved using only ZnO NWs.

A strain gated transistor (SGT) is made of a single ZnO NW with its two ends, which are the source and drain electrodes, being fixed by silver paste on a polymer substrate (Fig. 13a). Once the substrate is bent, a tensile/compressive strain is created in the NW since the mechanical behavior of the entire structure is determined by the substrate. Utilizing the piezopotential created inside the NW, the gate input for a NW SGT is an externally applied strain rather than an electrical signal. I_{DS} - V_{DS} characteristic for each single ZnO-NW SGT is obtained as a function of the strain created in the SGT (Fig. 13a) before further assembly into logic devices. A NW SGT is defined as forward biased, if the applied bias is connected to the drain electrode (Fig. 13a). For a SGT, the external mechanical perturbation induced strain (ϵ_g) acts as the gate input for controlling the “on”/“off” state of the NW SGT. The positive/negative strain is created when the NW is stretched/compressed. The SGT behaves in a similar way to an n -channel enhancement-mode MOSFET, apparently indicating the working principle of the SGT.

The working principle of a SGT is illustrated by the band structure of the device. A strain free ZnO NW may have Schottky contacts at the two ends with the source and drain electrodes but with different barrier heights of Φ_S and Φ_D , respectively (Fig. 13b-a). When the drain is forward biased, the quasi-Fermi levels at the source ($E_{F,S}$) and drain ($E_{F,D}$) are different by the value of eV_{bias} , where V_{bias} is the applied bias (Fig. 13b-b). An externally applied mechanical strain (ϵ_g) results in both band structure change and piezoelectric potential field in a ZnO NW. The former leads

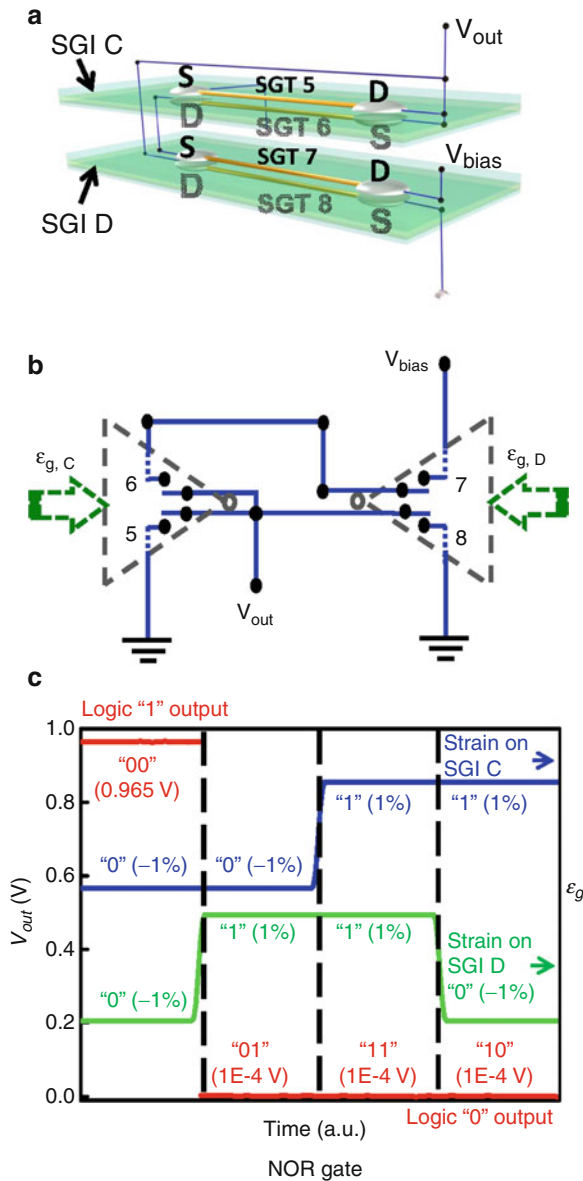


Piezoelectric Effect at Nanoscale, Fig. 13 Strain-gated transistor (SGT). (a) Current – strain ($I_{DS} - \epsilon_g$) transfer characteristic for a ZnO SGT device with strain sweeping from $\epsilon_g = -0.53$ – 1.31% at a step of 0.2% , where the V_{DS} bias values were 1, 0.75, and 0.5 V, respectively; (b) The band structures of the ZnO NW SGT under different conditions for illustrating the mechanism of SGT. The crystallographic c -axis of the nanowire directs from drain to source. (a) The band structure of a strain-free ZnO NW SGT at equilibrium with different barrier heights of Φ_S and Φ_D at the source and drain electrodes, respectively;

(b) The quasi-Fermi levels at the source ($E_{F,S}$) and drain ($E_{F,D}$) of the ZnO SGT are split by the applied bias voltage V_{bias} ; (c) With tensile strain applied, the SBH at the source side is reduced from Φ_S to $\Phi'_S \cong \Phi_S - \Delta E_p$. (d) With compressive strain applied, the SBH at the source side is raised from Φ_S to $\Phi''_S \cong \Phi_S + \Delta E'_p$ (Adapted with permission from Ref. 29 of [12] “Wu, W.Z., Wei, Y.G., Wang, Z.L.: Strain-gated piezotronic logic nanodevices. *Adv. Mater.* **22**, 4711–4715 (2010).” Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission)

to the piezoresistance effect, which is a non-polar and symmetric effect at the both source and drain contacts. Since ZnO is a polar structure along c -axis, straining in axial direction (c -axis) creates a polarization of cations and anions in the NW growth direction, resulting in a piezopotential drop from V^+ to V^- along the NW,

which produces an asymmetric effect on the changes in the SB heights (SBHs) at the drain and source electrodes. Under tensile strain, the SBH at the source side reduces from Φ_S to $\Phi'_S \cong \Phi_S - \Delta E_p$ (Fig. 13b-c), where ΔE_p denotes the effect from the locally created piezopotential and it is a function of the strain,



Piezoelectric Effect at Nanoscale, Fig. 14 Strain gated piezotronic logic NOR gate. (a) Schematic of the ZnO NW strain-gated NOR logic gate, which is composed of two SGIs, SGI C and SGI D. The strain input C for SGI C is defined in reference to the strain applied to SGT 5 and the strain input D for SGI D is defined in reference to the strain applied to SGT 8; (b) Layout for ZnO NW strain-gated NOR logic gate connecting two ZnO NW SGIs; (c) Logic operations and experimental truth table of the ZnO NW strain-gated NOR logic gate. Red line is the electrical output of the NOR gate. Blue and green lines represent the strain inputs applied on SGI C and SGI D, respectively (Adapted with permission from Ref. 29 of [12] “Wu, W.Z., Wei, Y.G., Wang, Z.L.: Strain-gated piezotronic logic nanodevices. *Adv. Mater.* **22**, 4711–4715 (2010).” Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission)

resulting in increased I_{DS} . For the compressively strained SGT, the sign of the piezopotential is reversed, thus the SBH at the source side is raised from Φ_S to $\Phi'_S \cong \Phi_S + \Delta E'_p$ (Fig. 13B-d), where $\Delta E'_p$ denotes the piezopotential effect on the SBH at source side, resulting in a large decrease in I_{DS} . Therefore, as the strain ϵ_g is swept from compressive to tensile regions, the I_{DS} current can be effectively turned from “on” to “off” while V_{DS} remains constant. This is the fundamental principle of the SGT.

Logic operations of NW strain-gated NOR gate have been realized by integrating SGTs, which are gated individually by the applied strains, according to corresponding layout rules (Figs. 14a–b). The output voltages of NOR gate versus the input gate strains are shown in Fig. 14c. It can also be seen that NW strain-gated NOR gate with active loads exhibit better overall performance, such as larger logic swing, compared to passive-load NOR gate. Using the SGTs as building blocks, other universal logic components such as inverters and NAND and XOR gates have been demonstrated for performing piezotronic logic calculations, which have the potential to be integrated with the current NEMS technology for achieving advanced and complex functional actions in nanorobotics, microfluidics, and micro/nanosystems.

Summary

With reduction in materials size, its mechanical robustness, elastic modulus, and toughness improve dramatically. It can easily go way beyond the nonlinear but still elastic mechanical deformation without causing plastic damage. The piezoelectric effect at nanoscale is expected to be largely enhanced, so that they can be effectively used for ultrasensitive sensors/actuators and energy converters.

The two most recent important advances in using the nano-scale piezoelectric effect have been introduced here. One is the nanogenerator that converts mechanical energy into electricity utilizing tiny physical motion. Besides its potential for large-scale applications such as harvesting ocean wave and wind energy, it finds application in driving wireless nanodevices and nanosystems for medical science, environmental/infrastructure monitoring, defense technology, and even personal electronics. It is a key step for building self-powered nanotechnology.

The second important concept in science is the piezotronics. Electronics fabricated by using inner-crystal piezopotential as a “gate” voltage to tune/control the charge transport behavior is named *piezotronics*. *Piezo-phototronic effect* is a result of three-way coupling among piezoelectricity, photonic excitation and semiconductor transport, which allows tuning and controlling of electro-optical processes by strain induced piezopotential. These designs are different from the CMOS field effect transistor by replacing the externally applied gate voltage using a crystal generated inner potential (piezopotential) for controlling the charge transport. It is a transistor controlled by a force/strain. It will find applications in human-computer interfacing, mechanical triggered electronics, nanomachine, electronic signature, medical diagnostics, and sensor.

Cross-References

- ▶ [Ab Initio DFT Simulations of Nanostructures](#)
- ▶ [AFM Probes](#)
- ▶ [Biosensors](#)
- ▶ [Carbon Nanotube-Metal Contact](#)
- ▶ [CMOS MEMS biosensors](#)
- ▶ [Electrospinning](#)
- ▶ [Finite Element Methods for Computational Nano-Optics](#)
- ▶ [Flexible Electronics](#)
- ▶ [Hybrid Solar Cells](#)
- ▶ [Nanorobotics](#)
- ▶ [Nanomaterials for Excitonic Solar Cells](#)
- ▶ [Nanostructures for Energy](#)
- ▶ [Nanotechnology](#)
- ▶ [NEMS Piezoelectric Switches](#)
- ▶ [Piezoelectric MEMS Switch](#)
- ▶ [Sol-gel Method](#)

References

1. Cady, W.G.: Piezoelectricity: An Introduction to the Theory and Applications of Electromechanical Phenomena in Crystals. Dover, New York (1964)
2. Haertling, G.H.: Ferroelectric ceramics: history and technology. *J. Am. Ceram. Soc.* **82**(4), 797–818 (1999)
3. Bhalla, A.S., Guo, R., Roy, R.: The perovskite structure – a review of its role in ceramic science and technology. *Mat. Res. Innovat.* **4**, 3–26 (2000)
4. Johnsson, M., Lemmens, P.: Crystallography and chemistry of perovskites. In: Kronmüller, H. (ed.) *Handbook of Magnetism and Advanced Magnetic Media*. Wiley, New York (2006)

5. Gruverman, A., Kholkin, A., Ramesh, R.: Nanoscale ferroelectrics: processing, characterization and future trends. *Rep. Prog. Phys.* **69**, 2443–2474 (2006)
6. Xu, S., Hansen, B.J., Wang, Z.L.: Piezoelectric-nanowire-enabled power source for driving wireless microelectronics. *Nat. Commun.* **1**, 93 (2010)
7. Turner, R.C., Fuierer, P.A., Newnham, R.E., Shrout, T.R.: Materials for high temperature acoustic and vibration sensors: a review. *Appl. Acoust.* **41**, 299–324 (1994)
8. Wang, Z.L.: Nanopiezotronics. *Adv. Mater.* **19**, 889–892 (2007)
9. Murayama, N., Nakamura, K., Obara, H., Segawa, M.: The strong piezoelectricity in polyvinylidene fluoride (PVDF). *Ultrasonics* **14**, 15–24 (1976)
10. Yang, J.S.: Introduction to the Theory of Piezoelectricity. Springer, New York (2005)
11. Wang, Z.L.: Towards Self-Powered Nanosystems: From Nanogenerators to Nanopiezotronics. *Adv. Funct. Mater.* **18**, 3553–3567 (2008)
12. Wang, Z.L.: Piezopotential gated nanowire devices: piezotronics and piezo-phototronics. *Nano Today* **5**(6), 540–552 (2010)
13. Yan, Z., Hu, Y.F., Xiang, S., Wang, Z.L.: Effects of piezopotential spatial distribution on local contact dictated transport property of ZnO micro/nanowires. *Appl. Phys. Lett.* **97**, 033509 (2010)
14. Bühlmann, S., Dwir, B., Baborowski, J., Murali, P.: Size effect in mesoscopic epitaxial ferroelectric structures: increase of piezoelectric response with decreasing feature size. *Appl. Phys. Lett.* **80**, 3195–3197 (2002)
15. Wang, Z.L.: Self-powering nanotech. *Sci. Am.* **298**, 82–87 (2008)
16. Wang, Z.L., Song, J.H.: Piezoelectric nanogenerators based on zinc oxide nanowire arrays. *Science* **312**, 242–246 (2006)
17. Thorsen, T., Maerkl, S.J., Quake, S.R.: Microfluidic large-scale integration. *Science* **298**, 580–584 (2002)

Piezoelectric MEMS Switches

R. G. Polcawich¹, J. S. Pulskamp¹ and R. M. Proie²

¹US Army Research Laboratory RDRL-SER-L, Adelphi, MD, USA

²US Army Research Laboratory RDRL-SER-E, Adelphi, MD, USA

Synonyms

[AIN](#); [Aluminum nitride](#); [Lead zirconate titanate](#); [PZT](#); [Relay](#); [Relay logic](#)

Definition

A piezoelectric switch is an electromechanical relay or switch relying on piezoelectric actuation to provide

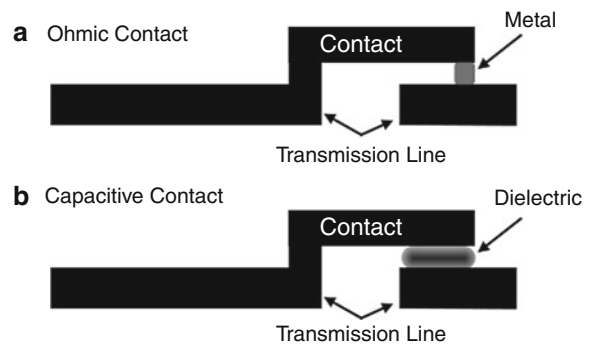
the force necessary to close or produce a gap between electrical contacts. In practical use, these switches can be designed for systems operating with carrier frequencies from DC to tens or hundreds of gigahertz.

Overview

The routing of electrical signals is controlled by switches of various types. For mechanical switches, a number of transduction techniques have been utilized including electrostatic, electromagnetic, thermomechanical, and piezoelectric. Fundamental to most radio frequency (RF) circuits, a switch is used not only to control the path of electrical circuits but also the phase and timing of the circuits. The continuous miniaturization of military radar and communication systems requires development of smaller, more energy-efficient switches for continuous control of a wide variety of electronic signals [1]. Additionally, switches are required for ever increasing frequencies, making existing solid state technologies less attractive to the system designer because of larger power losses and decreases in linearity as the frequency increases [2]. In short, microelectromechanical systems (MEMS) switches provide a very low loss, low power, highly linear, with respect to input power, alternative to existing solid state switch technologies. Moreover, they are significantly smaller than macro-switch technologies. Unlike electronic switches, PIN diodes, and field effect transistors (FETs), a well-designed MEMS switch is extremely linear with respect to input power and does not create distortion, such as harmonics or intermodulation products, in the electrical signal [2]. Next generation missile seeker radar, smart cellular phone technologies, and communication-on-the-move systems require highly linear, low-power, broadband RF switches as a fundamental building block [1, 3].

History and Brief Background of MEMS Switch Development

The development of MEMS switches began with the effort of Petersen at IMB in 1979 [4] and then continued with Hughes Research Laboratories on micromachined microwave actuators (MIMACs) that



Piezoelectric MEMS Switches, Fig. 1 Schematic illustration of (a) ohmic and (b) capacitive switch contacts

produced the first switch capable of operating up to 50 GHz [5]. Following this research, a concentrated effort began in developing switches in both series and shunt configurations for a wide variety of applications, including lower frequency commercial cell phones and high-frequency military radar and communication systems [1–3].

All MEMS switches have a few basic components for their operation. This includes the signal line, switch contact(s), and actuator(s). The signal line can be designed for use with signals ranging from DC to sub-THz. For low-frequency operation, a simple electrical trace commonly comprised of gold or other conductive metal can be utilized. For higher frequency operation, the signal line requires greater design consideration. For MEMS switches, the signal line is commonly configured using coplanar waveguide or microstrip transmission line geometries.

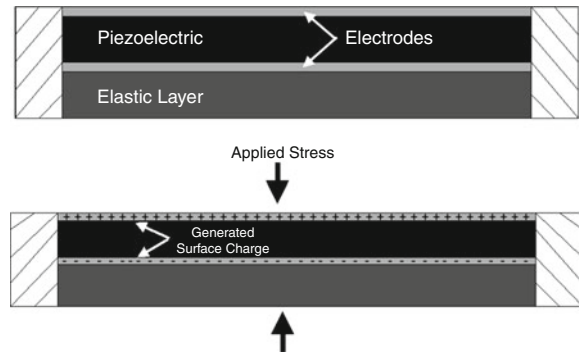
MEMS switches are defined by two different types of contacts: ohmic and capacitive (see Fig. 1). Capacitive contact switches utilize a dielectric thin film between two metals with adjustable gap. At elevated frequencies, the RF energy is capacitively coupled through the dielectric to the switch contact. As a result, capacitive switches are generally limited to higher frequency operation (>6 GHz) where capacitive coupling is feasible with reasonably sized capacitive contacts. In contrast, ohmic contacts use two metals to create the switch contact. Ohmic switches provide the opportunity for very broadband operation using carrier frequencies from DC on upward. It should be noted that the type of contact material is crucial for high-performance switches with long lifetime. Although beyond the scope of this entry,

dielectric materials with low charge trapping densities are preferred for capacitive contacts (e.g., silicon nitride and alumina) and metals with high hardness and low probability of creating frictional polymers are preferred for ohmic contacts (e.g., ruthenium and ruthenium oxide).

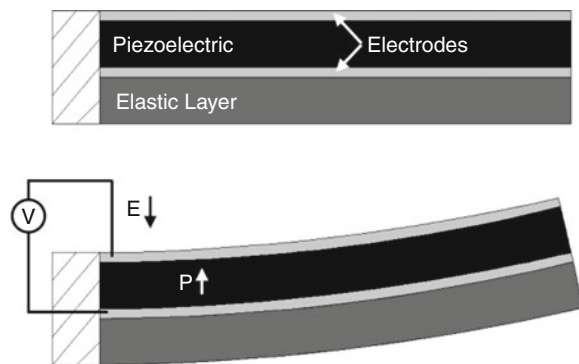
The final critical component is the actuator. Historically, most reported MEMS switches have utilized electrostatic actuators in which a metal beam is suspended over both a bias electrode and one of the switch contacts [2]. The remainder of this entry, however, will focus on an alternative actuation mechanism, piezoelectricity, and the piezoelectric actuators' role in MEMS switch technologies. When compared with switches actuated by electrostatic force, one critical advantage of piezoelectric switches is the full control of their displacement when the switch is making contact. In contrast to piezoelectric switches, electrostatic switches experience a “pull-down voltage” wherein there is an instability in the gap between the actuator and its bias electrode. Once the pull-down voltage is reached and the gap of the electrostatic capacitor is reduced to two-third of the gap at 0 V, a positive feedback is created in the device. A smaller gap creates larger closing force, which in turn moves the switch in the direction of even smaller gaps. This process of constant acceleration ends when the contact is closed against a typically immovable transmission line. In contrast, piezoelectric MEMS switches move only in response to the magnitude of the DC actuation voltage. They do not exhibit positive feedback and their dynamic behavior is well controlled. As a result, switches with piezoelectric actuation can be operated where the contacts close with only a modest acceleration resulting in minimal contact bounce thus improving switch cycle reliability [6].

Piezoelectricity

Piezoelectricity is comprised of two electromechanical responses: a direct effect and a converse effect. The direct effect refers to the generation of charge under the application of an applied stress and is governed by Eq. 1 where the induced electric displacement (D_i) is proportional to the piezoelectric coefficient (d_{ijk}) and the applied stress (T_{jk}). The converse effect refers to the generation of strain under an applied electric field



Piezoelectric MEMS Switches, Fig. 2 Illustration of the direct piezoelectric effect applied to a piezoelectric unimorph membrane



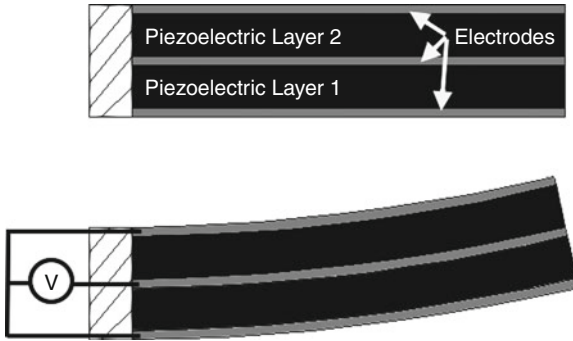
Piezoelectric MEMS Switches, Fig. 3 Illustration of the converse piezoelectric effect applied to a unimorph cantilever with the vertical scale exaggerated

and is governed by Eq. 2, where the induced strain (x_{jk}) is proportional to the piezoelectric coefficient and electric field (E_i).

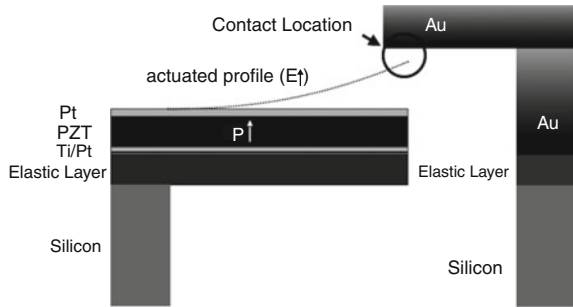
$$D_i = d_{ijk} T_{jk} \quad (1)$$

$$x_{jk} = d_{ijk} E_i \quad (2)$$

In application, the direct effect can be used for sensing and energy harvesting in which applied stresses are used to generate surface charges on the piezoelectric layer (see Fig. 2). For actuators, an applied electric field generates a strain in the piezoelectric layer producing flexure in unimorph actuators (see Fig. 3) and bimorph actuators (see Fig. 4). In switch applications, the piezoelectric unimorph can be combined with overhanging contact cantilevers to create a functioning switch (see Fig. 5).



Piezoelectric MEMS Switches, Fig. 4 Illustration of the converse piezoelectric effect applied to a bimorph cantilever with the vertical scale exaggerated



Piezoelectric MEMS Switches, Fig. 5 Cross section schematic of a PZT unimorph and electrical contacts configured for use as switch

The tensor notation for the aforementioned relationships can be simplified using matrix notation where the subscripts i, j , and k having values 1–3 are replaced with the subscripts p and q with values of 1–6 [7]. As a result, d_{ijk} is transformed to d_{ip} . The d_{ip} (strain) and e_{iq} (stress) piezoelectric coefficients are the most useful for thin film sensor and actuator applications and are related by the elastic constant, c_{pq}^E , in Eq. 3:

$$e_{iq} = d_{ip}c_{pq}^E \quad (3)$$

For thin film composites where the piezoelectric layer is rigidly attached to a supporting elastic layer (i.e., thin film or substrate), the out of plane stress, X_3 , is generally set to zero while the longitudinal strain, x_3 , varies from a combination of the piezoelectric effect and elastic coupling from the Poisson effect. As a result, an effective thin film piezoelectric coefficient, $e_{31,f}$, can be related to the common piezoelectric

coefficients, d_{31} and e_{31} (see Eq. 4) by compensating for the clamping effect experienced by piezoelectric thin films in composite structures [8, 9].

$$e_{31,f} = e_{31} - \frac{c_{31}^E e_{33}}{c_{33}^E} = \frac{d_{31}}{s_{11}^E + s_{12}^E} \quad (4)$$

Piezoelectric Materials: Ferroelectrics and Non-ferroelectrics

The important piezoelectric and dielectric material properties for the three most common thin film piezoelectrics, ZnO, AlN, and lead zirconate titanate (PZT) are listed in Table 1. A few distinctive features to recognize are the differences in crystal structure and the effective thin film piezoelectric coefficient, $e_{31,f}$, of these three compounds. The difference in crystal structures can be summarized as a distinction between purely polar materials and ferroelectric materials. Pure polar materials such as AlN and ZnO do not permit the polarization vector to change orientations relative to the crystal lattice. The polar direction is crystallographically defined along the c -axis of the Wurtzite crystal structure. In contrast, ferroelectric materials, such as PZT, have a spontaneous electric dipole moment that can be reoriented between crystallographically defined stable states by a realizable electric field (i.e., before dielectric breakdown occurs) [10]. As a result, extremely large polarization values and subsequently large dielectric constants and piezoelectric coefficients can be achieved in ferroelectric materials.

Another important characteristic distinction between purely polar materials and ferroelectric materials is the use of the constitutive equations of piezoelectricity, Eqs. 1 and 2, which describe the *linear* response of piezoelectric materials. These, generally, can be applied successfully to polar piezoelectric thin films like AlN and ZnO in most circumstances. However ferroelectrics, like PZT, can display pronounced stress and electric field-induced nonlinear material responses. This is particularly true for thin films and actuators. Thin films permit the application of extremely large electric fields with modest operating voltages due to the film thicknesses involved. Piezoelectric MEMS actuator applications typically drive the material well beyond the coercive field. For a half

Piezoelectric MEMS Switches,**Table 1** Comparison of thin film piezoelectric materials

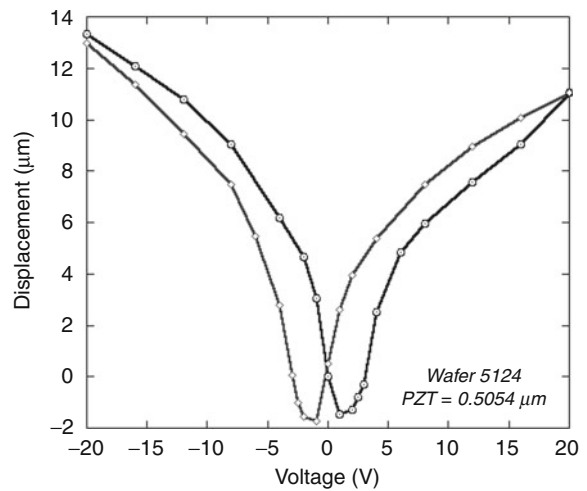
	ZnO	AlN	Pb(Zr _x Ti _{1-x})O ₃
Crystal structure	Wurtzite	Wurtzite	Perovskite
$e_{31,f}$ (C/m ²)	-0.4 → -0.8	-0.9 → -1.1	-8 → -18
$d_{33,f}$ (pC/N)	10 → 17	3.4 → 6.5	90 → 150
ϵ_{33}	8 → 12	10.1 → 10.7	800 → 1200
$\tan \delta$			0.02 → 0.05
Density (g/cm ³)	5.68	3.26	7.5 → 7.6
Young's modulus (GPa)	110 → 150	260 → 380	60 → 80
Acoustic velocity (m/s)	6.07×10^3	11.4×10^3	2.7×10^3
References	[11–14]	[11, 12, 15, 16]	[11, 12]

micron thin film of PZT (52/48) for example, the coercive field corresponds to an applied voltage of approximately 2.5 V. Ferroelectrics even show significant nonlinearity below the coercive field [17]. Despite this fact it is surprising that, for both bulk material and thin film applications, the linear piezoelectric equations are often used to describe the device response in ferroelectric materials. To accurately model ferroelectric MEMS device response, under high-operating-field (actuators) and/or bias field (sensors) conditions, these significant nonlinearities must be taken into account.

At the high fields encountered in MEMS actuators, the total strain response is due to the combination of the linear response, nonlinear piezoelectricity including saturation effects, domain wall motion, and electrostriction. All dielectrics display the property of electrostriction, whereby externally applied electric fields induce a strain response that is proportional to the square of the field strength. Equation 5 describes the total strain response in ferroelectric materials where Q is the appropriate electrostrictive coefficient, P_s is the spontaneous polarization, E is the applied electric field, and κ is the appropriate dielectric constant. The first term defines the remnant strain. The second term is equal to the linear piezoelectric coefficient (d_{ij}) multiplied by the electric field. The last term is the strain due to electrostriction. As the applied field strength increases, the electrostrictive term contributes more to the overall strain response in the material.

$$x = QP_s^2 + 2\epsilon_0\kappa QP_s E + Q(\epsilon_0\kappa E)^2 \quad (5)$$

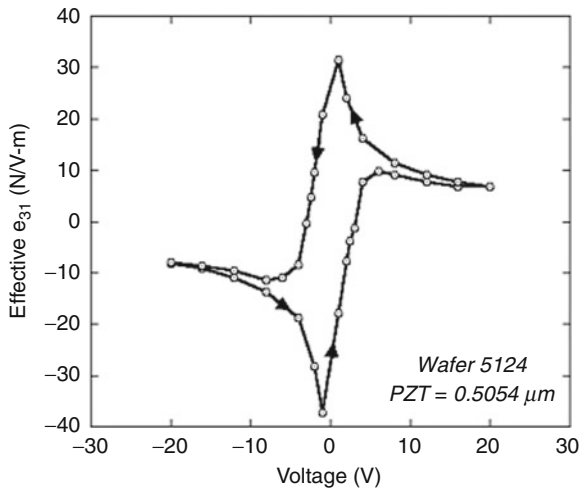
At large applied stresses and electric fields, significant contributions to the total strain response (Eq. 5) are due to nonlinear piezoelectricity and the extrinsic response. These nonlinear material properties used for



Piezoelectric MEMS Switches, Fig. 6 Displacement versus voltage plot for a PZT thin film cantilever comprised of a 0.5 μm elastic layer ($\text{SiO}_2/\text{Si}_3\text{N}_4/\text{SiO}_2$)/0.1 μm Ti/Pt bilayer, 0.505 μm PZT (52/48), and a 0.1 μm Pt top electrode. Key aspects of this diagram include the hysteretic displacement characteristics upon bipolar voltage swings. Additionally, all but a small fraction of the displacement occurs in one direction

design purposes are best measured for specific materials and processing conditions. Due to these complex nonlinear effects, the ratio of total strain to applied electric field is referred to as the effective electroactive coefficients. These coefficients can be used with the models described above using field-dependent functions.

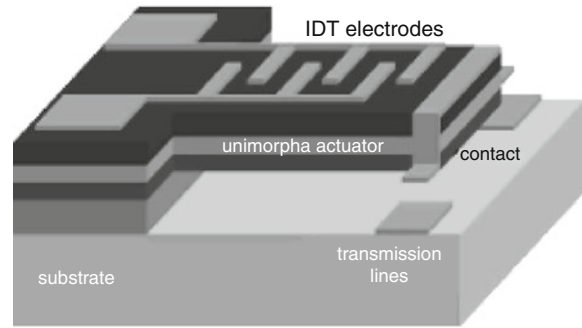
Perhaps the most important consequence of these nonlinear effects is the absence of a large bipolar strain response when operated above the coercive field. Figure 6 shows voltage-displacement data from a simple unimorph PZT MEMS-actuated cantilever, illustrating the so-called butterfly-loop. The plot was obtained by sweeping the actuation voltage up to



Piezoelectric MEMS Switches, Fig. 7 Effective piezoelectric stress constant, $e_{31,\text{eff}}$, versus voltage extracted from the displacement-voltage measurements for the PZT thin film unimorph described in Fig. 6

a positive voltage, then reducing to zero, then sweeping down to a negative voltage, and finally returning to zero field. The extracted effective electro-active piezoelectric stress constant, $e_{31,\text{eff}}$, corresponds to the displacement data in Fig. 6 and is illustrated in Fig. 7. This effective piezoelectric constant includes the small signal piezoelectric effect, electrostriction, and strain induced by large polarization charges associated with overcoming a pinched ferroelectric hysteresis loop.

The response of the device is seen to be bipolar only in the regime between the positive and negative coercive voltages where the negative displacement reverts to a positive displacement (the “V” regions of the lower curves). Thus, for large electric fields (i.e., in excess of the coercive field) applied to the piezoelectric material, the sense of the piezoelectric strain is independent of the applied field polarity; therefore, only a single sense of the piezoelectric strain is possible. In terms of the device, the in-plane contraction of the piezoelectric material at large fields gives a negative sense to the piezoelectric actuation force. The standard configuration of the d_{31} mode MEMS unimorph, with the neutral axis below the midplane of the piezoelectric layer gives a positive sense of the moment arm. Piezoelectric MEMS devices based on ferroelectric materials are most often operated under unipolar conditions and hence avoid much of the hysteresis observed in Fig. 6.

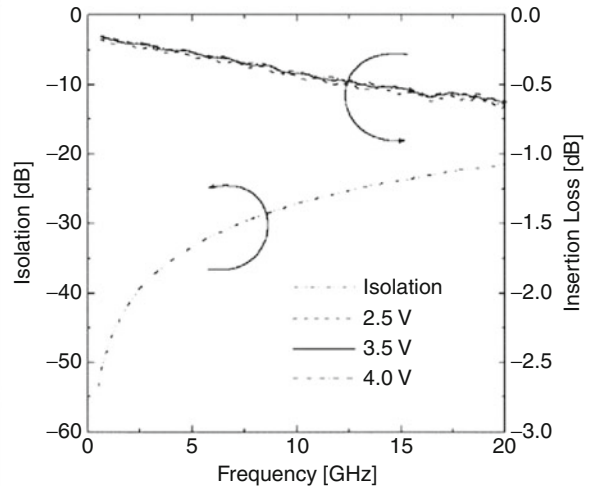
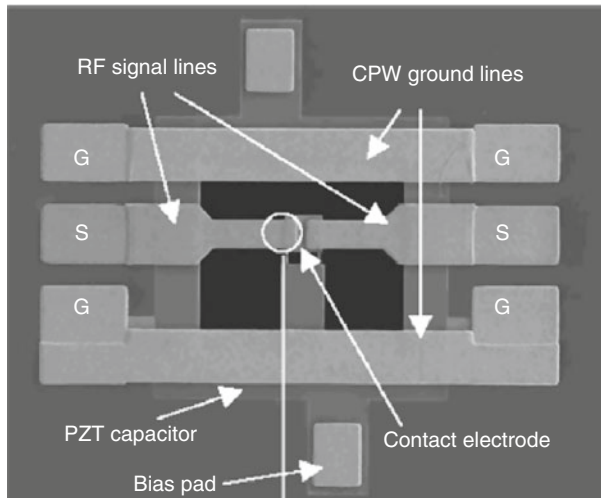


Piezoelectric MEMS Switches, Fig. 8 Illustration of a thin film PZT unimorph switch developed at Penn State University [20] (Reprinted with permission. Copyright 2003 American Institute of Physics)

Piezoelectric Switch Actuator History

One of the key limitations of electrostatically actuated MEMS switches is the inability to combine a low actuation voltage with excellent DC and RF performance (i.e., contact force and insertion loss, respectively) and reliability. Attempts to lower the actuation voltage with electrostatic switches generally rely on reducing either the mechanical stiffness of the released structures or the gap between the mechanical bridge/cantilever and the corresponding biasing pad [18]. Reductions in the stiffness can limit the restoring force of the switch and can lead to stiction failures [19]. Decreasing the electrode gap limits the high-frequency RF performance (in particular, the isolation in the open state for series switches) [2].

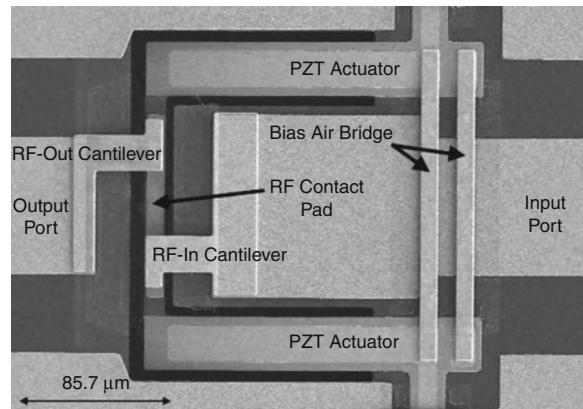
The first functioning piezoelectric MEMS DC relay was demonstrated by Gross et al. using in-plane poled PZT film actuators [20]. These devices used a cantilever beam with a supporting elastic layer, a PZT thin film as the actuator, and patterned gold structures to act as the electrodes (see Fig. 8). This device operated as low as 20 V with a switching on time as low as 2 μs . The first reported working RF MEMS switch using PZT thin films used bulk micromachining to release a unimorph actuator and a suspended transmission line [21]. This design utilized a cantilever that is perpendicular to the wave propagation direction along the RF conductor of the coplanar waveguide (CPW). The switch operated as low as 2.5 V with an insertion loss better than 0.7 dB and isolation better than -21 dB up to 20 GHz (see Fig. 9). The first surface micromachined approach for creating a PZT-actuated RF MEMS switch was developed by



Piezoelectric MEMS Switches, Fig. 9 SEM image of a bulk micromachined RF MEMS series switch using a thin film PZT actuator to complete the signal path [21] (Reprinted with permission. Copyright 2005 IEEE)

Polcawich et al. [22]. These devices operated at less than 10 V and demonstrated temperature stable performance from -25°C to 100°C for normally open series switches with isolation less than -20 dB and insertion loss of better than -1.0 dB up to 40 GHz [22, 23] (Fig. 10).

The interest in RF-compatible and cell-phone-frequency-compatible AlN-based filters has motivated the development of integrated piezoelectric switches and contour mode filters, both operated with AlN thin films. Mahameed et al. created the first AlN-based MEMS switches demonstrating good RF performance up to the low GHz regime [24]. These devices feature an innovative approach to address residual stress deformation, temperature sensitivity, and the modest piezoelectric coefficients of AlN. A symmetric “dual-beam” design features pairs of actuators that experience similar residual stress- and temperature-induced deformations that minimize variations in the open state contact gap despite the mechanical fluctuations (see Fig. 11). The device shows immunity to residual stresses, fast switching speeds (~ 2 μs), moderate actuation voltage (~ 20 V), and good low-frequency isolation (>26 dB) and insertion loss (<0.7 dB) at 2 GHz. These devices have also been successfully co-fabricated on the same wafer as contour-mode AlN RF MEMS resonators. The device fabrication leverages the process used for AlN contour-mode resonators. Once these features are defined, including multiple AlN and Pt electrode deposition and patterning steps, an amorphous silicon sacrificial layer is deposited and patterned with liftoff. This

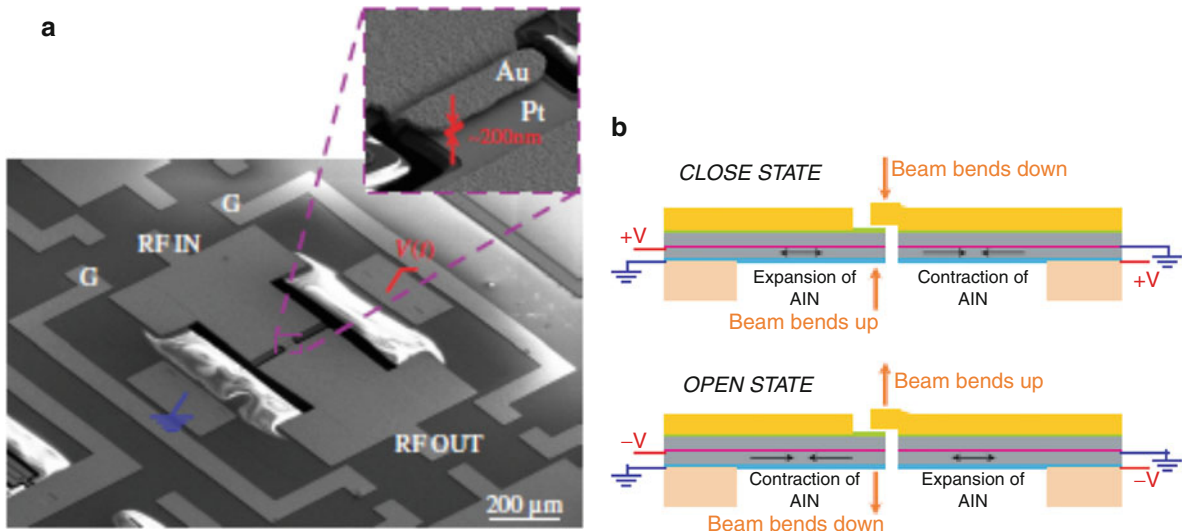


Piezoelectric MEMS Switches, Fig. 10 SEM image of a PZT-actuated RF MEMS switch [22] (Reprinted with permission. Copyright 2007 IEEE)

followed by the deposition and patterning of a thick electroplated Au layer to provide RF transmission lines and unimorph actuator features. The switch is released using a Xenon difluoride dry release to complete the process.

Switch Scaling

A number of groups are investigating mechanical logic technologies to address the limitations imposed in complementary metal-oxide-semiconductor (CMOS) transistors by gate oxide leakage. Static power



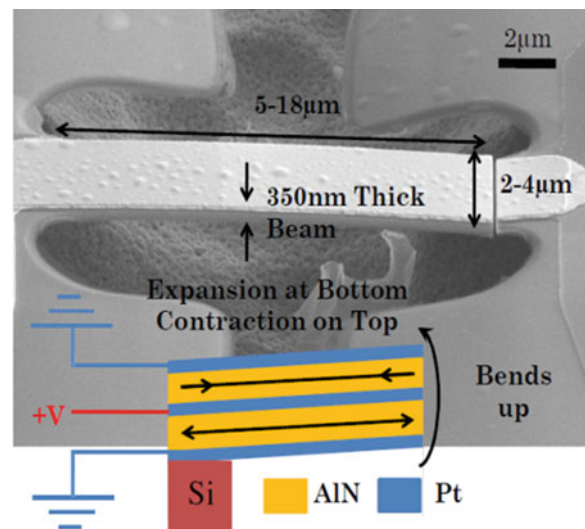
Piezoelectric MEMS Switches, Fig. 11 (a) SEM image of AlN RF MEMS switch with *inset* indicating the nano-gap contact location; (b) Illustration of principle of operation of

“dual-beam” architecture [24] (Reprinted with permission. Copyright 2008 Institute of Physics)

consumption now rivals dynamic power consumption in CMOS and presents a significant challenge to the continued scaling of CMOS logic devices. Motivated by low-power nano-mechanical logic applications, electromechanical relays based on either AlN or PZT have been investigated to provide practical solutions for low power-embedded microcontrollers. Switches have recently been demonstrated on both technologies using combinations of electron-beam lithography and focused ion beam milling [25–27].

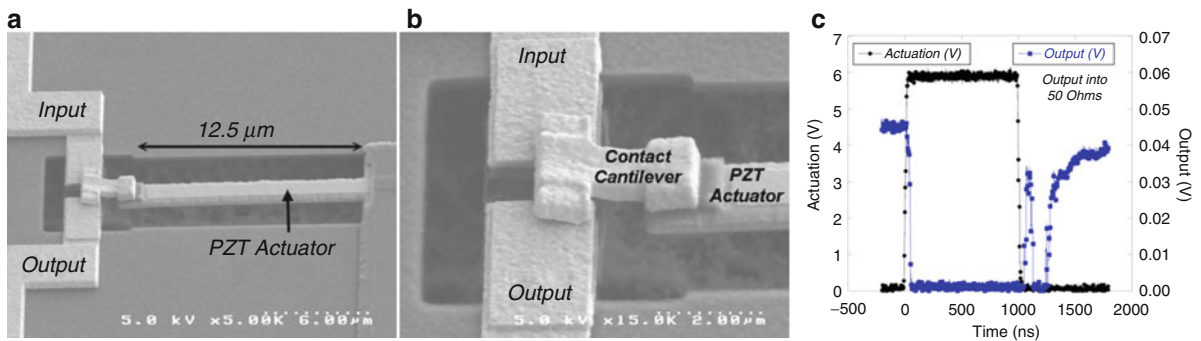
The AlN-based devices consisted of a 100 nm AlN layer and were implemented in nano-scale bimorph cantilevered actuators similar to those described previously. The actuators were fabricated using a five mask post-CMOS-compatible process similar to the process outlined for the fabrication of the AlN RF MEMS switch discussed earlier. Displacements of 40 nm were achieved at 2 V in 18 μm long cantilever devices. A key difference in the process, however, is the use of a focused ion beam (FIB) tool. The nearly complete suspended clamped-clamped AlN beam is severed near one of anchors with the FIB to create the cantilevered device seen in Fig. 12.

Similar to the scaled version of the AlN switch, a 3-D-scaled version of the aforementioned PZT RF MEMS switch has been demonstrated. Figure 13 was fabricated using an 11 step aligned electron beam lithography process following the same process flow



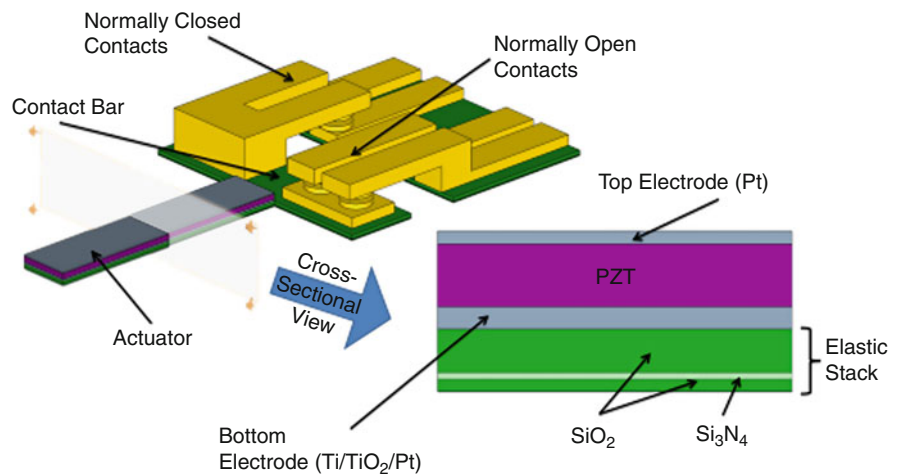
Piezoelectric MEMS Switches, Fig. 12 SEM image of AlN cantilevered bimorph nano-actuator [25] (Reprinted with permission. Copyright 2008 IEEE)

as its larger relative. The key differences were reductions in film thickness and gap along with reductions in lateral dimensions for the switch actuators and contact cantilever. In the normally closed configuration, the switch operated at 6 V demonstrated an opening response ranging from 17 to 33 ns with a corresponding closing response ranging from 33 to 43 ns.



Piezoelectric MEMS Switches, Fig. 13 Images (a and b) of a PZT nanoscaled relay fabricated from an 11 step electron beam lithography process and (c) electrical output response during one

cycle of operation [27] (Reprinted with permission. Copyright 2008 IEEE)



Piezoelectric MEMS Switches, Fig. 14 Single-pole double-throw relay schematic

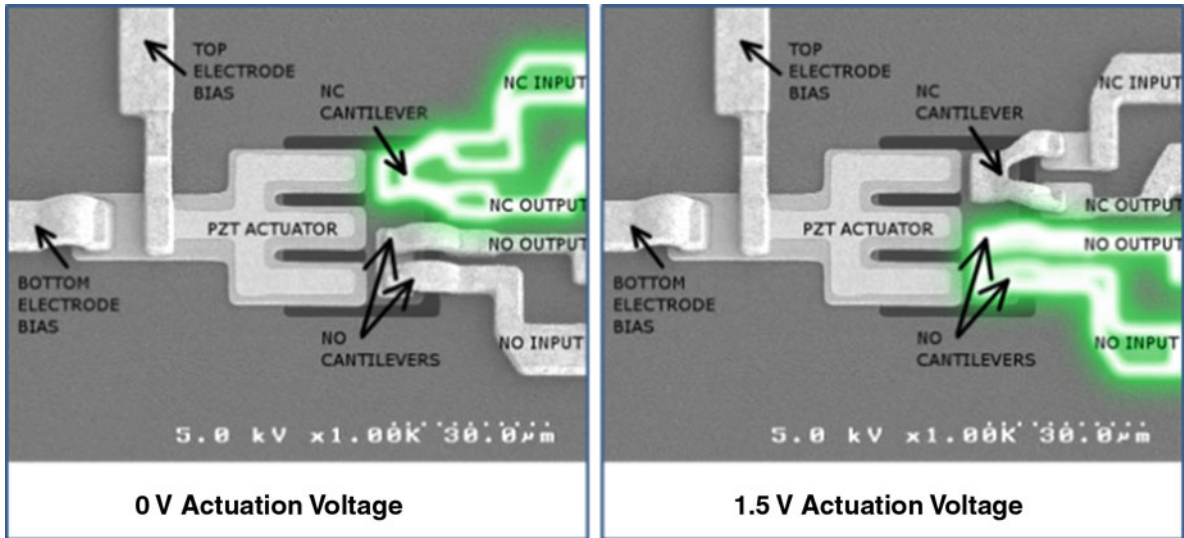
Piezoelectric-Enabled Mechanical Logic

Piezoelectric RF switches have also become enabling in very low-power digital logic circuits. Leveraging the results highlighted in the switch-scaling section, one can extend the technology to enable a fully functional electromechanical-based logic. Both AlN and PZT thin films have successfully demonstrated mechanical logic functionality. To date, the AlN logic relays have been assembled to yield both NAND and inverter gates using a body bias on the actuator [28, 29]. The body bias is used to reduce the contact gap and enable dynamic operation at less than 1 V.

