Nanostructure Science and Technology

Anatoli Korkin
Predrag S. Krstić
Jack C. Wells *Editors*

Nanotechnology for Electronics, Photonics, and Renewable Energy

Springer

Nanostructure Science and Technology

Series Editor

David J. Lockwood

For further volumes:
http://www.springer.com/series/6331

Anatoli Korkin · Predrag S. Krstić · Jack C. Wells
Editors

Nanotechnology for Electronics, Photonics, and Renewable Energy

Springer

Editors
Anatoli Korkin
Nano and Giga Solutions, Inc.
Gilbert, AZ 85296, USA
korkin@nanoandgiga.com

Predrag S. Krstić
Physics Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831, USA
krsticp@ornl.gov

Jack C. Wells
Oak Ridge National Laboratory
Center for Engineering Science
Advanced Computational
Nanotechnology Group
P.O. Box 2008
Oak Ridge, TN 37831, USA
wellsjc@ornl.gov

ISSN 1571-5744
ISBN 978-1-4419-7234-7 e-ISBN 978-1-4419-7454-9
DOI 10.1007/978-1-4419-7454-9
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010939003

Preface

Tutorial lectures given by world-renowned researchers have become one of the important traditions of the *Nano and Giga Challenges* (NGC) conference series. Soon after preparations had begun for the first forum, NGC2002,[1] in Moscow, Russia, the organizers realized that publication of the lectures notes would be a valuable legacy of the meeting and a significant educational resource and knowledge base for students, young researchers, and senior experts. Our first book was published by *Elsevier* and received the same title as the meeting itself—*Nano and Giga Challenges in Microelectronics.*[2] Our second book, *Nanotechnology for Electronic Materials and Devices,*[3] based on the tutorial lectures at NGC2004[4] in Krakow, Poland, the third book from NGC2007[5] in Phoenix, Arizona, and the current book from joint NGC2009 and CSTC2009[6] meeting in Hamilton, Ontario, have been published in Springer's *Nanostructure Science and Technology* series. Hosted by McMaster University, the meeting NGC/CSTC 2009 was held as a joint event of two conference series, Nano and Giga Challenges (*Nano & Giga Forum*) and Canadian Semiconductor Technology Conferences (CSTC), bringing together the networks and expertise of both professional forums.

Informational (electronics and photonics), renewable energy (solar systems, fuel cells, and batteries), and sensor (nano and bio) technologies have reached a new stage in their development in terms of engineering limits to cost-effective improvement of current technological approaches. The latest miniaturization of electronic devices is approaching atomic dimensions. Interconnect bottlenecks are limiting circuit speeds, new materials are being introduced into microelectronics manufacture at an unprecedented rate, and alternative technologies to mainstream CMOS are being considered. The low cost of natural energy sources and ignorance of the limits and environmental impact from use of natural carbon-based fuels have been long-standing economic barriers to the development of alternative and more efficient solar systems, fuel cells, and batteries. Nanotechnology is widely accepted as a source of potential solutions in securing future progress in information and energy technologies.

Nanotechnology as the art (i.e., science and technique) of control, manipulation, and fabrication of devices with structural and functional attributes smaller than 100 nm is perfectly suited to advanced CMOS technology. This technology

holds the capacity for massive production of high-quality nanodevices with an enormous variety of applications from computers to biosensors, from cell phones to space shuttles, and from large display screens to small electronic toys. Driven by scaling electronic devices to smaller and smaller sizes, the electronic industry has developed a set of sophisticated methods for deposition of ultra-thin films with very precise composition and nanoscale lithographic patterning. Enormous investment in nanotechnology for electronics R&D from companies and governments of many developed countries now finds a "new source of revenue"—solar energy applications. Based on the same core materials—silicon and other semiconductors—photovoltaics has for a long time been a "poor relative" of micro- and optoelectronics. That is not the case any more. The need for clean and renewable energy is becoming evident for the global community from the limited resources of mineral fuel and from the growing environmental impact of our current highly wasteful use of natural resources. Moreover, the world-wide energy crisis may spark a global political crisis and threaten the existence of human civilization.

The success of the NGC/CSTC 2009 conference, which resulted in the publication of this book and in other contributions making up special issues of *Nanotechnology*[7] and *Nanoscale Research Letters,*[8] would not have been possible without generous support from many sponsors and research institutions. We gratefully acknowledge the contributions and support of McMaster University (conference host and co-organizer), Arizona State University, Springer, Canadian Ministry of Foreign Affairs and Trade, Russian Nanotechnologies Corporation, Nokia, Oak Ridge National Lab, Ontario Centres of Excellence, Sematech International, CMC Microsystems, and many other local, national, international, and individual supporters.

Gilbert, Arizona — Anatoli Korkin
Oak Ridge, Tennessee — Predrag S. Krstić
Oak Ridge, Tennessee — Jack C. Wells

Notes

[1] http://asdn.net/ngcm2002/
[2] *Nano and Giga Challenges in Microelectronics*, eds. J. Greer, A. Korkin, J. Labanowski (Elsevier, Amsterdam, Netherlands, 2003)
[3] *Nanotechnology for Electronic Materials and Devices,* eds. A. Korkin, E. Gusev, J. Labanowski, S. Luryi (Springer, New York, USA, 2007)
[4] http://asdn.net/ngcm2004/
[5] http://asdn.net/ngc2007/
[6] http://asdn.net/ngc2009/
[7] Selected and invited papers from NGC/CSTC 2009 symposium on semiconductor technology (guest editors Stephen Goodnick, Anatoli Korkin, Predrag Krstic, Peter Mascher, John Preston, and Alexander Zaslavsky) published in *Nanotechnology* **21**(13) (2010)

[8]Nanoscale science and technology for electronics, photonics and renewable energy applications: Selected papers from NGC2009 & CSTC2009 conference. (guest editors Anatoli Korkin, Predrag Krstic, Zoran Miskovic, Hongbin Yu, and Igor Zhitomirsky) published in the open-access journal, *Nanoscale Research Letters* **5**(3) (2010).

Contents

Contributors

G.F. Cerofolini CNISM and Department of Materials Science, University of Milano–Bicocca, 20125, Milano, Italy

Gyorgy Csaba Institute for Nanoelectronics, Technische Universität München, D80333, Munich, Germany

Gyula Eres Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Christoph Erlen Institute for Nanoelectronics, Technische Universität München, D80333, Munich, Germany

D.B. Geohegan Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Dominic F. Gervasio Department of Chemical and Environmental Engineering, University of Arizona, Tucson, AZ 85721, USA

Adam R. Krause Nanoscale Science and Devices Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Predrag S. Krstić Physics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Simone Locci Institute for Nanoelectronics, Technische Universität München, D80333, Munich, Germany

D. J. Lockwood Institute for Microstructural Sciences, National Research Council Canada, Ottawa, Ontario K1A 0R6, Canada

Paolo Lugli Institute for Nanoelectronics, Technische Universität München, D80333, Munich, Germany

Paul Maguire Nanotechnology and Integrated BioEngineering Centre, University of Ulster, Newtownabbey, Co Antrim, BT37 OQB, UK

Charles Mahony Nanotechnology and Integrated BioEngineering Centre, University of Ulster, Newtownabbey, Co Antrim, BT37 OQB, UK

Gordana Malović Institute of Physics, University of Belgrade, 11080, Zemun, Serbia

A. Manthiram Materials Science and Engineering Program, The University of Texas at Austin, Austin, TX 78712, USA

Dragana Marić Institute of Physics, University of Belgrade, 11080, Zemun, Serbia

T. Muraliganth Materials Science and Engineering Program, The University of Texas at Austin, Austin, TX 78712, USA

Zoran Lj. Petrović Institute of Physics, University of Belgrade, 11080, Zemun, Serbia

Nevena Puač Institute of Physics, University of Belgrade, 11080, Zemun, Serbia

A.A. Puretzky Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Maja Radetić Faculty of Technology and Metallurgy, University of Belgrade, 11000, Belgrade, Serbia

Marija Radmilović-Radjenović Institute of Physics, University of Belgrade, 11080, Zemun, Serbia

C.M. Rouleau Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Larry R. Senesac Nanoscale Science and Devices Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Thomas Thundat Nanoscale Science and Devices Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

L. Tsybeskov Department of Electrical and Computer Engineering, New Jersey Institute of Technology, University Heights, Newark, NJ 07102-9895, USA

Charles W. Van Neste Nanoscale Science and Devices Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Chapter 1
Molecular Electronics: Challenges and Perspectives

Paolo Lugli, Simone Locci, Christoph Erlen, and Gyorgy Csaba

Abstract Molecular electronics has lately attracted increasing attention due to some appealing features such as possibly very higher integration capabilities, their low production cost, flexibility in the substrate choice, and possibility for large-area deployment. Two parallel approaches characterize this field: on one side molecules can be contacted and their transport characteristics exploited to achieve electronic functionalities; on the other side existing device structures, as well as novel ones, can be realized using organic layers instead of or together with inorganic materials. While in the latter case theoretical investigations on such devices can be carried out on adapting conventional simulators to the new materials and physics involved, completely new tools have to be developed in the former case. In this chapter, the operational principles of molecular systems will be presented based on a series of theoretical results obtained from our groups. Challenges and perspectives are also discussed.

Introduction

Molecular electronics has witnessed increased interest in recent years, triggered by the forecast that silicon technology might reach its scalability limits in a few years [1–4]. In order for molecular electronics to become a valuable alternative to silicon technology, it will not be sufficient to fabricate molecular electronic devices with outstanding characteristics, but appropriate circuit and architectural solutions will also be needed. While a lot of effort has been dedicated to the demonstration of electronic functionalities of single molecules and organic films, research at the circuit and system level is still in its infancy [5–9].

Investigation on single molecules or nanotube-based devices promises to keep Moore's law alive once miniaturization of silicon-based structures becomes

P. Lugli (✉)
Institute for Nanoelectronics, Technische Universität München, Arcistrasse 21, D80333 Munich, Germany
e-mail: lugli@tum.de

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_1,

impractical. As first proposed by Aviram and Ratner [10], one can imagine to squeeze entire nonlinear circuit elements (such as diodes or transistors) into single molecules. In principle, such devices could be significantly faster and smaller than end-of-the-roadmap solid-state electron devices. Despite an enormous progress on the experimental characterization of single-molecule conduction in the last years, only few device concepts have emerged and it is still unclear whether individual molecular devices could be integrated into a larger-scale computing circuit. In any case, we can be sure that molecular circuits and architectures will be very different from what we are used to in today's systems. Several architectures, which would be suitable for the realization of electronic logic circuits or memory cells based on molecular devices, have been proposed. One possibility is to synthesize complex molecules whose arms can be separately contacted to provide the same electrical input/output of conventional logic gates. Such fascinating idea, suggested by Ellenbogen and Love [11], is unfortunately extremely challenging from the chemical-synthesis point of view, and it has not been realized up to now. Another possibility is to create a programmable interconnected network of nanoparticles and molecular entities ("nanocells") [12]: conducting metallic nanoparticles are randomly deposited on a substrate (which can be a silicon one) and subsequently bridged via molecular connections, to create electrical pathways between previously patterned metallic leads. The molecular linkers should exhibit nonlinear IV characteristics, in the form of negative differential resistance or hysteretic behavior and the network is connected to a limited number of input/output pins at the edges of the nanocell, which could then be accessed, configured, and programmed from the edges. Simulations have demonstrated the capability of the nanocell to act as logic gate [12, 13].

Integration with CMOS technology will certainly be the first step for devices based on single molecules, with solutions that follow a road that has been called "More than Moore" [14], in an attempt to extend the standard chip functionalities already offered by silicon technology. Such hybrid systems would benefit from the speed and reliability of CMOS, while offering at the same time the versatility and intrinsically nanometer footprint of molecular devices. Two architectures which are in principle compatible with silicon technology are the "Quantum Cellular Automata" [15] and the "cross bar" [16, 17]. While the latter employs solutions based on standard interconnections, the former relies on a highly innovative approach. In fact, devices and their interconnections can no longer be separated in a straightforward way in single-molecule electronics and it is not clear if the paradigm of classical circuit theory for building a complex system can still be followed [18, 19]. It is very timely to explore alternative ideas to interconnect molecular building blocks inside the molecular circuit and the molecules to the outer world without metal wires [20], using electric or magnetic fields. One very interesting hybrid architecture is the semiconductor/nanowire/molecular one ("CMOL"), proposed by Likharev and coworkers [21, 22]. The basic difference with respect to crossbars is that in CMOL the interface between CMOS circuitry and nanowires is provided by pins distributed all over the circuit area. Thus, a much larger integration should be achieved, with relaxed requirement concerning alignment of the wire interconnects.

Besides single-molecule structures, molecular electronics refers also to the employment of organic materials as substitutes for the solid-state layers of traditional semiconductor devices. Since their carrier mobility is relatively low, organic devices cannot compete with silicon-based circuits in high-speed, high-performance computing applications, but can become competitive in the market of large-area devices, because their process technology is potentially inexpensive and circuits can be realized on virtually any substrate (e.g., on plastics, paper, or textiles) [23].

We will present our theoretical analysis on organic transistors, inverters, and ring oscillators, describing the simulation tools used and its applications to device and circuit modeling. In the areas of chemical sensing, light emission, and light sensing, organic materials could make far more versatile devices as their semiconductor counterparts, because of the wide selection of possible sensor and light-emitting materials [24]. The first products have already entered the marketplace (e.g., OLED-displays), whereas others will follow soon. Due to their versatility in terms of fabrication and performance, organic films could also be integrated with silicon circuitry, embracing again the "More than Moore" paradigm.

Modeling and simulation of these device classes, which imitate traditional solid-state devices with organic materials, seems quite straightforward, as the existing semiconductor simulation approaches, like the well-established drift-diffusion method [25], can be applied. In spite of many open issues, mainly related to the charge transport in organic semiconductors as well as to their morphology [26], simulating organic devices do not require the development of fundamentally new approaches. In parallel to numerical simulation, analytical approaches have also been developed, which can help in the analysis of transport phenomena (e.g., trapping) and device operation [27]. Circuits employing organic thin-film transistors (OTFTs) are based on the same building block of conventional electronics, namely the inverter, but, due to the difficulty of realizing n-type organic transistors, typical circuits rely only on p-type devices, without being of complementary types [28].

A quite different approach is required for single-molecule devices, where quantum mechanical tools are needed. Quantum chemistry methods have reached a high level of maturity and are credited with good predictive power for the determination of formation energies and molecular geometries. It seems natural to extend these methods beyond their original purpose to treat non-equilibrium transport processes through single (or few) molecules and to analyze their device performance. To such purpose, density functional theory (DFT) [29] usually provides the description of the electronic properties of the molecular system, while the non-equilibrium Green's function (NEGF) [30] allows one to compute the transport characteristics. With such tools, it is possible to calculate the current–voltage characteristics of a molecule connected to two metallic electrodes [31, 32]. Due to the computational complexity of the DFT–NEGF algorithms, it is basically impossible to extend them to the design of even very elementary circuits and memory cells. It is therefore necessary to develop approaches which allow us to go from the atomistic description of the single molecular element to design tools capable of full circuit simulations. Construction of equivalent circuit models and circuit-level simulations serve this purpose [33], as will be discussed afterwards.

Single Molecules

Electronic conduction through a variety of different molecules has been studied experimentally by many research groups [31, 32, 34–36] and novel design architectures for memories and logic circuits have been explored. However, the exact nature of the transport mechanisms in many of such systems remains still open to scientific debate. To correctly treat these problems, complex ab initio computations are needed. First-principles computations [29, 37, 38] are normally based on density-functional methods, which are, however, limited to molecular systems made up of a small number of atoms. In addition, a precise estimation of energy band gaps requires sophisticated many-body corrections. Since we are interested in the transport properties of molecules, which are not isolated entities but rather connected to metallic electrodes, we need first to clarify how the interactions between molecule and contacts determine the energy levels of the complete system. We have used the density-functional tight-binding (DFTB) method [39, 40–45], which allows a first-principles treatment of systems comprising a large number of atoms. This method, implemented in the code gDFTB, has been extended to the non-equilibrium Green's function (NEGF) approach [30], by generalizing the calculation of the atomic charges using the non-equilibrium (DFT) density matrix and solving a three-dimensional Poisson solver [40]. The implementation of NEGF is similar to other codes based on first-principles approaches [46, 29]. Thus, the computation of the tunneling current between two contacts is possible in a manner consistent with the open boundaries and non-equilibrium conditions that naturally arise in coherent transport problems.

The main approximations of gDFTB are (i) the use of limited basis functions (mainly the minimal basis) and (ii) the use of two-center approximation in the calculation of the Hamiltonian matrix elements. The use of such approximations makes the scheme computationally extremely fast and makes the approach suitable to treat systems with several thousands of atoms, thus applicable to nanostructured devices. The gDFTB scheme can be considered as a highly efficient high-level computational tool which works surprisingly well despite the large number of simplifying assumptions. It enables, for instance, the computation of the tunneling current flowing between two or more contacts in a manner consistent with the open boundaries and non-equilibrium conditions that naturally arise in coherent transport problems. Molecular dynamics simulation can be performed within the present approach, and time-dependent currents can be calculated accounting for molecular vibration effects [40, 47]. Recently, the approach has also been extended to account for electron–phonon interactions [48].

The density-functional tight-binding (DFTB) formalism has been described in detail in many articles and reviews [49]. All matrix elements and orbital wavefunctions are derived from density-functional calculations. The advantage of the method relies on the use of a small basis set and the restriction to two center integrals, allowing extensive use of look-up tables. What distinguishes our approach from empirical methods is the explicit calculation of the basis wavefunctions, which allows deeper physical insights into and better control of the approximations used.

The method solves the Kohn–Sham equations self-consistently using a Mulliken charge projection [43]. In the traditional DFTB code a minimal basis set of atomic orbitals is used in order to reduce the matrix dimensions for diagonalization speed-up. This approach has proved to give transferable and accurate interaction potentials, and the numerical efficiency of the method allows molecular dynamic simulations of large super-cells, containing several hundreds of atoms, particularly suitable to study the electronic properties and dynamics of large mesoscopic systems and organic molecules such as CNTs, DNA strands or adsorbates on surfaces, and semiconducting heterostructure [50].

We briefly describe here the self-consistent DFTB method. The electronic density is expanded as a sum of a reference density $n^0(r)$ (that can be chosen as the superposition of neutral atomic densities) and a deviation, $\delta n(r)$, such that $n(r) = n^0(r) + \delta n(r)$. The total energy of the system can be described, up to second order in the local density fluctuations, as

$$E_{\text{tot}}[n] = \sum_k n_k \left\langle \psi_k \left| H^0 \right| \psi_k \right\rangle + E_{\text{rep}}\left[n^0\right] + E^{(2)}[\delta n]. \tag{1.1}$$

The first term in Eq. (1.1) can be written in terms of the TB Hamiltonian, which is given by

$$\begin{cases} H^0_{\mu\nu} = \varepsilon_\mu^{\text{free-atom}}, & \mu = \nu \\ \left\langle \phi_\mu \left| T + v_{\text{eff}}\left[n_i^0 + n_j^0\right] \right| \phi_\nu \right\rangle, & \mu \in i, \nu \in j \end{cases}, \tag{1.2}$$

where ϕ_μ and ϕ_ν are the atomic orbitals localized around the atomic centers i and j; T is the kinetic energy operator; and v_{eff} is the effective one-particle potential, which depends on the density of the two atomic centers i and j.

The term $E_{\text{rep}}\left[n^0\right]$ in Eq. (1.1) is the repulsive energy between the ions, screened by the electronic distribution and the exchange energy. The third term in Eq. (1.1) is the second-order correction that can be written as

$$E^2[\delta n] = \frac{1}{2}\iint \left[\frac{1}{|r - r'|} + \frac{\delta^2 E_{\text{xc}}}{\delta n(r)\delta n(r')}\right] \delta n(r)\delta n(r')\,\mathrm{d}r\,\mathrm{d}r', \tag{1.3}$$

where the Hartree and exchange–correlation potentials have been separated. This quantity is greatly simplified by retaining only the monopole term in the radial expansion of the atom-centered density fluctuations [51].

Within the LDA approximation the exchange contribution vanishes for large atomic distances. Hence in Eq. (1.3) the second-order correction to E_{xc} can be neglected with respect to the Coulomb interaction. For short ranges Coulomb and XC contributions are accounted with on-site Hubbard parameters, which are calculated for any atom type within LDA-DFT as the second derivative of the total energy of the atom with respect to the occupation number of the highest occupied atomic orbital. These values are therefore neither adjustable nor empirical parameter [43]. By applying the variational principle to the energy functional of Eq. (1.1),

it is possible to obtain a modified Hamiltonian for the Kohn–Sham equations:

$$H_{\mu\nu} = H^0_{\mu\nu} + \frac{1}{2} S_{\mu\nu} \sum_k \left(\gamma_{ik} + \gamma_{jk}\right) \Delta q_k, \forall \mu \in i, \nu \in j, \tag{1.4}$$

where $S_{\mu\nu}$ is the overlap matrix between atomic wavefunctions, γ_{ik} is a shorthand for the interatomic potential [51], computed from Eq. (1.3), and Δq_k are the atomic charges. Since the atomic charges depend on the one-particle wave functions ψ_k, a self-consistent procedure is required. The improvement of the self-consistent over the non-self-consistent procedure is considerable in determining structural and energetic properties of molecular systems [43].

The DFTB approach has been extended to the NEGF formalism by generalizing the calculation of the atomic charges using the non-equilibrium (DFT) density matrix and computing the functional (1.3) using a three-dimensional Poisson solver [47, 49, 52]. Despite its mathematical complexity, the non-equilibrium Green's function (NEGF) for calculations of quantum transport has gained great popularity in recent years, mostly because of the versatility and numerical stability of the method, in contrast to wavefunction or transfer matrix approaches. The open boundary conditions can be elegantly included by exactly mapping the contacting leads into a finite and small part of the system [53]; furthermore, the Green's function approach can be generalized to many-body quantum theory, allowing the inclusion of electron–phonon [54] as well as electron–electron interactions [55] within a unified and systematic formalism. Good references in many-body quantum theory can be found in [56, 57] and an exhaustive review on NEGF can be found in [58].

Transport in OPVs

We have applied the gDFTB code to the study of molecular transport through oligopolyvinylene (OPV) chains anchored to metallic contacts [59]. OPV molecules consist of benzene rings which are interconnected by vinylene groups (Fig. 1.1 for a 5-ring OPV structure). A DFT-based investigation of the transport properties of short OPV wires (up to five rings) has been presented in [60]. We concentrate here on longer structures. The electronic transport characteristics of the OPV molecules

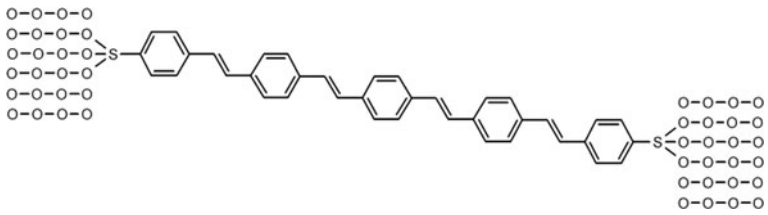


Fig. 1.1 OPV molecule with five rings bonded to gold contacts by a sulfur atom. From: [59] Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission

are governed by the electronic states available in the molecule. Such states can be effectively described by the local density of states (LDOS), which represents the equilibrium density of states existing in the molecular device broken down into the contributions of single atoms or ensembles of atoms of the device. The gDFTB code calculates the LDOS of each atom by projecting the density of states of the total structure on the molecular orbitals of this atom.

Figure 1.2 shows the LDOS for OPVs of different lengths. The broadening of the molecular levels (with respect to the sharp peaks that would be found for isolated molecules) is a result of the covalent bonding between the molecular device and the contacts or the adjacent sulfur atoms and is related to a finite lifetime of electrons on the corresponding orbitals. Such finite lifetime is crucial for current flow to appear, as will be emphasized in the following. The LDOS peaks of the molecules are getting sharper with increasing wire length. The associated longer lifetime of the electrons occurs because an electron that is delocalized over a larger molecule has a smaller probability to tunnel out into the contacts. By comparing wires of different lengths, we can observe that the longer wires display not only sharper peaks in their LDOS but also a larger number of peaks. In addition, the HOMO–LUMO gap continuously gets smaller when going from a 1-ring to a 4-ring OPV (not shown in the figure) but converges to a given value for the longer wires of Fig. 1.2. Such effect is related to the negligible effects that contacts play the longer the wire becomes.

The coherent current flowing through a molecule can be calculated from Landauer formula [30] by integrating the transmission function $T(E)$ in the energy

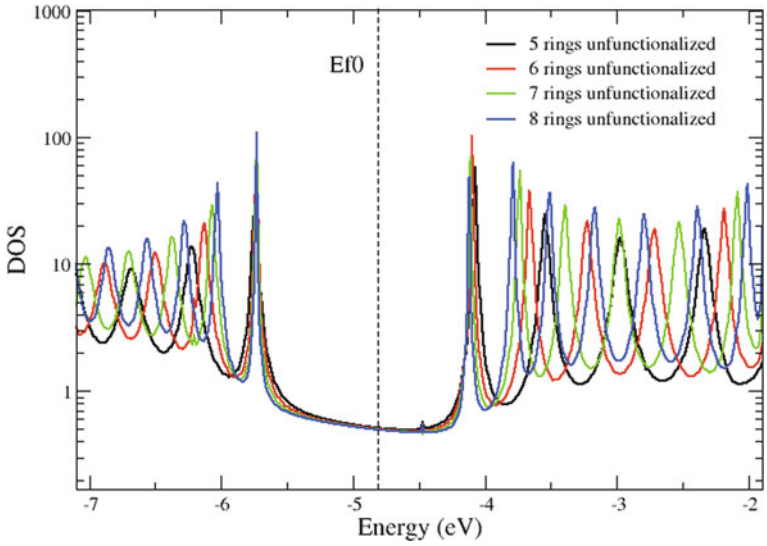


Fig. 1.2 Local DOS of four different unfunctionalized OPV molecules. The Fermi level (Ef0) of the gold contacts at equilibrium is also shown. From: [59] Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission

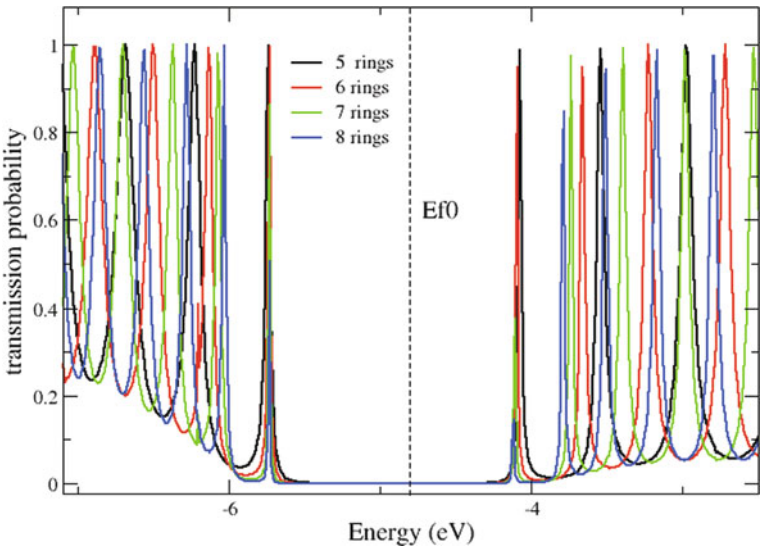


Fig. 1.3 Transmission probability of the OPV5 unfunctionalized as a function of energy for different applied voltages *V*. Ef0 denotes the equilibrium Fermi level of the electrodes. For better clarity, the non-equilibrium transmission functions have been vertically shifted for rising bias voltages. One vertical step corresponds to a difference in applied bias of 0.20 V. The *vertical line* represents the location of the constant Fermi level Ef0 of one contact, while the Fermi energy of the other contact $E\text{f}0 - eV$ is visualized by the *tilted line*. From: [59] Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission

range created by the voltage applied between the two electrodes. A typical example of the bias dependence of the transmission function is shown in Fig. 1.3. There we see that the peaks in the transmission functions exhibit a change in height, form, and relative position as a function of bias. In particular the amplitude of the transmission peaks decreases for increasing voltages. The reason is that an applied bias voltage at the contacts leads to an additional potential drop over the molecule, which in turn enhances the localization of the electronic states on the molecule. Since state delocalization is the prerequisite for these states to act as a channel for electronic transport, the applied voltage results in diminished transmission probabilities through these states. For longer molecules, the lowering of the transmission peaks for applied biases is even stronger. In such regime, the electrons go from one contact to the other only by "hopping" from one localized state to another, with a reduced transmission probability. The linear potential drop is also responsible for the shifting of the transmission peaks toward lower energy values.

Figure 1.4 compares the *I–V* characteristics of the four OPV5–OPV8 unfunctionalized wires. All structures display higher currents in the resonance regime as opposed to the tunneling one. The threshold voltage does not vary significantly as the transmission gap for such wires is relatively constant. The current decreases as a function of wire length. The threshold behavior in the current characteristics is a

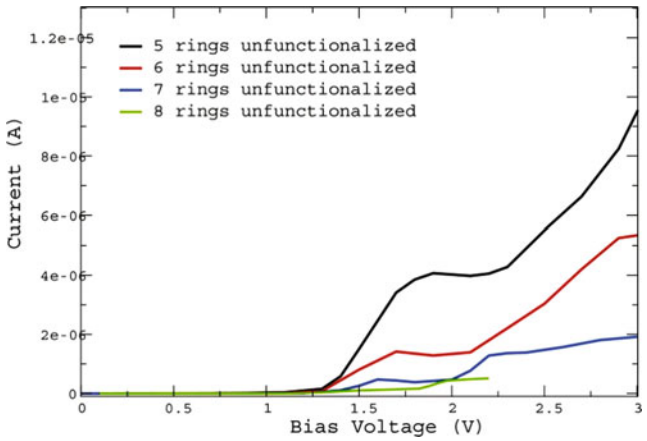


Fig. 1.4 Current–voltage characteristics for OPV5–OPV8. From: [59] Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission

consequence of the transmission gap. Examination of Fig. 1.3 clarifies that, as long as the bias window only comprises energies with negligible transmission probabilities, the current remains low and stems from tunneling. At sufficiently high bias, the current starts increasing due to the resonances appearing in the window (around 1.4 V for the OPV5). Plateaus appear in the *I–V* curve at higher bias (as for instance around 2 V for the OPV5), due to decrease of the transmission peak connected to the conducting states as the voltage increases. As soon as a new resonance is available for conductance, the current increases again. Despite the capability of theoretical approaches like gDFTB to model transport in single molecules, such tools can only be applied to molecular systems with a limited number of atoms. Furthermore, the metal contacts need to be greatly simplified. Thus, in order to model realistic molecular devices, where dissipation and parasitic effects play a fundamental role, different approaches have to be followed. One possibility that will be discussed below is to develop an equivalent circuit description of the molecular element and to use SPICE-like approaches to model complex molecular systems. Such equivalent circuit can originate either from an atomistic calculation as provided by gDFTB or directly from experimental results.

Power Dissipation in Molecular Junctions

So far, transport in single molecules has been treated completely ballistic, that is neglecting scattering and dissipation phenomena. In reality, molecules possess their own motion (in the form, e.g., of vibrational or rotational degrees of freedom) which causes interaction processes with the charge carriers moving across. In this part, we illustrate how to take such interactions into account and what effect scattering has in molecular transport [61]. The usual procedure to treat molecular vibrations is to

expand the effective nuclear potential up to the harmonic term and to decouple the Hamiltonian as a superposition of independent one-dimensional oscillators corresponding to the normal modes of vibrations. Each vibrational mode will be labeled by q. The harmonic oscillators can be quantized following the usual prescriptions, by making use of the standard relationships between the position operator and the Bose field operator. The calculation of the electron–phonon scattering requires an explicit evaluation of the electron–vibration couplings matrices, obtained by expanding the TB Hamiltonian to first order in the atomic displacements [48, 62].

The electron–phonon interaction is treated within perturbation theory of the non-equilibrium Green's function formalism and the current through the junction is computed by the Meir–Wingreen formula using the appropriate self-energies for electrons and holes [63].

The thermally averaged phonon population, N_q, can be obtained to first approximation by the Bose–Einstein distribution function, using the environment temperature. A more sophisticated calculation requires a self-consistent calculation of such a distribution as a function of applied bias, as discussed in the next section. The electron–phonon interactions can be viewed as an exchange of particle with a *virtual* contact [30] which adsorbs electrons at a given energy and emits them at another energy, accounting for the energy loss in the phonon quanta. In perfect analogy to the real contacts, the virtual contact brakes the phase coherence of the wavefunction. The amount of power dissipated in the molecule, due to inelastic phonon emission, can be obtained by considering the virtual contact current, as discussed for instance in [48, 30]. The power dissipated is given by the net rate of energy transferred to the molecule and can be calculated using the virtual contact current. The power dissipated can then be used to compute the rate of phonon emission.

In order to compute the non-equilibrium phonon population under bias, a phenomenological rate equation is used. The rate of phonon emission can be defined as $R_q\left(N_q\right) = W_q/\hbar\omega_q$, representing the ratio between the energy emitted per unit time in the oscillator q divided by its energy. This is generally a function of bias and depends on the phonon population itself. The rate R_q is actually the net rate of phonon emission, also including the absorption rate due to assisted tunneling and electron–hole pair production, very important in metal contacts. In order to compute the phonon population of the vibrational modes a rate equation can be written, including the rate of emission and dissipation into the leads, as

$$\frac{\mathrm{d}N_q}{\mathrm{d}t} R_q\left(N_q\right) - J_q\left(N_q - N_q^0(T)\right), \tag{1.5}$$

balancing the rate of quanta emitted with the rate of phonon decay to the bath, J_q, discussed below. This tends to restore the equilibrium Bose–Einstein distribution, characterized by the contact temperature, T. Under stationary condition the phonon distribution on the molecule is given by $N_q = N_q^0(T) + R_q\left(N_q\right)/J_q$. The phonon dissipation rate is by no means a simple quantity to be predicted from first-principles calculations, since a large number of different mechanisms participate in phonon relaxations. For this reason J_q can be regarded in many cases as a fitting parameter of the model. Nonetheless we compute an upper bound of this

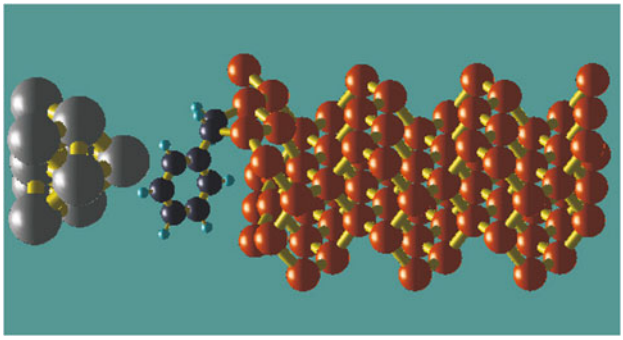


Fig. 1.5 Representation of the system simulated comprising a styrene molecule adsorbed on a Si(100) 2×1 substrate. From: [61] Copyright 2007 by Taylor & Francis Group, LLC

parameter by considering the elastic coupling of the molecular modes with the contact vibrations.

We consider as example the hybrid system shown in Fig. 1.5, comprising a Si(100) substrate, reconstructed 2×1, with an adsorbed styrene molecule in a bridge position and an Ag metal contact which models an STM tip. The position of the tip was chosen in order to obtain a relatively large direct coupling between the tip and the π orbitals of the styrene which guarantees a sufficiently high conductance. The figure shows a unit cell of the system considered, but periodic boundary conditions have been imposed. The Si substrate is heavily p-doped in order to make it conducting and the Fermi level is assumed at the valence band edge. While electrons cross the molecule, they interact with the molecular ionic vibrations from which they can be inelastically scattered. The electron–phonon scattering within the leads is not considered. In order to study the electron–vibron coupling we first relax very accurately the structure (this is done under no applied bias). Then we compute the vibrations of the molecule constraining the Si and Ag atoms.

In order to calculate the rate of phonon dissipation into the contacts the molecular modes are coupled with the bulk vibrations. Such a calculation is performed using a Green's function approach similar in concept to the electronic calculation. This formulation is possible because the atomic coupling, expressed by the dynamical matrix, is restricted to few neighbors. The imaginary part of the self-energy is used to extract the vibron lifetime. This approach includes first-order one phonon to one phonon decay processes, but obviously neglects many other mechanisms that may take place when the direct decay is not allowed by energy conservation. This happens, for instance, to the high-frequency modes, characteristic of molecular vibrations, generally lying well beyond the vibrational bandwidth of the bulk reservoirs which cannot decay other than via one-to-many phonon channels. Other decay mechanisms may involve coupling with the surrounding molecules and generally depend on the environment or multistep processes and anharmonic couplings. The result of this calculation is shown in Fig. 1.6, reporting the superposition of the phonon density of states of the coupled system (gray area) with the uncoupled

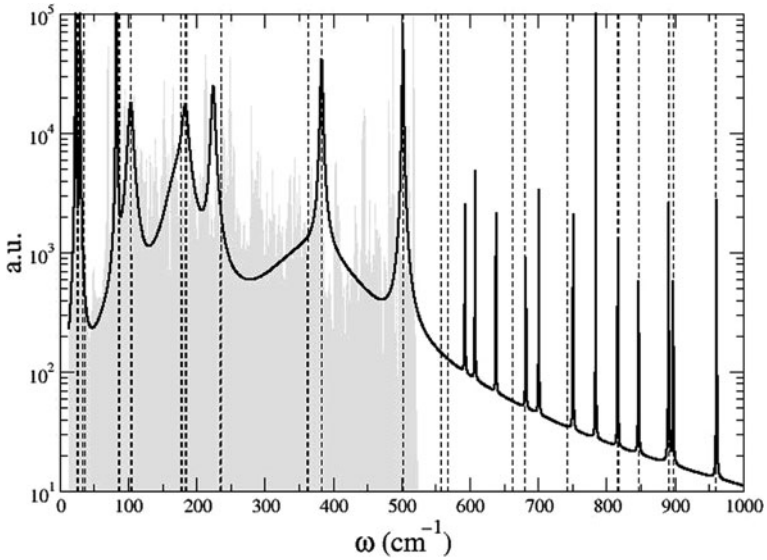


Fig. 1.6 Vibrational density of states of the system Si–styrene–Ag. Gray area represents the Si DOS, *vertical lines* the uncoupled molecular modes, and *solid line* the DOS projected on the molecular modes showing their broadening due to coupling with the substrate. From: [61] Copyright 2007 by Taylor & Francis Group, LLC

molecular frequencies shown as vertical dashed lines. The cut-off frequency of the Si vibrations at 500 cm^{-1} clearly appears in this plot. The solid line represents the projected density of states obtained by approximating the self-energies just as a broadening and frequency shift.

As discussed above, the calculation gives realistic results for those molecular modes lying within the Si phonon bands, where the broadening is sizable. The corresponding phonon decay rates are in the range $10^{11} - 10^{12}$ Hz. For those energy modes lying beyond the contact bandwidth the broadening sharply decreases, leading to decay rates as slow as 10^2 Hz for the highest modes, which is rather unrealistic. Therefore, the lowest limit for the decay rate has been fixed to $J_{q,\min} = 10^6$ Hz, in all subsequent calculations.

We find that the out-of-equilibrium phonon population is strongly bias dependent. Figure 1.7 reports the behavior of the average molecular temperature as a function of applied bias for two different contact temperatures (10 K and 300 K). The molecular temperature, T_{mol}, is computed by setting the energy balance expressed with

$$\sum_q \hbar\omega_q N_q = \sum_q \hbar\omega_q n_q \left(T_{\mathrm{mol}}\right) \tag{1.6}$$

where $n_q\left(T_{\mathrm{mol}}\right)$ is the Bose–Einstein distribution for the temperature T_{mol}. As the bias increases, the temperature rises since more heat is dissipated in the molecular vibrations. The figure reports three different calculations obtained for different

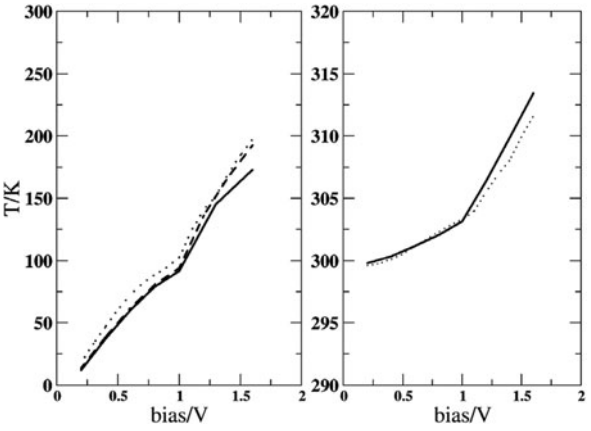


Fig. 1.7 Molecular temperature as a function of applied bias for different types of approximations in the electron–phonon couplings (see text). From: [61] Copyright 2007 by Taylor & Francis Group, LLC

approximations in the treatment of the electron–phonon coupling. The dotted lines correspond to the simplest lowest order calculation of the electron–phonon self-energy, not including the renormalization of the electron propagator; the dashed line corresponds to the first-order correction, commonly known as first-order Born approximation (BA), including renormalizations of the propagator to first order. The solid line corresponds to the self-consistent Born approximation (SCBA) where self-energy and propagator are computed self-consistently. We can conclude that, because of the small incoherent tunneling current, the inclusion of higher orders in the electron-phonon coupling brings only small corrections to the final result.

In Fig. 1.7 it is possible to appreciate the change in slope at the applied bias of 1.0 V. The total tunneling current at this bias is 1.2 nA. The origin of this effect comes from a molecular resonant state entering the injection window at the applied bias of 0.95 V. This has the effect of strongly increasing the coherent and incoherent currents. On average all vibrational modes are excited as the bias increases, although some of them are particularly favored, such as the lowest vibrational mode, at the frequency of 10.55 cm^{-1}, corresponding to a rigid oscillation of the whole benzene ring, away and toward the Ag tip. This mode leads to a strong variation of the molecule–metal matrix elements and, therefore, to a large electron–phonon coupling. Once the first resonance is reached, most of the phonon emission occurs into the lowest energy mode, leading to a sharp increase of molecular temperature.

Perspectives

The conduction observed in single molecules connected to electrodes opens the way for a variety of electronic applications, rectification being one of the first properties that researchers tried to study for molecular devices, with the first investigations dating back to 1974 [10]. One example is given in [64], where a molecular rod that consists of two weakly coupled electronic π-systems with mutually shifted energy levels is realized. There, the resulting asymmetry manifests itself in a

current–voltage characteristic which depends on the sign of the bias voltage like in a conventional diode. Other studies on the rectification properties of molecules can be found in [65] and [66], where a the oretical view is also offered.

Adding a third electrode to obtain transistor action has been proven difficult so far, and the roadmap toward the development of single molecule transistor is just at the beginning [67]. In 1998, Sander et al. [68] reported a field effect transistor working at room temperature, using a carbon nanotube as semiconductor. In 2009 Song et al. [69] realized a solid-state molecular device, in which transport current is directly modulated by an external gate voltage, which changes the energy of the molecular orbital. A further application of single-molecule devices is as active element of a memory, because they can be intrinsically good hysteretic switches and exhibit highly nonlinear current–voltage characteristics [70]. For the above reasons it seems promising to exploit special circuit architectures, like crossbar memories, which have a very simple and regular geometry and which are built exclusively from two-terminal switches and nonlinear elements [71–73]. Next we will describe the properties of these particular molecular devices.

Crossbar Memories

Crossbar passive memories [71, 33] are probably the simplest functional electrical circuits. They are built from horizontal and vertical conducting wires where a switching element is placed at each wire crossing. If the switches are made from molecular monolayers, then the size of the storage elements can be scaled down to molecular dimensions, e.g., in the few nanometer regime. Some nanofabrication methods (such as nanoimprint lithography) may enable the fabrication of the contact wires on a similar size scale [74]. For this reason, molecular crossbar memories received a lot of attention recently and there is a number of experimental [71] and theoretical [75–78] publications devoted to their study. Logic devices can also be built using similar principles [72]. In this work we assume that the read-out and write in circuitry of the memory (i.e., multiplexers/decoders) are realized with (standard) silicon technologies and only the storage array is molecular-based.

A molecular crossbar memory circuit should be integrable with conventional silicon microtechnologies and would require the introduction of additional fabrication steps [79–81]. Only a relatively large-scale molecular memory would make such technology transition to pay off: if the size of the molecular memory banks is too small, the silicon-based interface circuitry occupies most of the chip area. Crossbar memories are passive circuits and they contain no amplifying/signal-restoring elements. This can severely limit their scalability, and parasitic current paths can interfere with the read-out process.

In this section, we investigate molecular crossbar memories from a circuit design point of view. We will investigate what properties the molecular $I(V)$ curve should have in order to be a useful element of a scalable crossbar memory and we will show that "real" molecules exactly show such thresholding-type nonlinearities. The sketch of a crossbar circuit is shown in Fig. 1.8a. It is built from horizontal and

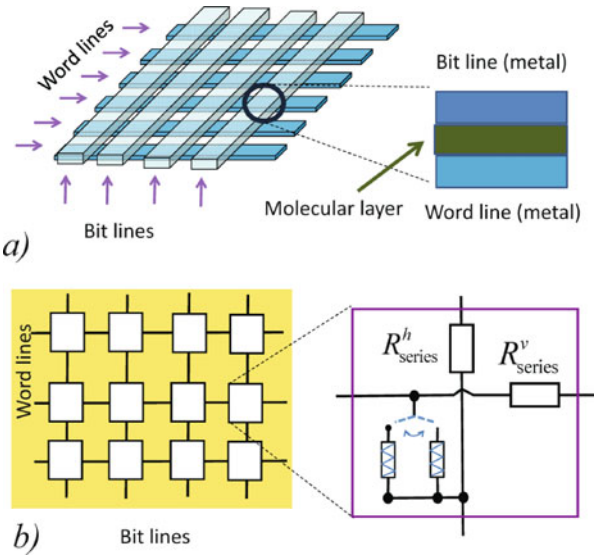


Fig. 1.8 Physical layout of the molecular crossbar (**a**) and a cellular model for the simulation of the circuit (**b**). From: [33] Reproduced with permission. Copyright 2009 IEEE

vertical conducting wires and at each wire crossing a switching element is placed. The circuit representation of the array is sketched in Fig. 1.8.

The junction can have two distinct states, the *on* and the *off* states, corresponding, respectively, to a low $R = R_{\text{on}}$ and a high $R = R_{\text{off}}$ resistance, which depends on the history of the junction. This is modeled by setting the switch of Fig. 1.8b to the on or off position. The series resistance of the interconnecting wires is also included in Fig. 1.8; however, the small (couple of tens of ohms) resistance of metallic interconnections is negligible compared to the typically several megaohms resistance of molecular junctions. We allow the junction resistances to be nonlinear (i.e., function of the applied voltage).

To get a qualitative picture of the crossbar operation, we consider junctions with linear $I(V)$ characteristics and zero serial resistances. The most straightforward read-out scheme [71] biases the interrogated junction and measures the current flowing from the junction toward the ground, as illustrated in Fig. 1.9a. All the unselected bit and word lines are put to ground. The schematics of a read-out circuitry (i.e., a resistor and a comparator) are also shown.

By redrawing the circuit in the way it is illustrated in Fig. 1.9b, it is apparent that the current or sense voltage is influenced only by the resistances connected to the same bit line as the interrogated bit. They are connected in parallel to R_{sens}, effectively changing its value.

It is clear from this circuit model that the parasitic resistances connected to the same bit line pull down the sensing voltage. Very small values of R_{sens} are required to maintain a reasonably large noise margin. Sense resistances in the kiloohm regime

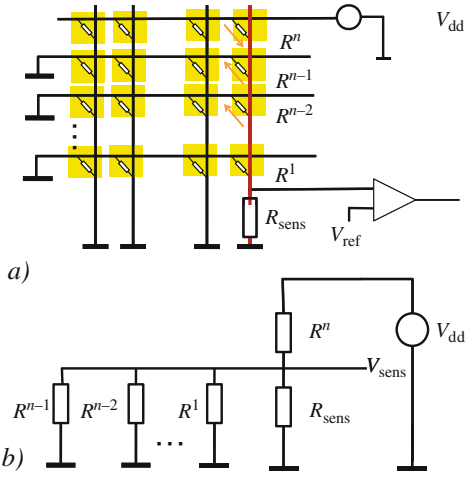


Fig. 1.9 Resistor model of the crossbar circuitry. **a** the resistance state of R^n is probed. **b** a simplified circuit diagram, with the resistances that influence the value of i_{bit}/Vsense. The sense current is measured as a voltage drop on R_{sens} and fed into a comparator. From: [33] Reproduced with permission. Copyright 2009 IEEE

or below become increasingly difficult to realize if one takes into account the resistance of the wires that connect the nanoscale crossbar array to a silicon-based read-out circuitry.

Another disadvantage of this read-out method is that it drives a current through a large number of unaccessed junctions thus dissipating power unnecessarily—unless the bit values along this word line are read out in parallel. The frequent access may cause early degradation of the molecular junctions and the driven word line is loaded with a very high current, potentially destroying it. The "wasted" current and power are proportional to the size of the array n.

Another possible read-out scenario involves leaving the unselected bit and/or word lines floating. This minimizes power dissipation and current inflow. Unfortunately in that case the parasitic current paths no longer end in a shortcut to the ground, but they flow into the signal path and interfere with the read-out process.

In a realistic case a crossbar memory is built from nonlinear, non-rectifying elements, such as the molecules investigated in [70]. We assume a self-assembled monolayer built from 100 molecules with each molecule having the same current–voltage characteristics as a stand-alone molecule in the break-junction experiment of [70]. Our choice is a rough estimation for the number of molecules in a self-assembled monolayer, covering few tens of nanometer squared area junction. The static $I(V)$ curve of the molecules (i.e., the two branches of the hysteretic switch) was implemented by piecewise linear voltage-controlled current sources in SPICE, as shown in Fig. 1.10. This example also serves as a model system for a large class of realizable molecular switches. We assumed a symmetric $I(V)$ curve ($I(V_0) = I(-V_0)$), even if the real molecular $I(V)$ is slightly asymmetrical.

The figure also shows the current corresponding to $R_{on} = 3.5\,\mathrm{M\Omega}$ and $R_{off} = 23\,\mathrm{M\Omega}$ values that we used in the linear model. When the junction is biased to

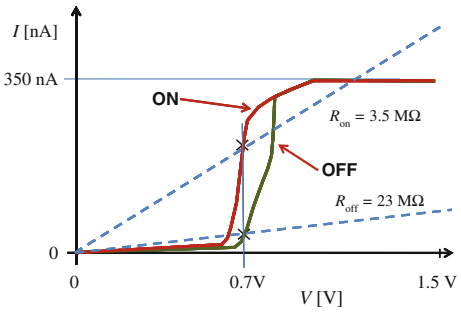


Fig. 1.10 The $I(V)$ curve of the investigated molecular monolayer, based on [70]. From: [33]. Reproduced with permission. Copyright 2009 IEEE

$V = 0.7$ V, the currents flowing through the "real" molecular layer and these resistors is the same, which allows us to compare our results with (linear) resistive and nonlinear junctions and understand the benefits or disadvantages of nonlinearities.

The scalability diagram (relative noise margin) of the crossbar is shown in Fig. 1.11. For 100×100 arrays the read-out margin for floating junctions is about 10 times larger with the nonlinear junctions. The voltage of the word lines stabilizes slightly below $V_{dd}/2$ and the bit lines are settled a bit above $V_{dd}/2$. The unselected junctions get about half of the voltage of the selected junction and due to the steep nonlinearity only very small current flows on them. The finite parallel resistance of molecular junctions still results in a "background" current, making the read-out more difficult at larger array sizes. From Fig. 1.11 we can extrapolate though that noise margin will not be a crucial problem up to many hundred kilobit arrays. The dashed curve of Fig. 1.11 shows the relative noise margin if all the unselected junctions are biased to $V_0 = 0.1$ V—biasing significantly improves the noise margin but it will also increase dissipation and unnecessary current inflow. For the 100×100 array the current inflow in the biased word line is only three times bigger than the sense current on the interrogated bit line. It is clearly possible to find a good compromise between the read-out margin and the extra current inflow/power.

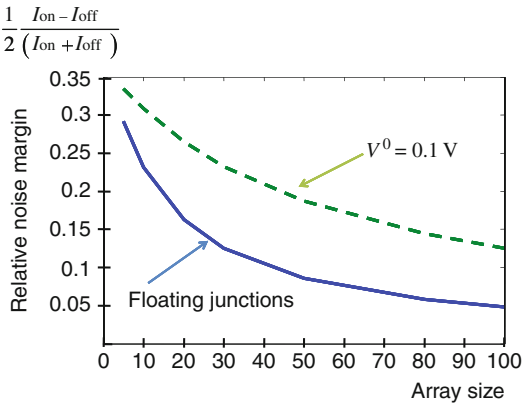


Fig. 1.11 Scalability of the crossbar with non-linear junctions if the unselected wires are left floating or they biased to a $V_0 = 0.1$ V voltage. From: [33]. Reproduced with permission. Copyright 2009 IEEE

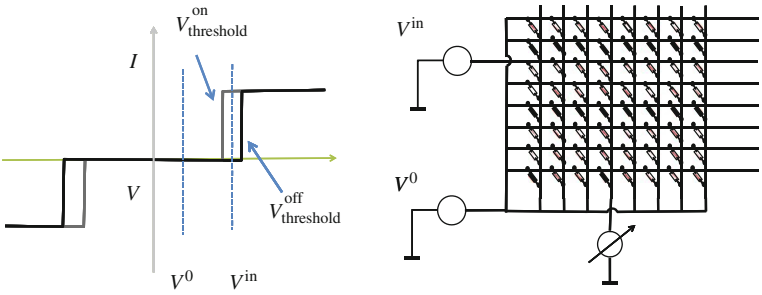


Fig. 1.12 Current–voltage characteristics of an idealized molecular switch and its corresponding read-out scheme. From: [33] Reproduced with permission. Copyright 2009 IEEE

An ideal, step-like nonlinearity with distinct $V^{on}_{threshold}$ and $V^{off}_{threshold}$ voltages (see Fig. 1.12) in the current–voltage characteristics of the molecule clearly would eliminate the problems of read-out. Biasing the chosen word line to $V^{in} = (V^{off}_{threshold} - V^{on}_{threshold})/2$, grounding the selected bitline, and putting all the unselected connections to $V^0 = (V^{off}_{threshold} - V^{on}_{threshold})/4$ voltage, the parasitic currents cease, as long as $V^0 < V^{on}_{threshold}$. For non-ideal step-like nonlinearities the leakage currents below the threshold voltage add up, resulting in a background current. This eventually reduces the noise margin, as shown in Fig. 1.11.

Organic Thin-Film Transistors

In was the year 1977, when Alan J. Heeger, Alan G. MacDiarmid, and Hideki Shirakawa discovered a carbon-based, highly conductive polymer: the oxidized, iodine-doped polyacetylene. For their discovery, which was one of the most important milestones for organic electronics, they were jointly awarded the chemistry Nobel Prize in 2000. Since then, carbon-based semiconductors have allowed scientists and engineers to develop devices with novel features and reduced costs, making this technology a premier candidate for several sectors of electronics. Organic semiconductors can be manufactured and processed at room temperature, making their production easier and cheaper than for conventional silicon and inorganic semiconductors; they can be transparent, flexible, and developed over large areas or non-planar geometries; completely plastic devices can be realized.

Their performances, compared to silicon, are, however, inferior: carrier mobility is orders of magnitude lower, even if this largely increased in the last years; they are highly sensitive to environmental conditions, especially to the atmosphere (oxygen and humidity); their performances decrease over time. Despite these drawbacks, the great potential of organic electronics has led to extensive theoretical and experimental research since the first discoveries during the late 1970s.

Organic transistors are three terminal devices in which the current flow going between the source and drain is modulated by a gate potential. A major difference

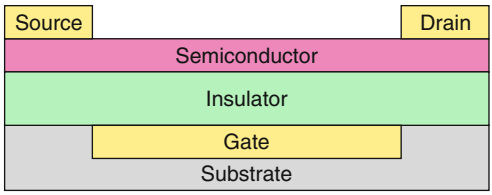


Fig. 1.13 Structure of a top contact transistor

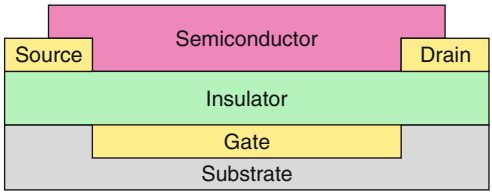


Fig. 1.14 Structure of a bottom contact transistor

with commonly used inorganic transistors is that no inversion layer is formed, but the conduction occurs by means of the majority carriers, which accumulate at the semiconductor/insulator interface. Possible transistor structures are shown in Figs. 1.13 and 1.14. Both have a substrate which acts as mechanical support for the structure. Over it, there is an insulating dielectric film, under which a gate contact is realized. Top contact structures have the semiconducting layer all over the insulator, with the source and drain contacts lying on top of the semiconductor; on the contrary, bottom contact devices have their contacts under the semiconductor layer. Moreover, both structures can also be realized without the substrate, if the mechanical support to the structure is provided by the insulating film itself (*free standing* devices).

In this section we will analyze organic transistors and simple circuits that can be realized with them, like inverters and ring oscillators. We will describe the structure and physical behavior of the transistors, as well as their electrical characteristics and associated models. Since the physics of organic semiconductors is quite different from that of silicon, because of their molecular structure, we will first describe how charge transport occurs inside the semiconductor.

Charge Transport in Conjugated Polymers

Materials employed as organic semiconductors are conjugated polymers, i.e., sp^2-hybridized linear carbon chains. This kind of hybridization is the one responsible for giving (semi)-conducting properties to organic materials. Starting from the electron configuration of a single carbon atom, which is $1s^2\,2s^2\,2p^2$, when multiple carbons bond together different molecular orbitals originate, as the electron wavefunctions mix together [82, 83].

While $1s$ orbitals of the carbons do not change when the atoms are bonded, $2s$ orbitals mix their wavefunctions with two of the three $2p$ orbitals, leading to the

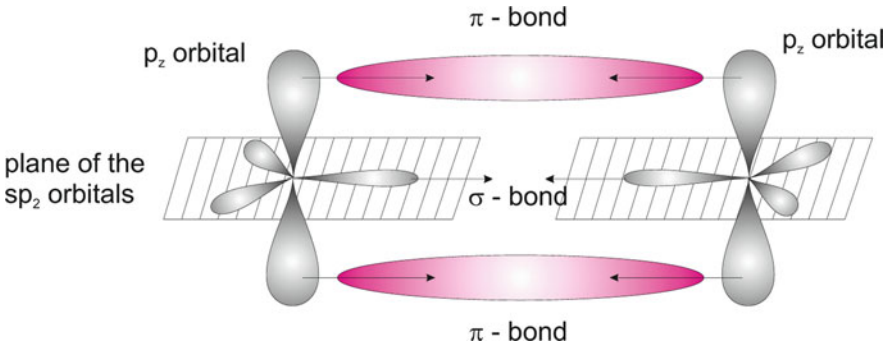


Fig. 1.15 sp^2 hybridization of two carbon atoms. sp^2 orbitals lie on the same plane and bond into a σ bond, p_z orbitals are orthogonal to the plane and bond into a π bond

electron configuration that is reported in Fig. 1.15. Three sp^2 orbitals are formed and lie on the molecular plane, at 120° to each other, leaving one p orbital which is orthogonal to the molecular plane. Hybrid sp^2 orbitals can then give origin to different bonds. sp^2 orbitals from different carbon atoms which lie on the molecular plane form strong covalent bonds, which are called σ bond, whereas p orbitals which are orthogonal to the molecular plane can mix to give less strong covalent bonds, which are called π bonds. The π bonds constitute a delocalized electron density above and below the molecular plane and are responsible for the conductivity of the molecule, as the charge carriers move through these bonds. This transport process is usually called *hopping transport* [84, 85] and is described by means of quantum mechanical tunneling.

The energy levels of the orbitals, when multiple molecules are considered, give origin to energy bands, in analogy to what happens in inorganic semiconductors: the edge of the valence band corresponds then to the highest occupied molecular orbital (HOMO), whereas the edge of the conduction band corresponds to the lowest unoccupied molecular orbital (LUMO). The difference between the energy of the HOMO and the energy of the LUMO is the energy gap E_g of the organic material and usually $1.5 < E_g < 4\,\text{eV}$ [86].

Charge transport in organic semiconductor is, however, quite different with respect to silicon and other mono-crystalline inorganic semiconductors. The periodic lattice of these materials and their very low density of defects allow one to accurately describe the charge transport by means of delocalized energy bands separated by an energy gap. Most of the organic semiconductors, on the other hand, are amorphous and rich in structural and chemical defects, therefore requiring conventional models for charge transport to be adapted and extended; moreover, charges can move, with different mobilities, within the molecular chain (intra-chain), between adjacent molecules (inter-chain), or between different domains, generally referred to as grains (inter-grain).

In reality, transport in organic semiconductors is a much more complex process than described above [87]. One possible description can be obtained in terms of

hopping processes along molecular chains and controlled by quantum mechanical tunneling. In the following, we will describe a simplified approach similar to those used in inorganic semiconductors. Hopping transport has been simulated with the Monte Carlo method: analysis of the temperature and electric field dependence of the mobility can be found in [88], and later in [89, 90], where the dependence on the temperature and the electric field of the carrier mobility is studied; in [91] Olivier et al. study the impact of the molecular parameters on hopping rates and charge mobilities in organic π-conjugated materials.

Drift Diffusion Model

Simulations for organic devices and more complex structures, which thoroughly account for the hopping theory, can be too complex and computationally onerous, so simpler models have been adapted from the physics of the inorganic semiconductors. One of the most widely used models for semiconductor simulations is the drift diffusion model, which consists of a set of nonlinear partial differential equations that relate three physical quantities, namely the electrostatic potential V, the density of electrons n, and the density of holes p inside the device structure. The first of these equations is the Poisson equation:

$$\nabla \cdot (\varepsilon \nabla V) = q\left(n - p + N_{\mathrm{A}} - N_{\mathrm{D}} + n_{\mathrm{tr}}\right), \tag{1.7}$$

where ε is the electric permittivity of the material; q is the elemental charge; N_{A} and N_{D} are the concentrations of ionized acceptors and donors, respectively; and n_{tr} is the density of occupied traps. The continuity equations are

$$\nabla \cdot \vec{\mathbf{J}}_n = +qR_n + q\frac{\partial n}{\partial t}, \tag{1.8}$$

$$\nabla \cdot \vec{\mathbf{J}}_p = -qR_p - q\frac{\partial p}{\partial t}, \tag{1.9}$$

where $\mathbf{J}_n$ and $\mathbf{J}_p$ are the current densities and R_n and R_p are the net generation–recombination rate.

In the DD model the currents of electrons and holes are described as the sum of two contributions, namely a drift component, proportional to the electrostatic field $\mathcal{E} = -\nabla V$, and a diffusion component, proportional to the gradient of the carrier density.

$$\vec{\mathbf{J}}_n = -qn\mu_n \nabla V + qD_n \nabla n, \tag{1.10}$$

$$\vec{\mathbf{J}}_p = -qp\mu_p \nabla V - qD_p \nabla \tag{1.11}$$

The charge transport properties are described using the mobilities μ_n and μ_p and the correspondent diffusivities D_n and D_p, which are assumed to be related by the Einstein relation $D = kT/q\mu$. As already mentioned previously, they can

Table 1.1 Mobility of p-type and n-type semiconductors

Compound	*Mobility* (cm^2/Vs)
p-Type	
Pentacene	3.0
Picene	1.1
2,7-Diphenyl[1]benzothieno[3,2-*b*]benzothiophene	2.0
DNSS	2.9
Rubrene (single crystal)	15
Bis-5'-alkylthiophen-2'-yl-2,6-anthracene	0.5
Thiophene oligomers	1.1
Phenylene systems	0.32
n-Type	
Perlylenediimides	0.6
Fullerene	0.56
Perfluoronated pentacene	0.11
Thiophene oligomers with perfluorohexyl groups	0.048

be field-and temperature-dependent. Table 1.1 shows the highest (so far) measured mobilities for some organic p and n semiconductors. Reference [92] contains a complete and in-depth review of organic semiconductor mobility.

The density of carriers can be described by means of Boltzmann statistics:

$$n = N_{\mathrm{C}} \exp\left(\frac{E_{\mathrm{qF},n} - E_{\mathrm{C}}}{kT}\right), \tag{1.12}$$

$$p = N_{\mathrm{V}} \exp\left(\frac{E_{\mathrm{V}} - E_{\mathrm{qF},p}}{kT}\right), \tag{1.13}$$

where $E_{\mathrm{qF},n} = -q\Phi_n$ and $E_{\mathrm{qF},p} = -q\Phi_p$ are the quasi-Fermi energies for electrons and holes and Φ_n and Φ_p are the quasi-Fermi potentials, respectively. N_{C} and N_{V} are the effective densities of states and are usually assumed to be equal to the density of molecules, since all the states are thermally accessible [93].

E_{C} and E_{V} are the conduction and valence band edges, defined as

$$E_{\mathrm{C}} = -q\chi - q\left(V - \phi_{\mathrm{ref}}\right), \tag{1.14}$$

$$E_{\mathrm{V}} = -q\chi - E_{\mathrm{g}} - q\left(V - \phi_{\mathrm{ref}}\right), \tag{1.15}$$

where $q\chi$ is the electron affinity and E_{g} the band gap. The reference potential ϕ_{ref} can be set equal to the Fermi potential of an intrinsic semiconductor. Then Eqs. (1.12) and (1.13) become

$$n = n_{\mathrm{i}} \exp\frac{q\left(V - \phi_n\right)}{kT}, \tag{1.16}$$

$$p = n_{\mathrm{i}} \exp\frac{q\left(\phi_p - V\right)}{kT}, \tag{1.17}$$

where $n_{\mathrm{i}} = \sqrt{N_{\mathrm{C}} N_{\mathrm{V}}} \exp\left(-E_{\mathrm{g}}/2kT\right)$ is the intrinsic density.

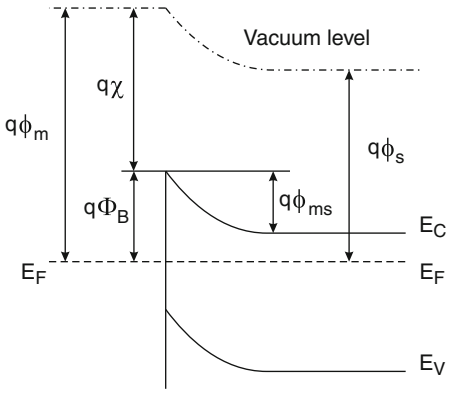


Fig. 1.16 Band diagram for a Schottky contact between an n-type semiconductor and metal at the thermal equilibrium. $q\phi_s$ is the workfunction of the semiconductor, E_C is the conduction band level (LUMO in organic semiconductors), E_V is the valence band level (HOMO)

Boundary conditions are required to solve the DD equations for the device structure. If a metal/semiconductor interface is present, it has to be specified whether it constitutes an ohmic or Schottky contact, as well as for metal/insulator interfaces (e.g., a gate contact in a transistor).

For ohmic contacts charge neutrality and equilibrium are assumed at the metals/semiconductor interface. For Schottky contacts, a barrier height is considered (the difference between the metal workfunction $q\phi_m$ and the electron affinity $q\chi$), given by the difference $q\Phi_B = q\phi_m - q\chi$, as reported in Fig. 1.16, which shows a band diagram for a Schottky contact between an n-type semiconductor and metal, including also the difference between the metal and the semiconductor workfunction $q\phi_{ms}$.

Analytical Models for OTFTs

The determination, by analytical means, of the drain current for any applied voltages has been object of extensive research over the years. In this section, we will describe a model based on constant mobility and doping, in linear and saturation region. It is mostly based on [94]. Extensions of this model, accounting for space charge limited currents and Poole–Frenkel mobility can be found in [95]. Other contributions to the modeling of the organic transistors, including the effects of traps and ambipolar devices, come from Stallinga et al. in [96, 97].