The low leakage current observed on a PZT-based actuator has led to the development of a digital, single-

pole, double-throw (SPDT) mechanical relay with zero leakage current in one state and under 30-fA of leakage current in the other [30]. The device is shown schematically in Fig. 14 and with the two independent current paths highlighted in green via electron micrograph in Fig. 15. In addition to low leakage currents, the structure's material composition provides innate radiation hardening [31], and the SPDT configuration of this device, highlighted in Fig. 15, makes it optimally suited for complimentary logic circuits.

Due to PZT's large $e_{31,f}$ coefficient, large deflections, several hundred nanometers for nanomechanical logic, can be achieved at modest voltages. This enables the relay to have an open-state contact gap sufficient to support up to 40-V hot switched across its contacts, while still allowing for fully functional



Piezoelectric MEMS Switches, Fig. 15 Electron micrograph of a fabricated relay highlighting the two paths of current flow

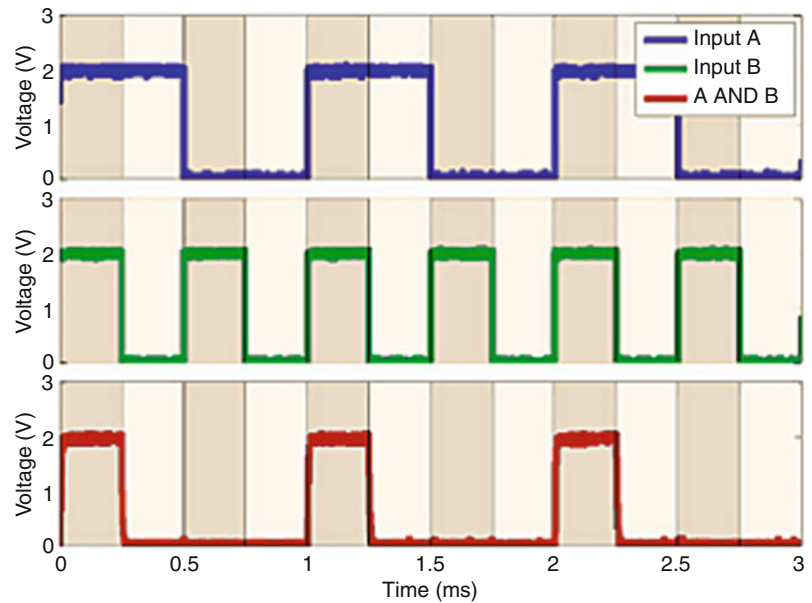
Piezoelectric MEMS Switches, Table 2 Boolean logic functions possible with PiezoMEMS SPDT relay

Boolean function	Top pt	Bottom pt	NC input	NC output	NO output	NO input
Always 0	–	–	GROUND	OUTPUT		INPUT B
A NOR B	VDD	INPUT A	GROUND	OUTPUT		NOT INPUT B
(NOT A) AND B	GROUND	INPUT A	INPUT B	OUTPUT		GROUND
NOT A	GROUND	INPUT A	VDD	OUTPUT		GROUND
A AND (NOT B)	GROUND	INPUT B	INPUT A	OUTPUT		GROUND
NOT B	GROUND	INPUT B	VDD	OUTPUT		GROUND
A XOR B	INPUT A	INPUT B	GROUND	OUTPUT		VDD
A NAND B	GROUND	INPUT A	VDD	OUTPUT		NOT INPUT B
A AND B	GROUND	INPUT A	GROUND	OUTPUT		INPUT B
A XNOR B	INPUT A	INPUT B	VDD	OUTPUT		GROUND
B	–	–	INPUT B	OUTPUT		INPUT B
(NOT A) OR B	GROUND	INPUT A	VDD	OUTPUT		INPUT B
A	–	–	INPUT A	OUTPUT		INPUT A
A OR (NOT B)	GROUND	INPUT B	VDD	OUTPUT		INPUT A
A OR B	GROUND	INPUT A	INPUT B	OUTPUT		VDD
Always 1	–	–	VDD	OUTPUT		VDD

Boolean logic operations at fewer than 2 V (see Table 2). Additionally, due to the ambipolar nature of metal-to-metal contacts, the device is well suited to functioning in a pass-gate logic configuration, allowing for a reduction in gate count without harming the signal integrity. Figure 16 demonstrates this behavior with a Boolean AND operation. Here, the top electrode of Fig. 15 is grounded, the

bottom electrode is connected to “Input A,” the normally closed (NC) input is grounded, the normally open (NO) input is connected to “Input B,” and both the NC and NO outputs are connected together. In this configuration, when “Input A” is false, ground is shorted to the output and when “Input A” is true, “Input B” is passed to the output. The additional Boolean functions and the related device

Piezoelectric MEMS Switches, Fig. 16 Measured results of Boolean AND function being performed via a PiezoMEMS relay



configurations are listed in Table 2. The device can also operate as either of the two classes of digital memory: dynamic and static (for additional information see References [32, 33]).

Conclusion

This entry provides a general overview of piezoelectric MEMS-based switch technologies and their application for RF circuits and mechanical digital relays. Both AlN and PZT thin films have demonstrated successful operation of switches for both RF and digital applications. The fundamental material differences between the AlN and PZT result in the use of bimorph actuators for AlN and unimorph actuators for PZT. The key advantages of a piezoelectric switch can be summarized as low operating or drive voltage (typically sub-10 V) reducing the need for large area and power-consuming charge pumps for integrated drive applications and full control of actuator displacement when the switch is making contact to improve contact dynamics and improve reliability.

Acknowledgment The authors would like to thank Dr. Tony Ivanov, Daniel Judy, Dr. Roger Kaul, Joel Martin, and Brian Power of the US Army Research Laboratories for their useful comments and their roles in device fabrication and characterization.

Cross-References

- ▶ [Capacitive MEMS Switches](#)
- ▶ [NEMS Piezoelectric Switches](#)
- ▶ [Piezoelectric Effect at Nanoscale](#)

References

1. Brown, E.: RF-MEMS switches for reconfigurable integrated circuits. *IEEE Microw. Theory Techn.* **46**(11), 1868–1880 (1998)
2. Rebeiz, G.: *RF MEMS Theory, Design, and Technology*. Wiley, Hoboken (2003)
3. Rebeiz, G., Tan, G., Hayden, J.: RF MEMS phase shifters: device and applications. *IEEE Microw. Magazine*, pp. 72–81 (2002)
4. Petersen, K.E.: Micromechanical membrane switches on silicon. *IBM J. Res. Dev.* **23**, 376–385 (1979)
5. Larson, L.E., Hackett, R.H., Melendes, M.A., Lohr, R.F.: Micromachined microwave actuator (MIMAC) technology – a new tuning approach for microwave integrated circuits. In: *Millimeter-Wave Monolithic Circuits Symposium Digest*, Boston, MA, June 1991, pp. 27–30
6. McCarthy, B., Adams, G.G., McGruer, N.E.: A dynamic model, including contact bounce, of an electrostatically actuated microswitch. *J. Microelectromech. Syst.* **11**, 276–283 (2002)
7. Setter, N.: ABC of piezoelectricity and piezoelectric materials. In: *Piezoelectric Materials in Devices*, Ceramics Laboratory EPFL, Lausanne, Switzerland, pp. 1–28 (2002)
8. Murali, P.: Piezoelectric thin films for MEMS. *Integr. Ferroelectr.* **17**, 297–307 (1997)

9. Dubois, M., Muralt, P.: Measurement of the effective transverse piezoelectric coefficient e_{31} , f of AlN and Pb(ZrxTi1-xO3) thin films. *Sens. Actuat.* **77**, 106–112 (1999)
10. IEEE Standard on Piezoelectricity, ANSI/IEEE Std 176 (1987)
11. Muralt, P.: PZT thin films for microsensors and actuators: where do we stand? *IEEE Trans. Ultrason. Ferro. Freq. Control* **47**, 903–914 (2000)
12. Polcawich, R.G., Pulskamp, J.S.: Additive processes for strain-dependent materials: piezoelectric MEMS. In: Reza, G., Pinyen, L. (eds.). *MEMS Processing Handbook*, pp. 273–353. Springer, New York (2011)
13. Fang, T.H., Chang, W.-J., Lin, C.-M.: Nanoindentation characterization of ZnO thin films. *Mat. Sci. Eng. A* **452–453**, 715–720 (2007)
14. Yang, P.-F., Wen, H.-C., Jian, S.-R., Lai, Y.-S., Wu, S., Chen, R.-S.: Characteristics of ZnO thin films prepared by radio frequency magnetron sputtering. *Microelect. Reliab.* **48**, 389–394 (2008)
15. Tonisch, K., et al.: Piezoelectric properties of polycrystalline AlN thin films for MEMS application. *Sens. Actuat. A* **132**, 658–663 (2006)
16. Andrei, A., Krupa, K., Jozwik, M., Delobelle, P., Hirsinger, L., Gorecki, C., Nieradko, L., Meunier, C.: AlN as an actuation material for MEMS applications. The case of AlN driven multilayered cantilevers. *Sens. Actuat. A* **141**, 565–576 (2008)
17. Damjanovic, D.: Ferroelectric, dielectric and piezoelectric properties of ferroelectric thin films and ceramics. *Rep. Prog. Phys.* **61**, 1267–1324 (1998)
18. Peroulis, D., Pacheco, S.P., Sarabandi, K., Katehi, L.: Electromechanical considerations in developing low voltage RF MEMS switches. *IEEE Trans. Microw. Theory Techn.* **51**, 259–270 (2003)
19. Majumber, S., et al.: Measurement and modeling of surface micromachined, electrostatically actuated microswitches. In: *International Conference on Solid-State Sensors, Actuators and Microsystems*, Chicago, pp. 1145–1148 (1997)
20. 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**, 174–176 (2003)
21. Lee, H.C., Park, J.Y., Bu, J.U.: Piezoelectrically actuated RF MEMS DC contact switches with low voltage operation. *IEEE Microw. Wireless Comp. Lett.* **15**(4), 202–204 (2005)
22. Polcawich, R.G., Pulskamp, J.S., Judy, D., Ranade, P., Trolier-McKinstry, S., Dubey, M.: Surface micromachined microelectromechanical ohmic series switch using thin-film piezoelectric actuators. *IEEE Trans. Microw. Theory Techn.* **55**, 2642–2654 (2007)
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: *Proceedings of IEEE International Microwave Symposium*, Honolulu, pp. 2083–2086 (2007)
24. Mahameed, R., et al.: Dual-beam actuation of piezoelectric AlN RF MEMS switches monolithically integrated with AlN contour-mode resonators. *J. Micromech. Microeng.* **18**, 105011 (2008)
25. Sinha, N., et al.: Ultra thin AlN piezoelectric nano-actuators. In: *Proceedings of the Transducers*, Denver, CO, pp. 469–472 (2009)
26. Judy, D., Polcawich, R.G., Pulskamp, J.: Low voltage complementary MEMS logic using piezoelectric actuators. In: *Solid State Sensor, Actuator and Microsystems Workshop*, Hilton Head, SC, pp. 328–331 (2008)
27. Judy, D.C., Pulskamp, J.S., Polcawich, R.G., Currano, L.: Piezoelectric nanoswitch. In: *Proceedings of the 22nd IEEE MEMS Conference*, Sorrento, Italy, pp. 591–594 (2009)
28. 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: *Proceedings of 2010 Microelectromechanical Systems Conference (MEMS 2010)*, Wanchai, Hong Kong (2010)
29. Sinha, N., Jones, T.S., Guo, Z., Piazza, G.: Body-biased complementary logic implemented using AlN piezoelectric MEMS switches. In: *Proceedings of 2009 IEEE International Electron Devices Meeting (IEDM 2009)*, Baltimore, MD (2009)
30. Proie Jr., R.M., Polcawich, R.G., Pulskamp, J.S., Ivanov, T., Zaghoul, M.: Development of a PZT MEMS switch architecture for low-power digital applications. *J. Microelectromech. Syst.* **20**(4), 1032–1042 (2011)
31. Benedetto, J.M., Moore, R.A., McLean, F.B., Brody, P.S., Dey, S.K.: The effect of ionizing radiation on sol-gel ferroelectric PZT capacitors. *IEEE Trans. Nucl. Sci.* **37**(6), 1713–1717 (1990)
32. Proie, R., Polcawich, R.G., Pulskamp, J.S., Ivanov, T., Zaghoul, M.: Nano-electromechanical storage element for a low power complimentary logic architecture using PZT relays. In: *Proceedings of Transducers*, Beijing, China (2011)
33. Proie, R., Pulskamp, J.S., Polcawich, R.G., Ivanov, T., Zaghoul, M.: Low power 3-bit piezoelectric MEMS analog to digital converter. In: *MEMS 2011*, Cancun, Mexico, 23–27 Jan 2011, pp. 1241–1244

Piezoelectric Switches

► NEMS Piezoelectric Switches

Piezoresistance

► Piezoresistivity

Piezoresistive Effect

► Piezoresistivity

Piezoresistivity

Woo-Tae Park
Seoul National University of Science and Technology,
Seoul, Korea

Synonyms

Piezoresistance; Piezoresistive effect; Semiconductor piezoresistance

Definition

Piezoresistivity is defined as change of resistance of semiconductors from the applied mechanical stress.

History

The discovery of piezoresistance dates back to the late nineteenth century (1856), when Lord Kelvin discovered the change of resistance in copper and iron wires. He compared the difference in the change of resistance of the two materials, while the elongation was the same [1]. The term “piezoresistance” was coined by Cookson in 1935, when he defined the term as the change in conductivity with stress [2]. The term “piezo” is rooted from the term “piezen” from the Greek meaning to *press*. “Piezoelectricity” sounds similar, but with a very different effect, where charge is generated from applied stress.

When Nobel laureates, Bardeen and Shockley invented the transistor, they also predicted relatively large conductivity change with deformation in single crystal semiconductors [3]. Later C.S. Smith, who was a visiting scientist at Bell labs, reported his pioneering results in the large piezoresistance in semiconductors [4]. Smith’s work marked the start of semiconductor piezoresistive sensors (Table 1).

In the 1960s, at least a couple of companies started manufacturing commercial silicon strain gauges, and in 1962, an “integrated” device using diffused piezoresistors was used to make pressure sensors at Honeywell [5].

From the early piezoresistive strain gauges to the present implanted piezoresistive pressure sensors and accelerometers, this technology required the

Piezoresistivity, Table 1 Values of the derived piezoresistance coefficients in units of $10^{-12} \text{ cm}^2 \text{ dyn}^{-1}$. Resistivities are in ohm-cm

Type, material, and resistivity	$-\frac{1}{P} \frac{\delta\rho}{\rho} = \Pi_{11} + 2\Pi_{12}$			
	Π_{44}	$\Pi_{11}-\Pi_{12}$	Adiabatic	Isothermal
<i>n</i> Ge 1.5	−138.1	+1.0	(−8.7) ^a	−7.0
<i>n</i> Ge 4.2 ^b				−8.3
<i>n</i> Ge 5.7	−136.8	+1.2	−10.5	−8.8
<i>n</i> Ge 8.4 ^b				−9.8
<i>n</i> Ge 9.9	−137.9	+0.3	−14.7	−13.0
<i>n</i> Ge 16.6	−138.7	+0.3	−16.2	−17.9
<i>n</i> Ge 18.5 ^b				−27.1
<i>p</i> Ge 0.9–3.3 ^c				+2.5
<i>p</i> Ge 1.1	+96.7	−6.9	+2.7	+5.2
<i>p</i> Ge 15.0	+98.6	−15.6	−0.6	+1.9
<i>n</i> Si 11.7	−13.6	−115.6	+4.6	+5.7 ^d
<i>p</i> Si 7.8	+138.1	+7.7	+4.4	+6.0

After Smith [4]. © 1954 American Physical Society, http://www.prola.aps.org/abstract/PR/v94/i1/p42_1

^aThis value extrapolated from *n* Ge 5.7 and *n* Ge 9.9

^bP. W. Bridgman (see Ref. [13]) and by latter. Values are from the initial pressure effect

^cP. W. Bridgman (see Ref. [13]). There were six specimens in this resistivity range, five of which were consistent at the value shown and one of which showed zero pressure effect

^dMeasured directly

three-dimensional micromachining of silicon. This requirement led to the development of bulk etching technology, and the development of MEMS technology.

Piezoresistance Essentials

Theory

Electrical resistance can be defined by a function of the resistivity of the material and the dimensions, as given in the following equation:

$$R = \rho \frac{l}{a}$$

where R = resistance, ρ = resistivity, l = length, and a = cross-sectional area.

So, the change of resistance is a function of the dimensional change and the resistivity change.

$$\frac{\Delta R}{R} = (1 + 2\nu) \varepsilon + \frac{\Delta\rho}{\rho}$$

where ν = Poisson’s ratio and ε = strain.

The former term comes from the dimensional change, and the latter term is the resistivity change. The change of resistance to the applied stress is defined as the gauge factor.

$$GF = \frac{\Delta R/R}{\varepsilon}$$

where GF = gauge factor.

For metals, geometric effect alone approximately accounts for 1.4–2.0 of gauge factor, and the change in resistivity is almost zero. However, for single crystal semiconductors, such as silicon and germanium, the change in resistivity can be 50–100 times larger than the geometric term. Because of this exceptional large resistivity change in semiconductors, a new term called “piezoresistive coefficient” denoted by “ π ” is defined as the following:

$$\frac{\Delta\rho}{\rho} = \pi_t\sigma_t + \pi_l\sigma_l$$

where π = piezoresistive coefficient, σ = stress, and the subscripts t = transverse, and l = longitudinal. Transverse stress and piezoresistive coefficient are defined as mechanical stress and electrical field being orthogonal, and longitudinal stress and piezoresistive coefficient are defined as mechanical stress and electrical field being parallel.

The theories to explain this large piezoresistive property of semiconductors are based on one-dimensional descriptions of electron and hole transport in crystalline structures under strain. The models are based on bandgap energy models, wave mechanics, and quantum effects; interested reader should refer to [6] for detailed explanation of the models.

N-type and p-type piezoresistors exhibit opposite trends in resistance change and different direction dependence under strain. The magnitude and sign of the piezoresistive coefficient depend on a number of factors including doping concentration, temperature, crystallographic direction, and relation between the current and stress direction. For n-type silicon, Herring explained piezoresistance in silicon or germanium using the Many-Valley model [7].

Fabrication

The most commonly used fabrication processes for doping piezoresistors are diffusion and ion

implantation. Although both methods have its advantage for doping silicon, ion implantation is the dominant method used for doping piezoresistors.

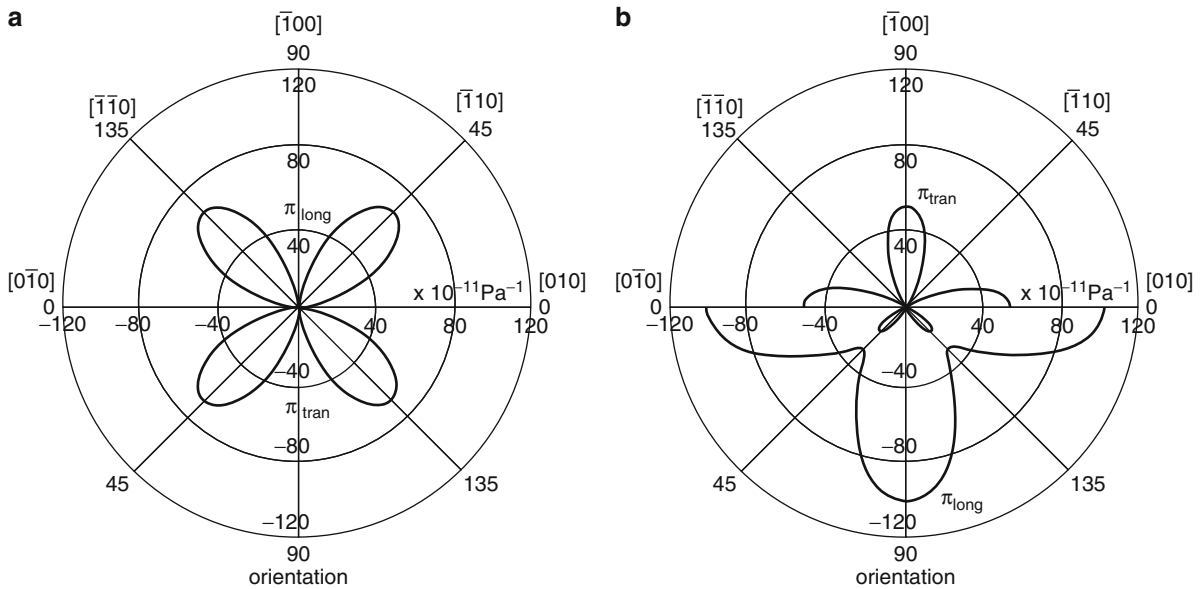
1. Diffusion: Diffusion is the migration of dopant atoms from a high-concentration region to a low-concentration region. The fabrication steps for diffusion, most commonly involve pre-deposition and drive-in step. During pre-deposition, the wafers are placed in high temperature (900–1,300°C) with a gas-phase or a solid-phase dopant source. The dopant source (e.g., diborane, phosphine, or arsine) is carried in an inert gas (N₂ Ar). The resulting dopant profile is approximated by a complementary error function. After the pre-deposition step, the dopants are driven in deeper with high-temperature annealing. Diffusion method is still useful for high-dose doping (near solid solubility).
2. Ion implantation: Ion implantation provides better control of the dopant dose, which is why it is the preferred method to dope piezoresistors today. The process starts by accelerating the dopant ions at high energy (KeV to MeV) into the substrate. The implanted ions leave a cascade of damage in the crystal structure. Dopant distribution is approximated by a symmetric Gaussian distribution. During implantation, 7° tilt angle used to avoid ion channeling effect. After implant, a short high-temperature annealing is used to restore the crystal lattice. Typical masking material is photoresist, silicon dioxide, or silicon nitride.

Design and Process Effects on Performance

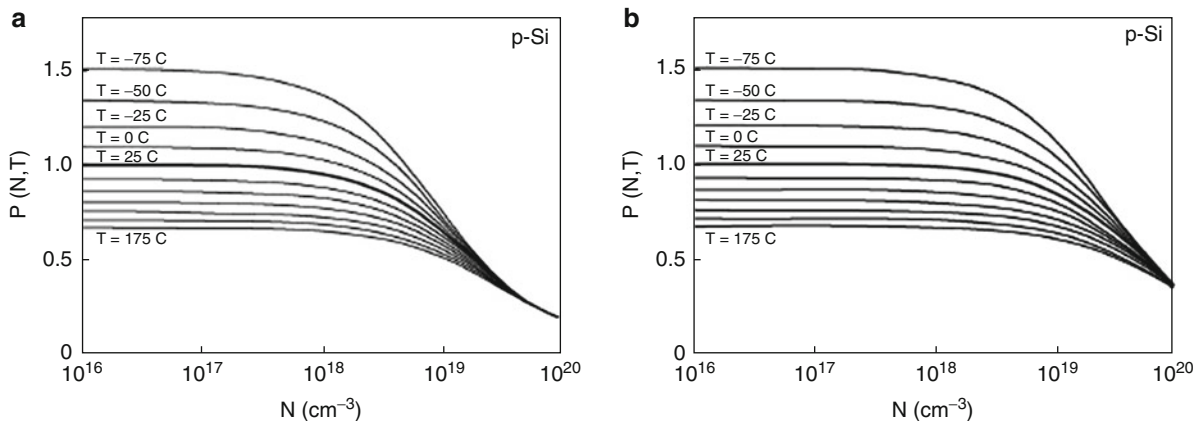
Design and process conditions strongly affect the sensitivity and noise of the piezoresistor. Doping concentration and piezoresistor crystallographic orientation strongly affect the sensitivity. Choosing the device geometry, doping, and anneal conditions, the piezoresistive device designer must also consider the temperature coefficients of sensitivity and resistance, nonlinearity with strain and temperature, and noise limits.

Kanda established the graphical representations of the piezoresistive coefficients in arbitrary orientations in the commonly used crystal substrates (Fig. 1) [8]. Kanda also showed theoretical calculations of the piezoresistive coefficients and dopant concentration (Fig. 2).

Resistors are generally sensitive to temperature. Silicon-doped resistors are also sensitive to pressure.



Piezoresistivity, Fig. 1 Piezoresistive coefficients in the popular (100) plane for (a) p-type silicon (b) n-type silicon (After Kanda [6], © 1982 IEEE)



Piezoresistivity, Fig. 2 The piezoresistance factor $P(N, T)$ as a function of doping concentration and temperature for (a) p-type silicon and (b) n-type silicon (After Kanda [6], © 1982 IEEE)

Temperature affects the absolute resistor values, but also affects the sensitivity. General trend is that the higher the sensitivity, the higher the temperature coefficient of the sensitivity. However, the rise in temperature sensitivity is faster than that of piezoresistive sensitivity. Sensor manufacturers use the Wheatstone configuration of sensors and conditioning circuits to compensate this temperature sensitivity.

Piezoresistive sensitivity also shows nonlinearity at high strain. So designers should avoid exerting strain higher than 0.1%.

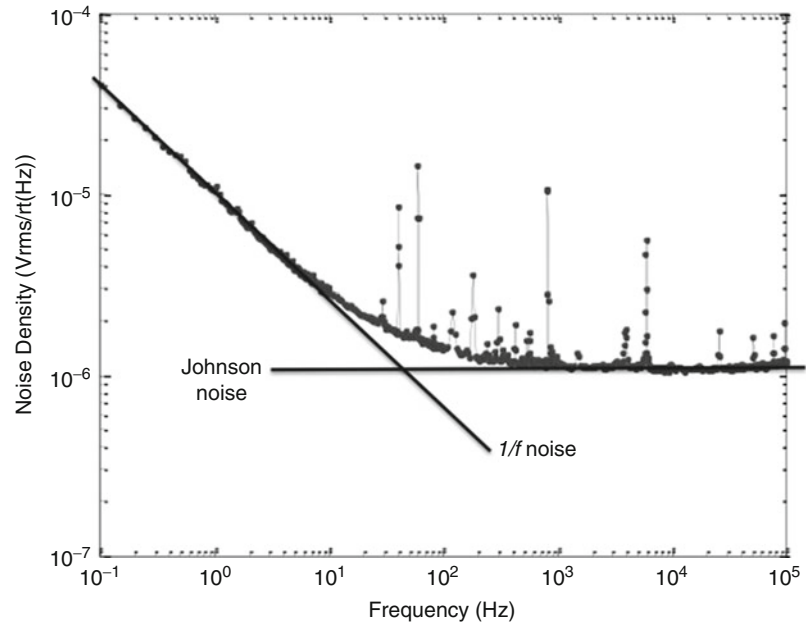
Noise

Noise in piezoresistors manifests as random fluctuations of the potential. Apart from other noise sources in the system, noise in the piezoresistor itself sets the fundamental limit in the resolution of the piezoresistor. Noise contributions come from two main components: Johnson noise (thermal noise) and $1/f$ noise (flicker noise). Figure 3 shows a representative example of piezoresistor noise.

The Johnson noise is determined by resistor value. So higher resistors have higher Johnson noise [9].

Piezoresistivity,

Fig. 3 Typical noise curve of a piezoresistor. The sloped *solid line* is the noise dominated by $1/f$ noise component plus the underlying Johnson noise. The *flat line* represents the inherent Johnson noise



$1/f$ noise is, as the name implies, inversely proportional to the frequency. So this noise component is the main culprit for low-frequency measurements. This is important since piezoresistors are used in mostly low-frequency static measurements. Although there are still debates on the source of this noise, there are two main popular theories. Hooge model views the fluctuation as carriers' mobility fluctuation by scattering centers [10]. Since the scattering centers depend on the quality of the crystal, annealing to repair internal damage reduces the noise. McWhorter model views the fluctuation as a surface effect scattering [11].

Alternative Piezoresistive Materials and Structures

Today, most commercially available piezoresistive sensors use silicon as the material. With the advances in material science, researchers studied new materials or scaling of the silicon to enhance the properties in sensitivity, and higher melting temperature.

Silicon Carbide (SiC)

Silicon carbide was attractive because of its superior mechanical properties over silicon such as higher

modulus (424 GPa), higher sublimation temperature (1,800°C), and inertness to corrosive environments. Although the fabrication is not as straightforward as silicon, and growing oxide is a challenge due to the carbon in the film, SiC-based piezoresistive devices were able to achieve gauge factors as large as -35 [12] and withstand harsh operating environments, such as high temperatures (200–500°C) [13] and high impact (40,000 g) [14].

Diamond

Diamond has higher Young's modulus, yield strength, and thermal conductivity than silicon. The reported piezoresistive gauge factors of single crystal diamond range up to 3,836. However, it is extremely difficult to grow single crystal diamond, and poly diamond has similar sensitivities as silicon carbide [15].

Carbon Nanotube (CNT)

Carbon nanotubes are the strongest and most resilient materials known in nature. Electronically, based on the manner in which the graphene films are rolled, it can be metallic, semiconducting, or insulator. Gauge factors range from 400 to 850 [16] for CNT response adhered to pressurized membranes. However, the difficulty is in the ohmic connection to the CNT. A more

manufacturable process was used mixing CNT to SU8, but the attained sensitivities were lower than what is possible with bare CNTs [17].

Silicon Nanowires

Silicon nanowires are simply reducing the dimension of conventional single crystal silicon components. Very high gauge factor, higher than 6,000, was reported for silicon nanowires in dimensions under 100 nm [18]. Reducing the dimensions down to 100 nm is not a big challenge for modern lithography and further shrinking by self-limiting oxidation is possible. Neuzil reported that top-down manufactured (by standard clean room process) nanowires can reach up to 5,000 of gauge factor [19]. However, recently Milne reported that his new measurement techniques revealed that the high change of signal is most likely due to long-term drift rather than high gauge factor [20].

Applications

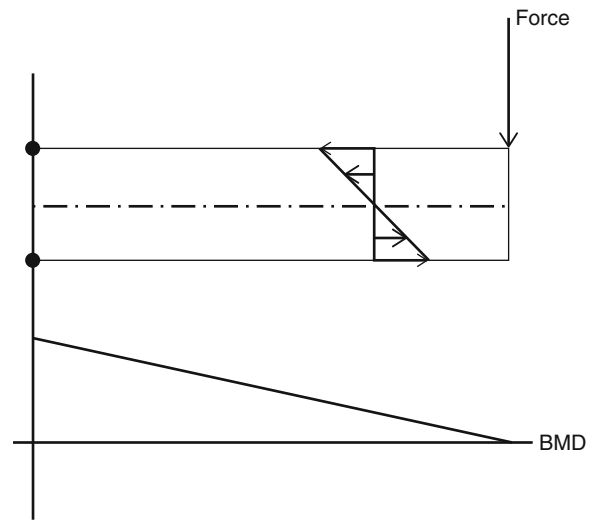
Piezoresistors have the longest history of usage as MEMS devices since the 1950s [4]. Since they measure strain, piezoresistors have wide applications such as cantilevers, pressure sensors, and accelerometers.

Cantilevers

Cantilevers are used as force sensors, biosensors, or imaging (atomic force microscopy (AFM)). Figure 4 is showing where the piezoresistor should be placed, which is the maximum stress location of the cantilever. For force sensing applications, scaling is important to measure smaller cantilevers to measure smaller force. Piezoresistors do have advantage in this area, since they can more easily scale if compared to capacitive or optical detection methods. Tortonese et al. first used piezoresistive sensing for cantilevers and achieved 0.1 Å resolution [21] (Fig. 4).

Pressure Sensors

Pressure sensing is the most reported application for piezoresistive MEMS sensors. With four piezoresistors at the four edges forming a Wheatstone bridge configuration, the output is maximized in a differential sensing configuration. Upon applied



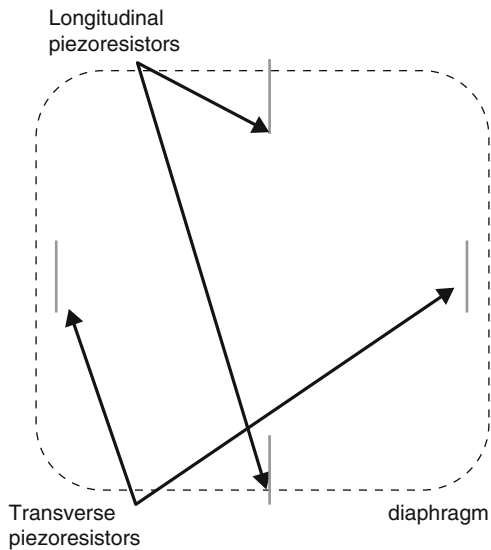
Piezoresistivity, Fig. 4 A cantilever sensor representation. Upon applied normal force, the bending moment amplitude is shown by the bending moment diagram (BMD). The internal stress/strain due to the bending moment is shown by the *arrows*. Therefore, the maximum stress/strain points are at the root of the cantilever and at the *top* and *bottom* surfaces of the beam

pressure, the two longitudinal piezoresistors change in opposite direction as the two transverse piezoresistors (Fig. 5).

Wheatstone bridge also improves the temperature sensitivity by canceling the temperature coefficient in each neighboring resistor. Piezoresistive sensing is still the most dominant sensing type for all MEMS pressure sensors. The only downside of piezoresistive sensing for pressure sensors is the inability to adapt to polysilicon surface micromachining. However, recent work from Bosch using porous silicon showed a very low-profile manufacturing process without the need of backside etching [22].

Accelerometers

One of the very first MEMS accelerometers reported used piezoresistive transduction [23]. Numerous research articles reported the use of piezoresistive sensing in accelerometers. Recent products from Hitachi metals and Hokuriku use a novel three-axis sensing structure for piezoresistive accelerometers. However, the most dominant sensing method used for commercial accelerometers is capacitive. This is due to the fact that capacitive sensing can be more sensitive



Piezoresistivity, Fig. 5 Typical piezoresistive pressure sensor layout. The *dotted line* is representing the micromachined diaphragm. Upon applied pressure, the two longitudinal piezoresistors change resistance in opposite direction as the two transverse piezoresistors. Forming a complete Wheatstone bridge with these four resistors, this configuration allows the maximum output

by employing very flexible springs without the need to have large strain concentration on the springs themselves. Capacitive sensing has proven to be more scalable for accelerometers.

Conclusions

Short after the discovery of transistors in 1947, piezoresistive property has been discovered in 1954 [4]. Through more than 50 years of steady development, piezoresistive sensors have transformed the way we live. It is used in numerous medical, automotive, aerospace, and consumer markets. While it is one of the easiest sensing techniques to construct, the fundamental physics is still not well understood. With the advances in material science, new materials are being proposed and studied. Yet, still the simplest implanted piezoresistors are the most common forms in use today. Piezoresistive sensing will always be in the MEMS designers' toolbox and employed in many devices in the years to come.

Cross-References

- ▶ [Atomic Force Microscopy](#)
- ▶ [Piezoresponse Force Microscopy and Spectroscopy](#)

References

1. Thomson, W.: On the electro-dynamic qualities of metals: effects of magnetization on the electric conductivity of nickel and of iron. *Proc. R. Soc. Lond.* **8**, 546–550 (1856)
2. Cookson, J.W.: Theory of the piezo-resistive effect. *Phys. Rev.* **47**, 194–195 (1935)
3. Bardeen, J., Shockley, W.: Deformation potentials and mobilities in non-polar crystals. *Phys. Rev.* **80**, 72–80 (1950)
4. Smith, C.S.: Piezoresistance effect in germanium and silicon. *Phys. Rev.* **94**, 42–49 (1954)
5. Tufte, O.N., Chapman, P.W., Long, D.: Silicon diffused-element piezoresistive diaphragms. *J. Appl. Phys.* **33**, 3322–3327 (1962)
6. Sun, Y., Thompson, S.E., Nishida, T.: Physics of strain effects in semiconductors and metal-oxide-semiconductor field-effect transistors. *J. App. Phys.* **101**, 104503 1–104503 22 (2007)
7. Herring, C.: Transport and deformation potential theory for many-valley semiconductors with anisotropic scattering. *Phys. Rev.* **101**, 944 (1956)
8. Kanda, Y.: A graphical representation of the piezoresistance coefficients in silicon. *IEEE Trans Electron. Dev.* **29**, 64–70 (1982)
9. Nyquist, H.: Thermal agitation of electric charge in conductors. *Phys. Rev.* **32**, 110–113 (1928)
10. Hooge, F.N.: $1/f$ noise is no surface effect. *Phys. Lett. A* **29**, 139–140 (1969)
11. McWhorter, A.L.: *Semiconductor Surface Physics*. University of Pennsylvania Press, Philadelphia (1957)
12. Okojie, R.S., Ned, A.A., Kurtz, A.D., Carr, W.N.: Characterization of highly doped n- and p-type 6H-SiC piezoresistors. *Electron. Dev. IEEE Trans.* **45**, 785–790 (1998)
13. Eickhoff, M., Moller, H., Kroetz, G., Berg, J.V., Ziemann, R.: A high temperature pressure sensor prepared by selective deposition of cubic silicon carbide on SOI substrates. *Sens. Actuators A Phys.* **74**, 56–59 (1999)
14. Atwell, A.R., Okojie, R.S., Kornegay, K.T., Roberson, S.L., Beliveau, A.: Simulation, fabrication, and testing of bulk micromachined 6H-SiC high-g piezoresistive accelerometers. *Sens. Actuators A* **104**, 11–18 (2003)
15. Taher, I., Aslam, M., Tamor, M., Potter, T.J., Elder, R.C.: Piezoresistive microsensors using p-type CVD diamond films. *Sens. Actuators A* **45**, 35–43 (1994)
16. Grow, R.J., Wang, Q., Cao, J., Wang, D., Dai, H.: Piezoresistance of carbon nanotubes on deformable thin-film membranes. *Appl. Phys. Lett.* **86**, 93104–93107 (2005)
17. Dai, H.: Carbon nanotubes: opportunities and challenges. *Surf. Sci.* **500**, 218–241 (2002)

18. He, R., Yang, P.: Giant piezoresistance effect in silicon nanowires. *Nat. Nano* **1**, 42–46 (2006)
19. Milne, J.S., Rowe, A.C.H., Arscott, S., Renner, C.: Giant piezoresistance effects in silicon nanowires and microwires. *Phys. Rev. Lett.* **105**, 226802 (2010)
20. Neuzil, P., Wong, C.C., Reboud, J.: Electrically controlled giant piezoresistance in silicon nanowires. *Nano Lett.* **10**, 1248 (2010)
21. Tortorese, M., Barrett, R.C., Quate, C.F.: Atomic resolution with an atomic force microscope using piezoresistive detection. *App. Phys. Lett.* **62**, 834–836 (1993)
22. Armbruster, S., Schafer, F., Lammel, G., Artmann, H., Schelling, C., Benzel, H., Finkbeiner, S., Larmer, F., Ruther, P., Paul, O.: A novel micromachining process for the fabrication of monocrystalline Si-membranes using porous silicon. In: *The 12th International Conference on Solid-State Sensors, Actuators, and Microsystems*, Boston, pp. 246–249 (2003)
23. Roylance, L.M.: A miniature integrated circuit accelerometer for biomedical applications. Ph.D., Electrical Engineering Department, Stanford University (1978)

Piezoresponse Force Microscopy and Spectroscopy

Sergei V. Kalinin¹, Brian J. Rodriguez² and Andrei L. Kholkin³

¹Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, USA

²Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland

³Center for Research in Ceramic and Composite Materias (CICECO) & DECV, University of Aveiro, Aveiro, Portugal

Definition

Piezoresponse force microscopy and spectroscopy (PFM and PFS, respectively) refer to a family of scanning probe microscopy (SPM) techniques based on the coupling between electrical bias and mechanical response (electromechanical detection). PFM and PFS are broadly used for the characterization of piezoelectric and ferroelectric properties (domain structure, switching dynamics, etc.) at the nanometer scale, and are also applicable to polar biological and macromolecular systems. PFM is closely related to electrochemical strain microscopy (ESM), an SPM

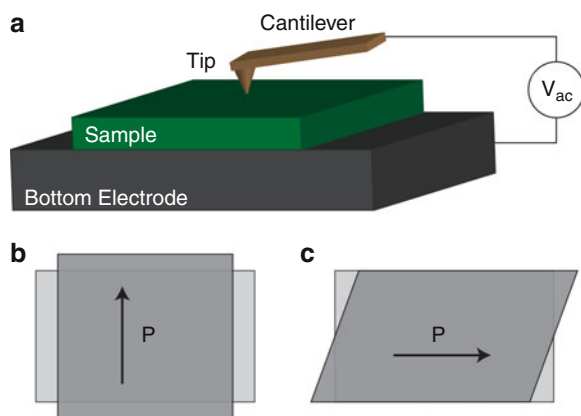
method used for the characterization of ionic transport and electrochemical reactivity on the nanoscale in ionic conductors.

Introduction

PFM and PFS were originally developed for the characterization of ferroelectric materials. The history of PFM can be traced to the first studies of local bias-induced sample deformations in piezoelectric materials using scanning acoustic microscopy [1] and scanning tunneling microscopy [2]. The seminal work by Kolosov and Gruverman [3] gave rise to the broad application of this method for imaging and manipulating domain structures in ferroelectric materials. Recent advances in PFM and PFS have allowed a wide variety of materials to be studied, including piezoelectric semiconductors, ferroelectric polymers, biological polymers, and energy materials. Many historical aspects of PFM development are discussed in recent books and reviews [4–10]. The unique aspect of PFM compared to other SPM techniques is the detection principle, which utilizes the coupling between bias and strain, often allowing for quantitative measurements.

Principles of PFM

PFM is implemented using an atomic force microscope (AFM) operated in contact mode. A nanometer-sized probe tip attached at the end of a microfabricated cantilever is brought into contact with a sample. The interaction force between the sample and the tip causes the cantilever to deflect, which can be detected using an optical beam method, i.e., through the displacement of a laser beam reflected off the backside of the cantilever and into a photodetector. Typically, a feedback loop is employed which adjusts the tip–sample separation to keep the cantilever deflection and hence the (static) loading force constant. In PFM, a conducting probe is used and an alternating current bias is applied to the conducting cantilever to create a highly localized time-dependent electric field at the probe tip apex. When the sample under investigation is piezoelectric, the material will deform, i.e., expand and contract, in response to the applied field. The tip can



Piezoresponse Force Microscopy and Spectroscopy, Fig. 1 Schematics of (a) the PFM experimental setup and (b) out-of-plane and (c) in-plane sample deformations which result in cantilever deflection and torsion, respectively

simultaneously be used as the sensor to detect this bias-induced deformation. The frequency of the ac-bias (typically 3 kHz–3 MHz) is chosen to be much higher than the scan rate of the AFM (~ 1 Hz), and thus, using a lock-in amplifier technique, the sample topography and piezoelectric response can be recorded simultaneously. The piezoelectric response is generally measured as an amplitude and phase signal. The PFM amplitude signal corresponds to the distance the tip is deflected, and is thus directly related to the deformation of the sample. The latter is determined by piezoelectric tensor coefficients (in the laboratory coordinate system), which in turn depend on the crystal structure of the material under investigation and the crystal orientation. The PFM phase signal is related to the phase lag of the piezoelectric response of the material compared to the sinusoidal ac-bias used to induce the deformation. When the direction of the applied electric field is parallel to the polarization direction of the ferroelectric domain under investigation, the domain will expand. Conversely, when the electric field and polarization are antiparallel, the domain will contract. Thus, the PFM phase provides information about the orientation of the ferroelectric polarization. The nanoelectromechanics (tip bias- and strain-induced phenomena) of PFM is reported in detail elsewhere [10]. A simple schematic of PFM is shown in Fig. 1. Note that similar to contact mode AFM, PFM can be operated in air, liquid, and vacuum environments.

For materials with a general crystallographic orientation, the surface deformation is not necessarily purely out-of-plane, and the surface can shift laterally (shear deformation) depending on the polarization direction and the piezoelectric coefficients. This shear motion results in cantilever torsion for a tip in contact with the sample surface. Thus, in order to develop a full picture of the polarization orientation, it is important to measure both the tip deflection (vertical PFM) and torsion (lateral PFM) [11]. In such a way, it is possible to reconstruct the three-dimensional real-space polarization direction in the material. An example of such Vector PFM [12] imaging of an epitaxial bismuth ferrite film is shown in Fig. 2. Note that torsion can occur on domains walls, and, e.g., even 180° domain walls in tetragonal (001)-oriented materials can be visible in lateral PFM amplitude images.

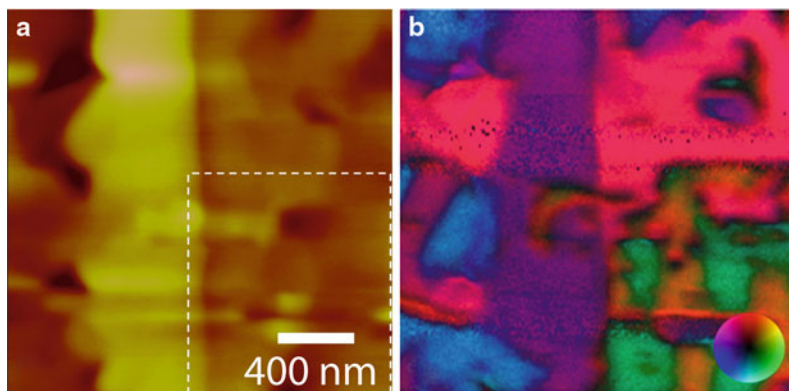
Domain Writing by PFM

In PFM imaging, an ac-bias is applied to the tip. This allows for local probing of the piezoelectric properties and for ferroelectric domain orientation imaging. The application of a dc-bias to the sample via the tip can cause local polarization reversal if the highly localized electric field at the tip exceeds the local coercive field of the material. Thus, PFM can be used for local domain writing and for subsequent imaging of the polarization reversal. Dense arrays of written domains have been envisioned for polarization-based information storage. Information bits can be written as small domains with opposite polarization, and the read cycle can be achieved using the same tip by measuring the amplitude and the phase of the piezoelectric response. These memory devices are expected to achieve high densities, limited only by the size of the written ferroelectric domains and the presence of possible cross talk [13].

Polar ferroelectric surfaces are characterized by the presence of surface charge and electric potential and can be used in a broad set of applications, including local molecular self-assembly, physical adsorption/desorption, chemical reactions, charge transfer reactions, and for the manipulation of individual organic molecules. Exploring the effect of ferroelectric polarization on chemical reactivity, an approach for assembling complex nanostructures consisting of oxide

Piezoresponse Force Microscopy and Spectroscopy, Fig. 2

(a) Topography and (b) vector PFM images of a BFO sample. The *dashed line* in (a) indicates a region subjected to a dc-bias prior to imaging. The *color wheel* in (b) indicates the ferroelectric polarization orientation and magnitude (Reproduced with permission from [26], copyright 2010, Wiley-VCH Verlag GmbH & Co. KGaA)



substrates, metal nanoparticles, and organic/biological molecules [14, 15] has been demonstrated (ferroelectric lithography). Controllable and selective deposition of molecular and metal species onto nanoscale domain-patterned ferroelectric templates may provide an alternative route to lithographic fabrication methods.

PFM Versus Other SPMs

The unique aspect of PFM is that for piezoelectric materials in the classical continuous mechanics regime, the electromechanical surface displacement is independent on the contact area. In other words, the displacement produced by the application of 1 V at the center of a 1 m, 1 mm, and 10 nm electrode will be the same. This behavior is fundamentally different compared to, e.g., contact mode mechanical AFMs, in which adhesion force, contact stiffness, and other parameters describing the properties of the tip–surface junction scale linearly with the tip–sample contact size. Correspondingly, on uneven surfaces, the effects of materials properties and topography variation cannot be unambiguously decoupled (direct topographic cross talk). In PFM, the response is intrinsically quantitative, albeit it can still be affected by the position-dependent transfer function of the cantilever (indirect crosstalk), the tip sliding along the sample surface, etc.

Similarly to other force-based SPM techniques, the signal-to-noise ratio in PFM can be improved if imaging is performed close to the contact resonance frequency of the cantilever. The quality factor of the

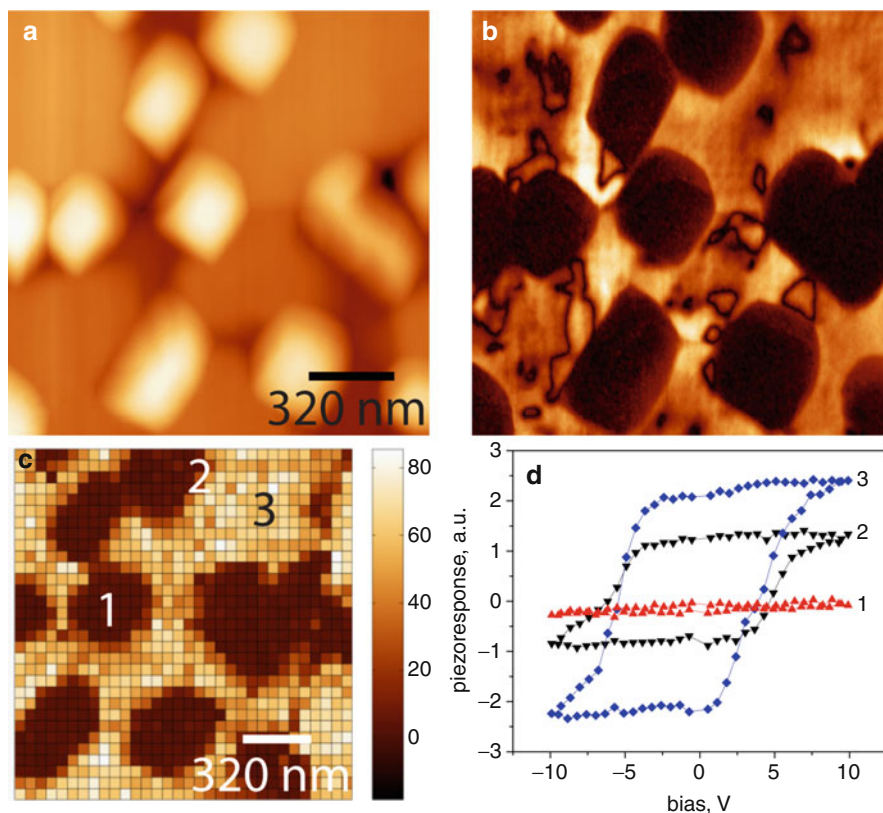
resonances can be as high as ~ 100 , allowing for a two order increase in signal strength compared to the off-resonance case. Quantitative PFM measurements can be performed if the frequency of the electric signal driving the cantilever is adjusted to follow the contact resonance at each point on the surface. However, classical strategies based on phase-locked loop detection are inapplicable to PFM, since the phase of the local response depends on the local polarization direction. In the past several years, techniques have been developed to enhance the signal-to-noise ratio in PFM and to quantify the piezoresponse signal, including dual amplitude resonance-tracking (DART) PFM [16], and band excitation (BE) PFM [17]. In DART-PFM, two ac-biases are sent to the tip, with frequencies chosen such that one is on each side of the contact resonance. A feedback loop is employed to keep the difference in the amplitude measured at each frequency equal to zero, allowing the contact resonance to be tracked, i.e., remain in-between the two frequencies during scanning. In BE-PFM, the cantilever is electrically excited across a band of frequencies, typically chosen to contain the contact resonance peak at all surface locations, allowing the resonant frequency and the quality factor of the system to be measured directly.

PFM Spectroscopies

PFM can be extended to spectroscopic modes, in which the response is measured as a function of (1) time during or after the application of a bias pulse (time spectroscopy), (2) bias during a triangular sweep

Piezoresponse Force Microscopy and Spectroscopy,

Fig. 3 (a) Topography of BFO-CFO and (b) corresponding PFM amplitude. (c) SS-PFM map of the work of switching and (d) representative loops from CFO (1), the CFO-BFO heterostructure interface (2), and BFO (3) (Reproduced with permission from [27], copyright 2007, Institute of Physics)



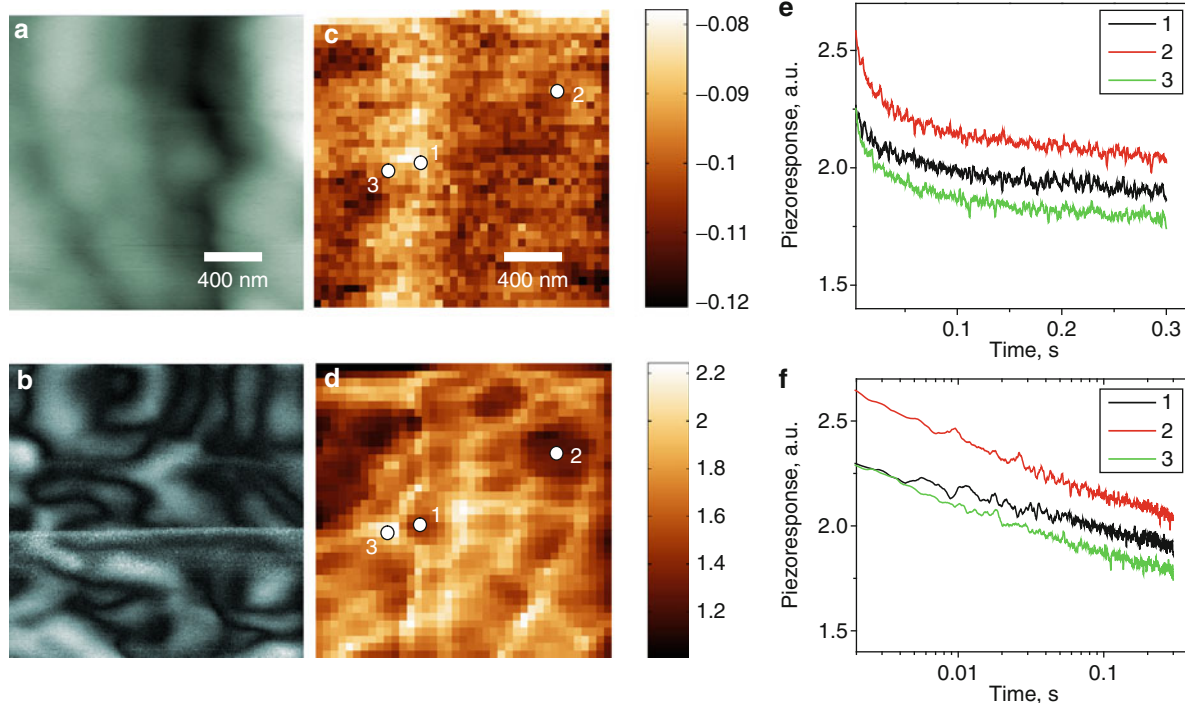
(voltage spectroscopy), and (3) ac-bias amplitude. More complex measurement sequences (e.g., first-order reversal curve (FORC) measurements) have also been demonstrated. These spectroscopic modes can in turn be implemented as spectroscopic imaging modes (similar to, e.g., force volume mode imaging in AFM), in which the spectroscopic sweeps are performed at every point in an image.

Switching Spectroscopy Piezoresponse Force Microscopy

In switching spectroscopy PFM (SS-PFM) the ac-bias is superimposed onto a dc-bias waveform, which cycles between a maximum and minimum dc voltage value (i.e., a triangular voltage sweep). SS-PFM allows local piezoelectric hysteresis loops to be measured at each pixel in an image, providing a nanoscale analog to macroscopic ferroelectric polarization-voltage loops. Each loop can be analyzed to determine the width of the loop, the nucleation biases, the remanent polarization, etc. These extracted switching parameters can then be plotted as two-dimensional maps, allowing correlations between switching properties and

local topography to be identified. Note that SS-PFM can be operated in single frequency, DART, and BE-PFM modes.

As an example of the SS-PFM imaging mode, maps of piezoelectric response and switching properties for $\text{BiFeO}_3\text{-CoFe}_2\text{O}_4$ (BFO-CFO) heterostructures are shown in Fig. 3. The BFO film is ferroelectric, whereas the CFO nanopillars are non-ferroelectric (and ferromagnetic). The ferroelectric behavior is clearly seen in the PFM amplitude image in Fig. 3b, in which the BFO regions are bright (i.e., have high piezoresponse) and show clearly visible domain walls, whereas the response from the CFO pillars is negligible. In the SS-PFM mode, the hysteresis loops are acquired over the same region on a rectangular grid. Shown in Fig. 3c is the work of switching map (i.e., the area under the hysteresis loop). Note that the loops are open (Fig. 3d) when measured on the BFO matrix and are closed when measured on the CFO nanopillars. The SS-PFM image illustrates that the polarization switching within BFO is uniform, indicative of a highly homogeneous material, which is not affected by proximity to the second-phase inclusions.



Piezoresponse Force Microscopy and Spectroscopy, Fig. 4 (a) Topography and (b) PFM amplitude images of a PMN-PT ferroelectric relaxor crystal. Spatially resolved maps of (c) slope, B_1 and (d) offset, B_0 of the logarithmic

relaxation law. Relaxation curves in selected locations in (e) linear and (f) logarithmic scale (Reproduced with permission from [28], copyright 2009, American Institute of Physics)

Time Spectroscopy in PFM

PFM can be extended to study time-dependent dynamic properties of materials. In time-resolved (TR)-PFM, the local piezoelectric response is measured during the application of a dc-bias pulse and upon the completion of the pulse. As with SS-PFM, the response curves can be modeled to reveal information about local relaxation behavior of polymers and ferroelectric relaxors. Note that TR-PFM can also be operated in single frequency, DART, and BE-PFM modes.

An example of TR-PFM is shown in Fig. 4 for the ferroelectric PMN-PT (solid solution of lead magnesium niobate [PMN] and lead titanate [PT]) relaxor. Here, the polarization response is collected at each point following the application of a voltage pulse. The response curves are fitted by the logarithmic function, $PR(t) = B_0 + B_1 \ln t$, and the spatial maps of offset, $B_0(x, y)$, and slope, $B_1(x, y)$, are shown in Fig. 4e, f, respectively. Note that both maps indicate the presence of mesoscopic structures and contain

a number of uncorrelated features, indicative of the validity of the analysis and demonstrating the presence of intrinsic *dynamic* inhomogeneities within the material.

More Complex Spectroscopies

The characteristic aspect of hysteretic systems is that the response is generally history dependent. For example, for a linear or nonlinear resistor, each value of current corresponds to only one value of applied external voltage. In a hysteretic system, the instantaneous value of current depends on the voltage history. However, the knowledge of the hysteresis loop for one type of field sweep does not allow for the prediction of a different one – e.g., knowledge of the hysteresis behavior in the $(-10, 10)$ V bias window is insufficient to describe the $(0, 10)$ or $(-5, 10)$ bias windows. For certain classes of systems, this limitation can be bypassed using FORC measurements that allow one to reconstruct the response for an arbitrary field history. Another aspect of these materials is that the

response can be time dependent, necessitating the study of the frequency or time dependence of the signal. To address these phenomena locally, more complicated time- and voltage-dependent spectroscopies in PFM are being developed.

PFM of Functional Materials

The materials that can be imaged by PFM are by definition functional materials in that they exhibit electromechanical coupling – they mechanically deform in an applied field. The classes of materials that have been investigated by PFM include, but are not limited to ferroelectrics, biopolymers and polysaccharides, and piezoelectric polymers.

Ferroelectrics

Ferroelectric materials are a subset of pyroelectric materials, which in turn are a subset of piezoelectric materials. Piezoelectric materials exhibit linear electromechanical coupling, while pyroelectric materials have a permanent spontaneous polarization. In ferroelectric materials, the direction of this spontaneous polarization can be changed, i.e., switched between stable directions corresponding to different ferroelectric domains. These domains are readily imaged by PFM in various ferroelectric families such as perovskites, Aurivillius-layered structures, tungsten bronze ferroelectrics, hexagonal manganites, organic materials and, recently, in a number of multiferroics (i.e., materials having both ferroelectricity and ferromagnetism simultaneously) and various composites [8, 18]. These measurements are indispensable for understanding the functionality of ferroelectrics at the nanoscale (e.g., within a single grain in polycrystalline ferroelectrics) and various size effects in as-grown and structured materials [19]. A ferroelectric relaxor is a special case of a ferroelectric material where long-range polarization is disrupted by defects and very small mesoscopic structures are observed on the surface on a length scale much smaller than the “normal” micron-size ferroelectric domains in these materials. These properties have recently been accessed by time-, temperature-, and depth-dependent PFM measurements, demonstrating the power and capabilities of PFM to uncover the polarization mechanism in complex materials with coexisting order parameters and disorder [20].

Biosystems

The piezoelectric properties of biopolymers and polysaccharides have been studied on the macroscopic scale for some time, and several researchers have speculated on a relation between piezoelectricity in these materials and biological processes such as bone remodeling. PFM is ideally suited for such studies and has been applied to a wide variety of biomaterials. These measurements are found to be especially important in macromolecules such as collagen fibers and, recently, in bioinspired peptide nanotubes [21]. The exceptional piezoelectric properties observed by PFM of the latter has renewed the interest in artificial biomaterials for biological sensor and actuator-based applications. PFM has also been implemented as a high-resolution nanoscale tool for uncovering new features in complex biological materials, such as cells and tissues.

Polymers

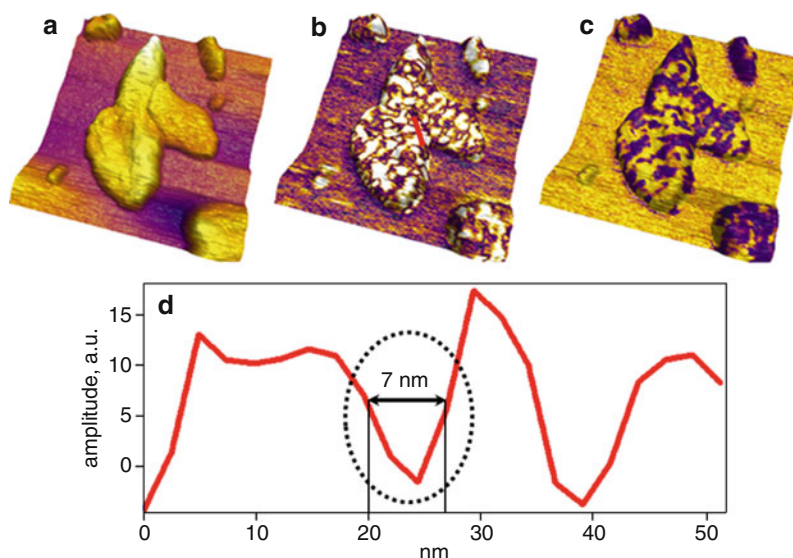
Many synthetic polymers, including polypropylene, polystyrene and poly(methyl methacrylate), semicrystalline polyamides, and amorphous polymers such as vinyl acetate are piezoelectric. However, piezoelectric effects in these materials are relatively weak, often unstable and, therefore, are of limited practical significance. Strong and stable piezoelectric properties have primarily been found in the synthetic polymer polyvinylidene fluoride (PVDF) and PVDF copolymers. PFM has been found to be extremely useful for studying local piezoelectric properties in PVDF because it allows one to relate the surface topography to the local polarization response in these semicrystalline materials where a mixture of material phases is often observed. It has also been demonstrated that PFM can effectively probe nanostructured PVDF nanomesas (Fig. 5). Studying polarization switching in these soft ferroelectric materials is one of the challenges for PFM because of the possible coupling present between electrostatics, piezoelectricity, and elasticity.

PFM of Device Structures

PFM is essentially a nanoscale polarization probe and, as such, can be used to study ferroelectric devices based on polarization dynamics and switching, including nonvolatile ferroelectric random access memories

Piezoresponse Force Microscopy and Spectroscopy, Fig. 5

(a) Topography, (b) PFM amplitude, and (c) PFM phase images of a PVDF nanomesa. (d) The resolution in (b) is found to be 7 nm (Reproduced with permission from [29], copyright 2011, American Chemical Society)



(NV-FRAM), thin-film capacitors, piezoelectric elements for MEMS, and novel multiferroic devices with multiple memory states [22]. As for NV-FRAM, the bit value is represented by one of the two distinguishable remanent polarization states of a ferroelectric material in a capacitor configuration. The spatial variation of the ferroelectric properties, e.g., the static and dynamic domain configurations, becomes critical as the NV-FRAM capacitor size decreases. If the ferroelectric properties of a material vary on a micron scale, then micron-sized capacitors would exhibit a significant capacitor-to-capacitor variation in their electrical properties, often resulting in a failure of the device. These properties have been extensively studied by a number of groups [23–25] and it has been shown that the polarization dynamics differ significantly as the lateral size of the ferroelectric capacitors is scaled down. Analogously, ferroelectric-based MEMS have been found to be sensitive to scaling in both lateral and vertical dimensions. This makes PFM a useful industrial tool, e.g., for quality control in the production of memory chips and micromechanical devices based on the piezoelectric effect.

Electrochemical Strain Microscopy

Bias-strain coupling can be driven by electrochemical mechanisms as a consequence of ionic motion in

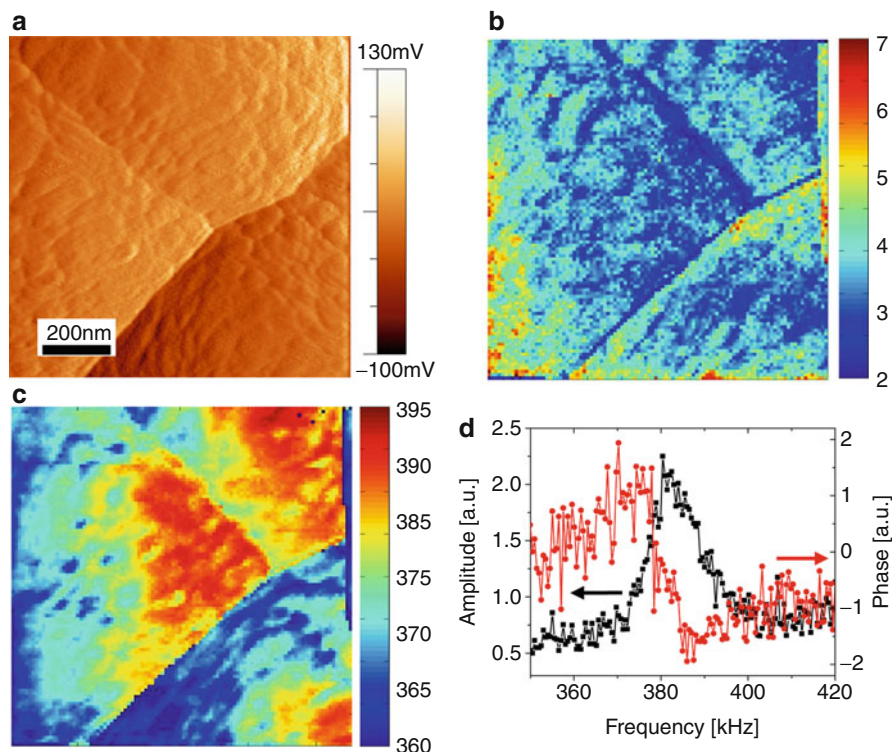
applied electric fields. PFM can be used to probe ac-bias-induced ionic motion in energy materials, giving rise to electrochemical strain microscopy (ESM). In ESM, a biased SPM tip concentrates an electric field in a nanometer-scale volume of the material, inducing interfacial electrochemical processes at the tip–surface junction and ionic currents through the solid. The intrinsic link between the concentration of ionic species and/or the oxidation states of the host cation and the molar volume of the material, results in electrochemical strain and surface displacement.

The ESM imaging mode is based on detecting the strain response of a material to an applied electric field through a blocking or electrochemically active SPM tip. The high-resolution (<10 nm) data shown in Fig. 6 demonstrates ESM imaging on a Li-intercalated Si-anode surface. A full suite of spectroscopic imaging modes can also be implemented in ESM. To date, ESM has been demonstrated for a variety of lithium-ion materials, including layered transition metal oxide cathodes, silicon anodes, and electrolytes such as LISICON; oxygen electrolytes, including yttria-stabilized zirconia and samarium-doped ceria; mixed electronic-ionic conductors for fuel cell cathodes; and some proton conductors.

The ability to probe electrochemical processes and ionic transport in solids is invaluable for a broad range of applications for energy generation and storage

Piezoresponse Force Microscopy and Spectroscopy, Fig. 6

(a) Deflection signal of a $1 \times 1 \mu\text{m}^2$ area showing a triple boundary point. (b) Contact resonance amplitude. (c) Resonance frequency map showing heterogeneous strain response and a strong correlation between resonance frequency and topography. (d) A single-point contact resonance peak from the boundary region



(e.g., batteries and fuel cells). The viability of electric vehicles and grid storage as key components of a renewable energy technology hinges on advances in battery energy densities and lifetimes. ESM has the potential to aid in these advances with two major improvements over existing technologies: (a) the resolution to probe nanometer-scale volumes and (b) imaging capability extended to a broad range of spectroscopy techniques. In addition, because electrochemical strains are ubiquitous in virtually all solid-state ionics, ESM is applicable to all battery and fuel cell materials in energy technologies and electroresistive/memristive materials in information technologies.

Summary

Electromechanical coupling is ubiquitous in systems ranging from piezoelectrics and ferroelectrics to polar macromolecules and solid-state ionics. Piezoresponse force microscopy and spectroscopy have opened these phenomena to scientific inquiry, providing advances ranging from sub-10 nm domain imaging and lithography, to probing intricate polarization

structures in ferroelectric relaxors, polar polymers and biosystems, to revealing details of electrochemical processes in battery materials. The future of this field clearly holds many more serendipitous discoveries.

Acknowledgment This research was supported in part (SVK) by the Center for Nanophase Materials Sciences, which is sponsored at Oak Ridge National Laboratory by the Office of Basic Energy Sciences, US Department of Energy. BJR gratefully acknowledges support from Science Foundation Ireland (grant no. 10/RFP/MTR2855) and UCD Research. ALK acknowledges Portuguese projects PTDC/CTM/73030/2006 and PTDC/FIS/81442/2006 for financial support. The work was partly done within the EU Atlantic Area Interregional project, Enermat.

Cross-References

- ▶ [AFM](#)
- ▶ [AFM in Liquids](#)
- ▶ [AFM Probes](#)
- ▶ [Friction Force Microscopy](#)
- ▶ [Kelvin Probe Force Microscopy](#)
- ▶ [Piezoelectric Effect at Nanoscale](#)
- ▶ [Scanning Tunneling Microscopy](#)

References

- Günther, P., Glatz-Reichenbach, J., Dransfeld, K.: Investigation of local piezoelectric properties of thin copolymer films. *J. Appl. Phys.* **69**, 7895 (1991)
- Birk, H., Glatz-Reichenbach, J., Jie, L., Schreck, E., Dransfeld, K.: The local piezoelectric activity of thin polymer films observed by scanning tunneling microscopy. *J. Vac. Sci. Technol. B* **9**, 1162 (1991)
- Kolosov, O., Gruverman, A., Hatano, J., Takahashi, K., Tokumoto, H.: Nanoscale visualization and control of ferroelectric domains by atomic force microscopy. *Phys. Rev. Lett.* **74**, 4309 (1995)
- Gruverman, A., Auciello, O., Tokumoto, H.: Imaging and control of domain structures in ferroelectric thin films via scanning force microscopy. *Annu. Rev. Mater. Sci.* **28**, 101 (1998)
- Alexe, M., Gruverman, A. (eds.): *Nanoscale Characterization of Ferroelectric Materials*. Springer, Heidelberg (2004)
- Hong, S. (ed.): *Nanoscale Phenomena in Ferroelectric Thin Films*. Kluwer, Dordrecht (2004)
- Kalinin, S.V., Gruverman, A. (eds.): *Scanning Probe Microscopy of Electrical and Electromechanical Phenomena at the Nanoscale*. Springer, Berlin (2007)
- Gruverman, A., Kholkin, A.: Nanoscale ferroelectrics: processing, characterization and future trends. *Rep. Prog. Phys.* **69**, 2443 (2006)
- Kalinin, S.V., Rodriguez, B.J., Jesse, S., Karapetian, E., Eliseev, E.A., Morozovska, A.N.: Nanoscale electromechanics of ferroelectric and biological systems: a new dimension in scanning probe microscopy. *Annu. Rev. Mat. Res.* **37**, 189 (2007)
- Kalinin, S.V., Morozovska, A.N., Chen, L.-Q., Rodriguez, B.J.: Local polarization dynamics in ferroelectric materials. *Rep. Prog. Phys.* **73**, 056502 (2010)
- Eng, L.M., Abplanalp, M., Günter, P.: Mapping the domain distribution at ferroelectric surfaces by scanning force microscopy. *Appl. Phys. A* **66**, S679 (1998)
- Kalinin, S.V., Rodriguez, B.J., Jesse, S., Shin, J., Baddorf, A.P., Gupta, P., Jain, H., Williams, D.B., Gruverman, A.: Vector piezoresponse force microscopy. *Microscopy Microanalysis* **12**, 206 (2006)
- Tybell, T., Paruch, P., Giamarchi, T., Triscone, J.-M.: Domain wall creep in epitaxial ferroelectric $\text{Pb}(\text{Zr}_{0.2}\text{Ti}_{0.08})\text{O}_3$ thin films. *Phys. Rev. Lett.* **89**, 097601 (2002)
- Kalinin, S.V., Bonnell, D.A., Alvarez, T., Lei, X., Hu, Z., Ferris, J.H., Zhang, Q., Dunn, S.: Atomic polarization and local reactivity on ferroelectric surfaces: a new route toward complex nanostructures. *Nano Lett.* **2**, 589 (2002)
- Heredia, A., Bdikin, I.K., Machado, M., Yudin, S., Fridkin, V.M., Delgadillo, I., Kholkin, A.L.: Preferred deposition of phospholipids onto ferroelectric P(VDF-TrFE) films via polarization patterning. *J. Phys. D* **43**, 335301 (2010)
- Rodriguez, B.J., Callahan, C., Kalinin, S.V., Proksch, R.: Dual-frequency resonance-tracking atomic force microscopy. *Nanotechnology* **18**, 475504 (2007)
- Jesse, S., Kalinin, S.V., Proksch, R., Baddorf, A.P., Rodriguez, B.J.: The band excitation method in scanning probe microscopy for rapid mapping of energy dissipation on the nanoscale. *Nanotechnology* **18**, 435503 (2007)
- Balke, N., Bdikin, I.K., Kalinin, S.V., Kholkin, A.L.: Electromechanical imaging and spectroscopy of ferroelectric and piezoelectric materials: state-of-the-art and prospects for the future. *J. Amer. Ceram. Soc.* **92**, 1629 (2009)
- Kholkin, A.L., Bdikin, I.K., Kiselev, D.A., Shvartsman, V.V., Kim, S.-H.: Nanoscale piezoelectric characterization of polycrystalline ferroelectrics. *J. Electroceramics* **19**, 81 (2007)
- Kholkin, A.L., Kiselev, D.A., Bdikin, I.K., Rodriguez, B.J., Wu, P., Bokov, A.A., Ye, Z.-G., Dkhil, B., Chen, L.-Q., Kosec, M., Kalinin, S.V.: Surface domain structures and mesoscopic phase transition in relaxor ferroelectrics. *Adv. Func. Mat.* **21**, 1977 (2011)
- Kholkin, A.L., Amdursky, N., Bdikin, I., Rosenman, G., Gazit, E.: Strong piezoelectric activity in peptide nanotubes. *ACS Nano* **4**, 610 (2010)
- Kalinin, S.V., Setter, N., Kholkin, A.L.: Electromechanics on the nanometer scale: emerging phenomena, devices, and applications. *MRS Bulletin* **34**, 634 (2009)
- Gruverman, A., Rodriguez, B.J., Kingon, A.I., Nemanich, R.J., Cross, J.S., Tsukada, M.: Spatial inhomogeneity of imprint and switching behavior in ferroelectric capacitors. *Appl. Phys. Lett.* **82**, 3071 (2003)
- Stolichnov, I.A., Tagantsev, A.K., Setter, N., Cross, J.S., Tsukada, M.: Top-interface-controlled switching and fatigue endurance of $(\text{Pb}, \text{La})(\text{Zr}, \text{Ti})\text{O}_3$ ferroelectric capacitors. *Appl. Phys. Lett.* **74**, 3552 (1999)
- Kim, D.J., Jo, J.Y., Kim, T.H., Yang, S.M., Chen, B., Kim, Y.S., Noh, T.W.: Observation of inhomogeneous domain nucleation in epitaxial $\text{Pb}(\text{Zr}, \text{Ti})\text{O}_3$ capacitors. *Appl. Phys. Lett.* **91**, 132903 (2007)
- Rodriguez, B.J., Choudhury, S., Chu, Y.H., Bhattacharyya, A., Jesse, S., Seal, K., Baddorf, A.P., Ramesh, R., Chen, L.-Q., Kalinin, S.V.: Unraveling deterministic mesoscopic polarization switching mechanisms: spatially resolved studies of a tilt grain boundary in bismuth ferrite. *Advanced Functional Materials* **19**, 2053 (2009)
- Rodriguez, B.J., Jesse, S., Baddorf, A.P., Zhao, T., Chu, Y. H., Ramesh, R., Eliseev, E.A., Morozovska, A.N., Kalinin, S. V.: Spatially resolved mapping of ferroelectric switching behavior in self-assembled multiferroic nanostructures: strain, size, and interface effects. *Nanotechnology* **18**, 405701 (2007)
- Kalinin, S.V., Rodriguez, B.J., Jesse, S., Morozovska, A.N., Bokov, A.A., Ye, Z.-G.: Spatial distribution of relaxation behavior on the surface of a ferroelectric relaxor in the ergodic phase. *Appl. Phys. Lett.* **95**, 142902 (2009)
- Sharma, P., Reece, T.J., Ducharme, S., Gruverman, A.: High-resolution studies of domain switching behavior in nanostructured ferroelectric polymers. *Nano Lett.* **11**, 1970 (2011)

Plasma

► Physical Vapor Deposition

Plasma Etching

- ▶ [Dry Etching](#)

Plasma Grafting

- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)

Plasma Polymerization

- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)

Plasma-Enhanced Chemical Vapor Deposition (PECVD)

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

Plasmon Resonance Energy Transfer from Metallic Nanoparticles to Biomolecules

- ▶ [Plasmon Resonance Energy Transfer Nanospectroscopy](#)

Plasmon Resonance Energy Transfer Nanospectroscopy

Gang Logan Liu

Micro and Nanotechnology Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Synonyms

[Plasmon resonance energy transfer from metallic nanoparticles to biomolecules](#)

Definition

Plasmon resonance energy transfer is the energy stored in the collective movement of free electrons in metallic nanoparticles being transferred to the adsorbed chemical and biomolecules with match electronic transition energy.

Overview

Nanoparticle Plasmon Resonance

A plasmon is defined as the collective and periodic free electron movement in metallic structures under electromagnetic excitation. Free electrons in the metal conduction band are not bounded to the nucleus, so they can move in the applied alternating electromagnetic field within the physical boundary of metal materials producing an alternating electron current. Like the resonance modes in mechanical movements, there exist plasmon resonance modes with minimal energy damping. If the plasmon resonance modes fall into the optical frequency range, especially the visible frequency range, the resonating free electrons in metallic nanostructures form an optical dipole to reemit light.

The plasmon resonance modes are dependent on the dielectric constants of the metallic nanoparticle and the surrounding material. Within the optical frequency range, the complex dielectric constants or the permittivity of the metallic materials such as gold and silver changes rapidly with the light wavelength while the dielectric constants of dielectric materials remain nearly unchanged. The plasmonic polarizability has to be modified when the metallic structure approaches sub-100 nm scale. At this scale, the dimension of the metallic nanostructure becomes comparable to the mean free path length of the conduction electrons in metal. The physical boundary of the metallic nanostructures poses a boundary condition for the plasmon resonance mode. Nanoplasmonic structures are thus referred to the metallic nanostructures such as nanoparticles with discrete plasmon resonance modes which are dependent on the nanostructure geometries and surrounding medium.

Plasmon Resonance Energy Transfer

Nanoparticle plasmon resonance [1, 2] is distinctive from the propagating surface plasmon resonance on metallic thin film [3]. It is spatially confined within the physical boundary of the nanoparticle; however it can

be continuously transferred to adjacent metallic nanoparticles [4, 5] through plasmon coupling. It has also been conjectured for a long time that the plasmon resonance energy can be possibly transferred to chemical or biological molecules adsorbed on metallic nanostructures [6, 7]. This entry will review the highest-sensitivity-ever nanoscopic biomolecular absorption spectroscopy enabled by plasmon resonance energy transfer (PRET) from a single metallic nanoparticle to a conjugated biomolecule and chemical molecule. Due to PRET, quantized plasmon quenching dips are observed in single nanoparticle scattering spectra which correspond to the absorption spectral peaks of conjugated biomolecules. The hybrid nanoparticle–biomolecule PRET system allows near single molecular sensitivity absorption spectroscopy with nanoscale spatial resolution for in vivo functional molecular imaging.

Although not directly observed before, PRET was accounted for one of possible explanations for surface-enhanced Raman scattering [8, 9] and fluorescence [10] on single nanoparticles. Experimentally the plasmon resonance of gold and silver nanoparticles conjugated with various biomolecules such as DNA [11], peptide [12], and biotin-streptavidin [13] has been studied by single particle Rayleigh scattering spectroscopy [14, 15]. All these previous studies demonstrated the shift of plasmon resonant wavelength by changing dielectric medium due to structural changes of biomolecule conjugated on the surface of single metallic nanoparticles. Since most of the previous cases have conjugated biomolecules with optical absorption or electronic resonance peaks in ultraviolet (UV) or far-infrared range on gold and silver nanoparticles with visible plasmon resonance peaks, only the shift of plasmon resonance peak was observed. However in this chapter, metalloprotein Cytochrome c (Cyt c) was conjugated on a single 30 nm gold nanoparticle in order to observe the direct quantized plasmon resonance energy transfer from the nanoplasmonic particle to Cyt c, as seen in Fig. 1.

Theory

The strict modeling of plasmon resonance energy transfer between metallic nanoparticle and adsorbed resonance chemical and biological molecules requires quantum mechanical calculation

of electron interaction Hamiltonian between nanoparticle and molecules.

$$H_I = \frac{1}{4\pi\epsilon_0} \times \int dr_{\text{molecule}} \int d^3r_{\text{metal}} \frac{\rho(r_{\text{molecule}})\rho(r_{\text{metal}})}{|R - r_{\text{metal}} + r_{\text{molecule}}|} \quad (1)$$

The above equation describes the interaction Hamiltonian where ρ is the electron density distribution in molecule or metallic nanoparticle and r is the radial distance of the electrons from the center of the molecule or nanoparticle and R is the distance between the molecule center and nanoparticle center. Apparently it is extremely difficult to precisely describe the three-dimensional spatial distribution of electron density; therefore some assumption and approximation have to be made to carry out the analytical calculation. Fortunately near the matched nanoparticle plasmon resonance and molecule electronic resonance the electron distribution may be approximated as dipolar distributions. Then the electron interaction becomes dipole–dipole interactions as described in Eq. 2.

$$H_I^{\text{dip}} = \frac{1}{4\pi\epsilon_0} \times \left[\frac{\mu_{\text{molecule}} \cdot \mu_{\text{metal}} - 3(\mu_{\text{molecule}} \cdot \hat{d})(\mu_{\text{metal}} \cdot \hat{d})}{d^3} \right] \quad (2)$$

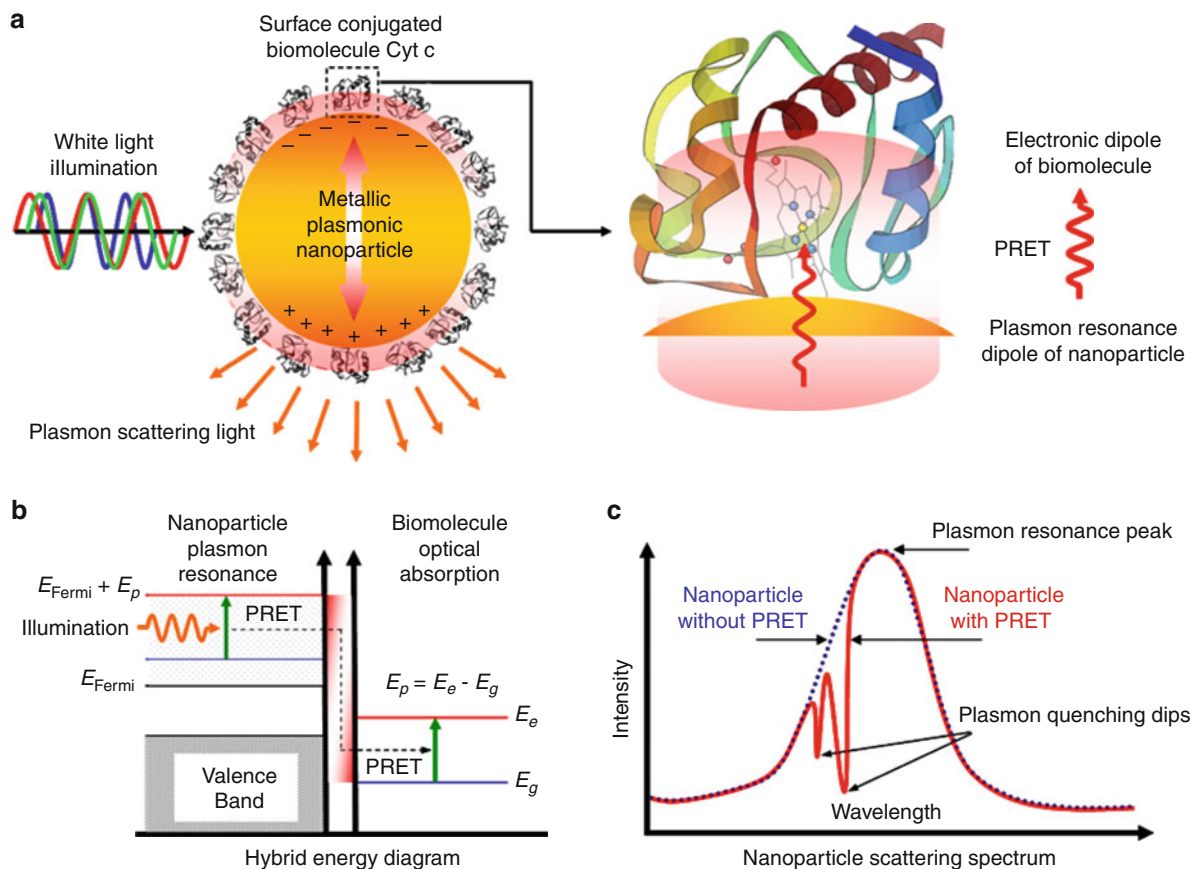
Here μ is the dipole moment operator and defined as

$$\begin{aligned} \mu_{\text{metal}} &= \int r_{\text{metal}} \rho_{1,0}(r_{\text{metal}}) d^3r_{\text{metal}} \\ \mu_{\text{molecule}} &= \int r_{\text{molecule}} \rho(r_{\text{molecule}}) dr_{\text{molecule}} \end{aligned} \quad (3)$$

Where ρ is simplified as the radial distribution of the electron density and especially for metallic nanoparticle $\rho_{1,0}$ becomes the dipolar plasmon charge density which is defined as

$$\nabla^2 \rho_{1,0}(r_{\text{metal}}) + \alpha^2 \rho_{1,0}(r_{\text{metal}}) = 0 \quad (4)$$

The constant α is the function of fundamental plasmon resonance mode frequency $\omega_{1,0}$ and bulk plasmon resonance frequency ω_p as well as the



Plasmon Resonance Energy Transfer Nanospectroscopy,

Fig. 1 Schematic diagram of PRET-enabled nanoscopic biomolecular absorption spectroscopy. (a) PRET from a single metallic nanoparticle to surface-conjugated biomolecules. The wavelength-specific plasmon resonance (collective free electron oscillation) in metallic nanoparticle is excited by white light illumination. The plasmon resonance dipole can interact with the biomolecular dipole and transfer energy to biomolecules. (b) Hybrid energy diagram showing quantized energy transfer process. With optical excitation, the free electrons in the

conduction band of metallic nanoparticle are elevated from Fermi to higher energy level forming resonating plasmon. The plasmon resonance energy is transferred to the metalloprotein biomolecules (i.e., Cyt c) conjugated on the nanoparticle surface when matched with the electronic transition energy in biomolecule optical absorption. (c) Typical Rayleigh scattering spectrum of the single PRET probe. The energy transition in PRET is represented as quenching dips in nanoparticle scattering spectrum, and the dip positions correspond to the biomolecule optical absorption peaks

electron velocity at Fermi energy v_F . The expression is as the following:

$$\alpha^2 = \frac{\omega_{1,0}^2 - \omega_p^2}{3v_F^2/5} \quad (5)$$

The dipole moment operator of the metallic nanoparticle at the fundamental plasmon resonance mode is then

$$\mu_{\text{metal}} \approx \left(\frac{4\pi^3 \epsilon_0 \hbar \omega_p}{9\sqrt{3}} \right)^{1/2} \alpha^{3/2} \quad (6)$$

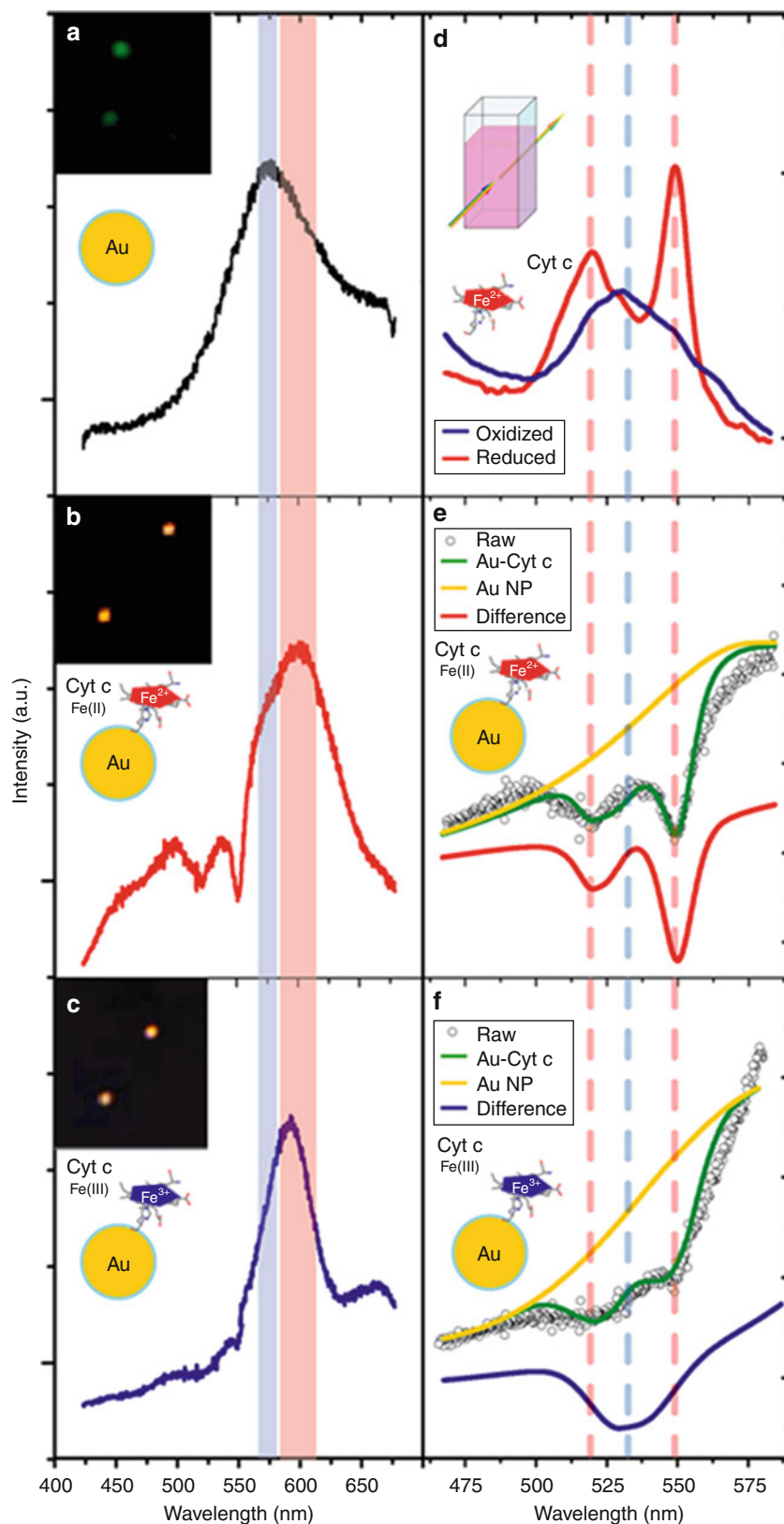
Basic Methodology

PRET-Enabled Biomolecular Absorption Spectroscopy

The intentional overlap of the absorption peak positions of desired biomolecules with the plasmon resonance peak of the metallic nanoparticle generates distinguishable spectral quenching dips on the Rayleigh scattering spectrum of single nanoparticle, which also allows near single molecular level nanoscopic absorption spectroscopy (Figs. 1 and 2). Cyt c, a metalloprotein in cellular mitochondria membrane, acts as the charge transfer mediator [16] and

Plasmon Resonance Energy Transfer Nanospectroscopy,

Fig. 2 Experimental results of PRET from single gold nanoparticle to conjugated Cyt c molecules. The Rayleigh scattering spectrum of a single gold nanoparticle coated with (a) only Cysteamine coating, (b) Cysteamine cross-linker and reduced Cyt c and (c) Cysteamine and oxidized Cyt c. The Rayleigh scattering spectrum was obtained using 1 s integration time. (d) The visible absorption spectra of Cyt c bulk solution in reduction form (blue solid line), and in oxidation form (red solid line) measured in conventional UV-vis absorption spectroscopy. (e) The fitting curve for the spectrum in (b). *Black open circle*: raw data, *Green solid line*: fitting curve of the raw data, *Yellow solid line*: Lorentzian scattering curve of bare gold nanoparticle, *Red solid line*: processed absorption spectra for the reduced conjugated Cyt c by subtracting yellow curve from the green curve. (f) The fitting curve for the spectrum in (c). *Black open circle*: raw data, *Green solid line*: fitting curve of the raw data, *Yellow solid line*: Lorentzian scattering curve of bare gold nanoparticle, *Blue solid line*: processed absorption spectra for the oxidized conjugated Cyt c by subtracting yellow curve from the green curve



plays a crucial role in bioenergy generation, metabolism, and cell apoptosis [17]. Unlike many other proteins, Cyt c has several optical absorption peaks in visible range around 550 nm coinciding with the 30 nm gold nanoparticle plasmon resonance, and more importantly it is a natural energy acceptor with electron tunneling channels [18]. Similar to the donor–acceptor energy matching in fluorescent (or Förster) resonance energy transfer (FRET) between two fluorophores, the critical matching of the localized resonating plasmon kinetic energy E_p in gold nanoparticles with the electron transition energy from ground to excited state $E_e - E_g$ in Cyt c molecules permits the PRET process (Fig. 1b). The quantized energy is likely transferred through the dipole–dipole interaction between the artificial alternating dipole – resonating plasmon in nanoparticle and the biomolecular dipole. Previous work on surface plasmon–mediated FRET process [19, 20], superlens imaging [21], surface plasmon resonance shift of redox molecules [22], and very recent work on bulk optical extinction spectroscopy of nanoplasmonic particle clusters with conjugated resonant molecules [23] also indicate the possibility of such dipole–dipole interaction. Distinctively in this experiment, the plasmon energy quenching of nanoparticle due to PRET is represented as the “spectral dips” in the single nanoparticle scattering spectrum (Figs. 1c and 2) and the positions of dips match with the molecular absorption peak positions (Figs. 1c and 2). PRET is a direct energy transfer process and thus much more efficient and faster than optical energy absorption [10], so the absorption spectral peaks of conjugated Cyt c molecules on single gold nanoparticles can be detected with a simple optical system, which is otherwise impossible using conventional visible absorption spectroscopic methods.