Figure 1.17 shows the basilar structure for the transistor to be modeled. It is a free-standing bottom contact structure. The length of the channel is L, its width is Z. The dielectric layer has thickness d_i and dielectric permittivity ε_i, with an associated capacity $C_i = \varepsilon_i/d_i$, whereas the semiconductor film has thickness d_s and dielectric permittivity ε_s, with an associated capacity $C_s = \varepsilon_s/d_s$. We will assume an n-type semiconductor with doping N and a density of free carriers $n_0 \simeq N$; an analogous model can be derived for p-type semiconductors changing the sign of the current and the applied voltages.

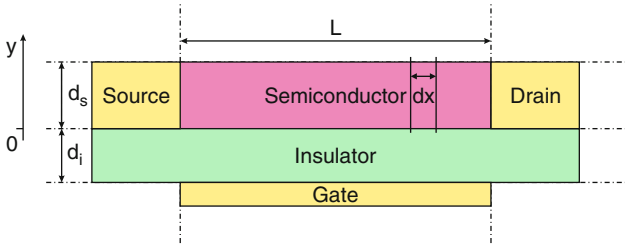


Fig. 1.17 Geometry of the bottom contact transistor

The threshold voltage of the device, which accounts for different workfunctions between the semiconductor and the gate and for possible charges located in the insulator or in the insulator/semiconductor interface, is set to be zero, to simplify notation. Since it just determines a shift of the gate voltages, non-zero threshold voltages can be included in the model substituting V_g with $V_g - V_{th}$.

The total drain current I_d in the linear region ($V_d < V_g$) can be expressed as the sum of the bulk current originating from the free carriers in the semiconductor and the current determined by the carriers in the accumulation layer. Therefore

$$\frac{I_d}{Z\mu} = -\left[Q_g(x) + Q_0\right]\frac{dV}{dx}, \tag{1.18}$$

where $Q_0 = qn_0d_s$ is the surface density of the free carriers and

$$Q_g(x) = -C_i\left[V_g - V_s(x) - V(x)\right]$$

is the surface density of the carriers in the accumulation layer, which forms as a capacitive effect. $V_s(x)$ is the ohmic drop in the bulk of the semiconductor, which can be generally neglected [98], and $V(x)$ is the potential with respect to the source at the coordinate x. Equation Eq.(1.18) can be integrated from $V(x = 0) = 0\,\mathrm{V}$ to $V(x = L) = V_d$ to obtain, for a constant mobility,

$$I_d = \frac{Z}{L}\mu C_i\left[\left(V_g + V_0\right)V_d - \frac{V_d^2}{2}\right]. \tag{1.19}$$

The transconductance in the linear zone, from which the field effect mobility can be extracted, can then be defined as

$$g_m = \left(\frac{\partial I_d}{\partial V_g}\right)_{V_d=\text{const}} = \frac{Z}{L}\mu C_i V_d. \tag{1.20}$$

When the drain voltage exceeds the effective gate voltage $V_d > V_g$ the accumulation layer near the drain changes to a depletion layer. The depletion layer will extend from a coordinate x_0 such that $V(x_0) = V_g$ to the drain contact ($x = L$). At those coordinates where there is the depletion layer, the only contribution to the current is

given by the free carriers, which are only in the volume of the semiconductor which has not been depleted. Then

$$I_{\mathrm{d}}\mathrm{d}x = Zq\mu n_0 \left[d_{\mathrm{s}} - W(x)\right]\mathrm{d}V, \tag{1.21}$$

wherein $W(x)$ is the width of the depletion region, which can be determined solving the Poisson equation [98] to get

$$W(x) = \frac{\varepsilon_{\mathrm{s}}}{C_{\mathrm{i}}}\left\{\sqrt{1 + \frac{2C_{\mathrm{i}}^2\left[V(x) - V_{\mathrm{g}}\right]}{qN\varepsilon_{\mathrm{s}}}} - 1\right\}. \tag{1.22}$$

Since we have assumed that the accumulation layer extends up to a point where $V(x_0) = V_{\mathrm{g}}$ and since beyond x_0 only the non-depleted free carriers add to the current, we can write the latter as the sum of these two contributions:

$$I_{\mathrm{d}} = \frac{Z}{L}\mu C_{\mathrm{i}}\int_0^{V_{\mathrm{g}}}\left(V_{\mathrm{g}} + V_0 - V\right)\mathrm{d}V + \frac{Z}{L}\mu qn_0\int_{V_{\mathrm{g}}}^{V_{\mathrm{dsat}}}\left(d_{\mathrm{s}} - W\right)\mathrm{d}V. \tag{1.23}$$

The second integral on the right-hand side means that we account for all the free carriers which are in the semiconductor out-of-the-depletion region, whose extension is $W(V_{\mathrm{g}}) = 0$ at the end of the accumulation layer and is $W(V_{\mathrm{dsat}}) = d_{\mathrm{s}}$ when the device reaches the saturation voltage. Changing the integration variable in Eq. (1.23) from V to W leads to

$$\begin{aligned} I_{\mathrm{d}} &= \tfrac{Z}{L}\mu C_{\mathrm{i}}\textstyle\int_0^{V_{\mathrm{g}}}\left(V_{\mathrm{g}} + V_0 - V\right)\mathrm{d}V + \tfrac{Z}{L}\mu\tfrac{q^2n_0N}{\varepsilon_{\mathrm{s}}}\int_0^{d_{\mathrm{s}}}\left(d_{\mathrm{s}} - W\right)\left(W + \tfrac{\varepsilon_{\mathrm{s}}}{C_{\mathrm{i}}}\right)\mathrm{d}W = \\ &= \tfrac{Z}{L}\mu\left[C_{\mathrm{i}}\left(\tfrac{V_{\mathrm{g}}^2}{2} + V_0V_{\mathrm{g}}\right) + \tfrac{q^2n_0N}{\varepsilon_{\mathrm{s}}}\tfrac{d_{\mathrm{s}}^3}{6}\left(1 + \tfrac{3C_{\mathrm{s}}}{C_{\mathrm{i}}}\right)\right], \end{aligned} \tag{1.24}$$

The pinch-off voltage V_{p}, i.e., the gate voltage for which the depletion region extension is equal to d_{s}, can be obtained by substituting $V_{\mathrm{p}} = V_{\mathrm{g}} - V(x)$ in Eq. (1.22). This leads to

$$V_{\mathrm{p}} = \pm\frac{qNd_{\mathrm{s}}^2}{2\varepsilon_{\mathrm{s}}}\left(1 + 2\frac{C_{\mathrm{s}}}{C_{\mathrm{i}}}\right) \simeq \frac{qNd_{\mathrm{s}}}{C_{\mathrm{i}}}, \tag{1.25}$$

wherein in the approximated expression the flat band voltage has been neglected and $d_{\mathrm{s}} \ll d_{\mathrm{i}}$ is reasonably assumed. If $n_0 = N$ is also assumed, then $V_{\mathrm{p}} = V_0$ and

$$I_{\mathrm{dsat}} = \frac{Z}{2L}\mu C_{\mathrm{i}}\left(V_{\mathrm{g}} - V_0\right)^2. \tag{1.26}$$

Figure 1.18 shows a simulation of OTFT output curves based on Eqs. 1.19 and 1.26.

In the previous calculations, the mobility has always been considered constant; however, this assumption is rather simplistic, as it usually depends on the temperature and the electric field in the semiconductor. Common models are the

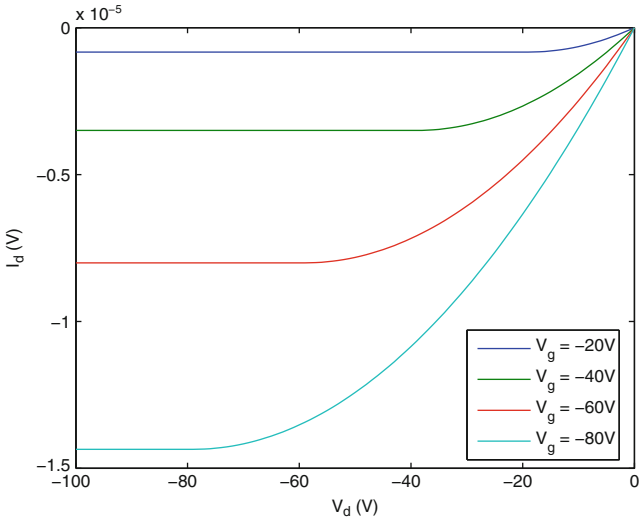


Fig. 1.18 Output curves simulated with Eq. (1.19) and (1.26). $Z/L = 200$, $\mu = 10^{-3}\text{cm}^2/\text{s}$, $C_i = 2.3 \cdot 10^{-8}\text{F/cm}^2$, $V_0 = -1\text{V}$

multiple trap and release (MTR) mobility model [99] or the variable range hopping (VRH) [100], for which the effective mobility is

$$\mu_{FET} = \frac{L}{C_\text{i} W V_\text{d}} \frac{\partial I}{\partial V_\text{g}} = \mu_0 \left(V_\text{g} - V_\text{th}\right)^\alpha , \tag{1.27}$$

wherein V_g is the applied gate voltage and μ_0 and α are parameters usually extracted by means of fitting of the experimental data.

Another mobility model was theorized in 1938 by Frenkel [101], to explain the increase of conductivity in insulators and semiconductors when high fields are applied. Its form is

$$\mu = \mu_0 \exp\left(\gamma \sqrt{\mathcal{E}}\right), \tag{1.28}$$

where μ_0 and γ are parameters which depend on the physics of the considered system and $\mathcal{E}$ is the electric field. Poole–Frenkel mobility has been extensively studied for modeling charge transport in organic LEDs with disordered semiconductors [102–104], but it can also describe several experimental results for organic field effect transistors which show a variation of the mobility with the source-drain electric field [105–107].

Simulation of Organic Thin Film Transistors

We simulate an organic, bottom contact, thin-film transistor. The equations of the drift diffusion model are solved using the finite elements method [108]. The Poisson equation, together with the continuity equations, is solved as a system of partial

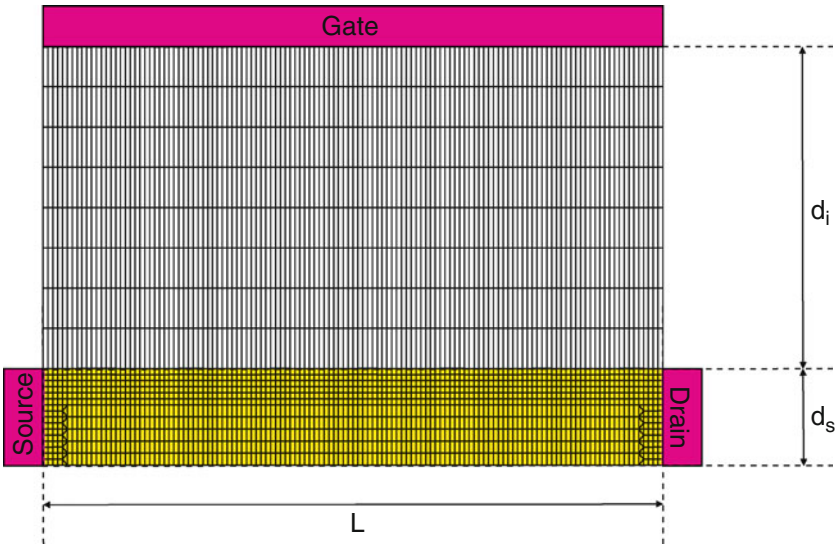


Fig. 1.19 Device mesh

differential equations on a discrete mesh grid. The solution is determined for each node of the grid. Figure 1.19 shows a possible grid for a device with channel length $L = 50\mu$m, channel width $W = 5$mm, semiconductor thickness $d_s = 30$nm, and insulator thickness $d_s = 100$nm. Even if it is not shown in the picture, air over the device is also simulated, because the electric field exceeds the semiconductor boundaries. The complete set of parameters used in the simulations is shown in Table 1.2.

The output curves are shown in Fig. 1.20. For a deeper understanding of the device behavior, one can look at the potential or charge profiles inside the structure. For example, from Fig. 1.21 one can see how the potential in the semiconductor follows the gradual channel approximation, whereas the charge profile in Fig. 1.22 shows how, under the application of the gate bias, an accumulation layer of holes forms (and respectively, there is a depletion of electrons).

Table 1.2 Material parameters used in the simulation

Parameter	*Symbol*	*Value*
Dielectric permittivity	ε_i	4
Semiconductor permittivity	ε_s	4
Electron affinity	$q\chi$	2.5 eV
Energy gap	E_g	2.5 eV
Valence band density	N_C	$2.8 \cdot 10^{21}\text{cm}^{-3}$
Conduction band density	N_V	$2.8 \cdot 10^{21}\text{cm}^{-3}$
Electrons mobility	μ_n	$10^{-20}\text{cm}^2/\text{Vs}$
Holes mobility	μ_p	$0.0381\text{cm}^2/\text{Vs}$
Metal workfunction	$q\phi_m$	5 eV

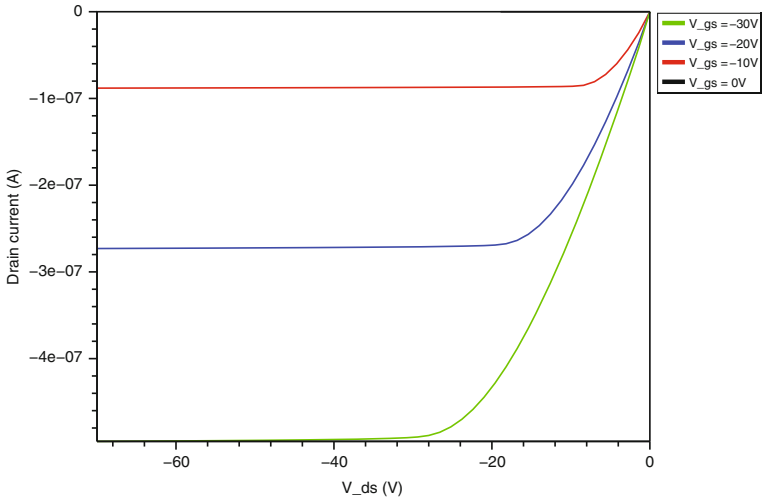


Fig. 1.20 Output curves for bottom contact OTFT device. Channel length is $L = 50\mu$m, channel width $W = 5$mm, semiconductor thickness $d_s = 30$nm, insulator thickness $d_s = 100$nm

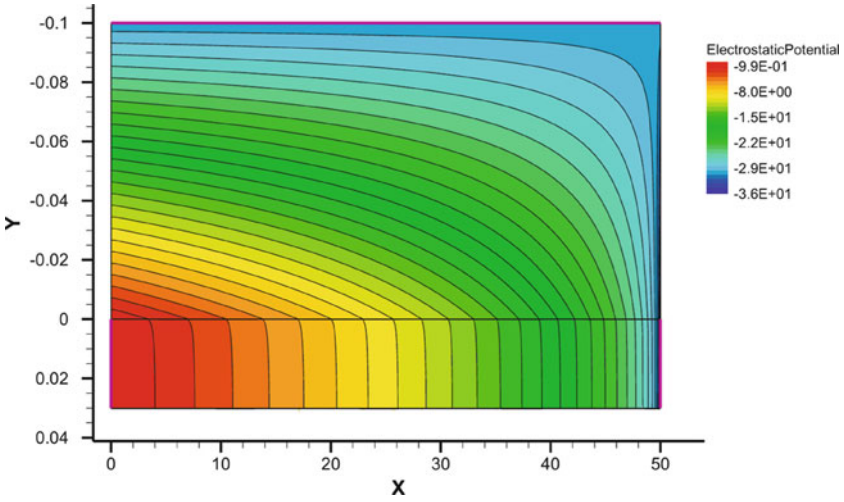


Fig. 1.21 Potential distribution in the structure. Gate potential $V_g = -30$V, drain potential $V_d = -35$V

Experimental results in literature, for example the work of Schmechel and von Seggern [109, 110], show that the performances of organic transistors are usually affected by trap states located in the energy gap of the material.

The origin of these localized states is still not completely clear: for example, shallow traps that can arise in vapor-deposited pentacene from the sliding of the molecules along the molecular axis, while two-dimensional crystallinity is

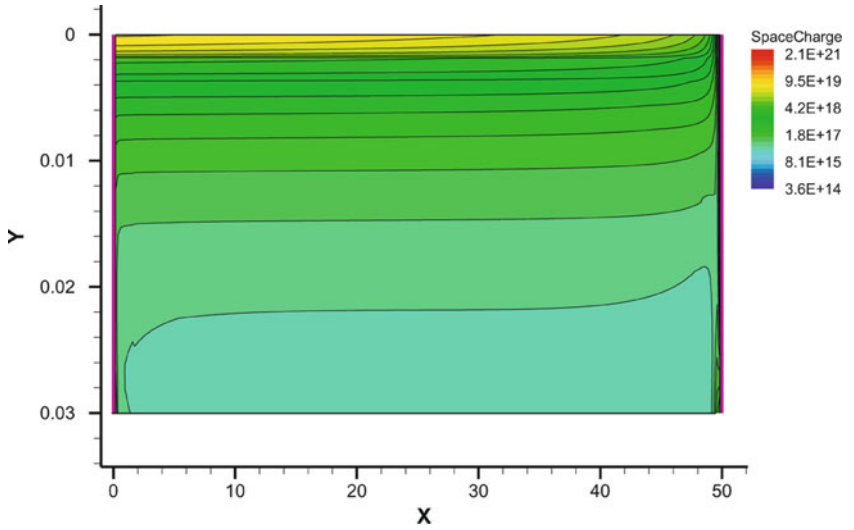


Fig. 1.22 Net charge distribution in the semiconductor. Gate potential $V_g = -30$V, drain potential $V_d = -35$V

maintained [111]. Northrup et al. [112] discuss the effects of hydrogen-related defects in crystalline pentacene, i.e., when the pentacene molecule $C_{22}H_{14}$ has one more hydrogen atom and becomes $C_{22}H_{15}$, but also analyze the effects of oxygen-related defects, i.e., when one oxygen atom replaces one hydrogen atom to give $C_{22}H_{13}O$. Both defects originate on exposure of the semiconductor to air humidity, but hydrogen defects could also come from bias stress in the device. Water adsorption on the surface of the insulator has also been suggested for the generation of discrete traps states in single-crystal pentacene [113].

The theory of Schockley–Read–Hall [114, 115] can be applied also to organic semiconductor, including in the drift diffusion model exposed in section "Drift Diffusion model" one more equation that describes the trap dynamics:

$$\frac{\mathrm{d}n_{\mathrm{tr}}}{\mathrm{d}t} = \frac{1}{\tau_n}\left[\exp\frac{(E_{\mathrm{tr}} - E_{\mathrm{i}})}{kT} f_{\mathrm{tr}} - n\,(1 - f_{\mathrm{tr}})\right] + \frac{1}{\tau_p}\left[\exp\frac{(E_{\mathrm{i}} - E_{\mathrm{tr}})}{kT}\,(1 - f_{\mathrm{tr}}) - p f_{\mathrm{tr}}\right], \tag{1.29}$$

where $n_{\mathrm{tr}} = N_{\mathrm{tr}} f_{tr}$ is the density of occupied centers, proportional to the density of traps N_{tr} and to the probability of occupation f_{tr}; $\tau_n = 1/v_{\mathrm{th},n}\sigma_n N_{\mathrm{tr}}$ and $\tau_p = 1/v_{\mathrm{th},p}\sigma_p N_{\mathrm{tr}}$ are the lifetime for holes and electrons, respectively; $v_{\mathrm{th},n}$, σ_n, $v_{\mathrm{th},p}$, σ_p are the carrier thermal velocity and the capture section, for electrons and holes, respectively. $E_{\mathrm{tr}} - E_{\mathrm{i}}$ is the difference between the energy levels of the trap center and the intrinsic semiconductor.

Figure 1.23 shows a simulation for a bottom contact transistor, in which the drain electrode has a constant bias $V_d = -25$V, while the gate electrode potential is swept between $V_g = +40$V and $V_g = -40$V, then again to $V_g = +40$V.

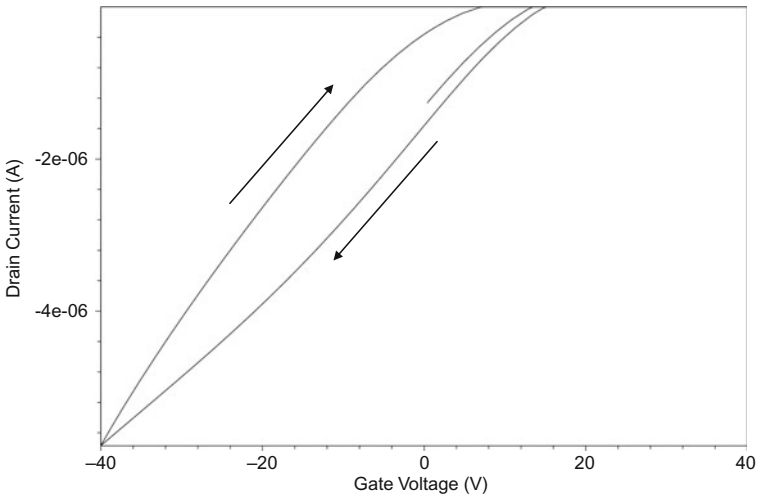


Fig. 1.23 Transient simulation of an organic transistor with $W = 5\text{mm}$, $L = 75\mu\text{m}$, $d_\text{i} = 1.6\mu\text{m}$, $d_\text{s} = 50\text{nm}$. Acceptor traps at the interface have a surface density $N_{\text{tr}_\text{s}} = 10^{12}\text{cm}^{-2}$, $E_\text{tr} = E_\text{V} + 0.3\text{eV}$, $v_\text{th}\sigma_n = v_\text{th}\sigma_p = 2 \cdot 10^{19}\text{cm}^3/\text{s}$

The main macroscopic consequence of the presence of traps in a device is hysteresis [116, 117], because traps are charged and uncharged with specific dynamics which can lead to different operating conditions, even if the applied bias is the same.

Initially, with a positive gate bias, the traps are filled with electrons, positively increasing the threshold voltage in the device. For Gauss's law, the shift ΔV_th in the threshold voltage is equal to

$$\Delta V_\text{th} = -\frac{Q_\text{tr}(t)}{C_\text{i}}. \tag{1.30}$$

wherein Q_tr is the surface-trapped charge and C_i the insulator capacitance. Once the gate potential becomes negative, holes are attracted at the interface and start to neutralize the negative charges, reducing the threshold voltage. If the trap charging and de-charging times are slow enough, a memory effect arises, resulting in a hysteresis. Analytical models [118], as well as numerical simulations [119, 120], have been proposed to study these phenomena, which still constitute a hot topic in the field of organic electronics.

Circuits Based on OTFTs: Inverters and Ring Oscillators

A major restriction in the design of organic electronics results from the limitation to p-type transistors. Although research aims at realizing n-type devices [121], the respective material systems are still immature. Additionally, the patterning of n- and p-type OTFTs on the same substrate poses great challenges in the fabrication processes. Therefore, it is of interest to design exclusively p-based circuits for digital

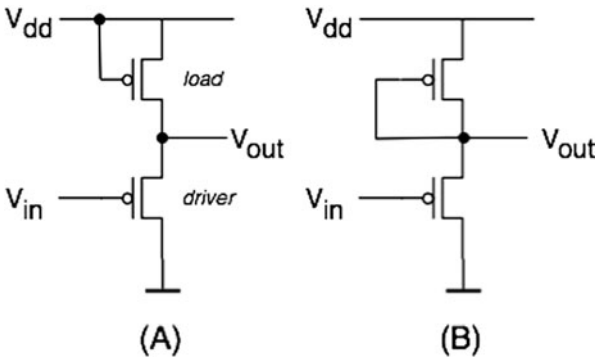


Fig. 1.24 Inverter layout with two p-type OTFTs in (**a**) saturated and (**b**) depleted load configuration

and analog purposes [122]. Figure 1.24 shows two possible inverter layouts with depleted and saturated load.

Generally, a great advantage of MOSFET-based circuits is the flexibility with which they can be tuned by adjusting the $\lambda = W/L$ ratios. Figures 1.25 and 1.26 demonstrate the strong influence of the geometry factor

$$\Gamma = \frac{\lambda_{\text{driver}}}{\lambda_{\text{load}}} = \frac{W_{\text{driver}} L_{\text{load}}}{L_{\text{driver}} W_{\text{load}}}$$

on the inverter transfer characteristic. The depleted load circuit exhibits superior performance with respect to inverter gain and the matching of input and output voltage

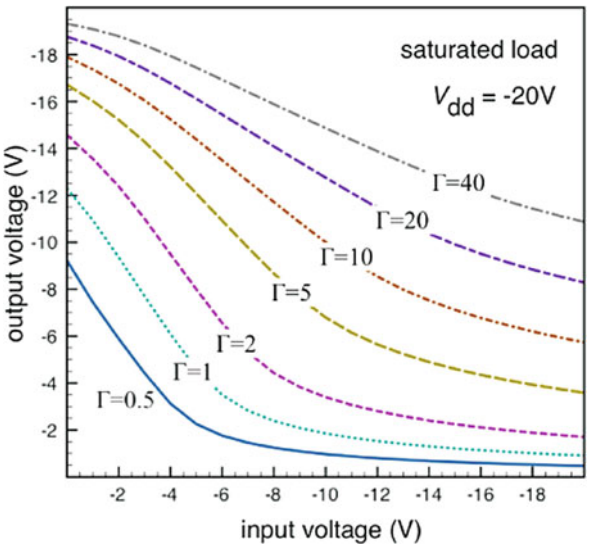


Fig. 1.25 Influence of the geometry factor on voltage transfer characteristic of an organic inverter with saturated load

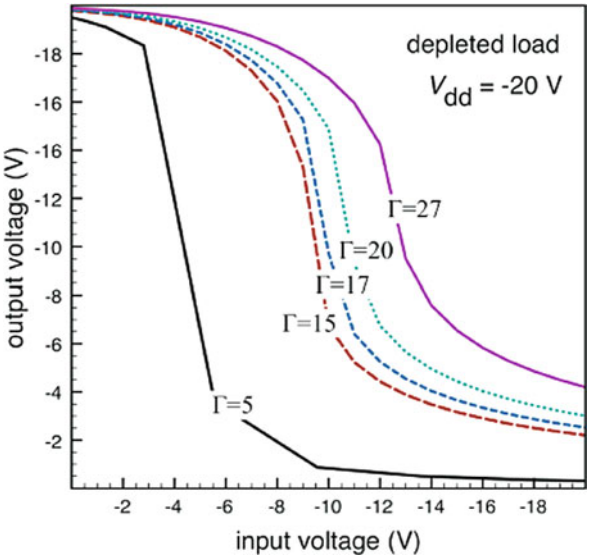


Fig. 1.26 Influence of the geometry factor on voltage transfer characteristic of an organic inverter with depleted load

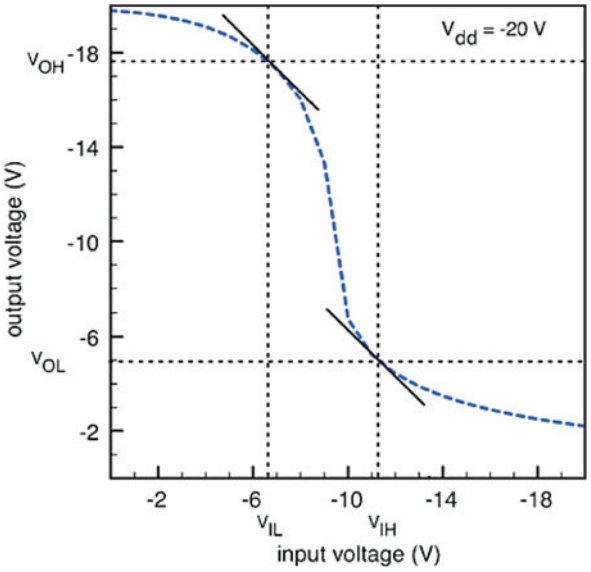


Fig. 1.27 Simulated voltage transfer characteristic of organic inverter with depleted load Γ=17, gain of 4.6, and static noise margin of $N_{ML} = V_{IL} - V_{OL} = -1.6V$ and $N_{MH} = V_{OH} - V_{IH} = -6.4V$

ranges. An inverter with reasonable symmetry and gain is the one with Γ=17 in Fig. 1.26.

Figure 1.27 shows the associated transfer characteristics. It has a gain of 4.6 and a static noise margin of $N_{ML} = V_{IL} - V_{OL} = -1.6V$ and $N_{MH} = V_{OH} - V_{IH} = -6.4V$, respectively.

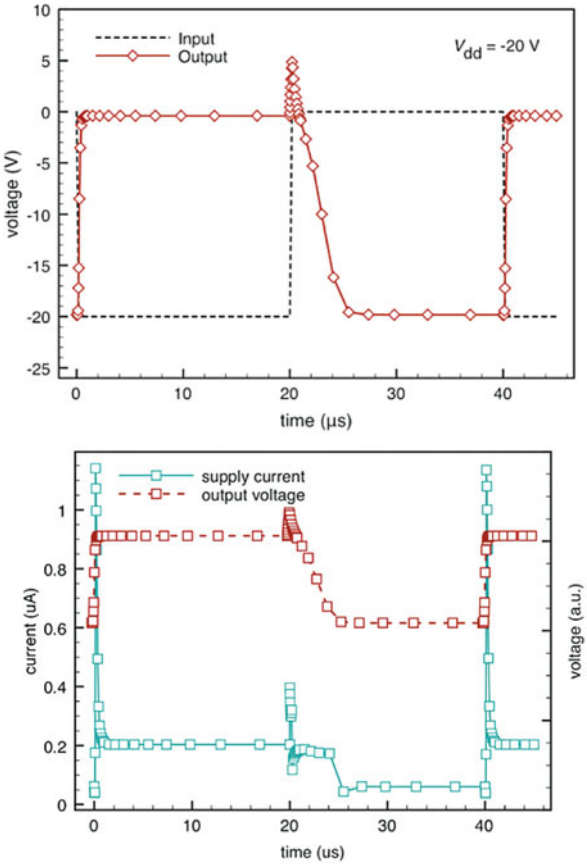


Fig. 1.28 Response of inverter with depleted load ($\gamma = 17$) and a load capacity $C_L = 20$fF to input pulse of $T = 40\mu$s. The *upper panel* shows the simulated output voltage transient and the *lower panel* the associated current consumption

To determine the switching behavior, we simulated this inverter in combination with a load capacity of $C_L = 20$fF. The transients are plotted in Fig. 1.28. In practical circuit applications, C_L is the sum of the input capacitance of the consecutive logic stage plus the parasitic capacitances of, for example, the wiring. The output voltage in Fig. 1.28 shows distinct asymmetry for rising and falling edges. The rise time going from 0 to -20V is approximately 5 μ s, while the fall time is close to 0.5 μ s. This asymmetry is a characteristic feature of inverters with depleted load and is also observed experimentally [123]. During the state $V_{in} = 0$V, the load capacitance C_L is fully charged and $V_{out} = -20$V. Changing the input to -20V instantly opens the driver OTFT. The resulting current discharges the capacitance. When the driver OTFT is subsequently closed with $V_{in} = 0$V, C_L needs to recharge in order to obtain the desired output level of -20V. For the charging process, current must flow through the load transistor, which still has a gate voltage of approximately 0 V. Therefore its impedance is initially high and the charging relatively slow. As demonstrated by the simulations, driving the output to -20V is as a result slower than shifting it in the opposite direction.

The switching times depend on the magnitude of the load capacitance. The overall speed decreases with increasing C_L. It is therefore important in the optimization of organic device and circuits to reduce unnecessary capacitances. Particularly critical are overlap capacitances at source and drain, which exist due to poor control on the positioning of the gate electrode. By inspecting the reported circuit layouts, it can be estimated that in terms of speed, most circuit realization are up to now rather limited by the layout than by the properties of the organic semiconductor. Current consumption is another critical circuit property. Non-complementary inverters have the disadvantage of constantly consuming power. The lower panel of Fig. 1.28 visualizes the supply current during pulse operation ($T = 40\mu$s). The switching of the input voltage causes strong peaks. But the invert also draws significant current from the supply, when the input is held at a constant level. The reduction of supply current is another interesting optimization task in which the presented simulation framework can fruitfully be employed.

The realization of organic ring oscillators has been a key milestone [28, 124]. Operation frequencies of up to 220 kHz have been reported [125]. The great attention on these systems has resulted from the fact that for logic applications, it is not sufficient to simply build inverters. It is required that they amplify the input signal so that it is not lost after a number of stages. Furthermore the output of one logic stage must be able to drive the input of another one. This requires that input and output voltage ranges match. A ring oscillator consists of an odd number of inverter stages. The output of one stage is connected to the input of the next stage. The last inverter connects to the input of the first one closing the ring. The odd number of stages prevents that system from obtaining a stable operation point. What makes the oscillator particularly interesting is the fact that each inverter must obey all the previously mentioned requirements in order to allow persistent oscillation. The realization of ring oscillators has therefore been the proof that large integrated organic circuits are feasible.

To test the capabilities of our simulation approach, we have set up a three-stage ring oscillator using the optimized inverter treated in the previous section. The overall circuit is shown in Fig. 1.29. We again included a load capacitance of $C_i = 20$fF

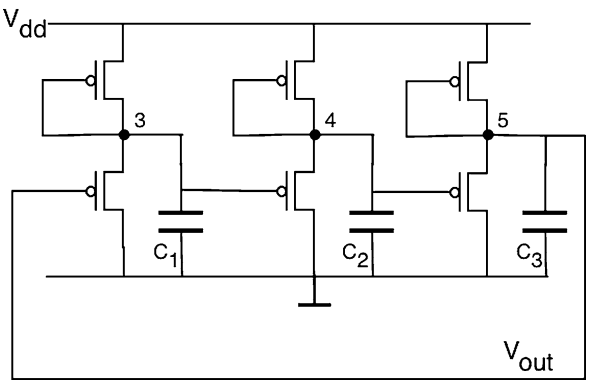


Fig. 1.29 Three-stage ring oscillator

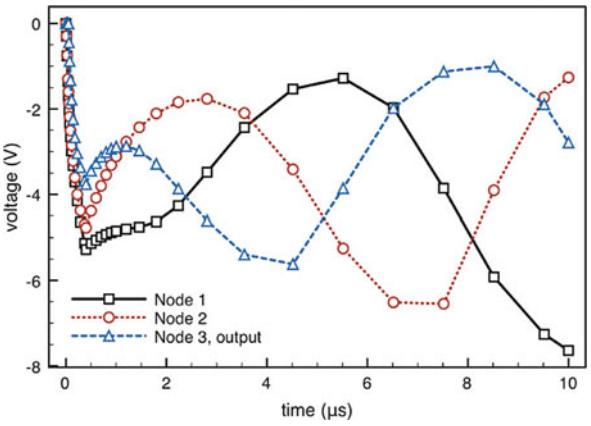


Fig. 1.30 Transient startup of oscillator circuit. V_{dd} is ramped to -20V during the first 0.5 µs

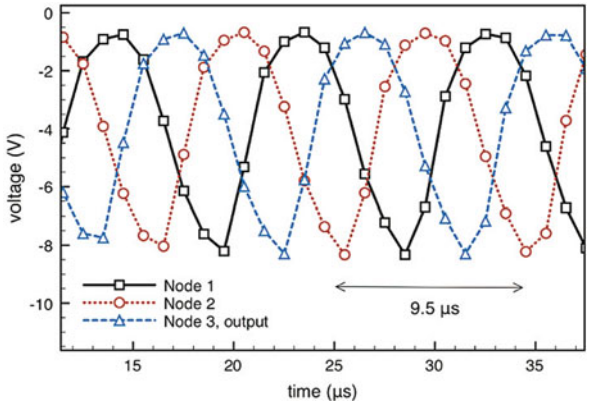


Fig. 1.31 Steady-state operation of simulated three-stage organic oscillator with $V_{dd} = -20$V

in each stage to explicitly account for parasitics. To start the oscillation, the supply voltage V_{dd} is ramped to -20V during the first 0.5 ms of the simulation. The node voltages at each inverter stage during startup are given in Fig. 1.30. After the initial V_{dd} ramp, the oscillator successfully reaches steady-state operation (compare Fig. 1.31). The simulated oscillation frequency equals 105 kHz and strongly depends on the magnitude of the load capacitances C_i. The stage delay, which results for the presented oscillator, has consequently a value of 3.16 ms. It is interesting to note that a half period of the oscillation is significantly shorter than the rise time of the transient inverter response in Fig. 1.28. Additionally, the output voltage range of the oscillator is limited to an interval of approximately -8 to -1V. Simulations have shown that the mismatch of rise and fall times in the inverter results in a timing deficiency within the oscillator. Effectively, the fast transmission of the falling signal in one inverter stage does not allow enough time for the slow rising edge in the following stage to reach a strongly negative value.

In conclusion, we have shown that relatively complex circuits can be realized with organic transistors. Simulation is a valuable tool for the optimization of such circuits.

Perspectives

The performances of organic semiconductor are continuously increasing, since their discovery in 1977. The mobility of the carriers has reached values comparable to those of the amorphous silicon, as shown by many promising compounds [92].

Integrated circuits have been developed; for example, ambipolar transistors and high gain inverters (output voltage gain as high as 30 V) have been reported 126, as well as unipolar inverters and multi-stage ring oscillator with single-stage delay of $t = 712$ns [127]. The IEEE Computer Society has already defined standard test procedures for the characterization of organic transistors and materials [128] and organic transistor-based ring oscillators [129]. Moreover novel architectures have emerged, setting the bases for new application fields: among the most promising, transistor-based sensors [130, 131], to detect chemical and biological species, organic RFID (Radio frequency IDentification) systems that can operate at 13.56 MHz at 128 bit [132] and can already reach a bit rate of 50 kbps [133]. Last, but not least, the promising printed electronics [134] has already become a hot research topic for companies [135], and its market, which was estimated by Idtechex in $1.92 billion in 2009, could rise up to $57.16 billion in 2019 [136].

References

1. Heath, J.R., Ratner, M.A.: Phys. Today **56**, 43 (2003)
2. Reimers, J.R., Picconnatto, C.A., Ellenbogen, J.C., Shashidhar, R. (eds.): Ann. N Y Acad. Sci. **1006**, 227 (2003)
3. Tour, J.: Molecular Electronics: Commercial Insights, Chemistry, Devices, Architecture and Programming. World Scientific, Singapore (2003)
4. Cuniberti, G., Fagas, G., Richter, K. (eds.): Lecture Notes in Physics, vol. 680. Springer, Berlin (2005)
5. Heath, J.R., Kuekes, P.J., Snider, G.S., Williams, R.S.: Science **280**, 1716 (1998)
6. Wada, Y., Proc. IEEE **89**, 8 (2001)
7. Tour, J.M., Zandt, W.L., Husband, C.P., Husband, S.M., Wilson, L.S., Franzon, P.D., Nackashi, D.P.: IEEE Trans. Nanotech. **1**, 2 (2002)
8. Macucci, M., Iannacone, G., Bonci, L., Girlanda, M.: IEE Proc. Circuit Dev. Syst. **151**, 5 (2004)
9. Chaudhary, A., Chen, D.Z., Hu, X.S., Niemeier, M.T., Ravichandran, R., Whitton, K.: IEEE Trans. Comp. Aided Des. Int. Cir. Sys. **26**, 11 (2007)
10. Aviram, A., Ratner, M.A.: Chem. Phys. Lett. **29**, 277 (1974)
11. Ellenbogen, J.C., Love, J.C.: Proc. IEEE **88**(3), 386 (2000)
12. Husband, C.P., Husband, S.M., Daniels, J.S., Tour, J.M.: IEEE Trans. Electron Devices **50**(9), 1865 (2003)
13. Skoeldberg, J., Önnheim, C., Wendin, G.: IEEE Trans. Circiuts. Syst. I **54**, 2461 (2007)
14. Kouchi Zhang, G.Q., Graef, M., van Roosmalen, F.: Proceedings IEEE 56th Electr. Comp. Tech. Conf. San Diego, CA, 151–157 (2006)
15. Lent, C.S., Tougaw, P.D., Porod, W., Bernstein, G.: Nanotechnology **4**, 49 (1993)

16. Stan, M.R., Franzon, P.D., Goldstein, S.C., Lach, J.C., Ziegler, M.M.: Proc. IEEE **91**(11), 1940 (2003)
17. Kuekes, P.J., Stewar, D.R., Williams, R.S.: Appl. J. Phys. **97**, 034301 (2005)
18. Walus, K., Jullien, G.A.: Proc. IEEE **94**(6), 1225 (2006)
19. Porod, W., Csaba, G., Csurgay, A.: Proceedings of the 2002 7th IEEE International Workshop on Cellular Neural Networks and their Applications, 2002. Frankfurt/Main CNNA (2002)
20. Burke, P.J., Li, S., Yu, Z.: IEEE Trans. Nanotech. **5**(4) (2006)
21. Strukov, D.B., Likharev, K.K.: Nanotechnology **16**, 888 (2005)
22. Strukov, D.B., Likharev, K.K.: IEEE Trans. Nanotech. **6**(6), 696 (2007)
23. Proceedings of the IEEE (vol. 97, no. 9) on Organic Electronics, September 2009
24. Forrest, R.: IEEE J. Sel. Top. Quant. Elect. **6**(6), 1072 (2000)
25. Haensch, W.: The Drift Diffusion Equation and its Applications in MOSFET Modeling, Springer, Wien (1991)
26. Erlen, C., Lugli, P.: IEEE Trans. Electron Devices **56**, 546 (2009)
27. Klauk, H., Halik, M., Zschieschang, U., Eder, F., Rohde, D., Schmid, G., Dehm, C.: IEEE Trans. Electron Devices **52**(4), 618 (2005)
28. Fix, W., Ullmann, A., Ficker, J., Clemens, W.: Appl. Phys. Lett. **81**(9), 1735 (2002)
29. Brandbyge, M., Mozos, J.L., Ordejon, P., Taylor, J., Stokbro, K.: Phys. Rev. B **65**(165) 401(2002)
30. Datta, S.: Electronic Transport in Mesoscopic Systems. Cambridge University Press, Cambridge, UK (1995)
31. Tian, W., Datta, S., Hong, S., Reifenberger, R., Henderson, J.I., Kubiak, C.P.: J. Chem. Phys. **109**, 2874 (1998)
32. Reed, M.A., Chen, J., Rawlett, A.M., Prince, D.W., Tour, J.M.: Appl. Phys. Lett. **78**, 3735 (2001)
33. Csaba, G., Lugli, P.: IEEE Trans. Nanotech. **8**(3), (May 2009)
34. Chen, Y., Jung, G.-Y., Ohlberg, D.A.A, Li, X., Stewart, D.R., Jeppesen, J.O., Nielsen, K.A., Stoddart, J.F., Williams, R.S.: Nanotech. **14**, 462 (2003)
35. Kushmerick, J.G., Holt, D.B., Pollack, S.K., Ratner, M.A., Yang, J.C., Schull, T.L., Naciri, J., Moore, M.H., Shshidhar, R.: J. Am. Chem. Soc. **124**, 10654 (2002)
36. Reichert, J., Ochs, R., Beckmann, D., Weber, H.B., Mayor, M., Loehneysen, H.v.: Phys. Rev. Lett. **88**, 176804 (2002)
37. Xue, Y., Datta, S., Ratner, M.A.: J. Chem. Phys. **115**, 4292 (2001)
38. Di Ventra, M., Pantelides, S.T., Lang, N.D.: Phys. Rev. Lett. **84**, 979 (2000)
39. Frauenheim, T., Seifert, G., Elstner, M., Niehaus, T., Kohler, C., Amkreutz, M., Sternberg, M., Hajnal, Z., Di Carlo, A., Suhai, S.: J. Phys. Condens. Matter **14**, 3015 (2002)
40. Pecchia, A., Gheorghe, M., Di Carlo, A., Lugli, P., Niehaus, T.A., Frauenheim, Th., Scholz, R.: Phys. Rev. B. **68**, 235321 (2003)
41. Di Carlo, A., Gheorghe, M., Lugli, P., Stenberg, M., Seifert, G., Frauenheim, Th.: Physica B **314**, 86 (2002)
42. Lugli, P., Pecchia, A., Gheorghe, M., Latessa, L., Di Carlo, A.: Semicond. Sci. Technol. **19**, S357 (2004)
43. Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, Th., Suhai, S., Seifert, G.: Phys. Rev. B **58**, 7260 (1998)
44. Porezag, D., Frauenheim, Th., Koehler, Th., Seifert, G., Kaschner, R.: Phys. Rev. B **51**(12), 947 (1995)
45. Frauenheim, T., Seifert, G., Elstner, M., Niehaus, T., Kohler, C., Amkreutz, M., Sternberg, M., Hajnal, Z., Di Carlo, A., Suhai, S.: J. Phys. Condens. Matter. **14**, 3015 (2002)
46. Xue, Y., Datta, S., Ratner, M.: Chem. J. Phys. **281**, 151 (2002)
47. Di Carlo, A., Pecchia, A., Latessa, L., Frauenheim, T., Gotthard Seifert, Introducing Molecular Electronics, Lecture Notes in Phys. 680/2005, 153–184 (2005) DOI: 10.1007/3-540-31514-4_6

48. Pecchia, A., Di Carlo, A., Gagliardi, A., Sanna, S., Frauenheim, T., Gutierrez, R.: Nano Lett. **4**, 2109 (2004)
49. Pecchia, A., Di Carlo, A., Rep. Prog. Phys. **67**, 1497 (2004)
50. Maragakis, P., Barnett, R.L., Kaxiras, E., Elstner, M., Frauenheim, T.: Phys. Rev. B **66**(241) 104 (2002)
51. Porezag, D., Pederson, M.R., Frauenheim, Th., Köhler, Th.: Phys. Rev. B **52**, 14963 (1995)
52. Pecchia, A., Latessa, L., Gagliardi, A., Frauenheim, Th., Di Carlo, A.: Chapter 8 The gDFTB tool for molecular electronics. In: Molecular and Nano Electronics: Analysis, Design and Simulation. Theoretical and Computational Chemistry, Vol. 17, Pages 205–232 (2007). doi:10.1016/S1380-7323(07)80026-4
53. Caroli, C., Combescot, R., Nozieres, P., Saint-James, D.: J. Phys. C Solid State Phys. **5**, 21 (1972)
54. Galperin, M., Ratner, M., Nitzan, A.: Nano Lett. **4**, 1605 (2004)
55. Faleev, S.V., Stockman, M.I.: Phys. Rev. B **66**(085) 318 (2002)
56. Fetter, A.L., Walecka, J.D.: Quantum Theory of Many Particle Systems. Dover, New York, NY (1971)
57. Mahan, D.M.: Many Particle Physics. Plenum Press, New York NY (1981)
58. Haung, H., Jauho, A.P.: Quantum Kinetics in Transport and Optics of Semiconductors, vol. 123. Springer Series in Solid State Science Springer, NY (1996)
59. Schuster, S., Scarpa, G., Latessa, L., Lugli, P.: Phys. Stat. Solid C **5**, 390 (2008)
60. Crljen, Z., Grigoriev, A., Wendin, G., Stokbro, K.: Phys. Rev. B **71**, 165316 (2005)
61. Erlen, C., Lugli, P., Pecchia, A., Di Carlo, A.: Chapter 19 of Nano and Molecular Electronics Handbook. CRC Press (2007)
62. Solomon, G.C., Gagliardi, A., Pecchia, A., Frauenheim, T., Di Carlo, A., Reimers, J.R., Hush, N.S.: J. Chem. Phys. **124**, 094704 (2006)
63. Meir, Y., Wingreen, N.S.: Phys. Rev. Lett. **68**, 2512 (1992)
64. Elbing, M., Ochs, R., Koentopp, M., Fischer, M., von Hänisch, C., Weigend, F., Evers, F., Weber, H.B., Mayor, M.: PNAS **102**(25), 8815–8820 (2005)
65. Oleynik, I.I., Kozhushner, M.A., Posvyanskii, V.S., Yu, L.: Phys. Rev. Lett. **96**, 096803 (2006)
66. Stadler, R., Geskin, V., Cornil, J.: J. Phys. Condens. Matter **20**, 374105 (2008)
67. Tao, N.J.: Nature Nanotechnol. **1**, 173 (2006)
68. Tans, S.J., Verschueren, A.R.M., Dekker, C.: Nature **393** (7 May 1998)
69. Song, H., Kim, Y., Jang, Y.H., Jeong, H., Reed, M.A., Lee, T.: Nature **462**(24/31) (December 2009)
70. Lortscher, E., Ciszek, J.W., Tour, J., Riel, H.: Small **2**(8–9), 973–977 (2006)
71. Chen, Y., Jung, G.-Y., Ohlberg, D.A.A., Li, X., Stewart, D.R., J.O, Jeppesen, K.A, Nielsen, Stoddart, J.F., Stanley Williams, R.: Nanotechnology **14**, 462–468 (2003)
72. Kuekes, P.J., Stewart, D.R., Williams, R.S.: J. Appl. Phys. **97**, 034301 (2005)
73. Stan, M.R., Franzon, P.D., Goldstein, S.C., Lach, J.C., Ziegler, M.M.: Proc. IEEE **91**(11), 1940–1957 (2003)
74. DeHon, A., Goldstein, S.C., Kuekes, P.J., Lincoln, P.: IEEE Trans. Nanotechnol. **4**(2), 215–228 (March 2005)
75. Coker, A., Taylor, V., Bhaduri, D., Shukla, S., Raychowdhury, A., Roy, K.: IEEE Trans. Nanotechnol. **7**, 202–208 (March 2008)
76. Ziegler, M.M., Stan, M.R.: IEEE-NANO 2002. Proceedings of the 2002 2nd IEEE Conference on Nanotechnology, Washington, D.C. 323–327 (2002)
77. Ziegler, M.M., Stan, M.R.: IEEE Trans. Nanotechnol. **2**(4), 217–230 (December 2003)
78. Amsinck, C., Di Spigna, N., Sonkusale, S., Nackashi, D., Franzon, P.: 3rd Workshop on Nonsilicon Computation Munich, Germany (NSC-3,2004)
79. Cerofolini, G.F., Romano, E.: Appl. Phys. A **91**, 181–210 (2008)
80. Cerofolini, G.F., Arena, G., Camalleri, C.M., Galati, C., Reina, S., Renna, L., Mascolo, D.: Nanotechnology **16** 1040–1047 (2005)

81. Cerofolini, G.F., Mascolo, D.: Semicond. Sci. Technol. **21**, 1315–1325 (2006)
82. Vollhardt, K.P.C.: Organische Chemie. VCH-Verlag, Weinheim (1990)
83. Bruetting, W.: Physics of Organic Semiconductors. Wiley-VCH, Weinheim (2005)
84. Hirsch, J.: J. Phys. C: Solid State Phys. **12** (1979)
85. Epstein, A.J., Lee, W.P., Prigodin, V.N.: Synt. Met. **117**, 9–13 (2001)
86. Zaumseil, J., Sirringhaus, H.: IEEE Trans. Semiconduct. Manufact. **14**(3), 281–296 (2001)
87. Pope, M., Swenberg, C.E.: Electronic Processes in Organic Crystals and Polymers. Oxford University Press, Oxford, New York, NY (2002)
88. Baessler, H., Schoenherr, G., Abkowitz, M., Pai, D.M.: Phys. Rev. B **26**(6), 3105–3113 (1982)
89. Bolognesi, A., Di Carlo, A., Lugli, P., Conte, G.: Synth. Met. **138**, 95–100 (2003)
90. Li, L., Meller, G., Kosina, H.: Microel. J. **38**, 4751 (2007)
91. Olivier, Y., Lemaur, V., Brédas, J.L., Cornil, J.: J. Phys. Chem. A **110**, 6356–6364 (2006)
92. Yamashita, Y.: Sci. Technol. Adv. Mater. **10**, 024313 (2009)
93. Alam, M.A., Dodabalapur, A., Pinto, M.R.: IEEE Trans. Electron Devices **44**(8), 1332–1337 (1997)
94. Horowitz, G., Hajlaoui, R., Bouchriha, H., Bourguiga, R., Hajlaoui, M.: Adv. Mat. **10**(12), 923–927 (1998)
95. Locci, S., Morana, M., Orgiu, E., Bonfiglio, A., Lugli, P.: IEEE Trans. Electr. Dev. **55**(10), (October 2008)
96. Stallinga, P., Gomes, H.L., Biscarini, F., Murgia, M., deLeeuw, D.M.: J. Appl. Phys. **96**(9), 5277–5283 (2004)
97. Stallinga, P., Gomes, H.L.: Synt. Met. **156**, 1316–1326 (2006)
98. Horowitz, G., Hajlaoui, R., Kouki, F.: Europ. Phys. J. Appl. Phys. **1**, 361–367 (1998)
99. Horowitz, G., Hajlaoui, M.E., Hajlaoui, R.: J. Appl. Phys. **87**(9), 4456–4463 (2000)
100. Vissenberg, M.C.J.M., Matters, M.: Phys. Rev. B **57**(20), 964–967 (1998)
101. Frenkel, J.: Phys. Rev. **54**, 647–648 (1938)
102. Dunlap, D.H., Parris, P.E., Kenkre, V.M.: Phys. Rev. Lett. **77**(3), 542–545 (1996)
103. Schein, L.B., Peled, A., Glatz, D.: J. Appl. Phys. **66**(2), 686–692 (1989)
104. Novikov, S.V., Dunlap, D.H., Kenkre, V.M., Parris, P.E., Vannikov, A.V.: Phys. Rev. Lett. **81**(20), 4472–4475 (1998)
105. Wang, L., Fine, D., Basu, D., Dodabalapur, A.: J. Appl. Phys. **101**(054515), 1–8 (2007)
106. Hamadani, B.H., Natelson, D.: J. Appl. Phys. **95**(3), 1227–1232 (2007)
107. Hamadani, B.H., Richter, C.A., Gundlach, D.J., Kline, R.J., McCulloch, I., Heeney, M.: J. Appl. Phys. **102**(044503), 1–7 (2007)
108. Snowden, C.M., Snowden, E.: Introduction to Semiconductor Device Modelling. World Scientific, Singapore (1987)
109. Schmechel, R., von Seggern, H.: Phys. Stat. Sol. (a) **201**(6), 1215–1235 (2004)
110. Yang, Y.S., Kim, S.H., Lee, J.I., Chu, H.Y., Do, L.M., Lee, H., Oh, J., Zyung, T., Ryu, M.K., Jang, M.S.: Appl. Phys. Lett. **80**(9), 1595–1597 (2002)
111. Kang, J.H., da Silva Filho, D., Bredas, J.L., Zhu, X.Y.: Appl. Phys. Lett. **86**(152115), 1–3 (2005)
112. Northrup, J.E., Chabinyc, M.L.: Phys. Rev. B **68**(041202), 1–4 (2003)
113. Goldmann, C., Gundlach, D.J., Batlogg, B.: Appl. Phys. Lett. **88**(063501), 1–3 (2006)
114. Schockley, W., Read, W.T.: Phys. Rev. **87**(5), 835–842 (1952)
115. Hall, R.N.: Phys. Rev. **87**, 387 (1952)
116. Singh, Th.B., Marjanovi`c, N., Stadler, P., Auinger, M., Matt, G.J., Günes, S., Sariciftci, N.S., Schwödiauer, R., Bauer, S.: J. Appl. Phys. **97**(083714), 1–5 (2005)
117. Gu, G., Kane, M.G., Doty, J.E., Firester, A.H.: Appl. Phys. Lett. **87**(243512), 1–3 (2005)
118. Li, L., Meller, G., Kosina, H.: Solid State Electron. **51**, 445–448 (2007)
119. Lindner, Th., Paasch, G., Scheinert, S.: J. Appl. Phys. **98**(114505), 1–9 (2005)

120. Paasch, G., Scheinert, S., Herasimovich, A., Hörselmann, I., Lindner, Th.: Phys. Stat. Sol. (a) **3**, 534–548 (2008)
121. Ahles, M., Schmechel, R., von Seggern, H.: Appl. Phys. Lett. **85**, 4489 (2004)
122. Erlen, C., Lugli, P., Fiebig, M., Schiefer, S., Nickel, B.: J. Comput. Electron. **5**(4) (December 2006)
123. Berliocchi, M., Manenti, M., Bolognesi, A., Di Carlo, A., Lugli, P., Paolesse, R., Mandoy, F., Di Natale, C., Proietti, E., Petrocco, G., D'Amico, A.: Semicond. Sci. Technol. **19**, 354, (2004).
124. Baude, P.F., Ender, D.A., Haase, M.A., Kelley, T.W., Muyres, D.V., Theiss, S.D.: Appl. Phys. Lett. **82**, 3964 (2003).
125. Hoon Han, S., Cho, S.M., Kim, J.H., Won Choi, J., Jang, J., Hwan Oh, M.: Appl. Phys. Lett. **89**(093504) (2006).
126. Kim, F.S., Guo, X., Watson, M.D., Jenekhe, S.A.: Adv. Mater. **21**, 15 (2009)
127. Smith, J., Hamilton, R., Heeney, M., de Leeuw, D.M., Cantatore, E., Anthony, J.E., McCulloch, I., Bradley, D.D.C., Anthopoulos, T.D.: Appl. Phys. Lett. **93**(253301) (2008)
128. IEEE Standard for Test Methods for the Characterization of Organic Transistors and Materials, IEEE Std 1620–2008. http://ieeexplore.ieee.org/servlet/opac?punumber=4695965 (2008)
129. IEEE Standard for Test Methods for the Characterization of Organic Transistor-Based Ring Oscillators, IEEE Std 1620.1–2006. http://ieeexplore.ieee.org/servlet/opac?punumber=4014407 (2006)
130. Mabeck, J., Malliaras, G.: Anal. Bioanal. Chem. **384**(34353) (2006)
131. Roberts, M.E., Mannsfeld, S.C.B., Queralt o, N., Reese, C., Locklin, J., Knoll, W., Bao, Z.N.: Proc. Natl. Acad. Sci. USA **105**(1213439) (2008)
132. Myny, K., Beenhakkers, M.J., van Aerle, N.A.J.M., Gelinck, G.H., Genoe, J., Dehaene, W., Heremans, P.: IEEE ISSCC Dig. Tech. Papers **52** (2009)
133. Elfrink, R., Pop, V., Hohlfeld, D., Kamel, T., Matova, S., de Nooijer, C., Jambunathan, M., Goedbloed, M., Caballero Guindo, L., Renaud, M., Penders, J. and van Schaijk, R.: First autonomous wireless sensor node powered by a vacuum-packaged piezoelectric MEMS energy harvester. In: IEEE International Electron Devices Meeting—IEDM, pp. 543–546. Baltimore, MD, USA, 7–9 Dec 2009
134. Gamota, D.R., Brazis, P., Kalyanasundaram, K., Zhang, J.: Printed Organic and Molecular Electronics, Kluwer Academic Publishers, New York (2004)
135. International Electronics Manufacturing Initiative (iNEMI) consortium. http://thor.inemi.org/webdownload/RI/2009_Research_Priorities.pdf (2009)
136. Das, R., Harrop, P.: Printed and Thin Film Transistors and Memory 2009–2029, Updated in Q3 2010. IDTechEx, United Kingdom. http://www.idtechex.com/research/reports/printed_and_thin_film_transistors_and_memory_2009_2029_000221.asp?viewopt=showall (2010)

Chapter 2
Three-Dimensional Silicon–Germanium Nanostructures for CMOS-Compatible Light Emitters

D.J. Lockwood and L. Tsybeskov

Abstract The present status of light emitters based on SiGe nanostructures is reviewed. To be commercially valuable, these light emitters should be efficient, fast, operational at room temperature, and, perhaps most importantly, compatible with the "main stream" CMOS technology. Another important requirement is in the emission wavelength, which should match the optical waveguide low-loss spectral region, i.e., 1.3–1.6 μm. Among other approaches, epitaxially grown Si/SiGe quantum wells and quantum dot/quantum well complexes produce efficient photoluminescence (PL) and electroluminescence (EL) in the required spectral range. Until recently, the major roadblocks for practical applications of these devices were strong thermal quenching of the luminescence quantum efficiency and a long carrier radiative lifetime. The latest progress in the understanding of physics of carrier recombination in Si/SiGe nanostructures is reviewed, and a new route toward CMOS-compatible light emitters for on-chip optical interconnects is proposed.

Introduction

Optical interconnects in the form of fiber optics have been used for many years in different long-distance communication applications [1, 2]. With the microprocessor clock speed approaching 10 Gbps, optical interconnects are now being considered for on-chip interconnects as an alternative to metal wires with their unavoidable RC delay, significant signal degradation, high power dissipation, and electromagnetic interference [2–6]. Two major avenues toward optical interconnects on a chip include a hybrid approach with densely packaged III–V optoelectronic components [7–16] and the all-group-IV approach (mainly Si, Ge, and SiGe), where all the major components, e.g., light emitters, modulators, waveguides, and photodetectors, are monolithically integrated into the CMOS environment [17–20]. For almost five

D.J. Lockwood (✉)
Institute for Microstructural Sciences, National Research Council Canada, Ottawa, Ontario K1A 0R6, Canada
e-mail: david.lockwood@nrc-cnrc.gc.ca

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_2,

decades, the efforts for obtaining light emission from group IV semiconductors have been mainly focused on porous silicon [21–24], silicon/silicon dioxide superlattices [25–29], silicon nanoprecipitates in silicon dioxide [24, 30], erbium in silicon [31], silicon/germanium quantum wells [32–34], and, more recently, iron disilicide [35]. However, no approach has so far been applied commercially. There are several reasons, including the lack of a genuine or perceived compatibility with conventional CMOS technology, the long carrier radiative lifetime in Si-based nanostructures (NSs), and, especially in the case of near-infrared emitters, the significant thermal quenching of the luminescence quantum efficiency [10].

Light Emission in Crystalline Silicon and Germanium

Single-crystal Si (c-Si) and Ge are both indirect band gap semiconductors, and, in their pure state, direct electron–hole recombination without momentum conservation is forbidden (Fig. 2.1). To conserve momentum, an interaction with a phonon of appropriate wave vector is needed, and in low-temperature c-Si photoluminescence (PL) spectra one can see a complete set of characteristic Si phonons (Fig. 2.1). This type of electron–hole recombination is known as a phonon-assisted process [36–40]. This phonon-assisted, slow ($\geq 10^{-3}$ s) radiative recombination competes with a variety of much faster non-radiative carrier recombination mechanisms, including recombination via structural defects, impurities, surface states, and Auger recombination. Also, the low free exciton binding energy of ~15 meV is responsible for a fast decrease in the PL quantum efficiency as temperature increases. Despite this unfavorable situation, low-temperature luminescence in c-Si could be quite efficient, because there are several possibilities to effectively suppress non-radiative recombination. For example, epitaxially grown c-Si has an extraordinarily low concentration

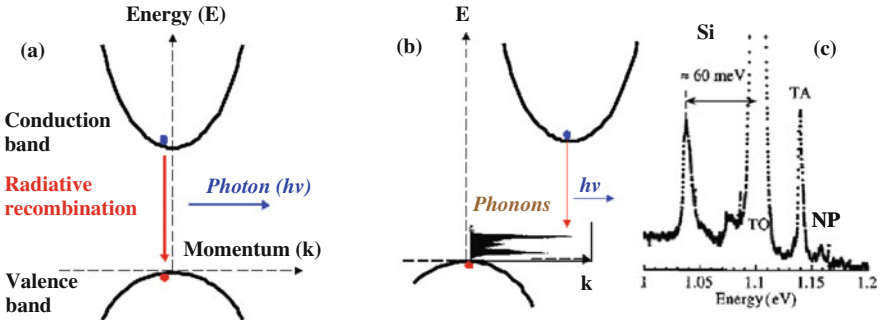


Fig. 2.1 Schematic of electron–hole recombination in (**a**) a direct band gap semiconductor and (**b**) an indirect band gap semiconductor with phonon assistance (the Si phonon spectrum is used as an example). (**c**) Low-temperature Si PL spectrum with the characteristic phonon sidebands indicated: NP—zero-phonon (also called the no-phonon (NP)) or direct exciton recombination, TA—transverse acoustic phonon, and TO—transverse optical phonon. The second replica of the TO phonon is separated from the main TO PL peak by ~60 meV

of unwanted impurities and structural defects (~10^{12} cm^{-3}). By decreasing the temperature and forming excitons, especially excitons bound to impurities, carrier diffusion is suppressed and the probability for a carrier to find a non-radiative recombination center is strongly reduced. Therefore, high-quality Si samples at low temperature might have a luminescence quantum efficiency close to 0.1% or even better [40–42]. A possibility to achieve a higher luminescence quantum efficiency is to passivate the c-Si surface with hydrogen, which leads to a strong reduction of the surface state density [43]. A similar effect can be obtained by electrochemical etching and formation of small-size Si crystallites with nearly perfectly hydrogen-passivated surfaces, as has been shown for anodically etched porous Si [21, 22]. It is also important to remember that the refractive index of c-Si is relatively high ($\geq$3.5), and some appropriate sample preparation is necessary to extract light with low losses [44]. One of the possibilities is to use a photonic crystal fabricated using a submicron-thick Si-on-insulator (SOI) layer; it has been shown that the luminescence external quantum efficiency increases 70 times compared to that in bulk, ~100 μm thick surface polished c-Si samples [45]. Thus, despite the fact that the radiative lifetime in c-Si remains long, different sample preparation techniques could result in a dramatic difference in the PL external quantum efficiency.

One quite different approach is to increase the internal luminescence quantum efficiency: mainly by relaxing the carrier recombination selection rules and accelerating the radiative recombination rate. In principle, this can be done by several methods: by changing the Si crystal structure (high pressure, strain, heavy doping, etc.) [40, 46]; by using radiative impurities such as rare earth, deep-level isoelectronic complexes (see Fig. 2.2 and [31, 46]); by forming extremely small Si

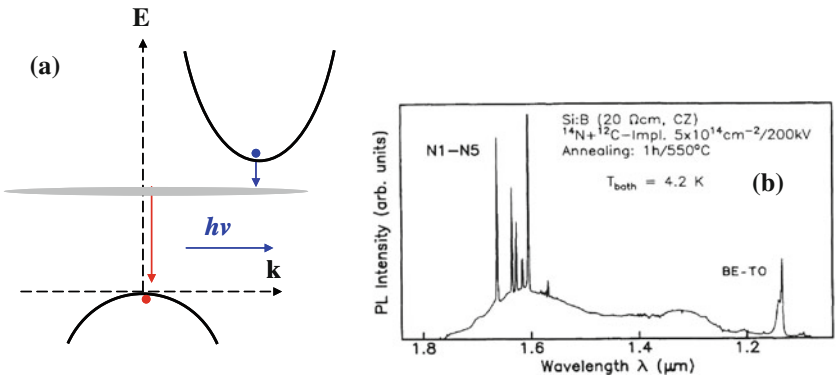


Fig. 2.2 (**a**) Schematic of carrier recombination in an indirect band gap semiconductor involving a deep energy level associated with an impurity. Note that for a localized carrier (e.g., an electron at an isoelectronic trap), the momentum is not well defined and it is stretched over the entire Brillouin zone; thus, electron–hole recombination no longer needs phonon assistance. (**b**) Low-temperature PL spectra in c-Si with nitrogen–carbon impurity complexes responsible for deep-energy levels and sharp PL features around 1.6 μm. (Reprinted with permission from [46], Copyright 1989, Elsevier)

crystallites comprised of just a few atoms and exploiting the quantum confinement effect [47–49]. This work has received much attention; it is well covered in [17, 18] and will not be discussed here. An alternative approach is to create carrier recombination conditions involving strong electron–electron interactions. Compared to the other approaches, the latter one has received relatively little attention, and it will be briefly presented next.

The most well-known feature of strong carrier–carrier interactions in group IV semiconductors is the electron–hole condensation and the formation of electron–hole droplets (EHDs) [50]. This process is mainly possible due to a long carrier radiative lifetime in pure Si and Ge crystals, but it also requires a combination of low temperature and high carrier concentration. At low temperature, electrons and holes quickly form free excitons (FE). The process of EHD condensation from the electron–hole FE gas into the electron–hole liquid (EHL) is a type I phase transition [50–53] (Fig. 2.3). In practice, a high level of excitation often produces an electron–hole plasma (EHP), a thermodynamic phase where the carrier–carrier interaction becomes comparable with the electron–lattice interaction. Thus, the condensation of the EHL and formation of EHDs could be achieved from both thermodynamic states (Fig. 2.3). The EHD luminescence is quite efficient, despite the fact that it remains a phonon-assisted process with typical phonon features in the PL spectra (Fig. 2.3). The phonon features are shifted to lower photon energies (due to additional carrier–carrier interactions), strongly broadened, and the measured PL transient characteristics indicate a carrier lifetime as short as ~10^{-8} s [54–56]. The details of the recombination mechanisms will not be discussed here, but a qualitative model suggests that the carrier–carrier interaction within EHDs is responsible for partial "whitening" of the phonon spectra (i.e., c-Si lattice field screening), and it produces a relaxation of selection rules and a shortening of the carrier radiative lifetime [56]. One of the limiting factors of the EHD condensation is exciton diffusion, and it has been shown that in ultra-thin (<100 nm) layers of SOI, EHD

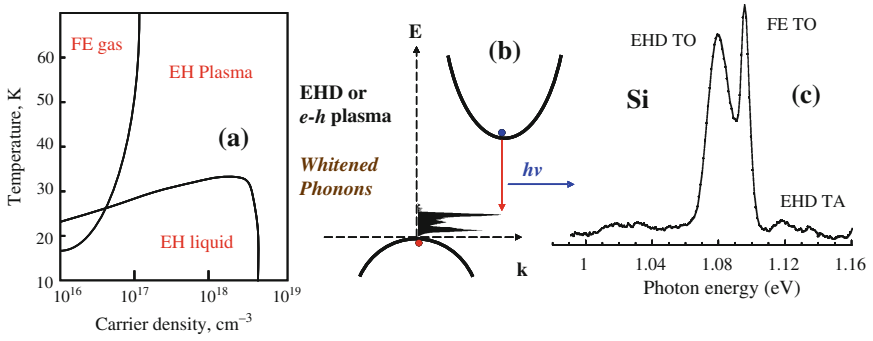


Fig. 2.3 (**a**) Phase diagram showing transitions from electron–hole (EH) free exciton (FE) gas to EH liquid and EH plasma in c-Si. (**b**) Phonon-assisted electron–hole recombination is shown with partial whitening of the phonon dispersion. (**c**) The PL spectrum in c-Si at 10 K and carrier density $\geq 10^{16}$ cm^3 showing coexisting FE and electron–hole droplet (EHD) PL peaks

condensation occurs at a lower carrier concentration compared to that in bulk c-Si [57]. Also, it has been shown that at a temperature higher than the thermodynamic limit for electron–hole condensation, an EHP can be produced instead of EHDs [50, 53]. EHP radiative recombination in c-Si is also much faster ($\tau_{\mathrm{EHP}} \approx 10^{-7}$ s) compared to low-excitation electron–hole recombination ($\tau_{\mathrm{e-h}} \sim 10^{-3}$ s), but at this level of excitation it competes with Auger recombination. Auger recombination has a lifetime $\tau_{\mathrm{Auger}} = 1/Cn^2$, where n is the carrier concentration, and for bulk silicon $C \approx 10^{-30(-31)}$ for electrons (holes), respectively [58]. Thus, at a carrier concentration approaching the EHP regime (~10^{18}–10^{19} cm^{-3}), Auger recombination quickly becomes the dominant recombination channel.

The majority of text books consider Auger recombination as a strictly non-radiative process where the recombining electron and hole pair (e–h) transfers its energy and momentum to a third particle, which could be either an electron (e) or a hole (h). Thus, Auger processes could be divided into two major categories: (i) two electrons–one hole (e–h, e) with generation of a "hot" electron in the conduction energy band and (ii) two holes–one electron (e–h, h) with generation of a "hot" hole in the valence energy band. Depending on the energy band structure, these hot carriers can dissipate their energy by generating multiple phonons (i.e. heat generation), or they could be transferred into another energy band inside the Brillouin zone. In c-Si, the bottom of the conduction energy band is in the X-valley. The second lowest conduction energy band is on the other side of the Brillouin zone in the L-valley, and it is separated from the bottom of the conduction energy band by an energy of 0.89 eV, which is comparable with the energy gap $E_G = 1.1$ eV at room temperature

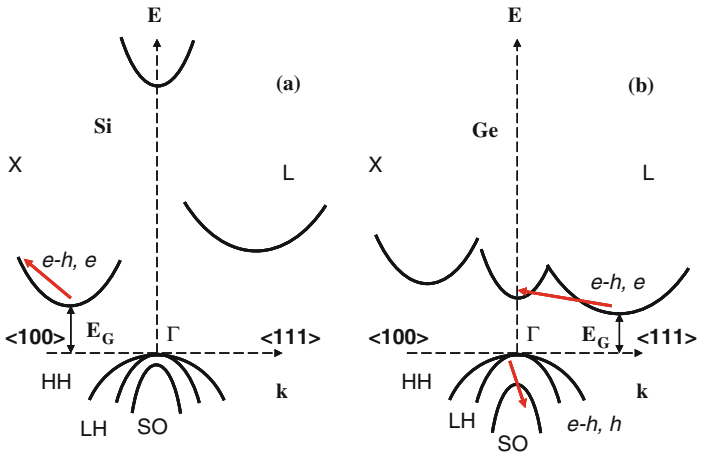


Fig. 2.4 Schematic of the energy band structure for Γ–X (in <111> direction) and Γ–L (in <100> direction) vicinities in (**a**) c-Si and (**b**) c-Ge with possible Auger-mediated electron (**e**) and hole (**h**) inter-valley transitions. Heavy-hole (HH), light-hole (LH), and split-off (SO) valence bands are shown and the fundamental energy gap is denoted by E_G

(Fig. 2.4a). At the Γ-point (i.e., the center of the Brillouin zone), the c-Si conduction band minimum is 4.2 eV higher than the maximum of the valence energy band. This conduction energy band structure makes electron transfer between different valleys very unlikely. However, hot electrons in c-Si produced by high electric field in a reverse bias p–n junction or a metal-oxide-semiconductor (MOS) device have a wide range of energies and momenta; they might recombine with holes and produce "white" light with photon energies ranging over 1.1 eV $\leq h\nu <$ 4.2 eV but with a low quantum efficiency [42, 59].

Single-crystal Ge (c-Ge) is also an indirect band gap semiconductor, but with a very different energy band structure (Fig. 2.4b). The indirect energy gap is $E_G =$ 0.66 eV (at temperature $T = 300$ K) with the conduction energy band minimum in the L-valley, but the direct energy gap at the Γ-point is $E_G^{\Gamma} = 0.8$ eV with an energy separation between the two neighboring conduction bands of only 0.14 eV (Fig. 2.4b). Thus, the probability of electron transfer from indirect (L) to direct (Γ) band is much higher compared to that in c-Si. Another important feature of the c-Ge energy band structure is the energy separation of 0.29 eV between the Γ-point and split-off (SO) band (Fig. 2.4b). Therefore, a variety of the carrier transfer processes between different energy valleys in both (conduction and valence) energy bands are possible [60, 61]. One of them is the electric field induced electron transfer from indirect to direct energy band—the L–Γ process, which was first discovered by Gunn in n-type Ge [62]. In p-type Ge, the Auger process could mediate hole transfer from the Γ-point to the bottom of the SO band [63] (Fig. 2.4b). Carrier transfer in c-Ge followed by radiative electron–hole recombination produces light emission in the desired spectral region of 1.2–1.6 μm (Fig. 2.5). Unfortunately, the c-Ge conduction energy band near the Γ-point has a density of states approximately 10 times lower compared to that in the L-valley, and fast electron back transfer suppresses the direct band electron–hole radiative recombination [64, 65]. Several possibilities to reduce the back transfer, including the use of degenerate n-type c-Ge and strain have been discussed, but none of them has so far produced efficient light emission associated with c-Ge direct band gap radiative electron–hole recombination.

Light Emission in SiGe Bulk Alloys and Composition-Controlled Quantum Wells

Since both c-Si and c-Ge are indirect band gap semiconductors, it is a common assumption that SiGe crystalline alloys should also exhibit properties of an indirect band gap semiconductor. Actually, the entire concept of *k*-space is applicable to a crystal lattice with an atomically perfect long range order. SiGe alloys are crystalline material but with strong compositional disorder (i.e., the absence of long range order in the position of Si and Ge atoms within a SiGe crystal lattice). As a result of this compositional disorder, carrier recombination selection rules in SiGe alloys are partially relaxed. This conclusion is supported by Fig. 2.6, where instead of a totally suppressed zero-phonon transition, which is normally observed in low-temperature PL spectra of high-quality single-crystal Si and Ge, SiGe bulk alloys

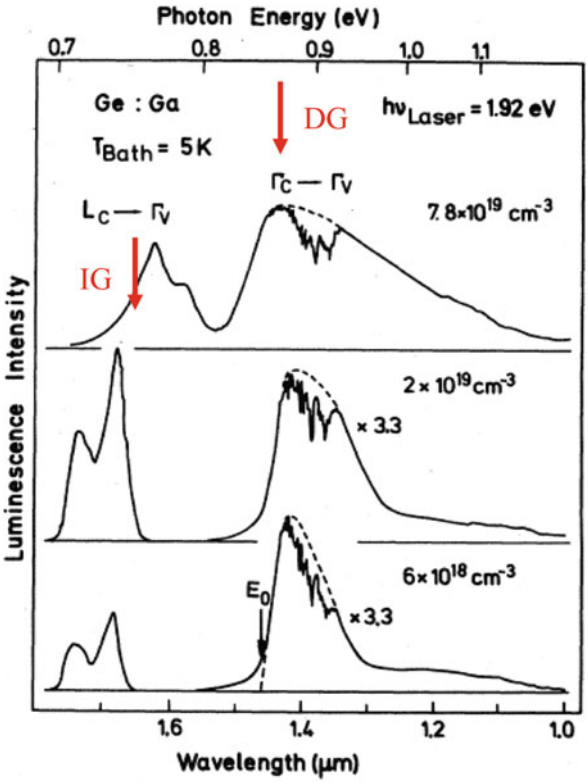


Fig. 2.5 Low-temperature PL spectra in p-type Ge in the vicinity of fundamental, indirect energy gap (IG), direct gap (DG), and above-the-gap electronic transitions. (Reprinted with permission from [63], Copyright 1984, American Physical Society)

exhibit PL spectra showing both phonon-assisted and zero-phonon radiative transitions in the spectral region of 1.3–1.6 μm (see [66]). The relaxation of carrier recombination selection rules in SiGe alloys produces PL that is faster compared to that in Si but still orders of magnitude slower compared to the recombination time in direct band gap semiconductors. As the temperature increases, the PL intensity quickly decreases with a thermal quenching activation energy of ~10–15 meV, which has been attributed to the binding energy of an exciton localized on SiGe stoichiometry fluctuations [66–68]. Therefore, this type of luminescence in SiGe bulk crystalline alloys has little practical importance.

Epitaxial growth of pure Ge on Si is complicated by the 4.2% lattice mismatch (see, for example, a review article [69]). To avoid formation of structural defects (mainly dislocations) with a large number of non-radiative recombination centers, controlled-composition $Si_{1-x}Ge_x$ alloys with $0.1 < x < 0.2$ can be grown by using techniques such as molecular beam epitaxy (MBE) or ultra-high vacuum (UHV) chemical vapor deposition (CVD). These heterostructures exhibit a low defect density at interfaces, and $Si/Si_{1-x}Ge_x$ multilayers with $0.1 < x < 0.2$ were intensively studied as a possible method to confine electron–hole pairs in double heterojunction or quantum well (QW) configurations and to reduce the luminescence thermal

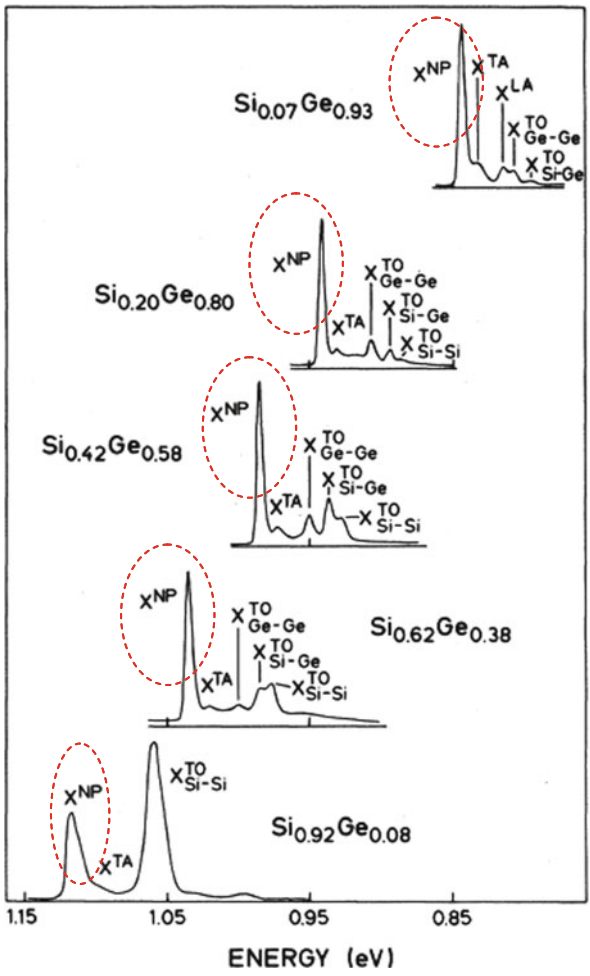


Fig. 2.6 Low-temperature PL spectra in bulk $Si_{1-x}Ge_x$ alloys for a wide range of compositions x. (Reprinted with permission from [66], Copyright 1989, American Physical Society)

quenching. An intense luminescence, but still restricted to low temperature, has been demonstrated in $Si/Si_{1-x}Ge_x$ quantum wells and superlattices with $x \leq 0.2$ (see Fig. 2.7 and also [70]). Several studies have been specifically focused on the determination of the energy band alignment (i.e., type I or type II, see Fig. 2.8) in these nanostructures (NSs), but so far no final conclusion has been reached [71–73]. It has also been shown that SiGe interdiffusion during growth of $Si/Si_{1-x}Ge_x$ QWs can be minimized using a relatively low growth temperature (<550°C). On the other hand, SiGe epitaxial growth at temperatures below 500°C usually produces a significant density of dislocations [73–75]. Thus, high-quality $Si/Si_{1-x}Ge_x$ QWs with $x < 0.2$ can be grown using only a relatively narrow processing growth "window," but again the PL quantum efficiency thermal quenching remains the unresolved problem for practical light emitters at 1.3–1.5 μm.

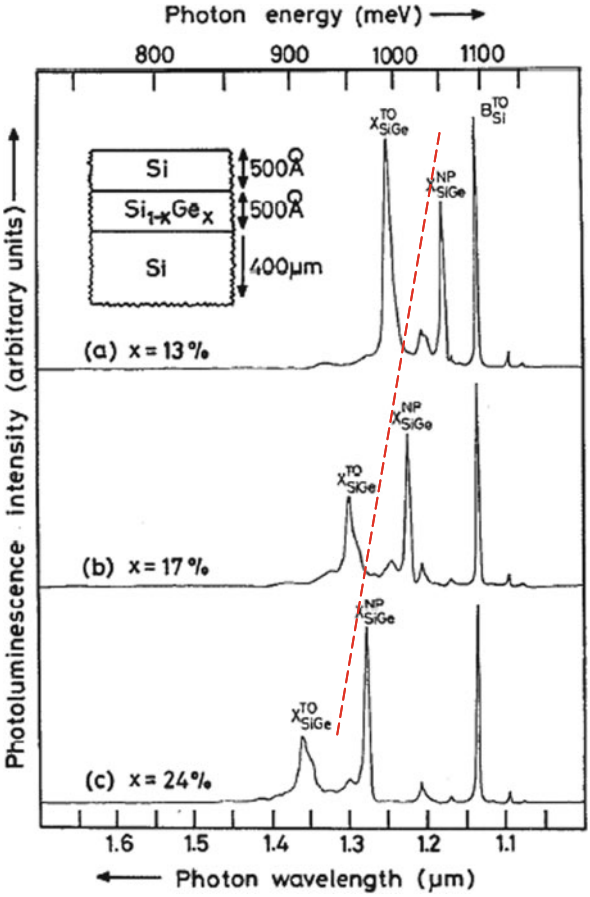


Fig. 2.7 Low-temperature PL spectra in $Si/Si_{1-x}Ge_x/Si$ single quantum wells (see *inset*) with composition $0.1 < x < 0.25$. (Reprinted with permission from [70], Copyright 1992, American Institute of Physics)

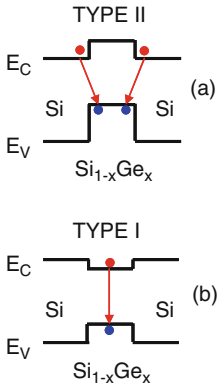


Fig. 2.8 Schematic representation of (**a**) type I and (**b**) type II energy band alignments supposed to exist in Si/SiGe NSs. E_C and E_V are the energies of the conduction and valence bands, respectively, in Si and $Si_{1-x}Ge_x$

Properties of SiGe Three-Dimensional Nanostructures

By the 1990s, a different form of SiGe NS, namely the three-dimensional (3D) self-assembled system produced by the well-known Stranski-Krastanov, or cluster layer, growth mode in lattice-mismatched materials, had been demonstrated [76–80]. It has been shown that dislocation-free SiGe growth can be achieved using a higher temperature (≥600°C), and that the non-planar geometry is mainly responsible for the significant increase of the SiGe critical layer thickness. It has also been found that, compared to two-dimensional (2D) Si/SiGe NSs, the PL and electroluminescence (EL) quantum efficiency in 3D Si/SiGe NSs is higher (up to 1%), especially for $T > 50$ K [81–84]. Despite many successful demonstrations of PL and EL in the spectral range of 1.3–1.6 μm, which is important for optical fiber communications, the proposed further development of 3D Si/SiGe-based light emitters was discouraged by several studies indicating a type II energy band alignment at Si/SiGe heterointerfaces [72, 85–87], where the spatial separation of electrons (located in Si) and holes (localized in SiGe) (see Fig. 2.8) was thought to make carrier radiative recombination very inefficient. Later, it was also shown that 3D Si/SiGe NSs exhibit an extremely long (of the order of 10^{-2} s) luminescence lifetime [88], which is of the order of a million times longer than in III–V semiconductors and their NSs. Thus, according to this analysis, 3D Si/SiGe NSs cannot be used to achieve efficient and commercially valuable light-emitting devices. Recently, however, it has been shown that despite the fact that the Si/SiGe heterointerface most likely exhibits type II energy band alignment, it is still possible to obtain conditions favorable for an efficient carrier radiative recombination. The rest of this review presents the revised understanding of basic physics in such systems and demonstrates that the radiative carrier recombination lifetime can successfully be reduced from 10^{-2}s to 10^{-7} s, which is only ~10 times longer than that found in direct band gap III–V semiconductors.

Growth and Structural Properties of Si/SiGe Three-Dimensional Nanostructures

The standard fabrication of Si/SiGe NSs is in the form of a Si/SiGe multilayer structure, which could be a planar layer–layer or 2D structure, or a layer–cluster system which by the definition is a 3D NS. Both types of fabrications are based on the sequential physical sputtering of Si and Ge (SiGe) in MBE or thermal decomposition of SiH_4 and GeH_4 in CVD, and it takes place at a temperature in the range $T = 500 - 650$°C. At those temperatures, both the high Ge solid solubility in Si and strain-induced SiGe interdiffusion due to the ~4% lattice mismatch between Si and Ge are important. The MBE growth provides a better control over the average SiGe cluster composition, although, because of interdiffusion during growth, the composition is not uniform within the cluster volume [89–92]. Figure 2.9 shows cross-sectional transmission electron micrographs (TEM) and top view scanning electron micrographs of Si/SiGe NSs formed by MBE at different temperatures and

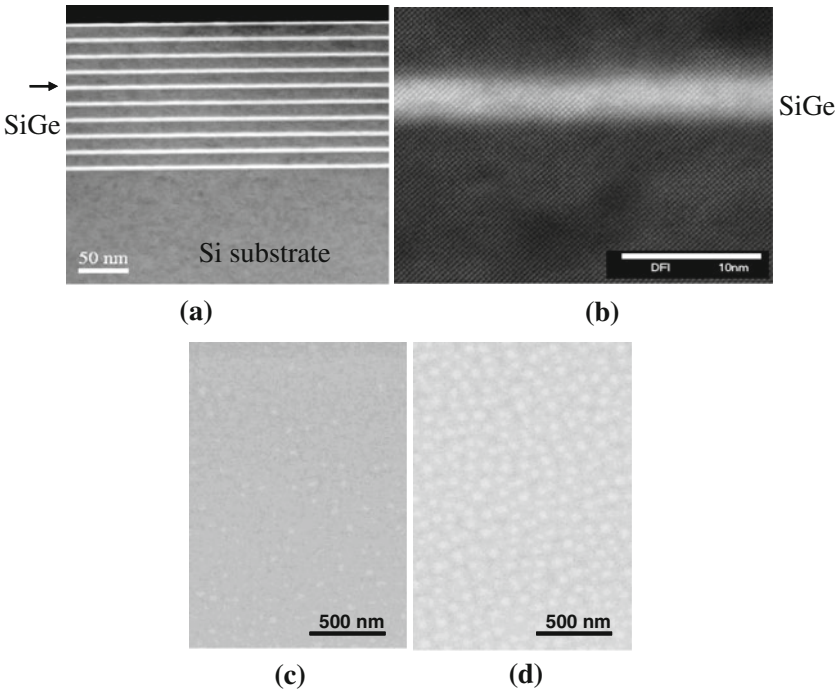


Fig. 2.9 (**a**) Cross-sectional TEM showing MBE grown Si/SiGe multilayer 2D NSs with a Ge composition close to 30% and grown at $T = 550°C$ and (**b**) a cross-sectional TEM micrograph focusing on a single SiGe layer showing its epitaxial nature and some "waviness" of the upper SiGe/Si interface. A larger scale SEM top view of a Si/SiGe single layer with a Ge composition close to 50% and grown at temperatures of (**c**) 570°C and (**d**) 600°C showing a clearly observable increase in density of SiGe clusters at higher temperature

depicts the transition from 2D (layer) to 3D (a cluster) growth modes. In general, SiGe cluster growth at temperatures $T_G \geq 550°C$ commences with the spontaneous development of a $Si_{1-x}Ge_x$ planar, <1 nm thick, wetting layer where x varies, mainly due to uncontrolled SiGe interdiffusion. With the further influx of Ge and Si, the growth mode then switches from 2D to 3D, and this helps release some of the lattice mismatch-induced strain [92, 93]. The fully grown, 3–10 nm high, and initially nearly pyramidal-shaped SiGe clusters (Fig. 2.10a) have a Ge-rich (~50% depending on the Ge flux) core, although the exact final cluster shape and composition strongly depend on the fabrication conditions, mainly deposition temperature and time [89–91]. If the SiGe cluster is covered by a Si cap, the initial pyramid-like cluster top is smoothed out, and the capping Si layer is locally strained, mostly near the top of the SiGe cluster. In multilayer Si/SiGe cluster samples with a thin (<10 nm) Si layer between SiGe cluster layers, this strain field propagates in the growth direction, and it induces vertical SiGe cluster self-ordering (Fig. 2.10b). In these $Si_{1-x}Ge_x$ clusters, the chemical composition x has been recently studied using the technique of analytical TEM (see Fig. 2.11 and also [89]). Detailed structural analysis also

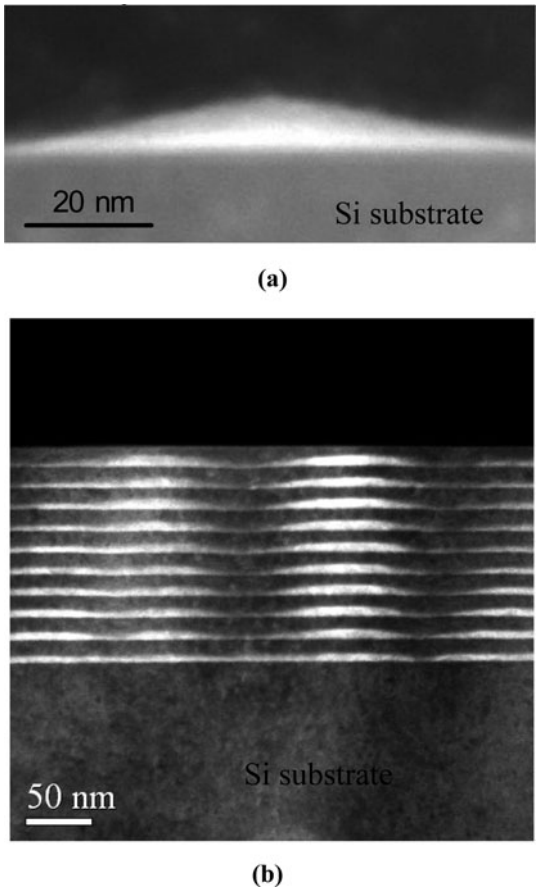


Fig. 2.10 (**a**) Cross-sectional TEM of a pyramidal-shaped SiGe cluster grown by MBE on Si without a Si capping layer. (**b**) Cross-sectional TEM of a multilayer 3D Si/SiGe NS clearly showing cluster vertical self-alignment. The growth was performed at 650°C

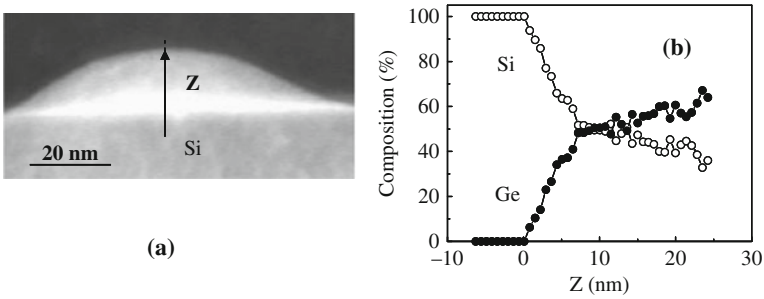


Fig. 2.11 (**a**) TEM micrograph showing the vertical direction of the compositional analysis of a SiGe cluster and (**b**) the distribution of Si/Ge atomic composition in that cluster. The Z direction is the direction of growth and $Z = 0$ corresponds to c-Si substrate surface

indicates that Si in the valleys between SiGe clusters is slightly compressed [49]. An increase in SiGe cluster height results in an increase of their lateral size and change of the cluster shape known as the pyramid–dome transition [79, 80, 89, 94]. These SiGe dome-shaped clusters have multiple facets (Fig. 2.11) and, therefore, might have some structural defects at SiGe/Si interfaces. To summarize, buried SiGe clusters with the highest Ge composition of near 50% in the middle of the clusters and typical cluster height of <10 nm are surrounded by Si, which is tensile strained above each cluster and compressed laterally between clusters to maintain a low overall strain. Each SiGe cluster consists of $Si_{1-x}Ge_x$ crystalline alloys with x increasing toward the cluster center. Thus, despite being fully crystalline materials, 3D Si/SiGe NSs with a high (~50%) Ge atomic concentration exhibit significant embedded strain and compositional disorder.

Further insight into the structural properties of the various SiGe multilayers grown by MBE discussed here is revealed by X-ray and Raman scattering experiments. Figure 2.12 compares the (004) rocking curves from three $Si_{1-x}Ge_x/Si$ multilayers with $0.09 < x < 0.55$. The diffraction data from the sample with the lower Ge composition is very well reproduced by the calculation using the structural parameters of Table 2.1. The agreement for the sample of intermediate composition is also good, but a damping of the higher order satellites (near $2\theta = 66°$) indicates a loss of coherence of the superperiodicity due to either thickness irregularities or loss of interface sharpness. No good simulation of the data could be obtained for samples with high Ge composition ($x > 0.5$). Here the satellites remain relatively sharp, but they are broadened at the base and the usual thickness fringes between the peaks are not observed. The result shown here is typical of superlattices exhibiting undulated interface morphology such as that depicted in Fig. 2.10b [80]. Although the sample shown here has a nominal Ge composition of 0.53, a better fit was obtained using a lower Ge composition; here, in Fig. 2.12c, a value of $x = 0.42$ was used. Although the conventional diffraction theory of layered media cannot accurately model structures with non-planar interfaces, the lower Ge composition found here points to Si–Ge atomic exchange during the island formation. This aspect has been discussed in detail elsewhere [90, 91].

Figure 2.13 presents the specular reflectivity curves measured on the same samples as in Fig 2.12. As for the rocking curves, reflectivity curves exhibit satellites arising from the periodic nature of the multilayers. Here, however, the contrast is primarily caused by the modulation of the atomic number in the growth direction rather than by the strain, and analysis of the reflectivity provides complementary information on the surface and interface roughness. Also shown in Fig. 2.13 are simulated profiles obtained with the same values for the thicknesses and compositions used in the rocking curve simulations of Fig. 2.12 (see Table 2.1). Here, the interface abruptness is an additional fitting parameter that is taken into account in the calculations by introducing a phenomenological Debye–Waller interface roughness parameter σ at each individual interface. Simulations obtained by assigning constant values for the interface roughnesses σ_{Si} and σ_{Ge} to all Si to $Si_{1-x}Ge_x$ and $Si_{1-x}Ge_x$ to Si interfaces, respectively, are in qualitative agreement with the experiment. In all cases, a departure between experiment and calculation is seen at higher

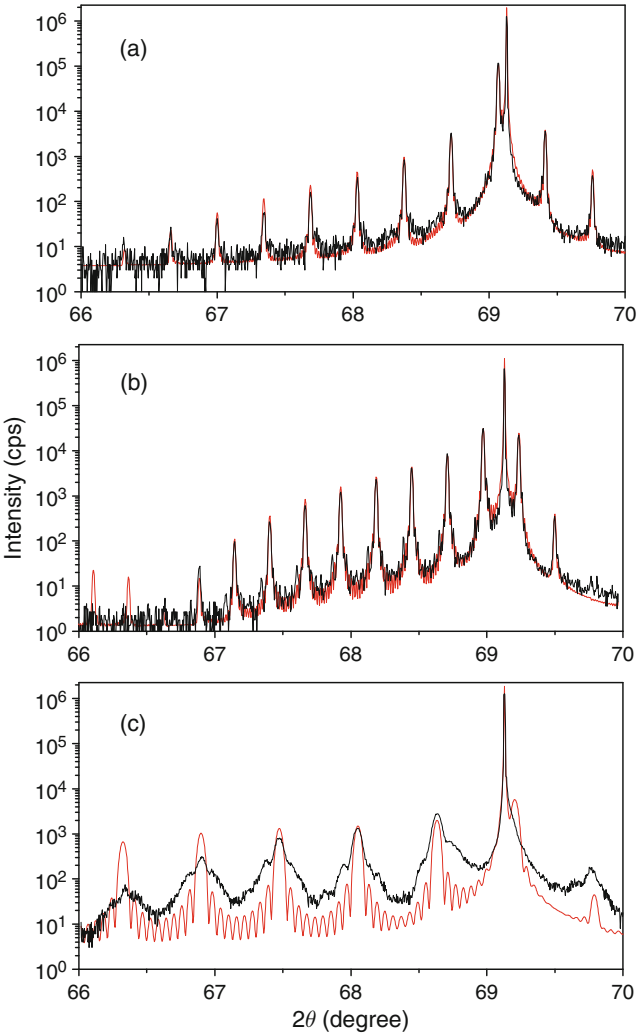


Fig. 2.12 Experimental (*black*) and simulated (*red*) (004) rocking curves from the $Si_{1-x}Ge_x/Si$ multilayers with (**a**) $x = 0.09$, (**b**) $x = 0.17$, and (**c**)$x = 0.53$. Note that in the latter case, the calculation shown here corresponds to $x = 0.42$

angles of incidence and is the result of superlattice irregularities and/or interface blurring. Based on the simulations, the value of σ_{Ge} is systematically larger than that of σ_{Si} and the difference increases for higher Ge concentrations, consistent with the evolution of the alloy layers to a 3D morphology in Ge-rich structures.

Additional information on the spatial coherence of the interfaces in multilayers can be obtained by recording the X-ray scattered intensity about a strong diffraction peak in a reciprocal space map (RSM). Figure 2.14 compares the (004) RSM from

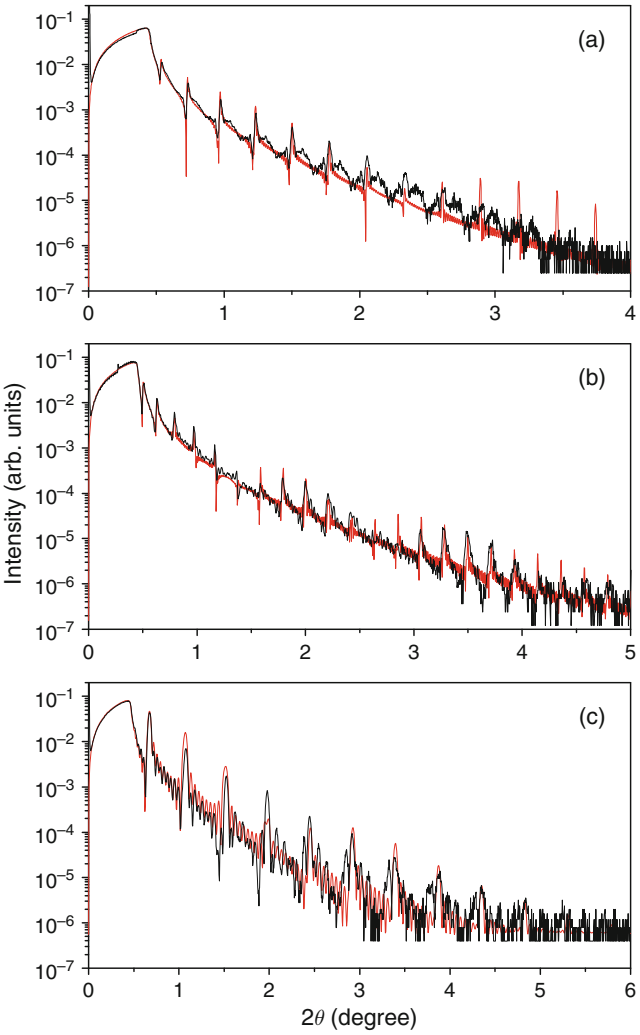


Fig. 2.13 Experimental (*black*) and simulated (*red*) (004) rocking curves from the $Si_{1-x}Ge_x/Si$ multilayers with (**a**) $x = 0.09$, (**b**) $x = 0.17$, and (c) $x = 0.53$. The simulated reflectivity curves were calculated using the structural parameters determined from the high-resolution X-ray rocking curve analysis. Distinct roughness parameters $\sigma_{Si}$and σ_{Ge}, as given in Table 2.1, were assigned to the Si to $Si_{1-x}Ge_x$ and $Si_{1-x}Ge_x$ to Si interfaces. These were assumed to be constant throughout the structures. The simulation also included a surface oxide layer of 0.6 nm roughness and 2 nm thickness for $x = 0.09$ and 0.17 and 1 nm thickness for$x = 0.53$

the various samples. In these maps, the horizontal direction h corresponds to the X-ray wave vector component in a direction parallel to the surface, while the vertical direction l is the wave vector component in the growth direction. While the rocking curves of Fig. 2.12 correspond to the integrated intensity in a vertical direction, the

Table 2.1 Growth and physical data on the $Si_{1-x}Ge_x/Si$ multilayer samples grown by MBE, where x is the nominal Ge composition, t_{Si} is the nominal thickness of the Si spacer layers, t_{SiGe} is the nominal thickness of the $Si_{1-x}Ge_x$ layers, N is the number of periods, $T_{growth}$is the growth temperature, R_{Si} is the Si deposition rate, R_{Ge} is the Ge deposition rate, and σ_{Si} and σ_{Ge} are the Si to $Si_{1-x}Ge_x$ and $Si_{1-x}Ge_x$ to Si interface roughness, respectively, as determined from X-ray reflectance studies

x	t_{Si} (nm)	t_{SiGe} (nm)	N	T_{growth} (°C)	R_{Si}(nm/s)	R_{Ge}(nm/s)	σ_{Si}(nm)	σ_{Ge}(nm)
0.091	27	3.8	20	595	0.2	0.05	0.5	0.7
0.17	34	7.0	15	595	0.2	0.05	0.5	0.9
0.53	15	3.5	10	625	0.075	0.09	0.6	0.9
0.56	13.6	3.0	10	640	0.07	0.09	0.7	1.0
0.61	14.1	3.0	10	650	0.07	0.09	–	–

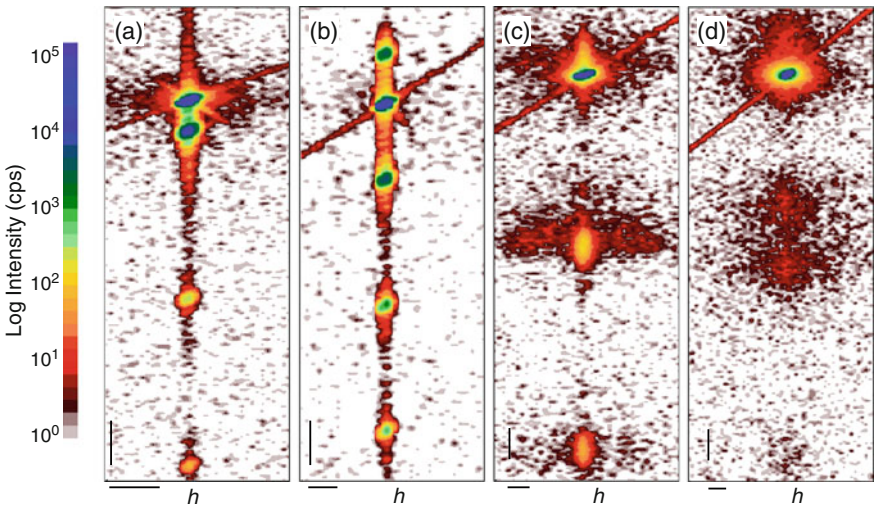


Fig. 2.14 Reciprocal space maps about $(h, k, \ell) = (0, 0, 4)$ for four $Si_{1-x}Ge_x/Si$ multilayers with (**a**) $x = 0.09$, (**b**) $x = 0.17$, (**c**) $x = 0.53$, and (**d**) $x \approx 0.6$. Here the intensity is shown on a logarithmic scale. The horizontal and vertical scale markers correspond to $\Delta h = 0.001$ and $\Delta \ell = 0.005$, respectively

RSM provides details on the actual shape of the various satellite features and reveals additional X-ray scattering away from Bragg conditions. In the RSM, the superperiodicity is revealed by the presence of equally spaced satellite spots aligned in a vertical direction. Interesting differences between samples can be noted by comparing the various RSMs of Fig. 2.14. First, the sample with $x = 0.091$ presents satellites that have the smallest full width $\Delta h = 0.0008$ in the h direction. Samples with $x = 0.17$ and 0.53 exhibit broader satellites with $\Delta h = 0.001$. More interestingly, the latter sample displays side lobes arising from a periodic modulation of the

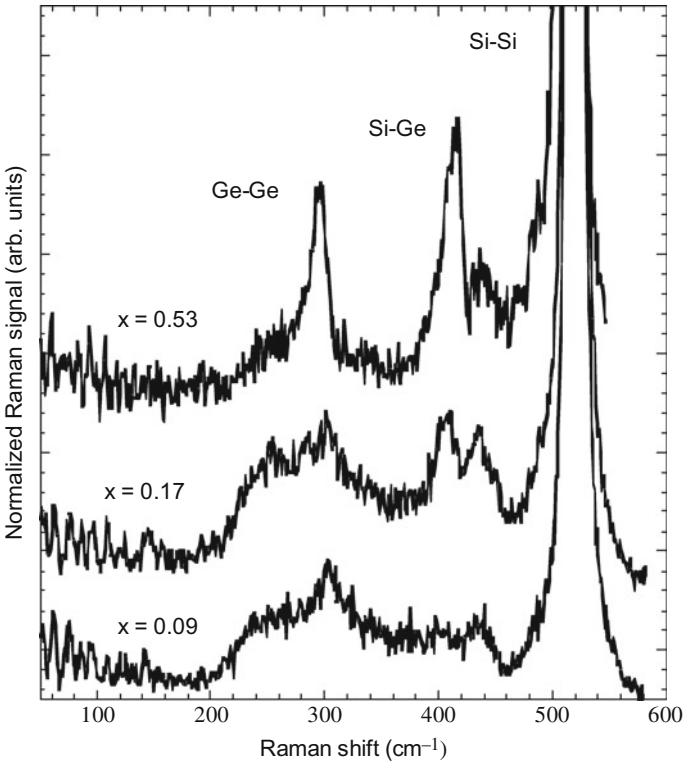


Fig. 2.15 Evolution of Raman spectra in 3D $Si_{1-x}Ge_x$/Si samples with x increasing from 9 to 53%

interfaces similar to that depicted in Fig. 2.10a and this is discussed in detail elsewhere [90, 91]. The sample with $x = 0.61$ exhibits broader diffuse satellites with $\Delta h = 0.002$. This RSM is typical of Ge-rich ($x > 0.6$) multilayers and indicates a significant loss of interfacial perfection and shorter coherence length both parallel and perpendicular to the interface.

Figure 2.15 compares Raman spectra in samples with $Si_{1-x}Ge_x$/Si 3D nanostructures with $0.09 < x < 0.55$. In $Si_{1-x}Ge_x$ islands with a low Ge content ($x = 0.091$) we observe a Raman signal attributed to Si–Si vibrations with a main, zone-center optical phonon peak at ~520 cm^{-1} and a weaker signal at ~300 cm^{-1} associated with Si acoustic phonons (see, for example, [80]). No significant Raman peaks associated with Ge–Ge vibrations at 290 cm^{-1}, Si–Ge vibrations at ~420 cm^{-1} [80], and an amorphous Si phase at ~ 480 cm^{-1} were found. An increase in Ge content up to $x = 0.17$ increases the intensity of the Raman peak associated with Si–Ge vibrations at ~420 cm^{-1}. However, the Raman signal associated with Ge–Ge vibrations at ~310 cm^{-1} is still quite weak. Further increase in the Ge content ($x = 0.53$) in 3D $Si_{1-x}Ge_x$ nanostructures produces strong Raman signals at ~310 cm^{-1} related to Ge–Ge vibrations and at 420 cm^{-1} related to Si–Ge vibrations. Note that in the

sample with $x = 0.53$, the Si–Si vibration-related peak at ~520 cm^{-1} broadens significantly, due to an increase in built-in strain, which in $Si_{1-x}Ge_x/Si$ samples with thicker Ge layers can be entirely localized within the Si spacer layers [93]. Another important observation is the absence of a-Si and a-Ge phases, showing that the islands in $Si_{1-x}Ge_x/Si$ multilayers are true crystalline alloys.

Light-Emitting Properties of Si/SiGe Three-Dimensional Nanostructures

In a system with strong selection rule relaxation (e.g., semiconductor bulk alloys), carrier recombination provides a higher PL quantum efficiency compared to that in an indirect band gap semiconductor, e.g., single-crystal Si and Ge. PL measurements in Si/SiGe NSs reveal a significantly enhanced intensity ratio between no-phonon (NP) luminescence and phonon-assisted luminescence compared to c-Si (Fig. 2.16). Thus, systematic studies of the PL spectra in 3D $Si/Si_{1-x}Ge_x$ NSs with control over the average Ge atomic concentration x provide very important information regarding changes in the carrier recombination mechanism (e.g., selection rule relaxation and conduction and valence band alignment) as x increases from 0 (c-Si) to ~55%, which is the highest Ge composition in structural defect-free SiGe clusters. Fig. 2.16

Figure 2.17a compares PL spectra in MBE samples with $x = 0.096$, 0.16, and 0.53. For samples with $x = 0.096$, we find relatively narrow PL bands at 1.05 and

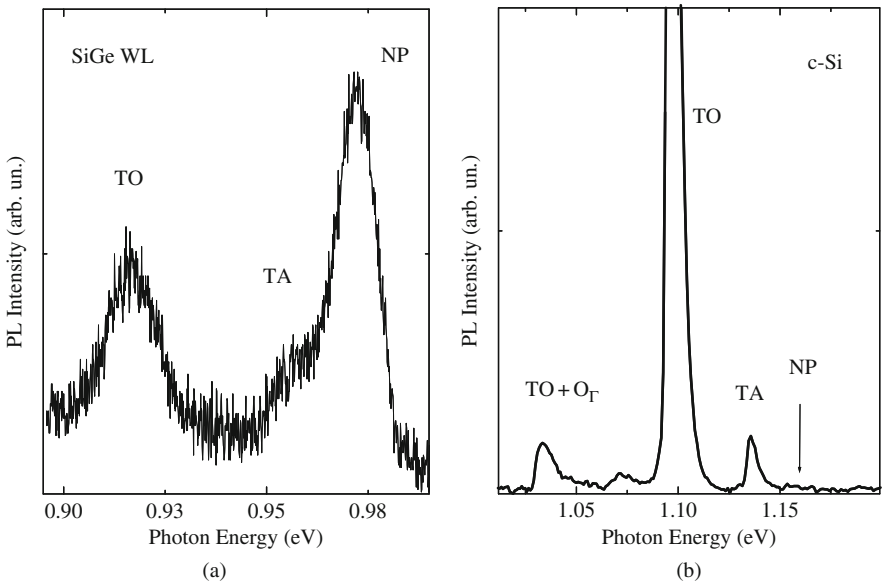


Fig. 2.16 Comparison between low-temperature ($T = 4$ K) PL spectra in (**a**) a SiGe wetting layer with a Ge concentration of ~20% and (**b**) undoped crystalline Si (c-Si). Note the dramatically different intensity ratios between the NP and TO phonon PL lines of the two samples.

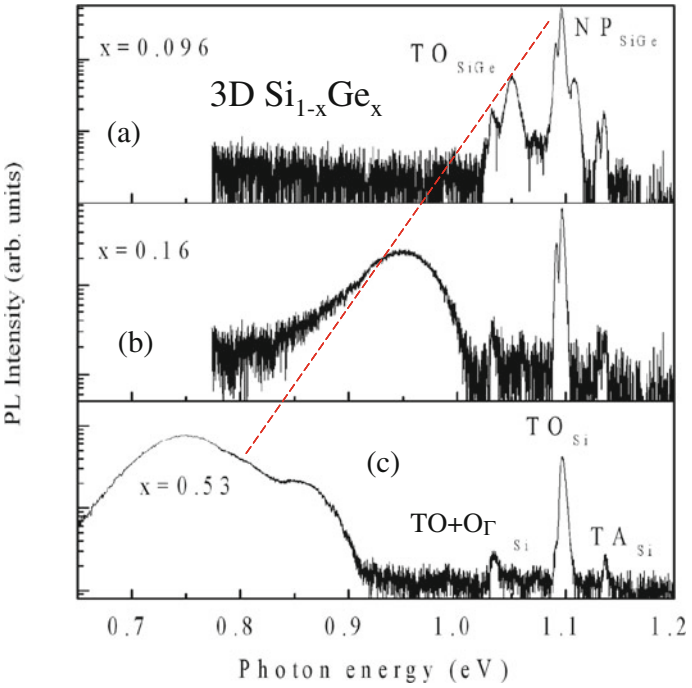


Fig. 2.17 Low-temperature ($T = 4$ K) PL spectra of MBE-grown $Si/Si_{1-x}Ge_x$ 3D NSs with the indicated average Ge atomic concentration x. The characteristic phonon-assisted transitions and the PL spectral shift due to an increase in Ge concentration are indicated by the labels and dashed line, respectively

1.11 eV (Fig. 2.17a) attributed to NP and TO phonon PL bands in Si-rich SiGe alloys [56–58]. Note that these PL bands are in the vicinity of c-Si luminescence, and that the intensities of these two sets of PL bands are comparable. Therefore, we conclude that a small amount of Ge (<10%) slightly reduces the SiGe band gap and mainly relaxes selection rules, increasing the ratio of NP/TO PL band intensities. The observed broadening of these two PL bands, compared to c-Si-related PL, is apparently due to the composition disorder resulting from the introduction of substitutional Ge atoms (~10%) into the c-Si matrix. Increasing the average Ge composition within $Si_{1-x}Ge_x$ clusters up to $x = 0.16$ results in significant changes in the PL spectrum (Fig. 2.17b). An intense PL band peaked at 0.95 eV has appeared, showing an effective SiGe band gap reduction of ~150 meV compared to c-Si. This broad and featureless PL band with a full width at half maximum (FWHM) of ~70 meV indicates a much stronger compositional disorder compared to 3D Si/SiGe NC samples with $x = 0.096$. The PL spectrum in the samples with an average Ge composition close to 53% (Fig. 2.17c) depicts a broad feature with a major PL peak centered at a photon energy of 0.75 eV; this peak energy is close to the band gap of crystalline Ge (c-Ge) at 4 K. A second PL peak is found at ~0.85 eV. These samples

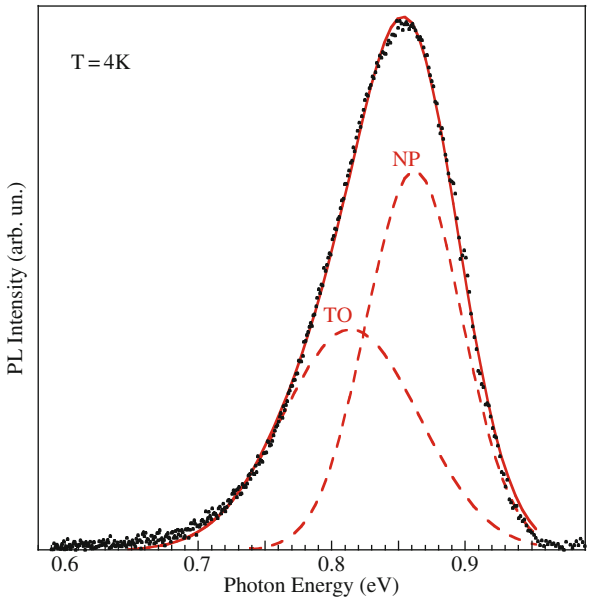


Fig. 2.18 Low-temperature ($T = 4$ K) PL spectra in Si/SiGe 3D NSs grown by CVD. The PL spectrum is fitted quite well with two Gaussian peaks. The PL peak separation energy is close to the SiGe characteristic phonon energy of ~48 meV

have higher PL quantum efficiency than samples with lower x values. Both PL bands are quite broad, most likely due to compositional disorder, which is in agreement with the Raman scattering measurements.

In contrast to MBE-grown samples, CVD growth of 3D Si/SiGe NSs does not provide precise control over the atomic composition, and, most likely, it produces more Si/SiGe interdiffusion at heterointerfaces [80]. This is well reflected in PL spectra, where no fine structure has been found (see Fig. 2.18). However, the broad and asymmetric PL peak is well fitted by two Gaussian bands, often identified as the NP and TO phonon lines [84] and separated by an energy of ~48 meV, which is close to the energy of characteristic SiGe phonons. Thus, there is at least a qualitative similarity between PL spectra in MBE- and CVD-grown 3D Si/SiGe NSs. Also, the PL efficiencies in MBE- and CVD-grown samples are comparable.

Figure 2.19 summarizes the PL intensity as a function of excitation intensity in MBE samples having different average Ge atomic concentrations. The same linear dependence (on a log–log plot) for PL associated with c-Si and a sublinear (close to square root) dependence for PL associated with Ge-rich SiGe clusters has been found in nearly all 3D Si/SiGe NSs grown by both CVD and MBE.

Studies of the PL spectra and intensity temperature dependence show that at low temperatures the PL intensity is nearly temperature independent (Figs. 2.20 and 2.21). At higher temperatures the PL intensity drops exponentially, and the activation energies of PL thermal quenching are shown in Fig. 2.21. There is a clear correlation between the Ge composition in $Si_{1-x}Ge_x$ 3D NSs and the PL intensity temperature dependence. For samples with a low ($x = 0.096$) Ge composition the activation energy E_a is ~10 meV, which is similar to that of the TO phonon peak

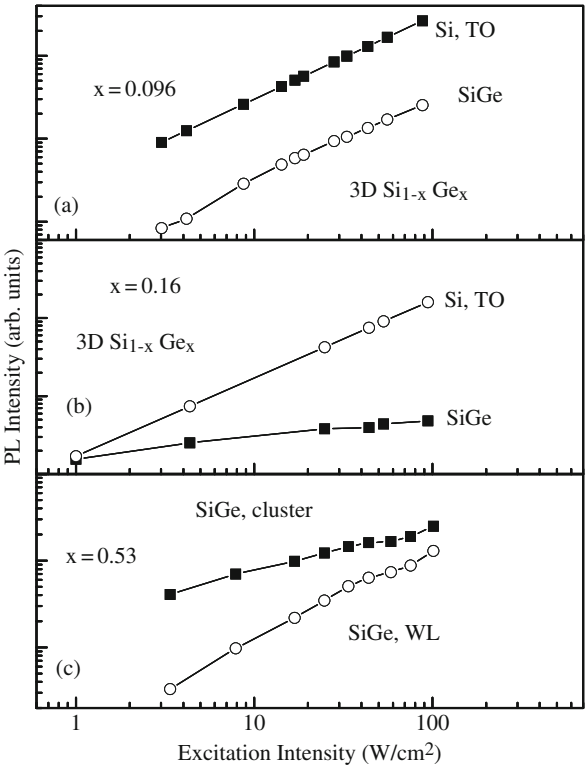


Fig. 2.19 The PL intensity as a function of excitation intensity in MBE-grown $Si/Si_{1-x}Ge_x$ 3D NS samples with the indicated average Ge concentration x. The PL bands associated with the cluster core, wetting layer (WL), and characteristic TO phonons are indicated

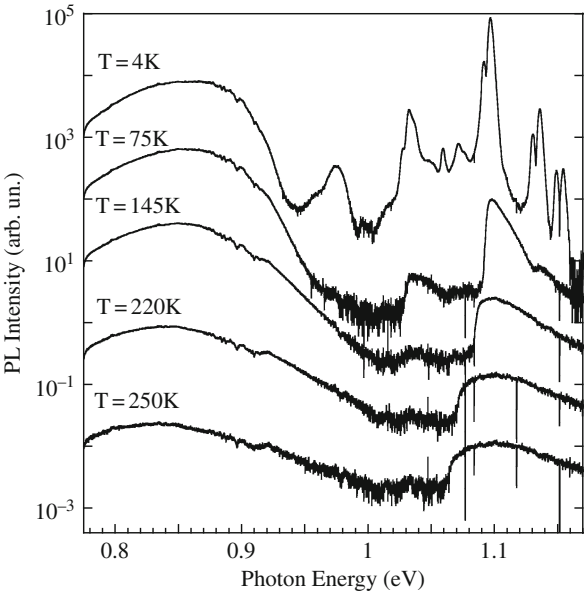


Fig. 2.20 PL spectra in CVD-grown Si/SiGe 3D NSs measured with an excitation intensity of 5 W/cm^2 at the indicated temperatures (the PL spectra have been shifted vertically for clarity)

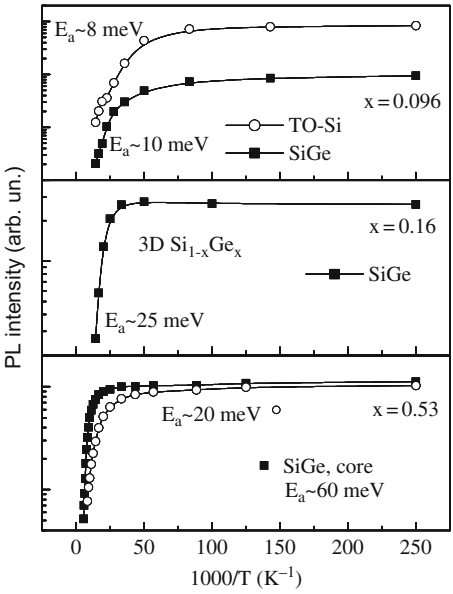


Fig. 2.21 The PL intensity temperature dependence in MBE-grown $Si/Si_{1-x}Ge_x$ 3D NS samples with the indicated average Ge concentration x measured for different PL bands

in c-Si (~8 meV), and the PL has almost vanished by 40 K. This result could be explained by the thermal dissociation of a nearly free exciton [84].

In samples with a higher ($x = 0.16$) Ge composition, the activation energy of PL thermal quenching is increased, and the PL is observable up to ~100 K (Fig. 2.21b). The broad PL band observed at 0.95 eV for the $Si/Si_{1-x}Ge_x$ sample of $x = 0.16$ exhibits an activation energy of ~25 meV, which combined with a sublinear dependence in PL intensity as a function of excitation intensity (Fig. 2.19b) is in contrast with the $x = 0.096$ sample. This observation suggests that the non-equilibrium carriers are spatially localized and that Auger recombination contributes to the overall recombination mechanism even at very low (~100 mW/cm^2) excitation intensity. Most likely, in 3D $Si/Si_{1-x}Ge_x$ NSs with $x = 0.16$, the carriers (possibly holes) are localized within 3D quasi-wells.

In 3D $Si/Si_{1-x}Ge_x$ NSs with $x = 0.53$ the PL spectrum contains two bands peaked at 0.85 eV and 0.75 eV (Fig. 2.17c). There are no characteristic phonons in the Si/SiGe system with an energy of ~100 meV, and it is thus reasonable to assume that the observed PL bands are associated with carrier recombination within two different regions of the 3D SiGe NSs. The PL band peaked at 0.85 eV has almost the same PL quenching activation energy (~20 meV) as in the sample with $x = 0.16$ while the PL intensity as function of excitation intensity is linear over a wide range of excitation intensities. The second PL band peaked at 0.75 eV has an activation energy of ~60 meV (Fig. 2.21c) and is nearly temperature independent up to 100 K. Because of its high quantum efficiency, it can be monitored almost up to room temperature. This data suggests that 3D $Si/Si_{1-x}Ge_x$ NSs with $x = 0.53$ contain coupled subsystems with different (lower and higher) Ge concentrations. It is quite

possible that a spatial localization of electron–hole pairs within 3D regions of higher Ge concentration, and thus having a lower band gap, could be responsible for the observed sublinear excitation dependence of the PL band at 0.75 eV (Fig. 2.19c). The remaining 3D regions with a lower Ge concentration (e.g., having a higher band gap) have a lower carrier concentration and the PL band (peaked at 0.85 eV) exhibits a linear excitation dependence.

The PL thermal quenching observed in 3D $Si/Si_{1-x}Ge_x$ samples grown by MBE could be associated with different mechanisms. As mentioned earlier, the activation energies of the PL thermal quenching in samples with $x = 0.096$ and $x = 0.16$ are close to the observed exciton binding energy in SiGe alloys and $Si/Si_{1-x}Ge_x$ NSs. In samples with $x = 0.53$, the greater activation energy could be attributed to carrier diffusion from a 3D potential well. This assumption is justified by the expected type II band alignment in SiGe NSs with a deep (>100 meV) potential well for holes and a relatively small potential barrier for electrons [88]. Thus, phonon-assisted carrier tunneling can produce the observed ~60 meV activation energy for thermal quenching of the PL intensity.

A continuous shift of the PL band from ~1 to 0.75 eV has been found previously in SiGe alloys with increasing Ge concentration [95–100]. Instead, in these 3D $Si/Si_{1-x}Ge_x$ samples with a Ge concentration higher than 50%, a simultaneous threshold-like appearance of two clearly resolved PL peaks at 0.85 and 0.75 eV is observed. This suggests that Ge segregation might take place as x increases up to ~0.5. Since the PL peak at 0.75 eV is close to the value of the band gap in pure c-Ge, we propose that such a segregation results in a Ge-rich core within a SiGe shell forming the 3D $Si/Si_{1-x}Ge_x$NSs embedded within a pure Si matrix.

It has been known for some time that the PL spectra in 3D Si/SiGe NSs, which are similar to that in III–V quantum wells with type II energy band alignment [101], exhibit a blue shift as the excitation intensity increases [63–64]. In our studies, this effect is found in both MBE- and CVD-grown samples. Figure 2.22 shows PL

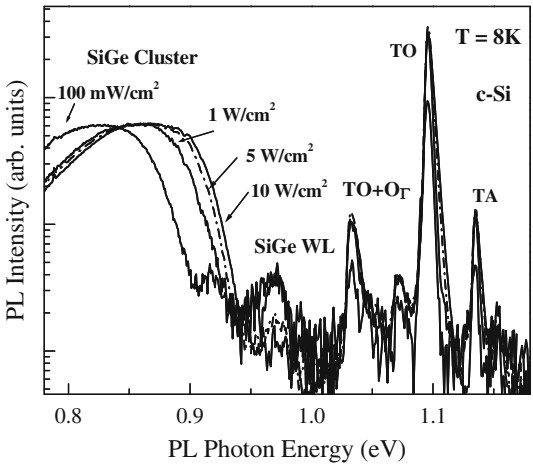


Fig. 2.22 PL spectra in CVD-grown Si/SiGe Si/SiGe 3D multilayer samples measured at $T = 8$ K under different (indicated) excitation intensities

spectra in a CVD-grown sample measured under excitation intensities varying from 0.1 to 10 W cm^{-2}. At the lowest excitation intensity used (0.1 W cm^{-2}), the PL peaks at ~0.8 eV. With increasing excitation intensity, a continuous almost-parallel PL blue shift of 30–40 meV per decade of excitation intensity increase is observed. At an excitation intensity of 10 W cm^{-2}, the PL peak reaches ~0.92 eV. Under photoexcitation of 1–10 kW cm^{-2}, the low-energy part of the PL spectrum does not shift further, while the high-energy part continues shifting toward higher energy.

Figure 2.20 compares PL spectra measured using a fixed excitation intensity (~5 W cm^{-2}) at different temperatures in the range 8–210 K. These measurements clearly show that, with increasing temperature, the PL peak associated with SiGe clusters at 0.85–0.9 eV shifts toward lower photon energies. This "red" PL spectral shift is, most likely, associated with the SiGe band gap decrease as the sample temperature increases. Thus, it is opposite to the "blue" PL spectrum shift observed with an excitation intensity increase, and thus this blue shift cannot simply be explained by sample heating due to the intense incident laser beam (Fig. 2.22). There was some sample heating in the measurements, but it was not significant except at the higher laser powers. Using the observed broadening of the PL spectral features associated with c-Si, the sample temperature is estimated to rise from 4 to 60 K under the highest excitation intensity used.

Figure 2.23 shows a modified (note the double logarithmic scale) Arrhenius plot of the normalized integrated PL intensity of a 3D Si/SiGe multilayer sample grown by CVD and measured at different excitation intensities. The normalized PL intensity temperature dependencies are fitted by a standard equation:

$$I_{PL}(T) = \frac{1}{\left[1 + C_1 \cdot \exp\left(-\frac{E_1}{kT}\right) + C_2 \cdot \exp\left(-\frac{E_2}{kT}\right)\right]}. \tag{2.1}$$

(see, for example, [41]) with two thermal quenching activation energies E_1 and E_2. Here T is the temperature, k is the Boltzmann's constant, and C_1 and C_2 are scaling coefficients. In all measurements of all samples, the PL thermal quenching activation energy $E_1 \approx 15$ meV and E_1 is independent of excitation intensity. In contrast, the activation energy E_2 depends significantly on the excitation intensity: the PL temperature dependence shows a step-like behavior, and E_2 increases dramatically from ~120 to 340 meV as the excitation intensity increases from 0.1 to 10 W cm^{-2}.

The PL dynamics, i.e., the PL intensity decay under pulsed laser excitation, is an important technique for studying the carrier recombination mechanism. Figure 2.24 shows the normalized low-temperature ($T = 4$ K) PL decays collected from a CVD-grown Si/SiGe 3D sample. The initial PL decay is fast—close to the resolution of our detection system (<20 ns). The longer-lived PL shows a strong dependence on the detection photon energy, as summarized in Fig. 2.25: the PL lifetime at photon energies below 0.8 eV is found to be ~20 μs, and it drastically decreases to ~200 ns for the PL component measured at 0.89 eV.

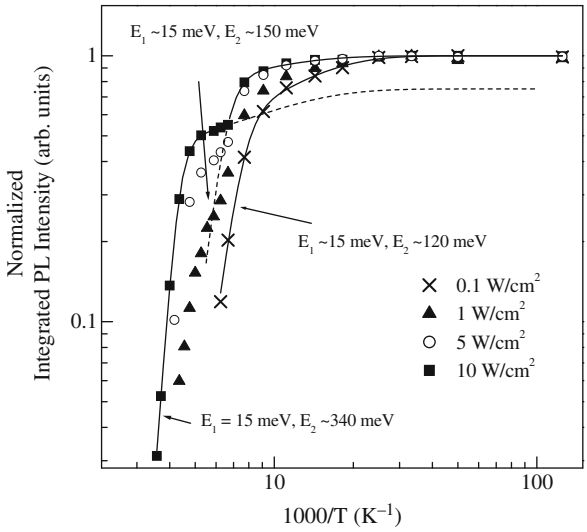


Fig. 2.23 Typical integrated PL intensity behavior for CVD-grown Si/SiGe 3D multilayer samples as a function of the reciprocal temperature measured with different excitation intensities, as indicated. The activation energies obtained from fits (shown by the different lines) to the data (shown by the different points) are also given

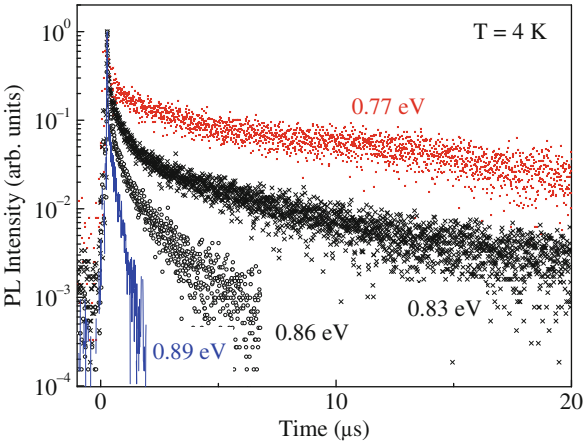


Fig. 2.24 Typical low-temperature ($T = 4$ K) PL dynamics for CVD-grown Si/SiGe 3D multilayer samples measured at the indicated photon energies using a short (~6 ns) excitation pulse

Carrier Recombination in Si/SiGe Nanostructures

In our discussion of carrier recombination in Si/SiGe NSs, we focus on MBE- and CVD-grown 3D (or cluster morphology) samples with an average Ge atomic composition close to 50%. These samples show the highest observed PL quantum efficiency (better than ~1% at low excitation intensity) with a PL peak wavelength close to 1.5–1.6 μm. The PL spectral distribution extending well below the band gap of pure Ge (Figs. 2.17 and 2.18) and the extremely long carrier radiative lifetime of ~10 ms [88], as well as the ~30 meV per decade PL spectral shift toward higher photon energies as the excitation intensity increases [102], point out strong

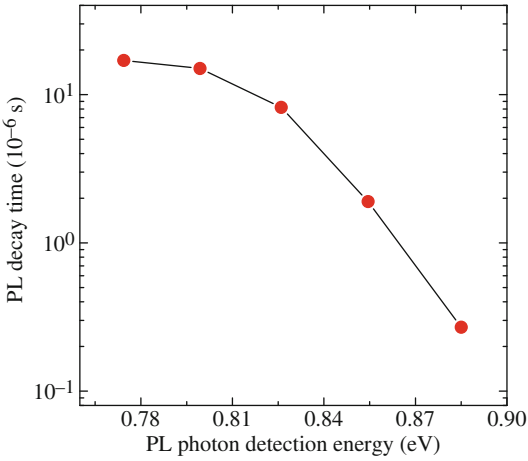


Fig. 2.25 Summary of PL lifetimes ($T = 4$ K) for CVD-grown Si/SiGe 3D multilayer samples at different photon energies

similarities between the PL in 3D Si/SiGe NSs and the PL in III–V quantum wells with type II energy band alignment [101]. Generally, a type II energy band alignment at the heterointerface is a strong disadvantage for light-emitting devices due to a weak overlap between spatially separated electron and hole wave functions. In reality, however, the critical limitation in the efficiency of light-emitting structures is rather the presence of competing non-radiative recombination channels for excess electrons and holes. The most important non-radiative mechanism is carrier recombination via defects, especially heterointerface structural defects such as propagating dislocations and dislocation complexes. The 3D Si/SiGe NSs investigated here, grown by both MBE and CVD processes, show an almost undetectable density of dislocations. Thus, in these properly grown 3D Si/SiGe NSs at low excitation intensities, we find a high quantum efficiency of PL with photon energy <0.9 eV, which is associated with carrier recombination in the vicinity of SiGe clusters.