The Cyt c conjugated 30 nm gold nanoparticles are dispersedly tethered on the surface of a transparent glass slide. The glass slide is mounted on a white light darkfield microscopy system with a true-color camera and a spectrometer to characterize the scattering image and spectrum of individual gold nanoparticles as well as the hybrid PRET probes (i.e., specific metallic nanoparticles with conjugated Cyt c molecules).

In comparison with the visible scattering spectrum of gold nanoparticles coated with only Cysteamine cross-linker molecules (Fig. 2a), the raw scattering

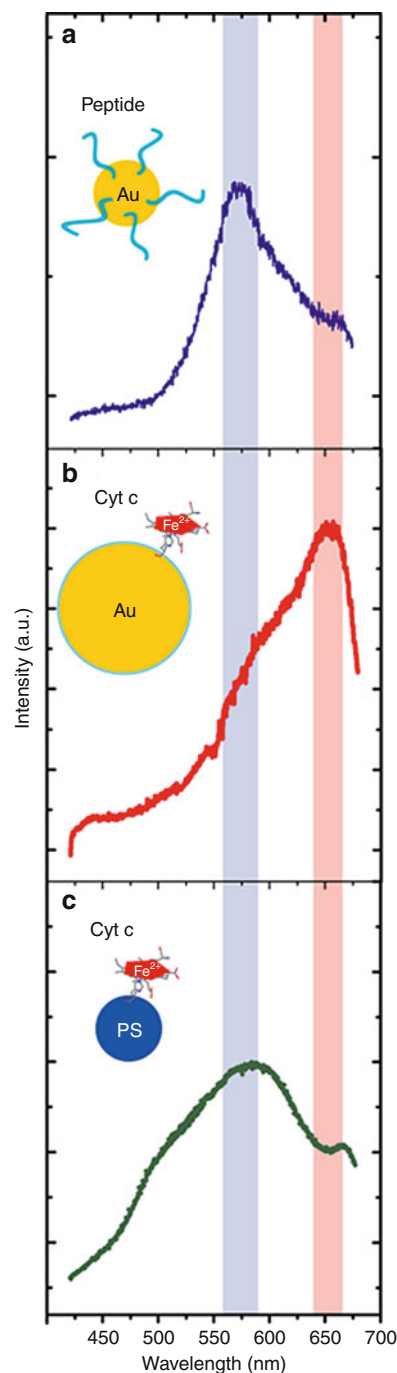
spectra of gold nanoparticles conjugated with reduced (Fig. 2b) and oxidized Cyt c (Fig. 2c) show not only a scattering peak (plasmon resonance peak) but distinctive dips next to it. The spectral dips modulated on the nanoparticle scattering spectrum can be decoupled and converted to the visible absorption peaks of the reduced and oxidized Cyt c molecules (Fig. 2e and f). In accordance with the conventional visible absorption spectrum of bulk Cyt c solutions (Fig. 2d), the processed spectra have matched absorption peaks of reduced Cyt c around 525 and 550 nm, and oxidized Cyt c at 530 nm [24]. Although the near-field optical excitation efficiency from the nanoparticle scattering light is much higher than far field optical excitation (i.e., the photon scattered from the nanoparticle is more likely to transmit through and be absorbed by the surface conjugated biomolecules than those far away from the nanoparticle), the optical absorption at 550 nm by the ferrocyclochrome c molecule monolayer only accounts for 0.03% of the nanoparticle scattering light even for 100% excitation efficiency (The absorption coefficient ϵ of horse heart Cyt c at 550 nm is $20.4 \text{ mM}^{-1} \text{ cm}^{-1}$. For a completely packed monolayer of ~ 3 nm-in-diameter Cyt c molecules on 30 nm gold nanoparticle surface, the local concentration C of Cyt c on single nanoparticle is 47 mM. The optical path length L of a Cyt c monolayer is ~ 3 nm, so the optical absorbance is $A = \epsilon CL \sim 3 \times 10^{-4}$ assuming 100% of scattering photon passing through the biomolecule monolayer.); therefore, the dramatic spectral dips are not a result of the direct optical absorption of Cyt c molecules.

The energy matching condition in PRET is further confirmed by three negative control experiments. For the first control experiment, synthesized peptides which have absorption peaks out of wavelength range of the plasmon resonance of a 30 nm gold nanoparticle are intentionally conjugated. As expected, the scattering spectrum of this hybrid system shows only the scattering peak because the absorption peaks of peptide do not coincide with the plasmon resonance spectrum of the nanoparticle (Fig. 3a). For the second control experiment, the importance of matching resonant frequency and molecular absorption peaks was tested by using a large gold nanoparticle cluster which has a plasmon resonance wavelength beyond 650 nm. As anticipated, the conjugated Cyt c absorption peaks can be hardly observed, including the 525 and 550 nm

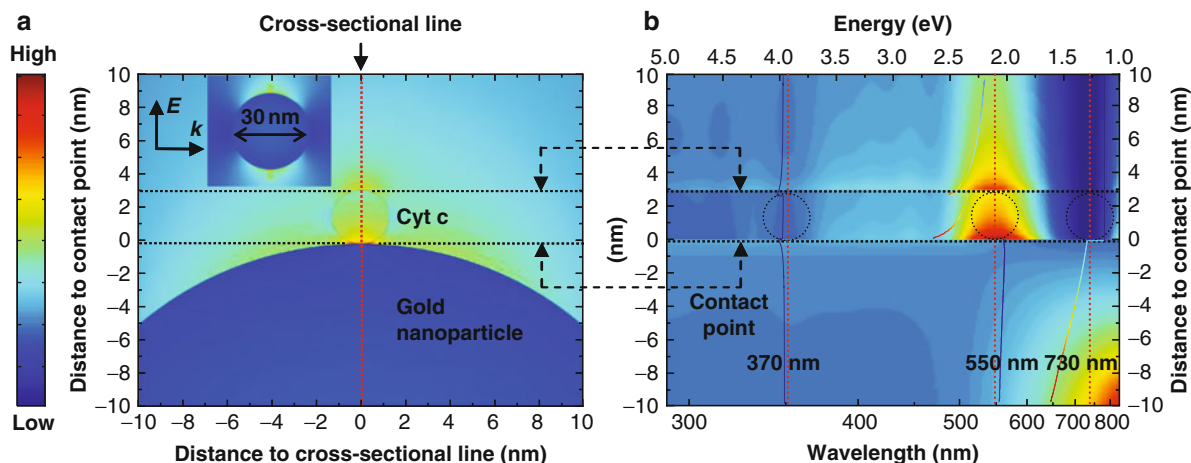
peaks for reduced Cyt c (Fig. 3b). For the third control experiment, dielectric polystyrene nanoparticles were conjugated with Cyt c and characterized. As expected, the plasmon quenching spectral dips cannot be found on the scattering spectrum of single dielectric polystyrene nanoparticle without plasmon resonance even though the dielectric nanoparticle also scatters light, which indicates the presence of excited free electrons is necessary for the PRET process (Fig. 3c).

The average surface density of Cyt c molecules on an individual gold nanoparticle is controlled by the molar concentration ratio used in the conjugation process. Considering the effective cross-sectional area of single Cyt c molecules and the surface area of single 30 nm gold nanoparticle, maximally around 400 Cyt c molecules can be tethered on a 30 nm gold nanoparticle. The scattering spectrum of many individual 30 nm nanoparticles was measured and extracted the reduced Cyt c visible absorption peaks. Due to the nonuniformity of the surface molecule numbers on each particle and nanoparticle plasmon resonance wavelength, the plasmon quenching or Cyt c absorption peak intensity shows variations from particle to particle (Fig. 4a), whereas the spectral measurement on each individual nanoparticle is repeatable and stable (Fig. 4b) and no photochemical changes are observed. Unpolarized white light source is used in all the above experiments.

Similar to the energy transfer process in FRET, the PRET efficiency is possibly dependent on the distance from the spectrally active moiety of biomolecules, e.g., Heme group for Cyt c, to the plasmonic nanoparticle surface as well as the relative orientations between the polarized plasmon resonance dipole and molecular dipoles (Fig. 5a). On the other hand, the simulated single nanoparticle PRET spectra (Fig. 5b) shows that the strongest plasmon resonance mode for a single Cyt c conjugated 30 nm gold nanoparticle occurs around 550 nm. This resonant frequency (i.e., energy) matching condition potentially explains why the plasmon quenching peak amplitude at 550 nm is relatively higher than at 525 nm (Fig. 2e) compared to the peak intensity ratio in Cyt c bulk solution absorption measurement (Fig. 2e). The scattering peak wavelength of the gold nanoparticle in experiments is higher than the simulated results due to larger numbers of Cyt c and Cysteamine molecules conjugated on surface.



Plasmon Resonance Energy Transfer Nanospectroscopy, Fig. 3 Negative control results showing the importance of critical energy matching for PRET. (a) The scattering spectrum of a 30 nm gold nanoparticle coated with Cys-(Gly-Hyp-Pro) [6] peptides. (b) The scattering spectrum of a single large gold nanoparticle cluster conjugated with Cysteamine and Cyt c. (c) The scattering spectrum of a 40 nm amine-modified polystyrene bead conjugated with Cyt c



Plasmon Resonance Energy Transfer Nanospectroscopy, Fig. 4 Simulation of nanoparticle plasmon resonance coupling to a single Cyt c molecule. (a) Time-averaged total electromagnetic (EM) energy at 550 nm vertically polarized light excitation around the interface of single 30 nm gold nanoparticle and single 3 nm spherical molecules. The dielectric nanosphere is used to simulate single reduced Cyt c molecule with a wavelength-dependent complex refractive index (16). The EM energy is transferred to the single molecule and forms the dipolar energy distribution across the molecule. The inset image of the whole nanoparticle shows the energy coupling only occurs in the light

polarization direction. (b) Time-averaged total energy profile at the cross-section line in (a) as the function of the excitation wavelength or energy. The energy distribution at each wavelength is normalized to the energy at the inner surface of the nanoparticle. The representative line plots of the energy profile at 370, 550, and 730 nm are superposed on the 2D color-coded energy distribution at corresponding wavelength positions. The dipolar energy of nanoparticle plasmon resonance is electromagnetically coupled to the single molecule around 550 nm, while at other wavelengths much less energy transfer is observed

The PRET effect can be also observed for hemoglobin molecules conjugated with single silver nanoparticles. Hemoglobin molecule has a distinctive absorption peaks around 407 nm which is the sorbet band of the heme group. The absorption spectra of hemoglobin with and without binding to oxygen are shown in Fig. 6. Using PRET nanosensor, plasmon quenching dips were observed and corresponding to the Soret band (407 nm) of hemoglobin molecules on the surface of a single nanoparticle, as shown in Fig. 7.

Key Research Findings

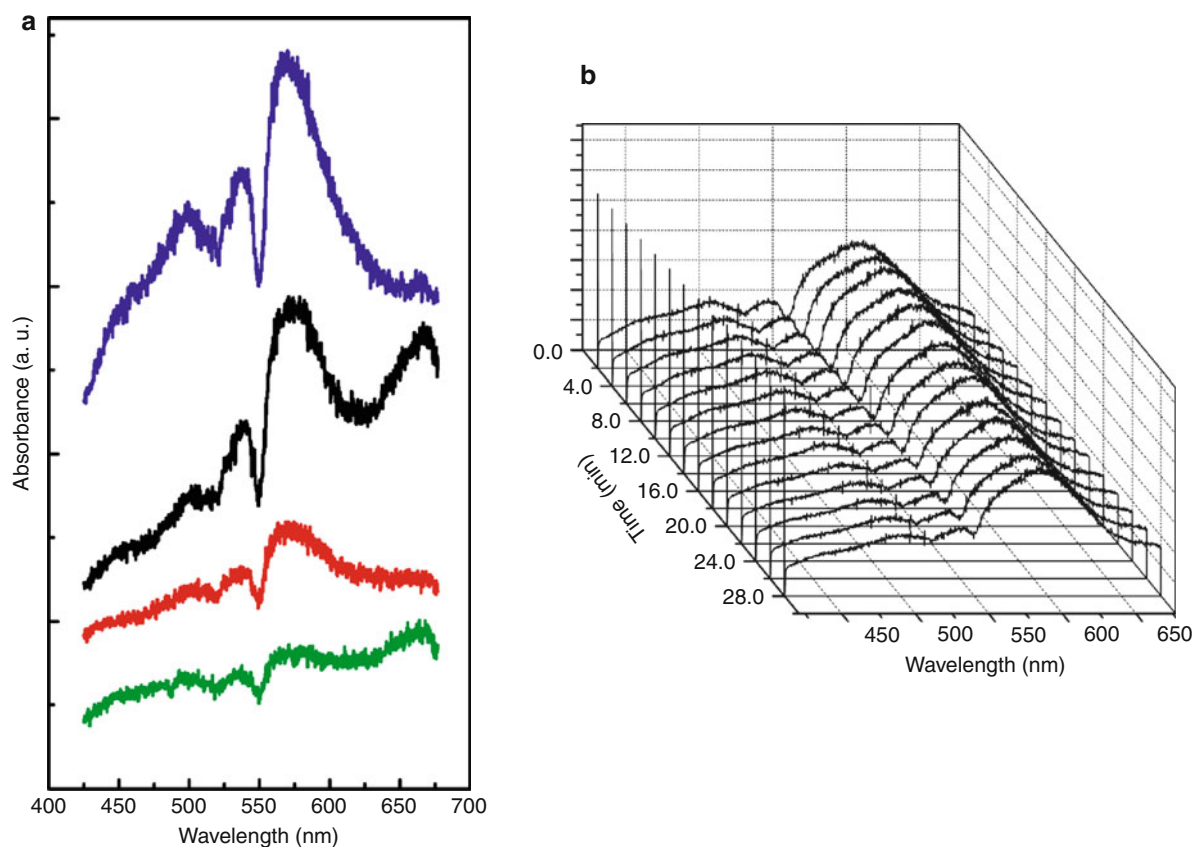
PRET Imaging of Nanoplasmonic Particles in Living Cells

The nanoplasmonic particles can be internalized into living cells via multiple ways, either by natural endocytosis of cells or assisted endocytosis such as lipotransfection technique. The nanoplasmonic particles can be effectively internalized in living cancer cells

using Fugene 6 (Roche Diagnostics) transfection reagent [25]. The nanoplasmonic particles can be clearly imaged using our nanospectroscopy imaging system. Figure shows the dark field scattering image of individual liver cancer cells with many internalized nanoplasmonic particles. Recent study also shows that intracellular Cyt c molecules can be imaged using the PRET nanospectroscopy in a long-term continuous measurement [26] as shown Fig. 8.

Near-Infrared PRET

The aforementioned examples primarily occur in visible light regime due to the visible plasmon resonance wavelength of sub-100 nm gold or silver nanoparticles. As the matter of fact, the PRET effect is not limited to visible wavelength range. When using long gold nanowires with the microscale lengths, the PRET effect can be observed even in mid-infrared region [27]. As shown in Fig. 9a, the gold nanowire with 100 nm in diameter and 1.31 μm in length has the fundamental plasmon resonance mode along the long



Plasmon Resonance Energy Transfer Nanospectroscopy,

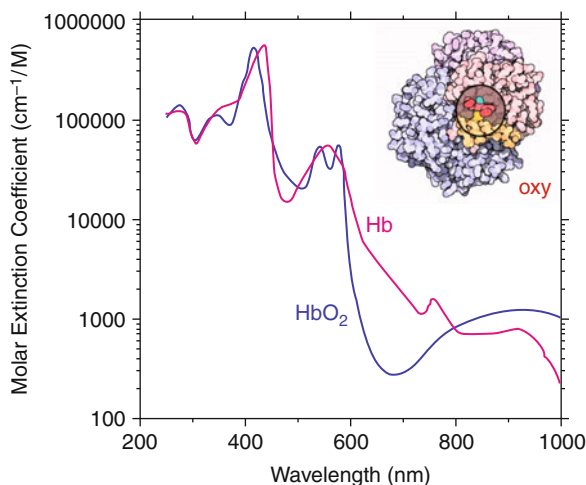
Fig. 5 (a) Raw scattering spectra of four representative gold nanoparticles conjugated with reduced Cyt c molecules. The nanoparticle plasmon resonance peaks and PRET-induced plasmon quenching dips have variable intensities from particle to particle due to the nonuniformity of the conjugated molecule number and particle geometry; however the plasmon quenching

peak positions are consistent. More than 50 individual nanoparticles were tested and PRET can be consistently observed. (b) Time-lapse measurement of scattering spectra of a single gold nanoparticle conjugated with reduced Cyt c molecules. The plasmon quenching spectral dips remain nearly constant during the whole time period of measurement. No photobleaching effect was observed

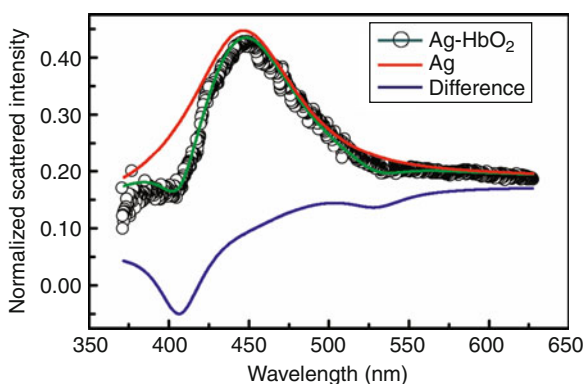
axis at about $3.41 \mu\text{m}$. The gold nanowire will show very different absorption spectra when excited by -perpendicularly and parallel polarized light (Fig. 9b). Only the excited light polarized along the long axis of the gold nanowire will be absorbed by the nanowire due to the fundamental plasmon resonance mode at $3.41 \mu\text{m}$.

When certain chemical molecules are adsorbed on the surface of the gold nanowire, a few spectral dips can be clearly observed and they overlap on top of the nanowire absorption peak. The positions of the spectral dips exactly match with the infrared absorption peak wavelengths of the chemical molecule at 2855 and 2927 cm^{-1} . Even if the plasmon

resonance wavelengths shift due to the nanowire length variation, the PRET spectral dip positions at infrared region remain identical and show high specificity to molecule species. However, the amplitude of the infrared PRET spectral dips is dependent on the extent how the infrared plasmon resonance peak of gold nanowire matches with the infrared absorption peaks of the adsorbed molecules. Since most of biological and chemical molecules have distinctive infrared absorption peaks, the PRET nanospectroscopy on gold and silver nanowires at infrared wavelength has even wider applications to detect small amount of adsorbed molecules without labeling process.



Plasmon Resonance Energy Transfer Nanospectroscopy, Fig. 6 Absorption spectra of oxyhemoglobin and deoxyhemoglobin bulk high-concentration solutions



Plasmon Resonance Energy Transfer Nanospectroscopy, Fig. 7 Absorption spectrum of hemoglobin molecules on single nanoparticle surface and the spectrum is acquired and can be only acquired using our PRET nanomolecular sensor

Applications

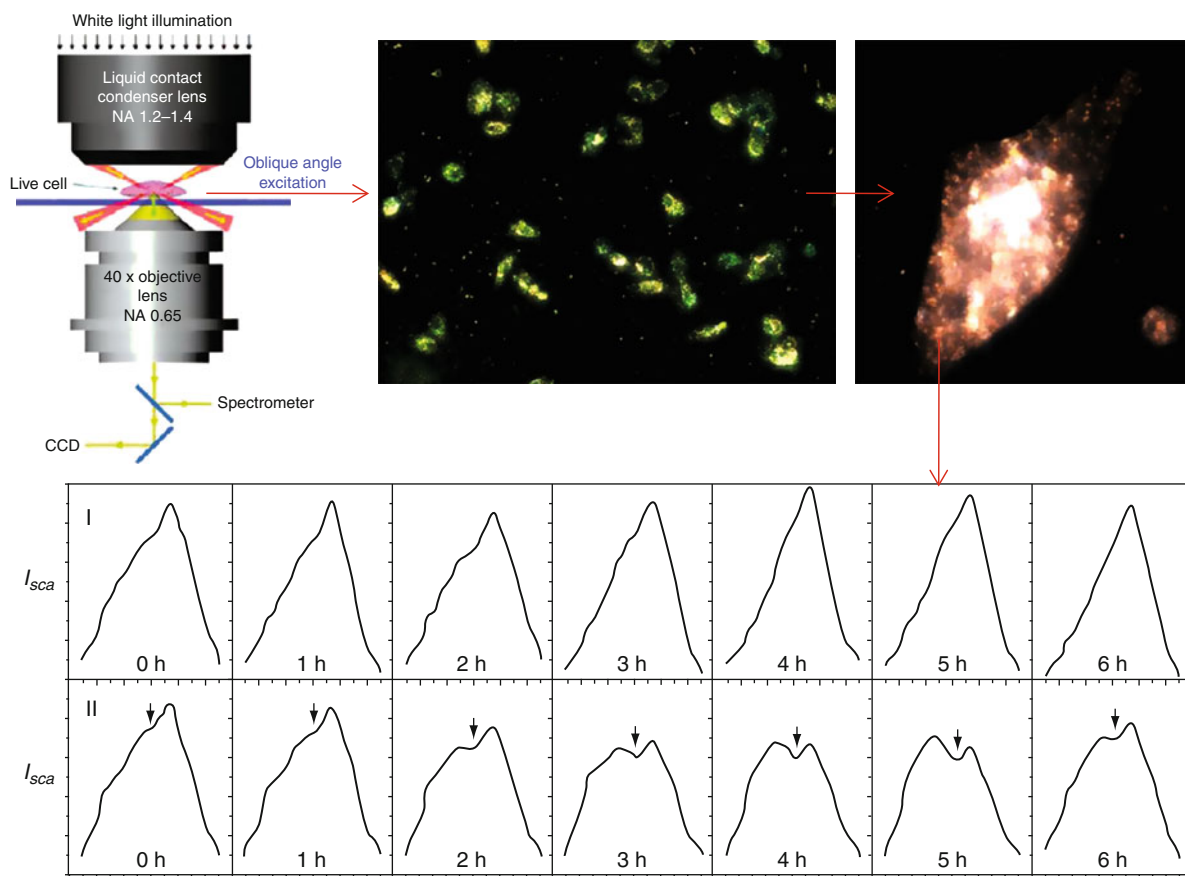
Whole-Field PRET Imaging

All previous PRET imaging works have been limited to the spectroscopy of several particles at most, and whole-field intracellular PRET imaging has not been implemented to date due to the limitations imposed by current spectral imaging system. All existing imaging

and spectroscopy systems for metallic nanostructures are based upon a microscopy system with a white light illumination source, a darkfield condenser, and a polychromator spectrograph, as shown before. In such a system, the field of view is limited to several microns by the width of the entrance slit to the polychromator even though the field of view of the objective lens is on the order of several hundred microns; therefore only a few nanoparticles or nanostructures that are aligned in a column can be imaged spectrally at once.

The polychromator entrance slit is required to limit data pollution for the high-resolution spectroscopy of each individual nanoparticle. Hence, in the case of spectrally imaging hundreds or even thousands of sparsely distributed nanostructures, the sample stage of the microscopy system must be translated either manually or automatically, which has two disadvantages. First, the spectra of separated nanostructures cannot be simultaneously captured and the acquired data will be from different instants in time, e.g., 20 ~ 30 min for a 100 × 100 array, which voids their applicability in the real-time monitoring of multiple biomolecular interactions. Second, the focal plane will tend to drift due to the stage translation which may change the condition of the immersion liquid between the sample and darkfield condenser lens, and thus manual refocusing is often required.

A new multispectral imaging system was constructed to simultaneously monitor the individual scattering spectra and plasmon resonance wavelength of large numbers of nanoplasmonic particles distributed within the field of view of a microscopy objective lens without mechanically scanning the sample [28]. The multispectral imaging system currently supports frame rates as high as 2 s per frame (wavelength) that could potentially be increased by using a light source with a higher power or an image detector with greater sensitivity, by which the image signal-to-noise ratio can also be increased. In contrast with the configurations of previous imaging systems, our whole field plasmonic imaging system consists of a multispectral illumination source synchronized with an intensity imaging camera rather than a polychromatic spectrograph. As shown in Fig. 10, the white light from a halogen lamp is coupled into a scanning monochromator which is

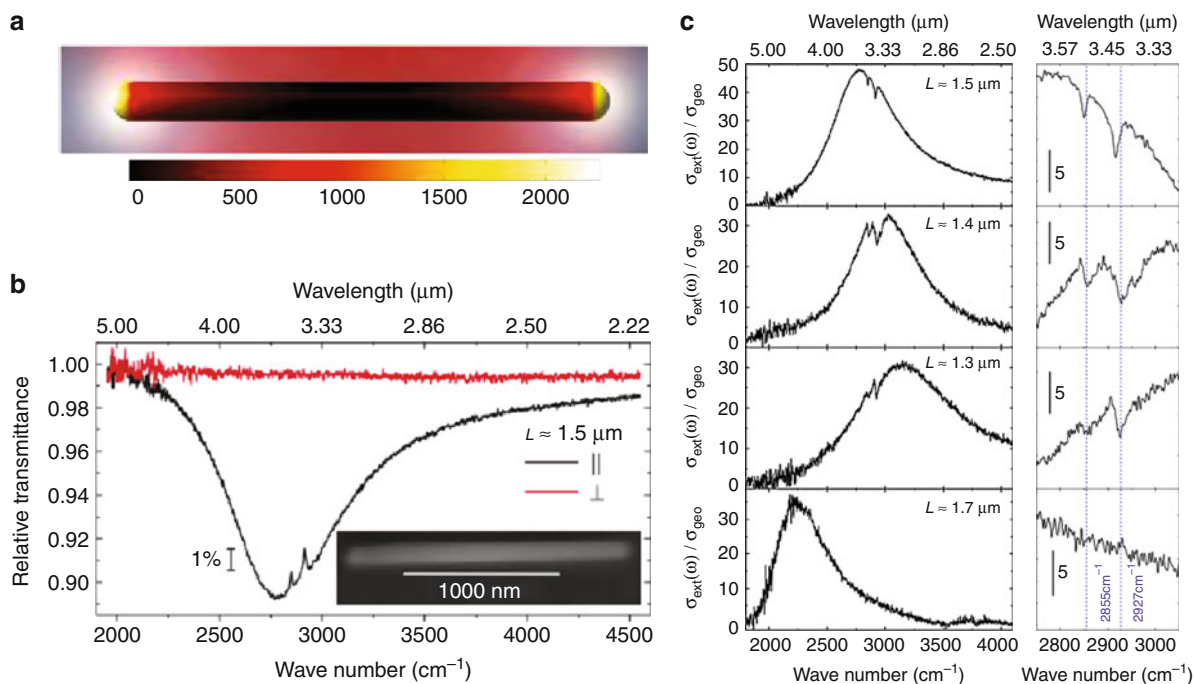


Plasmon Resonance Energy Transfer Nanospectroscopy, Fig. 8 Darkfield imaging of PRET nanoparticles in living cancer cells and the time-lapse observation of intracellular Cyt c molecules using PRET nanospectroscopy [26]

controlled by a computer program to output a monochromatic light beam with varying wavelength and specified spectral step size. The monochromatic light beam is then focused and illuminated on the sample by a darkfield condenser (N. A. = 1.2 ~ 1.4). A 40x microscopy objective lens (N. A. = 0.8) is used to collect the scattered light, which is captured by a CCD camera. The monochromator and image acquisition control software are integrated and they are synchronized to capture a single image at each wavelength of interest. The image acquisition can be finished in a few minutes and the acquisition time is dependent on the chosen spectral range and resolution.

All images are stored as grayscale data files and are analyzed by an image processing program.

The bright spot regions (typically 1 ~ 10 pixels) in each image are recognized by the analysis program as individual nanoplasmonic structures of interest. The mean intensity value of these small regions is extracted from the image at each wavelength as the raw scattering spectra data. The mean intensity value in a large, empty (black) region is also measured at each wavelength as the background spectrum, which is subsequently subtracted from the raw scattering spectra. The difference spectra are then scaled according to the previously stored spectrum of the light output from the monochromator to yield the final scattering spectra. The process of the image analysis and the spectral data reconstruction is completely automated by the computer program.



Plasmon Resonance Energy Transfer Nanospectroscopy, Fig. 9 Infrared PRET on gold nanowire. (a) Fundamental plasmon resonance mode along the long axis of the 1.34 μm long gold nanowire. (b) Relative transmittance spectrum of gold

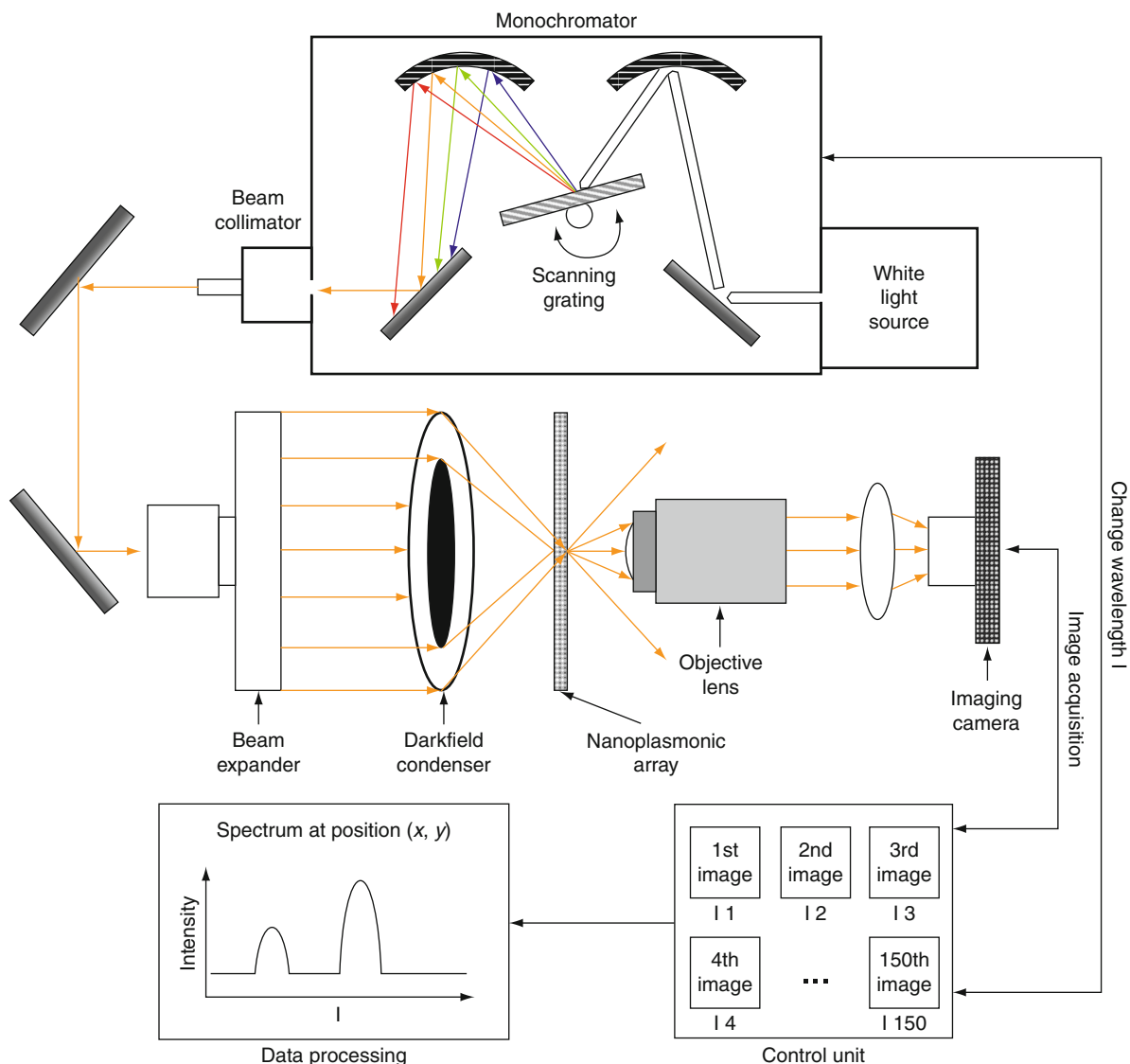
nanowire with (*black curve*) and without (*red curve*) molecule adsorption. (c) Spectral dips due to PRET in the infrared absorption spectra measured from gold nanowires in various lengths [27]

As a demonstration, randomly dispersed Au colloidal nanoparticles were used on a glass slide as the imaging sample. The diameters of the Au nanoparticles vary from 20 to 80 nm, so their plasmon resonance wavelengths, and thus their scattering colors are different. Figure 11a shows the true-color scattering image of the nanoparticles within $\sim 1/10$ of whole view field of the objective lens. The true-color image is taken in the same darkfield microscopy system but with a white-light illumination source and color camera. Figures 11b and 2c show the scattering intensity images of the same nanoparticles within the same field of view taken by our system at 550 and 630 nm, respectively. Figure 11c shows the scattering spectra of three representative nanoparticles marked in Figs. 11a, b, and c. The plasmon resonance wavelengths (spectral peaks) of these three nanoparticles are respectively 560, 580, and 630 nm, which agree well with their colors (green, yellow, and red) in Fig. 11a and their relative intensities in Figs. 11b

and 10c. The Au “particle” with red scattering color could be a cluster of a few Au nanoparticles, because the plasmon resonance wavelength of individual Au nanoparticle is shorter than 600 nm according to Mie scattering theoretical predictions. Although only the scattering spectra of three typical particles are shown here, the spectral information for all the other nanoparticles in the field of view are also stored at once and can be reconstructed in the same fashion.

Future Directions for Research

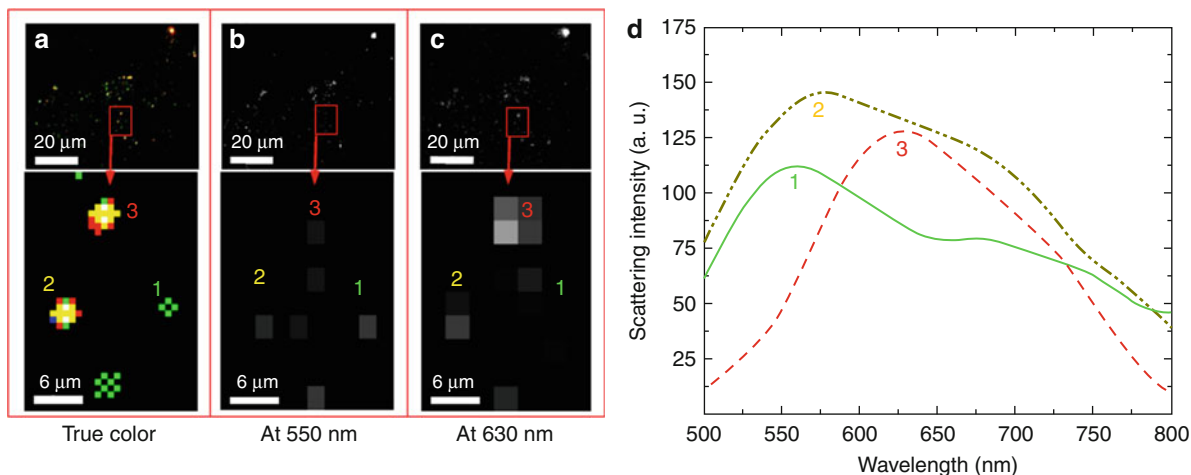
Although only the PRET in the visible wavelength range is observed here due to the optical properties of gold nanoparticles and Cyt c molecules, the PRET process at UV and near infrared range could be envisioned by using different properties (i.e., size, shape, free electron density, etc.) of metallic nanoparticles with UV or near infrared plasmon



Plasmon Resonance Energy Transfer Nanospectroscopy, Fig. 10 Configuration of the multispectral imaging system for the scattering spectra measurement of nanoplasmonic particles in the whole field of view [28]

resonance wavelength. The PRET-based ultrasensitive biomolecular absorption spectroscopy on single metallic nanoparticle could be used for molecular imaging such as genetic analysis of small copies of latent nucleotides, activity measurements of small numbers of functional cancer biomarker proteins, and rapid detection of little biological toxin, pathogen, and virus molecules.

Additionally PRET could be applied in intracellular biomolecule-conjugated nanoparticle sensors to detect localized in vivo electron transfer, oxygen concentration, and pH value changes in living cells with nanoscale spatial resolutions. Furthermore the optical energy in advanced plasmonic devices [29] could be potentially tuned by functional biomolecules taking advantage of the PRET process.



Plasmon Resonance Energy Transfer Nanospectroscopy,

Fig. 11 (a) True-color scattering image of thousands of dispersed Au nanoparticles. For the clarity of the image, a field of $100 \times 100 \mu\text{m}$ is cropped from the whole view field of $\sim 300 \times 300 \mu\text{m}$. The lower picture is the zoom-in image from the marked area (square) of the upper image. (b) and

(c) show the scattering intensity image of the same Au nanoparticles as in (a) with 550 and 630 nm monochromatic illumination, respectively. (d) Scattering spectra of three representative particles marked as 1, 2, and 3, respectively, in the images [28]

References

- Bohren, C.F., Huffman, D.R.: Absorption and Scattering of Light by Small Particles, pp. 335–336. Wiley, New York (1998)
- Mock, J.J., Smith, D.R., Shultz, S.: Local refractive index dependence of plasmon resonance spectra from individual nanoparticles. *Nano Lett.* **3**, 485–491 (2003)
- Otto, A.: Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Z. Phys.* **216**, 398–410 (1968)
- Maier, S.A., Kik, P.G., et al.: Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides. *Nat. Mater.* **2**, 229–232 (2003)
- Maier, S.A., Atwater, H.A.: Plasmonics: localization and guiding of electromagnetic energy in metal/dielectric structures. *J. Appl. Phys.* **98**, 011101 (2005)
- Moskovits, M.: Surface-enhanced spectroscopy. *Rev. Mod. Phys.* **57**, 783–826 (1985)
- Lombardi, J.R., Birke, R.L., Lu, T.H., Xu, J.: Charge-transfer theory of surface enhanced Raman spectroscopy – Herzberg-Teller contributions. *J. Chem. Phys.* **84**, 4174–4180 (1986)
- Nie, S., Emory, S.R.: Probing single molecules and single nanoparticles by surface-enhanced Raman scattering. *Science* **275**, 1102–1106 (1997)
- Kneipp, K., et al.: Single molecule detection using surface-enhanced Raman scattering (SERS). *Phys. Rev. Lett.* **78**, 1667–1670 (1997)
- Das, P., Metiu, H.: Enhancement of molecular fluorescence and photochemistry by small metal particles. *J. Phys. Chem.* **89**, 4680–4687 (1985)
- Liu, G.L., et al.: A nanoplasmonic molecular ruler for measuring nuclease activity and DNA footprinting. *Nat. Nano.* **1**, 47–52 (2006)
- Endo, T., Kerman, K., Nagatani, N., Takamura, Y., Tamiya, E.: Label-free detection of peptide nucleic acid-DNA hybridization using localized surface plasmon resonance based optical biosensor. *Anal. Chem.* **77**, 6976–6984 (2005)
- Rascheke, G., et al.: Biomolecular recognition based on single gold nanoparticle light scattering. *Nano Lett.* **3**, 935–938 (2003)
- Sönnichsen, C.: Plasmons in metal nanostructures. Ph.D. dissertation, University of Munich, Germany (2001)
- van Dijk, M.A., et al.: Absorption and scattering microscopy of single metal nanoparticles. *Phys. Chem. Chem. Phys.* **8**, 3486–3495 (2006)
- Wuttke, D.S., Bjerrum, M.J., Winkler, J.R., Gray, H.B.: Electron-tunneling pathways in cytochrome-c. *Science* **256**, 1007–1009 (1992)
- Yang, J., Liu, X.S., et al.: Prevention of apoptosis by Bcl-2: release of cytochrome c from mitochondria blocked. *Science* **275**, 1129–1132 (1997)
- Lange, C., Hunte, C.: Crystal structure of the yeast cytochrome *bc₁* complex with its bound substrate cytochrome c. *Proc. Natl Acad. Sci. U.S.A.* **99**, 2800–2805 (2002)
- Andrew, P., Barnes, W.L.: Energy transfer across a metal film mediated by surface plasmon polaritons. *Science* **306**, 1002–1005 (2004)
- Van Duyne, R.P.: Molecular plasmonics. *Science* **306**, 985–986 (2004)
- Fang, N., Lee, H., Sun, C., Zhang, X.: Sub-diffraction-limited optical imaging with a silver superlens. *Science* **308**, 534–537 (2005)

22. Bard, A.J., Faulkner, L.R.: *Electrochemical Methods: Fundamentals and Applications*, 2nd edn, p. 685. Wiley, New York (2001)
23. Ozbay, E.: Plasmonics: merging photonics and electronics at nanoscale dimensions. *Science* **311**, 189–193 (2006)
24. Boussaad, S., Pean, J., Tao, N.J.: High-resolution multiwavelength surface plasmon resonance spectroscopy for probing conformational and electronic changes in redox proteins. *Anal. Chem.* **72**, 222–226 (2000)
25. Lee, E.S., Liu, G.L., Kim, F., Lee, L.P.: Remote optical switch for localized control of gene interference. *Nano Lett.* **9**, 562–570 (2009)
26. Choi, Y., Kang, T., Lee, L.P.: Plasmon resonance energy transfer (PRET)-based molecular imaging of cytochrome c in living cells. *Nano Lett.* **9**, 85–90 (2009)
27. Neubrecht, F., Pucci, A., Cornelius, T.W., Karim, S., Garcia-Etxarri, A., Aizpurua, J.: Resonant plasmonic and vibrational coupling in a tailored nanoantenna for infrared detection. *Phys. Rev. Lett.* **101**, 157403 (2008)
28. Liu, G.L., Doll, J.C., Lee, L.P.: High-speed multispectral imaging of nanoplasmonic array. *Opt. Express* **13**, 8520–8525 (2005)
29. Haes, A.J., Zou, S., Zhao, J., Schatz, G.C., Van Duyne, R.P.: Localized surface plasmon resonance spectroscopy near molecular resonances. *J. Am. Chem. Soc.* **128**, 10905–10914 (2006)

Plasmonic Structures for Solar Energy Harvesting

Jao van de Lagemaat

National Renewable Energy Laboratory, Golden,
CO, USA

Renewable and Sustainable Energy Institute, Boulder,
CO, USA

Synonyms

[Plasmonics for photovoltaics](#); [Plasmonics for solar cells](#);
[Surface-plasmon-enhanced solar energy conversion](#)

Definitions

Plasmonic structures for solar energy harvesting describe the use of surface plasmons to increase the energy conversion efficiency of solar cells.

Introduction

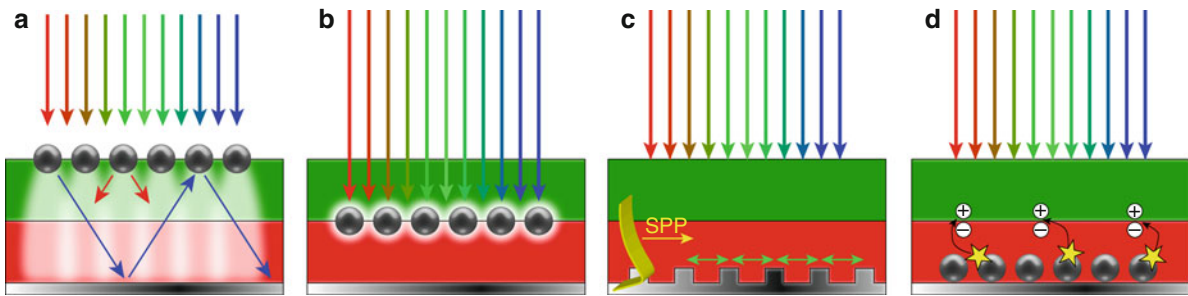
The field of plasmonics describes the interaction of light with matter in or near nanostructured metals and

owes its existence to so-called surface plasmons at the interface of metal nanostructures and dielectrics [1]. Plasmonics have found widespread use in analytical chemistry and are being researched for a variety of applications in photonics such as optical computers, nanoscale lasers, and more. Starting in the early 1990s, publications in the field of surface-plasmon (SP) research have seen geometric growth, with almost 4,000 papers published in 2010. In contrast, publications on the use of surface plasmons in photovoltaic (PV) applications have only recently taken off, with about 50 papers published per year currently.

Surface plasmons are collective oscillations of conduction electrons sustained on nanoscale metals. They are well known to enhance optical electromagnetic fields in their vicinity and show extraordinary light-scattering properties. This realization leads to the intriguing idea of using such metal nanostructures in PV applications. The energy of surface plasmons also depends strongly on the shape and size of the nanostructured metals. As a consequence, a solar cell designer can tune the spectral response of SP active systems in such a manner that the solar spectrum is harvested equally well at all wavelengths [2].

Numerous approaches are being pursued to use surface plasmons in PV applications. [Figure 1](#) illustrates these approaches. Firstly, SPs are used to more efficiently couple light into solar cells by efficient forward scattering, as well as efficient internal reflection of light ([Fig. 1a](#)). Secondly, the well-known property of SPs where they strongly enhance the optical electrical field is used to enhance optical absorption in the absorber ([Fig. 1b](#)). Thirdly, a grating pattern is used that couples light waves into SP polariton waveguide modes in the absorber region of the solar cell ([Fig. 1c](#)). Approaches a to c differ mostly by the placement of the plasmonic structures. Lastly, the creation of hybrid exciton plasmon states leads to new light-absorbing states or to altered excited-state pathways that can lead to increased overall efficiency ([Fig. 1d](#)). A fifth method to employ surface plasmons in solar energy conversion will not be extensively discussed here. In this method, the surface plasmon absorbs the solar energy directly and a “rectenna” structure is used to tap the electrical energy. Although intriguing, it is currently of little practical relevance but could become so in the future.

Surface plasmons are caused by the interaction of surface charges with the electromagnetic field incident



Plasmonic Structures for Solar Energy Harvesting, Fig. 1 Four different strategies proposed for plasmonically enhancing photovoltaics: (a) forward scattering and confinement illustrated by the lighter shaded areas, (b) direct field enhancement, (c) waveguiding, and (d) creation of surface plasmons

inducing altered excited-state dynamics leading to more efficient exciton creation. The photovoltaic device is shown as a generic p-n junction or an acceptor–donor bilayer that is represented by red and green layers, respectively. Graphic inspired by a figure in Ref. [4]

on the surface. They comprise collective oscillations of the conduction electrons in the surface. When one solves Maxwell's equation for the interface between a metal and a dielectric, the dispersion relation for surface plasmons is given by [1]:

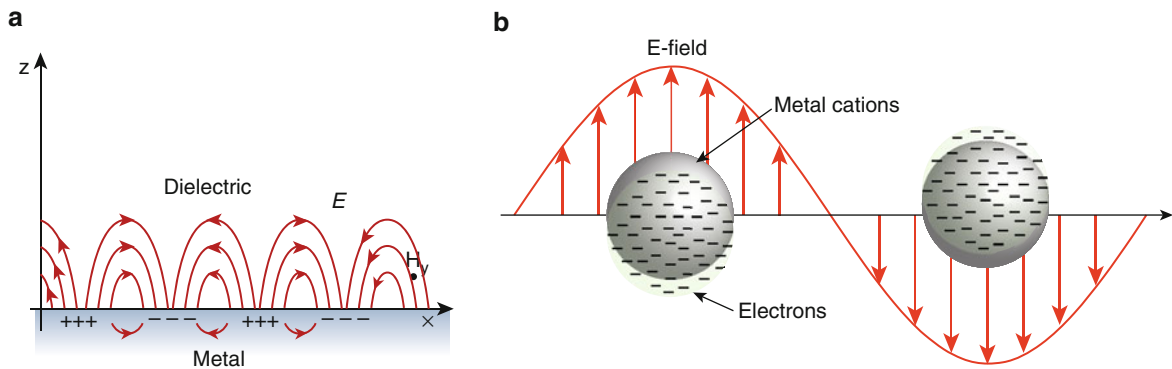
$$k_{SP} = \frac{\omega}{c} \sqrt{\frac{\epsilon_d \epsilon_m}{\epsilon_d + \epsilon_m}}. \quad (1)$$

For SPs to be possible, the complex dielectric constant of the metal ϵ_m must be opposite in sign of that of the dielectric ϵ_d on top. With metals, this is possible because ϵ_m is negative over a large frequency range. Equation 1 also shows that the momentum of the SP mode $\hbar k_{SP}$ is always greater than the momentum of the free-space photon of the same frequency $\hbar k_{fs} = \frac{\omega}{c}$. This leads to the realization that if one wants to generate SPs using light, extra momentum must be supplied. Such extra momentum can be supplied by three different methods. The first is by using prism coupling and is of little practical relevance to solar energy conversion. Secondly, extra momentum can be supplied by scattering by a nanoscale feature such as an isolated metal nanoparticle, or a topological defect such as a nanohole. Lastly, periodic structuring of a metal surface (e.g., a grating) at wavelength scales of visible light allows the light to be momentum matched to the SPs on the metal surface. The last two methods discussed above are of great significance to solar energy conversion and are the basis of all strategies for including SP formation in solar cells.