It has been suggested that SiGe Stranski-Krastanov (S-K) clusters with a small (3–5 nm) height and ~10:1 base–height aspect ratio can be modeled as NSs with a type II energy band alignment and possible SiGe cluster valence band energy quantization in the direction of growth [99]. Strained Si and Si-rich SiGe alloy regions near the base of the clusters (also called SiGe wetting layers) also need to be considered. Including the effect of strain, the observed PL bands at 0.916 and 0.972 eV indicate a composition of the $Si_{1-x}Ge_x$ transition region, which is located near the bottom of the Ge/Si pyramid-like clusters, to be close to $x \approx 0.2$, and this conclusion is supported by recent direct analytical TEM measurements [90, 91].

It has also been proposed that the broad PL band with a peak energy of ~0.8–0.9 eV is due to the recombination of carriers localized in the Ge-richest areas of the clusters, which is close to the center of "pancake"-shaped SiGe clusters [89]. We suggest that at the lowest excitation intensity, the PL arises from electron–hole recombination between holes localized in the Ge-richest regions of the cluster and electrons localized in the strained SiGe alloy region near the cluster base. This immediately explains the extension of the observed PL spectrum below the pure Ge band gap energy.

The observed excitation-independent PL thermal quenching activation energy of ~15 meV is close to the exciton binding energy in SiGe alloys and Si/SiGe superlattices. Thus we conclude that one of the mechanisms of PL thermal quenching is the thermal dissociation of excitons. The activation energy of ~15 meV can therefore be associated with exciton localization on specific regions of the clusters associated with variations of the SiGe composition. Hence, the non-uniform SiGe cluster composition and, perhaps, variations in SiGe cluster size and shape could be responsible for the observed relatively broad PL spectra.

Using the model discussed above of type II energy band alignment at the Si/SiGe heterointerface, we focus next on non-radiative carrier recombination and the different mechanisms of electron–hole separation. Electron transport in 3D Si/SiGe NSs is limited by a small ($\leq 10-15$ meV) conduction band energy barrier and SiGe compositional disorder [100, 103]. Thus, the PL thermal quenching activation energy of ~15 meV could also be associated with a small conduction band energy barrier for electrons in Si/SiGe 3D NSs.

In contrast to a low-energy barrier for electrons, hole diffusion in 3D Si/SiGe multilayer NSs with a high Ge content is controlled by large (>100 meV) valence band energy barriers at Si/SiGe heterointerfaces. In this system, we consider two major mechanisms of hole transport: (i) hole tunneling and (ii) hole thermionic emission. Hole tunneling in 3D Si/SiGe NSs with thin (5–7 nm) Si-separating layers and nearly perfect SiGe cluster vertical self-alignment could be very efficient. These NSs are usually grown by MBE and exhibit a PL thermal quenching activation energy of ~60 meV. The same PL thermal quenching activation energy is found in our CVD-grown samples with 7.5 nm thick Si-separating layers for the lowest excitation intensity. We suggest that in 3D Si/SiGe multilayer NSs with thin Si layers at low excitation intensity, the electron–hole separation and non-radiative carrier recombination are mainly controlled by hole tunneling between SiGe clusters. Due to significant variations in SiGe cluster size, shape, and chemical composition, the process of hole tunneling could be assisted by phonon emission and/or absorption [102, 104]. Therefore, the observed PL thermal quenching activation energy is close to the Si TO phonon energy. In 3D Si/SiGe multilayer samples with 20 nm thick Si layers, where SiGe cluster vertical self-alignment is practically absent [89, 91], the probability of hole tunneling is reduced, and hole thermionic emission over the Si/SiGe heterointerface barrier plays a bigger role. Thus, in these samples the PL thermal quenching activation energy is expected to be greater, as has been found in our experiments (see Fig. 2.23).

In this simple model, efficient hole tunneling between adjacent SiGe nanoclusters requires not only reasonably low and thin Si barriers but also a low carrier concentration (i.e., a large enough number of empty adjacent SiGe clusters). By increasing the photoexcitation intensity (i.e., the number of photogenerated carriers), hole tunneling can effectively be suppressed since fewer empty adjacent SiGe clusters can be found. At high excitation intensity, assuming (i) a negligible value of the conduction band offset compared to that in the valence band and (ii) a nearly pure Ge composition in the SiGe cluster core, the maximum anticipated PL thermal quenching activation energy should be $E_2 \leq E_g^{Si} - E_g^{Ge} \leq 400$ meV. This value sets the upper limit of the activation energy of PL intensity thermal quenching in 3D SiGe

multilayer NSs, and it is close to the activation energy of $E_2 \approx 340$ meV that has been found under the highest excitation intensity (see Fig. 2.23).

The same model also explains the experimentally found strong decrease in carrier radiative lifetime (~100 times) as the detection photon energy increases from 0.77 to 0.89 eV (Fig. 2.25). Under a low level of excitation intensity, holes are localized within the Ge-rich core of SiGe clusters, and the hole wave function does not penetrate into the Si barriers. Thus, in this quasi-type II energy band alignment, electron and hole wave functions do not overlap, causing a long carrier radiative lifetime. Under high-level excitation, holes occupying the excited energy states in small-size SiGe clusters, as well as holes leaking into the SiGe wetting layer with a lower Ge concentration, have their wave function significantly extended into the Si barriers. Therefore, a stronger overlap between electron and hole wave functions is responsible for a shorter radiative lifetime, i.e., a higher PL quantum efficiency at greater photon energy. Thus, this explanation is consistent with both the decrease of the PL decay time at greater photon energy and the PL spectral blue shift as excitation intensity increases.

Figure 2.26 depicts the quasi-type II energy band alignment at Si/SiGe heterointerfaces, where a gradual increase of Ge concentration toward the SiGe cluster core has recently been experimentally verified (see Fig. 2.11 and also [91]). It also shows five major processes controlling electron–hole recombination in 3D Si/SiGe NSs with a Ge-rich core:

(1) radiative recombination between electrons localized at the Si/SiGe heterointerface and holes localized within the Ge-rich core of a SiGe cluster—this slow recombination is associated with the lower photon energy part of the PL band;
(2) radiative recombination between electrons localized at the Si/SiGe heterointerface and holes localized in excited states in small-size SiGe clusters as well as holes leaking into the Si-rich outer part of the cluster or/and SiGe wetting layer—this faster carrier recombination channel is responsible for the higher photon energy part of the PL spectrum;

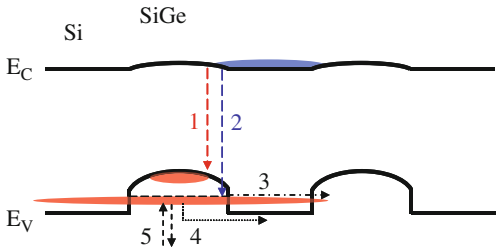


Fig. 2.26 Schematic of the Si/SiGe cluster multilayer energy band diagram with different electronic transitions: 1—slow electron–hole recombination due to spatial separation and a weak wave function overlap; 2—a faster recombination, presumably involving holes at excited energy states in SiGe clusters; 3—hole diffusion due to cluster-to-cluster tunneling; 4—hole diffusion due to thermionic processes; and 5—Auger hole excitation with possible charge transfer

(3) non-radiative carrier recombination due to electron–hole separation via hole tunneling in Si/SiGe NS samples with thin (<10 nm) Si barriers;
(4) non-radiative carrier recombination due to electron–hole separation via hole thermionic emission in samples with thicker (>15–20 nm) Si barriers;
(5) non-radiative carrier recombination due to Auger processes.

It has been pointed out that possible transformations of hole energy spectra due to quantization and/or strain might dramatically increase the rate of non-radiative Auger recombination in SiGe quantum wells by more than 100 times [105]. Thus, the "Auger limit" in Si/SiGe NSs could be considerably lower compared to that in c-Si and c-Ge. On the other hand, the Auger process competes with radiative processes, and their characteristic lifetimes are ~ 10^{-3} s in bulk Si and Ge and just ~10^{-7} s in the high photon energy part of the PL spectra in 3D Si/SiGe NSs. More work is being done (see below and [56]) to understand the important details of Auger processes in Si/SiGe NSs.

Figure 2.26 also reflects a continuous change in the Ge atomic concentration by a gradually increasing energy band gap from the cluster center toward the cluster edge and SiGe wetting layer where the Ge atomic concentration is estimated to be ~20%. The proposed energy band diagram and carrier recombination/diffusion mechanisms explain the experimental observations of (i) the PL spectral blue shift under increasing excitation intensity; (ii) the dramatic (~100 times) decrease in carrier radiative lifetime measured at photon energies from 0.77 to 0.89 eV; and (iii) the unusual PL intensity temperature dependence, which shows a different PL thermal quenching activation energy at different excitation intensities. This schematic representation also points out that moderate excitation intensity changes the overlap of electron–hole wave functions and allows a faster carrier radiative recombination. In other words, the type II energy band alignment at the Si/SiGe heterointerface can effectively be replaced by the "dynamic type I" alignment, where the electron and hole spatial separation no longer controls the recombination rate and the quantum efficiency of PL and EL.

Electroluminescence in Si/SiGe Nanostructures

Electroluminescence (EL) in Si/SiGe NSs was demonstrated almost simultaneously with the first investigations of the PL [33, 99]. The structures used for the EL studies usually are Si/SiGe p-i-n diodes with SiGe clusters, SiGe alloy layers, or SiGe QWs in the i-region. Figure 2.27a shows EL spectra with quite narrow EL peaks attributed to NP and TO phonon emission [33]. The proposed EL mechanism is quite similar to the PL mechanism, i.e., it is due to radiative electron–hole recombination in Si/SiGe NSs with electrically injected electrons and holes. Similar to that in PL, the EL spectrum depends on the SiGe layer (or cluster) composition, and a higher Ge concentration results in the EL spectrum red shift (i.e., a shift toward lower photon energy). Figure 2.27b shows the dependence of the integrated EL intensity on current. The EL intensity versus current is a power function $I_{EL} \sim J^n$ where I is the

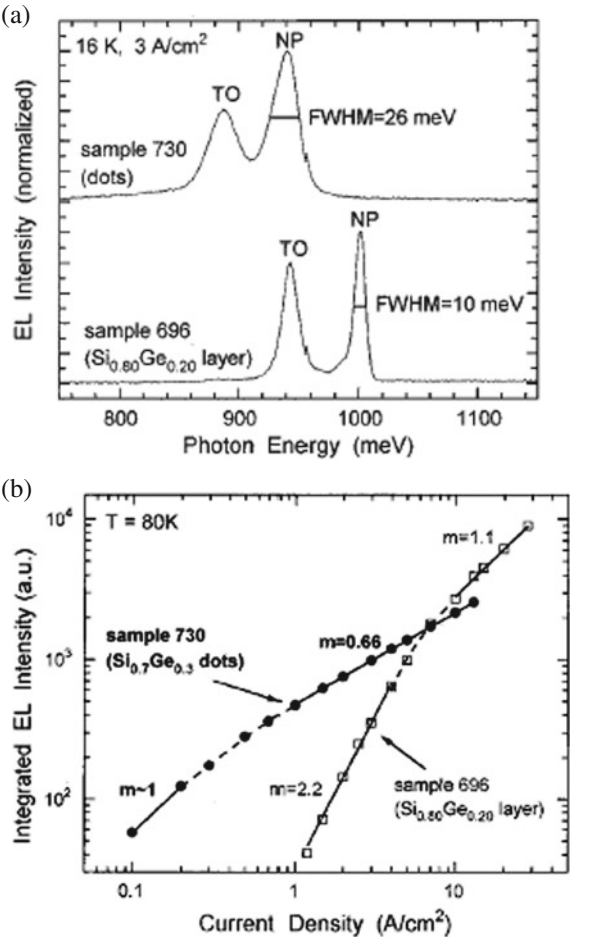


Fig. 2.27 (**a**) EL spectra of samples containing 3D SiGe NSs and (**b**) the EL intensity as a function of current under reverse bias. (Reprinted with permission from [33], Copyright 1995, American Institute of Physics)

intensity, J the current density, and $0.5 < n < 2$ [33]. A good-quality Si/SiGe p-i-n structure usually provides current flow associated with carrier diffusion in a forward biased p–n junction (i.e., an exponential current as a function of forward bias), and the diode ideal factor is less than 2 [33, 82, 106].

In Si/SiGe QW-based devices, the EL intensity temperature dependence, similar to that in PL, exhibits significant thermal quenching [33]. The EL thermal quenching activation energy is greater in Si/SiGe 3D nanostructures (especially with small-size SiGe clusters) compared to that in planar QWs, and the EL can be extended up to room temperature (Fig. 2.28) [106]. In [106], the EL thermal quenching activation energy is found to be as large as 200–250 meV, and this is, possibly, the largest reported value. The SiGe cluster EL spectra full width at half maximum is relatively narrow (~90 meV), possibly reflecting the cluster narrow size distribution. When the temperature increases from 40 K up to room temperature, the EL is reduced by a

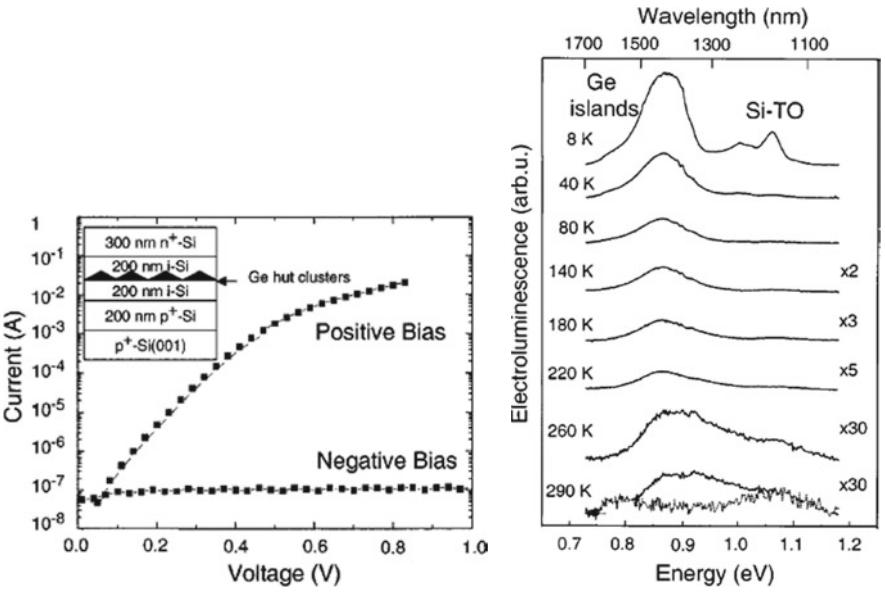


Fig. 2.28 (**a**) Room-temperature current–voltage characteristic of the p-i-n diode including one Ge hut cluster layer in the i-region. The device structure is given in the inset. (**b**) Temperature dependence of the EL spectra at constant current of 400 mA. (Reprinted with permission from [106], Copyright 2003, American Institute of Physics)

factor of ~40, but the EL peak is still observable at approximately the same photon energy. Considering the EL mechanism, the luminescence efficiency is believed to be limited (similar to that in PL) by type II radiative recombination between holes confined in the SiGe clusters and electrons localized in the silicon spacer layers above and below the islands. Also in the p-i-n structures with 3D Si/SiGe layers within the active region, the series resistance is found to be high (Fig. 2.28). Thus, the EL power efficiency is low, in part due to the high series resistance of the Si/SiGe p-i-n diode.

Typically, as already mentioned, for a current density of 1–100 A cm^{-2}, the EL intensity is a power function of current $I_{EL} \sim J^n$ with $0.5 < n < 2$. Thus, in p-i-n structures, the EL intensity should have an exponential dependence on the applied forward bias. In a simpler structure with two nearly ohmic contacts to Si/SiGe cluster multilayers, the integrated EL intensity is found to be nearly linear as a function of the applied voltage (Fig. 2.29). There are no energy barriers due to the p-i-n junction in these samples, and the EL intensity as a function of temperature shows thermal quenching with a smaller activation energy of ~130 meV (Fig. 2.30a). Interestingly, in the same sample, the device current as a function of temperature shows nearly an exact anti-correlation with the EL intensity and exhibits an activation energy of ~140 meV (Fig. 2.30b). Since both carrier transport and EL intensity exhibit a strong dependence on temperature, an accurate estimation of the EL quantum efficiency over a wide temperature region requires normalization of the EL

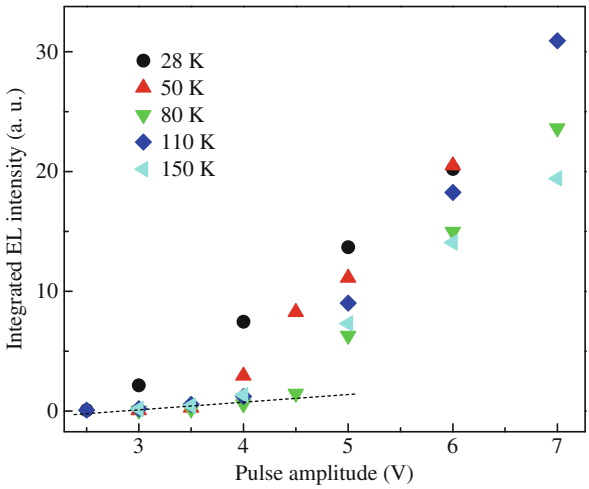


Fig. 2.29 Integrated EL intensity over the 0.75–0.95 eV spectral region for a CVD-grown Si/SiGe 3D multilayer sample as a function of the pulsed voltage amplitude measured at the indicated temperatures. The dashed straight line is a guide to the eye

intensity on current density, i.e., measurements of how many photons are emitted per one injected electron–hole pair at different temperatures.

The importance of the last statement is pointed out in [107], where the EL of light-emitting diodes with 5, 10, and 30 layers of Si/Ge clusters in the active region is studied. The enhanced integral EL intensity at temperature $T \geq 200\,\mathrm{K}$ in the samples with 30 layers of Si/Ge clusters presents a behavior opposite to the usually observed EL intensity thermal quenching, which is also found in the similar samples with 5 and 10 Si/SiGe layers. In [107], this EL intensity temperature dependence is attributed mainly to the EL intensity dependence on injection current, which according to Fig. 2.30 is temperature activated.

Photoluminescence in Si/SiGe Nanostructures Under High Excitation Intensity

Figure 2.31 shows PL spectra of 3D Si/SiGe nanostructures measured under different levels of photoexcitation [56]. Using relatively low ($\leq$1 W cm^{-2}) CW excitation, a broad PL feature is found at 0.75–0.9 eV, which is associated with SiGe clusters, together with a much weaker PL peak that is associated with Si exciton recombination at 1.096 eV. On increasing the excitation intensity there is a rapid saturation of the SiGe cluster PL. Under pulsed excitation of ~6 ns duration and an energy density of ~0.1 mJ cm^{-2}, which corresponds to a nearly ~1,000 times higher peak intensity, SiGe cluster PL becomes negligible compared to Si PL. At the same time, a broader PL peak at 1.079 eV appears. When the excitation intensity approaches 10 mJ cm^{-2}, this broad PL peak shifts to a slightly higher photon energy (~1.085 eV), and it quickly becomes the dominant PL feature (see Fig. 2.31). We find that the intensity of this PL band as a function of excitation intensity follows the dependence

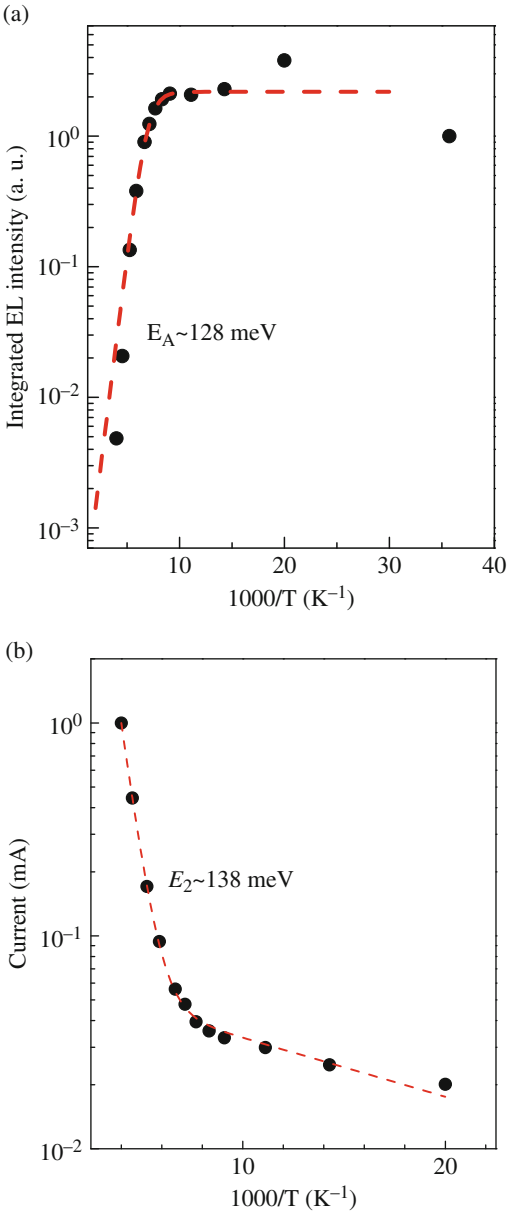


Fig. 2.30 (**a**) The EL intensity and (**b**) current as functions of the reciprocal temperature for a CVD-grown Si/SiGe 3D multilayer sample. The activation energies obtained from fits to the data (*dashed lines*) are also shown

$I_{PL} \sim (I_{\text{exc}})^n$, with $n \approx 1.2-2$, and it anti-correlates with the SiGe cluster PL where n correspondingly varies from 0.8 to 0.5. This faster growing, broader PL band is associated with the recombination of Si electron–hole condensates or electron–hole droplets (EHDs).

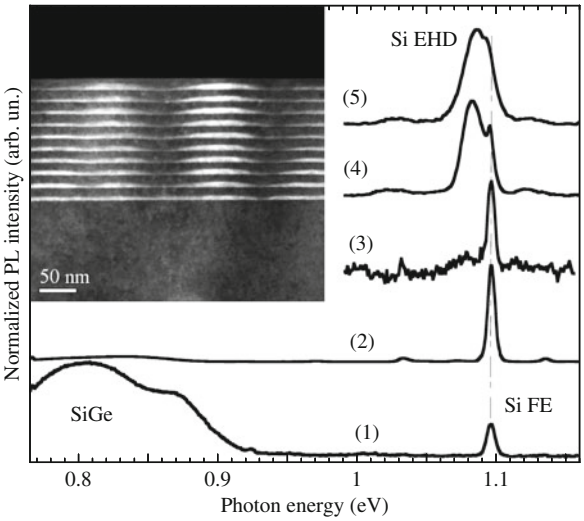


Fig. 2.31 The PL spectra in 3D Si/SiGe nanostructures at 15 K measured under CW excitation with intensities of (1) ~1 W cm^{-2} and (2) ~100 W cm^{-2}; and under pulsed excitation with 6 ns pulse duration and energy densities of (3) ~10^{-5} J cm^{-2}, (4) ~10^{-3} J cm^{-2}, and (5) ~10^{-2} J cm^{-2}. The PL spectra are shifted vertically for clarity, and PL peaks associated with SiGe clusters (SiGe), Si free excitons (FE), and Si EHDs are indicated. The inset shows a dark-field cross-sectional TEM micrograph with SiGe clusters as lighter areas separated by ~10–15 nm thick Si layers

Figure 2.32 shows low-temperature PL dynamics in a high resistivity ($\geq 10^4$ Ω·cm) bulk c-Si sample and focuses on Si exciton recombination at 1.096 eV and Si EHD recombination at 1.079 eV. There is a dramatic difference in the PL decay lifetimes, changing from ~3 μs in Si exciton PL to ~70 ns in Si EHD PL. For this

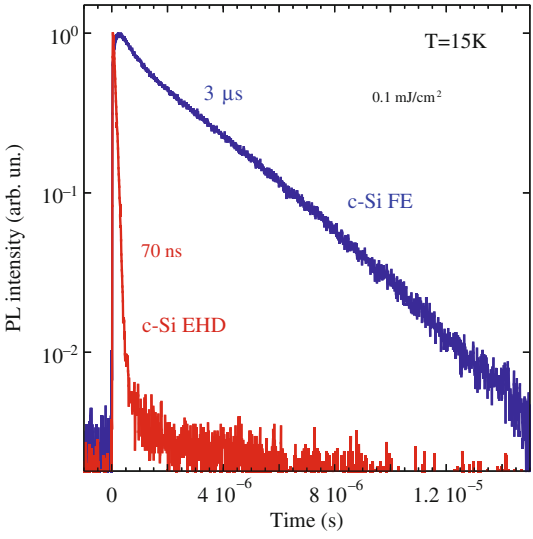


Fig. 2.32 Low-temperature PL dynamics in a bulk c-Si sample under 0.1 mJ cm^{-2} energy density excitation recorded at photon energies associated with the Si FE and Si EHD PL bands

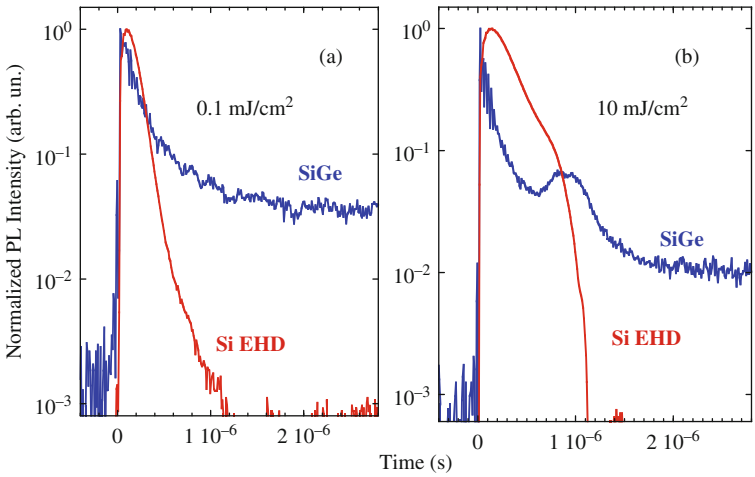


Fig. 2.33 The PL dynamics under pulsed laser energy densities of (**a**) 0.1 and (**b**) 10 mJ cm^{-2} recorded at photon energies associated with SiGe cluster PL (~0.82 eV) and Si EHD PL (~1.08 eV)

excitation intensity, both PL decays are nearly a single exponential. The PL rise time in c-Si EHDs is faster than the experimental time resolution of ~2 ns, while in c-Si exciton PL the rise time is ~20 ns.

Figure 2.33 depicts low-temperature PL dynamics in 3D Si/SiGe nanostructures, focusing on SiGe cluster PL (measured at photon energy 0.82 eV) and Si EHD PL with a peak at ~1.08 eV. The PL associated with SiGe clusters rises practically instantaneously, while the PL associated with Si EHDs, in contrast to that in bulk c-Si (Fig. 2.32), has a slower rise time of ~20–40 ns. For an excitation energy density of 0.1 mJ cm^{-2}, the SiGe cluster PL decay is non-exponential, with a fast initial characteristic lifetime of ~20 ns followed by a much slower ($\tau \geq 10^{-4}$ s) decaying PL. At the same time, the Si EHD PL has a nearly single exponential decay with a lifetime $\tau_{EHD} \sim 50$ ns (Fig. 2.33a). Using a 100 times higher excitation intensity, there is an acceleration of the fast component of SiGe cluster PL decay surprisingly followed by a non-monotonic, at first rising and then falling (i.e., N-shape), PL signal (Fig. 2.33b). Under the same excitation conditions, the Si EHD PL exhibits a nearly exponential decay with a lifetime of ~200 ns followed by a very fast (faster than 20 ns) decay (Fig. 2.33b). The Si exciton PL dynamics in Si/SiGe 3D nanostructures, like that in bulk c-Si, is much slower, with a microsecond characteristic lifetime.

Assuming a quasi-type II energy band alignment at Si/SiGe heterointerfaces (see Fig. 2.34), we propose that a 3D Si/SiGe nanostructure can be represented by a coupled electronic system, where Auger recombination in SiGe clusters not only is responsible for the fast saturation of the SiGe cluster PL intensity but also injects "Auger holes" into the Si nanometer-thick layers. This process facilitates the formation of Si EHDs. Thus, compared to bulk c-Si, where at an excitation energy density of $\geq$0.1 mJ cm^{-2} EHDs are formed practically instantly from hot electron–hole

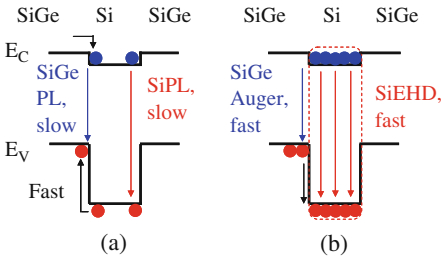


Fig. 2.34 Schematic representation of recombination processes in a Si/SiGe 3D nanostructure. (**a**) At low excitation intensity, because of the long carrier radiative lifetime in Si energy barriers, a hole falls into a SiGe energy well and recombines with an electron localized in Si. (**b**) At high excitation intensity, the "Auger fountain" ejects holes from SiGe clusters into Si layers and facilitates the formation of Si EHDs. The fast Si EHD recombination becomes the dominant recombination mechanism

plasma condensation, in Si/SiGe 3D nanostructures Si EHD formation is mediated by the Auger processes with a characteristic time of ~10^{-8} s.

At low excitation intensity, the Si electron–hole density is too low for EHD condensation and Si carrier recombination is slow. Under these conditions, photogenerated holes created in Si would rather fall into SiGe clusters. Thus, assuming a defect-free Si/SiGe heterointerface, slow radiative recombination between spatially separated electrons and holes at the Si/SiGe heterointerface is the only possible channel of radiative recombination. This conclusion is supported by the spectra given in Fig. 2.31, which show a dominant SiGe PL band only at low excitation intensity. At a high level of photoexcitation, SiGe Auger recombination overcomes the slow recombination of spatially separated electrons and holes at the Si/SiGe heterointerface (i.e., an indirect exciton). This process generates "Auger holes" with energies of 0.7–0.8 eV, while the valence energy barrier at the Si/SiGe heterointerface is <0.3–0.4 eV, and thus the fast Auger hole injection from SiGe clusters into Si barriers can be very efficient. The Auger-mediated hole transfer from SiGe clusters into Si barriers not only suppresses the SiGe cluster PL but also contributes to carrier accumulation in the nanometer-thick Si-separating layers resulting in EHD condensation. The EHD condensation from the electron–hole–gas is a first-order phase transition, and it depends on the excess carrier concentration and temperature. According to our data (Fig. 2.33b), it takes ~10^{-6} s after the short laser pulse for the carrier concentration to drop below the EHD condensation threshold and for the reverse EHD/electron–hole–gas phase transition to take place. This reverse phase transition creates recombination conditions very similar to that at low excitation intensity. Thus, instead of waiting a long time to recombine with an electron, a hole located in the Si barrier would be quickly recaptured by the SiGe cluster, where it will eventually recombine with an electron in Si. This hole capture by SiGe energy wells explains the observed rise (after ~10^{-6} s) in the SiGe PL intensity (Fig. 2.33b). The longer time (>10^{-6} s) part of the SiGe PL decay is governed by the continuously decreasing SiGe cluster carrier concentration.

SiGe Light Emitters: Integration Issues and Compatibility with CMOS

Besides the challenges associated with Si/SiGe nanostructure luminescence efficiency, speed, and thermal quenching, another critical question is their potential compatibility with the standard CMOS process. Fortunately, Si/SiGe technology, specifically SiGe heterojunction bipolar transistor (HBT) technology, is quite mature, and full SiGe heterostructure BiCMOS production lines were introduced in the late 1990s [69]. At that time, the most important issue in the HBT device design was not the individual device performance but rather the integration of the SiGe transistor into the standard silicon process with the minimum of additional cost. It was anticipated and later proved that the major increase in cost was the addition of SiGe epitaxy to the existing CMOS process: Several estimates showed that the cost increase was close to ~10%. Numerous new applications of SiGe transistors have emerged almost immediately, including advanced analog-to-digital converters and radiofrequency (RF) devices for local and world area networks. Indeed, the relatively recent explosive growth of wireless systems at the marketplace is proving that this investment provides a very high return. Assuming that the technical problems with Si/SiGe light emitters can be solved, can "opto-CMOS" repeat the success of BiCMOS and RF-CMOS? This question has been discussed in several early publications and in more recent papers including some published in a recent special issue of the Proceedings of the IEEE [108]. In early publications, the ambitious goal of creating a new computational environment is well summarized in Fig. 2.35, where Si/SiGe nanostructures are used in a variety of "high-end" electronic and photonic devices including unspecified quantum devices, resonant tunnel

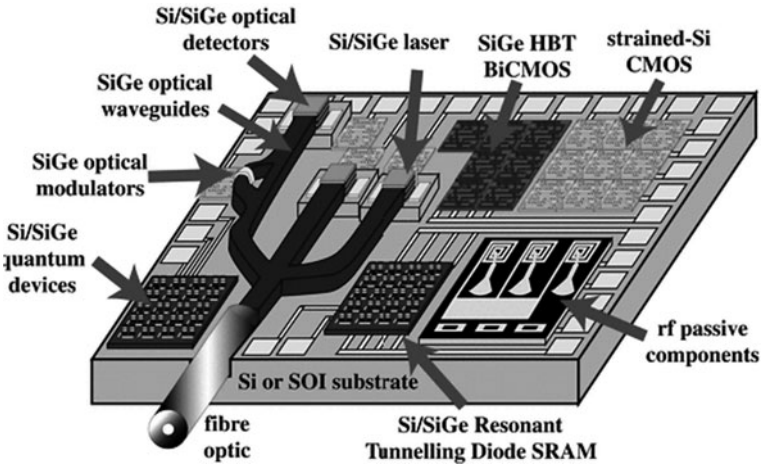


Fig. 2.35 Schematic of the suggested system-on-a-chip incorporating Si/SiGe nanoheterostructures with new functionalities as opposed to those of traditional CMOS. (Reprinted with permission from [69], Copyright 2004, Institute of Physics)

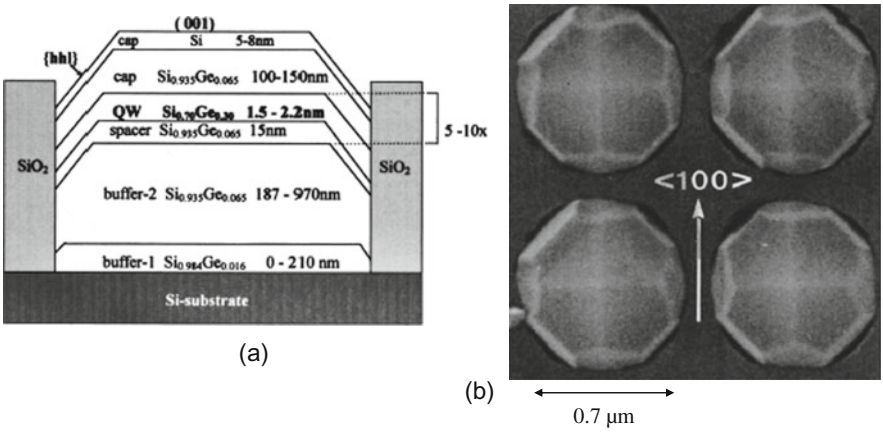


Fig. 2.36 (**a**) Schematic and (**b**) SEM view of selectively grown SiGe light emitters with the specified dimensions. (Reprinted with permission from [109], Copyright 1997, American Institute of Physics)

diode memories, high-index waveguides, modulators detectors, and even lasers [69]. This revolutionary approach will most likely require an enormous investment in the fabline infrastructure with a possible payoff in the quite distant future. Another, more conservative (or "evolutionary") approach considers limited applications of integrated optical interconnects, mainly for on-chip communication involving processing and transmission of relatively high volumes of data [108]. Also, the choice of the SiGe light emitter design is not currently clear (planar, vertical, directly integrated into a high-index waveguide, etc.). Figure 2.36 shows a proposed device design based on selective epitaxial growth using a lithographically defined SiO_2 mask and emphasizes a large device surface area [109]. Will the extremely high cost of the "wafer real estate" allow this particular device geometry? The answer, most likely, is "no." Also, in order to be combined with a high-index waveguide and a high-speed photodetector with minimum possible losses, all devices should probably be grown within the same processing step, and that requirement creates additional technological problems discussed in several review papers [17, 110–113].

Conclusion

In conclusion, the basic light-emitting properties of SiGe nanostructures have been reviewed with a major focus on new developments in 3D Si/SiGe NSs. These NSs emit light at the technologically important 1.3–1.6 μm wavelength region, and the highest PL and EL quantum efficiency is found in SiGe clusters with a ~50% Ge composition near the cluster core. A high luminescence quantum efficiency is observed at low excitation intensity, and the PL quantum efficiency decreases as the excitation intensity increases, most likely due to competition with faster non-radiative Auger recombination. Using time-resolved PL measurements, it has

been found that within the broad PL band, the part of the PL spectra at higher photon energies exhibits a ~100 times faster radiative transition, which is explained by a proposed model where radiative recombination occurs between holes (occupying excited states in SiGe clusters and/or outer, i.e., Si-rich, portions of SiGe clusters) and electrons mostly localized at Si/SiGe interfaces. This fast PL is less susceptible to Auger-induced intensity saturation as the excitation intensity increases. In addition, this PL exhibits an improved intensity temperature dependence: the PL can easily be extended up to room temperature. It has been shown that the EL in Si/SiGe nanostructures is very similar to the PL in its physical nature and that it can also be extended up to room temperature. In 3D cluster morphology multilayer Si/SiGe nanostructures, an anti-correlation exists between the PL originating from SiGe clusters and the PL associated with EHDs localized within nanometer-thick Si-separating layers. Auger processes eject holes from the SiGe clusters and facilitate the exciton/EHD phase transition in the Si layers. The Si EHD lifetime is less than 10^{-7} s and at high excitation intensity this is the dominant recombination channel. These experimental observations suggest that by controlling and modifying the composition of Ge-rich SiGe clusters (and possibly other types of Si/SiGe 2D and 3D NSs) it is possible to fabricate a more efficient SiGe light-emitting device. Despite the previously discussed challenges associated with integration into the traditional CMOS environment, the proven compatibility between Si/SiGe and conventional CMOS technology suggests that this task is not an impossible one.

Acknowledgments We would like to acknowledge the invaluable contributions over a number of years from our many collaborators on this work and whose names are given in the references to our work in this chapter and thank especially B. Kamenev of NJIT, J.-M. Baribeau and X. Wu of NRC Canada, and T. Kamins of HP Laboratories. We acknowledge the partial financial support for this research provided by US National Science Foundation, Intel Corporation, Semiconductor Research Corporation, and Foundation at NJIT.

References

1. Agrawal, G.P.: Fiber-Optic Communication Systems, p. 580. Wiley-Interscience, New York (2002)
2. Senior, J.M.: Optical Fiber Communications: Principles and Practice, 3rd edn. p. 1128. Prentice Hall, Englewood Cliffs, NJ (2008)
3. Miller, B.A.D.: Rationale and challenges for optical interconnects to electronic chips. Proc. IEEE **88**(6), 728–749 (2000)
4. International Technology Roadmap for Semiconductors, 2007 Edition, Interconnect, http://www.itrs.net/Links/2007ITRS/2007_Chapters/2007_Interconnect.pdf
5. Horowitz, M., Yang, K.K.C., Sidiropoulos, S.: High-speed electrical signaling: Overview and limitations. IEEE Micro. **18**(1), 12–24 (1998)
6. Plant, D.V., Kirk, A.G.: Optical interconnects at the chip and board level: challenges and solutions. Proc. IEEE **88**, 806–818 (2000)
7. Plant, D.V., Venditti, M.B., Laprise, E., Faucher, J., Razavi, K., Chateauneuf, M., Kirk, A.G., Ahearn, J.S.: 256-channel bidirectional optical interconnect using VCSELs and photodiodes on CMOS. J. Lightwave Technol. **19**(8), 1093–1103 (2001)
8. Savage, N.: Linking with light [high-speed optical interconnects]. IEEE Spectrum **39**(8), 32–38 (2002)

9. Neff, J.A., Chen, C., McLaren, T., Mao, C.-C., Fedor, A., Berseth, W., Lee, Y.C., Morozov, V.: VCSEL/CMOS smart pixel arrays for free-space optical interconnects. Proceedings of the third international conference on massively parallel processing using optical interconnections, 1996, pp. 282–289. 27–29 Oct 1996
10. Neff, J.A.: Optical interconnects based on two-dimensional VCSEL arrays. Proceedings of 1st international conference on massively parallel processing using optical interconnections, IEEE Computer Society Press, Washington, DC, 202–212 April 1994
11. Wada, H., Kamijoh, T.: Room-temperature CW operation of InGaAsP lasers on Si fabricated by Wafer Bonding. IEEE Photon. Technol. Lett. **8**, 73–175 (1996)
12. Kromer, C., Sialm, G., Berger, C., Morf, T., Schmatz, M.L., Ellinger, F., Erni, D., Bona, G.-L., Jackel, H.: A 100-mW 4×10 Gb/s transceiver in 80-nm CMOS for high-density optical interconnects. Solid-State Circuits **40**(12), 2667–2679 (2005)
13. Sieg, R.M., Carlin, J.A., Boeckl, J.J., Ringel, S.A., Currie, M.T., Ting, S.M., Langdo, T.A., Taraschi, G., Fitzgerald, E.A., Keyes, B.M.: High minority-carrier lifetimes in GaAs grown on low-defect-density Ge/GeSi/Si substrates. Appl. Phys. Lett. **73**, 3111–3113 (1998)
14. Razeghi, M., Defour, M., Blondeau, R., Omnes, F., Maurel, P., Acher, O., Brillouet, F., C-Fan J.C., Salerno, J.: First cw operation of a $Ga_{0.25}In_{0.75}As_{0.5}P_{0.5}$-InP laser on a silicon substrate. Appl. Phys. Lett. **53**, 2389–2392 (1988)
15. Park, H., Fang, A., Kodama, S., Bowers, J.: Hybrid silicon evanescent laser fabricated with a silicon waveguide and III-V offset quantum wells. Opt. Express **13**, 9460–9464 (2005)
16. Fang, A.W., Park, H., Cohen, O., Jones, R., Paniccia, M.J., Bowers, J.E.: Electrically pumped hybrid AlGaInAs-silicon evanescent laser. Opt. Express **14**, 9203–9210 (2006)
17. Soref, R.A.: Silicon-based optoelectronics. Proc. IEEE **81**(12), 1687–1706 (1993)
18. Pavesi, L., Lockwood, D.J.: Silicon Photonics, pp. 397. Springer, Berlin (2004)
19. Chatterjee, A., Mongkolkachit, P., Bhuva, B., Verma, A.: All Si-based optical interconnect for interchip signal transmission. Photonic Technol. Lett. **15**(11), 1663–1665 (2003)
20. Kobrinsky, M.J., Block, B.A., Zheng, J.-F., Barnett, B.C., Mohammed, E., Reshotko, M., Robertson, F., List, S., Young, I., Cadien, K.: On-chip optical interconnects. Intel Technol. J. **8**(2), 129–141 (2004)
21. Cullis, A.G., Canham, L.T.: Visible light emission due to quantum size effects in highly porous crystalline silicon. Nature **353**, 335–338 (1991)
22. Cullis, A.G., Canham, L.T., Calcott, J.D.P.: The structural and luminescence properties of porous silicon. J. Appl. Phys. **82**, 909–965 (1997)
23. Schuppler, S., Friedman, S.L., Marcus, M.A., Adler, D.L., Xie, Y.-H., Ross, F.M., Chabal, Y.J., Harris, T.D., Brus, L.E., Brown, W.L., Chaban, E.E., Szajowski, P.F., Christman, S.B., Citrin, P.H.: Size, shape, and composition of luminescent species in oxidized Si nanocrystals and H-passivated porous Si. Phys. Rev. B **52**, 4910–4925 (1995)
24. Kanemitsu, Y.: Light emission from porous silicon and related materials. Phys. Rep. **263**, 1–91 (1995)
25. Lu, Z.H., Lockwood, D.J., Baribeau, J.-M., Quantum confinement and light emission in SiO_2/Si superlattices. Nature **378**, 258–260 (1995)
26. Hirschman, K.D., Tsybeskov, L., Duttagupta, S.P., Fauchet, P.M.: Silicon-based visible light-emitting devices integrated into microelectronic circuits. Nature **384**, 338–341 (1996)
27. Tsybeskov, L., Hirschman, K.D., Duttagupta, S.P., Zacharias, M., Fauchet, P.M., McCaffrey, J.P., Lockwood, D.J.: Nanocrystalline-silicon superlattices produced by controlled recrystallization. Appl. Phys. Lett. **72**, 43–45 (1998)
28. Grom, G.F., Lockwood, D.J., McCaffrey, J.P., Labbe, H.J., Fauchet, P.M., White, B., Diener, J., Kovalev, D., Koch, F., Tsybeskov, L.: Ordering and self-organization in nanocrystalline silicon. Nature **407**, 358–361 (2000)
29. Tsybeskov, L. Lockwood, D.J.: Nanocrystalline silicon-silicon dioxide superlattices: structural and optical properties In: Efros, A.L., Lockwood, D.J., Tsybeskov, L. (eds.) Semiconductor Nanocrystals: From Basic Principles to Applications, pp. 209–229. Kluwer Academics/Plenum Publishers, New York (2003)

30. Pavesi, L., Dal Negro, L., Mazzoleni, C., Franzo, G., Priolo, F.: Optical gain in silicon nanocrystals. Nature **408**, 440–444 (2000)
31. Coffa, S., Franzò, G., Priolo, F.: High efficiency and fast modulation of Er-doped light emitting Si diodes. Appl. Phys. Lett. **69**, 2077–2079 (1996)
32. Fukatsu, S., Usami, N., Shiraki, Y., Nishida, A., Nakagawa, K.: High-temperature operation of strained $Si_{0.65}Ge_{0.35}$/Si(111) p-type multiple-quantum-well light-emitting diode grown by solid source Si molecular-beam epitaxy. Appl. Phys. Lett. **63**, 967–969 (1993)
33. Apetz, R., Vescan, L., Hartmann, A., Dieker, C., Luth, H.: Photoluminescence and electroluminescence of SiGe dots fabricated by island growth. Appl. Phys. Lett. **66**, 445–447 (1995)
34. Houghton, D.C., Noël, J.-P., Rowell, N.L.: Electroluminescence and photoluminescence from $Si_{1-x}Ge_x$ alloys grown on (100) silicon by molecular beam epitaxy. Mater. Sci. Eng. B **9**(1–3), 237–244 (1991)
35. Leong, D., Harry, M., Reeson, K.J. Homewood, K.P.: A silicon/iron-disilicide light-emitting diode operating at a wavelength of 1.5 μm. Nature **387**, 686–688 (1997)
36. Haynes, J.R. Briggs, H.B., Radiation produced in germanium and silicon by electron-hole recombination. Phys. Rev. **86**, 647–649 (1952)
37. van Roosbroeck, W., Shockley, W.: Photon-radiative recombination of electrons and holes in germanium. Phys. Rev. **94**, 1558–1560 (1954)
38. Haynes, J.R., Lax, M., Flood, W.F.: Analysis of intrinsic recombination radiation from silicon and germanium. J. Phys. Chem. Solids **8**, 392–396 (1959)
39. Cuthbert, J.D.: Recombination kinetics of excitonic molecules and free excitons in intrinsic silicon. Phys. Rev. B **1**, 1552–1557 (1970)
40. Pilkuhin, M.H.: Non-radiative recombination and luminescence in silicon. J. Luminescence **18–19**(1), 81–87 (1979)
41. Muss, D.R.: Injection luminescence in germanium. J. Appl. Phys. **35**, 3529 (1964)
42. Debay, G., Kolzer, J., Fundamentals of light emission from silicon devices. Semicond. Sci. Technol. **9**, 1017–1032 (1994)
43. Yablonovitch, E., Allara, D.L., Chang, C.C., Gmitter, T., Bright, T.B.: Unusually low surface-recombination velocity on silicon and germanium surfaces. Phys. Rev. Lett. **57**, 249–252 (1986)
44. Trupke, T., Zhao, J., Wang, A., Corkish, R., Green, M.A.: Very efficient light emission from bulk crystalline silicon. Appl. Phys. Lett. **82**, 2996–2998 (2003)
45. Zelsmann, M., Picard, E., Charvolin, T., Hadji, E., Heitzmann, M., Dal′zotto, B., Nier, M.E., Seassal, C., Rojo-Romeo, P., Letartre, X.: Seventy-fold enhancement of light extraction from a defectless photonic crystal made on silicon-on-insulator. Appl. Phys. Lett. **83**, 2542–2544 (2003)
46. Davies, G.: The optical properties of luminescence centres in silicon. Phys. Rep. **176**(3–4), 83–188 (1989)
47. Wilson, W.L., Szajowski, P.F., Brus, L.E.: Quantum confinement in size-selected, surface-oxidized silicon nanocrystals. Science **262**(5137), 1242–1244 (1993)
48. Read, A.J., Needs, R.J., Nash, K.J., Canham, L.T., Calcott, J.D.P., Qteish, A.: First-principles calculations of the electronic properties of silicon quantum wires. Phys. Rev. Lett. **69**, 1232–1235 (1992)
49. Zhao, X., Wei, C.M., Yang, L., Chou, M.Y.: Quantum confinement and electronic properties of silicon nanowires. Phys. Rev. Lett. **92**, 236805–236808 (2004)
50. Jeffries, C.D.: Electron-hole condensation in semiconductors. Science **189**(4207), 955–964 (1975)
51. Benoît à la Guillaume, C., Voos, M., Salvan, F.: Condensation of free excitons into electron-hole drops in pure germanium. Phys. Rev. B **5**, 3079–3087 (1972)
52. Thomas, G.A., Phillips, T.G., Rice, T.M., Hensel, J.C.: Temperature-dependent luminescence from the electron-hole liquid in Ge. Phys. Rev. Lett. **31**, 386–389 (1973)

53. Brinkman, W. F., Rice, T.M.: Electron-hole liquids in semiconductors. Phys. Rev. B **7**, 1508–1523 (1973)
54. Cuthbert, J. D.: Recombination kinetics of excitonic molecules and free excitons in intrinsic silicon. Phys. Rev. B **1**, 1552–1557 (1970)
55. Patel, N.K.C.: Stimulated effects in the radiative recombination from electron-hole liquid in semiconductors. Phys. Rev. **29**, 366–369 (1972)
56. Lee, E.-K., Lockwood, D.J., Baribeau, J.-M., Bratkovsky, A.M., Kamins, T.I., Tsybeskov, L.: Photoluminescence dynamics and auger fountain in three-dimensional Si/SiGe multilayer nanostructures. Phys. Rev. B **79**, 233307 (2009); Phys. Rev. B **80**, 049904 (2009)
57. Tajima, M., Ibuka, S.: Luminescence due to electron-hole condensation in silicon-on-insulator. J. Appl. Phys. **84**, 2224–2228 (1998)
58. See http://www.ioffe.rssi.ru/SVA/NSM/Semicond/Si/index.html
59. Tsang, J.C., Kash, J.A.: Picosecond hot electron light emission from submicron complementary metal-oxide-semiconductor circuits. Appl. Phys. Lett. **70**, 889–891 (1997)
60. Askerov, B.: Electron transport phenomena in semiconductors. World Sci., 389 (1994)
61. Klingenstein, W., Schmid, W.: Recombination of donor bound excitons in germanium. Phys. Rev. B **20**(8), 3285–3291 (1979)
62. Elliott, B.J., Gunn, J.B., McGroddy, J.C.: Bulk negative differential conductivity and traveling domains in n-type germanium. Appl. Phys. Lett. **11**(8), 253–254 (1967)
63. Wagner, J., Virla, L.: Radiative recombination in heavily doped p-type germanium. Phys. Rev. B. **30**(12), 7030–7036 (1984)
64. Auston, D.H., Shank, C.V., LeFur, P.: Picosecond optical measurements of band-to-band auger recombination of high-density plasmas in germanium. Phys. Rev. Lett. **35**, 1022–1025 (1975)
65. Othonos, A.: Probing ultrafast carrier and phonon dynamics in semiconductors. J. Appl. Phys. **83**, 1789–1830 (1998)
66. Weber, J., Alonso, M.I., Near-band-gap photoluminescence of Si-Ge alloys. Phys. Rev. B **40**(8), 5683–5693 (1989)
67. Sturm, J.C., Manoharan, H., Lenchyshyn, L.C., Thewalt, W.L.M., Rowell, N.L., Noël, J.-P., Houghton, D.C.: Well-resolved band-edge photoluminescence of excitons confined in strained $Si_{1-x}Ge_x$ quantum wells. Phys. Rev. Lett. **66**, 1362–1365 (1991)
68. Lenchyshyn, L.C., Thewalt, W.L.M., Sturm, J.C., Schwartz, P.V., Prinz, E.J., Rowell, N.L., Noël, J.-P., Houghton, D.C.: High quantum efficiency photoluminescence from localized excitons in $Si_{1-x}Ge_x$. Appl. Phys. Lett. **60**, 3174–3176 (1992)
69. Paul, D.J.: Si/SiGe heterostructures: from material and physics to devices and circuits. Semicond. Sci. Technol. **19**(10), R75–R108 (2004)
70. Robbins, D.J., Canham, L.T., Barnett, S.J., Pitt, A.D., Calcott, P.: Near-band-gap photoluminescence from pseudomorphic $Si_{1-x}Ge_x$ single layers on silicon. J. Appl. Phys. **71**, 1407–1414 (1992)
71. Houghton, D.C., Aers, G.C., Eric Yang, S.R., Wang, E., Rowell, N.L.: Type I Band Alignment in $Si_{1-x}Ge_x$/Si\(001\) quantum wells: photoluminescence under applied [111] and [100] uniaxial stress. Phys. Rev. Lett. **75**, 866–869 (1995)
72. Thewalt, W.L.M., Harrison, D.A., Reinhart, C.F., Wolk, J.A., Lafontaine, H.: Type II band alignment in $Si_{1-x}Ge_x$/Si(001) quantum wells: the ubiquitous type I luminescence results from band bending. Phys. Rev. Lett. **79**, 269–272 (1997)
73. Shiraki, Y., Sakai, A.: Fabrication technology of SiGe hetero-structures and their properties. Surface Sci. Rep. **59**, 153–207 (2005)
74. Savage, D.E., Liu, F., Zielasek, V., Lagaly, M.G.: Fundamental mechanisms of film growth. In: Hull, R., Bean, J.C. (eds.) Germanium Silicon: Growth and Materials, Semiconductor and Semimetals, vol. 56, pp. 49–96. Academic, New York, NY (1999)
75. Baribeau, J.-M., Pascual, R., Saimoto, S.: Interdiffusion and strain relaxation in $(Si_mGe_n)_p$ superlattices. Appl. Phys. Lett. **57**, 1502–1504 (1990)

76. Eaglesham, D.J., Cerullo, M.: Dislocation-free Stranski-Krastanow growth of Ge on Si(100). Phys. Rev. Lett. **64**, 1943–1946 (1990)
77. Mo, Y.-W., Savage, D.E., Swartzentruber, B.S., Lagally, M.G.: Kinetic pathway in Stranski-Krastanov growth of Ge on Si(001). Phys. Rev. Lett. **65**, 1020–1023 (1990)
78. Jesson, D.E., Pennycook, S.J., Tischler, J.Z., Budai, J.D., Baribeau, J.-M., Houghton, D.C.: Interplay between evolving surface morphology, atomic-scale growth modes, and ordering during Si_xGe_{1-x}epitaxy. Phys. Rev. Lett. **70**, 2293–2296 (1993)
79. Kamins, T.I., Carr, E.C., Williams, R.S., Rosner, S.J.: Deposition of three-dimensional Ge islands on Si(001) by chemical vapor deposition at atmospheric and reduced pressures. J. Appl. Phys. **81**, 211–219 (1997)
80. Baribeau, J.-M., Wu, X., Rowell, N.L., Lockwood, D.J.: Ge Dots and nanostructures grown epitaxially on Si. J. Phys. Condens. Mat. **18**, R139–R174 (2006)
81. Schittenhelm, P., Gail, M., Brunner, J., Nützel, J. F. Abstreiter, G., Photoluminescence study of the crossover from two-dimensional to three-dimensional growth for Ge on Si(100). Appl. Phys. Lett. **67**, 1292–1294 (1995)
82. Apetz, R., Vescan, L., Hartmann, A., Dieker, C., Lüth, H.: Photoluminescence and electroluminescence of SiGe dots fabricated by island growth. Appl. Phys. Lett. **66**, 445–447 (1995)
83. Schmidt, O.G., Lange, C., Eberl, K.: Photoluminescence study of the initial stages of island formation for Ge pyramids/domes and hut clusters on Si(001). Appl. Phys. Lett. **75**, 1905–1907 (1999)
84. Kamenev, B.V., Tsybeskov, L., Baribeau, J.-M., Lockwood, D.J.: Photoluminescence and Raman scattering in three-dimensional $Si/Si_{1-x}Ge_x$ nanostructures. Appl. Phys. Lett. **84**, 1293–1295 (2004)
85. Van de Walle, C.G., Martin, R.M.: Theoretical calculations of heterojunction discontinuities in Si/Ge system. Phys. Rev. B **34**, 5621–5634 (1986)
86. Schittenhelm, P., Engel, C., Findeis, F., Abstreiter, G., Darhuber, A.A., Bauer, G., Kosogov, A.O., Werner, P.: Self-assembled Ge dots: Growth, characterization, ordering, and applications. J. Vac. Sci. Technol. B **16**, 1575–1581 (1998)
87. El Kurdi, M., Sauvage, S., Fishman, G., Boucaud, P.: Band-edge alignment of SiGe/Si quantum wells and SiGe/Si self-assembled islands. Phys. Rev. B **73**, 195327–195336 (2006)
88. Kamenev, B.V., Tsybeskov, L., Baribeau, J.-M., Lockwood, D.J.: Coexistence of fast and slow luminescence in three-dimensional $Si/Si_{1-x}Ge_x$ nanostructures. Phys. Rev. B **72**, 193306–193309 (2005)
89. Baribeau, J.-M., Rowell, N.L., Lockwood, D.J.: Advances in the growth and characterization of Ge quantum dots and Islands. J. Mat. Res. **20**, 3278–3293 (2005)
90. Baribeau, J.-M., Wu, X., Lockwood, D.J.: Probing the composition of Ge dots and $Si/Si_{1-x}Ge_x$ Island superlattices. J. Vac. Sci. Technol. A **24**, 663–667 (2006)
91. Baribeau, J.-M., Wu, X., Picard, M.-J., Lockwood, D.J.: Characterization of coherent $Si_{1-x}Ge_x$ island superlattices on Si(100). In: Tsybeskov, L., Lockwood, D.J., Delerue, C., Ichikawa, M., van Buuren, A.W. (eds.) Group IV Semiconductor Nanostructures—2006, vol. 958, pp. 119–125. MRS, Pittsburgh, PA (2007)
92. Lockwood, D.J., Wu, X. Baribeau, J.-M.: Compositional redistribution in coherent $Si_{1-x}Ge_x$ islands on Si(100). IEEE Trans. Nanotech. **6**, 245–249 (2007)
93. Kamenev, B.V., Grebel, H., Tsybeskov, L., Kamins, T.I., Williams, R.S., Baribeau, J.-M., Lockwood, D.J.: Polarized Raman scattering and localized embedded strain in self-organized Si/Ge nanostructures. Appl. Phys. Lett. **83**, 5035–5037 (2003)
94. Henstrom, W.L., Liu, C.-P., Gibson, J.M., Kamins, T.I., Williams, R.S.: Dome-to-pyramid shape transition in Ge/Si islands due to strain relaxation by interdiffusion. Appl. Phys. Lett. **77**, 1623–1625 (2000)
95. Sunamura, H., Shiraki, Y., Fukatsu, S.: Growth mode transition and photoluminescence properties of $Si_{1-x}Ge_x$/Si quantum well structures with high Ge composition. Appl. Phys. Lett. **66**, 953–955 (1995)

96. Bozzo, S., Lazzari, J.-L., Bremond, G., Derrien, J.: Temperature and excitation power dependencies of the photoluminescence of planar and vertically self-organized $Si_{0.70}Ge_{0.30}$/Si strained superlattices. Thin Solid Films **380**, 130–133 (2000)
97. Schmidt, O.G., Eberl, K.: Multiple layers of self-assembled Ge/Si islands: Photoluminescence, strain fields, material interdiffusion, and island formation. Phys. Rev. B **61**, 13721–13729 (2000)
98. Wan, J., Jin, G.L., Jiang, Z.M., Luo, Y.H., Liu, J.L., Wang, K.L.: Band alignments and photon-induced carrier transfer from wetting layers to Ge islands grown on Si(001). Appl. Phys. Lett. **78**, 1763–1765 (2001)
99. Brunner, K.: Si/Ge nanostructures. Rep. Prog. Phys. **6**, 27–72 (2002)
100. Baier, T., Mantz, U., Thonke, K., Sauer, R., Schäffler, F., Herzog, H.-J.: Type-II band alignment in $Si/Si_{1-x}Ge_x$ quantum wells from photoluminescence line shifts due to optically induced band-bending effects: experiment and theory. Phys. Rev. B **50**, 15191–15196 (1994)
101. Hu, J., Xu, X.G., Stotz, H.A.J., Watkins, S.P., Curzon, A.E., Thewalt, W.L.M., Matine, N., Bolognesi, C.R.: Type II photoluminescence and conduction band offsets of GaAsSb/InGaAs and GaAsSb/InP heterostructures grown by metalorganic vapor phase epitaxy. Appl. Phys. Lett. **73**, 2799–2801 (1998)
102. Kamenev, B.V., Lee, E.-K., Chang, H.-Y., Han, H., Grebel, H., Tsybeskov, L., Kamins, T.I.: Excitation-dependent photoluminescence in Ge/Si Stranski-Krastanov nanostructures. Appl. Phys. Lett. **89**, 153106–153108 (2006)
103. Tilly, L.P., Mooney, P.M., Chu, J.O., LeGoues, F.K.: Near band-edge photoluminescence in relaxed $Si_{1-x}Ge_x$ layers. Appl. Phys. Lett. **67**, 2488–2490 (1995)
104. Qin, H., Holleitner, A.W., Eber, K., Blick, R.H.: Coherent superposition of photon- and phonon-assisted tunneling in coupled quantum dots. Phys. Rev. B **64**, 241302–241306 (2001)
105. Williams, C.J., Corbin, E., Jaros, M., Herbert, D.C.: Auger recombination in strained Si_xGe_{1-x}/Si superlattices. Physica B **254**(3–4), 240–248, (1998)
106. Stoffel, M., Denker, U., Schmidt, O.G.: Electroluminescence of self-assembled Ge hut clusters. Appl. Phys. Lett. **82**, 3236–3238 (2003)
107. Peng, Y.H., Hsu, C.-H., Kuan, C.H., Liu, C.W., Chen, P.S., Tsai, M.-J.: The evolution of electroluminescence in Ge quantum-dot diodes with the fold number. Appl. Phys. Lett. **85**, 6107–6109 (2006)
108. Special issue on Silicon Photonics of Proc. IEEE, **97**(7) (2009)
109. Vescan, L., Dieker, C., Souifi, A., Stoica, T.: Lateral confinement by low pressure chemical vapor deposition-based selective epitaxial growth of $Si_{1-x}Ge_x$/Si nanostructures. J. Appl. Phys. **81**, 6709–6715 (1997)
110. Wang, K.L., Karunasiri, G.P.R.: SiGe/Si electronics and optoelectronics. J. Vac. Sci. Technol. B **11**(3), 1159–1167 (1993)
111. Forbes, M., Gourlay, J., Desmulliez, M.: Optically interconnected electronic chips: a tutorial and review of the technology. Electron. Commun. Eng. J. **13**, 221–232 (2001)
112. Pinto, M.R.: Integrated communications microsystems. Proc. of 6th Int. C. Solid-State Integrated-Circuit Technol. **1**, 17–18 (2001)
113. Masini, G., Colace, L., Assanto, G.: Si based optoelectronics for communications. Mater. Sci. Eng. B **89**(1–3), 2–9 (2002)

Chapter 3
On Application of Plasmas in Nanotechnologies

Zoran Lj. Petrović, Paul Maguire, Marija Radmilović-Radjenović, Maja Radetić, Nevena Puač, Dragana Marić, Charles Mahony, and Gordana Malović

Abstract In this chapter we give a review of the application of non-equilibrium plasmas in the field of nanotechnologies, and nanotechnology-related science. The field of applications of plasma in general in nanotechnologies is extensive, and the field of non-equilibrium plasmas in the same context is almost as large with many possibilities covering both top-down and bottom-up fabrication approaches. Thus a single review cannot give it justice. We opted here, as may be expected, to give a review of possibilities in general, ranging from the growth and functionalization of aligned nanotubes, carbon walls and sheaths of graphene, through to the growth of gas-phase nanocrystals and dust particles; surface deposition of different structures, functionalization of the surface and conditions for developing means to prepare hyperhydrophobic surfaces; treatment of textiles, organic materials, living cells, to applications in nanoelectronics for the manufacture of future generations of integrated circuits to meet the Moore's law driven semiconductor roadmap. In a text like this it is not possible to cover all issues and even less to cover all sources. We thus focus on several topics and mostly on the plasma physics problems related to improving the plasma technologies. They include charging as a source of errors in integrated circuit fabrication to the generation of non-equilibrium plasmas at atmospheric pressure for even more convenient application and for applications in biomedicine. The message we hope to convey to all readers is an understanding as to the explicit advantages that non-equilibrium plasmas have in a large number of plasma-assisted nanotechnologies and other modern technologies—an advantage that other emerging technologies will struggle to match at the levels of integration, cost and quality required of future fabrication.

Introduction

Micro- and nanofabrication are important in many areas of modern economy, trade, science and technology and a lot of effort has been devoted to fabricating smaller

Z.Lj. Petrović (✉)
Institute of Physics, University of Belgrade, POB 68, 11080 Zemun, Serbia
e-mail: zoran@phy.bg.ac.yu

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_3,

and smaller patterns [1–4]. Nanotechnology (from the point of view of applications in electronics) seems to be nothing more than an extension of the micro-electronics technology, a step further towards miniaturization. On the other hand, nanotechnology is expected not only to reduce the dimensions of the existing microdevices, but, more importantly, to produce objects and materials with fundamentally different (new) properties and capabilities [5–7]. The expectations of what nanotechnology will achieve in the future range from quantum computers, ultra strong materials, self-cleaning fabrics and information storage and retrieval to drug delivery to individual cells and to implantable biosensors [8–14].

Nanotechnologies appear to be the current buzzword across a number of disciplines supported by policies of science funding agencies and editors of scientific journals. The field was certainly a novelty in science which, after a long and broadly focused initial period, bloomed when the numbers of scientists working in that field reached critical mass for a publication explosion. As the novelty is wearing off, the focus on research supporting the most promising applications is becoming a priority. This in turn increases the importance of plasma technologies related to the nanoscience and especially nanotechnologies.

These two are often identified although they differ considerably. While we could define nanoscience as the science about miniscule objects that belong in general to the nanometre world, the nanotechnology should consist of techniques that allow us to organize nanosized objects in predictable and eventually useful fashion. The nanotechnology also includes the relevant science to achieve these goals. There is a considerable difference between the two meanings. While plasmas could perhaps present a borderline case for belonging to the former group the case for belonging to the latter group is very strong, perhaps even essential. The situation is in a way similar to the micro (nano) electronic devices. While the physics of these devices does not require plasma physics, the main reason for their success lies in continuous miniaturization which not only increases the power of the existing devices and integrated systems but also opens new possibilities, products and markets. Some of the key processes in the miniaturization and production of integrated circuits are closely related to plasma physics (plasma etching, plasma cleaning, implantation). As a matter of fact we may regard the combination of photolithography and plasma etching as the standard technology for producing integrated circuits.

The connection of nanotechnologies and plasmas is even stronger as the discovery of some of the critical nano-objects occurred in plasmas [15–17]. Nevertheless, in order to justify the presence of plasmas within nanotechnology projects, conferences and strategies and priorities in science, one of the authors of this text had to resort to citing an old nursery rhyme about seven blind men who went to India to see an elephant. Each touched one part of an elephant and came to a conclusion about the elephant limited by his experience. None had seen the whole truth. In a similar fashion scientists tended to see nanotechnologies only through their own scientific interests and viewed other lines of research as less relevant or as a competition. Luckily, with the focus on practical applications being dictated by funding agencies, methods are sought that can bridge the gap between experiments with individual building blocks and massive, parallel production. In that respect plasmas

are viewed as one of the best possible methods to bring together nanoscience and nanotechnologies with industrial production.

This may be associated with the continuous progress in the standard technology of integrated circuits which, having mastered the 45 nm resolution, is progressing further to 32 nm and even higher resolutions (it has been reported that 32 nm production has begun this year). These technologies are developed as a complete set of processes to operate under industrial environment and to produce economically viable products. Thus the standard micro-nano electronic technology is a true nanotechnology developed all the way to the industrial level. More about that later on.

Two paradigms define the further development of nanotechnologies. The bottom-up approach is based on using the best of our abilities and facilities to produce some structures with extremely small dimensions, even at the level of single atoms, in fact to tease out *the self-building* properties of materials and structures. The aim is to use those structures as the building blocks of larger devices or materials. While that approach is based on major achievements in terms of human abilities, an even more challenging step will be to transfer this ability into industrially applicable and economically viable technology. For example, if we were able to develop a transistor that has a miniscule size equal to one molecule and we are able to put it on a surface in a precise position and prove its functionality we are still far from a useful technology. We need to put one billion transistors like that, in a short time, and to connect them in such a way that the whole system functions as a well-defined integrated circuit (IC). In order to keep up with the existing technology of integrated circuit production, all that should cost only a few hundred dollars per IC. It is thus obvious that the critical research in this approach will be not only in achieving a basic result but also in providing the technology for its practical implementation.

The second approach is to follow the path set out by Moore's law (more than 40 years ago) [18] and by the roadmaps outlined by the micro-electronic (now already nanoelectronic) industry. That is to carve the final product in a block of material where one may follow the existing and well-proven technologies and to allow them to achieve further miniaturization through research requiring a large degree of ingenuity. The latter approach is the strategy of the so-called top-down technologies.

While the bottom-up approach is in the domain from 1 nm to 10 nm, it is not yet well established in industrial applications. The top-down approach is at the moment at 45 nm (with recently introduced 32 nm production) in mass production while 22 nm technology is just round the corner. One is left with the research for future incarnations of miniaturized technologies. Those will, no doubt, require changes in the basic principles as well as in the materials that are used. On the other hand, one should be aware of the inertia of the industry whose favourite at the moment is definitely silicon substrate with the standard production technology.

Another issue is to maintain the growth of the associated industries that feed the research and further advances. While having in mind the holy grail of the modern applied research in physics is to maintain Moore's law, we can safely predict that most of the research that is carried under the banner of bottom-up nanotechnology

will actually yield applications in completely different, still unforeseen, areas. Yet both bottom-up and top-down technologies have to feed the support for the advances in electronics-related applications. One should thus bear in mind that the ICs of the near and not so near future will still be made by using the same principles that are implemented today and will include the two steps that are critical for further miniaturization: photolithography and plasma etching.

Whatever the approach, plasmas, in particular the low-temperature plasmas, will have a large role to play in bringing the technologies from the scientific laboratories to mass production. While that may be expected in the top-down approach the bottom-up achievements may also need plasma processing, in particular plasma growth of nanostructures on surfaces and in gas phase for achieving self-organized and functional assembly of nanostructures in industrial environment. In many ways plasmas are ideal for creating and allowing self-assembly of the building blocks in the bottom-up approach.

In this chapter we shall give a review of the application of plasmas in nanotechnologies. After briefly discussing the broader field we shall focus and give examples from the realm of our interests and experience. In other words, this would be just a report of one blind group that attempted to represent the big picture from its narrow and focused experience. Even the field of application of plasmas in nanotechnologies is too vast for one group of similarly minded people to represent. The need to inform a wider audience and the limited space mean that in some cases the coverage will be too brief and even trivial for specialists.

Introduction to Plasma Physics and Why Plasmas Are Useful

In a low-pressure gas, plasma is produced by the dissipation of electrical power to the gaseous medium. Electrons, to which most of the power is transferred, gain enough energy to initiate, in collisions with atoms and molecules, processes such as excitation, ionization and dissociation. Atoms, radicals and ions, which are produced this way, are at the origin of further reactions and, as a consequence, the plasma phase is generally a very complex mixture of chemical species [19].

Ultimately, the plasma species interact with the walls. Adsorption of species will occur depending on the chemical affinity and surface temperature. Adsorbed species may react with the surface to form a product or desorb without or after a surface reaction. If the product is volatile, it will desorb into the plasma phase and cause etching (removal) of the material from the surface. The material will either be eliminated through pumping or participate in the plasma chemistry. If the product is not volatile, it will contribute to the formation of a thin film at the surface. Ions, electrons and photons also interact with the surface. In plasma processing, the sample is, in general, negatively biased with respect to the plasma by means of the external power supply, so positive ions play a very important role in plasma surface interactions. Ions bring energy to the surface and, thus, can assist chemical reactions and desorption of weakly volatile species, or induce direct sputtering.

Plasmas are well defined in physics as gases with a large number of free electrons and ions that may be able to gain energy from the external electric fields and that satisfy certain criteria [20, 21]. One is that the size of the plasma should exceed other characteristic lengths such as the mean free path between collisions and the Debye length, the other is that they are quasi-neutral. In principle, plasmas maintain their prescribed electroneutrality by protecting themselves through generation of strong fields (in narrow volumes at their edge, known as sheaths) from the vessel that contains them (which is presumably a solid-state material that may be grounded if it is conducting). While one can often hear about non-neutral plasmas or even single-component plasmas, those are systems that are similar to plasmas (that follow the standard definition) due to strong coupling through Coulomb interaction of charged particles.

Ionized gases are not defined as narrowly as plasmas and merely are gases with free charged particles and without any specific conditions. The definition encompasses very weakly ionized gases that are also known as 'swarms' of charged particles (ensembles of independently moving charged particles driven by external field and collisions with the unperturbed molecules of the background gas), gas discharges (where charged particle densities may become sufficient to affect the distribution of the external electric field) and plasmas. As the charged particle density increases, properties of the ionized gas change dramatically. Two types of plasmas (and in broader sense ionized gases) may be defined. Those with lower densities of charged particles are known as the so-called non-equilibrium plasmas or low-temperature plasmas (sometimes even cold plasmas). When charged particle densities are higher usually one has equilibrium or thermal plasmas.

In non-equilibrium plasmas one uses the great mass ratio between electrons on one side and positive ions and background gas molecules on the other to decouple transfer of energy to electrons by electric field from the subsequent transfer of kinetic energy to ions and molecular particles. Thus electrons may achieve high energies and start dissociation and ionization while at the same time ions in the bulk of the plasma have low mean energies. This property of non-equilibrium plasmas is often described through the relation $T_e >> T_i = T_g$, where the respective temperatures of electrons, ions and gas molecules are a representation of their mean energies. Under the non-equilibrium conditions under discussion, the temperature of electrons, and even ions, are not well defined since the energy distribution function may be quite complex and different from a Maxwellian distribution, typical of gases [22]. Thus to sum up the intrinsic property of the non-equilibrium plasmas is that the energy from the external field may be transferred efficiently to electrons which initiate collisions leading to ionization to maintain plasma and dissociation which leads to a very reactive environment but that gas molecules and ions have low energies, which allows plasmas to treat thermally unstable materials including polymers, organic materials and even the living matter.

When charged particle densities become sufficiently high, Coulomb interaction couples electrons and ions more strongly and there is sufficient energy transfer between the two groups so that some of the energy gained by electrons is transferred to ions. Eventually their temperatures will become identical. At the same time ions

can transfer their momentum to the gas molecules so the energy eventually ends up in heating of the gas and the walls of the chamber. For the same given mean energy, thermal (equilibrium) plasmas are much more efficient in producing radicals (if the mean energy is sufficient). However, overall when total energy input is taken into account, non-equilibrium plasmas have a greater efficiency in using the energy as one may choose conditions so that electrons have a very high probability for a given process and at the same time no energy is wasted in heating of the gas and vessel.

Thermal plasmas are described by one parameter only, their temperature. Yet one could claim that only for the first 200,000 years of the universe did equilibrium plasma exist. We can, however, find numerous examples when plasma may be described by temperature albeit one that is variable across the plasma profile. In other words, those plasmas are not truly in equilibrium, and temperature is only useful in describing the properties locally. Spatial variation of the effective temperature reveals the degree of non-equilibrium. These temperature profiles are usually determined from spectroscopy, and to justify a claim for any degree of equilibrium, the temperature should be the same for all possible combinations of excited states (for the given position). If those temperatures, at a certain point, vary depending on the energy levels and processes, that is the sign of non-equilibrium. One usually does not calculate the plasma temperature from first principles; the temperature is a fitting parameter for a series of measurements.

One of the essential differences between equilibrium (thermal) and non-equilibrium (low-temperature) plasmas is in the fact that in the latter case individual collisional processes may change the electron energy distribution function (EEDF) considerably [21]. From a 'half empty glass perspective' we can consider this a problem; our optimistic 'half-full' interpretation is that we now have an opportunity to design the EEDF to fit the desired requirements. Adding a small amount of some gas with specific processes may achieve that goal, both by the shape of the cross sections and/or by getting the plasma to operate at a different E/N (electric field normalized to the gas number density usually given in units Townsend, 1 Td $= 10^{-21}$ V m^2).

Thermal plasmas, on the other hand, are to a large degree insensitive, even to the main buffer gas, as their properties would be similar for different gases, even for rare and molecular gases. These distinctions may certainly be observed in spectroscopy. There are plenty of plasmas where theory developed for equilibrium plasmas apply to the non-equilibrium case albeit with the use of an effective temperature that may be spatially dependent. In those plasmas the temperature is measured rather than derived from the energy balance. Spatially dependent effective temperature with local thermodynamic equilibrium (LTE) has all the properties of equilibrium plasmas even though their properties may not be uniform. Indication of a non-equilibrium nature would come from very different effective temperatures measured for different processes. For example, vibrational and rotational temperatures in most gas discharges may be few thousand degrees and few hundred degrees, respectively. This is an indication that temperature is a poor representation of all processes, even individual processes, and that EEDF may be quite different from a Maxwellian distribution.

Both low-temperature and thermal plasmas are being used in nanotechnologies, each with its own advantages and range of applications. Thermal plasmas were originally associated with the discovery of some of the key nanostructures, fullerenes and nanotubes. Discovery of the fullerenes [23] was made, in the 'one person's garbage is another person's Nobel prize' fashion, in the dust created by arc plasmas [24]. On the other hand, non-equilibrium discharges have been associated with anisotropic etching, as discovered empirically by Hosokawa in the early 1970 s [25], and these are becoming more and more important for nanotubes and nanocrystals since they provide a better control of properties of the created nanostructures.

Thermal plasmas are definitely more productive if large quantities are required, but at the same time they are much more costly in terms of energy invested as most of the energy goes to heat the chamber and the background gas. Non-equilibrium plasmas, on the other hand, may have a very efficient use of energy as one can have most of the energy used up by electrons. In addition, by tailoring the process one may even have a reasonably productive process or a very specific process with fine tuning that cannot be achieved in thermal plasmas. At the same time non-equilibrium plasmas, if used properly, have advantage for treatment of materials (and living tissues) that are sensitive to high temperatures.

Non-equilibrium plasmas are easily produced at low pressures (around 1 Torr) and under those conditions one may achieve large volumes, great uniformity and easily controllable properties. At high pressures, however, the ionization rate becomes quite high thereby producing large quantities of charged particles in a small distance crossed by charged particles. Thus thermal equilibrium plasmas are generally created. In order to achieve non-equilibrium operation of plasma at elevated pressures one needs to interrupt the charged particle production in some way and limit it. It can be achieved in very inhomogeneous fields where breakdown condition is satisfied only in a very small volume (corona discharges), by spatial interruption of the field with an inserted dielectric barrier where charges from plasma are deposited and may shield the field (dielectric barrier discharges), by high-frequency fields or localized microwave plasma production in a flow of rare gases mixing with the atmosphere or sometimes even in air. Finally, as the optimum gap for breakdown at 1 Torr is 1 cm, non-equilibrium operation at atmospheric pressure of around 760 Torr may be achieved with very small discharges with gaps of the order of 100 μm or less. Achieving flexibility and the ability to tailor the process and treat thermally sensitive targets, while still being able to operate at atmospheric pressure, is one of the holy grails of plasma physics, and all aspects of atmospheric pressure non-equilibrium plasmas are of primary interest for research and applications at the moment.

To some degree blinded by our interests and the need to cover mainly the technologies for electronic devices, we shall focus here on non-equilibrium plasmas, even though thermal plasmas are also of great interest for nanotechnologies. In addition, we shall only cover the technology itself briefly and devote more space to fundamental issues in plasma physics that need to be solved or optimized further. To summarize, one may view plasmas as a source of free radicals, charges

with different energies and ability to affect surface and basis and support for numerous gas-phase and surface chemical reactions. It can all be tuned and optimized but sometimes the parameter space is just too large and often empirical developments yield the best results. Nevertheless, with highly specific and technical requests of the current technological needs, a full understanding of plasma kinetics is required to optimize reactors.

Plasmas for Manufacturing of Integrated Circuits

Plasma Etching in Nanoelectronics

The physical basis for application of plasmas in IC manufacturing arises from the reactive ion etching (RIE) which was observed empirically by Hosokawa and coworkers in [25]. This empirical discovery led to significant advancement of microelectronics even before it was explained through direct experiment in 1979 [26]. In Fig. 3.1 we show a schematic representation of physical etching by Ar^+ ions and chemical etching by XeF_2 reactive molecules from [26]. The combined effects are at least 10 times faster as the two processes show a large degree of synergism.

Application of the non-equilibrium plasma requires that the mean energy of ions in the bulk of the plasma is low (close to thermal) and ions gain most of their energy in the sheath. Thus their trajectory is perpendicular to the surface of the wafer, allowing a high degree of anisotropy if the process is activated by high-energy ions. Properly chosen plasma chemistry will produce a large density of highly reactive radicals that cover all the surfaces and etch much faster if activated by high-energy ions. At the same time, selectivity of etching is achieved with a proper choice of

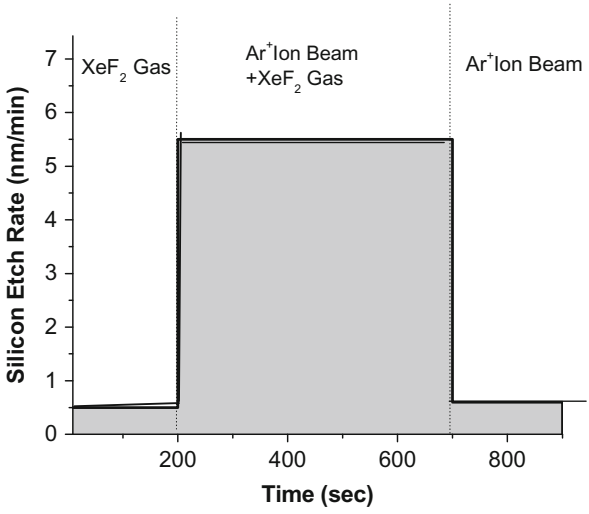


Fig. 3.1 Schematic representation of the synergism between reactive and physical etching by combined effects of XeF_2 reactive molecules and high-energy ions

radicals whereby a great difference in the etching rate exists between the mask (photoresist) and the substrate (poly-Si or SiO_2).

We may, however, state that the property of the non-equilibrium plasmas that the ions are at the room temperature while the electrons are at a much higher effective temperature (mean energy) is the basis for anisotropic etching that has enabled emergence of submicron technologies and has fuelled the growth of electronic industry for the past three decades according to the Moore's law [18]. Plasmas used for etching invariably operate at radio frequencies between 13.56 MHz and 200 MHz and in recent years two modes of operation have become dominant: capacitively coupled plasmas (CCP) and inductively coupled plasmas (ICP). Plasma sources may also be enhanced by application of external magnetic fields.

In principle, using reactive chemicals or thermal plasmas would lead to production of an isotropic etch as under equal temperatures directions and energies of all active particles, ions and radicals would be evenly (isotropically) distributed (see Fig. 3.2) over the surface. The isotropic etch is not sought as it reduces the resolution to a much larger value than that defined by photolithography and also it does not allow high aspect ratios (ratio of depth to width) of plasma-etched structures.

In addition, under those conditions the very high energies required to initiate efficient anisotropic etching of some materials such as SiO_2 are not available. As mentioned above and as shown in Fig. 3.3, in non-equilibrium plasma the radicals are distributed evenly while, in the high field of the sheath, electrons that start from the plasma with a relatively high energy slow down and barely reach the surface while ions are accelerated from the thermal (room temperature) energy and reach the surface with high directed energy and velocity and at right angle to the surface. High fields generated in the sheath and by the bias voltage add the possibility to adjust the energy to activate efficient etching in some materials. For example SiO_2 requires few hundreds of electron volts of energy per ion.

Etching is a much more complex process, involving redeposition, polymerization and charging of the dielectrics. In addition, ions and UV photons originating from the discharge deposit most of their energy deep inside the substrate which may lead to the damage of the substrate. Damage may also occur due to charging of

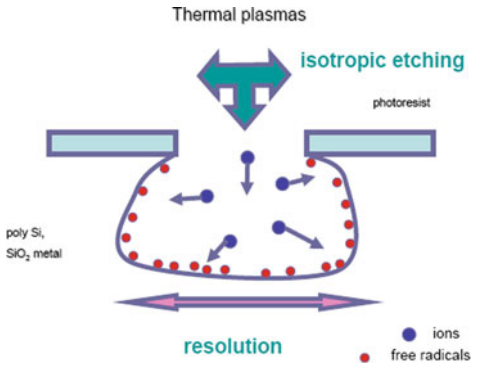


Fig. 3.2 Isotropic etch as a result of liquid chemical or thermal plasma with equal energies of all particles

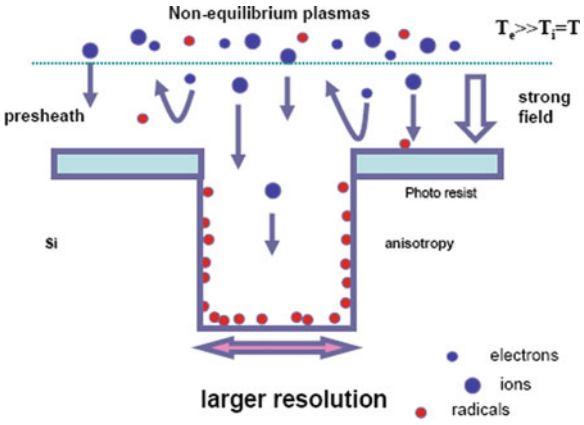


Fig. 3.3 Schematics of anisotropic plasma etching in non-equilibrium plasmas

the substrate especially in case of etching of the dielectrics as potentials develop inside the high aspect ratio structures that may divert trajectories of ions to side walls and even reflect the ions backwards. Problems caused by these processes will be discussed later on.

In addition to etching, plasmas are also versatile and lead to numerous related applications, such as cleaning, polymerization, thin-film deposition and ashing [27, 28].

What Are the Limits: Roadmaps

The industry defines goals for the next 10 years and those are given in well-defined roadmaps [29]. The idea that research may be defined so precisely seems far-fetched, especially if having in mind the complexity of the technology involved. Yet, as we have all witnessed in the past four decades, it has worked and, if anything, the progress has often been faster than previously expected. The plans are made for all aspects of the technology including etching and development of plasmas. Progress and plans for progress for dimensions of connection holes and gate lengths are shown in Fig. 3.4. The progress in employing plasmas for etching is also strongly tied to the progress in photolithography and in application of new materials. Improvements are required not only in reducing the resolution but also in increasing the speed of operation (thus changes in production of interconnects and materials involved may be required).

It has often been discussed just when the progress will depart from the Moore's law towards saturation. Of course, even in that case further improvements in processing power could be achieved by improved software and hardware design or by stacking the transistors in three dimensions (3D) provided that thermal dissipation issues can be resolved. And indeed, the thermal dissipation limitations may prove to be the critical issue to be solved in future years as the voltage that is involved cannot shrink as fast as the volume of transistors where the energy is being dissipated.

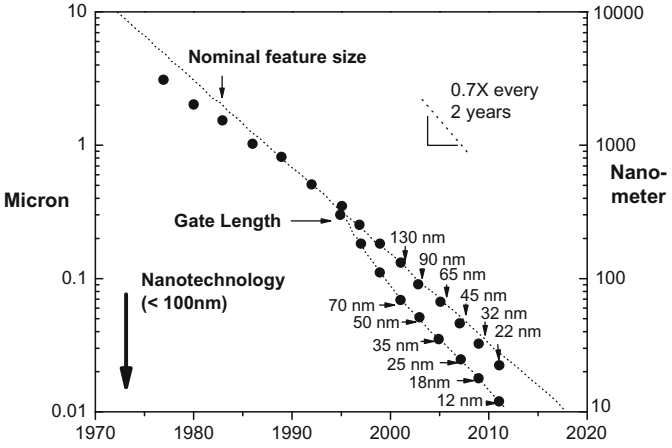


Fig. 3.4 Silicon technology reaches nanoscale

Another issue that is rarely mentioned is the problem that with shrinking sizes it is hard to see how far one can go in maintaining the properties of the bulk material in the small-size structures that are being developed. In addition to quantum tunnelling, one may be faced with the requirement for implantation of only a few atoms in order to achieve the required modification of the substrate. In that case, the location and statistical fluctuations of the numbers of implanted atoms may become an issue. In addition to those issues control of roughness of plasma-etched surfaces is going to be a limitation as it should not exceed 10–20% of the dimension of the etched structure.

Plasma Sources for the Standard Nanoelectronics Technology

Capacitively Coupled Plasmas

The capacitively coupled plasmas (CCPs) are mainly maintained by the input of energy to electrons from the moving sheath boundaries and by the effect of collisions on electron motion in RF field [21, 30, 31]. The electrons formed by bombardment of the instantaneous cathode dominate the maintenance of dc and low-frequency plasmas, but the so-called gamma electrons in RF plasmas contribute mainly at higher powers/current densities. It is possible to sustain plasma without their contributions. In the case where negative ions are the dominant negative particle, double-layer formation may become the most important source of ionization [32], while in the case of rare gas buffer gases, metastables may play a significant role in plasma maintenance [33].

Inductively Coupled Plasmas

Inductively coupled plasmas (ICPs) have also been used very much in recent times. Their main advantage is a high density of charged particles due to reduced losses

[34, 35]. As a result, radicals and excited states [36] are present at higher densities as compared to CCPs. This changes the properties of the plasma, facilitating ionization (by stepwise processes) and thereby reducing the mean electron energy [36]. ICPs allow special design of the coils that make it possible to achieve uniformity over large areas. On the other hand, complex geometry makes it more difficult to model ICPs unless simplifying assumptions are made. A very successful approach in this direction is the so-called 'non-local' model [37].

Typical issues that require further notice are the E–H transition and its understanding, the role of excited species and radicals, two-frequency operation and achievement of functional separation [38]. Nevertheless the ICP reactors seem to play a secondary but unavoidable role in the current plans of the semiconductor industry and IC manufactures.

Modelling Plasma Processing Devices

In order to be able to claim that one understands the basic kinetics of such complex systems as RF plasmas, one needs to achieve a very good qualitative and quantitative agreement between measurements and predictions of the models and such models may prove to be the basis for designing plasma reactor tools [32, 39], both CCPs and ICPs [40, 41]. Development of comprehensive tools for plasma modelling [42–45] is one of the goals which need to be completed in order to improve reactors for nanoelectronic applications. These modelling tools include the basic description of plasmas through either fluid equations coupled with calculated or measured swarm data or kinetic models. In fluid models some correction for non-local (non-hydrodynamic) behaviour of particles is required especially to describe the sheaths and regions of highly non-uniform field. The comprehensive set of data including electron molecule collisional cross sections and transport coefficient in the relevant energy range (for all energies covered by the plasma, atom and molecular collisional and reaction data for all gas molecules and their radicals and corresponding transport coefficients) is required for modelling. A number of both positive and negative ion species is usually generated and for each the cross sections and transport data are required (and usually lacking) again in the entire energy range covered by voltages applied to the plasma. In addition, parameters for surface collisions including reflection coefficients, accommodation coefficients, surface adsorption and desorption and related recombination coefficients, efficiency of inducing secondary electrons and other particles and etching and implantation rates are required for all particles. Finally, one needs to follow the kinetics of excited species, metastables and photons including their collisions with surfaces. When radicals and excited species are generated in large numbers, one needs cross sections for their collisions with active particles, mostly electrons, and their related chemical reactions. Finally, when dust particles are present their properties need to be included in the model as well. Comprehensive plasma models need to be developed to deal with complex geometries and also with external parameters such as temperatures, gas flow and loading and other issues.

With such developments the models have achieved a high degree of reliability with the ability to treat complex systems and provide a realistic representation of surface processes [46–48]. It has been possible for example to model development of the breakdown in the insulator of CMOS FETs during the process of manufacture due to charging of dielectrics and development of the corresponding potentials [49]. These problems are exclusively associated with nanodimension of the structures in new technologies where thickness of the oxide layer is only a few nanometres and size of the drain and source a couple of tens of nanometre. The ability to deal with issues such as charging and also to predict profiles of structures [49–51] seems to be essential for applying the same models to predict the properties and the development of nanostructures by bottom-up and top-down technologies.

These new challenges, however, require further improvements of the modelling and covering of some of the physics that has been poorly represented so far. Even in electron kinetics some assumptions are often made to simplify the models, that may fail to represent all aspects of important physical processes [52, 53]. Issues in basic development of the comprehensive models are the following: interpretation and accurate representation of fluid equations [54], interpretation of the transport data [53], inclusion of models relevant to describe negative differential conductivity [55], anomalous diffusion [56], the role of magnetic field in dc transport of electrons [57, 58] and resonant processes feeding the energy to electrons in complex electric and magnetic fields [59, 60]. In addition representation of time-dependent processes and spatially dependent processes not involving hydrodynamic approximation is required [53, 61]. Development of the new experiments is also needed to test complex time-dependent theories and provide the lacking data for fast neutral collisions and for negative (and also positive) ions. It is unlikely, however, that sufficient experimental data may be produced for electron radical collisions but even the small number of the available data [62] may be used as a benchmark for theories that could produce basis for modelling the role of radicals [63, 64].

Development of plasma technologies for nanomanufacturing will also open the need for the data for new gases that were not the subject of main interest of plasma modelling community, even the return to some of the standard gases that have a lot of data in the literature (N_2 and H_2 modelling of low-k solid-state electronics [65]).

Two-Frequency Plasmas, Pulsed Plasmas and Charging Damage

It seems that CCP reactors are the most interesting type of plasma reactors for the industry at the moment. Their operation under two frequencies and even three frequencies allows a good and functionally separated control of plasma maintenance (i.e. the flux) on one hand and biasing (i.e. the energy of ions hitting the surface) on the other [66, 67]. It was found that under some conditions the low-frequency power supply which was introduced mainly to control the energy of ions was able to change the properties of plasma significantly, even to turn it on and off. Thus in spite of being there to control the energy of ions, the flux and operating conditions could

change as well leading to a very complex and unpredictable pattern of operation. From the practical point of view one would prefer to have an independent control of the ion energy that does not affect the properties of the bulk plasma and flux of particles. In recent years some understanding of the pertinent physics of the multifrequency operation has been found although optimal choices still have not been made in the industry presumably due to existing patenting rights. It was found that optimal combination for CCPs was to have low frequency of the order of 1 MHz and high frequency of the order of 100 MHz.

Recent development of micro-technologies for micro-electro-mechanical systems (MEMS) and micro-optical electro-mechanical systems (MOEMS), as well as the reduction of critical dimensions in micro-electronics below 100 nm together with increased complexity of the processors requiring nine and more interconnect layers, has brought out the need to develop plasma processes able to etch small features with very high aspect ratio (10–100) [68]. Contacts between different levels and across the surface of the IC provide functionality to the IC. As a result of the need to produce very complex ICs and at the same time to achieve minimal dimensions, the connections have to be very narrow and deep, or in other words of a very high aspect ratio. Etching of the connections is the most difficult process determining the limits of the resolution and defining the technology.

Numerous 'aspect ratio' related problems may occur during the production of ICs. For example, due to the continuous reduction of the solid angle of the plasma from within the nanostructure (as the aspect ratio increases during the manufacture of the contact hole), the etch rate will be aspect ratio dependent and that is an undesirable effect. In addition, there are other associated issues, such as etch stop, notching and many more. Most of these problems have been strongly associated with charging of the bottom of the nanostructures [68, 69]. In principle we want to make a nanostructure of a very high aspect ratio in a dielectric. As the number of positive ions deposited at the bottom of the nanostructure increases, the potential at the bottom also increases and eventually it will become equal to the potential that corresponds to the energy of ions. At that point no ions will reach the bottom of the trench and further etching will stop. This effect is illustrated in Fig. 3.5 where

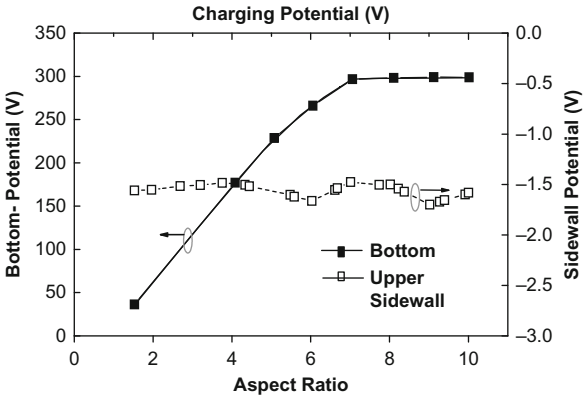


Fig. 3.5 Development of the charging potential (at the bottom of a trench and at the top of the side walls) as a function of the aspect ratio for structures in a dielectric etched (bombarded) by ions with 300 eV of the initial energy

we show the potentials at the bottom of a nanostructure [68, 69] as a function of the aspect ratio. It was found that for aspect ratios greater than 7 the potential at the bottom of the structure becomes equal to the potential corresponding to the ion energy. The potential at the side walls of the structure close to the top is small as it is defined by the energy of electrons reaching the side walls.

In modern IC manufacturing plasma charging damage is more prevalent and has grown into a significant field of its own [47, 68, 70–75]. In addition to the etch stop, other forms of damage exist. For example, charging of the bottoms of the wells leads to an increase of the potential on the dielectrics of the transistors and due to very thin dielectric layer even the smallest amount of charge may induce a high voltage and the resulting breakdown in the dielectric of the MOSFET transistor.

Understanding what causes the IC damage during the process of plasma etching necessarily requires both models and experiments but it may also provide a strategy to avoid charging-dependent and other problems. A solution to the charging in the high aspect ratio structures would be to have pulsed operation of plasma so that in the afterglow we may discharge the bottoms of nanostructures provided that such a relaxation of the accumulated charge is possible. After the collapse of the sheath electrons may be able to arrive to the surface and eventually reach the bottom of the structure. Temporal development of the charging potential gives the time constant of a few tens of milliseconds [68] for the potential to be developed. The real value of such prediction would be limited by the highly uncertain estimate of the charge losses in the bulk of the dielectric and on the surface. As the damage induced by charging through both etch stop and the breakdown of the very thin dielectrics is the principal cause of defects of ICs, techniques for avoiding these problems should be high in the list of priorities for further research. It will also be critical to reduce the charging damage if one wants to proceed to even smaller dimensions to true bottom-up plasma-aided nanotechnologies.

Experiments with pulsed plasmas gave an excellent improvement regarding the defects that may be associated with charging [76] where plasma sustaining source is pulsed while biasing voltage remains continuous. The principal limitation of the technique to discharge the dielectric by pulsing proved to be the poor conductivity of the dielectric. Makabe and coworkers have shown that in the afterglow during a short period of time, a double layer forms that may push negative ions into the nanotrench [77] which would improve neutralization of the deposited positive charges.

The influence of mask charging on profile evolution during the main etch has been considered [47]. Mask charging can influence profile evolution by directing ions to the side walls of the material thus causing bowing or leading to microtrenching. Figure 3.6 shows the electric field vectors on the profile surface in a dielectric for two different aspect ratios similar to those presented in [78].

The time necessary for the electric potential at the bottom of the feature to reach its steady-state value (Fig. 3.7) varies from less than 1 ms for the smallest aspect ratio up to 10 ms for the highest. These times combined with some additional techniques to remove the charges will determine the optimum pulsing periods that would reduce the charging problems.

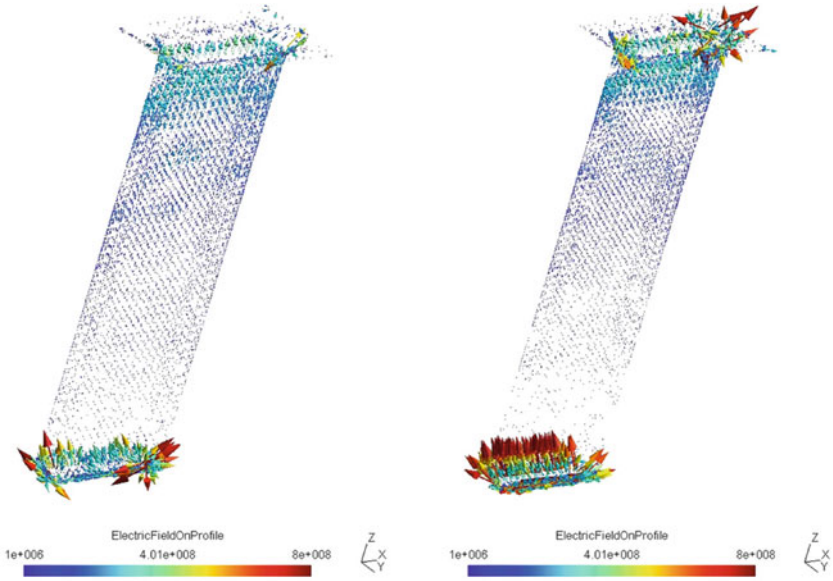


Fig. 3.6 Electric field at the profile surface for the aspect ratio of 4.5 (*left-hand side*) and 6 (*right-hand side*)

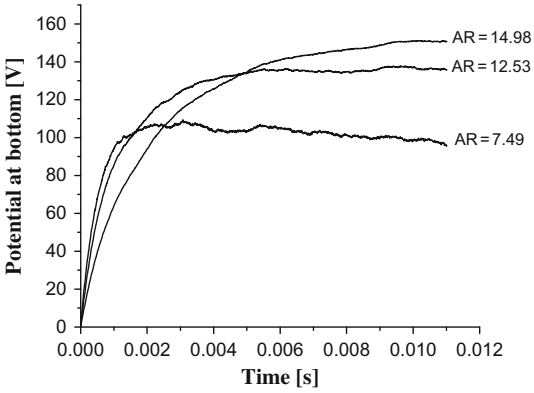


Fig. 3.7 Time dependence of the bottom potential for different aspect ratios (AR)

Fast Neutral Etching and Sources of Fast Neutrals

As the size of the opening decreases and as aspect ratios increase in order to accommodate more interconnect layers, it may prove that the pulsing strategy is not sufficient. A general solution to the problem of charging would be to use fast neutrals instead of ions. Those fast neutrals may be formed in the gas phase by charge transfer collisions or may be produced by surface neutralization (of both positive and negative ions). Contribution of hyperthermal neutrals with energies of

few electron volts to the etching has been recognized [79]. It has been proposed in 1991 that very high energy fast neutrals may be produced by the charge transfer in sheaths and may contribute significantly to plasma etching [80]. Subsequently, when charging damage became an issue in the manufacturing of ICs, it was proposed that fast neutral etching could solve these problems [81]. The idea of this proposal was to take advantage of charge transfer collisions to form a beam of fast neutrals and to apply a grid to stop the ions or to neutralize them even further. A similar proposal which relied more heavily on additional neutralization at surfaces of narrow tubes was given more recently [82, 83]. These techniques have been tested in production of ICs and were shown to reduce or even eliminate the damage. However, pulsing and other techniques were also sufficiently efficient for the same purpose and for current technologies (90–45 nm). Thus fast neutral etching has not been employed at the industrial level so far. It is, however, believed that it will have to be implemented for smaller dimensions as it also provides etched surfaces with smaller roughness. It was shown that 8 nm columns (quantum dots) could be made by using fast neutral beam etching [82]. This is a good indication that fast neutral beams may become essential in manufacturing of nanostructures smaller than 32 nm.

We have studied the efficiency of neutralization in a realistic system similar to that proposed by Samukawa and coworkers [82] by considering collisions in the gas phase along with the ion neutralization at the surface. Surface neutralization of ions and collisions of ions in the gas were first analyzed separately and then their joint effect was determined. Calculations [84] indicated that for grazing incidence conditions it is safe to assume that the neutralization efficiency is 100%. However, it was found that the effect of gas-phase charge transfer and surface collisions in converting ions into fast neutrals are approximately equal for the conditions of the system used by [82]. These calculations were coupled with a plasma model based on particle-in-cell (PIC) method [85].

Etching of Low-*k* Dielectrics

The need to increase the speed of the processors and memory units led to the need to reduce the effective dielectric constant of the dielectric that is used for interconnects. Transition from four to around two will suffice for the technologies that are developed now, but further improvements are required in order to meet the roadmaps beyond the end of this decade. Three approaches are considered: porous materials, low-*k* inorganic and low-*k* organic compounds [65].

New technologies for etching of such materials are required and one example would be etching of low-*k* organic materials where it was found that efficient anisotropic etching may be found in mixtures of nitrogen and hydrogen. While hydrogen provides etching even without special biasing, additional biasing increases the etch rate linearly. This implies that the etching process is a simple one-step process. On the other hand, nitrogen is required to provide anisotropy by passivation of the side walls. Thus the role of ions is to open the bottom of the trench to hydrogen radicals and also to provide additional treatment of the surface giving

effectively reactive ion beam etching but with different characteristics as opposed to SiO_2 etching by CF_4.

Including these processes into our code for simulation of the development of structures is required for low-*k* modelling and in addition simulations of the etching of porous structures would be of great interest.

Plasma as a Source of Nanomaterials

Nanomaterials

A critical element of nanotechnology is not only the discovery and creation of nanomaterials with advanced tailored properties but also their implementation in whatever form or substrate is most suitable for exploitation. Nanomaterials, so called because their dimensions (below 100 nm) lead to special properties or additional functionality over their macroscopic counterparts, present a major implementation and delivery challenge because of these dimensions. CMOS technology has, over the past decade, successfully delivered complex nanomaterial devices and much of this success has rested on the capabilities of plasma processing. We can be confident that, in the future, plasma technology will play an ever increasing role in the delivery of new nanomaterials and nanotechnologies. Nevertheless the exact nature of this role is difficult to predict. Plasmas offer exquisite control and versatility but are also considered complex and expensive. Their uptake will depend therefore on advances and requirements for both nanomaterials and plasma technology itself. Here we present a broad brush outline of current developments in nanomaterials with a few more specific examples to highlight the potential impact of plasma technology.

The term nanomaterials has been used to cover nanostructured materials with imposed nanoscale topography; nanoparticles and nanocomposites where the nanoparticles are embedded in a matrix, often a polymer; nanocapsules; nanoporous materials, e.g. membranes and scaffolds; nanofibres; fullerenes; nanowires; nanotubes, particularly carbon; dendrimers; quantum dots and thin films (<100 nm thick). Among the many synthesis techniques are self-assembly; chemical vapour deposition (CVD); chemical synthesis/solution methods (sol gel, colloidal chemistry); gas-phase methods (flame pyrolysis, electro-explosion, laser ablation); electrodeposition/electroplating; spin coating; spray coating; physical vapour deposition (PVD) including plasma techniques such as magnetron sputtering; plasma-enhanced CVD (PECVD).

The delivery or implementation of a given nanomaterial depends on the actual application and will vary from the production and handling of bulk quantities, for example in the addition of nanoparticles to sun cream, to the position-critical placement of, for example, nanowires and nanotubes as transistor and interconnect elements. The nature of the particular market is of course of critical importance in determining synthesis routes especially for volume markets where both cost and

quantity specifications are the challenge here. Nevertheless future markets for nanomaterials also include many high-value niches where simple bulk synthesis routes cannot deliver the required quality. Examples of high-value markets include displays ($90B): field emission displays, LCD backlighting with carbon nanotubes (CNTs), organic LEDs; photovoltaics ($50B): quantum dots (QD) for III–V solar cells; imaging ($10B): CMOS imager with plasmonics, IR imager with III–V quantum dots; lighting ($6B): LEDs with QD II–VI (ZnO) and high-index III–V nanostructures, II–VI QD organic LEDs; photonics: active fibre (amplifier) with nanoparticles; sensors ($4B): fluorescent markers with QD (II–VI), plasmonic biosensors, CNT-based biosensors.

Without any doubt a rapid development of nanotechnology was catalyzed by the discovery of carbon nanoclusters known as fullerenes [86, 87]. The term fullerene encompasses the entire class of closed-cage carbon clusters. Their applications include drug-delivery agents, fulleride-based superconductors and light-activated antimicrobial agents.

Carbon Nanotubes

Carbon nanotubes and associated nanostructures such as nanocones and nanowalls have been studied in great detail over the past decade due to their potential use in a large number of valued application areas. These range from transistors and interconnects for post-CMOS IC technology, field emission displays, high surface area electrodes in biosensors to incorporation in nanocomposites and bulk gas storage or filter materials. Originally derived from high-energy thermal (arc) plasmas, the synthesis routes have expanded to include laser ablation, chemical vapour deposition and PECVD in order to meet demanding specifications for both individual nanotube quality and, for example, positional growth. High-temperature arc and laser plasmas are under continued investigation for the production of bulk quantities of nanotubes but suffer the inclusion of high levels of metal and amorphous carbon impurities. Post-purification by chemical means is possible but can impact adversely on the nanotube quality. CVD synthesis, at high temperature (850–1100°C), is effective for producing moderate quantities at reasonable quality and more recently has been used to produce continuous fibres [88] of nanotubes held by van der Waals forces. Although CVD is simple and comparatively inexpensive, it is currently unable to produce high-specification carbon nanotubes, and the high temperatures are problematic for deposition onto many substrates, including semiconductors.

Plasmas used in PECVD cover a wide range of pressures and frequencies. dc, RF (CCP, capacitively coupled plasma, or ICP, inductively coupled plasma, which is preferred as it leads to less contamination) and microwave have been used [89–94]. Pressures of few hundred milliTorr are the standard for dc and RF plasmas but atmospheric pressure operation has also been tested.

Numerous issues are still open and worthy of scientific effort in relation to improvement of PECVD sources for CNT growth. These include the identification of the primary precursors and the role of other particles such as atomic hydrogen in

the growth, the role of ions in the growth and how it changes with the ion energy, the mechanism of alignment and how it is influenced by the electric field, how to optimize the substrate bias, how to achieve uniform growth over large areas compatible with IC wafers, how to control chirality of fibres and how to produce SWCNTs in PECVD production [89, 95].

Individual nanotubes are grown as single-wall (SWNT) or multiple concentric tubes (multi-wall MWNT) and quality parameters to be controlled are diameter, chirality, length and defects as well as incorporated impurities. They are grown vertically with density ranging from sparse to dense forests and here the quality of the vertical alignment and spacing between individual nanotube and nanotube bundles may need to be controlled. Applications include pillars of conducting nanotubes as IC vias or sensor probes and forests as biofluid filters, with field emission requiring relatively sparse growth to maximize the geometrical enhancement of electric field at the tip. Of course, there is a major drive to produce horizontally grown nanotubes for use as transistors or interconnects.

Plasma growth techniques for carbon nanotubes are dominated by microwave sources operated at high pressure in hydrogen- or nitrogen-diluted hydrocarbon gases (Fig. 3.8) [96]. Growth proceeds, in the majority of processes, from plasma sputter-deposited metal nanoparticles which often receive separate plasma pretreatment to enhance uniformity and surface activation prior to growth. Other plasma approaches include use of RF ICP and capacitive sources, although resultant quality is generally poorer.

The versatility of the structures that can be achieved from plasma growth is shown in Fig. 3.9. Here the density can be varied between sparse and dense forest with variations in microwave plasma-processing parameters and substrate conditions. Other structures such as pillars of nanotube bundles can be achieved by pre-patterned catalyst for biofluid filtering [97] while diamond spherules with high surface area can be selectively deposited along pre-patterned features [98].

For all nanotube synthesis methods, post-purification is required to remove the metallic remains of the catalyst which can be positioned at the top or bottom of each

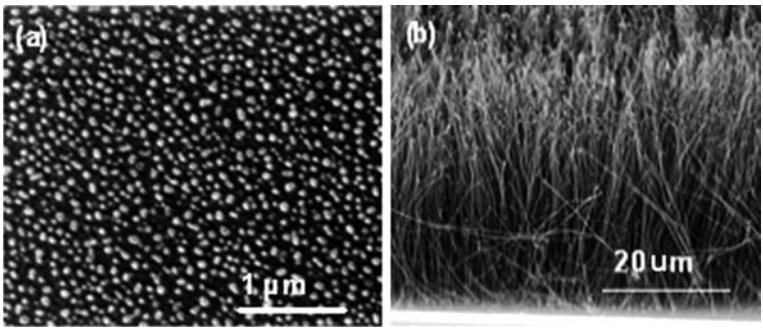


Fig. 3.8 (**a**) Nanoparticle islands formed from the metal catalyst layers and (**b**) carbon nanotube (CNT) growth at 750 C in microwave plasma (see [96])

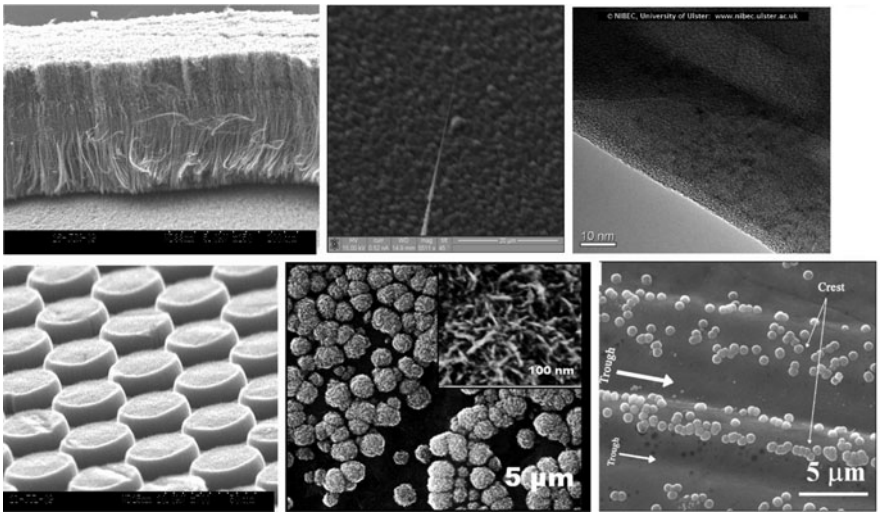


Fig. 3.9 *Top row from left*: dense vertical forest of carbon nanotubes deposited by microwave plasma-enhanced CVD, sparsely grown nanotubes and high-resolution TEM image. *Bottom row from left*: biofluid filter based on pillars of nanotube bundles, diamond spherules showing high surface area (*inset*) and selective deposition on copper tracks

grown nanotube, depending on the choice of deposition conditions. Also inefficient hydrocarbon conversion and/or hydrogen etching during growth leads to unwanted amorphous carbon phases which also have to be removed. Typical wet chemical techniques involve strong acid treatment, and increased defect density within the nanotubes has been observed after purification [99]. Plasma treatment to remove impurities shows considerable promise in this respect using ECR nitrogen plasmas (Fig. 3.10), while there are still issues regarding damage due to high bias using standard capacitive systems [100]. Nevertheless, exploration of the plasma

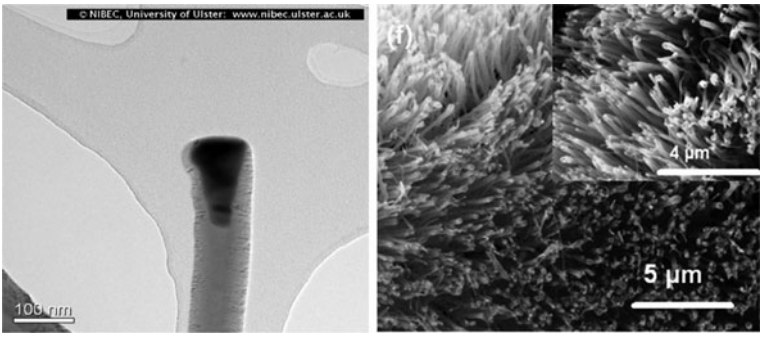


Fig. 3.10 TEM image of carbon nanotube with metal catalyst nanoparticle embedded in tip (*left*) and after plasma purification treatment (ECR, nitrogen) where the metal has been removed (*right*)

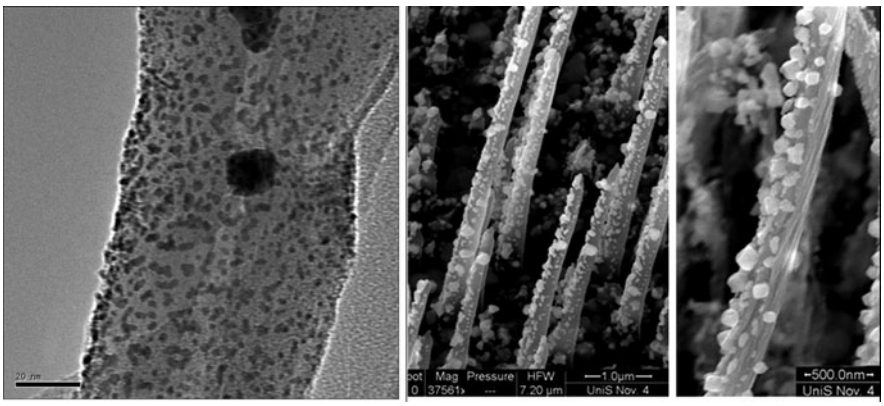


Fig. 3.11 Metal nanoparticle decoration of carbon nanotubes with Pt for fuel cells (*left*) and with AgCl, Ag for medical electrodes (*right*)

purification will undoubtedly lead to low damage solutions with systems capable of higher throughput than ECR and at lower cost.

After growth and purification, nanotubes can be further treated to enhance functionality. This functionalization step is often critical in device fabrication of, for example, biosensor electrodes or fuel cell elements (Fig. 3.11). Nitrogen functionalization (i.e. attachment or substitution of nitrogen) has been achieved in low-pressure RF plasmas [101] and at atmospheric pressure [102] as has oxidative functionalization [103] while more complex tailoring has also been achieved on nanotubes in a range of forms, from powder to vertically aligned forests [104]. More complex functionalization includes embedding metal nanoparticles onto nanotube surfaces, for example, to improve the surface contact area of thin-film medical electrodes [105] (Fig. 3.11).

Plasma processing, purification and functionalization are still at early stages. Plasmas will offer significant benefit once dedicated, and custom solutions (sources and processes) are developed for particular challenges in each market and application. The route to plasma success in CMOS fabrication technology provides an example. Here plasma source and process designs were constantly striving to meet an ever more challenging specification on an ever accelerating roadmap. With nanomaterials, once the current materials exploration phase has matured and priority exploitation paths determined, the process and equipment requirements can be more clearly defined. While most of the current research is carried out in general plasma systems, which draw on semiconductor processing (e.g. etching) for their design and operation, this is not necessarily appropriate for future nanomaterials processing. However, current developments in plasma science mean that sources can be tailored for precise ion energies or for selecting particular species with increasingly accurate control and can be made to operate across a much wider operating space.

While theories of carbon nanotube growth are still being developed, it is envisaged that once the carbon (and other) species dependencies are more fully

understood, then plasma sources can be more accurately designed to meet those exact requirements. This will provide reliable high-quality nanomaterials and hopefully will allow major advances in terms of significant reduction in process temperature, horizontal growth processes, high substrate sensitivity for positional deposition, catalyst reduction or even elimination and impurity rejection. Other major plasma technology developments in the pipeline include very large area plasma sources, e.g. for low-cost solar panels, and the intense research activity in (cold) atmospheric plasma materials deposition mean that existing barriers to plasma technology, namely the cost and complexity of vacuum processing, may be reduced or eliminated.

Thin-Film Nanomaterials

Thin-film processing has been around for a long time but thin films fall within the nanomaterials category when their function depends on thicknesses below 100 nm. In general, controlling the three key parameters of the surface chemical composition, thickness and topography allows the possibility of transferring specific or otherwise expensive properties to a nanometre-sized coating. This makes the list of potential applications almost endless with an impact in almost every industrial sector. Most thin-film applications are linked to developments in the semiconductors industry such as thin-film transistors (TFT); large-area displays (LCD, OLED); NEMS/MEMS; planar waveguides; magnetic RAM; among others. Here the use of plasma-based tools is common and, in many of these applications, plasma film deposition, functionalization and patterning are considered capital-intensive. One area that is less capital-intensive and therefore more attractive for new low-cost plasma designs is the creation of cheap, flexible, disposable electronic devices, for example RFID tags or smart packaging. Here low-temperature and high-pressure plasma designs will need to compete with less capable but also much less expensive competitor technologies. Other areas where high-quality nanocoating will become important are glass coatings, which could become ubiquitous; low-friction or wear-resistance coatings, which have almost endless applications from engine parts to mechanical or drilling tools; and even high-end applications such as hard disk manufacture; and solar cells, where nano-thin films could have a strong impact, and large returns. Here the design of efficient but very large area plasma tools is critical for enabling non-crystalline and hence low-cost silicon solar technology. Finally in the medical device sector, the creation of biocompatible coating materials and surface treatments is receiving considerable attention. These coatings can be for mechanical, e.g. low-friction hip joints, and/or biological ends, e.g. forming a barrier or anti-corrosion coating, eluting therapeutic agents, inhibiting or promoting specific cell growth. Whatever the application, providing a biocompatible and/or bioactive coating that remains viable within the body for a long time represents an extreme challenge and requires a costly and time-consuming approval process. Nevertheless the end result can have a dramatic impact on human health and there is considerable active research. The medical device industry has not embraced plasma technology

due to its perceived cost and the high degree of expertise required. Implantable medical devices are rarely planar in geometry. In fact they represent the most complex of 3D shapes from millimetre to many centimetre dimensions.

The most important challenges and bottlenecks in thin-film coatings are certainly price and production volumes. In general, most coating processes are intrinsically difficult to control (spray coating, PVD, spin coating, etc.) and if really precise control is needed there are probably limits to what can be obtained with these techniques. These drawbacks refer both to the size and (3D) shape of the substrates to be coated. Lack of a full understanding of the adhesion mechanisms between substrate and the coating as well as among multi-layers is another barrier. With chemical vapour deposition (CVD) the main barriers to success lie in the difficulty of correlation input parameters with the resulting properties which block the functionalization of the coatings to specific applications. Also the adhesion to the surface and process reproducibility, which is very dependent on the quality of the precursor material, can be problematic. As with most thin-film production methods, precise coating of 3D geometries is very difficult. The deposition rate (throughput per dollar) is considered low and the process requires very high operating temperatures. Plasma-enhanced CVD (PECVD) is considered in the same light but with the addition of expensive vacuum chambers and there is a perception that the CVD method, and hence the PECVD method, is too costly for large-area coatings and volume coatings. Whereas CVD processes, however, may not be flexible enough to achieve the required quality and are limited by the high temperature, the inclusion of the plasma-enhanced option, albeit with additional cost, may provide the only solution for many high-end applications. This is certainly the case for solar panels and here the issue of large-area coating (dollar per cm^2) is well on its way to solution with major suppliers operating 1 m^2 production plasmas and attempting to scale this further. Research into high-pressure RF glow discharge and cold atmospheric plasmas will be of major significance here. Another important technique, physical vapour deposition (PVD), includes plasma-based sputtering as well as, for example, thermal evaporation, laser ablation and electron beam techniques. There is a cost–quality issue here as in the case of CVD and PECVD, especially for large-area coating. Certainly beam and laser approaches are particularly expensive options while thermal evaporation, although suitable for low-cost large-area coating, cannot achieve highly functional coatings or, more importantly, sufficient control of thickness and adhesion control.

For all techniques, one main problem area that has yet to be seriously tackled is reliable and reproducible coating onto complex 3D shapes with uniform fidelity. Most deposition techniques are line of sight and in the case of plasmas, the size of the sheath, typically a few centimetres in low-pressure plasmas, effectively ‘smoothes’ its interaction with the 3D shape features. High-density plasma sources and highly collisional (elevated pressure) environments can theoretically ameliorate the line-of-sight effect and cause the plasma to faithfully follow the 3D shape at sufficient resolution to coat all important features evenly. However, high-pressure operation reduces the availability of energetic (ionic) species which are important for materials quality and adhesion. Again, recent materials research on atmospheric

plasma deposition has highlighted the potential advances to be gained here, although custom process solutions will probably be required for each material–substrate–shape set. Also current generalized plasma systems design strategy (e.g. parallel electrodes), which can be successfully adapted to multiple but planar processes, will probably not be as effective. At worst, each 3D part will need a customized plasma tool which coupled with the required plasma engineering expertise and custom processes implies a high cost per part, i.e. a shift from high to lower cost equipment but with additional design and process engineering costs in-house. However, the increasingly accurate design and process simulation tools for plasmas that are becoming available mean that these costs may be ameliorated and possibly open the full plasma technology option to small start-ups and SMEs, for lower volume applications. A key enabler in this respect would be a more detailed understanding of high-pressure deposition for a wider range of materials.

Medical Device Biocompatible Coatings

We can consider examples that illustrate the above points, especially with regard to 3D coating, coming from the implantable medical devices sector, which has a high added value and can therefore be able to consider a wider range of technology options. Current coatings technology is generally non-plasma; true long-life implantable coatings are very hard to achieve; and the impact of coating failure or degradation is very significant yet the introduction of customized plasma technology has still to overcome considerable inertia. As an example, coronary stents were originally constructed of bare metal which provoked an inflammation reaction (restenosis). A coatings strategy was pursued and this resulted in a winner combination of (electro-)spray-coated biocompatible polymer loaded with drugs. However, insertion and stent expansion within the body caused severe degradation of the polymer (Fig. 3.12) and considerable medical concern, leading to a re-evaluation of the procedure and the search for both new nanomaterials and new deposition techniques.

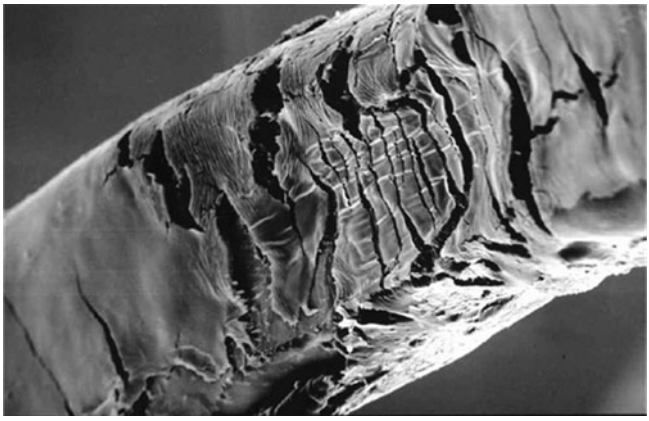


Fig. 3.12 Cracking and damage to polymer coating on stent after balloon expansion

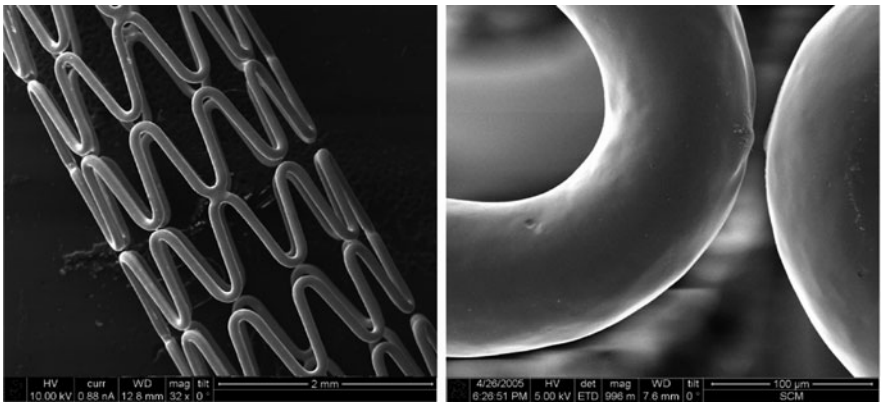


Fig. 3.13 Stent coated by a hard barrier film with graded elastic interface showing no cracking or delamination after expansion of the stent

Thin-film (<50 nm) stent barrier coatings with various levels of stress and cracking were deposited with plasmas with all the difficulties caused at the nanoscale by the non-uniform species bombardment. Those coatings displayed adhesion weakness on 3D surfaces [106, 107]. The final high-integrity (Fig. 3.13) solution used a complex multilayer film comprising an elastic interface, which can expand with the stent, and a hard corrosion impervious top layer. This technology can provide support for subsequent bioactive or drug-loaded layers which are stable with variable density cross-linking. However, there is still a considerable way to go in scaling these techniques to significant volume throughput.

From a 3D perspective, stents are hollow tubes about 2 mm diameter and up to a few centimetre in length while dental implants are small studs, typically less than 1 mm. At the other extreme, arterial guidewires are about 1.5 m long and 1–2 mm diameter mandril wound steel, while hip joints are large bulky items of irregular shape. The fundamental requirements for coating are that an electrical and a mechanical (holder) contact must be made. While these can be the same, no region can be left uncoated otherwise rapid corrosion and failure will occur within the body.

Finally we can consider the work on developing load-bearing anti-wear low-friction coatings for hip joints. High-hardness amorphous carbons developed for planar processes, e.g. hard disks, are ideal candidates but suffer considerable stress when deposited on non-planar surfaces even for extremely low thickness (<2 nm) [108]. Pulsed plasma systems [109] are under development for stress relief while pulsed plasma ion implantation related techniques (PIII) have been considered for biomedical coatings [110]. This operates by creating a plasma surrounding the part to be coated and then with the application of a high-voltage pulse, high-energy ions bombard the surface uniformly. Coupling this distributed ion bombardment source to gas-phase precursors may provide a cost-effective route to 3D coating of the desired materials.

Growth of Nanocrystals and Nanostructures on Surfaces and Nanowalls

Plasmas, in particular non-equilibrium low-temperature plasmas, have some advantages as compared to the thermal growth (high temperature or laser-induced pyrolysis) of nanocrystals or nanostructures. Non-equilibrium plasmas allow growth of structures that cannot grow under thermal, equilibrium conditions. The nanoparticles formed in plasmas are monodisperse and, in addition, such plasmas provide means for further functionalization, coating and change of properties of those particles. In plasmas we may grow structures on the surface or in the gas phase and we may use processes in the gas phase and on the surface to organize deposition of those particles on surfaces. It is rarely considered that the production in the bulk of a low-temperature plasma may serve as the main source of material that is to be collected and mass produced. For that purpose thermal plasmas were considered more regularly but recently low-temperature plasmas were shown to be very effective reactors for production of bulk materials.

In many aspects the growth of structures that are unlikely under thermal growth conditions is the key advantage in application of non-equilibrium plasmas to the production of nanocrystals. Perhaps the best known example is that of cubic nanocrystals of Si that have been grown by Kortshagen [111]. Non-equilibrium plasmas may be used for formation of crystals with covalent bonds while ionic bond based crystals may be grown very efficiently in liquids. Unipolar (usually negative due to a higher mobility of electrons compared to ions) charging of smaller clusters and particles favour crystallization (rather than agglomeration which usually leads to particles of a wide range of sizes) and formation of monocrystals, such as cubic Si crystals shown in Fig. 3.14 [112]. Charging also allows trapping of particles which reduces losses to the walls and also opens a possibility of self-organization of particles in the gas. Finally the negative particles form a sheath that produces a

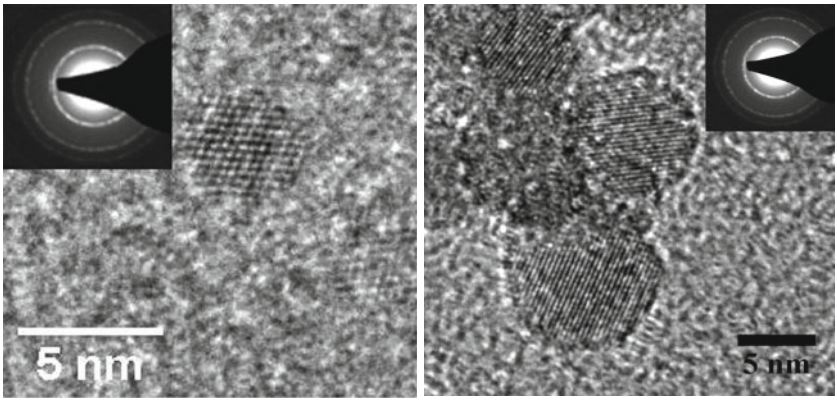


Fig. 3.14 Nanocrystals of Si grown in low-temperature plasma (see [112])

high field leading to acceleration of ions which, together with other energetic particles in plasma, allows localized heating of particle surfaces sufficient to produce crystals with covalent bonds while the surrounding gas is at the room temperature [111, 113].

One is then able to tune the electrical, mechanical, optical and thermal properties by adjusting the size, especially since the particles are monodisperse. For example doping may be performed and excellent nanosize transistors may be produced [114]. It has been shown that numerous applications may take advantage of these nanocrystals starting from nanoelectronic devices, luminescent markers for medicine, light-emitting devices and displays, including in particular the photovoltaic cells. Not only very high efficiency has been achieved but also a new technology for applying the photovoltaic cells has been developed whereby the nanoparticles are dissolved in an ink that may be applied to any surface [115].

Pyramidal growth on surfaces is also unlikely under thermal growth conditions and may be the result of some of the special plasma properties. For example it was shown that in growing of pyramidal CdO nanostructures the reactive oxygen atom flux from the plasma is a determining factor [116]. The effect, on the other hand, illustrates a synergism that may exist in development of nanotechnologies as the most sensitive catalytic sensor of oxygen atoms may be developed by using nanowires of Nb_2O_5 [117]. In a similar fashion carbon nanotips may be produced [118] by carefully selecting the plasma and other processing parameters.

Another example of how controlling the properties of plasmas may affect the deposited nanostructures is in the growth of single-crystalline Fe_2O_3 nanostructures. Narrow structures 50 nm in diameter and 1200 nm in length were grown in plasmas with a floating bias and much broader and tapered nanowires with average base diameters of 200–300 nm and lengths up to 1500 nm were grown for 150 V bias (and surface temperature of 585°C). The differences were related to microscopic flux distributions and how plasma properties affect them [119].

Finally we should discuss a new application of plasmas in the growth of graphene sheaths and nanowalls on surfaces and consisting of graphene. Using atmospheric pressure microwave plasma, Dato et al. [120] were able to generate single graphene sheaths. It is still uncertain in what way the optimization of the plasma source may proceed to improve the production yield. It is certain that low-pressure non-equilibrium plasmas would provide a wider range of controllable parameters and testing production under those conditions. On the other hand graphene-like structures may grow as walls deposited on the surface as may be seen in Fig. 3.15 (see also [121]). For that purpose PECVD was used with an additional flux of hydrogen atoms.