The surface plasmon that arises on a periodically structured metal interface is called a surface-plasmon

polariton (SPP). A SPP can be thought of as a photon caught at the interface between the metal and dielectric. These collective oscillations have a combined electromagnetic and surface charge character as shown in Fig. 2a [1]. The SPP waves are transverse magnetic (TM) in character and their electromagnetic field intensity is highest at the surface and decays exponentially away from the interface. The increased electrical field above the surface can be used to enhance optical absorption. They can also be used to guide light along the surface horizontally into the active layer of a device, causing increases of optical absorption because of the improved length of interaction of the light wave with the active layer of the solar cell.

The second type of surface plasmon is observed in isolated nanostructures and is visualized in Fig. 2b. It is easiest explained by visualizing a small metal nanoparticle in the electromagnetic field of an incoming light wave. The electromagnetic field moves the sea of conduction electrons up and down in unison at the frequency of the incoming light. The combination with the stationary positive ions of the metal causes a fluctuating dipole. The resonant oscillation frequency of such a nanoparticle where the movement of the conduction electrons is maximized is determined by several factors: the density of conduction electrons, their effective mass, and the shape and size of the distribution of charges (i.e., the particle size and shape). Metal nanoparticles can show multiple resonances such as the aforementioned dipole resonance, as well as higher modes such as a quadrupole mode. Not surprisingly, the oscillating dipole and multipole modes cause strong electromagnetic fields in the



Plasmonic Structures for Solar Energy Harvesting, Fig. 2 (a) Schematic representation of a surface-plasmon polariton (Reprinted by permission from Macmillan Publishers

Ltd: Nature, Ref. [1], copyright 2003). (b) Interaction of light with a spherical nanoparticle. Figure 2b was inspired by Ref. [3]

neighborhoods of such particles that can enhance optical absorption and strong light scattering (i.e., the particle reradiates the energy) that can direct light into light-absorbing structures. See Ref. [3] for a comprehensive discussion of nanoparticle resonances and their behavior.

Forward Light Scattering Using Metal Nanostructures

The strategy of using metal nanostructures on the front surface of solar cells is well established and is shown schematically in Fig. 1a. Metal particles embedded in a homogenous medium scatter light evenly in all directions. However, when such a particle sits at or in the interface between two dielectrics, it will scatter preferentially into the medium with the highest dielectric constant [4]. Analogous to structuring the front surface of solar cells, which is common especially for silicon-based cells, these nanostructures will scatter the light sideways into the active region of the solar cell, allowing a substantial increase of optical path length. At the same time, because a major fraction of the light is now traveling at an angle inside the active layer, it is less likely to escape from the active layer (it is trapped) because it is incident on the top surface again at an angle more shallow than the critical angle. For these reasons, the inclusion of particles at the top surface can cause strong increases in optical absorbance probability by the active layer.

The front-scattering strategy has been extensively studied, especially in silicon photodiodes and solar

cells and has been shown to be effective [5, 6]. One question that remains is whether metal particles in such a geometry are more effective than using simple light-scattering structures such as oxide nanoparticles, pyramids, and others. When designing an effective light-coupling strategy, one must weigh the benefit of improved forward scattering with the losses that occur inside the plasmonic nanoparticles due to simple ohmic heating upon absorption of light – something that does not occur in simple dielectric scattering structures. It is known that smaller particles, in principle, are more effective at scattering light into the active layer, but also that ohmic losses are considerable in the smallest particles. Since ohmic (absorption) losses scale more or less with the volume V of the particles, whereas light-scattering scales with V^2 , larger particles improve this trade-off [4]. Therefore, one would expect an optimum particle size for scattering efficiency. Catchpole and Polman recently carried out a study of different particle geometries and found that cylindrical and hemispherical geometries are more effective than simple spherical particles, and that parallel dipole orientation of the surface plasmon to the surface can lead to more effective light scattering than the Lambertian value [4]. On the other hand, a direct comparison of the enhancement in the case of metal particles versus that when using silica nanosphere scatterers showed that silica nanospheres provide a similar enhancement to the photocurrent as plasmonic gold nanoparticles. Such silica particles are also good candidates for increasing the light in-coupling in silicon solar cells [7]. This result indicates that possibly there is no convincing case from the perspective of cost for using plasmonic structures for forward-

scattering geometries because scattering oxide particles are more economically produced than metal nanoparticles and are also more stable to air exposure.

Direct Field Enhancement Using Embedded Nanostructures

Metal nanostructures can be used to enhance the optical field in their direct vicinity and can therefore be used to enhance optical absorption. The strategy to use this effect is shown schematically in Fig. 1b. The optical enhancement effect has been used widely to enhance the Raman cross section of analyte molecules by orders of magnitude, even enabling the detection of individual molecules. The enhancement of the optical fields around metal nanostructures is usually expressed as an enhancement of the square magnitude of this field because optical absorption processes to a first approximation scale linearly in this quantity. Enhancements of 4 orders of magnitude are attainable, but factors of more than 10 are more common [8].

The direct enhancement approach has been shown to work well in organic solar cell structures where silver particles or other silver nanostructures have commonly been sandwiched between layers or embedded in the active layer of the device [9–12], but also have been used in this manner in inorganic semiconductor-based devices [8, 13]. Figure 3 shows the enhancement of the photocurrent in a bulk heterojunction device fabricated by vapor depositing silver on top of an indium tin oxide conducting front electrode and subsequent deposition of PEDOT:PSS (a transparent conducting polymer), a P3HT/PCBM layer, and a Ca/Al top contact. In such devices, the conversion efficiency can be almost doubled by including silver nanoparticles [11]. In the structures discussed here, especially those where light is incident through the nanoparticle arrays, it is sometimes difficult to separate the effect of field enhancement and enhanced in-coupling of light due to enhanced forward scattering or decreased reflectance. It can be controversial whether any observed enhancement is actually due to field enhancement.

When there is direct contact between the metal nanostructures and the active layer, strong enhancements of the field are expected, but a major problem that can be encountered is that metal nanostructures can act like recombination centers or can quench excited states by energy transfer [10]. This can happen

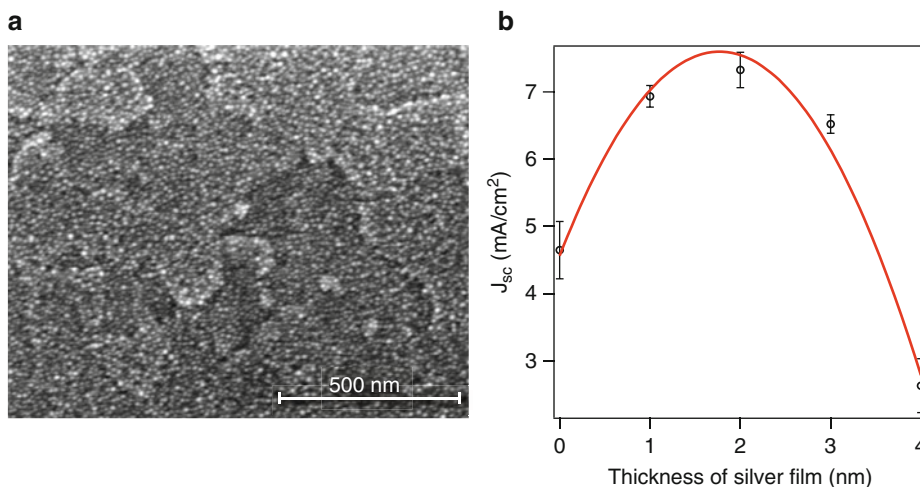
either by capturing both electrons and holes from the semiconductor or by energy transfer from an exciton in the semiconductor to the metal particle, followed by thermalization. Also, in certain PV technologies (such as organic solar cells, which rely on the semiconductor having a low dielectric constant to maintain a constant electrical field over the device to aid in charge separation), the presence of metal nanoparticles can collapse this field because of their very large negative dielectric constant. Direct contact is essential for the greatest absorption enhancements because field enhancement around a metal structure decays approximately exponentially with distance. Therefore, a maximum in the enhancement is expected somewhere in the neighborhood of 5–10 nm, where energy transfer is not efficient but where the evanescent field from the metal nanostructure has not yet fully decayed. This indicates that insulating spacers could be effective when the enhancing particles are in direct contact. Such spacers have typically been provided by buffer layers such as oxides and transparent polymers [11, 13, 14]. Because the spacer layer lowers the expected field enhancement, the benefit of the plasmonically active nanoparticles as direct field enhancers has been limited to small increases in efficiency.

An interesting use of silver nanoparticles that exploits both the optical field enhancement and the strong tendency to cause recombination is to use such particles in the interconnect between two subcells in a tandem device. In this region, charge carriers are supposed to recombine to allow the voltages of the subcell to add up and to mitigate the recombination problem. Such a use of silver nanoparticles has been demonstrated on small molecule organic solar cells [15].

The aforementioned applications are only scratching the surface of the possible enhancement in solar conversion applications by using direct field enhancement by metal nanoparticles. The enhancement that has been demonstrated to date has mostly been in non-optimized designs, and there is much room for further engineering of this enhancement strategy in organic PV, as well as many other types of solar cells.

SPP Waveguide-Enhanced Photovoltaics

Metallic, plasmonically active grating structures can be used to direct light into waveguide modes in



Plasmonic Structures for Solar Energy Harvesting, Fig. 3 (a) Silver nanoparticles created by vapor deposition on an ITO electrode. (b) Observed enhancement of the photocurrent

of an organic photovoltaic device deposited on top of the structure shown in Fig. 3a. Data previously published by the author in Ref. [11]

solar cell structures by generating SPPs (see Fig. 1c). This strategy basically rotates the travel direction of incoming light waves 90°, allowing it to travel as SPPs at the interface between a metal contact and the active layer of the device. This allows for extraordinary improvements in the optical path length of light inside the active area of a solar cell. When light is resonant with the plasmon resonance frequency, significant losses are expected from Joule heating of the metal. However for wavelengths to the red side of this frequency, there can be significant gains in absorption efficiency. An analysis of this trade-off can be found in [16]. To illustrate the potentially very large gains that can be made using the SPP waveguide approach, Fig. 4 shows the expected absorption enhancement in an organic solar cell structure that includes a grating structure, which enhances optical absorption of the active layer significantly [17] even if the effect of loss in the metal structures is taken into account. Such waveguide approaches are expected to be especially useful in solar cell designs where thin layers have to be used either because of charge transport and recombination limitations in the active layer, or because of economical or material supply reasons.

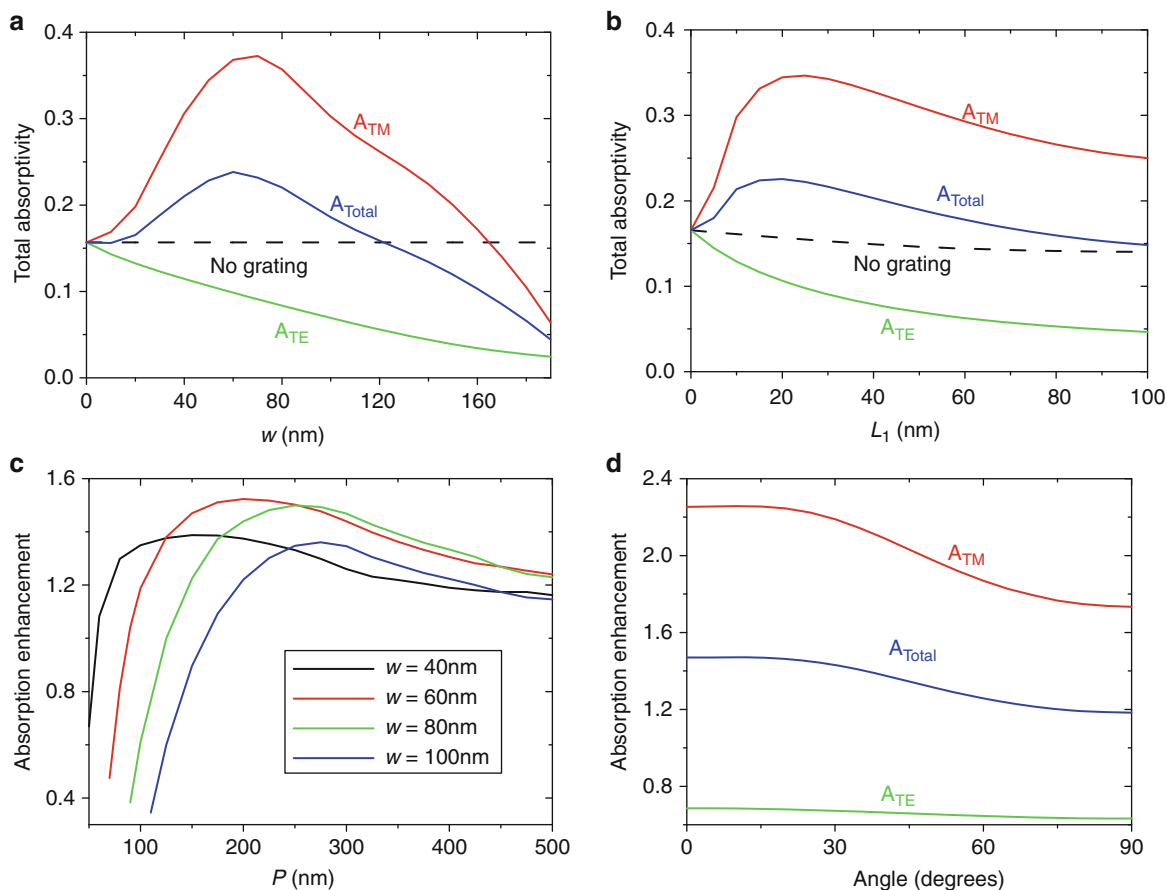
An intriguing idea for SPP waveguide-like modes in a tandem solar cell concept was recently proposed by Atwater and Polman [4]. Figure 5 shows this concept schematically. In each layer, a metallic grating is

employed that is tuned to enhance the absorption of only the wavelength of light that is desired in that layer. Such a structure could reduce the required thickness of the active layers of the device, which currently are often deposited using expensive vapor-phase methods and non-earth-abundant elements.

Contrary to the large amount of theoretical investigations into these structures, few have been demonstrated experimentally, and the potential gain in efficiency remains largely theoretical at the moment. Furthermore, a remaining question is whether such grating structures can be fabricated economically over large areas. SPP-active gratings are currently made using relatively expensive methods such as focused-ion-beam etching or e-beam lithography. More promising methods include conformal imprint lithography using rubber stamps that can generate large-area deposition mask patterns, as well as nanosphere lithography and laser interference lithography.

Enhancement of Third-Generation Approaches

A novel approach to employing plasmonics in solar cells is shown schematically in Fig. 1d. In this approach, the presence of surface plasmons



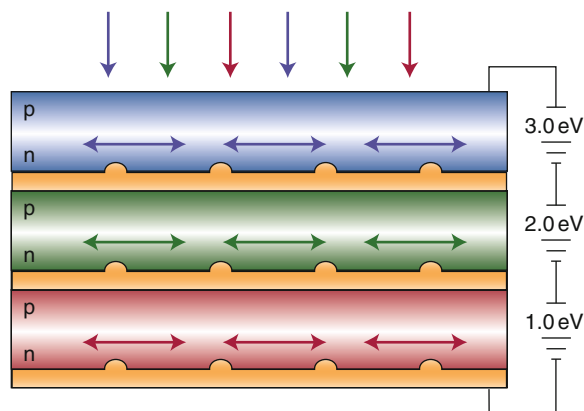
Plasmonic Structures for Solar Energy Harvesting,

Fig. 4 Absorption enhancement in a thin-film organic solar cell using CuPc and PTCBI as an active layer by including a metallic grating as a function of the width of the metal lines (w), their thickness (L_1), and period (P) for normally incident

light-polarized transverse magnetic (TM) and transverse electrical (TE) with respect to the grating repeat direction (Reprinted with permission from Ref. [17], Copyright 2010, American Institute of Physics)

fundamentally alters the excited-state dynamics of the photoactive system, driving a photosystem to a normally unlikely outcome. This approach can be especially significant in systems that employ third-generation photoconversion concepts. Such concepts potentially transcend the limits of single-bandgap photoconverters by converting energy that would normally be wasted as heat. Limiting theoretical efficiencies for such solar cells have been estimated to be around 50% under solar illumination, which is much higher than the limit of about 30% for a traditional, first-generation solar cell. Examples of such approaches where plasmonic nanostructures can be helpful are multiple-exciton generation (MEG) and singlet fission (SF). Both of these

processes are so-called quantum cutting processes, where a single incoming photon is converted into multiple excited states. In the case of MEG, a semiconductor nanoparticle (i.e., a quantum dot) is used that is so small that carriers in it behave like particles in a box, and the bandgap of the material widens. At the same time, the energy levels go from a continuum to a set of discrete levels. Under the right conditions, it becomes possible for a single highly excited state (i.e., an exciton) to break into multiple excitons at the bandgap energy. If one could design solar cell structures that separate these excitons faster than the recombination of the excitons, a commensurate increase in efficiency could be obtained. Although some promising structures have



Plasmonic Structures for Solar Energy Harvesting,

Fig. 5 Plasmonic tandem solar cell concept with layers tuned to different wavelengths of solar light. Semiconductors with different bandgaps are stacked on top of each other, separated by a metal contact layer with a plasmonic nanostructure that couples different spectral bands of the solar spectrum into the corresponding semiconductor layer (Reprinted by permission from Macmillan Publishers Ltd: Nature Materials, Ref. [4], Copyright 2010)

been demonstrated, it seems likely that new structures and concepts are needed. It has been theorized that surface plasmons can aid in the efficiency of long-range energy transfer between quantum dots and inorganic PV devices [18]. This could potentially lead to efficient energy funnel structures, aided by SP active nanostructures that lead the multiple excitations created by the MEG process away from the originating quantum dot before recombination occurs. At the moment, however, such structures are still purely theoretical.

Singlet fission is another process that can lead to much higher efficiency solar energy conversion. The process is observed in molecular systems and molecular crystals. In short, an excited molecule transfers about half of its energy to another similar molecule or another moiety on the same molecule. The molecule itself will go from the singlet excited state to a triplet state that is about halfway down in energy from the singlet excited state. At the same time and in a coherent fashion, the partner molecule undergoes a transition into the triplet state. The overall double-triplet state still has singlet spin character, making the transition spin allowed. Again, as in MEG, a single excitation is converted into multiple excited states, leading to

a decrease in the loss of energy to thermalization of charge carriers. Singlet fission is observed in a large array of different molecular systems, as well as in molecular crystals such as rubrene. It has been postulated that modification of excited-state dynamics can be responsible for more efficient SF processes. Recently, it was observed that pentacene deposited on a silver nanohole array shows more efficient production of triplets and slower triplet–triplet recombination [19]. Possibly, this is due to the surface plasmons creating an efficient pathway to decohere the spin states of the two triplets that start out in coherence.

As mentioned before, SPs can also be used to enhance long-range energy transfer. This effect was recently demonstrated in an organic solar cell structure where light absorbed in a rubrene antenna that had no electrical connection to the cell was effectively transferred to the active device using SPPs resonant in a purposely designed cavity [20]. Such a configuration could be used to design solar cells where light absorption is uncoupled from charge generation and transport.

It is likely that future third-generation photoconversion systems will rely on surface plasmons to increase their likelihood of leading to charge carriers not just by increasing the absorbance probability of light, but also, by fundamentally altering the excited-state dynamics. Using surface plasmons as partner states in such processes is a promising avenue to strongly increased photoconversion efficiencies in the future.

Acknowledgments The author wishes to acknowledge funding from the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences of the US Department of Energy. The submitted manuscript has been offered by an employee of the Alliance for Sustainable Energy, LLC (Alliance), a contractor of the US Government under Contract No. DE-AC36-08GO28308.

Cross-References

- ▶ [Active Plasmonic Devices](#)
- ▶ [Nanomaterials for Excitonic Solar Cells](#)
- ▶ [Nanostructures for Photonics](#)
- ▶ [Nanostructures for Energy](#)
- ▶ [Organic Photovoltaics: Basic Concepts and Device Physics](#)

- ▶ [Plasmon Resonance Energy Transfer Nanospectroscopy](#)
- ▶ [Surface Plasmon-Polariton-Based Detectors](#)

References

1. Barnes, W.L., Dereux, A., Ebbesen, T.W.: Surface plasmon subwavelength optics. *Nature* **424**(6950), 824–830 (2003)
2. Cole, J.R., Halas, N.J.: Optimized plasmonic nanoparticle distributions for solar spectrum harvesting. *Appl. Phys. Lett.* **89**(15), 153120 (2006)
3. Kelly, K.L., Coronado, E., Zhao, L.L., Schatz, G.C.: The optical properties of metal nanoparticles—The influence of size, shape, and dielectric environment. *J. Phys. Chem. B* **107**, 668–677 (2003)
4. Atwater, H.A., Polman, A.: Plasmonics for improved photovoltaic devices. *Nat. Mater.* **9**(3), 205–213 (2010)
5. Lim, S.H., Mar, W., Matheu, P., Derkacs, D., Yu, E.T.: Photocurrent spectroscopy of optical absorption enhancement in silicon photodiodes via scattering from surface plasmon polaritons in gold nanoparticles. *J. Appl. Phys.* **101**(10), 104309 (2007)
6. Pillai, S., Catchpole, K.R., Trupke, T., Green, M.A.: Surface plasmon enhanced silicon solar cells. *J. Appl. Phys.* **101**(9), 093105 (2007)
7. Sundararajan, S.P., Grady, N.K., Mirin, N.A., Halas, N.J.: Nanoparticle-induced enhancement and suppression of photocurrent in a silicon photodiode. *Nano Lett.* **8**(2), 624–630 (2008)
8. Mendes, M.J., Luque, A., Tobías, I., Martí, A.: Plasmonic light enhancement in the near-field of metallic nanospheroids for application in intermediate band solar cells. *Appl. Phys. Lett.* **95**(7), 071105 (2009)
9. Kim, K., Carroll, D.L.: Roles of Au and Ag nanoparticles in efficiency enhancement of poly(3-octylthiophene)/C-60 bulk heterojunction photovoltaic devices. *Appl. Phys. Lett.* **87**(20), 203113 (2005)
10. Yoon, W.-J., Jung, K.-Y., Liu, J., Duraisamy, T., Revur, R., Teixeira, F.L., Sengupta, S., Berger, P.R.: Plasmon-enhanced optical absorption and photocurrent in organic bulk heterojunction photovoltaic devices using self-assembled layer of silver nanoparticles. *Sol. Energy Mater. Sol. Cells* **94**(2), 128–132 (2010)
11. Morfa, A.J., Rowlen, K.L., Reilly, T.H., Romero, M.J., van de Lagemaat, J.: Plasmon-enhanced solar energy conversion in organic bulk heterojunction photovoltaics. *Appl. Phys. Lett.* **92**(1), 013504 (2008)
12. Kim, S.-S., Na, S.-I., Jo, J., Kim, D.-Y., Nah, Y.-C.: Plasmon enhanced performance of organic solar cells using electrodeposited Ag nanoparticles. *Appl. Phys. Lett.* **93**(7), 073307 (2008)
13. Rockstuhl, C., Lederer, F.: Photon management by metallic nanodiscs in thin film solar cells. *Appl. Phys. Lett.* **94**(21), 213102 (2009)
14. Standridge, S.D., Schatz, G.C., Hupp, J.T.: Distance dependence of plasmon-enhanced photocurrent in dye-sensitized solar cells. *J. Am. Chem. Soc.* **131**(24), 8407 (2009)
15. Rand, B.P., Peumans, P., Forrest, S.R.: Long-range absorption enhancement in organic tandem thin-film solar cells containing silver nanoclusters. *J. Appl. Phys.* **96**(12), 7519–7526 (2004)
16. Dionne, J., Sweatlock, L., Atwater, H.A., Polman, A.: Planar metal plasmon waveguides: Frequency-dependent dispersion, propagation, localization, and loss beyond the free electron model. *Phys. Rev. B* **72**(7), 075405 (2005)
17. Min, C., Li, J., Veronis, G., Lee, J.-Y., Fan, S., Peumans, P.: Enhancement of optical absorption in thin-film organic solar cells through the excitation of plasmonic modes in metallic gratings. *Appl. Phys. Lett.* **96**(13), 133302 (2010)
18. 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**(12), 1–16 (2007)
19. Johnson, J.C., Reilly, T.H., Kanarr, A.C., van de Lagemaat, J.: The ultrafast photophysics of pentacene coupled to surface plasmon active nanohole films. *J. Phys. Chem. C* **113**(16), 6871–6877 (2009)
20. Heidel, T., Mapel, J., Singh, M., Celebi, K., Baldo, M.: Surface plasmon polariton mediated energy transfer in organic photovoltaic devices. *Appl. Phys. Lett.* **91**(9), 093506 (2007)

Plasmonic Waveguides

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

Plasmonics

- ▶ [Integration of Nanostructures within Microfluidic Devices](#)

Plasmonics for Photovoltaics

- ▶ [Plasmonic Structures for Solar Energy Harvesting](#)

Plasmonics for Solar Cells

- ▶ [Plasmonic Structures for Solar Energy Harvesting](#)

Plasticity of Nanostructured Solids

- ▶ [Simulations of Bulk Nanostructured Solids](#)

Plasticity Theory at Small Scales

Lorenzo Valdevit¹ and John W. Hutchinson²

¹Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

²School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

Synonyms

Higher-order plasticity theory; Strain gradient plasticity theory

Definition

Plasticity theory is the mathematical formalism that describes the constitutive model of a material undergoing permanent deformation upon loading. For polycrystalline metals at low temperature and strain rate, the J_2 theory is the simplest adequate model. Classic plasticity theory does not include any explicit length scale, and as a result, the constitutive behavior is independent of the sample dimensions. As the characteristic length of a sample is reduced to the micro (and nano) scale, careful experimental observations clearly reveal the presence of a size effect that is not accounted for by the classical theory. Strain gradient plasticity is a formalism devised to extend plasticity theory to these smaller scales. For most metals, strain gradient plasticity is intended to apply to objects in the range from roughly 100 nm to 100 μm . Above 100 μm , the theory converges with the classical theory and below 100 nm surface and grain boundary effects not accounted for in the theory begin to dominate the behavior. By assuming that the plastic work (or in some theories, the yield strength) depends not only on strain but also on strain gradients (a hypothesis physically grounded in dislocation theory and, in particular, in the notion of geometrically necessary dislocations (GND) associated with incompatibility due to strain gradients), an intrinsic length scale is naturally introduced, allowing the theory to capture size effects. According to most theories, the intrinsic length scale is of the order of the distance between dislocation-clipping obstacles or cellular dislocation structures (typically, submicron to

tens of microns). This continuum theory is appropriate for length scales that remain large relative to the distance between dislocations. As the sample length scale is dropped below this level, dislocations must be modeled individually, and discrete dislocations simulations (DSS) are the preferred approach. At even smaller scales, molecular dynamics (MD) becomes the applicable tool. This article presents a brief overview of one of the simplest continuum strain gradient plasticity theories that reduces to the classical J_2 theory when the scale of the deformation becomes large compared to the material length scale. This simple theory captures the essence of the experimental trends observed to date regarding size effects in submicron to micron scale plasticity.

Plastic Flow of Materials and the Emergence of Size Effects at Small Scales

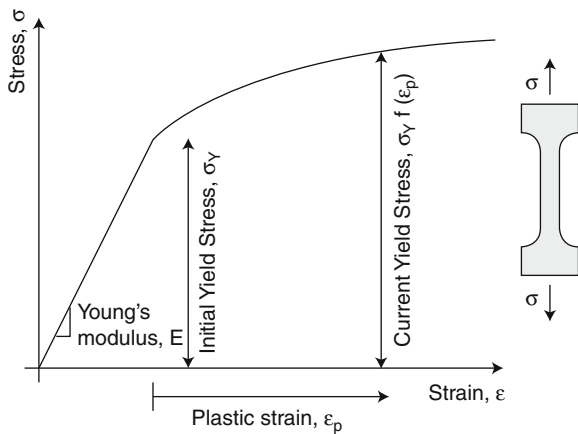
When ductile materials are loaded beyond their elastic limit, permanent deformation will result, involving long-range displacement of atoms (i.e., displacements much larger than the atomic spacing). In metals, this deformation occurs through the motion of dislocations. A typical stress–strain curve for uniaxial tensile loading is depicted in Fig. 1. In polycrystalline samples at low temperatures ($T < 0.5 T_m$, with T_m the melting point of the material on the absolute scale) and relatively low strain rates ($\dot{\epsilon} < 10/s$), a simple constitutive equation relating uniaxial stress to uniaxial strain is typically adequate. The flow stress can be defined as:

$$\sigma = \sigma_Y f(\epsilon_p) \quad (1)$$

with σ_Y the initial yield strength, ϵ_p as the uniaxial plastic strain, and $f(\epsilon_p)$ the hardening function. A power-law expression for f is often adequate. The plastic work per unit volume can be expressed as:

$$U_P(\epsilon_p) = \int_0^{\epsilon_p} \sigma d\epsilon_p = \sigma_Y \int_0^{\epsilon_p} f(\epsilon_p) d\epsilon_p \quad (2)$$

Generalization of this behavior to general three-dimensional states of stress and strain is the objective of plasticity theories and is reviewed in the next section.



Plasticity Theory at Small Scales, Fig. 1 Typical stress–strain curve for an elastoplastic material in uniaxial tension. Young modulus, initial yield strength, and hardening behavior are illustrated

For samples with characteristic dimensions larger than 10–100 μm (depending on the material), yield strength and hardening are independent of sample dimension, rendering (1) scale independent. Traditional plasticity theories embody this critical concept. Over the past two decades, though, as technological advances allowed measurements of mechanical properties in smaller and smaller systems, unexpected size effects became evident, typically suggesting that smaller samples are stronger. A large body of experimental literature supports this finding. Here, for the sake of conciseness, three key observations are reviewed. For a more exhaustive list, see [1].

- (a) *Materials are harder when indented at smaller depths* [2]. A number of careful nanoindentation experiments were conducted and compiled by Nix and Gao [2], revealing a linear dependence of hardness on indentation depth: $(H/H_0)^2 = 1 + (h^*/h)$, where H is the hardness for a given depth of indentation, H_0 the hardness in the limit of infinite depth, and h^* a characteristic length of order 1 μ that depends on the shape of the indenter, the shear modulus of the material, and H_0 (Fig. 2). The effect becomes negligible as the indentation depth is increased to a depth well beyond h^* .
- (b) *Thinner foils are stronger in bending than thicker ones* [3]. Experiments conducted by bending thin nickel foils around a mandrel revealed that thinner samples are stronger and strain harden more than thicker samples. Scaling arguments based on

classic plasticity theory predict that the normalized moment, M/bh^2 (with b and h the foil width and thickness, respectively), should depend uniquely on the plastic strain at the foil surface, $\epsilon_b = h/2R_0$, with R_0 the mandrel radius. For foils of different thickness but identical material, the results should superimpose in the absence of an intrinsic length scale, a conclusion clearly negated by experimental evidence for foil thicknesses between 12.5 and 50 μm (Fig. 3). Importantly, an opposite (albeit less significant) trend is observed in tension, i.e., thinner samples are weaker. Although this conclusion has not been fully rationalized, it clearly demonstrates that strengthening at small length scales entails the presence of strain gradients.

- (c) *Thinner wires are stronger in torsion than thicker ones* [4]. In this classic experiment by Fleck et al. spearheading intense activity in the development of strain gradient plasticity theories, thin copper wires in diameters ranging from 15 to 170 μ were tested in tension and torsion. Dimensional analysis based on conventional plasticity theory predicts a dependence of the normalized torque, T/a^3 (with a the wire radius), on the plastic shear strain at the surface of the wire, $\gamma_s = \kappa a$, with κ the angle of twist per unit length. Again, data from samples of different radius should superimpose, when plotted in these coordinates. Experimental results clearly contrast with this prediction (Fig. 4a). Once again, tensile tests do not show the same trend (Fig. 4b), confirming that the strengthening is associated with the presence of strain gradients.

A constitutive theory incorporating an explicit dependence of the stress on strain gradients (in addition to strain) could in principle explain all the size effects described above. An elementary theory encompassing this principle is discussed below. As explained above, observations (b) and (c) clearly indicate that the strengthening effect disappears as the gradients are removed. It is important to mention that this conclusion changes as the sample size is reduced to sub-micrometric dimensions: in these regimes, substantial size-dependent strengthening is observed under uniform stresses in single crystals (see entry on “► [Size-dependent Plasticity of Single Crystalline Metallic Nanostructures](#)” in this volume).

Conventional Plasticity Theory (J_2 Deformation Theory)

The goal of plasticity theories is the generalization of the one-dimensional model to general, three-dimensional states of stress and strain. For the sake of simplicity, in this article the attention is limited to plastically isotropic materials (i.e., materials for which the flow stress and hardening behavior are independent of the orientation of the coordinate system). It is also assumed that the loading is monotonic and proportional ($\sigma_{ij} = \lambda \sigma_{ij}^0$, with σ_{ij}^0 a constant stress state and λ a loading parameter; throughout this article, indicial notation is adopted: when two indices appear in one term of an equation, summation over those indices is implicitly assumed; see [5] for details) at least approximately. This condition justifies the use of deformation theory, a constitutive model relating plastic strain to stress independently of the loading path. For more general loading scenarios (including all cyclic loadings), a more elaborate theory relating increments of stress and plastic strain (flow theory) is required [5, 6].

The yield point in a uniaxial tensile test (Fig. 1) generalizes to a yield surface, defined in a six-dimensional stress space as the envelope of all stress states resulting in elastic deformation: $f(\sigma_{ij}) = 0$. If $f < 0$, the behavior is completely elastic, whereas if $f = 0$ plastic deformation occurs.

Under the common assumption that a hydrostatic state of stress ($\sigma_{ij} = -p \delta_{ij}$) can be neglected in characterizing plastic behavior, f is taken as a function of the deviatoric stress tensor ($s_{ij} = \sigma_{ij} - (1/3)\sigma_{kk}\delta_{ij}$). For isotropic materials, the most general dependence can be expressed as $f(J_2, J_3) = 0$, where $J_2 = (1/2)s_{ij}s_{ij}$ and $J_3 = (1/3)s_{ij}s_{jk}s_{ki}$ are the second and third invariants of the deviatoric stress tensor. In the J_2 theory of plasticity, the dependence of f on J_3 is ignored, and thus the yield surface can be simply represented as $f(J_2) = 0$. From (1), we have $J_2 = \sigma_Y^2/3$, with σ_Y the yield strength in uniaxial tension. This is the well-known Von Mises criterion [7]. Similarly, the hardening function f can be chosen as an isotropic function of the plastic strain tensor, ε_{ij}^p (isotropic hardening). In the simplest theories, it is also common to assume that hardening depends only on the second invariant of ε_{ij}^p , with $\varepsilon_{kk}^p = 0$ (plastic incompressibility). Thus, the yielding condition for an isotropic

material in a general state of stress can be still written in scalar form as:

$$\sigma_e = \sigma_Y f(\varepsilon_p) \quad (3)$$

with:

$$\sigma_e = \sqrt{\frac{3}{2}} s_{ij} s_{ij} \quad \text{and} \quad \varepsilon_p = \sqrt{\frac{2}{3}} \varepsilon_{ij}^p \varepsilon_{ij}^p \quad (4)$$

The numerical factors $2/3$ and $3/2$ are introduced so that the equivalent stress, σ_e , and the plastic strain magnitude, ε_p , coincide with stress and plastic strain for uniaxial tensile loading.

The plastic work per unit volume can then be expressed as:

$$U_p(\varepsilon_p) = \int_0^{\varepsilon_p} \sigma_e d\varepsilon_p = \sigma_Y \int_0^{\varepsilon_p} f(\varepsilon_p) d\varepsilon_p \quad (5)$$

which is identical to (2), as long as stress and plastic strain are defined as in (4).

As the Cauchy stress tensor σ_{ij} and the strain tensor ε_{ij} are work-conjugate, the plastic strain tensor can be expressed as:

$$\varepsilon_{ij}^p = \frac{\partial U_p}{\partial \sigma_{ij}} = \frac{\partial U_p}{\partial J_2} \frac{\partial J_2}{\partial \sigma_{ij}} = \frac{\partial U_p}{\partial J_2} s_{ij} = \frac{3}{2} \frac{\varepsilon_p}{\sigma_e} s_{ij} \quad (6)$$

where the relation $J_2 = \sigma_e^2/3$ was adopted. The implication is that according to J_2 deformation theory, the plastic strain tensor is codirectional with the stress deviator.

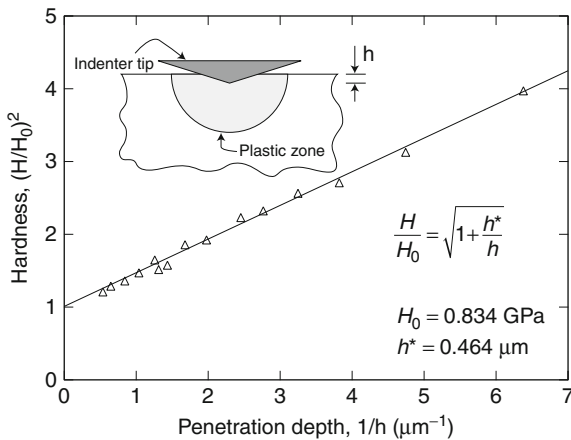
Strain Gradient Plasticity Theories

As mentioned earlier, the theory presented in the previous section does not contain any intrinsic length scale, and is thus incapable of predicting the size effects depicted in Figs. 2–4. The most natural way of introducing a length scale in the constitutive model is to assume that the plastic work is a function of an enhanced strain measure that incorporates both the plastic strain, ε_{ij}^p , and the plastic strain gradients, $\varepsilon_{ij,k}^p$. The deformation version intended for application to proportional or near-proportional loading was

introduced by Fleck and Hutchinson [4] and represents the application to plasticity theory of higher-order elasticity theories first introduced by Toupin [8] and Mindlin [9]. The simplest implementation defines the plastic work per unit volume as:

$$U_p(E_p) = \int_0^{E_p} \sigma_e d\varepsilon_p \quad (7)$$

where:



Plasticity Theory at Small Scales, Fig. 2 Dependence of hardness of cold worked polycrystalline copper on the indentation depth. The data fit a linear model almost perfectly (Modified from [2])

$$E_p = \left((\varepsilon_p)^\mu + (\ell \varepsilon_p^*)^\mu \right)^{1/\mu} \quad (8)$$

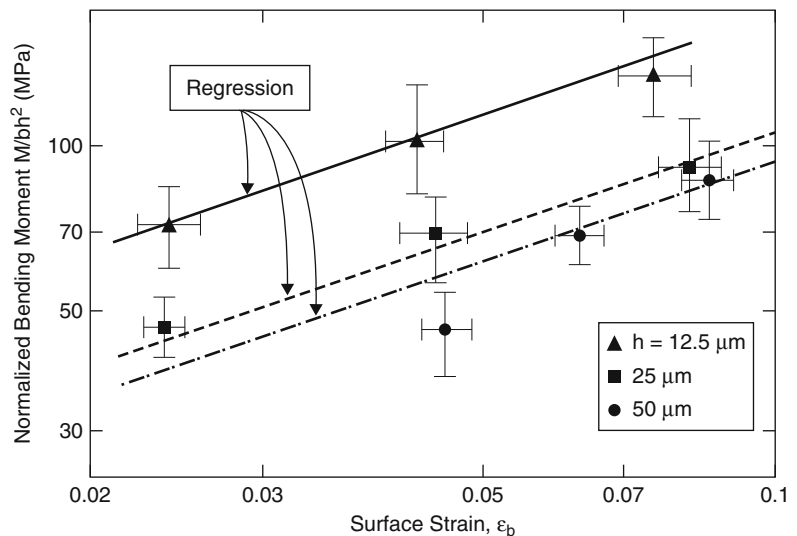
The measure of the plastic strain gradient, ε_p^* , can be defined in many different ways [10, 11]. Here the representation $\varepsilon_p^* = \sqrt{\varepsilon_{p,i} \varepsilon_{p,i}}$, based on the choice made by Zibb and Aifantis [12], is adopted for mathematical convenience. Based on dimensional considerations, the dependence of E_p on ε_p^* requires the introduction of an intrinsic length scale ℓ . This definition generates a family of theories, as a function of the parameter μ . The choice $\mu = 1$ appears to give the best agreement with experimental observations (e.g., it predicts the linear dependence of hardness on indentation depth, Fig. 2 [1]), but most formulations have used $\mu = 2$ for mathematical reasons. This section surveys the basic predictions and general trends of such theories. A physical interpretation of the intrinsic length scale ℓ is provided in the following section.

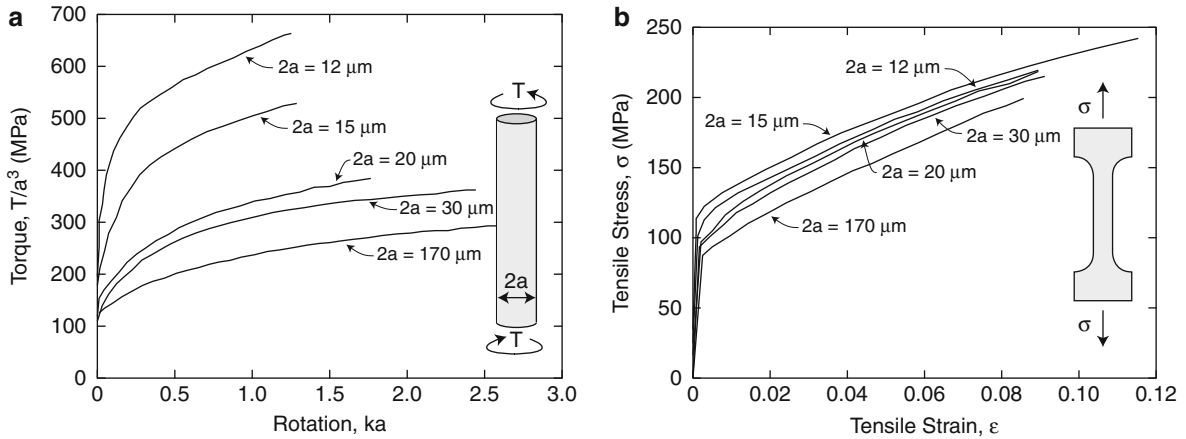
The general trends predicted by this theory are easily elucidated. Imagine a perfectly plastic object ($f(\varepsilon_p) = 1$) of characteristic dimension h , subject to an average plastic strain, ε_p . Dimensionally, the plastic strain gradient is given by $\varepsilon_p^* = c \varepsilon_p/h$, with $c \approx 1$. According to (7) and (8), the plastic work per unit volume can be expressed as:

$$U_p = \sigma_Y (1 + c\ell/h) \varepsilon_p \quad (9)$$

Comparing (9) to the classic J_2 theory prediction, $U_p = \sigma_Y \varepsilon_p$, one concludes that this strain gradient

Plasticity Theory at Small Scales, Fig. 3 Size effect in bending on thin nickel foils. Substantial strengthening is observed for thinner films, consistently with predictions from SGP theories. Importantly, the same effect is not observed in tension (Reprinted from [3])





Plasticity Theory at Small Scales, Fig. 4 Experimentally measured size effect in torsion of thin copper wires. (a) A significant size effect is reported for pure torsion. (b) Nearly

no size effect is observed in uniaxial tension, confirming that the strengthening entails the presence of strain gradients (Modified from [4])

plasticity theory predicts a yield strength elevation of $1 + c\ell/h$. As the length scale of the object, h , approaches the characteristic materials length scale, ℓ , the yield strength elevation becomes significant.

In general, the potential energy functional can be written as:

$$\Phi = \int_V \left(\frac{1}{2} C_{ijkl} \varepsilon_{ij}^e \varepsilon_{kl}^e + U_p(E_p) \right) dV - \int_{S_T} (T_i u_i + t \varepsilon_p) dS \quad (10)$$

with $\varepsilon_{ij}^e = \varepsilon_{ij} - \varepsilon_{ij}^p = \varepsilon_{ij} - \frac{3}{2} \frac{\varepsilon_p}{\sigma_e} s_{ij}$. Here, C_{ijkl} are the elastic moduli (isotropic here) with $\sigma_{ij} = C_{ijkl} \varepsilon_{kl}^e$ and where T_i and $t = \tau_i n_i$ are prescribed on S_T . This class of theories, referred to as higher-order theories, inevitably introduces a new stress measure, τ_i , that is work-conjugate to the plastic strain gradient, i.e., $\tau_i = \partial U_p / \partial \varepsilon_{p,i}$. Specification of an additional boundary condition is required, involving t , ε_p or a combination of both, thus allowing solution to a wider variety of problems, intractable with conventional plasticity theory. For example, passivation layers can be applied on the surface of a solid, forcing the plastic strain to vanish at the surface. (For more details, see [11].)

The solution to (10) represents a minimum with respect to u_i and ε_p , assuming a monotonic uniaxial stress–strain curve (1). The material length scale, ℓ , must be obtained by fitting one set of experimental data that brings in strain gradient effects.