Dust Particles and Their Control

Over the past decade, extensive studies have been carried out on formation and behaviour of dust particles in non-equilibrium plasmas. Both dust particles and thin films of amorphous carbon having a cauliflower structure at the surface are

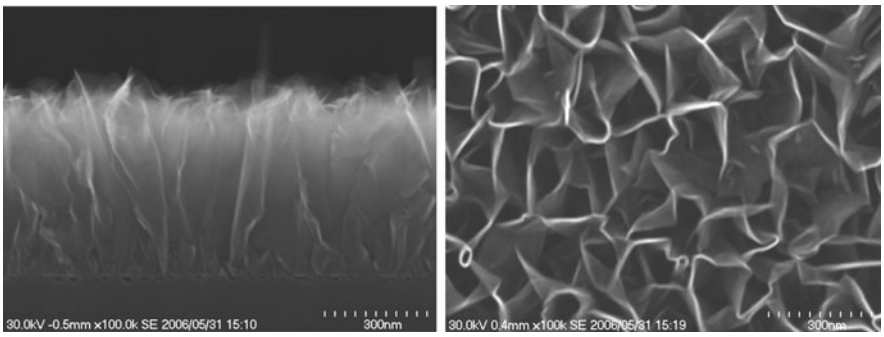


Fig. 3.15 Graphene nanowalls grown from plasma-enhanced CVD: end on (*left*) and side on (*right*) observation

readily created in RF plasmas in methane or acetylene or some other gaseous hydrocarbons [122, 123]. It was shown that charging of those particles leads to them being monodisperse. During their growth, conditions are reached when the particles become too heavy and cannot remain airborne. They then fall to the bottom of the vessel leaving the plasma suddenly void of negatively charged dust particles and at this point the plasma properties change abruptly (from the broader profile of Hα lines to greater electron density). The cloud of charged dust particles may become self-organized through Coulomb crystallization; voids may appear as well as complex structures. It is hoped that the fact that size may be controlled very accurately leads to efficient production, self-organization and then manipulation by electric fields of the particles which may then be deposited on the surface with regular geometric patterns.

As the balance of charges on the dust particles is the key to their behaviour we have focused on secondary electron yields of the dust particles material for the conditions of a gas breakdown [124]. We know very little about some of the basic properties of the particles, especially the secondary electron emission yield of the material that constitutes dust particles. Data exist for graphite but not for the carbon deposited from the plasma.

The dust particles were formed in capacitively coupled RF plasma through plasma polymerization of acetylene monomer diluted in argon [122, 124]. By controlling the discharge conditions, it was possible to deposit either thin polymer films or films with dust particles incorporated on the stainless steel plates (see Fig. 3.16a). The plates were then used as cathode surfaces in parallel-plate dc discharge, in order to determine secondary electron yields. The size of the dust particles was ~200 nm.

The dependence of the γ coefficient on the reduced electric field E/N is derived from the Paschen curves and profiles of emission for the breakdown conditions, for the three cathode surfaces (see Fig. 3.16b). Details of the procedure are given in [125]. The secondary yield for the polymer material is lower than that for the pure stainless steel. The main differences between the three materials occur at low E/N, where photon-induced electron emission dominates [125], i.e. the photoemission

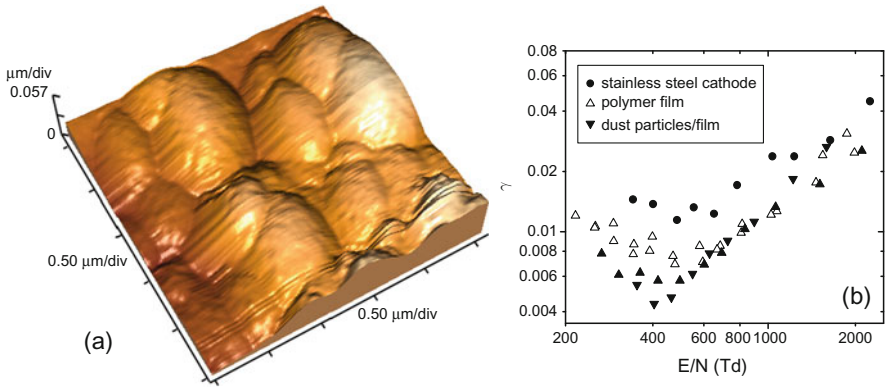


Fig. 3.16 (**a**) Cathode surface with dust particles incorporated into thin amorphous carbon film that is used as a cathode of a dc discharge. (**b**) Secondary electron emission yield γ as a function of the normalized electric field strength *E*/*N* for stainless steel cathode (*circles*), stainless steel covered by polymer film (*open triangles*) and by dust particles immersed into film (*solid triangles*)

is reduced for polymer surfaces, and even more for the surfaces with the particles incorporated in the thin film.

One might expect the particles at the surface to cause submicron roughness leading to a somewhat higher local field and through field emission to a more efficient secondary electron production. This is clearly not the case, at least not for the surface roughness covered here (~10–200 nm). The more complex structure of the film due to the presence of the particles will certainly lead to a lower value of the escape probability for electrons from the surface and consequently to lower yields. It is still not clear whether the difference between the metallic and the coated surfaces may be explained in the same fashion or the lower secondary yield is merely an intrinsic property of polymer material. The fact that the difference extends to higher *E*/*N* perhaps supports more the latter assumption but does not rule out the former one altogether.

Plasma Treatment of Surfaces

In the top-down approach, a plasma may be employed to create structures on the surface out of the bulk material, while in the bottom-up strategy one may generate the building blocks within the plasma and grow structures from the surface upwards as previously discussed for CNT or nanocrystals. In both cases one needs to define very well the pattern that needs to be manufactured and techniques to make the patterned nanostructure. In addition, the surface should not be damaged or the processing-induced roughness should not render the product non-functional. With transition to ever smaller dimensions and the need to mass produce nanomaterials, alternative

methods of transferring patterns onto the surface are sought and are subject to scientific interest.

Mariotti and coworkers [126] use oxides and electromagnetic fields to transfer patterns. Shirafuji et al. [127] use self-organization of filaments to transfer the pattern. One needs a better control of filaments and reduction of their dimensions (possibly higher-pressure operation) and spacing between them. Holländer and coworkers [128, 129] have produced very localized deposition at a very fast rate to generate micro-tips. It has been mentioned above that self-organization of dust particles in plasma may be used as a precursor for the pattern that may be transferred to the surface by manipulating the dust Coulomb crystal by electric fields.

Development of micro and subsequently nanodischarges opens a possibility of patterning of etching or deposition on μm or nm scales by appropriately designed discharges [130, 131].

Surface Roughness as a Limiting Factor

Roughness represents a good predictor of how a real object will interact with its environment [48, 132]. Decreasing the roughness of a surface will usually increase exponentially its manufacturing costs. This often results in a trade-off between the manufacturing cost of a component and its performance in application. In Fig. 3.17 we show examples of the modelled roughness development of a nanocomposite material that is plasma etched (see also [133]). Surface roughness should not exceed 10% of dimensions or otherwise the structure may lose its functionality in nanoelectronics.

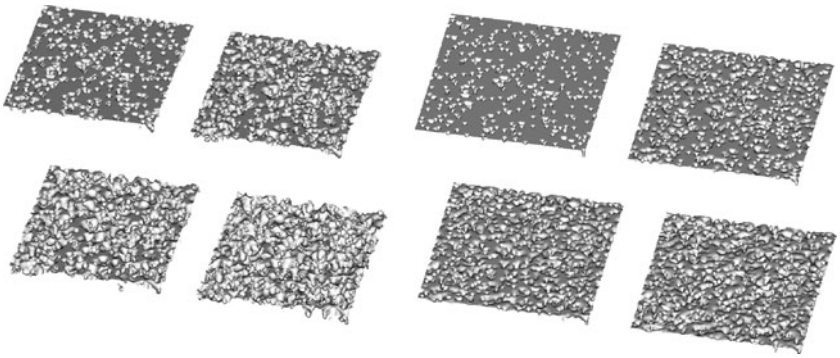


Fig. 3.17 The evolution of the surface roughness (during $t = 40$ arb. units) nanocomposite material for $s = 20$ and $p = 0.1$ induced by anisotropic (*left*) and isotropic etching (*right*)

It was found that the etching behaviour of nanocomposite materials depends strongly on the etch rate ratio s and depends on the volume fraction of the easily etched phase p [46, 133].

Practical Implications of the Treatment of Surfaces

Plasma Treatment of Textiles

The extensive research in last two decades indicated the great potential of plasma in advanced functionalization of textile materials. Plasma treatment can be a viable substitute to conventional processes and very often it can provide the advantageous effects that cannot be obtained by wet processing of textile materials [134]. It offers high efficiency, economical feasibility, environmental acceptability and flexibility. Additionally, plasma treatment is confined to the surface layers of fibres (10–100 nm), leaving the bulk properties unaltered [135–137]. This is an extremely important aspect of plasma processing since the fibre surface in most of the cases is responsible for the end-use properties of textile materials. Plasma treatment is also beneficial from the handling and time-consuming point of view.

Non-equilibrium plasma at low pressures, corona discharge, dielectric barrier discharge and atmospheric pressure glow discharge are types of the plasma that are of major relevance for textile processing [134–141]. Various effects can be achieved on different types of fibres (natural and manmade) and forms of textile materials (tops, yarns, woven, knitted and non-woven fabrics) but the selection of plasma system and optimum operating conditions (gas, pressure, power, exposure time, electrode geometry) is of crucial significance. Therefore, the specific functionalization of a wide range of textile materials with desired level of efficiency and reproducibility requires a great knowledge of the fibre nature and very complex plasma chemistry as well as an understanding of the interactions between the substrate and plasma species. Plasma treatment of textile materials involves numerous surface processes including cleaning, etching, activation, grafting and polymerization [135, 138, 142]. In general, all these processes strongly affect the morphology and chemical composition of the fibre surface, providing the desired effects. Hundreds of papers have been published and many patents have been granted on the plasma functionalization of textile materials. It is impossible to make a complete overview of all achievements in this field and thus only few selected applications involving reactive and non-polymerizable plasmas will be discussed in this section.

The major work so far was focused on the plasma modification of wool due to its very complex structure and properties [143–156]. The numerous problematic issues corresponding to conventional wet-finishing processes and maintenance of the wool products can be overcome or reduced by plasma processing. Concretely, plasma treatment of wool leads to an improvement of its dyeing and printing properties as well as to a reduction of its pilling and felting tendency. The severe attack of different plasma species causes the oxidation and etching of the fibre surface. The layer of covalently bound fatty acids on the wool fibre surface acts as a natural hydrophobic barrier for the diffusion of water and dye molecules into the fibre interior. In addition to oxidation and/or partial removal of fatty layer, plasma treatment induces a formation of new polar functional groups and conversion of the cystine to cysteic acid in the surface layers of the fibre [137, 143, 144, 146, 155]. Simultaneously, the fibre

surface becomes rougher as a result of plasma etching and sputtering. However, the extent of all these modifications is highly influenced by the type of applied gas and plasma system.

The increase in the amount of oxygen-containing groups on the fibre surface improves the wettability and swelling of wool [147]. In other words, the fibres become more accessible to water and dye molecules. The comparative analysis of the published results dealing with dyeability of wool indicates that plasma treatment only slightly enhances the final exhaustion of acid and chrome dyes [146, 150, 152, 153]. In contrast, the final exhaustion of reactive dyes increases after plasma treatment [150]. However, plasma treatment considerably contributes to increase in the dyeing rate independently of the dye class [146, 152–154]. Higher dyeing rates indicate the possibility of shortening the dyeing time and the reduction in energy consumption since less time is required to achieve the desirable colour shade. Plasma treatment can also replace the conventional chlorination process prior to printing of wool fabrics [141, 143, 144, 147, 156]. Improved printability is attributed to enhanced wettability and swelling of wool [147].

Probably the most complex issue corresponds to reduction of felting shrinkage of wool products. This phenomenon appears basically because of the specific scale-like, hydrophobic surface of fibre [137]. Plasma-induced increase of the fibre surface hydrophilicity, morphological and frictional changes reduce the tendency of wool to shrink [137, 143, 148]. Although the surface becomes significantly rougher and the fibre/fibre friction increases, the felting shrinkage is significantly reduced due to the decrease in differential frictional effect [143]. In spite of positive effects of plasma, to produce machine-washable wool, the treatment with polymers after plasma treatment is necessary [137,143,151,152]. Recent studies also revealed that plasma-treated recycle wool based non-woven material can be utilized as an efficient sorbent for removal of acid dyes and heavy metal ions from water [157,158].

Research on the plasma modification of cellulosic fibres has been mainly focused on cotton fibres. Electron spin resonance measurements indicate that considerable amount of free radical sites are created on the surface of cellulosic fibres [135, 159]. Therefore, unstable and extremely reactive free radicals with unpaired electrons can initiate different functionalization processes, depending on their location in the cellulose structure [159]. Early studies were mostly oriented towards improvement of the spinnability and strength of the cotton fibres [135]. However, much of the work so far also deals with wetting and dyeing properties of plasma-treated cotton fibres [160–166]. Similar to wool, the morphological changes induced by etching and higher content of carboxyl groups on the fibre surface positively affect the wettability, wickability and dyeability of cotton [159–164]. Formed cracks on the fibre surface provide easier access of the dye molecules during the ink-jet printing of cotton fabrics, leading to a higher dye uptake [165]. Additionally, the colour-fastness and the definition of the final print marks are also improved [165]. The increased capacity of the fibre surface to hold moisture enhances the conductivity, and consequently, the surface electrical resistivity of cotton fabrics decreases [161].

Plasma treatment can efficiently diminish the environmental impact of conventional high energy and water-consuming desizing of polyvinyl alcohol (PVA)

from cotton fabrics [166]. Air/He or air/O_2/He plasma at atmospheric pressure induces the PVA chain scission and formation of new polar functional groups, facilitating the swelling, dissolving and dispersing of PVA size [166]. This further leads to an improvement of percent desizing ratio by subsequent washing, particularly in cold water. At the same time, the tensile strength of cotton fabrics is not worsened.

Lately, a novel application of plasma treatment for obtaining the 'worn look' effect on indigo-dyed denim fabrics has been proposed [167, 168]. Achieved results can encourage further research on the possibility of plasma implementation in denim finishing because of water- and chemical-free processing, low energy costs and short treatment time in comparison with conventional bio-stoning (approximately 90 min) [152].

Plasma treatment of hemp fabrics causes an increase in water retention, dyeing rate, final dye exhaustion and colour yield of dyed samples [169]. In a similar manner, viscose fibres exposed to plasma exhibit higher water retention and wettability, while the tensile properties remain unaffected [170].

The enhancement of hydrophilicity of manmade fibres by plasma treatment was described in literature in detail [138, 141, 171–177]. Particularly the modification of polyester (PET), polyamide (PA) and polypropylene (PP) fibres is of great interest since they account for the highest share in total manmade fibre production and usage globally. The major drawback of these fibres is their strong hydrophobic nature, i.e. a low surface energy, which in turn makes adhesion, dyeing and printing of these materials more complex. Extensive X-ray photoelectron spectroscopy (XPS) measurements indicated that improved wettability, wickability and moisture retention can be attributed to formation of numerous polar functional groups (-OH, -OOH, -COOH, etc.) on the fibre surface during the plasma processing or through post-plasma reactions with environmental species [141, 171–175, 178]. Even very short exposure to plasma induces a substantial increase in water uptake and water uptake rate [141]. Additionally, the adhesion between plasma-treated fibres and a wide range of other polymers can be significantly enhanced [141, 72]. The level of hydrophilicity is strongly dependent on the applied gas and plasma system. As expected, air and oxygen plasmas are extremely efficient in incorporating the oxygen functionalities. Although less pronounced, the positive effect of argon or helium plasmas with air fractions lower than 0.1% should not be neglected [141, 178]. The benefits of helium and particularly argon plasma over air plasma become more prominent during the ageing process. Namely, plasma-treated polymers are prone to ageing during the storage in air. The ageing effect is the result of the reorientation of polar groups from the fibre surface towards its interior. A suppression of ageing can be achieved by cross-linking of the polymer chains during plasma processing and this is what exactly happens when the fibres are exposed to helium or argon plasma [141, 178]. The dyeing and printing properties of plasma-treated manmade fibres are highly influenced by the type of the fibre, dye class and plasma parameters [141, 146, 177]. Therefore, general conclusions cannot be made and broader discussion would be necessary.

Latest studies opened up some new perspectives on plasma that could be utilized for engineering of multifunctional, stable and durable textile nanocomposite materials. Namely, plasma activation particularly of hydrophobic textile surfaces positively affects the binding efficiency of different metal and metal oxide nanoparticles [179–183]. It was reported that low-pressure plasma or corona pretreatment of PET, PA and PP fabrics enhances the deposition of silver nanoparticles, providing an extraordinary antimicrobial activity with long-term durability [179, 181, 182]. Such effects are attributed to increased surface roughness and formation of new polar groups on the fibre surface. Consequently, better interaction between the substrate and hydrophilic colloidal silver nanoparticles are established. Figure 3.18 illustrates the positive effects of air RF plasma treatment of PP non-woven fabric on the deposition of silver nanoparticles. Scanning electron micrographs (SEM) clearly reveal plasma-induced morphological changes, i.e. the formation of very uniform nanostructured striations on the PP fibre surface. Plasma pre-treatment facilitates the binding of silver nanoparticles that are evenly distributed over the surface of the PP fibre. Similar effects are obtained with titanium dioxide nanoparticles which provide the UV-protective, antibacterial and self-cleaning properties to textile materials [180, 183].

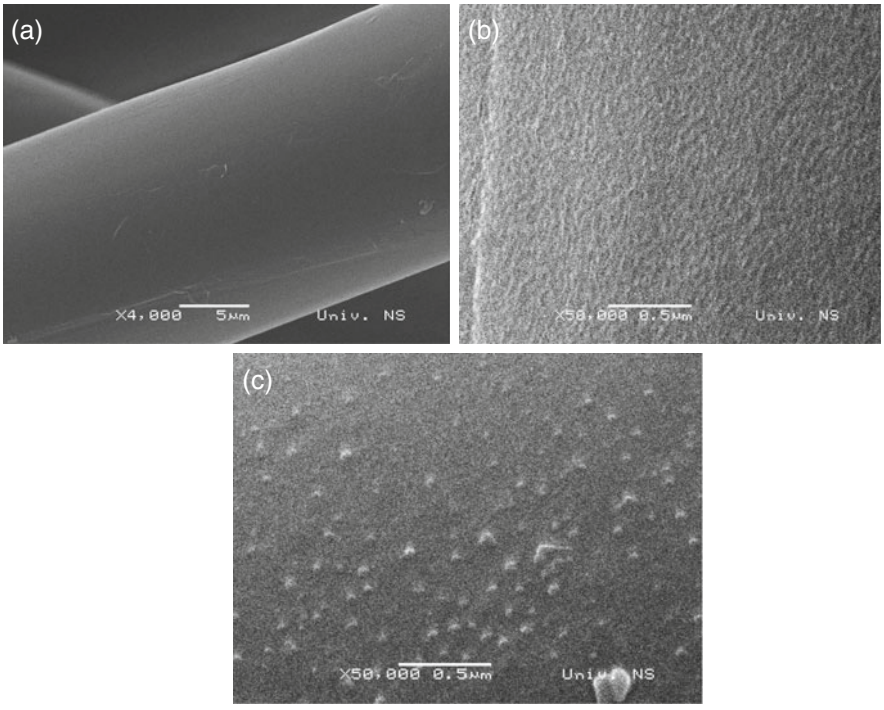


Fig. 3.18 SEM images of (**a**) untreated PP fibre, (**b**) PP fibre treated by air RF plasma and (**c**) PP fibre treated by air RF plasma and loaded with silver nanoparticles

Biomedical Applications

There is a large number of possible applications of nanostructures both free and associated with surfaces in biomedicine. In the production of some of these structures plasmas are useful, even more productive than other schemes. In some, applications plasmas are essential. Covering this field is beyond this article. We shall only briefly cover some of the possible applications where plasmas, nanotechnologies and biomedicine meet directly.

Some of the plasmas developed with specific nanotechnological applications in mind proved to be very effective in biomedical applications and vice versa.

One of the most promising future applications of nanowires and application of CNT is to act as an interface between electronic devices and living being, or as a biosensor [184].

Plasma Sterilization

The most common low-temperature sterilization methods are either based on chemical treatments, which call for toxic active agents and generate harmful effluents, or based on exposure to γ radiation, which can alter the bulk properties of the polymers being sterilized. In the last decades, the medical use of heat-sensitive tools has increased the need for efficient and safe way of sterilization by using low-temperature plasmas. Plasma sterilization ensures complete inactivation of microorganisms (e.g. vegetative bacteria, spores, viruses) and possibly removes all infectious residues (e.g. pyrogens (endotoxins), prion proteins) on used instruments. In particular plasmas have an advantage in destruction of spore.

Low-temperature discharges created over wide range of pressures (from low pressures to atmospheric pressures) [185–189] are used for plasma sterilization. Two generic plasma configurations are nowadays being investigated. One of them is when sterilization takes place within the discharge region [187] and the other is when it takes place in its flowing afterglow [190, 191].

Several mechanisms of sterilization have been considered and probably they all play some role with varying importance in different plasma sources. Those include UV radiation, effect of radicals, effect of high-energy particles and a number of combined synergistic effects [192]. Two separate issues related to sterilization must be considered: one for sterilization on living tissues as a part of healing of wounds and treatment of some diseases [192] and the other for sterilization on non-living materials. Application of some of the nanotechnologies in biology and medicine will require full sterilization on nanoscales relevant for the nanostructures that are being implemented.

New Sources of Non-equilibrium Plasmas

In addition to the sources discussed above in the context of the standard top-down technology as implemented in the current nanoelectronic industry, there are a number of plasma sources that may be relevant for nanotechnologies. Those include

microwave and surface wave plasmas, dc discharges, atmospheric pressure glow and other non-equilibrium atmospheric pressure discharges. Here we shall spend some time only on micro-discharges in general and one particular atmospheric pressure plasma source, the plasma needle.

Micro-discharges and Their Applications

Micro-discharges have grown into one of the most interesting topics in plasma physics, due to the needs in development of specialized applications of non-equilibrium plasmas in the fields of nanotechnologies and light sources. Generally, non-equilibrium plasmas are produced easily only at low pressures. At atmospheric pressure, rapid growth of ionization and strong coupling between ions and electrons lead to the tendency to produce thermal plasma. One of the ways to avoid thermalization of plasma at atmospheric pressure is to operate at very small gaps. Standard dimension discharges (~1 cm) at atmospheric pressures operate far from the minimum of the Paschen curve (breakdown voltages vs. pressure × gap (pd)). Thanks to the scaling laws, miniaturization of the discharge gives us a possibility to achieve non-equilibrium conditions (near the Paschen minimum, i.e. the lowest and thereby the optimum breakdown potential) even at atmospheric pressure.

Micro-discharge sources of various geometry and modes of operation have been developed for use in specific devices: dc plasma sources for optical emission detection [193] and localized micromachining of silicon; micro-hollow cathodes used as sources of UV light [194, 195]; dielectric barrier discharges for gas analysis [127]; capacitively coupled [196] and inductively coupled RF plasma sources [197] with numerous applications in mind, including micro-analytical applications; microwave plasma sources [198].

Scaling of the Properties of the Standard Size to Micro-discharges

The best way to establish a good understanding of the operating conditions in micro-discharges is to start from the low-pressure discharges and employ scaling laws [52, 199]. Standard scaling parameters in gas discharges are electric field to gas number density ratio (E/N), pressure times electrode gap (pd), current density normalized by the electrode gap to the square (jd^2) and frequency times gas number density (ωN) for RF discharges.

In the studies of the breakdown properties of gas discharges, it is customary to measure the Paschen curves [200, 201]. Most of the recent studies deal with micro-hollow cathode geometries, which lead to complex behaviour that does not resemble conditions in standard hollow cathode discharges [202]. Therefore, we pursue parallel-plate geometry, which is more prone to oscillations and instabilities, but the results are easier to interpret [52]. Based on modelling results, we can expect breakdown of scaling at gaps smaller than 10 μm [196, 203, 204], where field emission becomes a dominant mechanism of the breakdown and the discharge maintenance. However, some degree of departure from the standard dependence of the breakdown potential against pd is found even at distances of few 100 μm. It

was found that this is due to the long path breakdown, and any failure to comply with strict requirements to prevent the discharge behind the electrode will lead to the breakdown at effectively larger *pd* and therefore a smaller potential.

Plasma Needle

Medicine and biology are among the most promising, but at the same time most challenging, fields of application of nanotechnologies, plasmas and in particular micro-discharges. Four important areas of application in medicine and biology may be specified: (1) diagnostics, (2) drug delivery, (3) neutral prosthetics and tissue engineering and (4) minimally invasive surgery. Micro-discharges may play an important role especially in tissue engineering and in minimally invasive surgery.

One example of a versatile micro-discharge that has been developed recently for such applications is the so-called plasma needle [205]. This source operates at atmospheric pressure and meets all the necessary conditions for treatment of organic materials and living tissues. It is desired that the plasma should be non-aggressive, local, with small penetration depth and, at the same time, produce chemically active species at a low gas temperature. Plasma needle meets all those requirements.

The dimensions of the discharge are of the order of 1 mm and the plasma is sustained by an RF field. The plasma provides a facility to attack living matter by radicals formed in the plasma while gas heating does not exceed the acceptable limit for the living organisms (which is 420°C). The system has been used to induce localized sterilization, apoptosis and even separation of the cells from the tissue that may be assembled back into the tissue [206]. However, there are few data in the literature about plasma effects on plant cells and tissues [207]. At the same time the plasma may be used for localized treatment of the materials and more powerful versions are being used for minimally invasive surgery. In Fig. 3.19 we show a plasma needle developed at the Institute of Physics in Belgrade.

The benefits from the studies of micro-discharges may also be indirect. The understanding and control of electrical breakdown across the gap is critically

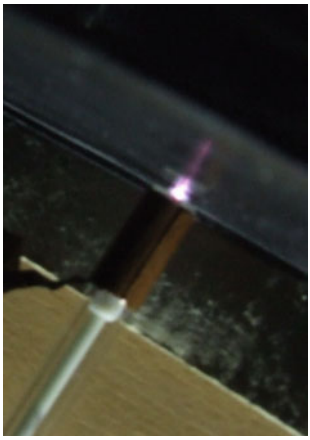


Fig. 3.19 Plasma needle (Institute of physics, Belgrade)

important for the micro-electronics industry. With each new generation of devices, the gap spacing is often reduced. The integrated circuit, MEMS, hard disk drive magnetic recording industry and flat panel display industries may benefit from the understanding of the breakdown mechanisms in micro gaps under atmospheric pressure and some of these processes may limit the deposition of energy and some applications of nano and microstructures.

Conclusion

Plasma processing is indeed one of the crucial modern technologies [17] that has made a significant impact on fabrication of integrated circuits (IC) and consequently in electronics and consumer goods in general. Damage to ICs during manufacture as a result of charging of the dielectrics during finalization of interconnects reduces the profitability and the ability to produce economically the larger size microchips that could house complex systems on a single chip such as a digital TV or complete PC.

All processes requiring energetic ion assistance, such as SiO_2 etching, suffer from differential charging since directionality and flux inside the feature will be affected by accumulated charges as the feature depth increases.

Development of a comprehensive modelling tool that includes both plasma model and a model of the development of the etched structures together with calculation of the surface potentials is a necessary aid in research and engineering. A large range of possible processes must be included if one wants to simulate processes of relevance for nanotechnologies. One example is the high aspect ratio low-k dielectric etching, the other possibility would be to study nanoroughness that would limit the applicability of plasmas in nanotechnologies. For further miniaturization it will be necessary to revisit the same old issues that have been resolved at the level of larger dimensions. Those would include etch stop, charging and other types of damage, non-uniform etching (especially close to the boundaries of reactor), increasing of the size of wafers while maintaining uniformity and many more. The new issues will be, on the other hand, dictated by the choice of materials. If all those problems are resolved the CCP reactors may continue to be the most useful tool even for dimensions of the order of 45 nm and less provided that lithography finds methods to maintain the reduction of dimensions. One should take into account the inertia of the industry in both preferred technology and material. Thus it seems that the standard top-down technology may be used whenever possible in producing nanoapplications in Si as the technology has reached resolutions that are in strong competition with other techniques especially if massive production and parallel and well-designed connectivity are sought.

Applications of plasmas in nanotechnologies are not limited to the specific cases of plasma etching, micro-discharges and applications that were mentioned above. For example production of some typical nanostructures such as nanotubes proceeds through application of plasmas, and in particular some applications such as that of the secondary electron sources involve gas discharges. In addition, the best way to achieve functionalization of nanotubes is by employing plasmas. Plasmas are

also used for thin-film deposition for analysis of surfaces and for treatment and changing of the properties of materials and surfaces. Nanostructured materials such as hyperhydrophobic surfaces rely on plasma etching of nanostructures and possible plasma polymerization. Applications in biology were shown to be numerous and to open new possibilities. Most importantly plasmas are well integrated in the present-day top-down technologies that have been tested in industrial environment.

Japan Society for Applied Physics has recently prepared a roadmap for plasmas in nanotechnologies (http://www.jsap.or.jp/english/index.html). The next decade is expected to yield drug-delivery systems, 3D displays and major advances in environmental detoxification while major electronics technologies will go beyond 25 nm; 2030 s are expected to master 10 nm and 5 nm technologies with increased biocompatibility of electronics and the 2040 s to reach 2.5 nm resolutions in production. In order to reach those goals it is necessary to improve control of species and surface reactions, selective reactions, non-invasive 3D diagnostics accurate predictive models in the near future. Diffusion of these achievements to development and further more to design of industrial processes is required to achieve all the promises that nanotechnologies may bring.

While applications in nanoelectronics have been the driving force behind the development of nanotechnologies, one may actually expect first applications in intelligent nanostructuring and in biomedical applications. The electronics will most probably continue along its top-down route and will be open to new possibilities.

The applicability of nanotechnologies will depend on the ability to achieve massive parallel production within the acceptable economical limits and that is where application of non-equilibrium plasmas will show its potential. Of course plasmas are not a panacea for nanotechnologies but are fully integrated in their development with their advantages and limitations.

Acknowledgement This work was supported by MNZZS project 141025. We thank all our colleagues who contributed to the work on topics covered here. We are grateful to Prof. U. Kortshagen, R. Gresback (Fig. 3.14), Prof. M. Hori (Fig. 3.15), Dr. V. Radmilović, Dr. B. Radjenović (Figs. 3.6, 3.7 and 3.17), Dr. U. Cvelbar and Prof. M. Mozetić for providing us with some of the materials used in this chapter.

References

1. Howard, R.E., Prober, D.E.: Nanometer-Scale Fabrication Techniques. In: Einspruch NG VLSI Electronics: Microstructure Science. Academic Press, New York (1982).
2. Flanders, C.D.: Microelectron. Eng. **2**, 82 (1984)
3. Ito, T., Yamada, T., Inao, Y., Yamaguchi, T., Mizutani, N., Kuroda, R.: Appl. Phys. Lett. **89**, 033113 (2006).
4. Yamaguchi, T., Yamada, T., Terao, A., Ito, T., Inao, Y., Mizutani, N. and Kuroda, R.: Microelectron. Eng. **84**, 690 (2007)
5. Chang, T.H.P., Kern, D.P., Kratschmer, E.: IBM J. Res. Develop. **32**, 462 (1988).
6. Torres, J.M., Dhariwal, S.R., Nanotechnology **10**, 102 (1999).
7. Ito, T., Izaki, T., Terashima, K., Thin Solid Films **386**, 300 (2001).
8. Krummenacker, M., Lewis, J., Prospects in Nanotechnology: Towards Molecular Manufacturing. Wiley and Sons, New York (1995)

9. Seeman, N.C.: Current Opinion Struct. Biol. **6**, 519 (1996).
10. Heller, A.: Current Opinion Biotechnol. **7**, 50 (1996).
11. Junno, T., Carlsson, S.B., Xu, H., Montelius, L., Samuelson, L.: App. Phys. Lett. **72**, 548 (1998).
12. Ben-Jacob, E., Herman, Z., Caspi, S., Phys. Lett. A **263**, 199 (1999).
13. Corbetta, J., McKeowna, P.A., Peggsb, G.N., Whatmore, R.: CIRP Ann.—Manufactur. Technol. **49**, 523 (2000)
14. Mamalis, A.G.: J. Mat. Process. Technol. **181**, 52 (2007).
15. Iijima, S.: Nature **354**, 56 (1991).
16. Ostrikov, K.: Rev. Modern Phys. **77**, 489 (2005).
17. Ostrikov, K., Murphy, A.B.: J. Phys. D. Appl. Phys. **40**, 2223 (2007).
18. Moore, G.E.: Electronics **38**, 1 (1965).
19. Cardinaud, C., Peignon, M.C., Tessier, Y.P.: Appl. Surf. Sci. **164**, 72 (2000).
20. Leiberman, M.A., Lichtenberg, J.A.: Principles of Plasma Discharges and Materials Processing. 2nd edn. Wiley, New York (2005).
21. Makabe, T., Petrović, Z.Lj.: Plasma Electronics: Application in Microelectronic Device Fabrication. Taylor & Francis, New York (2006).
22. Vrhovac, S.B., Stojanović, D.V., Jelenković, M.B., Petrović, Z.LJ.: J. Appl. Phys. **90**, 5871 (2001).
23. Smalley, R.E., Rev. Modern Phys. 69, 723 (1997).
24. Wang, X.K., Lin, X.W., Dravid, V.P., Ketterson, J.B., Changa, R.P.H.: Appl. Phys. Lett. **66**, 2430 (1995).
25. Hosokawa, N., Matsuzak, R., Asamaki, T.: Jap. J. Appl. Phys. Suppl. **2**, 435 (1974).
26. Coburn, J.W., Winters, F.H.: J. Appl. Phys. **50**, 3189 (1979).
27. Hippler, R., Pfau, S., Schmidt, M., Schoenbach, K.H. (eds): Low Temperature Plasma Physics. Wiley-VCH, Berlin (2001).
28. Makabe, T. (ed.): Advances in Low Temperature RF Plasmas: Basis for Process Design. North Holland, Amsterdam (2002).
29. Iwai, H.: Microelectron. Eng. **86**, 1520 (2009).
30. Kushner, M.J.: J. Appl. Phys. **58**, 4024 (1985).
31. Graves, D.B., Jensen, F.K.: IEEE Trans. Plasma Sci. **14**, 78 (1986).
32. Nakano, N., Shimura, N., Petrović, Z.LJ., Makabe, T.: Phys. Rev. E **49**, 4455 (1994).
33. Tochikubo, F., Chiba, T., Watanabe, T.: Jpn. J. Appl. Phys. **38**, 5244 (1999).
34. Hopwood, J.: Plasma Sources Sci. Technol. **1**, 109 (1992).
35. Okigawa, A., Tadokoro, M., Itoh, A., Nakano, N., Petrovic, Z.Lj., Makabe, T.: Jpn. J. Appl. Phys. **36**, 4605 (1997).
36. Makabe, T., Petrović, Z.LJ.: Appl. Surf. Sci. **192**, 88 (2002).
37. Kortshagen, U., Tsendin, D.L.: Appl. Phys. Lett. **65**, 1355 (1994).
38. Denda, T., Miyoshi, Y., Komukai, Y., Goto, T., Petrović Z.LJ., Makabe, T.: J. Appl. Phys. **95**, 870 (2004).
39. Nakano, N., Petrović Z.LJ., Makabe, T.: Jpn. J. Appl. Phys. **33**, 2223 (1994).
40. Ventzek, P.L.G., Hoekstra, J.R., Kushner, J.M.: J. Vac. Sci. Technol. B **12**, 461 (1994).
41. Collison, W.Z., Kushner, J.M.: IEEE Trans. Plasma Sci. **24**, 135 (1996).
42. Makabe, T., Maeshige, K.: App. Surf. Sci. **192**, 176 (2002).
43. Kolobov, V.I., Arslanbekov, R.R.: IEEE Trans. Plasma Sci. **34**, 895 (2006).
44. Yang, Y., Kushner, J.M.: J. Vacuum Sci. Techn. A **25**, 1420 (2007).
45. Van Dijk, J., Peerenboom, K., Jimenez, M., Mihailova, D., van der Mullen. J.: J. Phys. D: Appl. Phys. **42**, 194012 (2009)..
46. Zakka, E., Constantoudis, V., Gogolides, E.: IEEE Trans. Plasma Sci. **35**, 1359 (2007).
47. Radjenović, B., Radmilović-Radjenović, M., Petrović, Z.LJ.: IEEE Trans. Plasma Sci. **36**, 874 (2008).
48. Guo, W., Sawin, H.H.: J. Phys. D: Appl. Phys. **42**, 194014 (2009).
49. Hamaoka, F., Yagisawa, T., Makabe, T.: IEEE Trans. Plasma Sci. **35**, 1350 (2007).

50. Radjenović, B. Lee K.J., Radmilović-Radjenović, M.: Comput. Phys. Commun. **174**, 127 (2006).
51. Radjenović, B., Radmilović-Radjenović, M. Mitrić, M.: Appl. Phys. Lett. 89, 213102 (2006).
52. Lj Petrovic, Z., Škoro N., Marić D, Mahony C.M.O, Maguire P D, Radmilović-Rađenović M., Malović G.: Phys. D Appl. Phys. 41, 194002 (2008).
53. Lj Petrović, Z., Dujko, S., Marić, D., Malović, G., Nikitović, Ž., Šašić, O., Jovanović, J., Stojanović, V., Radmilović-Radjenović, M.: J. Phys. D: Appl. Phys. **42**, 194002 (2009).
54. Robson, R.E., White, R.D., Petrović, Z.LJ.: Rev. Mod. Phys. **77**, 1303 (2005)
55. Dujko, S., Raspopović, M.Z., Petrović, Z.LJ., Makabe, T.: IEEE Trans. Plasma Sci. **31**, 711 (2003).
56. Maeda, K., Makabe, T., Nakano, N., Bzenić, S., Petrović, Z.LJ.: Phys. Rev. E **55**, 5901 (1997).
57. Raspopović, Z., Sakadžić, S., Petrović, Z.LJ., Makabe, T.: Phys. D: Appl. Phys. **33**, 1298 (2000).
58. White, R.D., Ness, F.K., Robson, E.R.: Appl. Surf. Sci. **192**, 26 (2002).
59. Tadokaro, M., Hirata, H., Nakano, N., Petrović, Z.LJ., Makabe, T.: Phys. Rev. E **57**, R43 (1998).
60. Vasenkov, A.V., Kushner, J.M.: Phys. Rev. E **66**, 066411 (2002)
61. Dujko, S., White, R., Petrović, Z.LJ.: J. Phys. D: Appl. Phys. **41**, 245205 (2008).
62. Rozum, I., Mason, J.N., Tennyson, J.: New J. Phys. **5**, 155 (2003).
63. Nikitović, Ž.D., Stojanović, D.V., Booth, J.P., Petrović, Z.LJ.: Plasma Sources Sci. Technol. **18**, 035008 (2009).
64. Nikitović, Ž.D., Stojanović, D.V., Petrović, Z.LJ.: Acta Phys. Pol. A **115**, 765 (2009).
65. Miyauchi, M., Miyoshi, Y., Petrović, Z.LJ. Makabe, T.: Solid State Electron. **51**, 1418 (2007).
66. Kitajima, T., Takeo, Y., Petrović, Z.LJ., Makbe, T.: Appl. Phys. Lett. **77**, 489 (2000).
67. Turner, M.M., Chabert, P.: Phys. Rev. Lett. **96**, 205001 (2006).
68. Matsui, J., Maeshige, K., Makabe, T.: J. Phys. D: Appl. Phys. **34**, 2950 (2001).
69. Matsui, J., Nakano, N., Petrović, Z.Lj., Makabe, T.: Appl. Phys. Lett. **78**, 883 (2001).
70. Yoshida, Y., Watanabe, T.: Solid State Technol. **27**, 263 (1984).
71. Fang, S., McVittie, P.J.: IEEE Trans. Electron Devices **41**, 1848 (1994).
72. Hwang, G., Giapis, K.: J. Vacuum Sci. Technol. B **15**, 70 (1997).
73. Tsui, B.Y., Lin, S.S., Tsai, C.S., Hsia, C.C.: Microelectron. Reliabil. **40**, 2039 (2000).
74. Lin, H.C., Chen, C.C., Chien, C.H., Hsien, S.K., Wang, M.F., Chao, T.S., Huang, T.Y., Chang, C.Y.: IEEE Electron. Device Lett. **19**, 68 (1998).
75. Bhuva, B., Kerns, S.: Plasma Process-induced Damage. In: Buschow, K.H.J., Cahn, R.W., Flemings, M.C. and Ilschner, B.(eds.) Encyclopaedia of Materials: Science and Technology. Elsevier, Oxford, UK (2008).
76. Hioki, K., Itazu, N., Petrović, Z.Lj., Makabe, T.: Jpn. J. Appl. Phys. **40**, L1183 (2001).
77. Ohmori, T., Goto, T., Kitajima, T., Makabe, T.: Appl. Phys. Lett. **88**: 4637 (2003).
78. Radjenović, B., Radmilović-Radjenović, M.: J. Phys.: Conf. Series **86**, 012017 (2007).
79. Giapis, P.K., Moore, A.T., Minton, K.T.: J. Vac. Sci. Technol. A **13**, 959 (1995).
80. Petrović, Z.LJ., Phelps, A.V.: In: Goto, T. (ed.) Proceedings of the International Seminar on Reactive Plasmas. Japan Society of Applied Physics, Nagoya, Japan (1991).
81. Petrović, Z.LJ., Stojanović, D.V.: J. Vac. Sci. Technol. A **16**, 329 (1998).
82. Samukawa, S., Sakamoto, K., Ichiki, K.: Generating high-efficiency neutral beams by using negative ions in an inductively coupled plasma source. J. Vacuum Sci. Technol. A **20**, 1566 (2002).
83. Panda, S., Economou, J.D.: J. Vac. Sci. Technol. A **19**, 398 (2001).
84. Stojković, A., Radmilović-Radjenović, M. Petrović, Z.Lj.: Mater. Sci. Forum **494**, 297 (2005).
85. Radmilovic-Radjenovic, M., Radjenovic, B.: Plasma Sources Sci. Technol. **15**, 1 (2006).
86. Kroto, H.W., Heath, J.R., O′Brien, C.S., Curl, R.F., Smalley, R.E.: Nature **318**, 162 (1985).

87. Kroto, H.W.: Nature **329**, 529 (1987).
88. Li, L.Y., Kinloch, I.A., Windle, H.A.: Science **304**, 276 (2004).
89. Meyyappan, M.: J. Phys. D: Appl. Phys. **42**, 213001(2009).
90. Denysenko, I., Ostrikov, K., Cvelbar, U., Mozetic, M. Azarenkov, N.A.: J. Appl. Phys. **104**, 073301 (2008).
91. Ngo, Q., Cassell, A.M., Radmilović, V., Li, J., Krishnan, S., Meyyappan, M., Yang, C.Y.: Carbon **45**, 424 (2007).
92. Wang, Y.Y., Gupta, S., Nemanich, R.J.: Appl. Phys. Lett. **85**, 2601 (2004).
93. Bower, C., Zhou, O., Zhu, W., Werder, D.J., Jin, S.: Appl. Phys. Lett. 77, 2767 (2000).
94. Qin, L.C., Zhou, D., Krauss, A.R., Gruen, D.M.: Appl. Phys. Lett. **72**, 3437 (1998).
95. Levchenko, I., Ostrikov, K., Keidar, M., et al.: Modes of nanotube growth in plasmas and reasons for single-walled structure. J. Phys. D: Appl. Phys. **41**(132004), 6 pp. (2008)
96. Mathur, A., Tweedie, M., Roy, S.S. Maguire, D.P. McLaughlin, J.A.D.: unpublished; see also: Electrical and Raman Spectroscopic Studies of Vertically Aligned Multi-Walled Carbon Nanotubes. J. Nanosc. and Nanotechnology **9**, 4392 (2009).
97. Saha, A.A., Mitra, S.K., Tweedie, M., Roy, S.S., McLaughlin, J.A.D.: Microfluid Nanofluid **7**, 334 (2009).
98. Kumar, S., Yadev, P., Hamilton, J.W.J., McLaughlin, J.A.D.: Arrays of carbon nanoflake spherules realised on copper substrate. Diamond Relat. Mater. **18**, 1070 (2009)
99. Murphy, H., Papakonstantinou, P., Okpalugo, T.T.T.: J. Vacuum Sci. Techn. B **24**, 715 (2006).
100. Iyer, G.R.S., Papakonstantinou, P., Abbas, G., Maguire, P.D., Bakirtzis, D., e-J. Surf. Sci. Nanotech. **7**, 337 (2009).
101. Abbas, G.A., Papakonstantinou, P., Iyer, G.R.S., Kirkman, I.W., Chen, C.L.: LC Phys. Rev. B **75**, 195429 (2007).
102. Roy, S.S., Papakonstantinou, P., Okpalugo, T.I.T., Murphy, H.: J. Appl. Phys. **100**, 053703 (2006).
103. Okpalugo, T.I.T., Papakonstantinou, P., Murphy, H., McLaughlin, J., Brown, N.M.D.: Carbon **43**, 2951 (2005).
104. Vohrer, U., Holmes, J., Li, Z., Teh, A., Konstantinou, P., Ruether, M.: J. Nanotechnol. Online **3**, 1 (2007).
105. Escoffier, C., Maguire, D.P. Mahony, C., Graham, W.G., McAdams, E.T. McLaughlin, J.A.D.: J. Electrochem. Soc. **149**, H98 (2002).
106. McLaughlin, J. A.D., Maguire, D.P.: Diamond Relat. Mater. **17**, 873 (2008).
107. Maguire, P.D., McLaughlin, J.A., Okpalugo, T.I.T., Lemoine, P., Papakonstantinou, P.: Diamond Relat. Mater. **14**, 1277 (2005).
108. Lemoine, P., Quinn, J.P. Maguire, D.P.: Wear **257**, 509 (2004).
109. Md. Rahman, A., Maguire, P.D., Roy, S.S., McCann, R., McKavanagh, F., McLaughlin, J.A.D.: Diam. Relat. Mat. **18**, 1343 (2009).
110. Chu, P.K., Tang, B.Y., Wang, L.P., Wang, X.F., Wang, S.Y. Huang, N.: Rev. Sci. Instr. **72**, 1660 (2001).
111. Kortshagen, U.: J. Phys. D: Appl. Phys. **42**, 113001 (2009).
112. Gresback, R., Kortshagen, U., unpublished; see also: Mangolini, L., Thimsen, E., Kortshagen, U.: Nano Lett. **5**, 655 (2005).
113. Watanabe, Y.: J. Phys. D: Appl. Phys. **39**, R329 (2006).
114. Pi, X.D., Gresback, R., Liptak, W.R., Campbell, A.S., Kortshagen, U.: Appl. Phys. Lett. **92**, 123102 (2008).
115. Jurbergs, D., Rogojina E, Mangolini L Kortshagen, U., Appl. Phys. Lett. 88, 233116 (2006).
116. Cvelbar, U., Ostrikov, K., Mozetić, M.: Nanotechnology **19**, 405605 (2008).
117. Cvelbar, U., Ostrikov, K., Drenik, A., Mozetić, M.: Appl. Phys. Lett. **92**, 133505 (2008).
118. Wang, B.B., Ostrikov, K.: J. Appl. Phys. **105**, 083303 (2009).
119. Cvelbar, U., Ostrikov, K., Levchenko, I., Mozetić, M., Sunkara, M.K.: Appl. Phys. Lett. **94**, 211502 (2009).

120. Dato, A., Radmilovic, V., Lee, Z., Phillips, J., Frenklach, M.: Nano Lett. **8**, 201 (2008).
121. Hiramatsu, M., Shiji, K., Amano, H., Hori, M.: Appl. Phys. Lett. **84**, 4708 (2004).
122. Kovačević, E., Stefanović, J., Berndt, J., Winter, J.: J. Appl. Phys. **93**, 2924 (2003).
123. Tomčik, B., Jelenak, A., Mitrović, M., Petrović, Z.Lj.: Diamond Relat. Mater. **4**, 1126 (1995).
124. Stefanovic, I., Berndt, J., Maric, D., Samara, V., Radmilovic, M.-Radjenovic, Petrovic, Z.Lj. Kovacevic, E., Winter, J.: Phys. Rev. E **74**, 026406 (2006).
125. Phelps, A.V., Petrović, Z.Lj.: Plasma Sources Sci. Technol. **8**, R21 (1999).
126. Mariotti, D., Švrček, V., Kim, G.D.: Appl. Phys. Lett. **91**, 183111 (2007).
127. Shirafuji, T., Kitagawa, T., Wakai, T., Tachibana, K.: Appl. Phys. Lett. **83**, 2309 (2003).
128. Choukourov, A., Biederman, H., Slavinska, D., Trchova, M., Hollander, A.: Surf. Coat. Technol. **863**, 174–175 (2003).
129. Holländer, A., Abhinandan, L.: Surface Coatings Technol. 1175 174–175 (2003).
130. Sankaran, R.M., Giapis, P.K.: J. Phys. D: Appl. Phys. **36**, 2914 (2003) .
131. Sankaran, R.M., Giapis, P.K.: Appl. Phys. Lett. **79**, 593 (2001).
132. Constantoudis, V., Christoyianni, H., Zakka, E., Gogolides, E.: Phys. Rev. E **79**, 041604 (2009).
133. Radmilović-Radjenović, M., Radjenovic, B., Petrović, Z.Lj.: Thin Solid Films **517**, 3954 (2009).
134. Marcandalli, B., Riccardi, C.: Plasma treatments of fibres and textiles. In: Shishoo, R. (ed.) Plasma technologies for textiles. 1st edn. Woodhead publishing in textiles, Cambridge, (2007).
135. Ueda, M., Tokino, S.: Rev. Prog. Coloration **26**, 9 (1996).
136. Takata, T., Watanabe, H.: Jpn. Tribol. J. **41**, 1255 (1996).
137. Höcker, H.: Pure Appl. Chem. **74**, 423 (2002).
138. Wakida, T., Tokino, S.: Indian J Fibre Text **21**, 69 (1996).
139. Herbert, T.: Atmospheric-pressure Cold Plasma Processing Technology. In: Shishoo, R. (ed.) Plasma technologies for textiles, 1st edn. Woodhead publishing in textiles, Cambridge (2007).
140. Lippens, P.: Low-pressure Cold Plasma Processing Technology. I In: Shishoo, R. (ed.) Plasma technologies for textiles, 1st edn. Woodhead publishing in textiles, Cambridge (2007).
141. Morent, R., De Geyter, N., Verschuren, J., De Clerck, K., Kiekens, P., Leys, C.: Surf Coat Tech. **202**, 3427 (2008).
142. Grill, A.: Cold plasma in materials fabrication. IEEE Press, New York (1994).
143. Thomas, H.: Plasma Modification of Wool. In: Shishoo, R.(ed.) Plasma technologies for textiles, 1st edn. Woodhead publishing in textiles, Cambridge (2007).
144. Ryu, J., Wakida, T., Takagishi, T.: Textile Res. J **61**, 595 (1991).
145. Rakowski, W.: Fibre Text East. Eur. **3**, 45 (1995).
146. Wakida, T., Lee, M., Sato, Y., Ogasawara, S., Ge, Y., Niu, S.: J. Soc. Dyers Colour **112**, 233 (1996).
147. Radetić, M., Jocić, D., Jovančić, P., Trajković, R., Petrović, Z.Lj.: Textile Chem. Color Am. Dyest. Rep. **32**, 55 (2000).
148. Molina, R., Jovančić, P., Jocić, D., Bertran, E., Erra, P.: Surf Interface Anal **35**, 128 (2003).
149. Kan, C.W., Yuen, C.W.M.: J. Appl. Polym. Sci. **102**, 5958 (2006).
150. Kan, C.W., Yuen, C.W.M.: Plasma Proc. Polym. **3**, 627 (2006).
151. Jovančić, P., Molina, R., Bertran, E., Jocic, D., Julia, R.M., Erra, P.: Wool surface modification and its influence on related functional properties. In: Mittal, K.L.(ed.) Polymer Surface Modification: Relevance to Adhesion, 1st edn. VSP, Boston (2007).
152. Radetić, M., Jovančić, P., Puač, N., Petrović, Z.Lj.: J. Phys. Conf. Ser. **70**, 012017 (2007).
153. Cai, Z., Qiu, Y.: J. Appl. Polym. Sci. **109**, 1257 (2008).
154. Wakida, T., Cho, S., Choi, S., Tokino, S., Lee, M.: Textile Res. J. **68**, 848 (1998).

155. Jovančić, P., Jocić, D., Radetić, M., Topalović, T., Petrović, Z.Lj.: Mater. Sci. Forum **494**, 283 (2005).
156. Puač, N., Petrović, Z.Lj., Radetić, M., Đorđević, A.: Mater. Sci. Forum **494**, 291 (2005).
157. Radetić, M., Jocić, D., Jovančić, P., Rajaković, Lj., Thomas, H., Petrović, Z.Lj.: J. Appl. Polym. Sci. **90**, 379 (2003).
158. Radetić, M., Jocić, D., Jovančić, P., Petrović, Z.Lj., Thomas, H.: Indian J. Fibre Text. Res. **30**, 82 (2005).
159. Johansson, K.: Plasma modification of natural cellulosic fibres. In: Shishoo, R. (ed.) Plasma technologies for textiles, 1st edn. Woodhead publishing in textiles, Cambridge (2007).
160. Jung, H.Z., Ward, T.L., Benerito, R.R.: Textile Res. J **47**, 217 (1977).
161. Rashidi, A., Moussavipourgharbi, H., Mirjalili, M., Ghoranneviss, M.: Ind. J. Fibre Text. **29**, 74 (2004).
162. Karahan, H.A., Özdoğan, E.: Fiber Polym. **9**, 21 (2008).
163. Navaneetha, P.K., Selvarajan, V.: J. Mater. Process. Tech. **199**, 130 (2008).
164. Sun, D., Stylos, G.: Text. Res. J. **74**, 751 (2004)
165. Yuen, C.W.M., Kan, W.C.: J. Appl. Polym. Sci. **104**, 3214 (2007).
166. Cai, Z., Qiu, Y., Zhang, C., Hwang, Y.J., McCord, M.: Textile Res. J. **73**, 670 (2003).
167. Ghoranneviss, M., Bahareh, M., Shahidi, S., Anvari A., Rashidi, A.: Plasma Process. Polym. **3**, 316 (2006).
168. Radetic, M., Jovancic, P., Puac, N., Petrovic, Z.Lj. Šaponjic, Z.: Text. Res. J **79**, 558 (2009).
169. Radetić, M., Jovančić, P., Jocić, D., Topalović, T., Puač, N., Petrović, Z.Lj.: The influence of low-temperature plasma and enzymatic treatment on hemp fabric dyeability. Fibres Text. East Eur. **15**, 93 (2007).
170. Vrabič, U., Jesih, A., Svetec, G.D.: Fibres Text. East Eur. **15**, 124 (2007).
171. De Geyter, N., Morent, R., Leys, C.: Surf. Coat. Tech. **201**, 2460 (2006).
172. Leroux, F., Campagne, C., Perwuelz, A., Gengembre, L.: Surf. Coat. Tech. **203**, 3178 (2009).
173. Pappas, D., Bujanda, A., Demaree, J.D., Hirvonen, J.K., Kosik, W., Jensen, R., McKnight, S.: Surf. Coat Tech. **201**, 4384 (2006).
174. Canal, C., Gaboriau, F., Molina, R., Erra, P., Ricard, A.: Plasma Process Polym. **4**, 445 (2007).
175. Masaeli, E., Morshed, M., Tavanai, H.: Surf. Interface Anal. **39**, 770 (2007).
176. Yip, J., Chan, K., Sin, K.M., Lau, S.K.: J. Mater. Process Tech. **123**, 5 (2002).
177. Raffaele-Addamo, A., Riccardi, C., Selli, E., Barni, R., Piselli, M., Poletti, G., Orsini, F., Marcandalli, B., Massafra, M.R., Meda, L.: Surf. Coat. Tech. **886**, 174–175 (2003).
178. Morent, R., De Geyter, N., Leys, C., Gengembre, L., Payen, E.: Textile Res. J **77**, 471 (2007).
179. Yuranova, T., Rincon, A.G., Bozzi, A., Parra, S., Pulgarin, C., Albers, P., Kiwi, J.: J. Photochem. Photobio. A **161**, 27 (2003).
180. Bozzi, A., Yuranova, T., Kiwi, J.: J. Photochem. Photobio. A **172**, 27 (2005).
181. Radetić, M., Ilić, V., Vodnik, V., Dimitrijević, S., Jovančić, P., Šaponjić, Z., Nedeljković, M.J.: Polym. Advan. Technol. **19**, 1816 (2008) .
182. Ilić, V., Šaponjić, Z., Vodnik, V., Molina, R., Dimitrijević, S., Jovančić, P., Nedeljković, J., Radetić, M.: J. Mat. Sci. **44**, 3983 (2009).
183. Qi, K., Xin, J.H., Daoud, W.A., Mak, C.L.: Int. J. Appl. Ceram. Technol. **4**, 554 (2007).
184. Melechko, A.V., Desikan, R., McKnight, E.T.: J. Phys. D: Appl. Phys. **42**, 193001 (2009).
185. Kelly-Wintenberg, K. Montie, C.T., Brickman, C., Roth, J.R., Carr, A.K., Sorge, K., Wadworth, L.C., Tsai, P.P.Y.: J. Ind. Microbiol. Biotechnol. **20**, 69 (1998).
186. Herrmann, H.W., Henins, I., Park, J., Selwyn, S.G.: Physics of Plasmas **6**, 2284 (1999).
187. Roth, J.R., Sherman, D.M., Gadri, R.B., Karakaya, F., Zhiyu, T.C., Chen Montie, K., Kelly-Wintenberg, P.P.Y.: Tsai, IEEE Trans. Plasma. Sci. **28**, 56 (2000).
188. Laroussi, M.: IEEE Trans. Plasma Sci. **30**, 1409 (2002).
189. Fridman, G., Peddinghaus, M., Ayan, H., Balasubramanian, M., Gutsol, A., Brooks, D.A., Fridman, A., Friedman, G.: Plasma Chem. Plasma Process. **26**, 425 (2006).
190. Ricard, A., Moisan, M., Moreau, S. : J. Phys. D: Appl. Phys. **34**, 1203 (2001).

191. Kutasi, K., Saoudi, B., Pintassilgo, C.D., Loureiro, J., Moisan, M.: Plasma Process. Polym. **5**, 840 (2008) .
192. Fridman, G., Gutsol, A., Shekhter, A.B., Vasilets, V., Fridman, A.: Appl. Plasma Med. Plasma Process. Polym. **5**, 503 (2008).
193. Eijkel, J.C.T., Stoeri, H. Manz, A.: Anal. Chem. **72**, 2547 (2000).
194. Stark, R.H., Schoenbach, H.K.: Appl. Phys. Lett. **74**, 3770 (1999).
195. Kurunczi, P., Lopez, J., Shah, H., Becker, K.: Int. J. Mass Spectrom. **205**, 277 (2001).
196. Radmilović-Radjenović, M., Lee, K.J., Iza, F. Park, G.Y.: J. Phys. D: Appl. Phys. **38**, 950 (2005).
197. Minayeva, O.B., Hopwood, A.J.: J. Anal. At. Spectrom. **18**, 856 (2003).
198. Bilgic, A.M., Engel, U., Voges, E., Kückelheim M., Broekaert, J.A.C., Plasma Sources Sci. Technol. **9**, 1 (2000).
199. Marić, D., Hartmann, P., Malović, G., Donko, Z., Petrović, Z.Lj.: J. Phys. D: Appl. Phys. **36**, 2639 (2003).
200. Mariotti, D., McLaughlin, J.A., Maguire, P.: Plasma Sources Sci. Technol. **13**, 207 (2004).
201. Malović, G., Strinić, A., Živanov, S., Marić, D., Petrović, Z.Lj.: Plasma Sources Sci. Technol. **12**, S1 (2003).
202. Boeuf, J.P., Pitchford, L.C., Schoenbach, H.K.: Appl. Phys. Lett. **86**, 071501 (2005).
203. Radmilovic-Radjenovic, M., Radjenovic, B.: Plasma Sources Sci. Technol. **16**, 337 (2007).
204. Radmilović-Radjenović, M., Radjenović, B.: Plasma Sources Sci. Technol. **17**, 024005 (2008).
205. Stoffels, E., Flikweert, A.J., Stoffels, W.W., Kroesen, G.M.W.: Plasma Sources Sci. Technol. **11**, 383 (2002) .
206. Stoffels, E., Kieft, I.E., Sladek, E.R.J.: J. Phys. D: Appl. Phys. **36**, 2908 (2003).
207. Puač, N., Petrović, Z.Lj., Malović, G., Đorđević, A., Živković, S., Giba, Z., Grubišić, D.: J. Phys. D: Appl. Phys. **39**, 3514 (2006).

Chapter 4
All Carbon Nanotubes Are Not Created Equal

Gyula Eres, D.B. Geohegan, A.A. Puretzky, and C.M. Rouleau

Abstract This chapter presents the various factors that enter into consideration when choosing the source of carbon nanotubes for a specific application. Carbon nanotubes are giant molecules made of pure carbon. They have captured the imagination of the scientific community by the unique structure that provides superior physical, chemical, and electrical properties. However, a surprisingly wide disparity exists between the intrinsic properties determined under ideal conditions and the properties that carbon nanotubes exhibit in real-world situations. The lack of uniformity in carbon nanotube properties is likely to be the main obstacle holding back the development of carbon nanotube applications. This tutorial addresses the nonuniformity of carbon nanotube properties from the synthesis standpoint. This synthesis-related nonuniformity is on top of the intrinsic chirality distribution that gives the ~1:2 ratio of metallic to semiconducting nanotubes. From the standpoint of carbon bonding chemistry the variation in the quality and reproducibility of carbon nanotube materials is not unexpected. It is an intrinsic feature that is related to the metastability of carbon structures. The extent to which this effect is manifested in carbon nanotube formation is governed by the type and kinetics of the carbon nanotube synthesis reaction. Addressing this variation is critical if nanotubes are to live up to the potential already demonstrated by their phenomenal physical properties.

Introduction

Research on the science and application of carbon nanotubes (CNTs) illustrated in Fig. 4.1 will soon enter into its third decade [1, 2]. The initial flood of papers describing the extraordinary properties of CNTs has generated great excitement, promising

G. Eres (✉)
Materials Science and Technology Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831, USA
e-mail: eresg@ornl.gov

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_4,

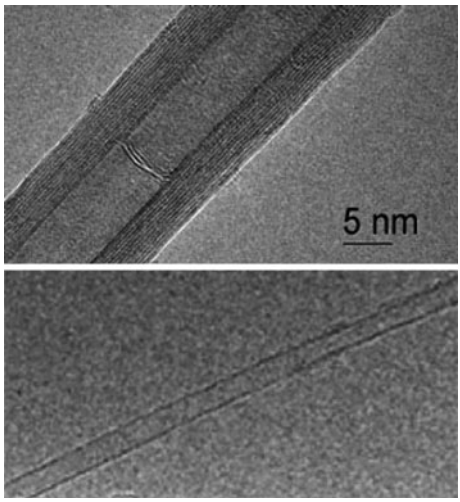


Fig. 4.1 The structures in these TEM images showing hollow tubes having outer diameters from 8-25 nm with multiple coaxial walls (top) consisting of single atomic sheets of graphitic carbon (bottom) became known as carbon nanotubes.

technological breakthroughs in fields as diverse as nanoelectronics, biomedicine, and aerospace [3–16]. Today the typical products containing nanotubes include paint, vehicle parts, sports equipment, and similar applications in which CNTs are playing an auxiliary role by reinforcing bulk mechanical, thermal, or electrical properties. But any anticipated revolution in nanotechnology products and applications based on the unique one-dimensional (1D) properties of CNTs has in fact been slow to materialize.

The main reason most of the promising technology applications remain stuck on laboratory benches is the lack of reproducibility in the structure and quality of the CNT samples available for application research. This nonreproducibility is related to the ability of carbon to form a large variety of structures from the same number of carbon atoms just by changing the bonding configuration with itself. The number of such structures, referred to as isomers, increases nonlinearly with the number of carbon atoms. For example, the best known fullerene C_{60} has one isomer, while C_{120} has around 14,000 [17]. The large number of isomers leads to a nonuniform product distribution that can dramatically change in response to relatively small changes in experimental parameters that are often difficult or impossible to account for. This is why the various methods used for the synthesis of CNTs produce such vastly different distributions of single-wall and multiwall CNTs along with other carbon structures that are undesirable side products of the synthesis process.

The slow progress in understanding how CNT synthesis works is currently the main obstacle to widespread application of CNTs. In a recent interview for *Nanotoday*, Harry Kroto noted that the field is lacking a systematic methodology for making carbon structures with predetermined properties [17]. By contrast, Kroto comments, we routinely use organic and polymer chemistry to make complicated molecules and structures in a highly predictable and reproducible fashion. Instead the current approach relies on trial and error synthesis methods followed by elaborate and laborious purification steps to separate and extract the CNT fraction from the numerous other reaction products.

Brief Overview of CNT Applications

The specific requirements of the applications as illustrated in the example in Fig. 4.2 dictate the quality and quantity of the CNTs needed [18]. The applications can be divided into two main categories. The first category includes the various applications that use CNTs in a bulk form or in combination with other materials to enhance their bulk thermal, electrical, and mechanical properties. For these applications the primary concern is the availability of CNTs in large quantities.

The second category of applications is based on exploiting the size-dependent and quantum properties of individual CNTs. These applications include novel molecular scale electrical and mechanical devices and probes capable of interacting with single molecules. The most exciting device applications are based on the 1D electrical and optical properties of metallic and semiconducting CNTs [7, 8, 10–15, 19]. Many of the unique nanoscale properties of CNTs and other nanostructures do not have bulk equivalents. The capture of these properties into macroscale applications critically depends on the uniformity of the CNT properties.

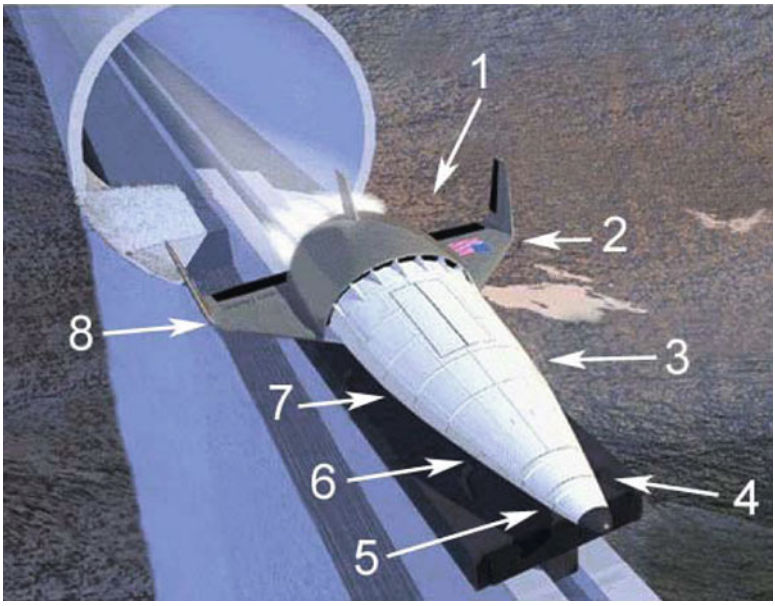


Fig. 4.2 The vision of NASA for faster, better, cheaper space transportation enabled by nanotechnology using materials and devices based on carbon nanotubes includes: 1) electronically operated flight surface, 2) nano-electromechanical sensors (NEMS), 3) fuel cells, 4) thermal protection system (TPS) elements, 5) digital nanoelectronics, 6) composite aeroshell, 7) hydrogen storage, 8) integrated aerospike engines (http://www.ipt.arc.nasa.gov/spacetransport.html).

The Origin of Many Forms of Carbon

The chemistry of carbon plays a critical role in the synthesis and processing of CNTs. However, this type of carbon chemistry occurs according to slightly different

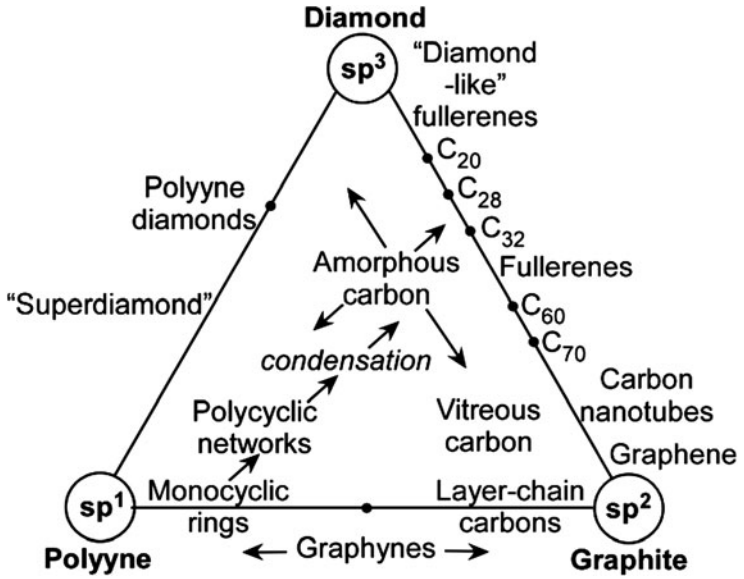


Fig. 4.3 A ternary phase diagram using sp^1, sp^2, and sp^3 hybridization of carbon as a basis for classification of known allotropes and intermediate forms of carbon (drawn after R.B. Heimann *et al*. Carbon **35**, 1654 (1997).)

rules than carbon bonding in organic chemistry that was developed over the past couple of centuries. The reason for such great variety in carbon structures is in the electronic structure of elemental carbon that can exist in three different electronic states determined by the hybridization of the carbon atomic orbitals [20].

Carbon is the sixth element of the Periodic Table. The six electrons of carbon occupy the 1 s^2, 2 s^2, and $2p^2$ atomic orbitals, making carbon the first element in Group IV. The 2 s^2 and $2p^2$ orbitals are close in energy resulting in atomic orbital configuration consisting of 2 s, $2p_x$, $2p_y$, $2p_z$ orbitals with wavefunctions that can readily mix with each other to minimize the bonding energy of carbon with its neighbors [20]. This intermixing of atomic orbitals gives rise to three new electronic configurations designated by sp^3, sp^2, and sp^1 that are known as hybrid molecular orbitals. Hybridization determines the type of bonding of the carbon atoms with its neighbors. The unique spatial orientation of the hybridized orbitals provides a convenient classification scheme for carbon-based materials shown in Fig. 4.3 [21, 22]. The sp^3 hybridization corresponds to the well-known tetrahedral configuration in which carbon binds to four neighbors giving rise to 3D interconnectivity of carbon atoms that is found in diamond. The sp^2 bonding in which carbon atoms bind to three neighbors, also known as trigonal hybridization, gives planar structures found in graphite and graphene. Diagonal or sp^1 hybridization that gives linear molecules in which carbon binds to only two neighbors is found in polyyne chains.

While this classification scheme based on the pure hybrid orbitals of carbon provides a general framework for understanding the structural stability of traditional carbon allotropes (diamond and graphite), the transformations from one form of

hybridization to another give rise to a great richness of intermediate structures, many of which still remain undiscovered. The best known examples of intermediate hybridization structures are fullerenes and CNTs with $sp^{2+\eta}$, where $\eta > 0$. The value of η depends on the degree of curvature of the sp^2 network. The existence of these intermediate degrees of hybridization greatly increases the complexity of the product distribution in the synthesis of carbon materials, including CNTs [23]. Understanding the synthesis of carbon materials is further complicated by the fact that except for graphite—the thermodynamically stable form of carbon at room temperature—all other structures are metastable [3–5, 8, 9, 14, 16].

The metastability of the structures means that the product distribution is controlled by the kinetics. In such reactions a rich variety of carbon structures, not just CNTs, form [3–5, 14]. The reason why a large number of structures form is that the reaction is prevented from exploring the entire parameter space to settle into the minimum energy state. The width of the product distribution will depend on the magnitude of departure from the ground state conditions reflected in the energetic content of the reactants illustrated in Fig. 4.4. That is the reason why different synthesis methods (arc, laser evaporation, CVD) produce different quality, quantity, and purity of CNTs. The current synthesis methods are poorly suitable for producing carbon nanotubes because they are based on trial and error. This tutorial takes the view that the way to developing a reproducible synthesis method for CNTs is to better understand the synthesis chemistry that governs the carbon bonding in these structures.

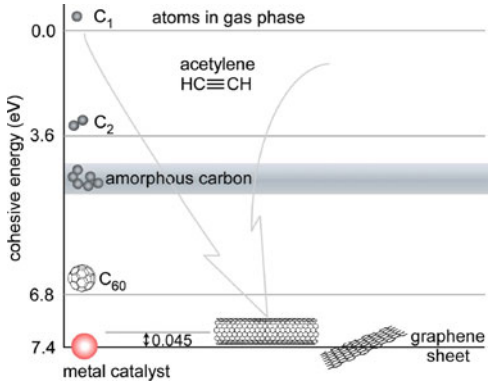


Fig. 4.4 The energetic stability of various carbon structures with respect to graphene. Introducing a curvature to form a (10,10) CNT has a low energy cost of 0.045 eV, compared to about 0.6 eV for C_{60}. Converting amorphous carbon and C_2 is even more energetically favorable. Acetylene is the primary hydrocarbon precursor for CVD growth of CNTs (drawn after B.Yakobson and R. Smalley, Fullerene Nanotubes: $C_{1,000,000}$ and Beyond, American Scientist, July-August 1997).

Atomic Structure and the Properties of Carbon Nanotubes

Understanding the atomic structure of CNTs is important because the properties of CNT materials are primarily determined by interactions that occur within a single

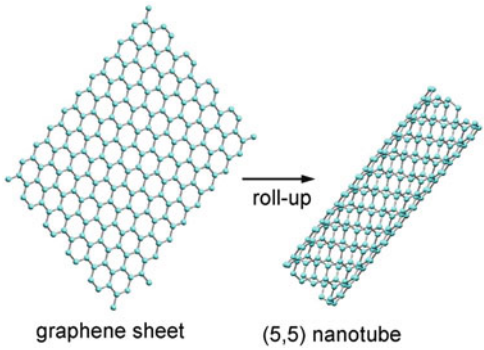


Fig. 4.5 Conceptual formation of a CNT by wrapping of a graphene sheet into a seamless cylinder.

graphene sheet (intralayer interactions) rather than between layers in multiwall CNTs, or between neighboring CNTs. Conceptually CNTs can be obtained by wrapping a single sheet of graphite (graphene) into a seamless cylinder illustrated in Fig. 4.5 [3–16]. As such, they are dominated by the properties of the sp^2 carbon that makes up the graphene sheets. In addition, small-diameter CNTs exhibit quantum confinement effects characteristic of 1D periodicity. The band structure and the electrical properties of CNTs are governed by the rotational and helical symmetry of the CNTs [6, 14, 15].

The properties of CNTs are classified in terms of the symmetry of their atomic structure. The starting point of this structure is the honeycomb lattice that is described by two lattice translation vectors. The chiral vector of the CNT connecting two equivalent points of the honeycomb lattice is given as a linear combination of the two basis vectors designated by a pair of integers (n, m). A CNT in general is described by the diameter of the tube and the angle of the chiral vector that specifies how a graphene sheet is rolled up into a CNT [6, 14, 15].

Ideally this cylindrical structure is terminated at both ends by a pair of caps that when joined together form a fullerene [3–6, 14]. This description represents a convenient way to illustrate the range of CNT structures using the example of a CNT derived from a well-known fullerene such as C_{60} shown in Fig. 4.6. The CNT is constructed by bisecting a C_{60} molecule at the equatorial line and using it to cap a cylindrical segment made from a wrapped graphene sheet that perfectly connects to the edges of the C_{60} caps. If the C_{60} molecule is bisected perpendicular to the fivefold axis the "armchair" tube designated by (5, 5) results, and if the C_{60} molecule is bisected normal to a threefold axis the "zigzag" tube designated by (9, 0) is formed. These two characteristic CNT structures determine a range in chiral angles that spans 30°. Of course many other armchair and zigzag CNTs with larger diameters having larger fullerene caps can exist. In addition, a large number of chiral CNTs exist with a chiral angle in the range $0 < |\theta| < 30°$, with a screw axis along the CNT axis with matching fullerene caps. The properties of the CNTs are classified according to the mod 3 rule. All armchair ($n = m$) CNTs are metals at room temperature. CNTs with $n - m = 3j$ ($j = 1, 2, 3, \ldots$) are semimetals. All

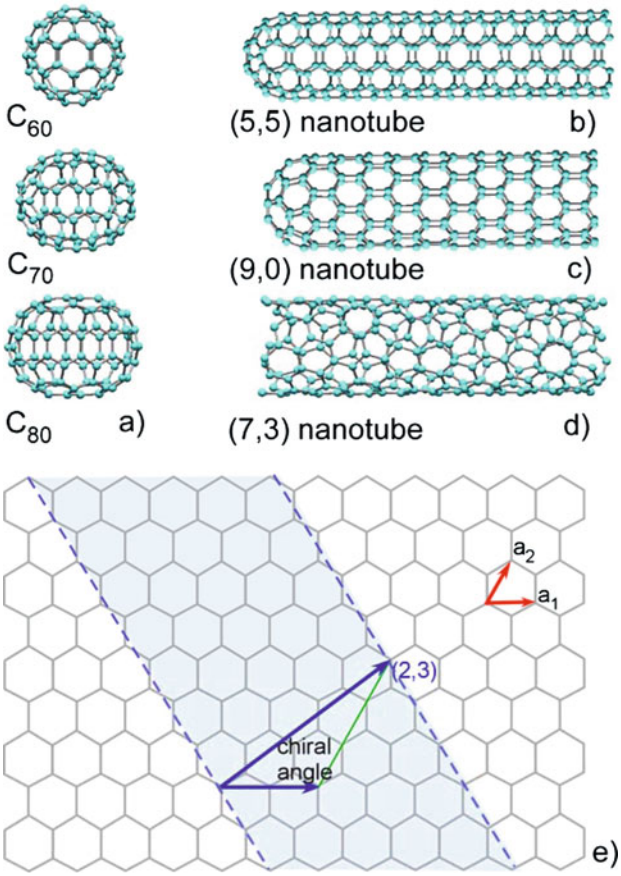


Fig. 4.6 Illustration of carbon nanotubes derived from C_{60} by addition of extra layers consisting of a belt of 10 carbon atoms **(a)** between the two end caps terminating the CNT. Depending on the orientation of the cut bisecting a single armchair (5,5) CNT **(b)** or a single zigzag (9,0) CNT **(c)** or a large number of chiral CNTs **(d)** can form. **(e)** Illustration of a graphene sheet and the parameters for rolling it into a CNT. (drawn after Dresselhaus et al. Science of Fullerenes and Carbon Nanotubes)

other CNTs with $n - m = 3j + 1$ and $n - m = 3j + 2\,(j = 0,\ 1,\ 2, \ldots)$ are semiconductors with a band gap that is inversely proportional to the diameter [3–16].

The General Principles of Carbon Nanotube Synthesis

All current synthesis methods for CNT production evolved through trial and error [3–5, 8, 14, 16]. The general approach is to use some energetic source to create a high concentration of reactive carbon species illustrated in Fig. 4.4 that are allowed to equilibrate at some rate that is governed by the temperature of the background. In

the process of thermalization the reactive carbon species assemble into carbon structures of increasing size that can form during the finite time period that the system spends while sweeping across the entire range of the energetic stability landscape.

The three main synthesis methods include arc discharge, laser vaporization, and chemical vapor deposition (CVD). The key elements of carbon reaction chemistry in these three methods are in principle similar. However, in practice the processes are different by the degree of the initial excitation that starts the nanotube growth process and the intermediate stages that the reactions go through during approaching the final state. A direct consequence of the high temperatures in arc discharge and laser vaporization that are required to evaporate solid carbon is the formation of a large variety of other carbon structures in the final product along with single-wall CNTs (SWCNTs). In addition, it was found that addition of a small amount (1%) of transition metal (Fe, Co, Ni) was necessary to obtain SWCNTs.

The typical byproducts of these methods include a large amount of fullerenes, a variety of graphitic structures, amorphous carbon in a form of particles or coating

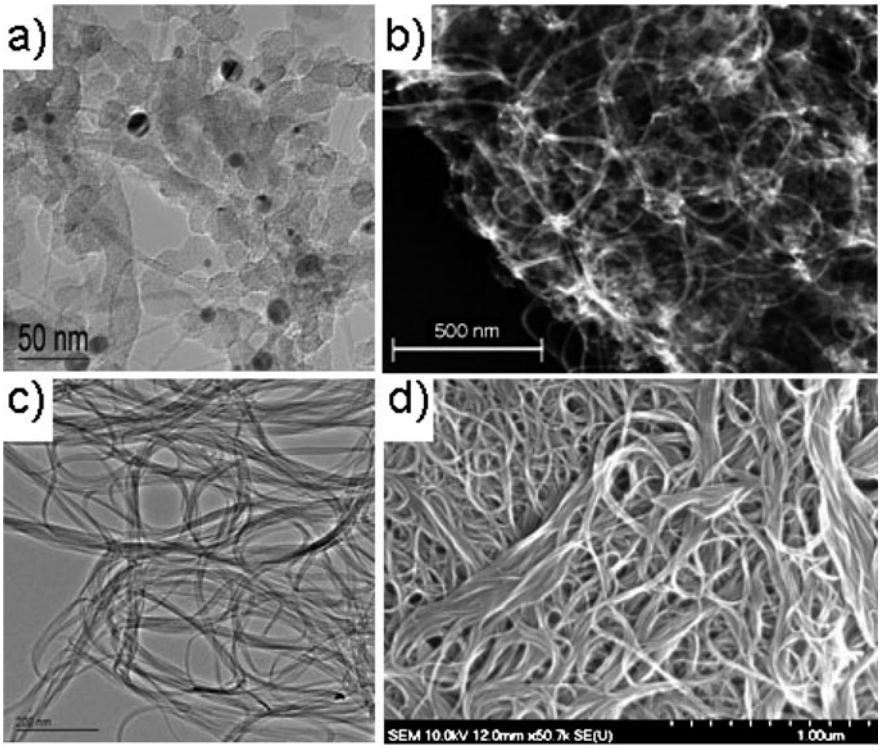


Fig. 4.7 Ultrahigh purity SWCNTs are an essential starting point for nanotube applications. Top row shows as grown CNTs using laser vaporization. Bottom row shows purified CNTs with less than 0.02 weight% metal content. TEM images **(a)** and **(c)**, SEM images **(b)** and **(d)**. (D.B. Geohegan unpublished)

on the walls of the CNTs. The main drawback of both arc discharge and laser vaporization is that a purification process must be performed to extract the SWCNTs [3–5, 8, 16]. The purification step illustrated by SEM and TEM images in Fig. 4.7 is aimed at removing the carbon side products and the residual metal catalyst particles. In addition to being labor-intensive the acids and harsh chemicals used in purification along with removing the side products often attack and damage the CNTs. The purification step results in a significant loss of the SWCNT material often combined with irreversible changes of the intrinsic CNT properties.

Over the past decade CVD emerged as the leading synthesis method [4, 8, 14, 16]. The most attractive feature of the CVD method is that it is easy to adapt for large-scale production [24]. Several manufacturing facilities that started operating recently use CVD for producing tens to hundreds of tons of CNTs annually. CNT growth in CVD processes occurs by catalytic decomposition of hydrocarbon molecules. This process occurs at much lower temperatures than arc discharge and laser evaporation. The merits of CVD include a higher CNT yield and higher purity. Both of these advantages are made possible by the milder reaction environment that is used in CVD. The details of the reaction mechanism of CVD will be discussed in a later section.

Chemical Vapor Deposition

The ability to grow high-quality SWCNTs has made CVD the method of choice for CNT synthesis [4, 8, 13, 14, 16]. Currently the number of papers published on CVD synthesis of CNTs by far exceeds all other methods combined and is still growing exponentially. This tutorial is focused on understanding how the mechanism of growth affects the product distribution specifically in CVD growth of CNTs. CVD has been adapted for CNT growth from the carbon fiber community. As a consequence of this connection the framework for understanding CVD growth of CNTs was not developed from the ground up but instead it was transferred from the carbon fiber community without critical validation [4, 9, 14, 16].

The mechanism of carbon fiber growth is based on the concept of the vapor–liquid–solid or VLS growth of whiskers. According to this mechanism illustrated in Fig. 4.8 carbon fiber growth occurs by precipitation of carbon from a molten metal particle [25]. Diffusion of carbon through the particle was identified as the rate-limiting step based on the similar value of the activation energy for fiber growth and carbon diffusion in the molten metal. This picture was subsequently used as the starting point for understanding CVD growth of CNTs where it is also known as the diffusion/precipitation model [3–5, 8, 14, 16]. The particle size in this model plays a central role. It is assumed that the particle size determines the diameter of the CNTs, and some versions of the model go as far as assuming that the particle is also capable of controlling the chirality of the CNT.

The main assumptions of the diffusion/precipitation model have never been directly proven in experiments and remain a source of controversy in CNT

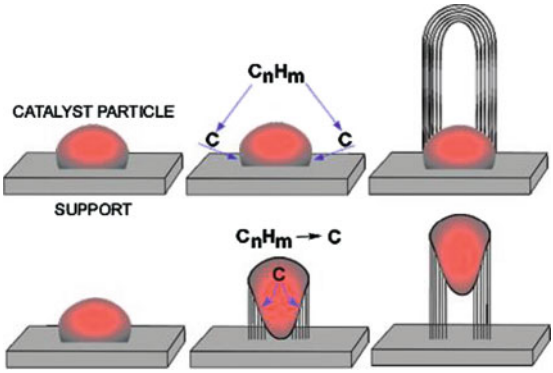


Fig. 4.8 Two possible growth modes according to the vapor-liquid-solid model in which CNTs form by dissolution and precipitation of carbon from metal particles. In the base mode (top row) the particle remains at the surface, while in the tip mode (bottom row) the particle is lifted up by the growing CNT. (drawn after R .T. K. Baker and P.S. Harris, Formation of Filamentous Carbon, Chemistry and Physics of Carbon, vol. 14, p. 83)

growth. The emergence of vertically aligned forest growth enables more controlled growth studies, which produced numerous observations and data that cannot be explained by the diffusion/precipitation model [26]. The main weakness of the diffusion/precipitation model is that it grossly oversimplifies the process by ignoring the role of carbon chemical bond formation. It is not surprising that a model based on the bulk parameters of the metal catalyst and carbon is unable to account for the great complexity of carbon structures that form and evolve on a molecular level in CNT growth.

The main parameters in CVD growth are the carbon source gas, the metal catalyst particles, the temperature, and the pressure of the process. There are numerous variations in the basic concept of CVD in which plasma or other means are used to break down partially or completely the molecular structure of the carbon source gas to increase its reactivity in the process. In addition, the catalyst particles can be supplied as free particles floating in the reactor or as metal catalyst films supplied on suitable substrates. These two modes of catalyst supply are used in different growth environments resulting in dramatically different carbon nanotube products.

The floating catalyst method has been the first adaptation of CVD for growing CNTs. We briefly describe the salient features of this method because it is the approach that is in some form used in all commercial production of CNTs by CVD. The carbon source gas and a chemical source that is capable of producing metal particles in the gas phase are introduced simultaneously in the CVD reactor. The decomposition of the carbon source gas provides carbon that is deposited on the metal catalyst particles. The floating catalyst method produces CNTs in a form of a random entangled mat (spaghetti) intermixed with other carbon structures still

containing the metal catalyst particles. The fraction and quality of the CNTs in each batch in commercial production is determined by how tightly the production parameters can be controlled.

The mechanism of carbon assembly in the floating catalyst CVD method is very similar and shares many common features with the direct carbon evaporation methods. The most remarkable feature of the carbon evaporation methods is that the carbon structures nucleate before the metal particles are formed [27]. Real-time spectroscopic and optical scattering studies shown in Fig. 4.9 were used to measure the time scales of CNT formation in expanding carbon laser plumes. These studies reveal a surprising finding that self-assembly of carbon structures precedes the formation of metal particles in the expanding and cooling laser plume. This sequence implies that the carbon structures must condense on the metal particles, a step that prevents premature closing of the growing carbon network and enables addition of carbon that is necessary for increasing the length of the CNTs. This extremely intriguing fact has important implications for understanding how CNTs form in general. It indicates that carbon bond formation is an important driving force that is key to self-assembly of carbon from highly reactive carbon species. However, the significance of these results has not been recognized and embraced by the CVD community.

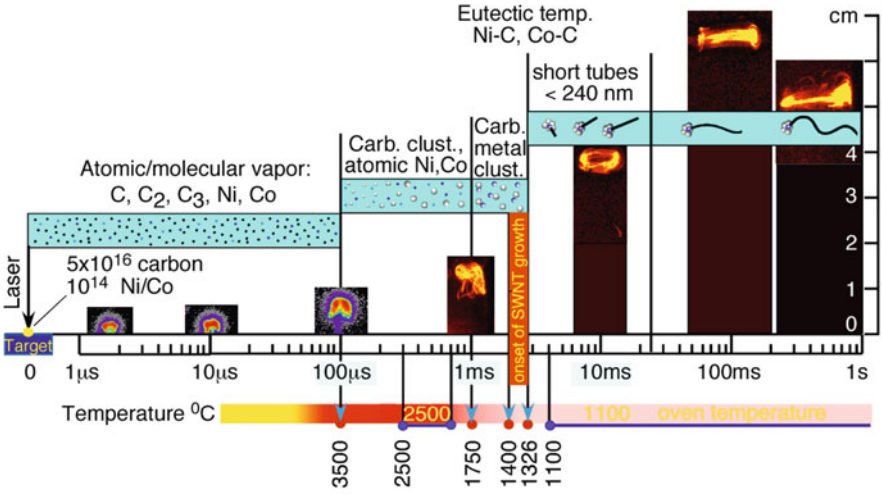


Fig. 4.9 Real-time imaging of the formation of CNTs from carbon vaporized by a pulsed laser. Carbon condensation and formation of clusters occurs 200 μs after the pulse. The metal catalyst atoms condensation starts much later, at 0.8 ms after the pulse, and is complete after about 2 ms. After 4 ms the plume is fully thermalized, but the CNTs are still very short, about 240 nm. The majority of growth, starting at the red vertical bar, occurs at thermal oven temperature over much longer times. (reused with permission from A.A. Puretzky *et al*. PRB **65**, 245425 (2002). ref. [27] Copyright 2002, The American Physical Society)

CVD Growth of Application-Specific CNT Structures

The greatest advance in CNT CVD has been the ability to grow arrays of aligned CNTs on substrates of choice. These structures enable applications that can take full advantage of the highly anisotropic thermal and electrical properties of CNTs. The CNTs are grown directly on the substrates of interest and the alignment can be controlled vertically or horizontally with respect to the substrate surface.

The advantage of this approach is that it eliminates the need for purification and processing of CNTs for a particular application. Of course, the growth of CNTs in a specific architecture brings new challenges for the growth process. However, an unanticipated benefit from growing such structures is that they enable more controlled measurements for probing the growth process that allows kinetic studies that lead to better understanding of the fundamental growth mechanisms [26, 28].

Vertically Aligned Nanotube Arrays

The most widely studied aligned CNT structures are the vertically aligned nanotube arrays (VANTA) [26, 28, 29]. VANTA growth is performed on a substrate coated with a catalyst film or a layer of catalyst particles. Upon heating, the film breaks up into particles that are the right size for nucleation and growth of CNTs [28]. The particle size distribution is affected by the temperature and composition of the catalyst films. A bilayer metal film consisting of 10 nm Al and 0.5–1 nm Fe is most often used [28]. VANTA can be grown by both thermal and plasma-induced CVD, containing pure single-wall, pure multiwall, or mixed CNTs. This growth mode has important advantages for practical applications in reducing dramatically the level of catalyst metal and undesirable carbon structures in the VANTA. The metal particles are absent because growth occurs by the base mode in which the catalyst particles are anchored to the substrate and VANTA can be lifted off from the substrate to avoid residual metal contamination. The highly localized carbon incorporation reaction also helps controlling or suppressing undesirable side products from spontaneous and nonselective gas-phase pathways such as amorphous carbon and other pyrolysis products [28, 29].

The breakdown of the particle-centric model and the shortcomings of the diffusion/precipitation concept are nowhere more evident than in the growth of VANTA. A particularly perplexing feature of CNT growth is a spontaneous self-termination of the growth process. However, it is not widely appreciated that growth termination occurs in all CNT growth processes. Since it is easy to observe and confirm directly, VANTA growth became a prototype for studying growth termination phenomena in CNTs [26, 28–31]. In the context of the particle-centric model, growth termination occurs because of poisoning of the catalyst particle surface. The concept of poisoning here is not as specific as it is in heterogeneous catalysis where it was borrowed from, and means that the access to catalyst sites is blocked by adsorption of inactive parasitic species. In CNT growth catalyst poisoning is typically attributed to

amorphous or graphitic carbon that coats the particle surface and blocks access to the active sites.

Understanding and preventing catalyst poisoning became an important direction of study in CVD growth of CNTs. Of the various schemes that have been proposed, the addition of water to CVD growth has generated most attention [29]. The model for explaining the role of water is based on removing the carbon coating by etching to regenerate the catalytic activity of the particle. Since VANTA growth terminates despite addition of water this interpretation of the role of water in growth termination does not represent the final answer. Another interesting outcome of the addition of water is that VANTA growth accelerates in the early stages of growth. However, similar growth acceleration occurs in the growth of CNTs without water as well as in growth of other carbon structures.

This behavior implies that an alternative mechanism illustrated in Fig. 4.10 must be considered in which VANTA growth is governed by the carbon bonding chemistry rather than by dissolution and precipitation of carbon from metal particles [32]. The growing number of reports showing that materials other than transition metal particles are effective for growing CNTs serves as additional support in favor of this picture. Many of these materials are not catalysts and most of them such as the oxides do not dissolve carbon [33]. Carbon self-assembly from reactive carbon

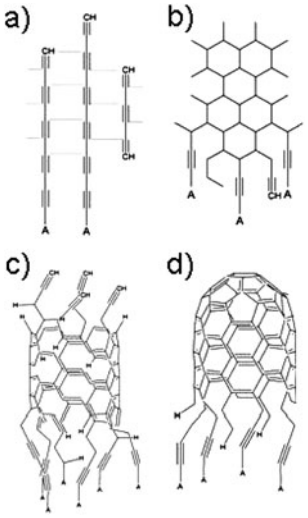


Fig. 4.10 A model for self-assembly of CNTs from acetylene. **(a)** Acetylene first forms high-temperature intermediates that have a short lifetime and a chain-like structure attached to the particle surface (A). **(b)** These intermediates cross link to form small graphene sheets that coalesce into incomplete tubular structures **(c)** that grow longer by addition of acetylene at the base of the growing CNTs. This model requires no metal to dissolve or precipitate carbon. The addition of carbon occurs at active sites in the carbon network. The nanotubes stop growing **(d)** when there are no more radical type carbon bonds available for further addition of carbon. (drawn after G. Eres *et al.* J. Phys. Chem. C **113**, 15484 (2009).)

species in reactions that are initiated by the action of the catalyst is a spontaneous result of energy minimization by formation of new carbon bonds. The reason that VANTA growth terminates in this picture is because the carbon incorporation reaction is carried to its completion when the active carbon species are exhausted [32]. The active carbon species include radicals that have incomplete coordination. These radical reactions propagate by chain reactions to produce more radicals until they are balanced by a stabilization mechanism referred to as termination. Specifically for VANTA, growth termination occurs when there are no more active sites directly at the base of growing CNTs for the incorporation of carbon.

The nature of the carbon source gas has been shown to also play an important role in the growth process. Small carbon containing molecules such as hydrocarbons or hydrocarbon derivatives are used typically as source gases [26]. Methane, ethane, ethylene, acetylene, and ethanol have been most widely used. But not all carbon molecules that can produce random CNTs can produce the high nucleation densities needed for VANTA growth [26]. The particle-centric model only considers carbon incorporation through the particle in a form of atomic carbon, ignoring the structure of the carbon source gas. However, recent experiments show that the structure of the carbon source molecule plays an important role in determining the density of the arrays, the type of CNTs in the arrays, and their structural perfection [26, 32]. The sensitivity of CNT growth to the molecular structure of the carbon source molecules indicates that the molecule is not breaking down all the way into atomic carbon in order to be incorporated into CNTs. Comparison of several families of small carbon containing molecules in a molecular beam VANTA growth experiments established that acetylene is the primary (closest) precursor for CNT growth [26]. In the molecular beam experiments the secondary gas-phase reactions are suppressed and growth occurs only from surface reactions of the incident molecules. In separate growth experiments it was also shown that if acetylene is not the starting molecule, it must form by intermediate reactions before CNT growth can occur.

The alignment in VANTA occurs by simple crowding among the growing CNTs, making the nucleation density the key factor that governs VANTA growth. The type of the carbon source gas, the catalyst composition, and the particle size are important factors for achieving a nucleation density that exceeds the threshold density for vertically aligned growth. The threshold for VANTA growth was found to be related to the incidence rate of the acetylene molecules [32]. Below some minimum acetylene flux there is no aligned growth. At the onset of growth the CNTs are intertwined much the same as in random mat growth. Aligned growth develops by lifting the random layer to the top of the arrays where it remains and continues to grow as a thin crust [31, 32]. The CNTs in VANTA are not an assembly of isolated tubes [26, 31, 34]. While there is an overall alignment of the CNTs in the growth direction, individual CNTs show a departure in order and alignment from the surface normal, and their properties, as Fig. 4.11 shows, vary with the distance from the base (thickness) of the arrays [26, 31, 34]. The order and alignment in the base region diminish with increasing thickness. Using scattering experiments it was shown that growth

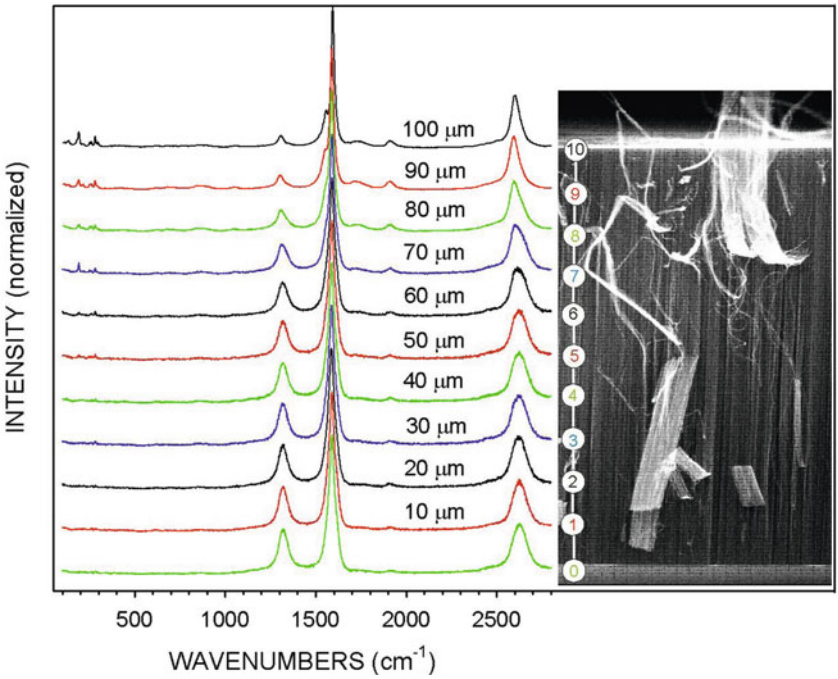


Fig. 4.11 Raman spectroscopy was used to characterize the VANTA grown by molecular beam CVD shown in the SEM image on the right. The spectra taken at 10 μm steps show weakening of the radial breathing modes and a falling G/D ratio as a function of distance from the base of the arrays where the catalyst layer is located. (drawn after G. Eres *et al*. J. Phys. Chem. B **109**, 16684 (2005).)

termination correlates with isotropic distribution of carbon and total loss of order and alignment in the base region of growth [31, 34].

The Role of Elements Other than Carbon

In addition to carbon containing species, CNT growth is also affected by the presence of other elements that serve to either enhance or hinder growth, and even alter the structure of the CNTs [26]. The effect of water is particularly intriguing [29]. The effects of water can be difficult to ascertain because it can interact with both the catalyst and the carbon incorporation reaction. Influence on the carbon incorporation reaction is based on the many reactions that are known to occur between water vapor and carbon gases at elevated temperatures. In these reactions water can act both as an inhibitor and as a promoter of molecular growth of carbon depending on other experimental factors.

Another way that water can affect CNT growth is by interacting with the catalyst particle. The recent work on water-assisted ethylene CVD finds that water is effective for keeping the catalyst particle surface clean by preventing overcoating of the particle surface by amorphous and graphitic carbon that is not incorporated into CNTs [29]. Whether particle overcoating is the cause or the consequence of growth termination is not clear. Water can also interact with the catalyst support [35]. Initially, catalyst support was considered an inert matrix that separates the active particles. However, recent work shows that the type of the catalyst support is critical for efficient CNT growth suggesting that it plays a more active role than previously thought [36].

It is intriguing that alumina (Al_2O_3) is found to be the most effective catalyst support [36]. This behavior is not unique to CNT growth [35]. Similar behavior of alumina is observed in heterogeneous catalytic processes, where it is used for creating large surface areas that enhance catalytic activity. In VANTA growth alumina is thought to be instrumental in preventing catalyst particle agglomeration and sintering. Monitoring the evolution of catalyst particles in real time by TEM imaging has revealed coarsening of the particle size distribution [37]. This ripening process is proposed as an alternative mechanism for explaining growth termination. The addition of water was found to slow down the ripening process, suggesting another mechanism by which water can enhance the growth of VANTA [35, 37].

Horizontally Aligned Nanotubes Arrays

In situ growth of a CNT bridging two metal electrodes is perhaps the most obvious application-specific CNT structure. These types of configurations are the key part of a device known as the CNT field effect transistor (CNT-FET) that was used to show that the conductivity of CNTs exceeds that of Si leading to predictions of unique CNT electronic devices [8, 11, 12, 14, 15, 19]. There are many challenges in routinely fabricating such single CNT devices in large quantities according to uniform specifications. To focus the discussion only on the CNT properties it is assumed that the electrical contacts to the CNTs are ideal. The following considerations are important in making such a device. The most critical issue is that currently there are no known methods for predicting whether a single CNT will be metallic or semiconducting [19]. Also, there are no known mechanisms to control the CNTs to bridge the electrodes in a predictable and reproducible fashion.

The need to simplify the manufacturing process of the CNT-FETs resulted in the development of a different device concept in which random CNTs are used to form a percolation network that connects the two opposite electrodes [38]. The advantage of this configuration is that the large variations in the individual CNT properties are replaced by the more uniform properties of the network. It was also shown that parallel paths increase the current-carrying capacity of such networks. A low-cost, solution-processing method is used for post-growth dispersion of the

CNTs on substrates. However, structures fabricated using these approaches exhibit vastly inferior conductivities in comparison with the intrinsic CNT conductivities [39]. Another problem is that the presence of metallic CNTs creates great complexity in the gating behavior of the devices. The simplest of these adverse effects is that the metallic tubes can short out parts of the network.

Direct growth of horizontally aligned CNTs on substrates by CVD is pursued as a potential way to get around most of these problems [40]. The CVD growth of horizontal CNT structures has some additional features that are different from those encountered in VANTA. First, as Fig. 4.12 shows, the CNTs grow parallel to the surface by crawling in molecular scale topological grooves that occur along specific crystallographic directions on a quartz surface. It is important to note that on thermally deposited, amorphous SiO_2 layers, only random CNT growth is reported. Second, a large variety of different elements including Fe, Co, Ni, Cu, Pt, Pd, Mn, Mo, Cr, Sn, Au, Mg, and Al were found to serve as catalysts in this growth method [40]. However, the mechanism of this alignment process is still hotly debated. The CNTs grow by the tip mode and, similar to the proposed mechanism for VANTA growth, the critical step for this growth behavior is also governed by carbon interactions and not by the properties of the metal particles. According to the proposed

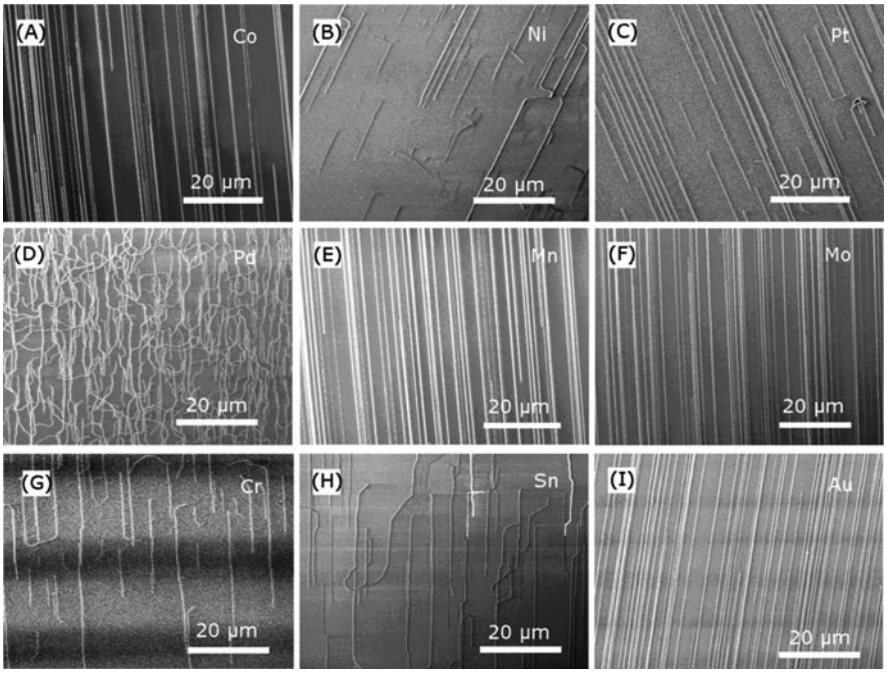


Fig. 4.12 SEM images show horizontally aligned CNT arrays on a quartz substrate formed by a variety of catalysts with the chemical symbol given in the image. (reused with permission from D. Yuan *et al*. Nano Lett. **8**, 2576 (2008). ref. [40] Copyright 2008, American Chemical Society)

mechanism it is the carbon shell that covers the outer surface of the particle which interacts with the substrate lattice to guide the growing CNTs along a preferential substrate lattice direction.

The Quest for Chiral Purity

Controlling the chirality of the CNTs is the ultimate goal in CNT synthesis [41–45]. Currently there are no synthesis methods that can directly produce CNTs with single predetermined chirality, also referred to as monodispersed CNTs. Instead, all CNT methods produce polydispersed CNTs, roughly according to the1:2 metal to semiconductor ratio dictated by the number of ways that a given number of carbon atoms can arrange into a CNT. However, steady progress toward this goal is being made from several different directions [41].

The most widely used approach is post-growth purification and separation of the CNTs into metal and semiconductor fractions [41]. These methods rely primarily on the delicate difference in the side wall chemistry of the metallic and semiconducting CNTs to generate selectivity. The diameter of the CNTs is another important factor that affects the selectivity. Following selective functionalization, the CNTs are physically sorted by using a variety of techniques including ultracentrifugation, electrophoresis, and chromatography. The success of separation is dictated by the selectivity of the functionalization reaction. These approaches are effective for producing research-scale quantities and are not expected to provide a solution for commercial production. In addition, the chemical reactions with the CNT side walls can be difficult to prevent from altering the intrinsic properties of the CNTs. Any defects introduced unintentionally into the side walls are typically not reversible, rendering the sorted CNTs suitable only for limited range of applications that are insensitive to such defects. Selective destruction is a separation concept that as its name suggests applies brute force methods such as selective oxidation, large currents, or intense optical irradiation to break down and eliminate the undesired fraction (for example, metallic CNTs) of the CNT population.

The ultimate test for the understanding of the growth mechanisms of CNTs is the ability to control the chirality of CNTs in the synthesis process [42–44]. These approaches are encouraged by findings that the chirality distribution in several methods has been skewed toward armchair structures [42, 44]. The justification for this trend is found in the atomistic models for self-assembly of carbon into CNTs showing that armchair CNTs are energetically more stable than zigzag CNTs. However, the chirality distribution at high temperatures is also affected by the energy spread in the vibrational degrees of freedom given by kT, where k is the Boltzmann constant and T is the temperature. The thermal energy at CNT synthesis temperatures is much larger than the energy spread among chiralities making selectivity at high temperatures doubtful.

Experimentally, diameter and chiral selectivity are currently performed by meticulously searching the growth parameter space [43, 44]. The particle-centric

model offers some guidance. This model implies that the nature and size of the catalyst particle are instrumental for achieving chiral selectivity. Accordingly, the development of new catalysts and methods for controlling the catalyst particle size and its reactivity have been actively pursued. Some success has been reported using Fe [43] and CoMo bimetallic catalysts [44] in CVD methods. Other growth parameters that were reported to affect the chirality distributions include the carbon source molecule, pressure, temperature, flow velocity, and the type of the inert carrier gas. However, in all cases increased selectivity was achieved at lower growth temperatures sacrificing production yield.

A different approach analogous to epitaxial growth has been explored for circumventing statistics from dictating the CNT distributions in synthesis [45–47]. This method illustrated in Fig. 4.13 is based on using CNT segments with specific chirality as seeds for growing long CNTs. The success of this cloning approach is again determined by understanding and controlling the growth mechanism, in particular the carbon incorporation step. The important steps here are restarting and continuing the growth to increase the length of existing CNTs [46, 47]. Some progress in cloning and continued growth has been reported using single CNTs and small bundles of CNTs. However, these methods have not produced the expected breakthrough needed for achieving monodispersed CNTs.

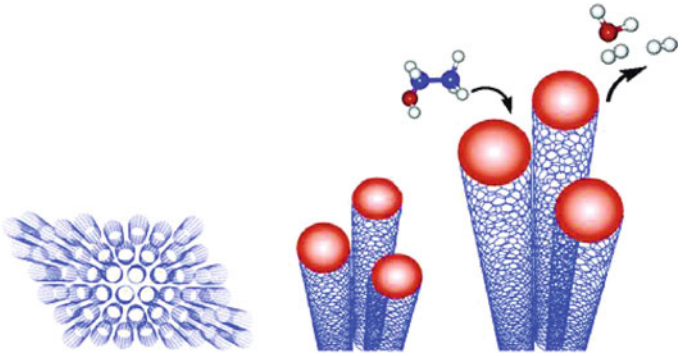


Fig. 4.13 Illustration of continued growth from ordered arrays of open-ended SWCNTs in a way analogous to epitaxy. Nanometer-sized metal catalysts were docked to the SWCNT open ends and subsequently activated to restart growth. SWCNTs thus grown inherit the diameters and chirality from the seeded SWCNTs. (reused with permission from Y. Wang *et al*. Nano Lett. **5**, 997 (2005). ref. [46] Copyright 2005, American Chemical Society)

Conclusions

The synthesis of CNTs is part black magic and part science. Over the past decade it became clear that mass production of high-quality CNTs needed for commercial applications is more challenging than initially expected. The main problem is in

the difficulty of producing individual nanotubes with exactly the same properties in large quantities. The product nonuniformity in the synthesis process is imposed by two different factors. First is the intrinsic distribution of nanotubes with different structures and different properties characterized by their chirality. Second, the innumerable structures that carbon can form by bonding to itself mean that other structures than CNTs must also form and makes selectivity harder to control and achieve. These two factors impose randomness on the outcome of the synthesis reactions that is manifested in unpredictable CNT distributions along with undesirable side products such as amorphous carbon and other carbon structures. The effect of these intrinsic nonuniformities is further exaggerated by the fact that there are no known quality standards that CNTs intended for the marketplace are expected to meet.

The main reason the synthesis process is poorly controllable is that the framework for understanding CNT growth based on bulk parameters of carbon and catalyst particles is woefully inadequate for treating the atomic- and molecular-level mechanisms of carbon network formation in CNT growth. In a radical departure from the current practice the focus should be on how to use catalysts to achieve selectivity of carbon structures on the level of individual carbon bond formation. The complexity of the intermediate steps might prevent determining the exact reaction mechanism. However, knowing the type of the rate-limiting reaction step can often be sufficient to determine the product distributions. The extreme reaction environments currently used are inadequate and are clearly not suitable for solving the quality and reproducibility problems in CNT production. There are numerous examples in organic and polymer chemistry, which demonstrate that carbon–carbon bond formation occurs and can be controlled in mild reaction environments [48]. The added advantage of this approach is that it can render unnecessary the extensive and laborious post-growth processing steps that are currently the only way for accessing the unique properties of CNTs.

Acknowledgment Research sponsored by the Division of Materials Sciences and Engineering, Office of Basic Energy Sciences, U.S. Department of Energy.

References

1. Monthioux, M., Kuznetsov, V.: Who should be given the credit for the discovery of carbon nanotubes? Carbon **44**, 1621 (2006)
2. Iijima, S.: Helical microtubules of graphitic carbon. Nature **354**, 56–58 (1991)
3. Dresselhaus, M.S. Dresselhaus, G., Eklund, P.C.: Science of Fullerenes and Carbon Nanotubes: Their Properties and Applications. Academic, San Diego (1996)
4. Ebbesen, T.W.: Carbon nanotubes: preparation and properties. CRC Press, Boca Raton (1997)
5. Endo, M., Iijima, S., Dresselhaus, S.M.: Carbon Nanotubes. Carbon, vol 33, Pergamon, Elsevier, Oxford (1996)
6. Saito, R.: Physical Properties of Carbon Nanotubes. World Scientific, London. Imperial college press (1998)
7. Tománek, D., Enbody, J.R.: Science and application of nanotubes. Springer, Newyork (2002)
8. Dresselhaus, S.M., Dresselhaus, G., Avouris, Ph.: Carbon nanotubes: synthesis, structure, properties, and applications. Springer, Berlin Heidelberg (2001)

9. Biró, P.L., Bernardo, C.A. Tibbetts, G.G., Lambin, Ph.: Carbon filaments and nanotubes: common origins, differing applications? Kluwer, Dordrecht (2001)
10. Meyyappan, M.: Carbon nanotubes: science and applications. CRC Press, Boca Raton (2005)
11. Reich, S., Thomsen, C., Maultzsch, J.: Carbon nanotubes: basic concepts and physical properties. Wiley-VCH, Weinhein (2004)
12. Rotkin, S.V., Subramoney, S.: Applied physics of carbon nanotubes: fundamentals of theory, optics and transport devices. Springer, Berlin Heidelberg (2005)
13. O′Connell, M.J.: Carbon Nanotubes: Properties and Applications. CRC Press, Boca Raton (2006)
14. Loiseau, A., Launois, P., Petit, P., Roche, S., Salvetat, J.-P.: Understanding Carbon Nanotubes: From Basics to Applications. Lecture Notes in Physics. Springer, Berlin Heidelberg (2006)
15. Léonard, F.: The physics of carbon nanotube devices. William Andrew, Norwich (2009)
16. Harris, J.P.F.: Carbon Nanotube Science: Synthesis, Properties and Applications. Cambridge, (2009)
17. Palmer, D.J.: Where nano is going. Nanotoday **3**, 46 (2008).
18. Baughman, R.H., Zakhidov, A.A., de Heer, W.A.: Carbon Nanotubes—the Route Toward Applications. Science **297**, 787 (2002).
19. Avouris, Ph.: Carbon Nanotube Electronics and Photonics. Phys Today, 34 (January 2009)
20. Pauling, L., The nature of the chemical bond, Cornell University Press, Ithaca (1960)
21. Heimann, B.R., Evsyukov, E.S., Koga, Y.: Carbon allotropes: a suggested classification scheme based on valence orbital hybridization. Carbon **35**, 1654 (1997).
22. Crespi, H.V.: The geometry of nanoscale carbon. In: Di Ventra, M., Evoy, S., Heflin, R.J. (eds.) Introduction to Nanoscale Science and Technology. Springer, Newyork (2004)
23. Ebbesen, T.W., Takada, T.: Topological and sp^3 defect structures in nanotubes. Carbon **33**, 973 (1995).
24. Cassel, M.A., Raymakers, A.J., Kong, J., Dai, J.H.: Large scale CVD synthesis of single-walled carbon nanotubes. J. Phys. Chem. B **103**, 6484 (1999).
25. Baker, R.T.K., Harris, S.P.: Formation of Filamentous Carbon. Chemistry and Physics of Carbon, vol. 14, p. 83. Marcel Dekker, (1978).
26. Eres, G., Kinkhabwala, A.A., Cui, T.H., Geohegan, B.D., Puretzky, A.A., Lowndes, H.D.: Molecular beam-controlled nucleation and growth of vertically aligned single-wall carbon nanotube arrays. J. Phys. Chem. B **109**, 16684 (2005).
27. Puretzky, A.A., Schittenhelm, H., Fan, X., Lance, J.M., Allard, F.L., Jr., Geohegan, B.D.: Investigations of single-wall carbon nanotube growth by time-restricted laser vaporization. Phys. Rev. B **65**, 245425 (2002).
28. Puretzky, A.A., Geohegan, B.D., Jesse, S., Ivanov, N.I., Eres, G.: In situ measurements and modeling of carbon nanotube array growth kinetics during chemical vapor deposition. Appl. Phys. A—Materials Sci. Process. **81**, 223 (2005).
29. Hata, K., Futaba, N.D., Mizuno, K., Namai, T., Yumura, M., Iijima, S.: Water-assisted highly efficient synthesis of impurity-free single-walled carbon nanotubes. Science **306**, 1362 (2004).
30. Puretzky, A.A., Eres, G., Rouleau, M.C., Ivanov, N.I., Geohegan, B.D.: Real-time imaging of vertically aligned carbon nanotube array growth kinetics. Nanotechnology **19**, 055605 (2008).
31. Bedewy, M., Meshot, R.E., Guo, C.H., Verploegen, A.E., Lu, W., Hart, A.J.: Collective mechanism for the evolution and self-termination of vertically aligned carbon nanotube growth. J. Phys. Chem. C **113**, 20576 (2009).
32. Eres, G., Rouleau, M.C., Yoon, M., Puretzky, A.A., Jackson, J.J., Geohegan, B.D.: Model for self-assembly of carbon nanotubes from acetylene based on real-time studies of vertically aligned growth kinetics. J. Phys. Chem. C **113**, 15484 (2009).
33. Homma, Y., Liu, H., Takagi, D., Kobayashi, Y.: Single-walled carbon nanotube growth with non-iron-group “catalysts” by chemical vapor deposition. Nano Res. **2**, 793 (2009).
34. Wang, H., Xu, Z., Eres, G.: Order in vertically aligned carbon nanotube arrays. Appl. Phys. Lett. **88**, 213111 (2006).

35. Hass, C.K., Schneider, W.F., Curioni, A., Andreoni, W.: The chemistry of water on alumina surfaces: reaction dynamics from first principles. Science **282**, 265 (1998).
36. Noda, S., Hasegawa, K., Sugime, H., Kakehi, K., Zhang, Y.Z., Maruyama, S., Yamaguchi, Y.: Millimeter-thick single-walled carbon nanotube forests: Hidden role of catalyst support. Jap. J. Appl. Phys. Part 2 **46**, L399 (2007).
37. Amama, B.P., Pint, L.C., McJilton, L., Kim, M.S., Stach, A.E., Murray, T.P., Hauge, H.R., Maruyama, B.: Role of water in super growth of single-walled carbon nanotube carpets. Nano Lett. **9**, 44 (2009).
38. LeMieux, C.M., Roberts, M., Barman, S., Jin, W.Y., Kim, M.J., Bao, Z.: Self-sorted, aligned nanotube networks for thin-film transistors. Science **321**, 101 (2008).
39. Topinka, A.M., Rowell, W.M., Goldhaber-Gordon, D., McGehee, D.M., Hecht, S.D., Gruner, G.: Charge transport in interpenetrating networks of semiconducting and metallic carbon nanotubes. Nano Lett. **9**, 1866 (2009).
40. Yuan, D., Ding, L., Chu, H., Feng, Y., McNicholas, T.P., Liu, J.: Horizontally aligned single-walled carbon nanotube on quartz from a large variety of metal catalysts. Nano Lett. **8**, 2576 (2008).
41. Hersam, C.M.: Progress towards monodisperse single-walled carbon nanotubes. Nature Nanotechnol. **3**, 387 (2008).
42. Ding, F., Harutyunyan, R.A., Yakobson, I.B.: Dislocation theory of chirality-controlled nanotube growth. PNAS **106**, 2506 (2009).
43. Harutyunyan, R.A., Chen, G., Paronyan, M.T., Pigos, M.E., Kuznetsov, A.O., Hewaparakrama, K., Kim, M.S., Zakharov, D., Stach, A.E., Sumanasekera, U.G.: Preferential growth of single-walled carbon nanotubes with metallic conductivity. Science **326**, 116 (2009).
44. Bachilo, M.S., Balzano, L., Herrera, E.J., Pompeo, F., Resasco, E.D., Weisman, R.B.: Narrow (n,m)-distribution of single-walled carbon nanotubes grown using a solid supported catalyst. J. Am. Chem. Soc. **125**, 11186 (2003).
45. Smalley, E.R., Li, Y., Moore, C.V., Price, B.K., Colorado, R., Jr., Schmidt, K.H., Hauge, H.R., Barron, R.A., Tour, M.J.: Single wall carbon nanotube amplification: en route to a type-specific growth mechanism. J. Am. Chem. Soc. **128**, 15824 (2006).
46. Wang, Y., Kim, J.M., Shan, H., Kittrell, C., Fan, H., Ericson, M.L., W.-F. Hwang, Arepalli, S., Hauge, H.R., Smalley, E.R.: Continued growth of single-walled carbon nanotubes. Nano Lett. **5**, 997 (2005).
47. Yao, Y., Feng, C., Zhang, J., Liu, Z.: Cloning of single-walled carbon nanotubes via open-end growth mechanism. Nano Lett. **9**, 1673 (2008).
48. Diederich, F., Rubin, Y.: Synthetic Approaches toward Molecular and Polymeric Carbon Allotropes. Angew. Chem. Int. Ed. **31**, 1101 (1992).

Chapter 5
Two Routes to Subcellular Sensing

G.F. Cerofolini

Abstract The current development of microelectronics (and its projected shift to nanoelectronics) makes possible to sense many metabolites inside or around the entire cell with spatial resolution on the length scale of 10 nm. Although reachable, this goal is, however, not easy. This chapter describes which new architectures, technologies, processes, and substances are presumably required for that. Two different routes are described, involving in one case (the 'boron route') existing electronic devices, or exploiting in the other case (the 'crossbar route') the expected shift of paradigm likely resulting from the development of the crossbar structure.

A (Rational) Dream

An ambitious long-term goal of medicine is to make analyses and deliver drugs at cellular level. The availability of nanorobots able to satisfy this goal would "give physicians the most potent tools imaginable to conquer human disease, ill-health, and aging" [1]. What is interesting here is that order-of-magnitude feasibility calculations indicate that nanorobots are physically possible [2] and their use for cell diagnosis and therapy seems at the reach of nanotechnology. Nanorobots by themselves would, however, be useless unless they are organized in a complex system providing all the required functions. In my view, this system should be inspired by the natural immune system.

The immune system is a heterogeneous collection of relatively homogeneous families of specialized cells spread throughout the entire organism and devoted to the surveillance, recognition, and termination of hostile guests. A drawback of this system is the poor recognition of endogenous pathological cells like those eventually responsible for cancer or self-immune diseases. As far as these pathological states are usually manifested after the fertile period, there has been no selective

G.F. Cerofolini (✉)
CNISM and Department of Materials Science, University of Milano–Bicocca, Via Cozzi 53, 20125 Milano, Italy
e-mail: gianfranco.cerofolini@mater.unimib.it

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_5,

pressure for the development of the immune system against them. In this context it should be interesting to develop artificial devices working as artificial white blood cells addressed to the recognition and eventually the destruction of endogeneous pathological cells. Attempts in this direction are known: drug carriers (nanoparticles functionalized with suitable ligands able to dock and carrying potentially lethal substances for the ill cells) have already been prepared and used in cancer diagnosis and therapy [3–6].

Building an *auxiliary immune system* seems beyond present possibilities; however, the construction of a distributed surveillance system for the diagnosis of endogeneous diseases poorly detected by the immune system seems at the reach of current technology. To the best of my knowledge, no attempt towards the definition of such distributed surveillance system has hitherto been tried. Of course, this goal is extremely ambitious and is expected to require several decades. Nonetheless, sketching a scenario is not a mere speculation, but rather a useful exercise to identify the nature of problems posed by the definition of an auxiliary immune system.

A *nanorobot* is an artificial machine with overall size on the order of a few micrometres or less in all spatial directions and constituted by nanoscopic components with individual dimensions in the interval $1–10^2$ nm and able to perform sophisticated functions like navigation, recognition, and data storage and handling. Starting from the reasonable assumption that in the next 10 years integrated circuits (ICs) will attain a density on the scale of 0.1 Tbit cm^{-2} [7], a circuit built on an area of $10^2 \mu m^2$ (comparable with that of small eukaryotic cells) can host 10^5 devices that, functionalized for operating as chemical sensors, would allow the neighbourhood of each cell to be chemically mapped on a smaller space scale than the one characteristic of the organelles forming the cell. Of course, the transformation of a space-resolved chemical sensor into a nanorobot requires the addition of the intelligence required to recognize whether the chemical pattern is characteristic of healthy or ill cells, to control the motion, to manage power, and so on. Assuming for the moment that these nanorobots can actually be built, they can be used for the constitution of a distributed surveillance system.

The *distributed surveillance system* is constituted by two parts: a *central unit*, externally accessible but permanently resident in the organism (e.g. as an earring), and a 'colony' of nanorobots, each swarm of the colony being specialized to the different tissues and organs.

Each nanorobot is a self-propelled machine, taking energy from the environment, able to recognize and dock the target cell, to sense its membrane and neighbourhood, to recognize its health state, to store the information, to transfer it to the central unit, and eventually (once allowed) to destroy the malignant cell.

At the present stage of knowledge, the described nanorobot is certainly far from being producible, but it is not an (irrational) dream because most of the critical steps required for its preparation have already been established.

Consider a silicon-based chip with area comparable with that of a blood cell, say $10^2 \mu m^2$, and thickness of $1 \mu m$ (obtained, for instance, by back-etching from a device built on a silicon-on-insulator substrate) so shaped as to explore all capillaries avoiding the obstruction of the capillary lumen and the formation of blood

clots,[1] whose outer surface is almost completely coated with a biomimetic material designed to not activate the immune response of the organism. This chip should be able to explore the whole organism through the circulatory system.

In the hypothesis of a bit density of the order of 10^{11} cm^{-2}, a chip of area $10^2\mu$m^2 can host a logic of 100 kbit. This logic is expectedly able to manage the information coming from a few (of the order of 10) sensing regions with space resolution on the length scale of 10 nm, each specialized to the identification of a different metabolite. The choice of the set of metabolites is characteristic of the target tissue. A single elementary charge is expected to be able to switch ON–OFF or vice versa an underlying transistor of area 10 nm^2. Exploring the whole cell surface with a step of 10^3 nm would imply the collection of 10^3 data (10^2 regions times 10 metabolites) for the chemical map of the cell surface. Assuming that each measurement requires 0.1 s (including positioning), mapping the entire cell would require approximately 10^2 s. The repetition of 10 mappings would thus require 10^3 s and their memorization should occupy 10 kbit, about 10% of the entire circuitry.

If the time evolution of a normal cell is known, exploring it for a duration lasting approximately 10^3 s (about 1% of the characteristic cellular lifetime [8]) is likely sufficient to recognize whether the cell is undergoing a physiological or pathological path. Assigning the comparison to the portion of logic not involved in memory, the device will be able to establish (with a certain accuracy) the health state of the cell.

If the analysis of a cell requires indeed 10^3 s, in 1 year the nanorobot will be able to test approximately 3×10^4 cells; in a tissue of 1 kg it will be able to explore approximately 3 ppm of the whole tissue. To compare this situation with that of current preventative medicine, consider that in ordinary blood analysis the concentration of 10–10^2 metabolites is measured with a period of about 10^7 s and is sampled in a fraction of about 10^{-5} of the body volume. The number of nanorobots forming the swarm must be so chosen as to have an adequate statistical coverage of the tissue.[2]

Although the production of nanorobots for subcellular sensing is beyond the reach of current technology, almost all critical points required for their preparation have singularly been overcome.

Biomimetic coatings designed not to activate the immune response of the organism are at the reach of the current technology, as demonstrated by several demonstrators of lipid monolayers or bilayers (mimicking biological membranes) supported on solid or polymer surfaces on even large areas [9]. In most cases the supported films are only weakly bound and are thus poorly stable; a larger stability,

[1] The shape could be an ellipsoidal disk with major axis of 50μm, minor axis of 2.5μm, and height h of 1μm.

[2] An adult human body contains approximately 10^{14} cells. Assuming that each nanorobot does really sense 3×10^4 cells per year, the exploration in 1 year of 0.1% of the whole cell population would require 3×10^6 nanorobots. Assuming that each of them has a mass of 3×10^{-9} g (as follows from their volume), the total mass of circulating nanorobots should be 10^{-2} g (that gives an idea of the 'invasiveness' of the auxiliary immune system) with a total area in the range 1–10 cm^2 (comparable with that of ICs with giga-scale complexity).

however, can likely be achieved by bonding covalently (e.g. via silanization) to the nanorobot surface molecules with carboxylic terminations, mimicking the outer surface of cells.

Supplying power to such a complex system is not trivial and most likely requires different solutions at different levels. The power required for the central unit is macroscopic and may be supplied either by external sources (batteries, electromagnetic coupling) or by implanted generators (stochastic microelectromechanical generators). More difficult is to supply power to the nanorobots that, in view of their size ($10^2 \mu m^3$), do not tolerate such features.

For fuelling and propelling nanorobots, hybrid solutions must necessarily be found. An arrangement where the motion is imparted by the derivatization of the nanorobot with biomotors can be hypothesized: for instance, Montemagno and Bachand have reported the construction of nanomechanical devices powered by biomolecular motors [10]; Kim and Breuer have described the successful use of live bacteria as mechanical actuators in microfabricated fluidic systems [11]; and Behkam and Sitti have exploited bacterial flagella for propulsion and motion control of microscale objects [12]. An enormous advantage of this solution is the fact that the motion does not require an alien input of energy—for that the chemical energy available in the organism (in the form of adenosine triphosphate, ATP) can be exploited. The same form of energy can also be exploited for powering the electrical circuitry, because biomotors are reversible and can operate as engines too. Other approaches involve the derivatization of the nanorobot with externally operated synthetic flagella; for instance, Zhang et al. have designed and constructed artificial flagella with size and shape comparable to their organic counterparts and which can swim in an externally controllable fashion via weak applied magnetic fields [13].

Wireless transmission of data from the nanorobot to the central unit is, at present, seemingly insurmountable: radio-frequency transmission requires indeed an antenna whose size is most likely on the millimetre length scale. This practical impossibility would require that the nanorobot periodically returns to the central unit to feed it with the information on cell health state. That solution, however, poses dramatic technological problems. While the passive transport through the circulatory system requires only passive (though sophisticated) appendices and docking may be driven by physical chemistry (through the conjugation of the nanorobot to suitable specific ligands), localizing the targeted cell and returning to the local unit requires a sophisticated navigation system providing the nanorobot with its coordinates. However sophisticated the algorithms might be, a logic of 10^5 bits can hardly be considered sufficient for all the requested functions. Although the use of the crossbar structure (as described in Section 6) might increase the maximum allowed bit density by one order of magnitude, even 10^6 functionalized cross-points seem insufficient for so many functions. Increasing further the bit density requires technological breakthroughs. Of the various architectures considered, none seems to match the constraint on density except for the one exploiting DNAs working as cellular automata [14–16] for the functionalization of cross-points. Caution reasons (avoid the use of gene-hybridized devices in vivo [17]) discard, however, this route, but

rather suggest alternative approaches like the exploitation of the nanorobot swarm intelligence.

Allowing the nanorobot to recognize the health state of the cell with relatively few parameters (say 10) on relatively few regions (say 10^2) over a relatively short time (say 10^2 s) requires in any case a sufficiently simple built-in model of 'standard cell'. Of course, the construction of such models requires an enormous preliminary work that I intend to outline in the following.

The Metabolic Pattern

That metabolic diseases at the organism level have often familiar bases and may thus be reduced to genetic disorder is too well known to deserve a discussion. That a metabolic pattern at the lowest functional level (the cellular one) may be used as a tool for sensing genetic disorder (possibly culminating in cancer) is instead not so trivial—but the opinion that information at sub-cellular level may be used for preventative and curative medicine is gaining increasing consensus [18, 19].

The scope of the metabolic pattern covers the temporal, spatial, and chemical domains. The following examples sketch such domains.

The Temporal Domain

The life of somatic cells is scanned by a clock dividing the vital cycle into two phases: mitosis M and interphase [20]. In mammalian tissues they last 1–2 h and 12–24 h, respectively. During mitosis the deoxyribonucleic acid (DNA) undergoes replication, while during the interphase the cell is constantly synthesizing ribonucleic acids (RNAs), producing proteins and growing in size (see Fig. 5.1).

The interphase can be divided into four steps: gap 0 (G0), gap 1 (G1), synthesis (S), and gap 2 (G2). G0 is characteristic of cells that leave the reproduction cycle and quit dividing; this may be a temporary resting period or more permanent (as it happens for neurons). In stage G1 cells increase in size, produce RNAs, and synthesize proteins; a control mechanism activated during this period (G1 checkpoint)

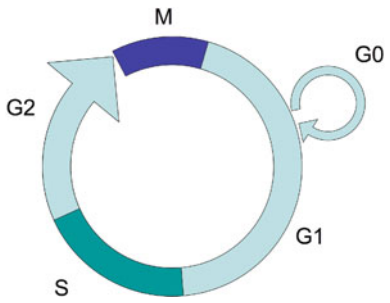


Fig. 5.1 The cellular clock

ensures that everything is ready for DNA synthesis. To produce two similar daughter cells, the complete DNA instructions in the cell must be duplicated; DNA replication occurs during S. During the gap G2 between DNA synthesis and mitosis, the cell will continue to grow and produce new proteins; at the end of this gap is another control checkpoint (G2 checkpoint) to determine whether the cell can now proceed to enter M and divide.

Cell growth and protein production stop at this stage in the cell cycle. All of the cell's energy is focused on the complex and orderly division into two similar daughter cells. As in both G1 and G2, there is a checkpoint in the middle of mitosis (metaphase checkpoint) that ensures the cell is ready to complete cell division. The relative durations of the various phases are characteristic of the cell; any deviation from these ratios is a candidate pathological marker.

The Spatial Domain

That the cell is a highly nonuniform and heterogeneous system is well known. This fact is reflected in a spatial variation of its surface properties. The characteristic length of such variations depends on the considered property: For instance, the membrane permeability of alkali cations is negligible over all the membranes except in special regions (ion channels) which is selectively high for sodium or potassium ions; the length scale for such variations is of tens of nanometres. If the attention is instead focused on the surface potential, the length scale is expectedly controlled by the Debye length of the aqueous medium to which the cell is exposed. Moreover, the cell diameter may be taken as an overall marker of the biological clock described in the previous part.

Even more structured is the internal organization of the cell: although the structure is matched to the function so that the same organism may be so specialized as to contain extremely branched cells (the neurons) as well as highly symmetric lens-like cells (red blood cells), all eukaryotic cells with nuclear materials contain the same kinds of organelles (nucleus, nuclear membrane, mitochondria, ribosomes, rough endoplasmatic reticulum, and smooth endoplasmatic reticulum, possibly specialized in the Golgi apparatus) whose spatiotemporal evolution controls the life functions.