Application to the torsion problem presented in Fig. 4 illustrates the predictive capability of the theory. Full details of the analysis are presented in [1]. Elastic strains constitute a very small contribution to the tensile and torsion data in Fig. 4, justifying the adoption of a power-law tensile stress–strain curve of the form $\sigma_e = \sigma_0 \varepsilon_p^N$. The strain measure ε_p must be replaced by E_p for all loading cases involving strain gradients, as explained above. The ensuing plastic work can then be expressed as:

$$U_p(E_p) = \frac{1}{N+1} \sigma_0 E_p^{N+1} \quad (11)$$

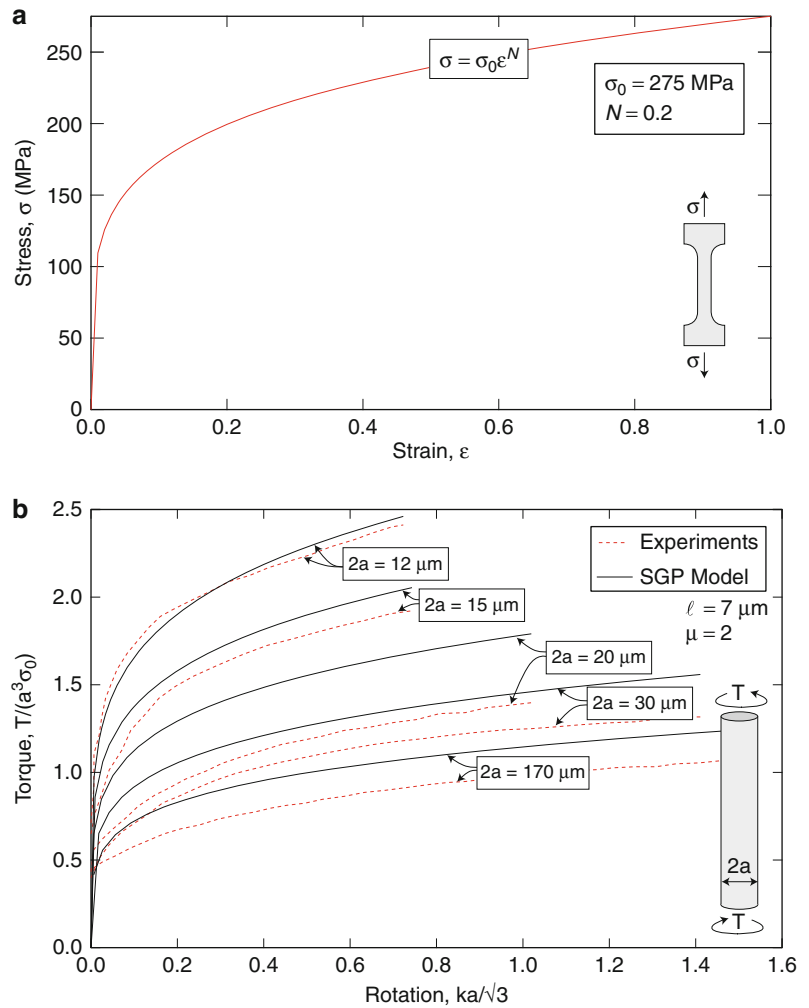
In torsion, with κ as the twist/length, the shear strain varies with distance, r , from the center of the wire according to $\gamma = \kappa r$, such that $\varepsilon_p = \kappa r / \sqrt{3}$ (with $\kappa > 0$), $\varepsilon_p^* = \kappa / \sqrt{3}$, and $E_p = \frac{\kappa}{\sqrt{3}} [r^\mu + \ell^\mu]^{1/\mu}$. The torque, T , under monotonically increasing κ is:

$$\frac{T}{\sigma_0 a^3} = \left(\frac{\kappa a}{\sqrt{3}} \right)^N \frac{2\pi}{\sqrt{3}} \int_0^1 [\zeta^\mu + (\ell/a)^\mu]^{N+1/\mu} \zeta d\zeta \quad (12)$$

where a is the radius of the wire.

The tensile stress–strain curve in Fig. 5a is plotted with $\sigma_0 = 275 \text{ MPa}$ and $N = 0.2$, giving an approximate fit to the experimental curves in Fig. 4a. With these same values and with the choice $\ell = 7 \mu\text{m}$, the trends of the experimental torsion data in Fig. 4a are approximately captured by the theoretical curves from (12) in Fig. 5b. $\mu = 2$ was chosen for all calculations.

Plasticity Theory at Small Scales, Fig. 5 (a) Power-law stress–strain curve in pure tension (extrapolated at large strains), giving an approximate fit to the data in Fig. 4a. (b) Strain gradient plasticity prediction ((12), with $\ell = 7\mu\text{m}$, $\mu = 2$) for the torsion data of Fig. 4b, based on the stress–strain law in (a)

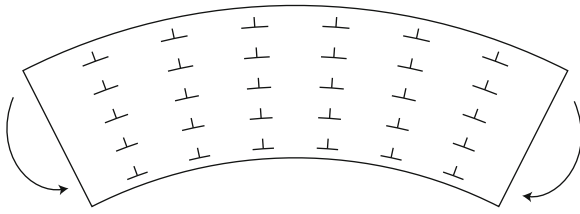


Although in this brief survey the attention was limited to the deformation theory version of strain gradient plasticity (both for mathematical convenience and transparency of scaling effects), more general flow theory versions (relating plastic strain increments to the increments of stress in a manner which has intrinsic strain-path dependence) have been developed [13–16]. The constitutive behavior based on flow theory is applicable to any loading history (including cyclic loading, in principle, although none of the existing versions have been vetted for such loadings) and almost always requires a finite element formulation to analyze problems of interest.

In the next section, the intrinsic length scale ℓ will be qualitatively related to basic notions from dislocation theory.

Physical Interpretation of the Intrinsic Material Length Scale

Conventional strain hardening and yield stress elevation are attributed to the interaction of mobile dislocations with relative immobile *statistically stored dislocations* (SSD). The density of SSD, ρ_{SSD} , increases proportionally to the accumulated plastic strain, according to $\varepsilon_p \approx \rho_{SSD}bd$, with b the burger's vector and d the average distance traveled by a dislocation (defined by the spacing of obstacles). As emphasized by Ashby [17], accumulation of plastic strain gradients requires the introduction of additional dislocations to ensure geometric compatibility (Fig. 6): unlike SSDs, these *geometrically necessary dislocations* (GND) must be arranged in space to



Plasticity Theory at Small Scales, Fig. 6 Cartoon illustrating the formation of geometrically necessary dislocations (GND) in pure bending

accommodate the incompatibility associated with the plastic strain gradient and scale as $\varepsilon_p^* \approx \rho_{\text{GND}} b$. The motion of both families of dislocations contributes to the plastic work. If one assumes that SSD and GND do not strongly interact, a simple additive contribution is appropriate. According to this model, the plastic work can be expressed as:

$$U_p \approx \sigma_Y (\rho_{\text{SSD}} b d + \rho_{\text{GND}} b d) \approx \sigma_Y (\varepsilon_p + d \varepsilon_p^*) \quad (13)$$

Comparison with (9) immediately reveals $\ell \approx d$. As the average distance traveled by a dislocation is comparable to the average spacing between obstacles, this simple analysis suggests that usually $\ell \approx 1 - 20 \mu\text{m}$, consistent with experimental results and modeling prediction (Fig. 5).

Modeling Plasticity of Very Small Samples

The strain gradient plasticity theory described in this entry is still a continuum theory (i.e., it does not model individual dislocations). As the size of the sample is reduced to the point that only a handful of dislocations exist across the characteristic length scale, this fundamental assumption breaks down. In this case, numerical approaches which explicitly model dislocation interactions are required. Molecular dynamics (MD) is a natural modeling strategy, but its application is limited to relatively small samples and extremely high strain rates. Discrete dislocations simulations (DDS) present an approach that bridges the two theories [18–20]. These models generally utilize a FE continuum framework to solve for stresses, strains, and displacements, and treat dislocations as singularities, which both affect and are affected by the global strain and stress fields. Dislocations motion is governed by the Peach–Koehler equation and standard dislocation–

dislocation interaction laws. Dislocation sources like Frank–Read and single-arm sources are introduced at statistically random locations. These models (both in two-dimensional and three-dimensional) are computationally more efficient than full-scale MD simulations, but have not yet fully succeeded in duplicating key aspects of the experimentally measured size effects in plasticity [21].

Open Issues, Future Developments, and Final Considerations

One reason why strain gradient plasticity theories have not been widely embraced is the ambiguity surrounding the definition of the intrinsic length scale. Although the scaling arguments presented in this article ascribe its physical basis and define its order of magnitude, a quantitative physical model allowing expression of ℓ as a function of measurable materials parameters is still lacking. Attempts to define such a model based on dislocation theory are still ongoing. At the same time, efforts are still underway in developing a fully consistent flow theory version of strain gradient plasticity and in implementing such theories in finite element codes.

Cross-References

- ▶ [Ab Initio DFT Simulations of Nanostructures](#)
- ▶ [Finite Element Methods for Computational Nano-Optics](#)
- ▶ [Nanoindentation](#)
- ▶ [Nanomechanical Properties of Nanostructures](#)
- ▶ [Size-Dependent Plasticity of Single Crystalline Metallic Nanostructures](#)

References

1. Evans, A.G., Hutchinson, J.: A critical assessment of theories of strain gradient plasticity. *Acta Mater.* **57**, 1675–1688 (2009)
2. Nix, W.D., Gao, H.: Indentation size effects in crystalline materials: a law for strain gradient plasticity. *J. Mech. Phys. Solids.* **46**(3), 411–425 (1998)
3. Stolken, J., Evans, A.: A microbend test method for measuring the plasticity length scale. *Acta. Mater.* **46**(14), 5109–5115 (1998)
4. Fleck, N., et al.: Strain gradient plasticity – Theory and experiments. *Acta. Metall. Mater.* **42**(2), 475–487 (1994)
5. Lubliner, J.: Plasticity theory. Dover Publications, New York (2008)

6. Hill, R.: The mathematical theory of plasticity. Oxford University Press, Oxford (1998)
7. Mises, R.V.: Mechanik der festen Körper im plastisch deformablen Zustand. Göttin. Nachr. Math. Phys. **1**, 582–592 (1913)
8. Toupin, R.A.: Elastic materials with couple stresses. Arch. Ration. Mech. An. **11**(5), 385–414 (1963)
9. Mindlin, R.D.: Micro-structure in linear elasticity. Arch. Ration. Mech. An. **16**(1), 51–78 (1964)
10. Fleck, N., Hutchinson, J.: Strain gradient plasticity. Adv. appl. Mech. **33**, 295–361 (1997)
11. Niordson, C.F., Hutchinson, J.W.: Basic strain gradient plasticity theories with application to constrained film deformation. J. Mech. Mater. Struct. **6**(1–4), 395–416 (2011)
12. Zibb, H., Aifantis, E.: On the gradient-dependent theory of plasticity and shear banding. Acta. Mech. **92**(1–4), 209–225 (1992)
13. Fleck, N.A., Willis, J.R.: A mathematical basis for strain-gradient plasticity theory. Part II: Tensorial plastic multiplier. J. Mech. Phys. Solids. **57**(7), 1045–1057 (2009)
14. Fleck, N.A., Willis, J.R.: A mathematical basis for strain-gradient plasticity theory-Part I: Scalar plastic multiplier. J. Mech. Phys. Solids. **57**(1), 161–177 (2009)
15. Gudmundson, P.: A unified treatment of strain gradient plasticity. J. Mech. Phys. Solids. **52**(6), 1379–1406 (2004)
16. Gurtin, M., Anand, L.: A theory of strain-gradient plasticity for isotropic, plastically irrotational materials. Part I: small deformations. J. Mech. Phys. Solids. **53**(7), 1624–1649 (2005)
17. Ashby, M.F.: Deformation of plastically non-homogeneous materials. Phil. Mag. **21**(170), 399 (1970)
18. Deshpande, V.S., Needleman, A., Van der Giessen, E.: Plasticity size effects in tension and compression of single crystals. J. Mech. Phys. Solids. **53**(12), 2661–2691 (2005)
19. Tang, H., Schwarz, K.W., Espinosa, H.D.: Dislocation-source shutdown and the plastic behavior of single-crystal micropillars. Phys. Rev. Lett. **100**(18), 185503 (2008)
20. Van der Giessen, E., Needleman, A.: Discrete dislocation plasticity – A simple planar model. Model. Simul. Mater. Sci. Eng. **3**(5), 689–735 (1995)
21. Uchic, M.D., Shade, P.A., Dimiduk, D.M.: Plasticity of micrometer-scale single crystals in compression. Annu. Rev. Mater. Res. **39**(1), 1–23 (2009)

Definition

Plating is a film coating process on a substrate through either chemical or electrochemical reactions in the solution. In chemical plating, several chemical reactions occur in the plating solution to deposit metal on a substrate. In electrochemical plating, known as electroplating, metal ionic species in the plating solution react with electrons supplied by an external electrical power supply to deposit metal onto a substrate.

Overview

Plating is used to change the surface properties of a substrate, for example, hardness, wear resistance, friction, adhesion, corrosion, and electrical or thermal conductivity, and also to decorate objects. Films are deposited by either chemical or electrochemical reactions which are distinguished by whether an outside source of electric current is involved in the reaction. Chemical plating is a non-galvanic plating method where crystallization occurs on the substrate, providing films with no upper limit in thickness. The electroless plating deposits a metal film on nonconductor such as synthetic plastics, which already makes this process important in the printed circuit industry. The list of synthetic plastics to be plated is growing because of the importance of the process for a variety of industrial applications. Electrochemical plating or electroplating is referred to as electrodeposition and is a fairly simple process. A negative charge is placed on the object that is electroplated, which is called cathode and a positive charge is placed on the anode. Both the cathode and the anode are immersed in an electrolyte solution containing the charged particles of dissolved metal that deposit metal on the substrate by passing current through an external circuit between the cathode and the anode. The surface of substrate must be cleaned thoroughly. Recessed areas plate less than other external surface due to the difference of proximity to plating source, often resulting in making a scratch more prominent, rather than being smoothed by the plating process. The thickness of the electroplated films is generally controlled by adjusting time and current in the electroplating process, enabling to control film thickness with nanometer scale accuracy. Three-dimensional nanostructures are produced by the electroforming process in the nanoscale templates such

Plating

Seajin Oh and Marc Madou
 Department of Mechanical and Aerospace
 Engineering & Biomedical Engineering, University of
 California at Irvine, Irvine, CA, USA

Synonyms

[Metal plating](#)

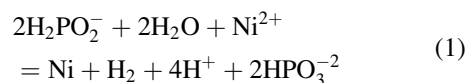
as a metal nanowire array grown through nanopores formed in a polymer matrix. The electroplating process is called anodizing when the object is anodic and an oxide film is formed on the surface.

Basic Methodology

Chemical Plating

Chemical plating is classified into (1) deposition by homogeneous reaction of a metal ion with a reducing agent in solution and (2) deposition by heterogeneous reaction on the substrate specified for the autocatalytic reaction [1].

1. Homogeneous reduction reaction takes place in a solution containing a metal ion to be deposited and a reducing agent that reacts with the metal ion and crystallizes to grow a metal film on the substrate. One old art is silver mirrors that are deposited on glass or synthetic plastics. This process involves a solution of silver salt (e.g., AgNO_3) as one solution and a reducing agent (e.g., sucrose) dissolved in a second solution. The two solutions are mixed before using, or sprayed from a two-nozzle gun on the substrate to be silvered. This method is also used to deposit Cu, Ni, Au, and PbS.
2. Autocatalytic chemical reaction takes place only on catalytic surfaces immersed in the solution where a chemical reducing agent provides electrons for the reaction $\text{M}^{n+} + ne = \text{M}_0$ with no use of external electric power source. This inhomogeneous process requires that the deposited metal layer itself must be catalytic to continue plating on the growing layer and enables continuous deposition of a metal or alloy simply by immersion in the plating solution. In that regard, at least theoretically there is no upper limit to the film thickness to grow. Compared to electroplating, the deposition process has advantages of simplicity and essentially less porous films than electroplating. There are two critical steps before plating. In the first step a reducing agent (e.g., SnCl_2) is adsorbed on the plastic surface (sensitization step), which in the following step reduces PdCl_2 in a dilute solution to Pd deposited on the surface at discrete sites [nucleation step]. This method deposits Ni, Co, Pd, Pt, Cu, Au, Ag, and some alloys containing P or B in the metals. Among them, Ni is the most important film that is deposited from hypophosphite baths, formulated as



Electroplating

Electroplating involves passing an electric current between two electrodes immersed in an electrolyte. When current enters or leaves an electrode, a chemical change takes place at the interface between electrode and electrolyte. For example, a metal cation in the electrolyte receives electron at the interface to deposit as a metal layer. An electrolyte is an ionic solution – solution of acids, bases, salts, fused salts, or some gases – where ions are moved by the electrical potential gradient applied between two electrodes. Ions in the solution are usually hydrated or complexed. Metals in the solution are either in the cationic or in the anionic form. In the latter, the metal is complexed with a ligand or coordinating agent [e.g., $\text{Zn}(\text{CN})_2^{-4}$]. The conductivity of electrolyte is far lower than metal electrode, for example, the conductivity of 1% solution of NaCl is about 5×10^7 times lower than that of copper. The electrode where reduction takes place is the cathode and the electrode where oxidation takes place is the anode. In the electrolyte cell where electroplating takes place, the cathode is the negative electrode and the anode is the positive electrode (Fig. 1).

In electroplating, the quantity of electricity that passes through the electrolyte cell measures the amount of chemical change. Theoretically if the chemical change is solely the deposition of metal and no current leakage occurs, the mass of metal deposit can be exactly calculated as follows:

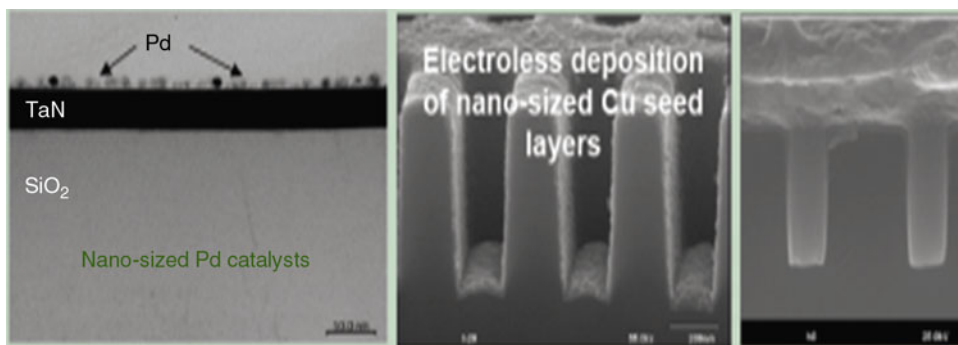
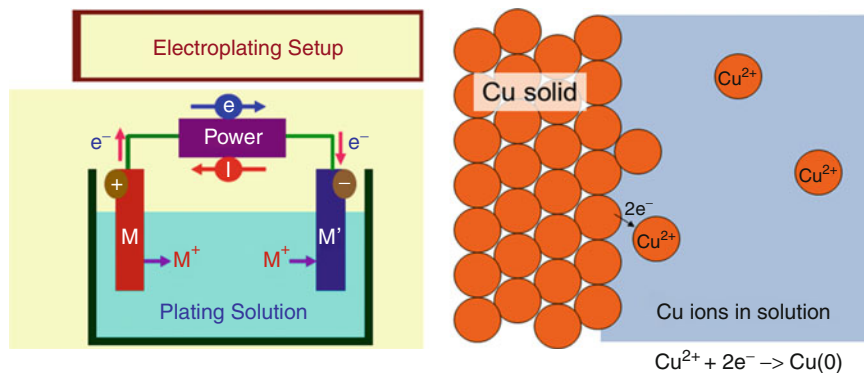
$$G = Iet/96,500 \quad (2)$$

where G is grams of metal, I the current (A), e is the electrochemical equivalent weight (gram) which is atomic weight divided by the valance charge, and t is the time (second). In most cases, the electroplated metal weight is less than predicted by Eq. 2 because of additional reactions involved, such as the co-deposition of hydrogen with metal at the cathode and the evolution of oxygen at the anode which is dissolving in electrolyte. Current efficiency in electroplating varies widely and depends on such factors as electrolyte bath composition, temperature, agitation, pH, and current density.

Anodization is the electroplating process that takes place at anode where the surface metal is converted to a form of its oxide while hydrogen evolves at the

Plating,

Fig. 1 Electroplating. Metal deposition involves reduction of a metal cation in solution by means of electrical current flow



Plating, Fig. 2 Copper electroless plating and electroplating. Trenches (200 nm width) in the dielectric layers are filled by copper electroplating [2]

cathode. The principal application is to Al for corrosion resistance, coloring, and electrical insulation. Other metals to be anodized are Mg, Cu, Ag, Cd, Ti, and steel mostly for decorative effects.

Key Research Findings in Nanotechnology

Copper metallization in damascene process. In the modern microprocessors, the wiring material changes from aluminum to copper to accommodate a large number of transistors interconnected in the restricted space. Copper electroplating takes place to fill the trenches and holes in the dielectric layers with copper in the damascene process for wiring (Fig. 2).

Porous anodic alumina [3]. Porous alumina layers are produced by the anodization of aluminum in an acidic electrolyte. The pores are circular or arranged in a close-packed array of square, triangle, and hexagonal opening. This process starts with an artificial laying-out of the initiating indentation sites in the arrays

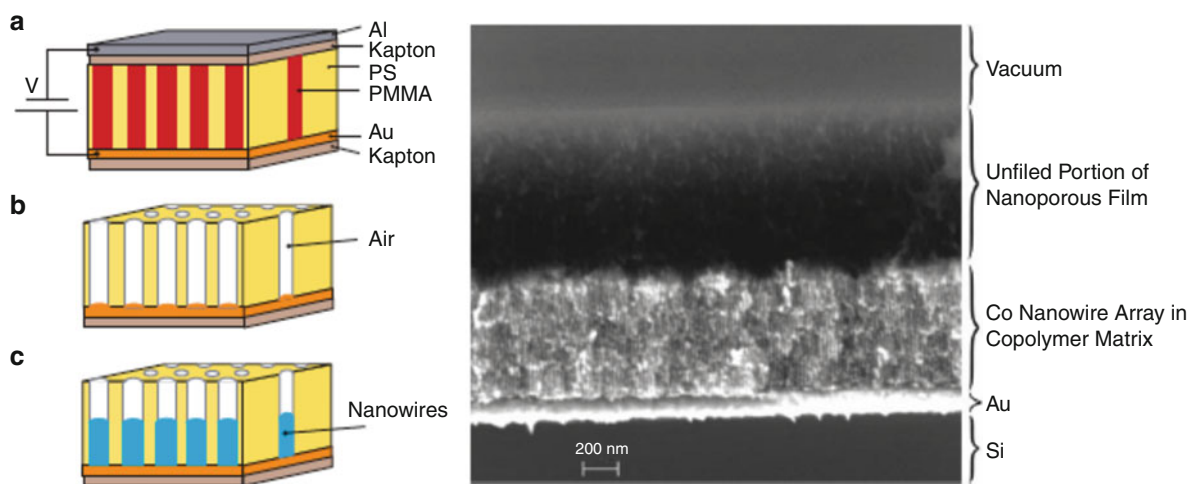
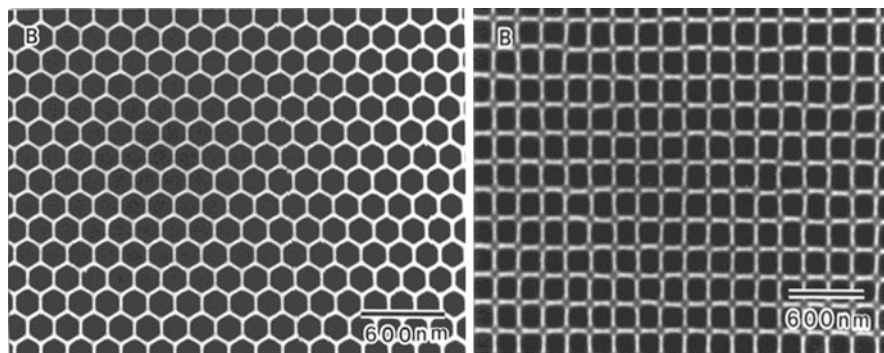
patterned by photolithography and from which corresponding shapes are produced as anodization proceeds in an acidic electrolyte. The pore size, pore spacing, and film thickness are all controllable (Fig. 3).

Electroforming through lithographic masks. Combining electroplating with novel lithography techniques becomes a powerful electroforming technique in the fabrication of microstructures or nanostructures. A representative example is the LIGA process where X-ray lithography creates high aspect ratio structures of photoresist mask, which allows for the fabrication of three-dimensional MEMS devices of a few hundred microns in height [4]. Using nanolithography techniques produces templates including nanoscale features that can produce high-fidelity replicas.

Electrodeposited nanowires. Nanowire arrays are produced by electroplating in diblock copolymer template [5] or porous anodic alumina template [6] (Figs. 4–6).

Selective plating on silicon nanowire. P-n junctions are grown in situ in semiconductor nanowires for

Plating, Fig. 3 The pore opening in the anodic alumina formed by indentation of triangular array (*left*) and square array (*right*) [3]



Plating, Fig. 4 Nanowire grown in diblock copolymer template [5]. (*Left*) A schematic representation of high-density nanowire fabrication in a polymer matrix. (a) An asymmetric diblock copolymer annealed above the glass transition temperature of the copolymer between two electrodes under an applied electric field, forming a hexagonal array of cylinders oriented normal to the film surface. (b) After removal of the minor

component, a porous film is formed. (c) By electrodeposition, nanowires are grown in the porous template, forming an array of nanowires in a polymer matrix. (*Right*) SEM image of a fracture surface of an array of Co nanowires grown within an array of nanopores formed from block copolymers. The growth of the nanowires is terminated before the template is completely filled

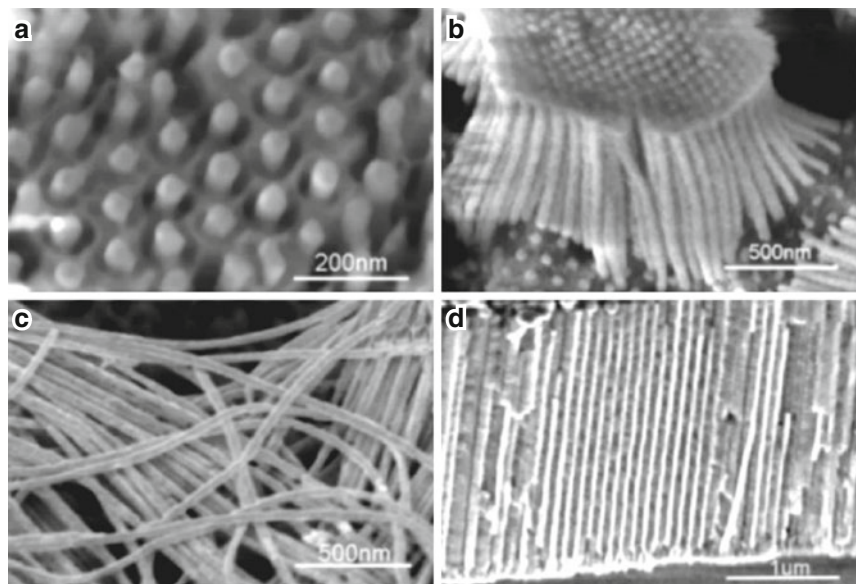
various electronic or optical applications. Au is selectively coated onto the n-type segments of silicon nanowire p-n junction which allows to measure the position of p-n junction and segment length [7].

Future Directions for Research

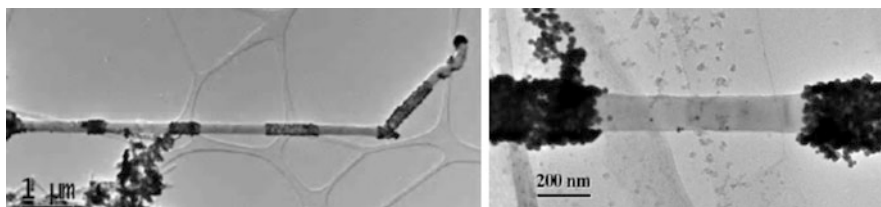
Plating is less expensive than any other thin and thick film deposition methods. Because of the economical advantage continuous effort will be made to produce plated films with desirable mechanical, physical, chemical, and optical properties that meet particular

application requirements. Electroplating is simple to control the deposition rate by controlling the current. In that regard, electroplating is a superior technique for the production of nanoscaled structures made of crystalline pure metals, alloys as well as nanocomposites. Electroplating of metal in templates enables the fabrication of miniaturized structures from submillimeter to nanometer-scaled dimensions with extremely high precision replicas. Research will be continued to provide various three-dimensional nanostructures (e.g., nanowires and nanofilm) in conjunction with the progress of innovative nanotemplate fabrication techniques.

Plating, Fig. 5 Nanowires grown in porous anodic alumina template [6]. (a) Top view of Ni nanowire arrays (b) bundle of Ni nanowires in a 2 mm anodized aluminum oxide film after aluminum oxide was partially dissolved (c) part of the Bi nanowires in a 25 mm anodized aluminum oxide template after aluminum oxide was partially dissolved (d) cross section of the Bi nanowire array, in which the bright lines are Bi nanowires exposed by cracking the template



Plating, Fig. 6 TEM image showing nanoparticle morphology of gold plated on n+ segments [7]



Cross-References

- ▶ [DUV Photolithography and Materials](#)
- ▶ [Nanoindentation](#)
- ▶ [Nanotechnology](#)
- ▶ [SU-8 Photoresist](#)

References

1. Lowenheim, F.A.: Deposition of inorganic films from solution. In: Vossen, J.L., Kern, W. (eds.) *Thin Metal Process*, pp. 209–257. Academic, New York (1978)
2. OM group Inc, power point presentation. www.omgi.com
3. Masuda, H., Asoh, H., Watanabe, M., Nishio, K., Nakao, M., Tamamura, T.: Square and triangular nanohole array architectures in anodic alumina. *Adv. Mater.* **13**, 189–192 (2001)
4. Romankiw, L.T.: A path: from electroplating through lithographic masks in electronics to LIGA in MEMS. *Electrochim. Acta* **42**(20–22), 2985–3005 (1997)
5. Thurn-Albrecht, T., Schotter, J., Kastle, G.A., Emley, N., Shibauchi, T., Krusin-Elbaum, L., Guarini, K., Black, C.T., Tuominen, M.T., Russell, T.P.: Ultrahigh-density nanowire

arrays grown in self-assembled diblock copolymer templates. *Science* **290**, 2126–2129 (2000)

6. Yin, A.J., Li, J., Jian, W., Bennett, A.J., Xu, J.M.: Fabrication of highly ordered metallic nanowire arrays by electrodeposition. *Appl. Phys. Lett.* **79**(7), 1039–1041 (2001)
7. Eichfeld, C.M., Wood, C., Liu, B., Eichfeld, S.M., Redwing, J.M., Mohny, S.E.: Selective plating for junction delineation in silicon nanowires. *Nano Lett.* **7**(9), 2642–2644 (2007)

Polarization-Induced Effects in Heterostructures

Katja Tonisch
 Institut für Mikro- und Nanotechnologien, Technische Universität Ilmenau Fachgebiet Nanotechnologie, Ilmenau, Germany

Synonyms

[Pyroelectricity](#); [Spontaneous polarization](#)

Polarization-Induced Effects in Heterostructures, Table 1 Overview of III-V nitride, silicon carbide, and II-IV oxide materials with wurtzite structure exhibiting both spontaneous and piezoelectric polarization. Values are given in (C/m²) for the main polar axis:

Material	Crystal class	Spontaneous polarization P _{SP} (C/m ²)	Piezoelectric polarization P _{PE} (C/m ²)		
			e ₁₅	e ₃₁	e ₃₃
AlN [1, 2]	P6 ₃ mc	−0.081	−0.48	−0.60	1.46
GaN [1, 2]	P6 ₃ mc	−0.029	−0.3	−0.49	0.73
InN [1, 2]	P6 ₃ mc	−0.032		−0.57	0.97
2H-SiC [3]	P6 ₃ mc	−0.007	0.08	−0.10	0.20
ZnO [4]	P6 ₃ mc	−0.0503	−0.489	−0.664	1.243
BeO [4]	P6 ₃ mc	−0.0363	−0.275	−0.436	0.804
MgO [4]	P6 ₃ mc	−0.134	−0.428	−0.763	2.004
CdO [4]	P6 ₃ mc	−0.107	−0.831	−0.790	1.584

Definition of Crystal Polarization

According to the number of rotational axes and reflection planes, crystals can be divided into 32 classes (point groups). Twenty of these classes exhibit a non-centrosymmetric structure leading to a direct piezoelectric behavior, i.e., an electric field is generated when the crystal is submitted to mechanical force and vice versa. Of these 20 piezoelectric classes, 10 are additionally polar, which implies that they possess a spontaneous polarization, having a dipole in their unit cell. These crystals are called pyroelectric. If this dipole can be reversed by an external electric field, the material is also called ferroelectric. Any dielectric material develops a dielectric polarization (electrostatics) when an electric field is applied, but only substances exhibiting a natural charge separation in the absence of external electric fields are polar. This material property is determined solely by its crystal structure.

Piezoelectric crystal classes: 1, 2, m, 222, mm², 4, −4, 422, 4 mm, −42 m, 3, 32, 3 m, 6, −6, 622, 6 mm, −62 m, 23, −43 m

Pyroelectric crystal classes: 1, 2, m, mm², 3, 3 m, 4, 4 mm, 6, 6 mm

Examples for materials exhibiting both, a spontaneous and a piezoelectric polarization, are given in Table 1. In most cases, two or more polytypes exist, meaning that the same composition can occur in different crystal structures. Here, only the polar crystal structure (i.e., wurtzite or hexagonal) is given, the corresponding polytype is indicated in case of SiC. The polar properties of ternary and quaternary alloys such as AlGaN, AlInN, and AlInGaN can be

interpolated from the involved binary composites. However, this interpolation is not strictly linear but exhibits a parabolic component.

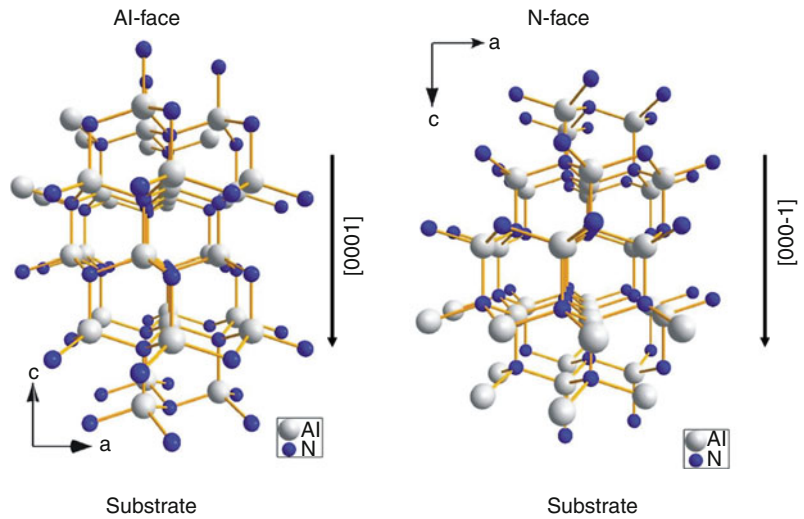
Origin of the Spontaneous Polarization

Crystallographic polarity along certain crystal axes can occur, when two different sequences of the atomic layering in the two opposing directions can be observed. For binary A–B compounds with wurtzite structure, the sequence of the atomic layers of the constituents A and B is reversed along the (0001) and (000−1) directions. The corresponding <0001> and <000−1> faces are the A-face and B-face, respectively (Fig. 1). Please note that they also differ in their chemical behavior.

Each A–B (cation-anion) bond has a dipole character; its strength depends on bond length and charge distribution of the binding electron due to the different electron negativities of cation and anion. In case of wurtzite crystals, the tetrahedral binding structure consists of one vertical and three in-plane cation-anion bonds between an atom and its nearest neighbors. In case of an ideal crystal, the dipoles which correspond to these four bonds compensate each other. However, in an ideal wurtzite crystal, the distance to the second nearest neighbor is by 13% shorter along the polar axis than the distance to the second nearest neighbors within the (basal) plane. This is not the case in cubic crystals. This in fact finally leads to an electric field (i.e., electric polarization) in the direction of the polar axis of the crystal [6]. Real crystals can still deviate from the

Polarization-Induced Effects in Heterostructures,

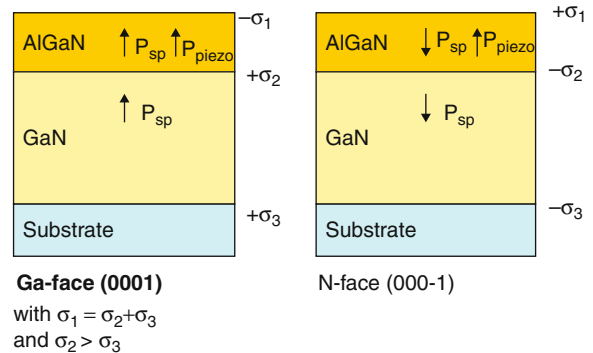
Fig. 1 Crystal structure of AlN illustrating the polarity along the main crystal axis in wurtzite binary A-B compounds [5]



ideal structure in bond lengths and angles, leading to an additional lattice distortion which in return changes the values of the spontaneous polarization.

Charge Carrier Accumulation at Heterostructure Interfaces

Typical applications of polar semiconductors require the epitaxial growth of heterostructures, i.e., a layer stacking with multiple layers of polar materials exhibiting a similar crystal structure is grown on a (hetero-) substrate. A multitude of possibilities exist to stack various layers of binary, ternary, or even quaternary compounds within one material class (e.g., combinations of AlN, InN, GaN, AlInN, InGaN, AlGaN, and AlGaInN layers) or even in combining different classes (ZnO/GaN, GaAs/GaN, AlN/SiC). Each single layer then possesses an individual internal electrical field which is the result of its polarization. Since pseudomorphically grown layers are strained in most cases, the total polarization of each layer is the sum of its spontaneous polarization due to the pure material property and of its piezoelectric-induced polarization due to strain. Such layer stackings can be achieved in both polarization directions (i.e., A- or B-face), which defines the direction of the internal electric field. In all cases, the polarization exhibits an abrupt change in its value or even a change in its sign (positive or negative) at the layer interfaces, resulting



Polarization-Induced Effects in Heterostructures,
Fig. 2 Scheme of a simple AlGaN/GaN heterostructure with its spontaneous and piezoelectric polarization; the GaN layer is assumed to be unstrained

in a net sheet charge. This effect is commonly used intentionally in devices based on polar materials. To simplify matters, a detailed explanation will be given for one material combination only.

A commonly used structure (compare section [Application](#)) is based on thin $\text{Al}_x\text{Ga}_{1-x}\text{N}$ layers grown on thick GaN buffer layers ([Fig. 2](#)). Without external electric fields, the total polarization of a GaN or AlGaN layer is the sum of the spontaneous polarization \mathbf{P}_{SP} of the ideal crystal lattice, and of the strain-induced or piezoelectric polarization \mathbf{P}_{PE} . The piezoelectric polarization can be calculated from the deviation of the real crystal from its ideal structure. In case of epitaxial

grown layers, a biaxial strain occurs due to the alignment of the lateral lattice constant of one layer to the layer beneath. Then, the lateral strain ε_x , ε_y is calculated from the difference between the as-grown lattice constant a_{real} and a_0 (depending on the composition ratio x):

$$\varepsilon_x = \varepsilon_y = \frac{a_{\text{real}} - a_0(x)}{a_0(x)} \quad (1)$$

The corresponding vertical strain ε_z represents the deviation of the vertical lattice constant c and can be calculated using the elastic constants C_{11} and C_{33} :

$$\varepsilon_z = \frac{c_{\text{real}} - c_0(x)}{c_0(x)} = -2 \frac{C_{13}}{C_{33}} \left(\frac{a_{\text{real}} - a_0(x)}{a_0(x)} \right) \quad (2)$$

The piezoelectric polarization is then given by:

$$P_{\text{PE}} = e_{33}(x) \cdot \varepsilon_z + e_{31}(x) \cdot (\varepsilon_x + \varepsilon_y) \quad (3)$$

$$P_{\text{PE}} = 2 \cdot \frac{a_{\text{real}} - a_0}{a_0} \cdot \left(e_{31} - e_{33} \cdot \frac{C_{13}}{C_{33}} \right) \quad (4)$$

The piezoelectric polarization is negative for tensile and positive for compressive strained AlGaN barriers over the whole range of compositions. Since the spontaneous polarization for group III-nitrides is negative, it always points toward the substrate in case of metal-face heterostructures. As a consequence, the alignment of the piezoelectric and spontaneous polarization is parallel in the case of tensile strain, and antiparallel in the case of compressively strained top layers. The piezoelectric, as well as the spontaneous polarization, changes its sign, when N-face structures are considered instead. The polarization field abruptly changes at all heterostructure interfaces resulting in a net sheet charge density σ , which is defined by the polarization difference ΔP of the adjacent layers:

$$\begin{aligned} \sigma(P_{\text{SP}} + P_{\text{PE}}) &= \Delta P \\ &= P_{\text{SP}}(\text{AlGaN}) - P_{\text{SP}}(\text{GaN}) \\ &\quad + P_{\text{PE}}(\text{AlGaN}) \end{aligned} \quad (5)$$

If this sheet charge density is positive as it is the case of the above-considered heterostructure, it will be compensated by free electrons. This results in a two-dimensional accumulation of electrons at the interface,

which is also called two-dimensional electron gas (2DEG). Its sheet charge density can be calculated according to [5]:

$$\begin{aligned} n_s(x) &= \frac{\sigma(x)}{e} - \frac{\varepsilon_0 E_F}{e^2} \left(\frac{\varepsilon_{\text{AlGaN}}(x)}{d_{\text{AlGaN}}} + \frac{\varepsilon_{\text{GaN}}}{d_{\text{GaN}}} \right) \\ &\quad - \frac{\varepsilon_0 \varepsilon_{\text{AlGaN}}(x)}{e^2 \cdot d_{\text{AlGaN}}} (e\phi_{\text{AlGaN}}(x) + E_F(x) - \Delta E_{\text{AlGaN}}^{\text{C}}(x)) \end{aligned} \quad (6)$$

with the Fermi energy E_F of GaN, the relative dielectric constants ε_{GaN} and $\varepsilon_{\text{AlGaN}}$, the Schottky barrier height $e\phi$ at the surface and the conduction band offset ΔE_{C} at the interface. Vice versa, a negative sheet charge density will cause an accumulation of holes at the interface. As the Schottky barrier height at the surface directly influences the sheet carrier density, the choice of the contact material as well as applied external voltages at the surface leads to an increase or depletion of the sheet carrier density. This phenomenon is the typically used effect in applications based on the field transistor configuration (see section “Application”).

In case of heterostructures containing more than one interface, electron and hole accumulations can occur at the respective interface, resulting in a coexistence of 2D electron and hole gases within the same structure. In Fig. 3, the simulation results of a triple heterostructure GaN/AlGaN/GaN is displayed, showing the course of the conduction band and a slightly asymmetric charge distribution at the interfaces.

This coexistence recently has been proven by electro-reflectance measurements [7]. Shifting the net internal electric field by applying an external voltage leads to a corresponding decrease or increase of the electron and hole density (Fig. 4).

Application

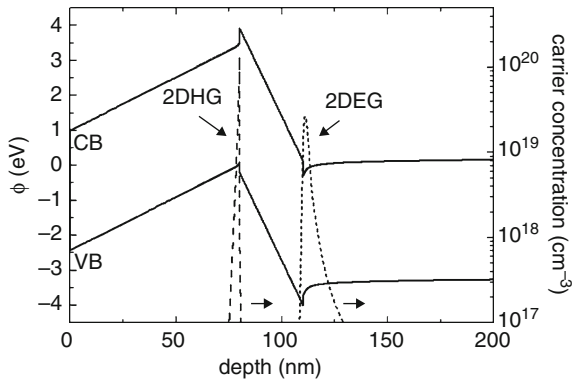
The electron transport properties at heterostructure interfaces in polar semiconductors provide a new field of pyroelectronics and pyrosensors. Thus, the pyroelectric character of such semiconductors yields a novel degree of freedom in designing and tailoring devices for modern micro- and nanoelectronic as well as for micro- and nanomechanical applications. The interface charges, which are induced by the above-described spontaneous and piezoelectric polarization,

respond very sensitively to changes of the overall electric field caused by mechanical strain, the adsorption of ions, wetting by polar liquids, or exposure to gases.

High Electron Mobility Transistors (HEMTs)

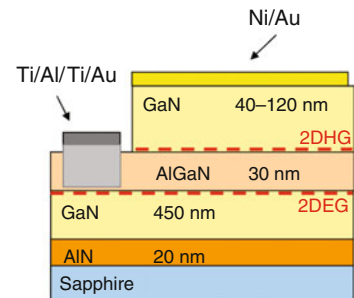
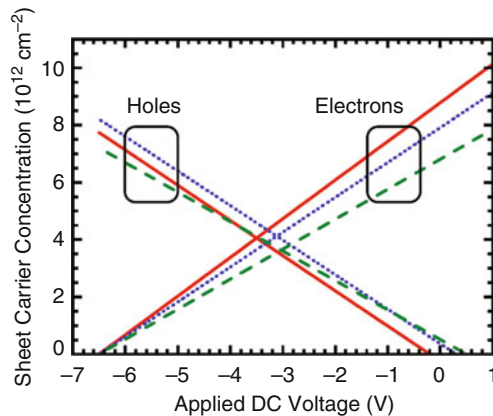
Generally, devices for high currents and high frequencies require a material with high charge carrier mobilities and high saturation velocities in the electron channel. The v - E characteristic defines the electron mobility μ_n , as the slope of the v - E characteristic at low-electric field, and the saturated velocity v_{sat} as the constant, field independent value of the electron velocity at high electric field. Compared to Ga(In)As as a classic semiconductor for rf applications, electron and hole low field mobilities in GaN are three times smaller, which results in lower maximum operation

frequencies. GaN, however, has a larger saturation electron velocity. The fairly good electron mobility and the outstanding high saturation velocity in connection with the high thermal stability and the large band gap of GaN makes AlGaN/GaN-based transistors promising candidates for microwave power amplification (frequencies > 1 GHz, power > 1 W/mm). For this application, high electron mobility transistors (HEMTs) are the most attractive kind of transistors. Originally, HEMTs were developed based on AlGaInAs/(In)GaAs heterostructures on GaAs or InP substrates [8]. These devices are also known as MODFETs (Modulation-doped FETs), HFETs (Heterojunction FETs), TEGFETs (Two-dimensional electron gas FETs), or SDHT (Selectively doped heterostructure transistors). While in GaAs-based HEMTs, the 2DEG is formed at the heterostructure interface as a consequence of strong asymmetrically doping, in GaN-based HEMTs, the interface accumulation is due to polarization charges as described before (Figs. 2 and 3) without intentionally doping. Electron confined in those channels might have a substantially higher mobility due to their separation from their heavily doped source region, and their accumulation in a low-doped region, where impurity scattering has no dominant influence. The mobility in heterostructures increases with Al content [9]. Due to the increased polarization gradient (caused by higher spontaneous polarization as well as higher strain-induced piezoelectric polarization), the 2DEG sheet charge density increases. The mobility peaks for an Al-mole fraction, which is dependent on the barrier thickness and decreases for both lower and higher mole fractions.



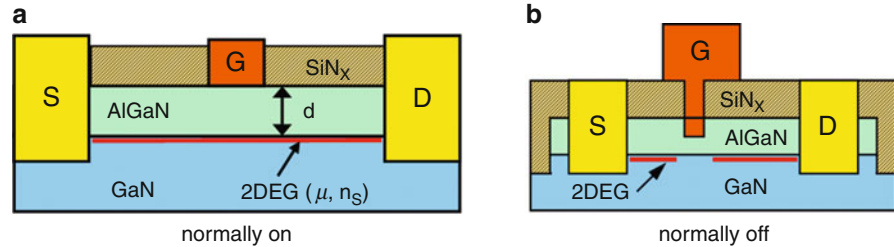
Polarization-Induced Effects in Heterostructures, Fig. 3 Simulation results for a GaN/AlGaN/GaN heterostructure. Displayed are valence band (VB), conduction band (CB), and the resulting 2D hole gas (2DHG) and 2D electron gas (2DEG) at the interfaces

Polarization-Induced Effects in Heterostructures, Fig. 4 Electron and hole concentration at the GaN/AlGaN and AlGaN/GaN interfaces depending on the vertically applied external voltage (left) [7]. Scheme of the analyzed structure (right)



Polarization-Induced Effects in Heterostructures, Table 2 Comparison of electron mobilities and velocities in AlGaAs/GaAs and AlGaIn/GaN heterostructures and bulk material

	μ_{bulk} (cm ² /Vs)	μ_{bulk} (cm ² /Vs)	v_{peak} (cm/s)	v_{sat} (cm/s)	Ref.
AlGaAs/GaAs	10,000 (10 K)	10,00,000 (10 K)	5×10^7 (77 K)		[11]
AlGaIn/GaN	2,070 (115 K)	75,000 (77 K)	3×10^7 (300 K)	2.5×10^7 (300 K)	[12, 14]

Polarization-Induced Effects in Heterostructures, Fig. 5 Basic structure of an AlGaIn/GaN HEMT

Typically, room temperature mobility in the range of $\mu_n \sim 1,000\text{--}1,500\text{ cm}^2/\text{Vs}$ is obtained for device applications. This mobility is affected by several scattering mechanisms. Limiting mechanisms in the AlGaIn/GaN heterostructures are scattering at (1) threading dislocations, (2) due to the interface roughness between AlGaIn and GaN, and (3) due to the random distribution of Al and Ga atoms within the alloy. The electron enhancement in AlGaIn/GaN heterostructures was first reported in 1991 [10]. It was attributed to the two-dimensional character of the electrons, which has been demonstrated by mobility enhancement at lower temperatures as well as Shubnikov-de-Haas oscillations. Now, sheet concentrations of 10^{13} cm^{-2} or higher are achievable close to the interface, well in excess of those achievable in other III-V material systems [5]. In addition, the piezoelectric polarization of the strained top layer is more than five times larger as compared to AlGaAs/GaAs structures, leading to a significant increase of the sheet carrier concentration at the interface (Table 2).

The application of AlGaIn/GaN HEMTs for microwave applications was proposed in 1994 [15]. Such HEMTs can be used in power as well as small-signal applications. In the first case, the device has to withstand high drain source voltages and drain currents. Here, the large breakdown voltage, the achievable high channel conductivity, and the good thermal properties of GaN are advantageous for the devices. A schematic cross section is given in Fig. 5. Typically, AlGaIn/GaN HEMT devices are established on sapphire or SiC

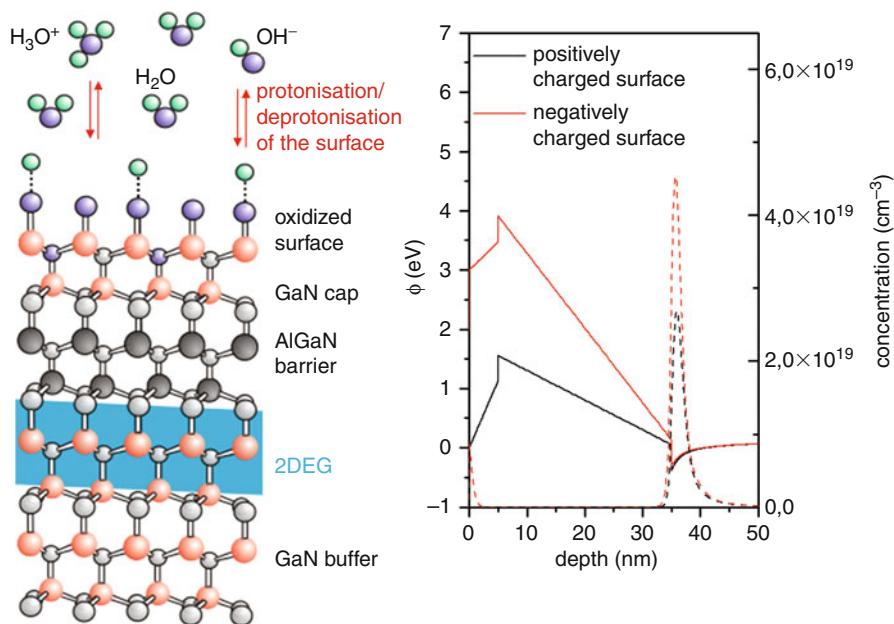
substrates, seldom also on silicon substrates [16]. First, thick GaN films are grown for reduction of growth defects and device isolation. The 2DEG channel is formed between this GaN buffer and an AlGaIn top layer with a thickness of 15–30 nm. Occasionally, thin GaN ($\sim 3\text{ nm}$) are grown on top of the heterostructure in order to improve the stability. Drain and Source contacts are usually Ti/Al based and are alloyed in a high temperature annealing step. The gate contact is crucial for operation at high frequency and has typically T- or Y-shaped cross sections. Additional field plates are employed to reduce the local field enhancement at the gate contact edges, thus enhancing the breakdown voltage. Finally, the device has to be passivated for stable operation, which is commonly achieved by deposition of SiN_x , SiO_2 , Sc_2O_3 , or related materials.

The achievable frequency roughly scales with the transconductance of the device. Thus, thin barriers (thickness d) as well as small gate lengths L_g are preferable for high-frequency operation. Small-signal amplifier transistors usually have small gate lengths ($< 200\text{ nm}$); the breakdown voltage, however, is low ($\sim < 10\text{ V}$). Power HEMTs, on the other hand, have rather uncritical gate lengths. Especially for blocking voltages exceeding 100 V, gate lengths $> 1\text{ }\mu\text{m}$ are sufficient.

The strong polarization in AlGaIn/GaN heterostructures is considered to be one source of some of the experienced reliability problems. To minimize such effects, alternate barrier materials such as InAlN have been employed, which are less polar. In addition, InGaN layers have been used instead of

Polarization-Induced Effects in Heterostructures,

Fig. 6 The hydroxyl groups on the GaN cap layer of an AlGaIn/GaN heterostructure reacting with (a) water, (b) acidic, and (c) alkaline solutions and the resulting impact on the 2DEG sheet carrier density [17]



a pure GaN buffer as well as AlN interlayers at the AlGaIn/GaN interface to improve the confinement of the 2DEG electrons.

Electrolyte-Gated Field Effect Transistors

The development of biosensors based on group III-nitrides received an increasing interest in the last years. In analogy to the development of biologically applied FETs (BioFETs) based on silicon enzymes, DNA, antibodies, and whole cells have already been successfully used as transducing element. Especially the strong bond strength and the resulting high band gap are responsible for the excellent chemical stability and biocompatibility of the nitride materials. Therefore, no functionalization is required for the adhesion and proliferation of cells on the surface. In order to attach different biomolecules, the modification methods, which have been developed for Si-based BioFETs, can easily be adopted. Sensitivity and limit of detection for the first reported AlGaIn/GaN-BioFETs were shown to be at least similar and often exceeding silicon-based devices and their background noise is significantly lower.

The strong responses of AlGaIn/GaN field effect structures have been explained by the site-binding model, which has previously been applied to Si-ISFETs. According to this model, the metal oxide

forms amphoteric hydroxyl groups (M-OH) at the surface when the surface is exposed to aqueous solutions. Depending on the pH value of the solution, these sites can either be protonated (M-OH₂₊), neutral (M-OH), or deprotonated (M-O⁻) (Fig. 6). At a low pH value, i.e., a high concentration of H₃O⁺, they act as a proton acceptor, and at high pH values, i.e., a high concentration of OH⁻, they act as proton donors. This results in a change of the surface charge which in turn enhances/depletes the sheet charge density n_s in the 2DEG at low and high pH, respectively. The pH response of the AlGaIn/GaN heterostructure can also be used to record a pH change induced by a biochemical reaction in the solution, e.g., of an enzyme assay.

For the functionalization of group III-nitrides with biomolecules, different methods of covalent immobilization were applied which can be categorized as follows:

- Direct chemisorption of biomolecules on GaN via amine chemistry
- Deposition on organosilanes, e.g., aminopropyltriethoxysilane (APTES) followed by immobilization of biomolecules via amine chemistry
- Using the high affinity of thiols to a Au-coated gate directly via thiol-modified DNA or indirectly binding biomolecules through ester groups

One of the main advantages of group III-nitrides based ISFETs is their chemical stability under physiological conditions. Silicon is known to be attacked by many biological important agents; thus, an appropriate surface modification by native oxidization or by deposition of other metal oxides such as TaO_5 is required. Also, other semiconductors require a thorough passivation (e.g., GaAs in order to achieve a biocompatible surface). In contrast to that, (Al)GaN-based surfaces are naturally chemically inert, which greatly simplifies the device technology.

Microelectromechanical Systems

With the increasing requirements for microelectromechanical systems (MEMS) regarding stability, miniaturization, and integration, novel materials such as wide band gap semiconductors receive more and more attention. However, only few applications in MEMS have been reported yet, since the commonly used sapphire and silicon carbide heterosubstrates exhibit an extraordinary chemical stability and hinder the realization of freestanding functional layers by selective undercutting.

Nevertheless, wide band gap semiconductor materials, namely SiC, AlN, GaN, and AlGaN/GaN heterostructures, offer superior properties concerning mechanical, thermal, chemical, and biochemical stability. They show no or very low reaction with molecules from the air. For MEMS and NEMS, this is particularly advantageous as surface influences increasingly compromise their performance with decreasing device dimensions. Wide band gap materials exhibit outstanding ratios of the Young's modulus to the mass density, which, for a given resonator geometry and size, effects in significant

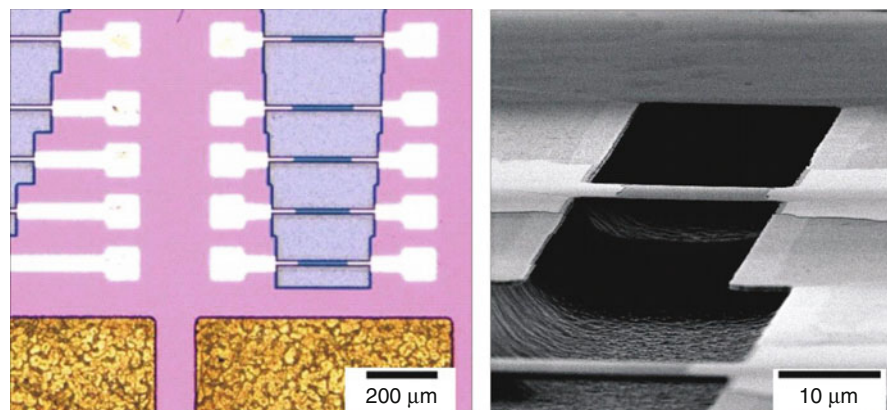
higher mechanical operation frequencies than other materials, e.g., silicon, would allow [18].

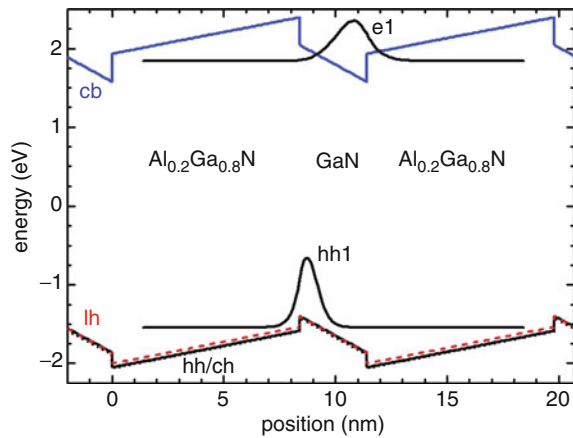
The piezo- and pyroelectric properties of the group III-nitrides provide an elegant way for an integrated electromechanical coupling mechanism (Fig. 7). Thereby, the 2DEG serves as back electrode for the piezoelectric coupling of both, actuation and read-out [19]. Although other materials, e.g., functional ceramics such as $\text{Pb}(\text{Zr},\text{Ti})\text{O}_3$ (PZT), offer higher piezoelectric constants, the group III-nitrides convince by their structural quality and the dimensional control down to the monolayer level. The above-described 2DEG at the interface of AlGaN/GaN heterostructures can also be used for designing sophisticated MEMS-based sensor concepts. Due to its dependence on the spontaneous polarization of the interfacing materials, the 2DEG is sensitive to mechanical stress and other environmental influences such as temperature or chemical surface modifications. Additionally, this material allows the combination of MEMS and HEMT structures, thus opening the way to sensing concepts with integrated signal amplification.

Unintentional Polarization Effects

The quantum-confined Stark effect (QCSE) describes the effect of an external electric field upon the light absorption spectrum of a quantum well (QW). Without electric fields, electrons and holes within the quantum well only occupy states within a discrete set of energy subbands. Consequently, only a discrete set of frequencies of light may be absorbed by the system. Under the influence of an external electric field, the electron states shift to lower energies, while the hole states shift to higher energies, thus reducing the permitted light absorption frequencies. Additionally, the external

Polarization-Induced Effects in Heterostructures, Fig. 7 Double clamped beam structures of AlGaN/GaN on silicon. *Left:* optical micrograph, the size varies from 275 to 350 μm , *right:* scanning electron micrograph of a 20- μm beam [20]





Polarization-Induced Effects in Heterostructures, Fig. 8 Quantum-confined Stark Effect illustrated at an AlGaIn/GaN/AlGaIn quantum well structure, as simulated by nextnano³ [20]

electric field shifts electrons and holes to opposite sides of the well, decreasing the overlap integral, which in turn reduces the recombination efficiency (i.e., fluorescence quantum yield) of the system due to a spatial separation. In case of polar materials, the above-described shifts in the energy subband levels and in the localization of charge carriers are not caused by external fields, but inherently caused by the polarization-induced internal electric field. Figure 8 shows the conduction (cb) and valence (heavy hole (hh), light hole (lh), and crystal-field split-off hole (ch)) band edges of an AlGaIn/GaN heterostructure, including the effects of strain, piezo- and pyroelectricity. The ground state electron (e1) and the ground state heavy hole wave functions (hh1) are shown. Due to the built-in piezo- and pyroelectric fields, the electron wave functions are shifted to the right and the hole wave function to the left.

Cross-References

- ▶ [Basic MEMS Actuators](#)
- ▶ [Biosensors](#)
- ▶ [NEMS Piezoelectric Switches](#)
- ▶ [Piezoelectric Effect at Nanoscale](#)

References

1. Ambacher, O.: Growth and applications of group III-nitrides. *J. Phys. D Appl. Phys.* **31**, 2653–2710 (1998)
2. Bernardini, F., Fiorentini, V., Vanderbilt, D.: Spontaneous polarization and piezoelectric constants of III-V nitrides. *Phys. Rev. B* **56**, R10024–R10027 (1997)
3. Goldberg, Yu, Levinshtein, M.E., Rumyantsev, S.L.: Silicon carbide (SiC). In: Levinshtein, M.E., Rumyantsev, S.L., Shur, M.S. (eds.) *Properties of Advanced Semiconductor Materials GaN, AlN, SiC, BN, SiC, SiGe*, pp. 93–148. Wiley, New York (2001)
4. Gopal, P., Spaldin, N.A.: Polarization, piezoelectric constants and elastic constants of ZnO, MgO and CdO. *J. Electron. Mater.* **35**, 538–542 (2006)
5. Ambacher, O., et al.: Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaIn/GaN-heterostructures. *J. Appl. Phys.* **85**, 3222–3233 (1999)
6. Ambacher, O., et al.: Pyroelectric properties of Al(In)GaIn/GaN hetero- and quantum well structures. *J. Phys. Condens. Matter* **14**, 3399–3434 (2002)
7. Buchheim, C., Goldhahn, R., Gobsch, G., Tonisch, K., Cimalla, V., Niebelschuetz, F., Ambacher, O.: Electric field distribution in GaN/AlGaIn/GaN heterostructures with 2DEG and 2DHG. *Appl. Phys. Lett.* **92**, 013510 (2008)
8. Yu, P.Y., Cardona, M.: *Fundamentals of Semiconductors: Physics and Materials Properties*. Springer, Heidelberg (2010)
9. Arulkumaran, S., Egawa, T., Ishikawa, H., Jimbo, T.: Characterization of different-Al-content AlGaIn/GaN heterostructures and high-electron-mobility transistors on sapphire. *J. Vac. Sci. Technol.* **B21**, 888–894 (2003)
10. Khan, M.A., Van Hove, J.M., Kuznia, J.N., Olson, D.T.: High electron mobility GaN/Al_xGa_{1-x}N heterostructures grown by low-pressure metalorganic chemical vapor deposition. *Appl. Phys. Lett.* **58**, 2408–2410 (1991)
11. Yngvesson, S.: *Microwave Semiconductor Devices*. Kluwer, Boston/Dordrecht/London (1991)
12. Manasreh, O. (ed.): *III-Nitride Semiconductors: Electrical, Structural and Defects Properties*. Elsevier, Amsterdam (2000)
13. Manfra, M.J., Pfeiffer, L.N., West, K.W., Stormer, H.L., Baldwin, K.W., Hsu, D.V., Lang, J.W.P., Molnar, R.J.: High-mobility AlGaIn/GaN heterostructures grown by molecular-beam epitaxy on GaN templates prepared by hydride vapor phase epitaxy. *Appl. Phys. Lett.* **77**, 2888–2890 (2000)
14. Khan, M.A., Kuznia, J.N., Olson, D.T., Schaff, W.J., Burm, J.W., Shur, M.S.: Microwave performance of a 0.25 μm gate AlGaIn/GaN heterostructure field effect transistor. *Appl. Phys. Lett.* **65**, 1121–1123 (1994)
15. Tonisch, K., Jatal, W., Niebelschuetz, F., Romanus, H., Baumann, U., Schwierz, F., Pezoldt, J.: AlGaIn/GaN-heterostructures on (111) 3C-SiC/Si pseudo substrates for high frequency applications. *Thin Solid Films* **520**, 491–496 (2011)
16. Cimalla, I., Gebinoga, M., Lübbers, B., Cimalla, V., Ambacher, O., Schober, A.: Group III nitride based sensors – a promising bio-sensing generation. In: Ho-Young, C. (ed.) *Advanced Semiconductor Materials and Device Research : III-nitrides and SiC*. Research Signpost, Kerala, (2011). ISBN: 978-81-7895-371-7
17. Brueckner, K., Niebelschuetz, F., Tonisch, K., Foerster, Ch., Cimalla V., Stephan, R. Pezoldt, J., Stauden, T., Ambacher, O., Hein, M.A.: Micro- and nano-electromechanical

- resonators based on SiC and group III-nitrides for sensor applications, *Phys. Status Solidi A*, 1–20 (2010)
18. Tonisch, K., Buchheim, C., Niebelschütz, F., Cimalla, B., Schober, A., Gobsch, G., Ambacher, O., Goldhahn, R.: Piezoelectric actuation of (GaN)/AlGaIn/GaN heterostructures. *J. Appl. Phys.* **104**, 084516 (2008)
 19. Brückner, K., Niebelschütz, F., Tonisch, K., Michael, S., Dadgar, A., Krost, A., Cimalla, V., Ambacher, O., Stephan, R., Hein, M.A.: Two-dimensional electron gas based actuation of piezoelectric AlGaIn/GaN microelectromechanical resonators. *Appl. Phys. Lett.* **93**, 173504 (2008)
 20. Birner, S.: Nextnano³ Tutorial. www.nextnano.de (2012)

direct exposure to their toxic components. Recent advancements in polymer technology have also produced polymers able to change properties over time or according to environmental conditions (temperature, pH, etc.). Both natural and synthetic polymers can be used to modify nanomaterials through wet and dry surface–surface coupling techniques. Polymer coatings are currently considered as a valuable tool for the improvement of nanomaterials due to the fine control and tuning of their properties and due to the ability to add multiple functionalities to the grafted surfaces.

Polyethylene Carbonate (PEC)

► [SU-8 Photoresist](#)

Polymer Coatings

Enrica De Rosa¹, Joseph Fernandez-Moure² and Ennio Tasciotti¹

¹Nanomedicine Department, The Methodist Hospital Research Institute, Houston, TX, USA

²The Methodist Hospital Research Institute, Houston, TX, USA

Synonyms

[Polymeric surface modifications](#)

Definition

A polymer coating is defined as the chemical or physical binding of polymeric molecules on the surface of a substrate material.

Summary

In the last two decades, the physical, chemical, and functional properties of polymers have been exploited to enhance the properties of nanomaterials. Polymer coatings are capable to shield nanoparticles from adverse physiological environments, retard their degradation or erosion, or to protect the body from the

Introduction

Natural and synthetic polymers have been used in medical applications for decades thanks to their biocompatibility, biodegradability, and tunable properties. They have been injected, implanted and used to host cells, to release drug, replace tissues, or drive tissue regeneration. Historically, polymer coatings have been used in the surgical setting to improve prosthesis biocompatibility, facilitate bone implant fixation, and reduce fibrosis and contracture of mammary implants. Since the introduction of polymer coating in the field of medicine and surgery, their use has applied to the treatment, diagnosis, and monitoring of more complex pathologies such as cancer, cardiovascular disorders, and degenerative or metabolic diseases. The ability to obtain thin polymeric coatings and to control their self-assembly, make their use advantageous and versatile especially in combination with nanomaterials and in particular with nanoparticles. This integration has served to not only protect nanoparticles from sequestration or degradation but also to lend them multiple functionalities. In essence, polymers can be considered the true enabler of the use of nanoparticles in the biomedical field and today multifunctional nanostructures combine the biomimetic and targeting properties enabled by the presence of complex coatings, with their unique properties like remote control and photothermal activation.

Polymer Coating of Nanoparticles

The crucial aspect of any therapeutic treatment is the identification of the cellular and molecular targets and the localization of the therapeutic molecules in the right

place and at the right time and dose. Nanotechnology offers solutions to address the biodistribution and bioavailability of the drug of choice and to achieve higher therapeutic efficacy and reduce systemic toxicity [1]. Some nano-based delivery systems lack the ability to control their trafficking inside the cell where biological barriers such as the lysosomal sequestration limit the advantage of using nanoparticles. To overcome these limitations, physical or chemical coatings of nanomaterials with biocompatible and biofunctional polymers have been the focus of many laboratories in the past two decades [2]. By chemically conjugating the appropriate ligands, antibodies, or targeting moieties to the reactive functional groups on the polymercoating, nanoparticles have been targeted to specific body districts as needed [3]. Coating nanoparticles with hydrophilic polymers such as poly(ethylene glycol) (PEG) reduces the interfacial energy in an aqueous environment, preventing unwanted aggregation. Nanoparticles coated with hydrophilic polymers are less recognized by macrophages and opsonins in the body. This allows the nanoparticles to stealthily circulate in the blood for a longer period of time, thus increasing their possibility to reach their target site [4]. Some of the substrate materials used to synthesize nanoparticles (iron oxide nanocrystals, gold nanoshells, quantum dots, and nanofibers), and nanodevices (nanosensors, micro-, and nanofluidic devices) [5], include metal and metal oxide particles (gold, silver, iron, zinc, and silicon) [6]. These compounds are ideally suited for many biomedical applications given their ability to be tracked in vivo, to be triggered by external fields or forces, and to react with the tissue microenvironment. The drawback of some of these nanoparticles (such as quantum dots made of an inorganic mixture of cadmium selenide and zinc sulfide) is their inherent toxicity [7]. Natural and synthetic polymers can be used to increase the biocompatibility of this class of nanoparticles and made possible their use in the clinical and preclinical settings.

Natural Biomaterials and Polymers

A variety of natural biomaterials and polymers can be adapted for biomedical applications. Because of their intrinsic biocompatibility, they represent an excellent

solution for in vitro and in vivo applications. They include polysaccharides, proteins, peptides, and DNA.

Polysaccharides Starch, chitosan, dextran, agarose, alginate, and cellulose are a class of long-chain sugars and carbohydrates consisting of repeated monosaccharide units. Coating using these materials not only improves biocompatibility and the colloidal stability of nanoparticles in a physiological environment, but also provides binding sites for the further functionalization of nanoparticles with various biological molecules such as peptides, antibodies, enzymes, DNA, and RNA [8]. Dextran coating was initially introduced to reduce particles' toxicity [9]. Glycochitosans, hydrophobically modified polysaccharides able to self-assemble in a core-shell structure, were used to reduce tumor growth through the delivery of RGB peptides and paclitaxel [10]. Heparin nanoparticles have prolonged the half-life and ID50 of the encapsulated drug and resulted in increased therapeutic efficiency [11]. DNA layer-by-layer capsules, able to shrink in response to environmental stimuli, has been proposed for the controlled release of drugs [12].

Proteins, peptides, and poly(amino acids) may provide hydrophobic, hydrophilic, or amphiphilic coatings to nanoparticles due to the secondary and tertiary arrangement of their moieties. Bovine serum albumin (BSA), collagen, gelatin, and silk are widely used as protein-based coatings for tissue engineering scaffolds and biomimetic approaches [13]. Proteins such as β -galactosidase were used to enhance immune response [14], while proteins that undergo solution pH-dependent association and dissociation could be used to induce reversible association-dissociation transitions by varying pH [15]. These coatings can be used to target the nanoparticles to a specific tissue or organ in the body and to respond to the local microenvironment to activate the release of the drug [16]. The main drawback of these applications is the long-term retention of the biological activity of the protein coating in circulation which still remains a major challenge of the field [17].

DNA, or deoxyribonucleic acid, is a complex polymer based upon nucleic acid units bound by ionic and hydrogen-bonds. Multilayered DNA coatings have been proven to be cyto- and histocompatible and to increase fibroblast proliferation without altering cell morphology [8]. It was reported that DNA coatings on magnetic Fe_2O_3 nanoparticles could decrease the magnetic interparticle interaction, while retaining individual magnetism of Fe_2O_3 nanoparticles [17]. DNA

coatings affected the osteoblast-like cell differentiation through an increase in osteocalcin deposition indicative of enhanced differentiation [18].

Synthetic Polymers

Numerous synthetic polymers have been developed to replace or mimic naturally occurring molecules for various bio-applications, most of which derive from hydrocarbon-based monomers. They have been used in the coating of nanoparticles for therapeutic delivery, evasion of cellular sequestration, and environmental interaction. Although varied in chemical formulation, the principles of biocompatibility and biodegradability inherent to these compounds are paramount to their success and to the success of other novel formulations.

Poly(ethylene glycol) (PEG) is a hydrophilic polymer that can be chemically/physically cross-linked via acrylate-modification to form a hydrogel. Based on the method of cross-linking, layer thickness, physical and chemical properties of PEG-based hydrogel may be precisely tuned. Its chemical structure allows for limited protein adhesion or immune response. PEG coatings can therefore improve the biocompatibility of nanoparticles while simultaneously preventing their aggregation, thereby enhancing their colloidal stability in body fluid [19]. PEG modification also provides protection for nanoparticles from being taken up by macrophages and monocytes, allowing them to temporarily avoid the mononuclear phagocyte system and have a prolonged circulation time in vivo. Such polymers prevent the particle surface from adsorbing plasma components due to the steric barrier, and therefore accumulation of blood proteins and recognition by macrophages is greatly reduced [20]. In vivo studies observed that PEG coating allowed formation of stealth particles that globally distributed much better through the animal, while the PEG molecular weight did not affect probe biodistribution [21]. A successful example of PEG-lipid conjugation now advanced into clinical testing phases is DOXIL, a liposomal formulation of the chemotherapeutic agent, doxorubicin (DOX). The in vivo data showed 90% suppression of 4T1 murine mammary carcinomas and reduced toxicity compared to DOX [22]. DOXIL is now FDA approved for ovarian and breast cancer treatment. However, recent studies also assessed that repeated intravenous injections of

PEGylated liposomes are characterized by reduced circulation time [23]. This finding suggests that prolonged treatment with PEGylated liposomes can generate antibodies to PEG and lead to unexpected pharmacokinetic behavior.

Another synthetic polymer is *polyvinyl alcohol (PVA)* prepared by hydrolysis of polyvinyl acetate. It is biocompatible, biodegradable, and readily functionalized. PVA is water soluble, but resistant to oils and organic solvents, thus it can protect nanoparticles from agglomeration and allow them to form uniform colloids in aqueous biological environment [24].

Polyethylenimine (PEI), another synthetic polymer, has been widely used for gene delivery. PEI can form complexes with DNA and condense them into compact nanoparticles [25]. It is also capable of destabilizing the endosomal membrane by triggering osmotic swelling in the endosomes, to increase DNA release [26]. PEI has become one of the most efficient nonviral transfection agents. The PEI coated nanoparticles have been used as a successful vector for gene delivery [27]. In the last decades, different PEI modifications, including thiol cross-linking, polymer size, covalent attachment, and branching degrees, achieved high transfection efficiency of plasmid DNA and siRNA into cells and live animal [28]. It is important to consider that high molecular weight PEI can cause toxicity. Recent studies have shown that it is possible to reduce toxicity maintaining effective nucleic acid delivery by using intermediate length polymers, and coating with low molecular weight PEI polymers, while inefficient for gene delivery, is nonetheless efficient to increase delivery of antitumor agent into cancer cells [29].

Methods of Surface Coating

The surface properties of nano and microparticles can be tuned by coating a thin film of polymeric materials to achieve the desired physical, chemical, optical, electrical, mechanical, and biomedical properties. There are two conventional approaches for nano- and microparticle coating, categorized as dry or wet methods depending on the environment of polymerization reaction [30]. Dry methods include: physical vapor deposition, chemical vapor deposition, plasma treatment, and pyrolysis of polymeric or nonpolymeric

polyethylene glycol organic materials for in situ precipitation of nanoparticles. Wet methods include: sol-gel processes, emulsification and solvent evaporation techniques, and layer-by-layer assembly methods (Table 1).

Dry Surface Coating

The two principal methods for dry coating of nanomaterials are *physical vapor deposition (PVD)* and *chemical vapor deposition (CVD)*. PVD is a technique that condenses vaporized materials in vacuum and at high temperature, to form a thin film coating onto the substrate surface. The vaporized materials are usually carried by H₂, N₂, or inert gases and there is no chemical reaction involved in this process [31]. CVD is a chemical method to deposit gaseous reactants to the surface of substrate materials to form a thin solid film through chemical reaction or decomposition of the precursor in an activated (heat, light, plasma) environment. CVD method has been used as a coating method for various materials including nano/microparticles, nanocrystals, nanowires, etc. [32].

Another option for dry coating is *plasma treatment* which can be used to directly attach polymers to a desired surface generating highly branched or cross-linked polymers. The principle of this technique is that the ionized and excited molecules and radicals activated by the electrical field collide and react with the substrate surface [33]. The plasma technique is environmentally friendly and usually performed at room temperature, which provides this technique with flexibility for surface modification or coating of almost all substrates. However, there are challenges to the deposition of thin films on nanoparticles due to their severe aggregation and high surface to mass ratio. In order to achieve uniform coating by this

technique, it is necessary to homogeneously expose the surface to the plasma as any unexposed area will be minimally coated. To solve this problem, a fluidized bed reactor has been used to serve as an ideal tool for gas-particle reactions. The combination of these two technologies provides an innovative way for nanoparticle surface modification [34].

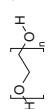

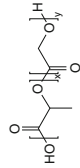
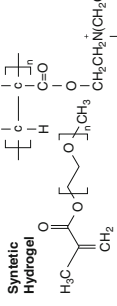
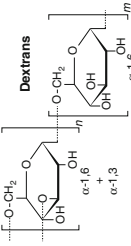
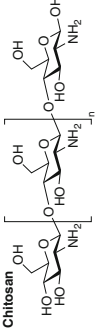
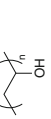
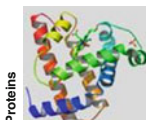

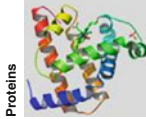
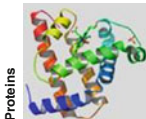
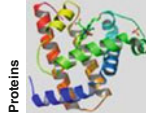
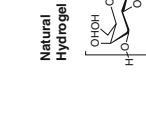
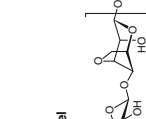
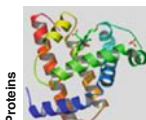

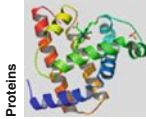
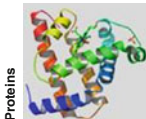
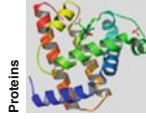
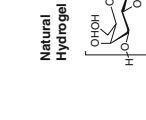
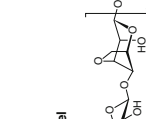
Wet Surface Coating

Wet surface coating refers to the coating of nanoparticles in a liquid (usually aqueous) solution. It can be achieved by *sol-gel processes*, emulsification/evaporation, or layered assembly. Sol-gel is a chemical technique where a chemical or colloid solution (sol) acts as a precursor for an integrated network nano/microparticles (gel). This method is characterized by a gel-like diphasic system containing both a liquid and a solid phase. It can be applied to the synthesis of solid materials (such as glass or ceramics) at lower temperatures than the conventional approaches or in nano/microparticle coating. A thin film is deposited onto the surface of the nano/microparticles and the composition of the film can be tuned by controlling the reaction parameters (temperature, composition, etc.) [35]. *Emulsification* involves two or more immiscible liquids: a liquid phase (dispersed phase), composed of the desired coating materials dispersed into small droplets in a second phase (continuous phase) by constant stirring. Surfactants can be added to the second phase to stabilize the newly formed droplets of the dispersed phase against the coalescence. By resuspending metal, inorganic, or organic nanoparticles into the dispersed phase, their surface is coated with thin polymer layers [36]. *Layer-by-layer (LbL) assembly*, a versatile and easy to use technique, has become a nanofabrication method for a broad range of

Polymer Coatings, Table 1 Summary of nanomaterial substrates, coating materials, methods, and applications

Nanomaterial	Coating material	Coating method	Function/application
Nanoparticles	Synthetic	Emulsion/solution	Controlled drug delivery
Nanoporous microparticles	PEG	CVD	Prolonged bioavailability
Nanowires	PVA	PVD	Improved stability/stabilized suspension
Implant	PEI	Chemical/physical binding	Enhanced loading efficiency
Microfluidic device	PLGA/PLLA	Electro-nanospray	Modified bio-interaction
	PCL	Plasma	
	Natural	LBL	
	Polysaccharide		
	Protein/peptide		
	DNA/RNA		

Polymer Coatings, Table 2 Scheme of polymeric coating applications and relative coating materials

Prolong bioavailability		Control payload delivery	Environmental response	Enhance biocompatibility	Targeting biosensing	Improve stability
<p>PEG</p> 	<p>PEI</p> 	<p>PLGA</p> 	<p>Synthetic Hydrogel</p> 	<p>Dextrans</p> 	<p>Chitosan</p> 	<p>PVA</p> 
 <p>Proteins</p>	 <p>Proteins</p>	 <p>Proteins</p>	 <p>Proteins</p>	 <p>Proteins</p>	 <p>Proteins</p>	 <p>Natural Hydrogel</p>
 <p>DNA</p>	 <p>DNA</p>	 <p>DNA</p>	 <p>DNA</p>	 <p>DNA</p>	 <p>DNA</p>	 <p>DNA</p>

applications. Basically, when LbL assembly is applied to nanoparticle coating, it is used to sequentially absorb positively or negatively charged species onto the nanoparticles by alternatively submerging the different materials into solutions. The multilayered coatings can alter the intrinsic surface characteristics of nanoparticles for desired physical, chemical, mechanical, thermal, and optical properties pertaining to diverse biomedical applications [37] as in the following examples. The electrostatic and covalent LbL assembly strategy of Dhuram et al. has been used to engineer nanoparticles with multilayer shells, small size distribution, high stability in physiological media, low toxicity, protection against macrophage uptake, and triggered release and activation only after endocytosis and enzymatic cleavage [38]. Another successful modification of the LbL method has been developed by Caruso and coworkers using colloidal particles to coat the surfaces of nanomaterials [39]. LbL assembly of a lipid bilayer and polyelectrolyte structures can be used in drug and gene delivery, thanks to the lipid bilayer's ability of accommodating biomaterials. Katsuhiko Ariga's group has developed mesoporous nanocompartment films through LbL assembly of mesoporous silica capsules. These films were able to modulate the release of the payload by means of a nonequilibrium evaporation process [40]. Mohwald's emulsion encapsulation combines the advantages of direct encapsulation of a liquid colloidal core with the accuracy and multifunctionality of LbL polyelectrolyte deposition [41]. Their capsules show two different destruction mechanisms upon interaction with solid surface and hence may find interesting application in pharmacy and medicine industry. Synthesis through LbL is a very flexible and easy coating method that may provide multiple functions and enhance particle functionality.

Future Perspectives

Polymer coating is a versatile powerful tool in nanomedicine and all evidence suggests that its use will continue to grow. The different polymer coatings developed so far are able to overcome several limitations, particularly in drug delivery and nanotechnology, such as biocompatibility, bioavailability, and gene delivery efficiency (Table 2). Furthermore, advances in polymer chemistry and nanotechnology will continue to provide additional compounds to

develop more effective nanostructures able to target the desired tissue and to release the entire payload in a successful fashion. Through the exploration of polymer combinations and through the advent of new coating techniques, it has been possible to pave a new avenue in the biomedical and commercial application of polymers. As polymer coating continues to add more functionalities to nanodevices and nanoparticles, the innovation in their synthesis and modification will continue to bring a fresh breath of life into the field and contribute to the use of nanomaterials to the growing biomedical field. In particular, environmental responsiveness offers a unique opportunity to add functionality to the already versatile nanosystems. By being able to interact with the surrounding environment, polymer coatings can generate a new paradigm of smart therapeutics that only act in the presence of a specific stimulus. Finally, the versatility and variety of polymer coatings could offer a solution to the issue of patient-to-patient variability encountered in any approach of personalized medicine. By providing an optimal therapeutic regimen for each individual patient in response to their specific condition and background, they might facilitate the introduction of the concept of individualized therapy in the clinical arena.

Cross-References

- ▶ [Chemical Vapor Deposition \(CVD\)](#)
- ▶ [Nanomedicine](#)
- ▶ [Nanoparticles](#)
- ▶ [Nanostructured Functionalized Surfaces](#)
- ▶ [Nanostructures for Surface Functionalization and Surface Properties](#)
- ▶ [Nanotechnology](#)
- ▶ [Physical Vapor Deposition](#)
- ▶ [Self-assembly](#)
- ▶ [Sol-gel Method](#)

References

1. Sahoo, S.K., Parveen, S., Panda, J.J.: The present and future of nanotechnology in human health care. *Nanomedicine* **3**, 20–31 (2007)
2. Mikhaylova, M., Kim, D.K., Bobrysheva, N., Osmolowsky, M., Semenov, V., Tsakalakos, T., Muhammed, M.: Superparamagnetism of magnetite nanoparticles: dependence on surface modification. *Langmuir* **20**, 2472–2477 (2004)

3. Wang, L., Zhao, W., Tan, W.: Bioconjugated silica nanoparticles: development and applications. *Nano Res.* **1**, 99–115 (2008)
4. Park, J.H., Lee, S., Kim, J.-H., Park, K., Kim, K., Kwon, I.C.: Polymeric nanomedicine for cancer therapy. *Prog. Polym. Sci.* **33**, 113–137 (2008)
5. Moghimi, S.M., Hunter, A.C., Murray, J.C.: Nanomedicine: current status and future prospects. *FASEB J.* **19**, 311–330 (2005)
6. Ju-Nam, Y., Lead, J.R.: Manufactured nanoparticles: an overview of their chemistry, interactions and potential environmental implications. *Sci. Total Environ.* **400**, 396–414 (2008)
7. Clapp, A.R., Goldman, E.R., Mattoussi, H.: Capping of CdSe-ZnS quantum dots with DHLA and subsequent conjugation with proteins. *Nat. Protoc.* **1**, 1258–1266 (2006)
8. van den Beucken, J.J.J.R., Walboomers, X.F., Vos, M.R.J., Sommerdijk, N.A.J.M., Nolte, R.J.M., Jansen, J.A.: Biological responses to multilayered DNA-coatings. *J. Biomed. Mater. Res. B* **81B**, 231–238 (2007)
9. Lemarchand, C., Gref, R., Passirani, C., Garcion, E., Petri, B., Muller, R., Costantini, D., Couvreur, P.: Influence of polysaccharide coating on the interactions of nanoparticles with biological systems. *Biomaterials* **27**, 108–118 (2006)
10. Kim, J.H., Kim, Y.S., Kim, S., Park, J.H., Kim, K., Choi, K., Chung, H., Jeong, S.Y., Park, R.W., Kim, I.S., Kwon, I.C.: Hydrophobically modified glycol chitosan nanoparticles as carriers for paclitaxel. *J. Control. Release* **111**, 228–234 (2006)
11. Park, K., Lee, G.Y., Kim, Y.S., Yu, M., Park, R.W., Kim, I.S., Kim, S.Y., Byun, Y.: Heparin-deoxycholic acid chemical conjugate as an anticancer drug carrier and its antitumor activity. *J. Control. Release* **114**, 300–306 (2006)
12. Johnston, A.P.R., Caruso, F.: Stabilization of DNA multilayer films through oligonucleotide crosslinking. *Small* **4**, 612–618 (2008)
13. Quaroni, L., Chumanov, G.: Preparation of polymer-coated functionalized silver nanoparticles. *J. Am. Chem. Soc.* **121**, 10642–10643 (1999)
14. Cui, Z., Mumper, R.J.: Coating of cationized protein on engineered nanoparticles results in enhanced immune responses. *Int. J. Pharm.* **238**, 229–239 (2002)
15. Mezziani, M.J., Sun, Y.-P.: Protein-conjugated nanoparticles from rapid expansion of supercritical fluid solution into aqueous solution. *J. Am. Chem. Soc.* **125**, 8015–8018 (2003)
16. Lutolf, M.P., Hubbell, J.A.: Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nat. Biotechnol.* **23**, 47–55 (2005)
17. Dutta, P., Sehra, M.S.: Effect of DNA coating on magnetic relaxation in gamma Fe₂O₃ nanoparticles. *IEEE Trans. Magn.* **43**, 2468–2470 (2007)
18. Van den Beucken, J.J., Walboomers, X.F., Leeuwenburgh, S.C., Vos, M.R.: DNA-coatings: bioactive properties and effects on osteoblast-like cells. *Key Eng. Mater.* **361–363**, 605–608 (2008)
19. Jiang, J.-S., Gan, Z.-F., Yang, Y., Du, B., Qian, M., Zhang, P.: A novel magnetic fluid based on starch-coated magnetite nanoparticles functionalized with homing peptide. *J. Nanopart. Res.* **11**, 1321–1330 (2009)
20. Mu, L., Seow, P.H., Ang, S.N., Feng, S.S.: Study on surfactant coating of polymeric nanoparticles for controlled delivery of anticancer drug. *Colloid Polym. Sci.* **283**, 58–65 (2004)
21. Maldiney, T., Richard, C., Seguin, J., Wattier, N., Bessodes, M., Scherman, D.: Effect of core diameter surface coating, and PEG chain length on the biodistribution of persistent luminescence nanoparticles in mice. *ACS Nano* **5**, 854–862 (2011)
22. Charrois, G.J., Allen, T.M.: Multiple injections of pegylated liposomal doxorubicin: pharmacokinetics and therapeutic activity. *J. Pharmacol. Exp. Ther.* **306**, 1058–1067 (2003)
23. Ishida, T., Masuda, K., Ichikawa, T., Ichihara, M., Irimura, K., Kiwada, H.: Accelerated clearance of a second injection of PEGylated liposomes in mice. *Int. J. Pharm.* **255**, 167–174 (2003)
24. Lin, C., Tsai, H., Tung, C.: Preparation of silver nanoparticles/pseudo-thermoplastic polyvinyl alcohol (PT-PVA) films by the synchronous chemical reduction method. *Polym. Plast. Technol. Eng.* **48**, 1171–1175 (2009)
25. Dunlap, D.D., Maggi, A., Soria, M.R., Monaco, L.: Nanoscopic structure of DNA condensed for gene delivery. *Nucleic Acids Res.* **25**, 3095–3101 (1997)
26. Remy-Kristensen, A., Clamme, J.P., Vuilleumier, C., Kuhry, J.G., Mely, Y.: Role of endocytosis in the transfection of L929 fibroblasts by polyethylenimine/DNA complexes. *Biochim. Biophys. Acta* **1514**, 21–32 (2001)
27. McBain, S.C., Yiu, H.H.P., ElHaj, A.J., Dobson, J.: DNA delivery using polyethyleneimine (PEI) coated iron oxide-silica mesostructured particles. *Stud. Surf. Sci. Catal.* **165**, 869–872 (2007)
28. Kircheis, R., Schuller, S., Brunner, S., Ogris, M., Heider, K.H., Zauner, W., Wagner, E.: Polycation-based DNA complexes for tumor-targeted gene delivery in vivo. *J. Gene Med.* **1**, 111–120 (1999)
29. Xia, T., Kovichich, M., Liong, M., Meng, H., Kabehie, S., George, S., Zink, J.I., Nel, A.E.: Polyethyleneimine coating enhances the cellular uptake of mesoporous silica nanoparticles and allows safe delivery of siRNA and DNA constructs. *ACS Nano* **3**, 3273–3286 (2009)
30. Wang, X., Zhou, L., Ma, Y., Li, X., Gu, H.: Control of aggregate size of polyethyleneimine-coated magnetic nanoparticles for magnetofection. *Nano Res.* **2**, 365–372 (2009)
31. Choy, K.L.: Chemical vapour deposition of coatings. *Prog. Mater. Sci.* **48**, 57–170 (2003)
32. Zhang, B., Liao, Y.C., Girshick, S.L., Roberts, J.T.: Growth of coatings on nanoparticles by photoinduced chemical vapor deposition. *J. Nanopart. Res.* **10**, 173–178 (2008)
33. Joan, C.M.: Plasma Polymerization, vol. 108. American Chemical Society, Washington, DC (1979)
34. Tatoulian, M., Brétagnol, F., Arefi-Khonsari, F., Amouroux, J., Bouloussa, O., Rondelez, F., Paul, A.J., Mitchell, R.: Plasma deposition of allylamine on polymer powders in a fluidized Bed reactor. *Plasma Processes Polym.* **2**, 38–44 (2005)
35. Klein, L.C.: Sol-gel Optics: Processing and Applications, vol. 80. Academic, Boston (1994)
36. Hielscher, T.: Ultrasonic production of nano-size dispersions and emulsions. In: Proceedings of European Nanosystems Conference ENS'05, Paris (2005)
37. Varahramyan, K., Lvov, Y.: Nanomanufacturing by layer-by-layer assembly - from nanoscale coating to device applications. *Proc. Inst. Mech. Eng. Part N J. Nanoeng. Nanosyst.* **220**, 29–37 (2006)
38. Schneider, G.G.F., Subr, V., Ulbrich, K., Decher, G.: Multifunctional cytotoxic stealth nanoparticles a model

- approach with potential for cancer therapy. *Nano Lett.* **9**, 636–642 (2009)
39. Katagiri, K., Caruso, F.: Functionalization of colloids with robust inorganic-based lipid coatings. *Macromolecules* **37**, 9947–9953 (2004)
40. Ariga, K., Vinu, A., Hill, J.P., Mori, T.: Coordination chemistry and supramolecular chemistry in mesoporous nanospace. *Coord. Chem. Rev.* **251**, 2562–2591 (2007)
41. Grigoriev, D.O., Bukreeva, T., Mohwald, H., Shchukin, D.G.: New method for fabrication of loaded micro- and nanocontainers: emulsion encapsulation by polyelectrolyte layer-by-layer deposition on the liquid core. *Langmuir* **24**, 999–1004 (2007)

Polymer Nanocapsules

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

Polymer Pen Lithography

- ▶ [Dip-Pen Nanolithography](#)

Polymer Sensors

- ▶ [Organic Sensors](#)

Polymeric Surface Modifications

- ▶ [Polymer Coatings](#)

Polymer–Metal Oxide Solar Cells

- ▶ [Hybrid Solar Cells](#)

Polymethyl Methacrylate (PMMA)

- ▶ [SU-8 Photoresist](#)

Polypropylene Carbonate (PPC)

- ▶ [SU-8 Photoresist](#)

Post-exposure Bake (PEB)

- ▶ [SU-8 Photoresist](#)

Precritical Clusters

- ▶ [Prenucleation Clusters](#)

Prenucleation Clusters

Denis Gebauer
Department of Chemistry, Physical Chemistry,
University of Konstanz, Konstanz, Germany

Synonyms

[DOLLOP](#); [Precritical clusters](#); [Stable ion clusters](#)

Definition

Prenucleation clusters are thermodynamically stable and small (~0.5–2 nm) aggregates of ions in aqueous solution, which are solutes and form prior to the nucleation of a solid phase, in undersaturated as well as supersaturated (metastable) solution states.

Overview

Relevance

How crystallization may be controlled is an important question, since many processes and applications depend on the control of crystal polymorphism, morphology, size, and other crystal characteristics. However, today, the design of bottom-up approaches to tailor-made crystalline materials is widely settled in

the realm of trial and error. In order to improve these approaches, (nano)technology depends on a basic understanding of the underlying processes. A look at living nature shows that organisms are able to control these processes, and form mineralized structures – biominerals like bone, teeth, or nacre [1] – that show outstanding material properties [2] based on sophisticated micro- and nanostructures. Many studies have shown that biomineralization processes can proceed via nonclassical and particle-mediated pathways [3], which may employ amorphous intermediates [4]. The physicochemical basis of these alternative pathways remains largely unknown, but it is evident that they do not conform to the classical theories presented in crystallization textbooks.

The recent discovery of stable prenucleation clusters [5, 6] may be an important step toward a novel understanding of the most fundamental step in crystallization: nucleation. The stable clusters are not taken into account within the classical concept, and may be the clue to explain the nonclassical and particle-mediated crystallization pathways as well as observations in bio- and biomimetic mineralization studies (cf. Key Research Findings). This may lead to a novel understanding of the nucleation and crystallization of minerals.

Background: Classical Nucleation Theory

Nucleation means the onset of a phase transition, that is, the formation of the first nuclei of a nascent phase in a system that has become supersaturated. The work of Becker and Döring is commonly known as classical nucleation theory (CNT), which was originally derived for the formation of nuclei from supersaturated water vapors [7]. The fundamental assumption of CNT is that the bulk energy of a forming nucleus drives nucleation, while the generation of a phase interface, and with it, of interfacial tension, is the antagonist, and impedes nucleus growth [8]. The bulk energy (favorable, negative sign) scales with the cube of the radius of the nuclei, and their surface energy (unfavorable, positive sign) increases with the square of the radius. Hence, the bulk energy begins to balance the energetic costs due to the generation of a phase interface at a certain point (Fig. 1). A nucleus of this size is the critical nucleus. Nuclei that are smaller are thermodynamically unstable and dissolve again; the critical size corresponds to a metastable state of equilibrium, and infinitesimal change will lead to either nucleus dissolution or growth

without limit. The change in free enthalpy for the formation of (pre-) critical nuclei is positive, hence, formation of these species is thermodynamically improbable, and can only occur through fluctuations on microscopic scales.

CNT makes several assumptions, among which the so-called capillary assumption is the most debatable: Nanoscopic nuclei of a nascent phase behave as if they were macroscopic [9]. There have been several empiric and semiempirical treatments of CNT that aimed at the improvement of quantitative predictions. Under purely qualitative considerations, though, CNT is successful in describing many nucleation phenomena. As to CNT, nucleation is a stochastic process that can hardly be controlled in a sophisticated manner.

Prenucleation Clusters: A Nonclassical Concept

In protein crystallization, the so-called two-step mechanism involving the formation of protein clusters is well known [10]. In case of minerals, stable prenucleation clusters were first described for calcium carbonate [5] and imaged for the first time by means of cryogenic Transmission Electron Microscopy (cryo-TEM) by Pouget et al. (Fig. 2) [6]. Prenucleation clusters have been qualitatively evidenced for calcium oxalate and calcium phosphate [5], where they appear to play an important role during the mineralization of organic matrices [11, 12]. This contribution is focused on prenucleation clusters of calcium carbonate, which have been physicochemically characterized, but the presented physicochemical aspects likely apply qualitatively also to other minerals that form prenucleation clusters. The present view on nucleation of these minerals, corroborated by the cryo-TEM studies, is shown in Fig. 3 and contrasted with the notion of CNT [13]. The prenucleation clusters coexist in thermodynamic equilibrium with the ions, and upon nucleation, amorphous nanoparticles are formed as an intermediate phase. From the amorphous intermediate, eventually, crystals are formed (Fig. 3, bottom). It appears also possible that crystals might be nucleated directly from clusters without the formation of an amorphous intermediate. On the other hand, following CNT, unstable nuclei form on the basis of stochastic fluctuations. The nuclei may reach a critical size, which represents a metastable crystal nucleus that may dissolve again or grow without limit (Fig. 3, top). As opposed to the

Prenucleation Clusters,

Fig. 1 Classical representation of the energy of a nucleus (*blue line*) in dependence of its radius. The surface energy gives a positive (unfavorable, *red line*) and the bulk energy a negative energetic contribution (*green line*), which are proportional to the square and the cube of the radius, respectively. Hence, at the critical size, the bulk term begins to balance the surface term

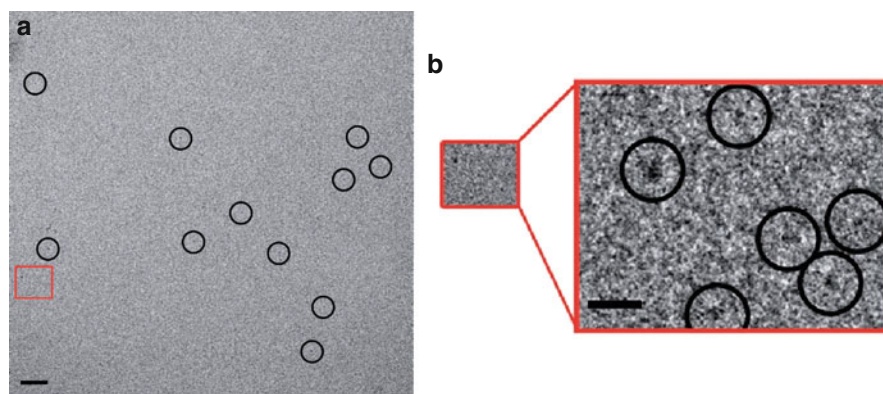
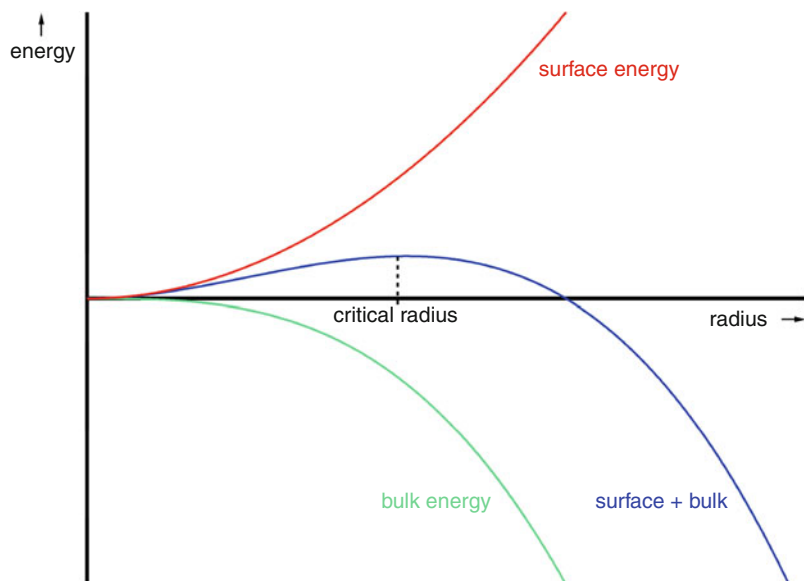
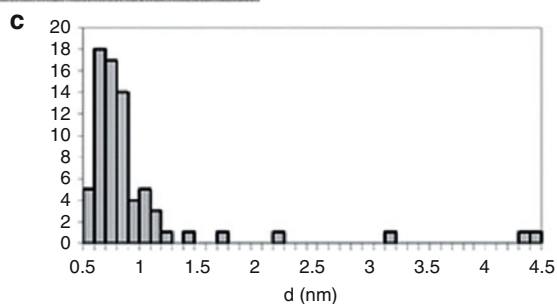
**Prenucleation Clusters,**

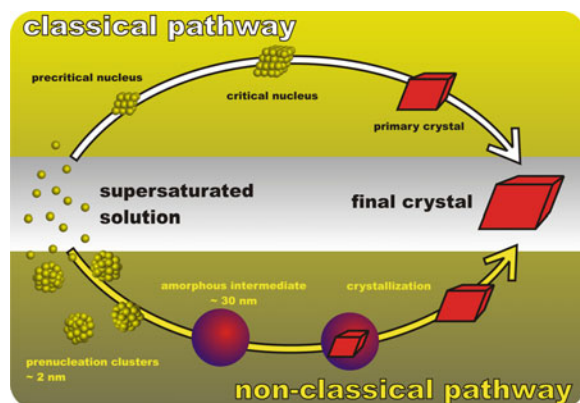
Fig. 2 (a) Cryo-TEM image of calcium bicarbonate solution, pre-nucleation clusters are highlighted with *black circles*; scale bar 20 nm. (b) Magnification of the zone indicated by a *red rectangle* in (a); scale bar 5 nm. (c) Cluster size distribution (diameter d) derived on basis of cryo-TEM data (From Ref. [6]. Reprinted with permission from AAAS)



classical considerations, the complex multistep scenario of the nonclassical pathway renders sophisticated control patterns possible. For example, (macro)molecules may interact with pre-nucleation clusters, amorphous intermediates, as well as growing crystals and influence the course of crystallization in multiple ways [14].

Key Research Findings

Prenucleation clusters are on average neutral in charge [15] and thermodynamically stable. This notion is based on the evaluation of the ion binding in clusters based on a multiple-binding equilibrium, which gives a free enthalpy change for the formation of a CaCO_3



Prenucleation Clusters, Fig. 3 Classical pathway (*top*) from supersaturated solution to crystals as opposed to the nonclassical concept involving pre-nucleation clusters (*bottom*). Possibly, pre-nucleation clusters may directly form crystals (not shown). Reprinted from Ref. [13] with permission from Elsevier.

unit within a cluster of $\Delta G \approx -18 \text{ kJ mol}^{-1}$ [5]. The stability of the clusters, though, depends on the pH value. More stable clusters form at lower (pH ~ 9) than at higher pH (pH ~ 10). Importantly, the physico-chemical characterization strongly suggests that the pre-nucleation clusters do not have a phase interface, and that they have to be considered to be solutes. This notion comes from the fact that the clusters can be detected in a solution state before nucleation that conforms to the solution thermodynamics of a single-phase system [5]. This underlines that the “capillary assumption” of CNT is an oversimplification. The clusters should hence not be called “nanoparticles” as this would imply the presence of a phase interface.

Computer simulation approaches have corroborated that small clusters of calcium carbonate are thermodynamically stable, based on strong hydration and a distinct entropic contribution through a high degree of conformational disorder in chain-like structural forms [16]. The modeling study virtually quantitatively supports the physicochemical characterization of the experimentally found pre-nucleation clusters [5].

As evident from Analytical Ultracentrifugation (AUC), the pre-nucleation clusters do not represent ion pairs [17], but bigger agglomerates of ions. The sedimentation coefficient of these species, which represents the sedimentation velocity per centrifugal field unit, is in the order of 1.4 S, corresponding to a diameter of $\sim 2 \text{ nm}$ assuming spherical clusters with a density that complies with the density of amorphous calcium carbonate. The cryo-TEM studies suggest

that the pre-nucleation clusters are smaller ($\sim 0.6 \text{ nm}$, Fig. 3); however, the cryo-TEM data do not represent the hydration layer, and it remains unclear whether the entire cluster structures exhibit sufficient contrast for imaging in cryo-TEM too [13]. Taken AUC and cryo-TEM together, the data suggest that pre-nucleation clusters contain (few or several) tens of ions on average.

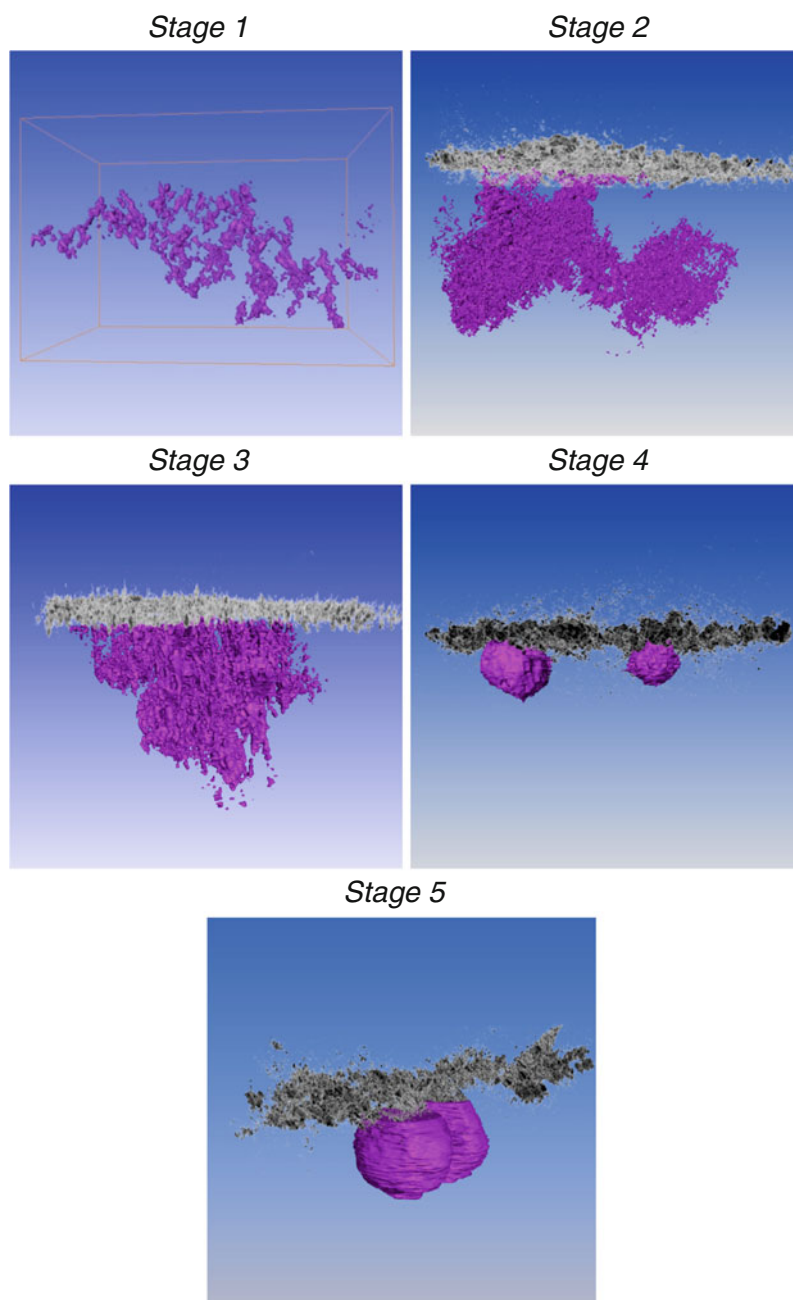
The AUC [5, 6] and cryo-TEM [11] data indicate that nucleation proceeds via the aggregation of clusters that leads to the formation of primary and amorphous nanoparticles of $\sim 30 \text{ nm}$ size. Interestingly, it is known from biomineralization studies that amorphous calcium carbonates may exhibit distinct short-range structural ordering that compares to the crystalline long-range order of different polymorphs [4, 18]. Indeed, the more stable clusters (lower pH) transform into more stable amorphous intermediates than the less stable clusters (higher pH) [5]. Structural analyses show that the more stable amorphous species exhibit short-range structural features that relate to the stable crystalline polymorph (calcite), while the less stable amorphous species relates to the unstable crystalline polymorph (vaterite) [19]. In reverse, these findings suggest that different structures are already present in pre-nucleation clusters, and that polymorphism may be encoded in solution. These findings may be highly important for the topical question of polymorphism control (e.g., [20]).

Cryo-TEM studies suggest that the pre-nucleation clusters play a central role during the heterogeneous nucleation of calcium phosphate from simulated body fluid, which is supersaturated with respect to amorphous calcium phosphate [11]. The study shows that the formation of amorphous mineral particles starts with the aggregation of the pre-nucleation clusters in the presence of an extrinsic interface, that is, the aggregation processes are surface induced (Fig. 4). With it, the nucleation of amorphous calcium phosphate particles occurs only when an organic interface is present; these results appear highly important for the understanding of pathological mineralization in, for instance, cardiovascular disease. On the other hand, pre-nucleation clusters stabilized against aggregation by the presence of an anionic polymer, an inhibitor of nucleation, appear to be sufficient for the internal mineralization of collagen fibers as in bones [12].

Not only extrinsic interfaces but also crystallization additives in solution can interact with pre-nucleation

Prenucleation Clusters,

Fig. 4 Computer-aided 3D visualization of tomograms from cryo-TEM of different stages of the surface-induced mineralization of calcium phosphate from simulated body fluid [11]. Prenucleation clusters (Stage 1) that are approximately 0.9 nm in size aggregate on an extrinsic surface (Stage 2) to form initially amorphous nanoparticles (Stage 4), which subsequently crystallize (Data kindly provided by A. Dey and N. A. J. M. Sommerdijk)



clusters and influence the course of crystallization in multiple ways, producing a “fingerprint” of additive action that is accessible in calcium potential measurements and titration during the early stages of mineralization [14]. This is a first step toward the understanding of the complex mechanisms that underly the control of crystallization by additives.

Future Directions for Research

As the field of prenucleation clusters and nonclassical nucleation [13] is relatively new, many open questions need to be addressed. The following list does not demand to be complete, but reflects some of the most important points and ideas.

- *Structure*: The stoichiometry, size distribution, or the exact arrangement of ions within the prenucleation clusters has to be thoroughly characterized experimentally and compared to the theoretical speciation [16]. Another question is whether distinct structures are actually present in prenucleation clusters and if/how they may relate to the structures of crystalline polymorphs. A related and important question concerns the dynamics of the prenucleation cluster structures.
- *Thermodynamics*: The cluster formation is endothermic as shown by isothermal titration calorimetry [13]; thus, it is evident that cluster formation is entropically driven. A respective physicochemical model is eagerly anticipated. The determination of thermodynamic parameters of prenucleation clusters is required.
- *Mechanism of nucleation*: The mechanism by which the prenucleation clusters transform into primary amorphous nanoparticles is unknown. A central question is how and when a true phase interface is established, upon the aggregation of the clusters into primary nanoparticles. It is also possible that the clusters directly nucleate crystalline particles. Possibly, the generation of the phase interface is triggered and accompanied by structural changes within prenucleation clusters [16]. It appears promising to carefully analyze the kinetics of nucleation taking stable prenucleation clusters into account. Modeling and simulation approaches appear crucial.
- *Nucleation and crystallization control*: Inhibition of nucleation may be achieved through colloidal stabilization of clusters against aggregation [14]. This is important in the view of antiscalants (that inhibit the formation of incrustations in laundry machines or dishwashers, for example) which are so far designed toward their ion-binding capacity. On the other hand, it is well known that additives induce the formation of certain crystal (nano/micro-) structures and orientations [3]. Multiple distinct functions of additives during the early stages of crystallization could be identified [14]. Their detailed understanding may render the design of tailor-made additives for the design of advanced functional materials possible. Synergetic effects of combinations of different additives are expected and may be resolved.
- *Biom mineralization*: It appears evident that prenucleation clusters play important roles during surface-induced crystallization in biomimetic approaches [11, 12]; however, their role in biological systems still remains obscure.
- *Generality*: The prenucleation clusters have been shown to exist for some, mainly sparingly soluble minerals. Likely, the formation of stable prenucleation clusters is a more general phenomenon that may apply to many minerals and salts [13, 16]. The generality of the prenucleation cluster concept should be tested.
- *Nucleation theory*: It is most important to include the stable prenucleation cluster concept into a new, quantitative theory of nucleation. Also in this context, modeling and simulation approaches are crucial [13, 16].

Cross-References

- ▶ [Biomimetics](#)
- ▶ [Nanoparticles](#)
- ▶ [Nanoscale Properties of Solid–Liquid Interfaces](#)
- ▶ [Nanotechnology in Cardiovascular Diseases](#)
- ▶ [Self-assembly of Nanostructures](#)
- ▶ [Transmission Electron Microscopy](#)

References

1. Lowenstam, H., Weiner, S.: On Biom mineralization. Oxford University Press, New York (1989)
2. Smith, B.L., Schaffer, T.E., Viani, M., Thompson, J.B., Frederick, N.A., Kindt, J., Belcher, A., Stucky, G.D., Morse, D.E., Hansma, P.K.: Molecular mechanistic origin of the toughness of natural adhesives, fibers and composites. *Nature* **399**, 761–763 (1999)
3. Cölfen, H., Antonietti, M.: Mesocrystals and Nonclassical Crystallization. Wiley, Chichester (2008)
4. Gower, L.B.: Biomimetic model systems for investigating the amorphous precursor pathway and its role in biom mineralization. *Chem. Rev.* **108**, 4551–4627 (2008)
5. Gebauer, D., Völkel, A., Cölfen, H.: Stable prenucleation calcium carbonate clusters. *Science* **322**, 1819–1822 (2008)
6. Pouget, E.M., Bomans, P.H.H., Goos, J.A.C.M., Frederik, P.M., de With, G., Sommerdijk, N.A.J.M.: The initial stages of template-controlled CaCO₃ formation revealed by Cryo-TEM. *Science* **323**, 1455–1458 (2009)
7. Becker, R., Döring, W.: Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Ann. Phys.* **24**, 719–752 (1935)

8. Kashchiev, D.: *Nucleation: Basic Theory with Applications*. Butterworth-Heinemann, Oxford (2000)
9. Nielsen, A.E.: *Kinetics of Precipitation*. Pergamon, New York (1964)
10. Vekilov, P.G.: The two-step mechanism of nucleation of crystals in solution. *Nanoscale* **2**, 2346 (2010)
11. Dey, A., Bomans, P.H.H., Müller, F.A., Will, J., Frederik, P.M., de With, G., Sommerdijk, N.A.J.M.: The role of prenucleation clusters in surface-induced calcium phosphate crystallization. *Nat. Mater.* **9**, 1010–1014 (2010)
12. Nudelman, F., Pieterse, K., George, A., Bomans, P.H.H., Friedrich, H., Brylka, L.J., Hilbers, P.A.J., de With, G., Sommerdijk, N.A.J.M.: The role of collagen in bone apatite formation in the presence of hydroxyapatite nucleation inhibitors. *Nat. Mater.* **9**, 1004–1009 (2010)
13. Gebauer, D., Cölfen, H.: Prenucleation clusters and non-classical nucleation, *Nano Today* **6**, 564–584 (2011)
14. Gebauer, D., Cölfen, H., Verch, A., Antonietti, M.: The multiple roles of additives in CaCO₃ crystallization: a quantitative case study. *Adv. Mater.* **21**, 435–439 (2009)
15. Wolf, S.E., Müller, L., Barrea, R., Kampf, C.J., Leiterer, J., Panne, U., Hoffmann, T., Emmerling, F., Tremel, W.: Carbonate-coordinated metal complexes precede the formation of liquid amorphous mineral emulsions of divalent metal carbonates. *Nanoscale* **3**, 1158–1165 (2011)
16. Demichelis, R., Raiteri, P., Gale, J.D., Quigley, D., Gebauer, D.: Stable prenucleation mineral clusters are liquid-like ionic polymers. *Nat. Commun.* **2**, 590 (2011)
17. Gal, J.-Y., Bollinger, J.-C., Tolosa, H., Gache, N.: Calcium carbonate solubility: a reappraisal of scale formation and inhibition. *Talanta* **43**, 1497–1509 (1996)
18. Addadi, L., Raz, S., Weiner, S.: Taking advantage of disorder: amorphous calcium carbonate and its roles in biomineralization. *Adv. Mater.* **15**, 959–970 (2003)
19. Gebauer, D., Gunawidjaja, P.N., Ko, J.Y.P., Bacsik, Z., Aziz, B., Liu, L.J., Hu, Y.F., Bergström, L., Tai, C.W., Sham, T.K., Edén, M., Hedin, N.: Proto-calcite and proto-vaterite in amorphous calcium carbonates. *Angew. Chem. Int. Ed.* **49**, 8889–8891 (2010)
20. Raw, A.: Regulatory considerations of pharmaceutical solid polymorphism in abbreviated new drug applications (ANDAs). *Adv. Drug Delivery Rev.* **56**, 397–414 (2004)

PRINT®

- ▶ [Nanoscale Printing](#)

Probe Technology of Scanning Probe Microscopy (SPM) and Its Fabrication Technology

- ▶ [Microfabricated Probe Technology](#)

Production of Gold Nanoparticles

- ▶ [Synthesis of Gold Nanoparticles](#)

Production of Graphene or Graphene Production

- ▶ [Synthesis of Graphene](#)

Propylene Glycol Methyl Ether Acetate (PGMEA)

- ▶ [SU-8 Photoresist](#)

Prostheses

- ▶ [MEMS Neural Probes](#)

Protein Nano-crystallization

- ▶ [Macromolecular Crystallization Using Nanovolumes](#)

Pseudoelasticity

- ▶ [Superelasticity and the Shape Memory Effect](#)

Pulsed Focused Ultrasound

- ▶ [Pulsed High-Intensity Focused Ultrasound](#)

Pulsed HIFU

- ▶ [Pulsed High-Intensity Focused Ultrasound](#)

Pulsed High-Intensity Focused Ultrasound

Brian E. O'Neill¹ and King C. Li²

¹Department of Radiology Research, The Methodist Hospital Research Institute, Houston, TX, USA

²Department of Bioengineering, Rice university, Houston, TX, USA

Synonyms

Pulsed focused ultrasound; Pulsed HIFU

Definition

Pulsed high-intensity ultrasound (pHIFU) is a technique used to physically modify the permeability and hydraulic conductivity characteristics of living tissue in order to increase penetration and accumulation of large molecular drugs or nanoparticles in the interstitial space. Pulsed HIFU is a form of *ultrasound enhanced drug delivery*; however, while the latter may be assisted by the use of exogenous agents such as *ultrasound contrast agents* or *microbubbles*, those who use pHIFU generally rely on the physical interaction of the ultrasound with the tissue alone. The use of pulsed energy deposition is specifically designed to enhance the mechanical over the thermal effects of the focused ultrasound by allowing for deposited heat to dissipate even at high powers. Microbubble contrast agents are avoided to prevent *cavitation* in the tissue.

Ultrasound and the Nanoparticle Delivery Problem

The transport of agents to tissues can be divided into steps based on the different physiological barriers presented, from administration to final interaction with individual cells. The relative importance of each step is dependent on the type of agent (size, surface chemistry), the type of tissue, and the type of administration (intravenous, direct injection). Assuming systemic injection, the amount of agent that ultimately reaches an individual cell depends, in some part, on eight factors: (1) agent-specific blood pool

pharmacokinetics, (2) tissue perfusion, (3) local vascular topography, (4) agent-specific local vascular permeability, (5) agent-specific interstitial conductivity, (6) interstitial topography, (7) cellular affinity and, finally, (8) vascular/lymphatic drainage efficiency. In the case of a direct (intra-tissue) injection, the first four factors are much less important, but probably still play a role in the distribution of agents whose transport is convection dominated.

Nanoparticles pose a specific challenge to the natural transport scheme because of their sizes and surface chemistry. Certain issues, such as blood pool pharmacokinetics, have been largely addressed with various surface coatings and proper particle size; however, the issue of getting nanoparticles across vascular, endothelial, and cell membrane barriers and into the tumor parenchyma and cell is still a challenge. First, within the vasculature of tumors there are pockets of stasis and poor perfusion, due in large part to the improperly constituted and chaotic network of blood vessels in these environments [1]. This situation might be improved by increases in perfusion due to ultrasound hyperthermia or other vascular effects of ultrasound. Huge increases in perfusion are known to accompany hyperthermia in temperature-sensitive tissue such as muscle. Increases in tumor are much more modest and limited [2], but still important. Second, most of the currently successful nanoparticles targeted to solid tumors rely heavily on the atypically high fenestration of the endothelial barrier in tumors compared to most other tissues. This means that particles in the size range of ~50–200 nm in diameter have a tendency to accumulate preferentially in tumor tissue even without further considerations (known as the enhanced permeability and retention or EPR effect). Unfortunately, the exact size cutoff is highly dependent on tumor type, and can be highly variable across tumors and even within the same tumor [3]. In tissues with poor fenestration, there is evidence that ultrasound treatment can enhance the movement of nanoparticles across the endothelium. Third, nanoparticles generally move most efficiently by convection; however, in the tumor microenvironment, convective flow beyond the vasculature is virtually nonexistent. The interstitial conductivity and topography that determine how an agent penetrates and distributes in the interstitial space by either convection or diffusion is not favorable to the movement of nanoparticles through tumors. Cellular crowding due to rapid growth and the production of

a collagenous extracellular matrix can both result in barriers to large particle movement. As well, the lack of lymphatics in tumors contributes to poor drainage, causing high interstitial fluid pressures and sluggish flow. A number of preliminary experiments involving pHIFU and direct intra-tumoral injection of agents have suggested improved distribution in the interstitial space. A final issue is the difficulty of getting nanoparticles into the cell through the cell membrane. Although many potential techniques exist to address this problem, many of them leave the nanoparticle sequestered inside the cell only to be ejected later, or are not practical for clinical work. The effects of ultrasound cavitation can transiently increase membrane permeability (sonoporation), and may be able to provide a physical means of achieving intracellular accumulation of nanoparticles. All these considerations suggest strongly that there is a need to remodel or re-engineer the transportation dynamics to be more favorable to nanoparticle traffic, and that ultrasound treatment provides some opportunity to do so.

Basic pHIFU Physics

Ultrasound is a coherent traveling pressure disturbance that easily transports mechanical and thermal energy through biological soft tissues [4]. As ultrasound travels, the coherent wave is attenuated by absorption and scattering processes. A traveling ultrasonic pressure plane wave might generally be written as

$$p(x, t) = p_0 \sin\left(\omega\left(\frac{x}{c} - t\right)\right) \exp(-\mu_a x),$$

where p_0 is the pressure amplitude, ω is the frequency, c is the speed of sound, and μ_0 is the attenuation of the wave. The intensity of the wave is given by $I = \frac{p_0^2}{2\rho c}$ with ρ the density of the medium. At frequencies around 1 MHz, traveling at a speed of 1,500–1,600 m/s in water or tissue, the ultrasonic wavelength is on the order of 1.5–1.6 mm, much too long to directly couple to objects on the cellular scale. The low compressibility of water means that ultrasonic strains remain quite low even when pressures reach tens of atmospheres. Thus, the pressure disturbance itself, even at powers as high as 10^4 W/cm², is generally regarded as having no direct biological effects, unless cavitation is involved (see below). In the absence of cavitation, part of the

wave energy is still continuously being reflected or absorbed as heat, a fact modeled with an attenuation coefficient. The attenuation coefficient, μ_a , determines how the wave transmits energy and momentum to the media through which it passes. Attenuation in general incorporates both scattering and absorption phenomena; however, often absorption is thought to be dominant. The absorption of the energy creates heat, and this heat per unit volume is given by:

$$\dot{Q} = \mu_a I$$

As well, both absorption and scattering result in a transfer of momentum from the wave to the medium. The resulting body force is known as the radiation pressure and is often written (ignoring scattering) as:

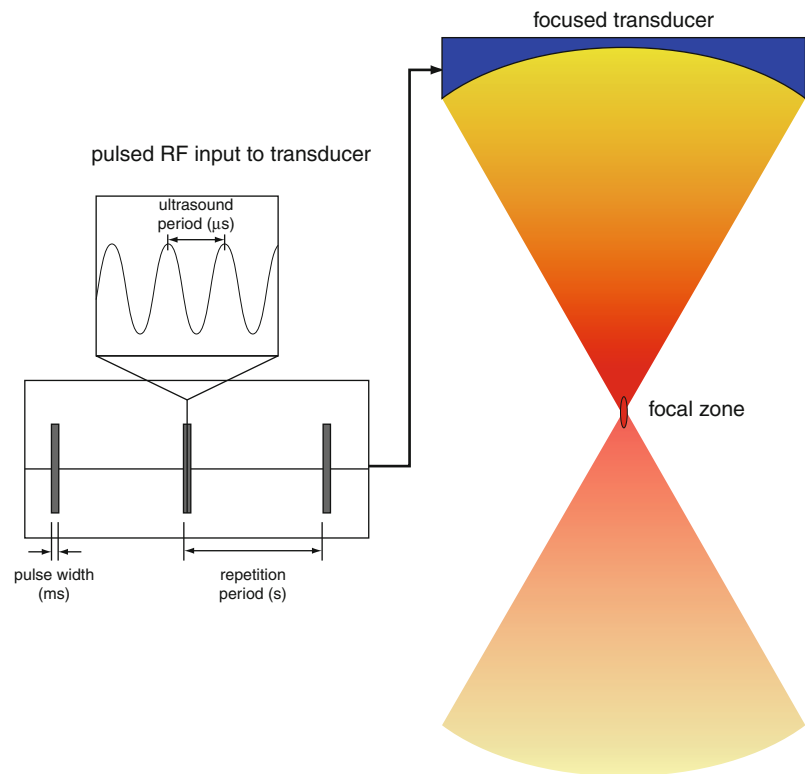
$$F = \frac{\mu_a I}{c}$$

Generally, in the low megahertz frequency range, attenuation is not very large; however, at very high intensity the effect is still biologically significant. Such high intensity may be achieved by concentrating the ultrasound from a large source into a small volume. Focused ultrasound may be achieved by means of a curved transducer or an acoustic lens. Either way, the result is a concentration of much of the acoustic energy produced by the transducer into a diffraction- and aperture-limited ellipsoidal focal spot (Fig. 1). In typical pulsed HIFU experiments, using a 1.0 MHz transducer with equal aperture and focal depth ($f = 1$), the spot size full-width half-maximum is about 1.5 mm in diameter and 7 mm in length. With this design, local peak intensities can range up to 10^4 W/cm². Pulsed HIFU powers are typically maintained in the range of 10^2 – 10^3 W/cm².

Ultrasonic heating is the basis of HIFU ablation and ultrasound hyperthermia treatment. The principal goal of using pulsed, rather than continuous HIFU is to avoid overheating the tissue by allowing the heat deposited in the focal zone to diffuse, while still maintaining high peak powers to induce the mechanical effects. In fact, there will always be some unavoidable heating, but because the thermal response is relatively slow (order of seconds), it is governed by the average power. In contrast, the mechanical response is rapid (order of tens of milliseconds), and thus is related more to the peak power.

Pulsed High-Intensity Focused Ultrasound,

Fig. 1 Schematic of pulsed high-intensity focused ultrasound. Pulsed input to focused ultrasonic transducer results in high peak intensity, low average intensity beams, with highest energy density in and around the focal zone. The focal zone is scanned across the tissue to enable treatment of large volumes

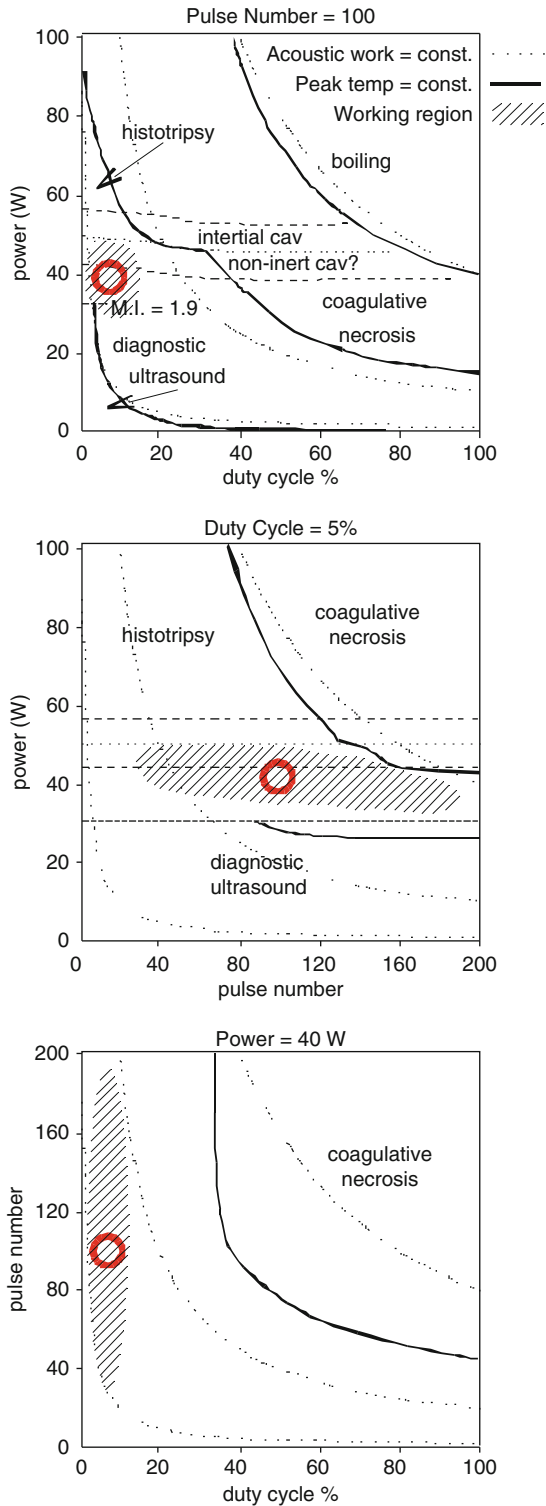


If sufficient levels of gas exist, either dissolved or in suspension, then acoustic cavitation becomes a possibility. In that case, rapid and large expansion of the gaseous phase can happen during the negative pressure portion of the cycle, a phenomenon known as acoustic cavitation [5]. Because of the huge acoustic impedance mismatch between gas and liquid and because of the resonant expansion and contraction of gas microbubbles in an acoustic wave, cavitation produces an enormous concentration of energy in and around the microbubbles that can have a number of biological effects. Acoustic cavitation phenomena are divided into two types: the first, known as non-inertial or stable cavitation, involves gas bubbles resonating around an equilibrium size, while the second, known as inertial or unstable, involves the rapid expansion and subsequent violent collapse of gas bubbles. Generally, the existence of cavitation in tissue is inferred from its effect on the spectral content of the wave; in particular, the non-inertial cavitation will create an unambiguous increase in half harmonic content, while inertial cavitation also causes an increase in broadband noise level. Both processes involve the concentration of acoustic, mechanical, and thermal energy in the neighborhood

of the bubble, except that in the second instance, this energy is focused even more into a tiny volume during collapse, with resulting extremes in temperature and pressure. These in turn can have profound effects on tissue, including enhanced heating, membrane poration, and cell lysis.

Biological Impact of pHIFU

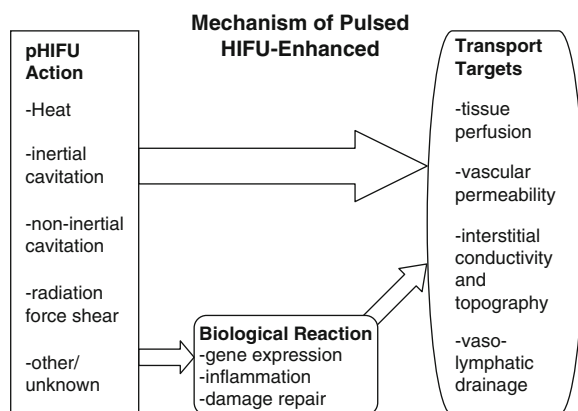
The set of acoustic parameters most often used for pHIFU applications are: 100 pulses at 1 MHz center frequency, 5% duty cycle, and 1 Hz repetition rate (50 ms on, 950 ms off), 40 W peak acoustic power, using an $f = 1$ transducer (approximately $2,200 \text{ W/cm}^2$ in water). Another common set of parameters employed roughly half the power and double the duty cycle. (The earliest studies described below used much higher total energy and peak powers, however, because the efficiency of the transducer was not measured, it is not clear precisely what these were.) Under these study conditions, mostly carried out in mouse models, researchers have demonstrated changes in normal and tumor tissues that effectively increase the rates of



Pulsed High-Intensity Focused Ultrasound, Fig. 2 (continued)

nanoparticle uptake by two- to three folds over untreated tissues. The effects appear to be accompanied by edema and often inflammation. While the precise mechanism remains uncertain, some possibilities exist. It is widely accepted that inertial and possibly non-inertial cavitation could cause such effects. Indeed, introducing cavitating microbubbles in the form of ultrasound contrast agents is a technique that has been employed for blood-brain barrier disruption [6]. However, under the conditions given, there is no evidence that the cavitation threshold is being reached during pHIFU in the absence of exogenous agents. Cavitation is very difficult to achieve in tissues, first, because the levels of available gas are small, and second, because the threshold for cavitation is increased with increasing viscosity. Heating is a more likely factor, although preliminary studies have generally concluded that the heating and thermal dose produced is rather low when compared to standard hyperthermia treatments. However, many of the effects seen do appear to mimic those of hyperthermia. Finally, tissue shear is created by the radiation forces and also when the ultrasonic wave reflects from a boundary at some angle and undergoes mode conversion. The resulting highly attenuating shear waves propagate at a much lower velocity, approximately 30–100 m/s, therefore having a more biologically relevant wavelength of 30–100 μm . It is generally accepted that shear forces can mediate biological effects, particularly in the vasculature [7], where normal blood flow shear is known to be a factor in the proper maintenance of the endothelial layer. Such processes might be disrupted or mimicked by radiation force induced streaming in the blood vessels. It is less clear how the relatively small shear strains of radiation forces, or the high frequency, rapidly attenuating shear waves, might affect the cells of bulk tissue.

Pulsed High-Intensity Focused Ultrasound, Fig. 2 Pulsed HIFU neighborhood in tissue “phase space,” in relation to cavitation, thermal ablation, and diagnostic ultrasound. The effect on tissue of ultrasound in this region is incompletely understood. Axes are the principle treatment parameters in common units (reflecting the choice of $f = 1, 1$ MHz focused transducer, and 1 Hz pulse repetition rate): acoustic power (50 W acoustic power implies $\sim 2,800$ W/cm^2 in the focal zone), duty cycle (5% implies pulse length of 50 ms), and total number of applied pulses



Pulsed High-Intensity Focused Ultrasound, Fig. 3 Pulsed HIFU acts to alter tissue transport either directly, by altering the physical properties of transport targets, or indirectly, by inducing biological changes that then influence transport properties

As well as the direct physical effects of pHIFU on tissues, there is always the possibility of triggering secondary biological processes, such as inflammation. Indeed, pHIFU often appears to trigger a clear inflammatory response, at least in muscle. It is not so clear whether this is also true in tumor or other tissue, nor whether the observed vascular and parenchymal permeability is the cause or the effect of inflammation.

If tissue changes are considered as comparable to material phase changes, then pHIFU effects lie in volatile and unexplored region of this phase space. This paradigm is illustrated in Fig. 2, which shows where treatment might be relative to a tissue phase map using the three fundamental treatment parameters as axes. Figure 3 lists the candidates for pulsed HIFU action and possible tissue targets.

Examples of Application

Pulsed HIFU has been explored in combination with a diverse set of agents, including small molecule chemotherapeutics, thrombolytics, large molecule therapeutics, antibodies, plasmid DNA, viral vectors, and nanoparticle drug carriers. The main goal of the technology since it was first conceived in the mid-1990s was to allow the uniform penetration of nanoparticle drug carriers throughout the tumor parenchyma. In that early study [8], the nanoparticle was a polymerized liposome containing gadolinium contrast agent of mean diameter around 100 nm. MRI was used to

monitor the delivery of the particle to rabbit muscle. The delivery of plasmid DNA (fluorescent reporter gene) was successfully demonstrated in a mouse squamous cell carcinoma more than 5 years later [9]. Later, the same technique was used to enhance delivery of TNF alpha plasmid and the liposomal nanoparticle Doxil and for tumor therapy [10, 11]. In 2006–2007, pHIFU was used to deliver radioactive antibodies [12] as well as genes via viral vector to mouse tumor models [13]. It was also combined with small molecular weight chemotherapeutics [14] and thrombolytic agents [15]. During various mechanistic studies, the delivery of nanoparticle-sized fluorescent dextrans and microspheres were used as tracking agents. Most of the mechanistic studies were done in mouse tumor models or mouse muscle. Comprehensive reviews of these and other related studies are available [16, 17].

Future Directions

Pulsed HIFU produced by ultrasound and magnetic resonance image guided systems will be used for assisted delivery of a variety of tissue targeted and untargeted nanoparticles. Virtually any large molecule or nanoscale therapeutic agent can take advantage of this technology for improved, targeted delivery. In fact, to be truly efficacious, it is anticipated that many nanoparticle systems will require the assistance of this or similar technology. Finally, it is possible to combine image-guided pulsed HIFU with various ultrasound and energy sensitive nanoconstructs for further control local and systemic therapeutic concentrations.

Cross-References

- ▶ [Liposomes](#)
- ▶ [Nanoencapsulation](#)
- ▶ [Nanoparticles](#)
- ▶ [Thermal Cancer Ablation Therapies Using Nanoparticles](#)

References

1. Jain, R.K.: Transport of molecules in the tumor interstitium: a review. *Cancer Res.* **47**(12), 3039–3051 (1987)

2. Song, C.W.: Effect of local hyperthermia on blood flow and microenvironment: a review. *Cancer Res* **44**, 4721s–4730s (1984)
3. Netti, P.A., Hamberg, L.M., Babich, J.W., Kierstead, D., Graham, W., Hunter, G.J., Wolf, G.L., Fischman, A., Boucher, Y., Jain, R.K.: Enhancement of fluid filtration across tumor vessels: implication for delivery of macromolecules. *Proc. Nat. Acad. Sci. USA* **96**(6), 3137–3142 (1999)
4. Hill, C.R., Bamber, J.C., ter Haar, G.R. (eds.): *Physical Principles of Medical Ultrasonics*, 2nd edn. Wiley, Chichester (2004)
5. Kimmel, E.: Cavitation bioeffects. *Crit. Rev. Biomed. Eng.* **34**(2), 105–161 (2006)
6. McDannold, N., Vykhodtseva, N., Hynynen, K.: Targeted disruption of the blood-brain barrier with focused ultrasound: association with cavitation activity. *Phys. Med. Biol.* **51**(4), 793–807 (2006)
7. Hillsley, M.V., Tarbell, J.M.: Oscillatory shear alters endothelial hydraulic conductivity and nitric oxide levels. *Biochem. Biophys. Res. Commun.* **293**(5), 1466–1471 (2002)
8. Bednarski, M.D., Lee, J.W., Callstrom, M.R., Li, K.C.: In vivo target-specific delivery of macromolecular agents with MR-guided focused ultrasound. *Radiology* **204**(1), 263–268 (1997)
9. Dittmar, K.M., Xie, J., Hunter, F., Trimble, C., Bur, M., Frenkel, V., Li, K.C.: Pulsed high-intensity focused ultrasound enhances systemic administration of naked DNA in squamous cell carcinoma model: initial experience. *Radiology* **235**(2), 541–546 (2005)
10. Yuh, E.L., Shulman, S.G., Mehta, S.A., Xie, J., Chen, L., Frenkel, V., Bednarski, M.D., Li, K.C.: Delivery of systemic chemotherapeutic agent to tumors by using focused ultrasound: study in a murine model. *Radiology* **234**(2), 431–437 (2005)
11. Frenkel, V., Etherington, A., Greene, M., Quijano, J., Xie, J., Hunter, F., Dromi, S., Li, K.C.: Delivery of liposomal doxorubicin (Doxil) in a breast cancer tumor model: investigation of potential enhancement by pulsed-high intensity focused ultrasound exposure. *Acad. Radiol.* **13**(4), 469–479 (2006)
12. Khaibullina, A., Jang, B.-S., Sun, H., Le, N., Yu, S., Frenkel, V., Carrasquillo, J.A., Pastan, I., Li, K.C.P., Paik, C.H.: Pulsed high-intensity focused ultrasound enhances uptake of radiolabeled monoclonal antibody to human epidermoid tumor in nude mice. *J. Nucl. Med.* **49**, 295–302 (2008)
13. Sun, H., Khaibullina, A., Li, K.: Bioluminescence studies: pulsed high-intensity focused ultrasound enhances the viral gene delivery in tumor models. *Mol. Imaging* **5**, 336 (2006)
14. Poff, J.A., Allen, C.T., Traughber, B., Colunga, A., Xie, J., Chen, Z., Wood, B.J., Waes, C.V., Li, K.C.P., Frenkel, V.: Pulsed high-intensity focused ultrasound enhances apoptosis and growth inhibition of squamous cell carcinoma xenografts with proteasome inhibitor bortezomib. *Radiology* **248**, 485–491 (2008)
15. Dromi, S., Frenkel, V., Luk, A., Traughber, B., Angstadt, M., Bur, M., Poff, J., Xie, J., Libutti, S.K., Li, K.C.P., Wood, B.J.: Pulsed-high intensity focused ultrasound and low temperature-sensitive liposomes for enhanced targeted drug delivery and antitumor effect. *Clin. Cancer Res.* **13**, 2722–2727 (2007)
16. Frenkel, V., Li, K.C.P.: Potential role of pulsed-high intensity focused ultrasound in gene therapy. *Future Oncol.* **2**, 111–119 (2006)
17. O'Neill, B., Li, K.C.P.: Augmentation of targeted delivery with pulsed high intensity focused ultrasound. *Int. J. Hyperthermia* **24**, 506–520 (2008)

Pulsed-Laser Deposition (PLD)

- ▶ [Physical Vapor Deposition](#)

Pyroelectricity

- ▶ [Polarization-Induced Effects in Heterostructures](#)

PZT

- ▶ [Piezoelectric MEMS Switches](#)