The Chemical Domain

The spatial heterogeneity is the result of the extreme chemical heterogeneity of the cell. This heterogeneity is manifested with respect to the material (substance and phase) and to the biological function.

Although only a score of elements are necessary (in ponderable amounts or as traces) for the life of the cell, in practice any cell contains not only them, but also all other elements contained in the environment to which the cell has been exposed to and at concentrations depending on ambient concentrations.

Although the folding properties of proteins go beyond current computability, there is a wide agreement that the conformations of proteins are not, in general, the equilibrium ones, but are rather controlled by the growth conditions in the aqueous phase where the synthesis was carried out. In turn, the structure of the medium is determined by the interactions with the 10^3-10^4 mutually interacting different proteins (dominated by the heterogeneity of backbone and side-chain polar groups and characterized by allosteric transition and hysteresis effects [21]) contained in the cell.

Even the simplest protein (myoglobin) appears to be formed by a heterogeneous collection of molecules with the same composition but in different conformations [22]. Since the functional properties of proteins are controlled by their conformations, it follows that each individual molecule may be regarded as a chemical species, whose biological activity is potentially different from those of the other molecules with the same structural formula but different conformation.

Sensing as a Key Tool for Systems Biology

Molecular biology has hitherto been focused on the genome. That cellular behaviour, however, is not totally expressed by its genetic code alone is immediately understood by the variety of forms and functions (ranging, for instance, from nucleus-free blood red cells to mononuclear nervous cells or to polynuclear muscular cells in mammals) sustained by the same genome. One of the most important expressions of the cell is its metabolism.

Starting from Turing's pioneering ideas [23] and basing on the extended analysis of Prigogine and his school [24], it has become clearer and clearer that open thermodynamic systems subjected to diffusion–reaction processes and sufficiently far from equilibrium may develop non-homogeneous spatial structures and undergo periodic time oscillations. Metabolic reactions run in the previous case, so that they may be viewed as responsible for the behaviour commonly referred to as 'life'.

Any cellular system is necessarily associated with a metabolic field (sustaining the transport of anabolites from the environment into the cell and of catabolites in the opposite verse) and undergoes nearly periodic cycles. The distribution of metabolites inside the cell is manifested outside too, although controlled by the (inner and outer) biomedium structure and mediated by the passive or active transport through the membrane.

Life Functionals

The diffusion field generated by the metabolism transports catabolites outside the cell and depletes anabolites from the nearby medium. Sensing the region surrounding the cell is expectedly able to provide information on the metabolic pathways inside the cell.

It is a common place in the description of cells to consider the *expressome* (i.e. the set of functional and structural molecules that are the final result of the central dogma of biology) and the *metabolome* (i.e. the set of substances destroyed or produced in the cell metabolism) as separate entities, the first one acting as a regulatory agent of the second one. Since both functional and structural molecules are produced or degraded during the cellular metabolism, the distinction is mainly quantitative, being related to the lifetime of the considered species—high for 'expressites', low for metabolites [25, 26]. In the following it is convenient to consider metabolites M_i and expressites E_k, separately.

The concentration $C_i(\mathbf{x'}, t)$ at time t of any metabolite M_i outside the cell is expected to depend on the concentrations $C_j(\mathbf{x}, t)(j = 1, \ldots, J)$ of all the species M_j inside the cell and $C_j(\mathbf{x'}, t)$ $(j \neq i)$ of all the other species M_j:

$$C_i(\mathbf{x'}, t) = F_i\left[C_j(\mathbf{x}, t), C_j(\mathbf{x'}, t)_{j\neq i} | \mathcal{E}_k(\mathbf{x}, t)\right], \tag{5.1}$$

where $\mathbf{x}$ and $\mathbf{x}'$ denote internal and external points, and $\mathcal{E}_k(\mathbf{x}, t)$ $(k = 1, \ldots, K)$ denotes the local concentration of the kth of the K internal substances (nucleic acids or expressites) interacting with the metabolites. It is noted that the above partition between metabolites and expressites implies that from the topological point of view the cell must be considered as a closed set (with its boundary, the membrane, belonging to the set).

Each functional $F_i\left[\cdot|\cdot\right]$ is extremely complex, although in ultimate analysis it is nothing but the solution of the coupled diffusion–reaction equations, where diffusion coefficients, reaction rates and orders, etc. contain the internal degrees of freedom $\mathcal{E}_k(\mathbf{x}, t)$ as parameters. In turn, these concentrations may similarly be written in terms of other K functionals $\mathcal{F}_k\left[\cdot|\cdot\right]$:

$$\mathcal{E}_k(\mathbf{x}, t) = \mathcal{F}_k\left[\mathcal{E}_l(\mathbf{x}, t)_{l\neq k} | C_j(\mathbf{x}, t)\right]. \tag{5.2}$$

Needless to say, functionals $\mathcal{F}_k\left[\cdot|\cdot\right]$ are extremely complex too.

The major goal of systems biology is the determination of $F_i\left[\cdot|\cdot\right]$ and $\mathcal{F}_k\left[\cdot|\cdot\right]$ specifying in detail the reaction–diffusion equations for all the species in the cell.

Determining the Life Functionals

Imagine that one is able to map all metabolic species and monitor their time variation outside the cell; assume in other words that $C_i(\mathbf{x'}, t)$ is known for all i. If one may formulate a reasonable guess of $\mathcal{E}_k(\mathbf{x}, t)$ (through the accumulated knowledge on cell structure and function), Eq. (5.1) for known $F_i\left[\cdot|\cdot\right]$ may thus be viewed as an equation for all $C_i(\mathbf{x}, t)$. Although most likely this problem is improperly posed, its solution is expected to allow *the recognition of the inner cellular state from the outer chemical–physical state*.

Even assuming that recognition is possible, the ill posedness of the problem will result at most in a fuzzy information. Moreover, recognizing the health state of the cell (the major goal of nanorobots) will be possible only in the presence of a cellular model, sufficiently accurate to produce a realistic estimate of $F_i\left[\cdot|\cdot\right]$ and $\mathcal{F}_k\left[\cdot|\cdot\right]$. In this context the fundamental difference between metabolite and expressite is that expressite concentration outside the cell should be identically null during the duration $\mathcal{T}$ of the measurement:

$$\forall k(k=1,\ldots,K)\forall t(t\in\mathcal{T})\forall \mathbf{x'}(\mathbf{x'}\in\mathcal{D})\left(\mathcal{E}_k(\mathbf{x'},t)=0\right), \tag{5.3}$$

where $\mathcal{D}$ denotes a suitable neighbourhood of a cell.

In this line of thought, any violation of condition (5.3) may be considered as a marker of pathological state. In other words, let O be an organism occupying a region $\mathcal{D}_{\mathrm{O}}$; if there exists a domain $\mathcal{D}$ and a time interval $\mathcal{T}_0$ such that

$$\exists k(k=1,\ldots,K)\exists\mathcal{T}_0(\mathcal{T}_0\subseteq\mathcal{T})\exists\mathcal{D}(\mathcal{D}\subseteq\mathcal{D}_{\mathrm{O}})\forall t(t\in\mathcal{T}_0)\forall\mathbf{x}'(\mathbf{x}'\in\mathcal{D})\ \left(\mathcal{E}_k(\mathbf{x}',t)>0\right), \tag{5.4}$$

then the organism O has undergone a *pathological state*; the degree of pathology is determined by the extension of $\mathcal{D}$, the duration of $\mathcal{T}_0$, and the strength of the inequality.

It is stressed that the pathological state is not a local property of the cell. Actually for certain expressites, like transaminase or the prostate-specific antigen, condition (5.4) holds true even at the organism level (being markers for liver or prostate disease, respectively).

Of course, condition (5.4) does not exhaust all pathological states and may seem somewhat too formal; however, succeeding in defining formally the possible pathological states means that this information can be wired in a medical nanorobot.

In view of the inherent complexity of the theoretical description of cell behaviour, all cell models are wrong, and some more than others. The selection of less wrong, potentially useful, models can be done sensing the cellular behaviour and selecting (essentially via genetic algorithms) the simplest model accounting for observed behaviour within an assigned accuracy.

Clearly enough, increasing the level of sensing accuracy (finer space localization, longer time scale, greater number of parameters, etc.) will select a smaller class of descriptions, so that one can imagine a roadmap where the model is improved (starting from a short time scale, poor space resolution, and few parameters) with a progressive

(i) greater duration of the observation,
(ii) refinement of the spatial resolution, and
(iii) increase of the number of parameters.

The *first goal* is devoted to the verification of the ability of the model to recognize whether from an observation on a short time base one can extract a sufficiently detailed information on the whole cellular life. This goal is attacked mainly on the

theoretical and computational grounds and does not seem to pose severe limits to technology. The *second goal* would be easier if the scaling law of the probe were known. In that case, indeed, one could imagine a relatively slow evolution, over a time of tens of years, somewhat similar to what has happened for the ICs. The *third goal* requires the ability to produce sensors that are sensitive to more and more cellular physico-chemical parameters (membrane potential, osmotic pressure, *p*H, ion nature and concentration, etc.). This requires the development of derivatization processes with more and more sophisticated functional molecules (e.g. bases for *p*H, crown ethers for alkaline ions, carcerands for anions, molecular motors for osmotic pressure).

Sensing and actuating neurons in hybrid combination with electronic circuits is expectedly one of the first applications. The major problems here come from the fact that mature neurons in vitro survive for very short time, so that most of the work has been carried out with partially undifferentiated neuroblastomas (see [27] for an early work in this field). Once one has learnt how to allow the neuron to survive in vitro, the sensing of its membrane potential with circuits (functionalized to allow the adhesion to, and to be sensitive to the ion distribution on, the neuron membrane) with resolution on the 10-nm length scale will succeed in its determination on a length scale approximately coinciding with the nearest neighbour distance between the ion channels that control cell excitability.

Subcellular sensing, however, has so far been the object of occasional interest only, and a roadmap indicating which actions are required to increase the level of sensing accuracy (finer space localization, greater number of parameters, etc.) is still missing. This work is targeted at the formulation of two strategies for attacking the problem of subcellular sensing.

The Boron Route to Subcellular Sensing

The IC technology allows the preparation of planar matrices of devices with deep-submicrometre resolution; this fact allows since now, at least in principle, subcellular sensing. However, the exploitation of the IC technology for that purpose has not been revealed so easy.

Early Attempts at Subcellular Sensing

Without pretending to assign any priority, one of the first examples where the high spatial resolution offered by the IC technology could be used for solving biological problems was perhaps proposed in [28]: In dynamic random-access memories (DRAMs) the information is stored as metastable charge packets in their memory cells. The passage of an α particle in semiconductors produces a column of electron–hole pairs along the particle track (in silicon, the α particle is stopped releasing an average energy of 3.6 eV per pair); in DRAMs this charge produces the destruction of the information in the memory cells crossed by the particle (*single*

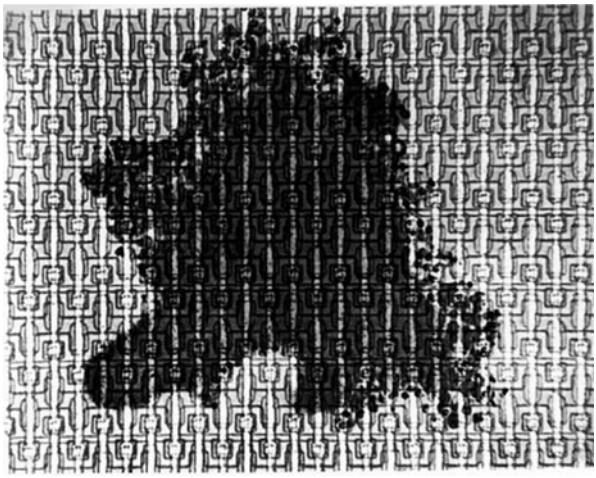


Fig. 5.2 Microphotograph of an amoeba covering approximately 40 memory cells of a DRAM

event upset). On the contrary, thermal neutrons have a negligible probability of interacting with semiconductor silicon, so that DRAMs are nearly insensitive to thermal neutrons. While the cross section for the interaction of thermal neutrons with silicon is fraction of barns, it is instead of the order of 10^3 barn for ^{7}Li and ^{10}B and in both cases the interaction results in the production of α particles:

$$^7\mathrm{Li} + {}^1n \longrightarrow 2\alpha, \tag{5.5}$$

$$^{10}\mathrm{B} + {}^1n \longrightarrow {}^7\mathrm{Li} + \alpha. \tag{5.6}$$

Starting from these observations, Cerofolini et al. suggested that information on boron or lithium distributions inside any cell could simply be obtained depositing the cell onto a DRAM, activating reactions (5.5) and (5.6) by exposing the resulting structure to a stream of thermal neutrons, and detecting which memory cells give single event upsets. Fig. 5.2 shows that this kind of radiography was able to locate (when the method was proposed, in the early 1980) the considered elements to within $\frac{1}{40}$th of amoeba.

Updating the Early Attempts

The biological relevance of lithium and boron is modest, and perhaps for this reason the suggestion of [28] had no practical implementation. This state of affair might perhaps change in the light of the following observations:

(i) there are pharmaceutical drugs containing boron whose biological activity is known (e.g. boronic acids $RB(OH)_2$, with R being an alkyl or aryl group, that are inhibitors of membrane β-lactamase enzymes);

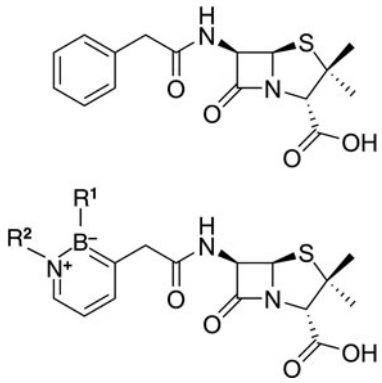


Fig. 5.3 Penicillin G (*top*) and its possible modification (*bottom*) with the insertion of boron on the side chain, however, preserving the biological function

(ii) pharmaceutical drugs with known biological activity may be modified via the insertion of boron in their structure without any a priori modification of their activity (e.g. penicillin G with the substitution of a $B(R^1)\cdots N(R^2)$ group for one $C(H)\cdots C(H)$ group, R^1 and R^2 being side-chain groups sufficiently bulky to protect the $B\cdots N$ bond from hydrolysis, see Fig. 5.3); and
(iii) there are substances, with assured biological activity but of uncertain mechanism, that can likely be modified with the insertion of boron in their structure without loss of biological activity (via the same substitution considered above for penicillin G).

In view of these observations, one can hypothesize a roadmap for subcellular sensing based on the exploitation of *existing* ICs. This roadmap can be imagined to develop along the following path:

(i′) validation and calibration of α-sensitive ICs with biological cells exposed to boron-containing pharmaceutical drugs with known docking sites to the cell;
(ii′) verification that the boron functionalization of the side-chain of pharmaceutical drugs with known docking sites does not modify them, via comparison of the measured α distribution from the expected one; and
(iii′) identification from the α autoradiography of the cellular loci where the boron-functionalized molecules with assured biological activity but unknown mechanism are accumulated. These loci are the candidate site for the localization of the true (non-modified) molecules.

This pathway has the advantage of exploiting existing circuits, leaving to bioinformatics the task of reconstructing the space distribution of biomolecules from the distribution of memory cells crossed by the α particles, and to chemistry the task of designing and synthesizing the molecules with the wanted behaviour. The boron route takes opportunistical advantage from the following occurrences:

- Storing the information in the original DRAM cells required 10^5 electrons, while today storing the information in flash memories requires about 50 electrons; this difference may be used to improve the sensitivity to α particles.
- A lot of potential boron carriers (such as boron-containing porphyrins, amino acids, carbohydrates, nucleic acid bases) have been synthesized and tested, especially having in mind cancer therapy based on neutron capture by boron [29].

Another approach is, however, possible, based on the development of ICs able to sense electrically the spatiotemporal and chemical patterns of the cell.

From ICs to Nanobiosensors

The driving force for the development of any technology is the existence of an adequate market niche. Subcellular sensing requires nanoscopic circuits with complexity exceedingly larger than that of microprocessors. A technology for subcellular sensing is thus motivated by a huge market only. Although the health market for the nanorobots considered in the Introduction certainly satisfies this need, the technology for subcellular sensing is nonetheless expected to have a higher chance of success if it can exploit the progress of the IC technology, taking advantage from the increasing complexity it allows in probing more and more parameters on shorter and shorter length scale.

The Incremental Increase of Complexity of ICs and Sensors

Consider the following classification:

Moore denotes the past and current IC technology.
More Moore indicates the near-future electronic devices and circuits hypothesized in the International Technology Roadmap for Semiconductors (ITRS), hereafter simply the 'Roadmap' [7]. Scaling electronic devices and circuits will continue into the next decade; it will be driven by Moore's law, according to the technologies described in the Roadmap, the main technological platform being based on the CMOS (complementary metal-oxide-silicon) structure.
More than Moore summarizes the extension of current CMOS processes to add new functionalities. Current CMOS processes can be used to develop new micro- and nanodevices, such as sensors, microelectromechanical systems, and light detectors. More-than-Moore technologies include heterogeneous integration, advanced packaging, three-dimensional integration, biomedical electronics, autonomous embedded appliances, and photovoltaics.

The above frames can be arranged in the diagram on the left of Fig. 5.4, where even in Moore and More Moore there is an increase of functions in addition to the

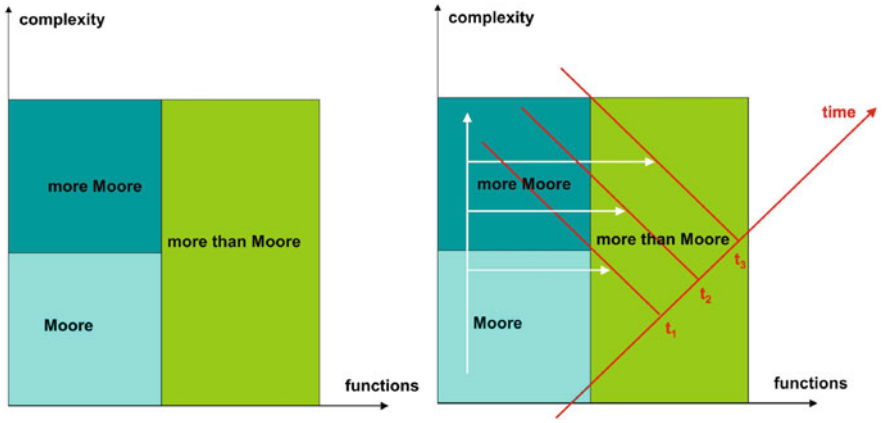


Fig. 5.4 Classification (*left*) and time evolution (*right*) of ICs according to complexity and functions

increase of density. The functions, however, are limited to electrical functions other than the digital ones: namely, the management of power (smart power circuits), the transmission of data (radio-frequency circuits), etc.

However, whereas the CMOS technology is characterized by the need of insulating the field effect transistors (FETs) and their interconnects from the environment (so that the FETs are separated from the environment by a number of layers with total thickness on the micrometre length scale), the sensor technology is characterized by the need of putting the sensitive element in contact with the environment (see Fig. 5.5). Because of these opposite needs, integrating logic with sensing is manifestly difficult.

The difficulties of integrating new functions into a certain technology is responsible for the delay in the development of high-density smart sensors. The time evolution of the technology is thus the one sketched on the right of Fig. 5.4.

The Shift of Paradigm

Sooner or later it will be impossible to overcome the limits of photolithography in an economically convenient way. Producing ICs with tera-scale integration (TSI) will only be possible employing methods going beyond Moore:

Beyond Moore sketches a future where CMOS technology will survive in hybrid combination with genuine nanotechnologies, possibly based on new devices (quantum dots, reprogrammable molecules, etc.), architectures (crossbar), or even principles (quantum computing).

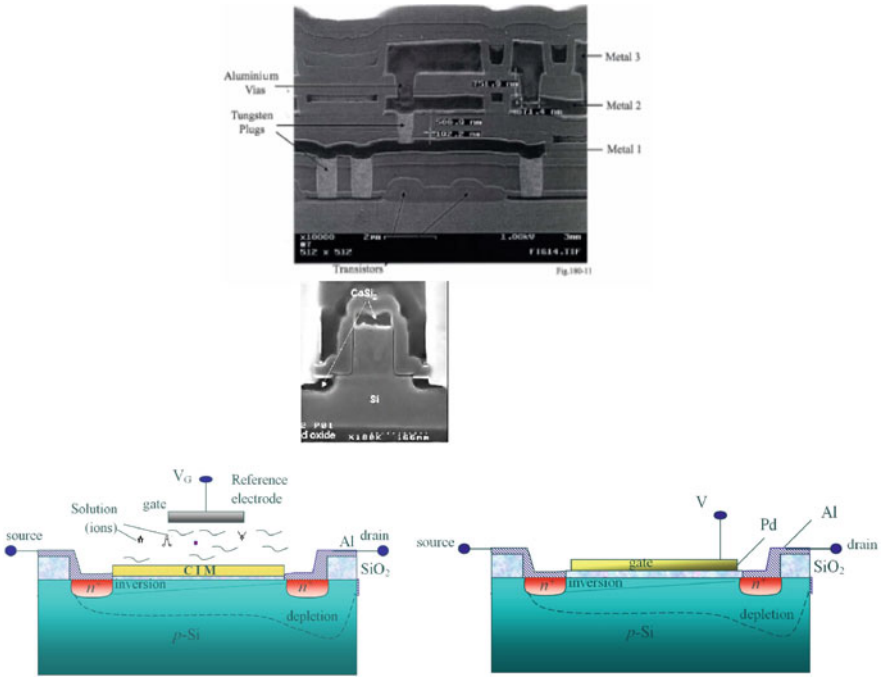


Fig. 5.5 Comparing the structure of a real IC (*top*) with those of ion-sensitive (*bottom and left*) and gas-sensitive (*bottom and right*) FETs of interest as sensors

Adding new functions to circuits with TSI complexity produced with non-lithographic methods will eventually bring the technology outside the walls—*extra muros*[3]:

> **Extra muros** hypothesizes a world where nanodevices with new functions are integrated in hybrid combinations with post-CMOS circuits for the full exploitations of their nanoscopic properties.

The positioning in the {functions, complexity} plane of circuits beyond Moore and *extra muros* is sketched in the left part of Fig. 5.6. The right part of this figure shows the expected time evolution.

Nanobiosensing (holding the circled zone in Fig. 5.6) is expected to be the final step of a gradual evolution eventually leading from current ICs to circuits of TSI

[3] *Extra muros* (read 'extra moorohs' and meaning, in cleric Latin, 'outside the walls') is used to denote the extreme development of the technology, however retaining its root ('Moore').

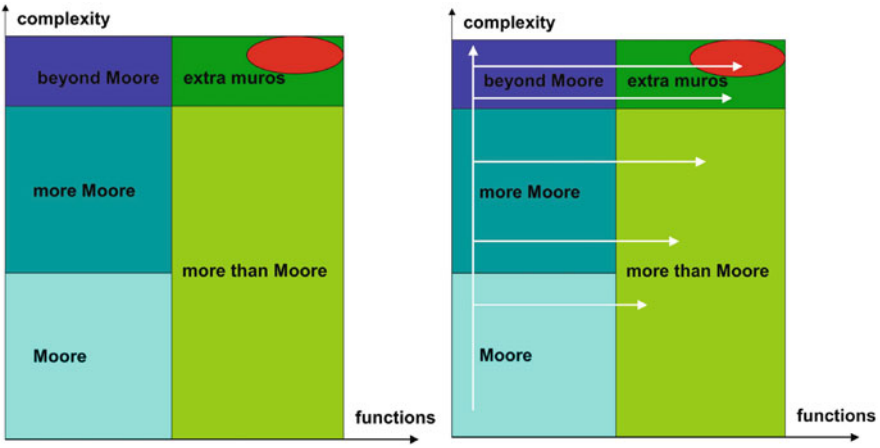


Fig. 5.6 Classification (*left*) and time evolution (*right*) of conventional (*lower boxes*) and new (*upper boxes*) devices according to complexity and functions (*the oval denotes* the region of nanobiosensors)

complexity with many functions and spatial resolution sufficient to sense living systems with deep subcellular resolution, embedded in silicon-based logic circuit.

As discussed in the next section, crossbars with cross-points of 20–30-nm pitch are already producible with current technologies. The crossbar is thus a candidate for subcellular biosensing provided that one is able to functionalize the cross-points and link them to the external world.

The Crossbar Structure

A crossbar is nothing but the superposition of an array of n parallel conductive wires on an array of m parallel wires; the arrays are oriented perpendicularly (within a non-critical accuracy) to one another. The $m \times n$ overlapping regions are referred to as cross-points and are usually filled with material with desired electrical properties. If the material displays suitable electrical properties (like hysteresis in electrical conductance) the crossbar may open a new paradigm for the design and production of electronic devices [30, 31].

Of course, the crossbar structure can be prepared using standard photolithographic methods in the frame of the planar technology. However, the extremely simple geometry of the crossbar structure—essentially two perpendicularly oriented arrays of parallel wires—has a dramatic advantage over the conventional ones: it allows a preparation with wire width on the nanoscale not only employing any advanced lithography but also with non-lithographic techniques (NLTs). Each NLT exploits the following features:

(V) it is possible to prepare highly homogeneous film and to control 'vertically' their thickness t down to the sub-nanometre length scale and

(V-to-H) it is possible to transform the 'vertical' thickness t into patterns with 'horizontal' width w:

$$t \xrightarrow{\text{NLT}} w.$$

These techniques are *superlattice nanowire pattern transfer* (SNAP) and two variants of the *multi-sidewall patterning technique* (S^nPT).

Superlattice Nanowire Pattern Transfer

In the SNAP properties (V) and (V-to-H) are exploited for the preparation of a contact mask for imprint lithography [32, 33]. The process is essentially based on the sequential alternate deposition of two films, A and B, characterized by the existence of a preferential etching for one (say A) of them. After cutting at 90°, polishing, and controlled etching of A, one eventually gets a mask formed by nanometre-sized trenches running parallel to one another at a distance fixed by the thickness of B [34, 35]. For instance, a contact mask for imprint lithography with pitch of 16 nm was prepared by growing on a substrate a quantum well via molecular beam epitaxy, cutting the sample perpendicularly to the surface, polishing the newly exposed surface, and etching selectively the different strata of the well [35]. The SNAP technique is reviewed in [36, 37].

A number of variants for transferring the pattern to the surface have been developed: molding, embossing, and stamping are the ones most frequently considered [32]. In one of them (molding), after filling the trenches with a suitable polymer, the mask is used as a stamp, pressing it onto the surface; if the polymer has a higher affinity for the surface than for the mask, the pattern is transferred to the surface when the mask is eventually removed [32]. The transfer of the polymer to the surface is possible without loss of geometry only if the trench is sufficiently shallow; this implies that the polymer must sustain a subsequent process where it is used as a mask for the definition (via directional etching) of the underlying structure with a high aspect ratio. Another method (embossing), sketched in the right-hand side of Fig. 5.7, involves the pressure-induced transfer of the pattern from the mask to a plastic film and its subsequent polymerization.

Sidewall Patterning Techniques

A totally different approach for the preparation of wire arrays with pitch on the 10-nm length scale is based on the multi-sidewall patterning technique (S^nPT). The S^nPT is essentially based on the repetition of the sidewall patterning technique (SPT), an age-old technology originally developed for the dielectric insulation of source-and-drain metal electrodes from the gate of metal-oxide-semiconductor (MOS) transistors.

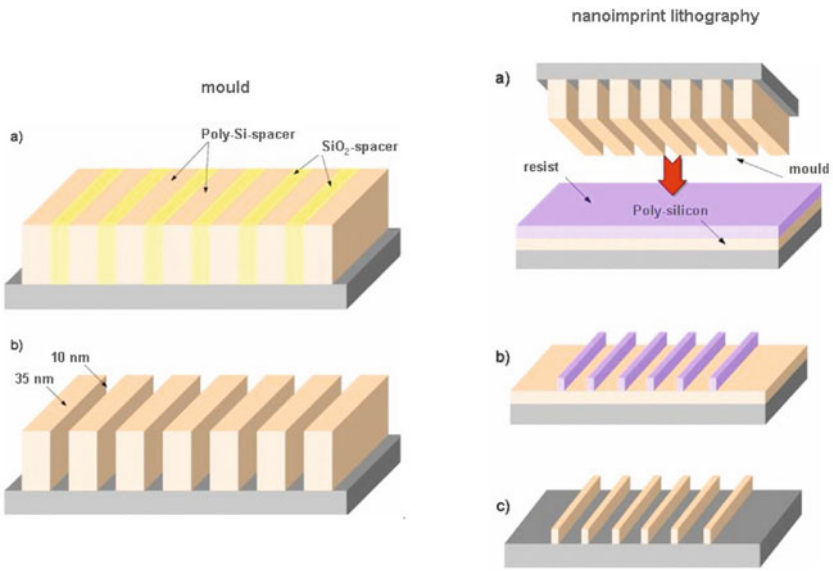


Fig. 5.7 Preparation of mold for imprint lithography (*left*) and its use as contact mask (*right*); the multilayer has been supposed to be produced with cycles of sequential depositions of silicon and SiO_2

The SPT involves the following steps:

SPT^0, the *lithographic definition* of a seed with sharp edge and high aspect ratio;
SPT^1, the *conformal deposition* on this feature of a film of uniform thickness; and
SPT^2, the *directional etching* of the film until the original seed surface is exposed.

If the process is stopped at this stage, it results in the formation of side walls aside the original seed; otherwise, if

SPT^3, the original seed is removed via a *selective etching*,

what remains at the end of this sequence is constituted only by the walls of the seed edges.

This technique has been demonstrated to be suitable for the preparation of features with minimum size of 7 nm [38, 39] and has already achieved a high level of maturity, succeeding in the definition, with yield very close to unity, of nanoscopic wires with high aspect ratio (see Fig. 5.8).

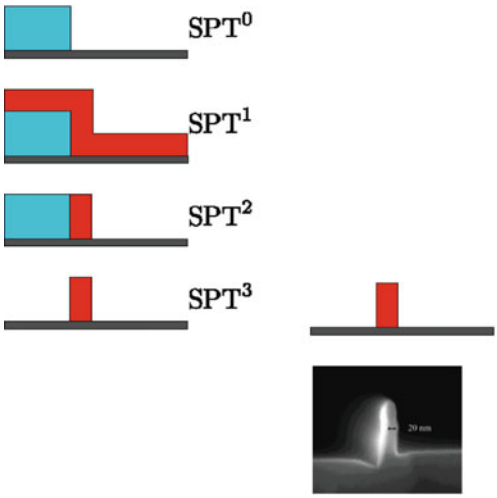


Fig. 5.8 (*Left*) The sidewall patterning technique: SPT^0, definition of a pattern with sharp edges; SPT^1, conformal deposition of a uniform film; SPT^2, directional etching of the deposited film up to the appearance of the original seed; and SPT^3, selective etching of the original feature. (*Right*) Cross section of a wire produced via SPT

Two S^nPT routes have been considered: the additive (S^nPT_+) and multiplicative ($S^nPT_\times$) routes. The S^nPT_+ is recent and was proposed having in mind the preparation of crossbars for molecular electronics [40–43]. The $S^nPT_\times$ is instead much older: The first demonstrators were developed for the generation of gratings with sub-lithographic period [44]; recently, however, this technique has been used for the preparation of wire arrays in biochips too [39].

Additive Route

The S^nPT_+ is substantially based on n SPT repetitions where *the original seed is not removed and each free wall of newly grown bars is used as a seed for the subsequent SPT*. Each SPT_+ cycle starts from an assigned seed and proceeds with the following steps:

$S^nPT^1_+$, conformal deposition of a conductive material;
$S^nPT^2_+$, directional etching of this material up to the exposure of the original seed;
$S^nPT^3_+$, conformal deposition of an insulating material; and
$S^nPT^4_+$, directional etching of this material up to the exposure of the original seed.

The basic idea of the S^nPT_+ is shown in Fig. 5.9: the upper part sketches the process; the lower part shows instead how poly-silicon arrays separated by SiO_2 dielectrics with sub-lithographic pitch (35 nm) can indeed be produced [40, 41]. Demonstrators with $n = 3$ [40, 41], $n = 4$ [45], and $n = 6$ [46] have been presented; the S^nPT_+ has recently been used to demonstrate the feasibility of a crossbar with cross-point density of 10^{10} cm^{-2} [46].

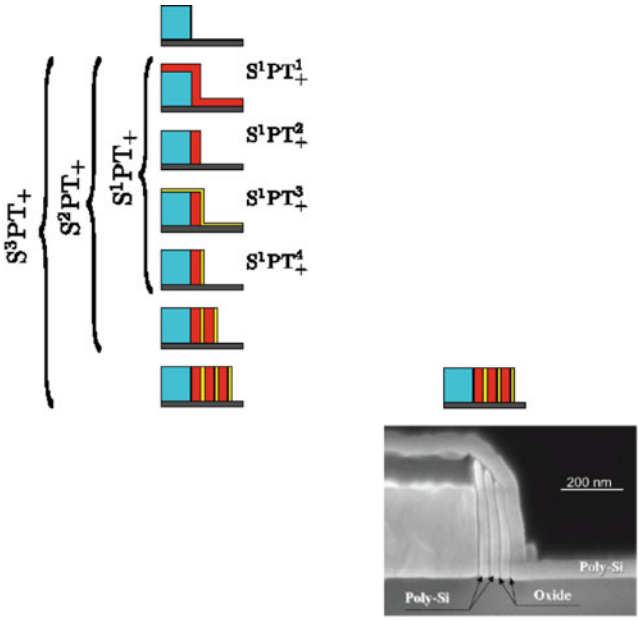


Fig. 5.9 (*Left*) The additive multi-sidewall patterning technique. (*Right*) an example of S^3PT_+ multi-sidewall (with pitch of 35 nm and formed by a double layer poly-Si|SiO_2) resulting after three repetitions of the SPT_+

The sketch in Fig. 5.9 shows a process in which lines are additively generated onto a progressively growing seed, preserving the original lithographic feature along the repetitions of the unit process. The unit process is based on two conformal depositions of uniform layers (poly-silicon and SiO_2) each followed by a directional etching. Variants of this process to reduce significantly the number of steps are discussed in [47].

Multiplicative Route

The SPT allows, starting from one seed, the preparation of *two* side-walls [38]; in principle, this fact allows another, multiplicative, growth technique—$S^nPT_\times$. The multiplicative generation requires that both sides of each newly grown side-wall are used as seeds for the subsequent growth—that is possible only if the original seed is etched away at the end of any cycle. In $S^nPT_\times$ each multiplicative $SPT_\times$ cycle involves therefore the following steps:

$S^nPT^1_\times$, conformal deposition of a film on the seed;
$S^nPT^2_\times$, directional etching of the newly deposited film up to the exposure of the seed; and
$S^nPT^3_\times$, selective etching of the original seed.

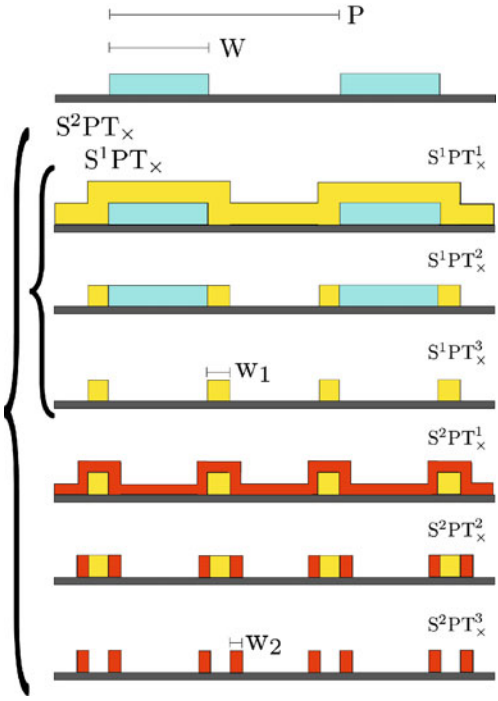


Fig. 5.10 Two $STP_{\times}$ steps for the formation of a sub-lithographic wire array starting from a lithographic seed array

Figure 5.10 sketches three $S^nPT_{\times}$ repetitions.

Addressing Sublithographic Features with Lithographic Tools

If the availability of nanofabrication techniques is fundamental in establishing a nanotechnology, not less vital is accessing the nanostructures with higher-level structures. This is especially difficult because the nanoworld is not directly accessible by means of standard lithographic methods—the difficulties in communication between the nanoworld and the macroworld are a bottleneck for the development of nanotechnology [48].

Several strategies have been adopted to attack the problem of addressing cross-points in a nanoscopic crossbar structure by means of externally accessible lithographic contacts. Many of them involve materials and methods quite far from, if not totally orthogonal to, those of the planar technology [49–54]. Moreover, a few of them are stochastic in nature and allow to access cross-points without knowing exactly which cross-point is addressed. Of course, these solutions, though potentially interesting for electronics, are useless for nanobiosensing. The strategy described in the following does not suffer from this limitation.

According to this strategy each line defining the crossbar extends beyond the crossing region and in this zone it is used for addressing. This region is then covered

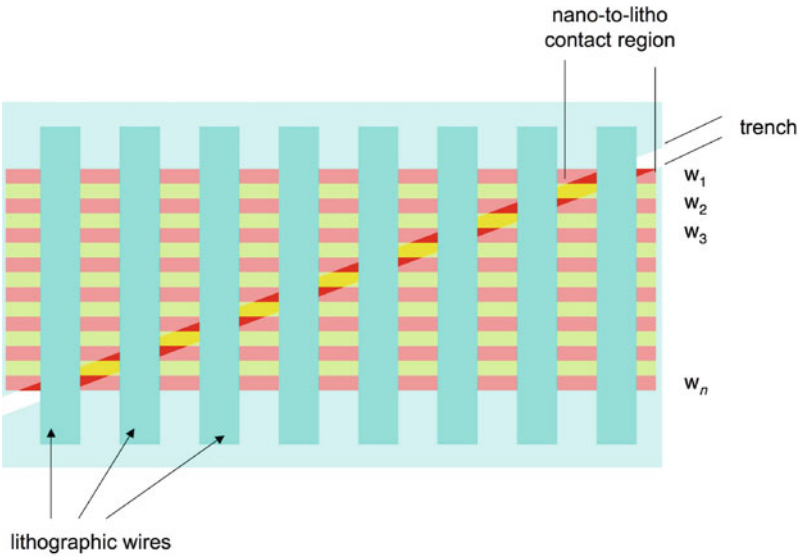


Fig. 5.11 The trick adopted in the horizontal bevelling technique for contacting each wire w_i ($i = 1, \ldots, n$) opening a trench slightly misoriented with respect to the wire array

with a cap protecting the cross-points, and the cap is etched away along a narrow (sub-lithographic) line misoriented with respect to the array by a small angle θ. In this way the zones where the wires are not covered are separated by a distance that diverges for $\theta \to 0$; thus, if θ is sufficiently small, the separation between the zones no longer protected makes them accessible to conventional lithography and suitable for contacting the external circuitry (Fig. 5.11). In this method each wire is linked to the external circuitry separately from the others—addressing n^2 cross-points requires therefore $2n$ contacts. Addressing a crossbar with 10^4 cross-points for sub-cellular sensing would require 200 contacts.

The Crossbar Route to Subcellular Sensing

Nanobiosensors for the subcellular analysis are possible, exploiting the potentials of the crossbar.

In line with the ideas sketched in [42], the crossbar is conveniently produced using heavily doped poly-silicon for the bottom and top wire arrays. After the definition of a bottom array of p-doped poly-silicon wires (Fig. 5.12a), a thin film of a material displaying highly non-linear conduction (SiO_2, with conduction controlled by the electric field at the poly-silicon surface, being a plausible candidate) is deposited (Fig. 5.12b), and another array of n-doped poly-silicon wires is defined on it (Fig. 5.12c). After that, the whole structure is covered by a relatively thick insulator (Fig. 5.12d) that is etched using a mask with the same pattern on the top electrode

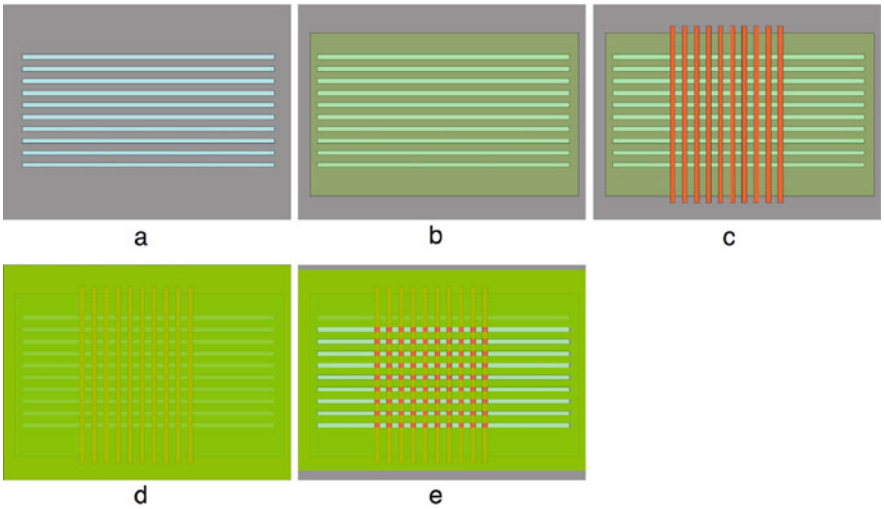


Fig. 5.12 The process for the preparation of a crossbar suitable for further functionalization: **a** definition of the bottom array of poly-silicon wires; **b** deposition of a material with strongly non-linear conductance; **c** definition of the top array of poly-silicon wires; **d** deposition of an insulating film; and **e** etching of the outer insulating layer with a mask for the exposure of a matrix of n^2 squares on the top wire array

but oriented perpendicularly. This will result in the exposure of n^2 of n-type doped poly-silicon zones with approximate square shape (Fig. 5.12e).

The immersion of the resulting structure in a solution of a salt of a reducible metal (like nickel and copper) will result in its electroless deposition in elemental form onto the exposed silicon. The deposited film may be thermodynamically or kinetically controlled to have an assigned thickness. A thermal treatment of the structure will result in the formation of one or more islands on the silicon with size controlled by surface and interfacial tension, as shown in the top of Fig. 5.13.

After that, the system is exposed to an ethylene or acetylene atmosphere at high temperature (say in the interval 700–1000 °C); in this environment the metal islands catalyse the formation of carbon nanotubes (CNTs) whose diameter is controlled by the diameter of the metal catalyst.[4] The process proceeds, as sketched in the bottom of Fig. 5.13, with the segregation of the metal island at the top of the growing CNT thus preserving the catalytic action.

The last stage, the functionalization of the nanotube, is notoriously difficult because of its poor reactivity. However, in addition to the chemistry targeted at direct

[4]With carbon nanotube it is denoted as an all-carbon hollow graphitic material with a high aspect ratio. Formally a CNT is obtained rolling up single or multiple graphene sheets to give a single-walled CNT (SWCNT) or a coaxial multiple-walled CNT (MWCNT), respectively. The CNT length is typically in the interval 10^2–10^4 nm and diameter 0.2–4 nm for SWCNTs or 2–100 nm for coaxial multiple-walled MWCNTs [55].

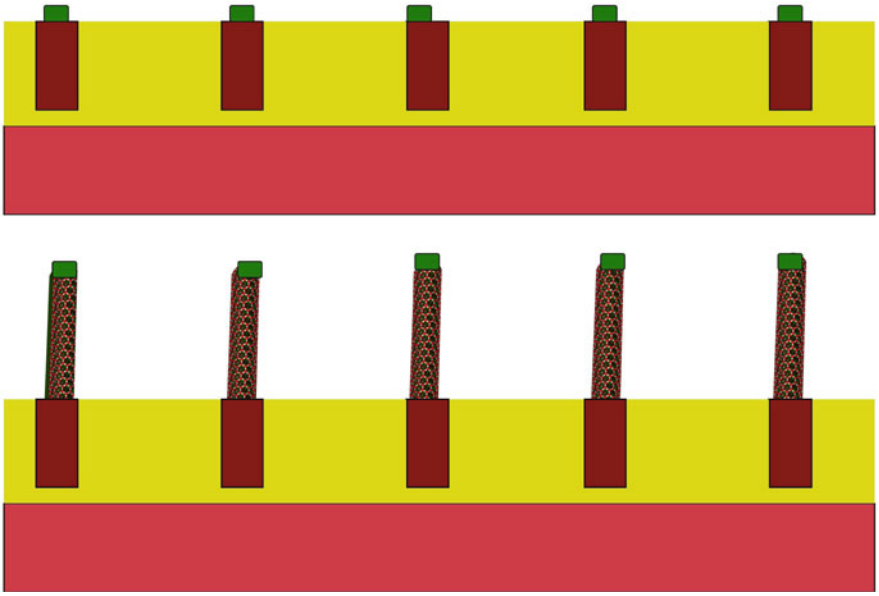


Fig. 5.13 Preparing the CNT matrix by the deposition of metal islands on the exposed n-type silicon (*top*) and the metal-catalysed growth of CNTs by exposure to a hydrocarbon atmosphere at high temperature (*bottom*)

functionalization, the metal cluster at the CNT top can be exploited as Trojan horse for that.

The choice of the functionalizing agent is a delicate matter. Two major approaches seem possible, both made possible by the fact that, due to its tiny capacitance, the conductance state of a cross-point is controlled by really few (perhaps only one) ions adsorbed on the nanotube.

The first approach is addressed to the detection of ionic species. For that the CNT (or its metal cup) is derivatized with a molecule with a very large specificity towards the considered ion. Several substances with this property are known for practically all cations of biological interest. Without pretending to be exhaustive, a few examples are the following:

Selective determination of alkaline ions The distribution of sodium and potassium between the cell and its medium controls the membrane potential and has been the matter of extended investigations. Crown ethers (see Fig. 5.14) are macrocyclic compounds displaying high selectivity towards alkaline ions [56]; for instance, 18-crown-6 is selective for K^+ whereas 15-crown-5 is selective for Na^+.

The derivatization of CNTs with one of them (e.g. 18-crown-6, see Fig. 5.15) will produce electrodes able to capture in highly selective way the corresponding alkali ion (K^+, in the considered case).

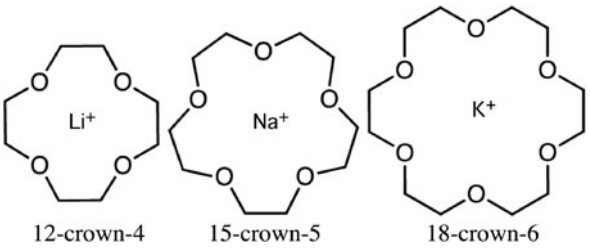


Fig. 5.14 A few common crown ethers and the ions to which they bind selectively

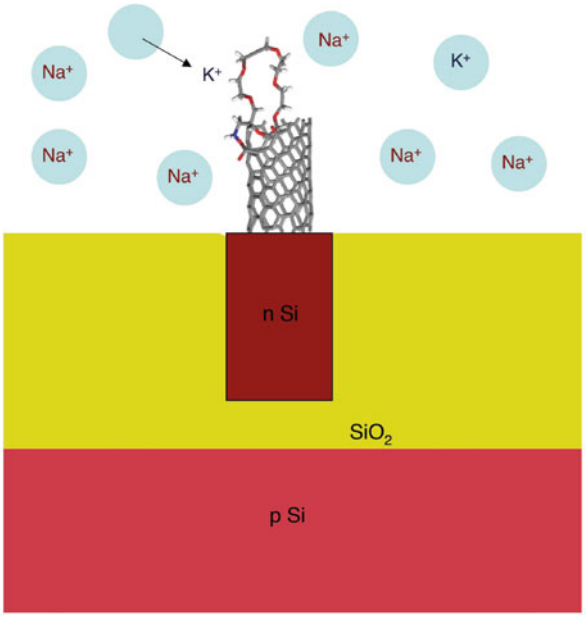


Fig. 5.15 How CNT functionalized with crown ethers may be used for the chemically selective and space-resolved detection of alkali ions

Determination of calcium The time variations of sodium and potassium are poor markers of cell functionality. In this respect, calcium is much more interesting and it has become the matter of growing interest. Moreover, its large variations with time suggest it as the first goal of subcellular sensing. Selective detectors of divalent cations, and especially Ca^{2+}, are known; among them the ethylenediaminetetraacetic acid, EDTA (a well-known ligand), and A23187 (an ion carrier with weak antibiotic activity produced at fermentation of *Streptomyces chartreusensis* [57], see Fig. 5.16) may be mentioned.

Somewhat more complex is the detection of anions. For that, a general route might involve the derivatization of the CNT with a cavitand—cyclodextrines [56] or cucurbiturils (see Fig. 5.17) are possible candidates. They should be so chosen

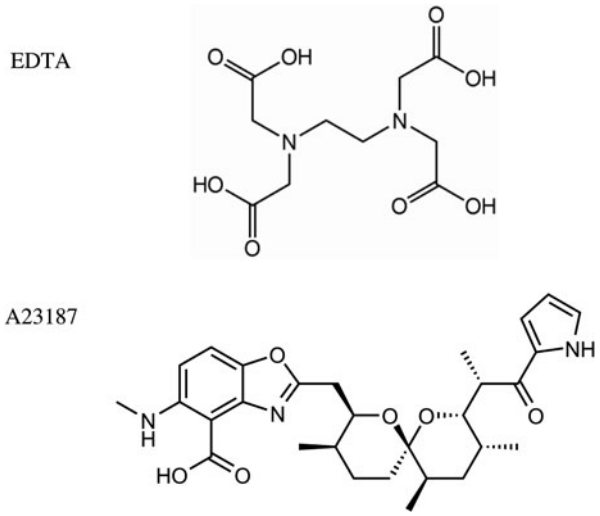


Fig. 5.16 Candidate molecules for CNT functionalization addressed for calcium detection

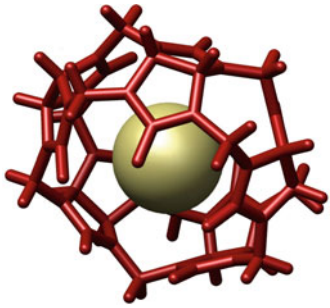


Fig. 5.17 A carcerand (cucurbit[5]uril) including a chlorine anion

as to allow an easy docking of the considered anion and modified (if necessary) to contain an electrophilic centre (or even a Lewis acid) allowing a strong bonding with the anion.

More difficult is the sensing of neutral species: for them, indeed, one must hypothesize that the CNT has been derivatized with suitable receptors reacting redox with the target species donating an electron to it (or vice versa). To allow the continuous operation of the electrode, however, the receptor must contain a sacrificial region that restores its pristine redox state with a longer time constant than the time required for the detection of the signal.

Observe now that, roughly speaking, any cell extends in the direction perpendicular to the crossbar by an amount of the same order as in the plane. The three-dimensional shape of the cell impacts seriously its structure and operation mode. In fact, assuming for simplicity a spherical shape and a specular distribution of species with respect to a plane passing through the centre of the sphere and

parallel to the crossbar, exploring the whole cell requires CNTs with length not smaller than the cellular radius. Moreover, to obtain the distribution of the various species both inside and outside the cell, the nanobiosensor must be arranged on the cantilever of a probe (like the ones characteristic of atomic force microscopy) able to control (= vary and measure) the distance from an assigned plane.

Of course, the penetration of any probe into the cell is expected to perturb appreciably the detection. This effect is, however, minimum for CNTs because no living system in the entire evolution has so far met such materials in amounts and duration sufficient to activate an immune response. It is stressed that this situation gives the human mankind an opportunity window (necessarily open for a limited time) for exploring the interior of cells with a minimum of perturbation.

Conclusions

The diagnostic use of nanorobots in preventative medicine is subordinated to a sufficiently accurate knowledge of the standard behaviour of healthy cells.

The construction of such a model depends strictly on the ability to sense cells at subcellular level. The preparation of devices able to do that goal is a major technological problem, whose solution is possible only in the presence of an adequate driving force. The huge potential market of nanorobots (in the long run expected to be of the order of 10^{10} persons times 10^6 nanorobots per person-year, at least) is sufficient for that, so that subcellular sensing is expected to become a viable research area as far as they are indeed physically possible.

This chapter has shown not only its possibility, but also the existence of ways to circumvent a few expected bottlenecks in the path towards nanobiosensing, identifying two routes in that direction.

The first route (the 'boron route') is based on the opportunistic exploitation of the sensitivity to α particles of electron devices and their insensitivity to thermal neutrons. Since α particles are generated (with large cross section) in the interaction of thermal neutrons with ^{10}B, information on boron distribution inside the cell can be obtained from the map of the α-sensitive events. Another opportunistic advantage of this route comes from the already existing large family of boron-containing substances developed for thermal-neutron cancer therapy.

The second route (the 'crossbar route') runs along the same path that will allow electronics to go *extra muros*. The starting point of this path is the current possibility to prepare matrices of electrically conductive and singularly addressable pixels with areal density close to 10^{11} cm^{-2}. If such pixels were suitably functionalized with receptors specifically sensitive towards metabolites, such a density would allow their chemical distributions to be mapped with spatial resolution of the order of 10 nm. The preparation of such matrices, although possible in principle, requires nonetheless a long and difficult path involving both technology and chemistry. In fact, overcoming the limits of current photolithography (allowing the definition of arrays with pitch of about 100 nm) requires tricks for the transformation of the thickness (usually controlled on the sub-nanometre length scale) into a width, as discussed for

the SNAP and S^{n}PT. Moreover, when a pixel matrix has been prepared, there is the need of putting the active elements (the cross-points) in electrical contact with the cell. This step is not trivial; the selective growth of CNTs has been identified as a key tool for that. At last, the preparation is completed with the functionalization of the CNTs, addressed to allow them to feel the assigned species.

Of course, although most of the barriers along the path to nanobiosensing do not appear to be insurmountable, it is nonetheless clear that all the proposed ways of attack to them are essentially working hypotheses instead of consolidated technologies.

References

1. Freitas, R.A., Jr.: Current status of nanomedicine and medical nanorobotics. J. Comput. Theor. Nanosci. **2**, 1–25 (2005)
2. Requicha, A.A.G.: Nanorobots, NEMS, and nanoassembly. Proc. IEEE **91**, 1922–1933 (2003)
3. Brigger, I., Duberne, C., Couvreur, P.: Nanoparticles in cancer therapy and diagnosis. Adv. Drug Deliv. Rev. **54**, 631–651(2002)
4. Brannon–Peppas, L., Blanchette, J.O.: Nanoparticle and targeted systems for cancer therapy. Adv. Drug Del. Rev. **56**, 1649–1659 (2004)
5. Jurgons, R., Seliger, C., Hilpert, A., Trahms, L., Odenbach, S., Alexiou, C.: Drug loaded magnetic nanoparticles for cancer therapy. J. Phys. Condens. Mat. **18**, S2893–S2902 (2006)
6. Praetorius, N.P., Mandal, T.K.: Engineered nanoparticles in cancer therapy. Recent Patents Drug Delivery Formul. **1**, 37–51 (2007)
7. International Technology Roadmap for Semiconductors, ITRS 2007 Edition
8. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P.: Molecular Biology of the Cell, Ch. 17. 5th edn. Garland Science, New York (2007)
9. Tanaka, M.: Polymer-supported membranes: artificial models of cell surfaces on polymer thin films. MRS Bull. **31**, 513–520 (2006)
10. Montemagno, C., Bachand, G.D.: Constructing nanomechanical devices powered by biomolecular motors. Nanotechnology **10**, 225–231 (1999)
11. Kim, M.J., Breuer, K.S.: Microfluidic pump powered by self-organizing bacteria. Small **4**, 111–118 (2007)
12. Behkam, B., Sitti, M.: Bacterial flagella-based propulsion and on/off motion control of microscale objects. Appl. Phys. Lett. **90**, 023902 (2007)
13. Zhang, L., Abbott, J.J., Dong, L., Kratochvil, B.E., Bell, D., Nelson, B.J.: Artificial bacterial flagella: fabrication and magnetic control. Appl. Phys. Lett. **94**, 064107 (2009)
14. Benenson, Y., Paz-Elizur, T., Adar, R., Keinan, E., Livneh, Z., Shapiro, E.: Programmable and autonomous computing machine made of biomolecules. Nature **414**, 43–434 (2001)
15. Benenson, Y., Gil, B., Ben-Dor, U., Adar, R., Shapiro, E.: An autonomous molecular computer for logical control of gene expression. Nature **429**, 423–429 (2004)
16. Reif, J.H., Sahu, S.: Autonomous programmable DNA nanorobotic devices using DNAzymes. Theor. Comput. Sci. **410**, 1428–1439 (2009)
17. Crichton, M., Prey. Harper Collins, New York, NY (2002)
18. Hood, L., Heath, J.R., Phelps, M.E., Lin, B.: Systems biology and new technologies enable predictive and preventative medicine. Science **306**, 640–643 (2004)
19. Stern, E., Klemic, J.F., Routenberg, D.A., Wyrembak, P.N., Turner–Evans, D.B., Hamilton, A.D., LaVan, D.A., Fahmy, T.M., Reed, M.A.: Label-free immunodetection with CMOS-compatible semiconducting nanowires. Nature **445**, 519–522 (2007)
20. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P.: Molecular Biology of the Cell, Ch. 17, 5th edn. Garland Science, New York (2007)

21. Cerofolini, G.F., Cerofolini, M.: Heterogeneity, allostericity and hysteresis in adsorption of water by proteins. J. Colloid Interface Sci. **78**, 65–73 (1980)
22. Frauenfelder, H., Wolynes, P.G., Austin, R.H.: Biological physics. Rev. Mod. Phys. **71**, S419–S430 (1999)
23. Turing, A.M.: A chemical basis for biological morphogenesis. Phil. Trans. Roy. Soc. B (London) **237**, 37–72 (1952)
24. Glansdorff, P., Prigogine, I.: Thermodynamic Theory of Structure, Stability and Fluctuations. Wiley-Interscience, London (1971)
25. Rose, R.: Life Itself. Columbia University Press, New York (1991)
26. A. Cornish-Bowden, Cárdenas, M.L., Letelier, J.-C., Soto-Andrade, J.: Beyond reductionism: metabolic circularity as a guiding vision of real biology of systems. Proteomics **7**, 839–845 (2007)
27. Grattarola, M., Cambiaso, A., Cenderelli, S., Parodi, G., Tedesco, M., Die, B., Cerofolini, G., Meda, L., Solmi, S.: ISFET-like devices coupled to neuroblastoma cells: cytometric and electrical characterization., In: Hong, F.T. (ed.) Molecular Electronics: Biosensors and Biocomputers, pp. 297–304. Plenum Press, New York, NY (1989)
28. Cerofolini, G.F., Ferla, G., Foglio, A. Para: Non-standard applications of silicon devices as nuclear radiation detectors (in Italian). Giornale di Fisica **23**, 201–211 (1982)
29. Leśnikowski, Z., Paradowska, E., Olejniczak, A.B., Studzińska, M., Seekamp, P., Schüßler, U., Gabel, D., Schinazi, R.F., J. Plešěk: Towards new boron carriers for boron neutron capture therapy: metallacarboranes and their nucleoside conjugates. Bioorg. Med. Chem. **13**, 4168–4175 (2005)
30. Heath, J.R., Kuekes, P.J., Snider, G.S., Williams, R.S.: A defect-tolerant computer architecture: opportunities for nanotechnology. Science, **280**, 1716–1721 (1998)
31. Cerofolini, G.F., Romano, E.: Molecular electronics in silico. Appl. Phys. A **91**,181–210 (2008)
32. Gates, D.B., Xu, Q.B., Stewart, M., Ryan, D., Willson, C.G., Whitesides, G.M.: New approaches to nanofabrication. Chem. Rev. **105**, 1171–1196 (2005)
33. Whitesides, G.M., Love, J.C.: The art of building small. Sci. Am. Repts. **17**, 12–21 (2007)
34. Natelson, D., Willett, R.L., West, K.W., Pfeiffer, L.N.: Fabrication of extremely narrow metal wires. Appl. Phys. Lett. **77**, 1991–1993 (2000)
35. Melosh, N.A., Boukai, A., Diana, F., Gerardot, B., Badolato, A., Heath, J.R.: Ultrahigh-density nanowire lattices and circuits. Science **300**, 112–115 (2003)
36. Wang, D., Sheriff, B.A., McAlpine, M., Heath, J.R.: Development of ultra-high density silicon nanowire arrays for electronics applications. Nano Res. **1**, 9–21 (2008)
37. Heath, J.R.: Superlattice nanowire pattern transfer. Acc. Chem. Res. **41**, 1609–1617 (2008)
38. Choi, Y.-K., Zhu, J., Grunes, J., Bokor, J., Somorjai, G.A.: Fabrication of sub-10-nm silicon nanowire arrays by size reduction lithography. J. Phys. Chem. B **107**, 3340–3343 (2003)
39. Choi, Y.-K., Lee, J.S., Zhu, J., Somorjai, G.A., Lee, L.P., Bokor, J.: Sublithographic nanofabrication technology for nanocatalysts and DNA chips. J. Vac. Sci. Technol. B **21**, 2951–2955 (2003)
40. Cerofolini, G.F., Arena, G., Camalleri, M., Galati, C., Reina, S., Renna, L., Mascolo, D., Nosik, V.: Strategies for nanoelectronics. Microelectr. Eng. **81**, 405–419 (2005)
41. Cerofolini, G.F., Arena, G., Camalleri, M., Galati, C., Reina, S., Renna, L., Mascolo, D.: A hybrid approach to nanoelectronics. Nanotechnology **16**, 1040–1047 (2005)
42. Cerofolini, G.F.: Realistic limits to computation. II. The technological side. Appl. Phys. A **86**, 31–42 (2007)
43. Cerofolini, G.F., Amato, P., Romano, E.: The multi-spacer patterning technique: a non-lithographic technique for terascale integration. Semicond. Sci. Technol. **23**, 075020 (2008)
44. Flanders, D.C., Efremow, N.N.: Generation of <50 nm period gratings using edge defined techniques. J. Vac. Sci. Technol. B **1**, 1105–1108 (1983)

45. Cerofolini, G.F., Casuscelli, V., Cimmino, A., Di Matteo, A., Di Palma, V., Mascolo, D., Romanelli, E., Volpe, M.V., Romano, E.: Steps farther towards micro-nanomole integration via the multispacer patterning technique. Semicond. Sci. Technol. **22**, 1053–1060 (2007)
46. Ben Jamaa, M.H., Cerofolini, G., De Micheli, G., Leblebici, Y.: Complete nanowire crossbar framework optimized for the multi-spacer patterning technique, International Conference on Compilers, Architecture and Synthesis for Embedded Systems—CASES 2009
47. Cerofolini, G.: Nanoscale Devices. Springer, Berlin (2009)
48. Roukes, M.: Plenty of room indeed. Sci. Am. Repts. **17**(3), 4–11 (2007)
49. Huang, Y., Duan, X., Cui, Y., Lauhon, L.J., Kim, K.-H., Lieber, C.M.: Logic gates and computation from assembled nanowire building blocks. Science **294**, 1313–1317 (2001)
50. Zhong, Z., Wang, D., Cui, Y., Bockrath, M.W., Lieber, C.M.: Nanowire crossbar arrays as address decoders for integrated nanosystems. Science **302**, 1377–1379 (2003)
51. DeHon, A., Lincoln, P., Savage, J.E.: Stochastic assembly of sublithographic nanoscale interfaces. IEEE Trans. Nanotechnol. **2**, 165–174 (2003)
52. Likharev, K.K., Strukov, D.B.: CMOL: Devices, circuits, and architectures. In: Cuniberti, G., Fagas, G., Richter, K. (eds.) Introducing Molecular Electronics, pp. 447–477. Springer, Berlin (2005)
53. Strukov, D.B., Likharev, K.K.: Prospects for terabit-scale nanoelectronic memories, Nanotechnology **16**, 137–148 (2005)
54. Beckman, R., Johnston-Halperin, E., Luo, Y., Green, J.E., Heath, J.R.: Bridging dimensions: demultiplexing ultrahigh-density nanowire circuits. Science **310**, 465–468 (2005)
55. Saito, R., Dresselhaus, G., Dresselhaus, M.S.: Physical Properties of Carbon Nanotubes Imperial College Press, London (1998)
56. Lehn, J.M., Ball, P.: Supramolecular chemistry. In: Hall N. (ed.) The New Chemistry, Ch. 12. Cambridge University Press, Cambridge (2000)
57. Abbott, B.J., Fukuda, D.S., Dorman, D.E., Occolowitz, J.L., Debono, M., Farhner, L.: Microbial transformation of A23187, a divalent cation ionophore antibiotic. Antimicrob. Agents Chemother. **16**, 808–812 (1979)

Chapter 6
Photothermal Sensing of Chemical Vapors Using Microcantilevers

Thomas Thundat, Charles W. Van Neste, Larry R. Senesac, and Adam R. Krause

Abstract Although microfabricated cantilevers have been used for detecting a variety of chemicals with high sensitivity, their selectivity appears to be poor. This selectivity problem is directly related to the poor selectivity of chemical interfaces immobilized on cantilever surfaces. Here we discuss two cantilever-based techniques that can be used for obtaining orthogonal signals for multimodal operation for enhanced selectivity. The first technique is based on photothermal deflection spectroscopy where selectivity is achieved through a mechanical response due to optical absorption by the adsorbed molecules on a cantilever. In the second technique, the position of the resonance frequency peak of the cantilever is monitored for shifting due to mass loading. This technique allows the precise measurement of the mass of the surface-adsorbed molecules. These two methods are demonstrated for adsorbed explosives.

Introduction

The detection of extremely small concentrations of chemical vapors is very important in many applications. Nanomechanical sensors based on the cantilever platform have been shown to have very high sensitivity in detecting adsorbed molecules [1]. A microfabricated cantilever sensor undergoes bending when molecules are preferentially adsorbed on one of its surfaces [2–8]. The adsorption-induced bending is directly related to changes in surface energy due to molecular adsorption on one of its surfaces. Simultaneously the resonance frequency of the cantilever varies sensitively due to mass loading. Therefore the bending of the cantilever is related to the adsorption energy while the resonance frequency variation is directly related to adsorbed mass. Both of these signals can be detected simultaneously for cantilevers that are around 100–400 μm in length with a spring constant of around 0.1–1 N m^{-1}.

T. Thundat (✉)
Nanoscale Science and Devices Group, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831, USA
e-mail: thundattg@ornl.gov

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_6,

Cantilever Sensors and Challenges

Many chemical vapor sensors with high sensitivity have been demonstrated using the cantilever sensor platform. The cantilever sensors have very high sensitivity [9]. Other advantages include miniature size, low power consumption, array-based detection, and potential low cost. Chemical selectivity with cantilever sensors is often accomplished by using chemically selective interfaces immobilized on their surfaces. The selectivity of detection is, therefore, directly related to the selectivity of the coatings. However, chemically selective coatings capable of reversible interaction with the analyte molecules do not have very high selectivity. Most reversible coatings are based on weak chemical interactions such as hydrogen bonding which are very generic. The response from an array of cantilevers coated with different chemical coatings and analyzed with pattern recognition can provide limited selectivity. However, when mixtures of chemical vapors are present, the array response fails to identify the individual mixture components in the targeted vapor due to the lack of orthogonal response.

Therefore, despite its many advantages, the cantilever platform remains unattractive for small molecule detection due to lack of selectivity. However, selectivity can be achieved for small molecules by incorporating a photothermal concept that can provide selectivity based on the well-established principles of vibrational spectroscopy. Microcantilever-based photothermal spectroscopy was first demonstrated by Gimzewski et al. [10]. Since then it has been used for the detection of a range of chemicals including DNA, explosives, and chemical warfare agents [11–17]. Cantilever-based photothermal deflection spectroscopy (PDS) utilizes the extremely high thermal sensitivity of a bimaterial cantilever. A bimaterial cantilever with adsorbed molecules shows bending when exposed to infrared (IR) radiation. The adsorbed molecules absorb the IR energy and heat the bimaterial microcantilever. The change in temperature causes the cantilever to bend due to differential thermal expansion.

Photothermal Deflection Spectroscopy

In this chapter we will discuss the cantilever-based photothermal technique in the context of detecting explosive vapors. Since explosives have very low vapor pressures, extremely high sensitivity is essential for detecting very small concentrations of explosive [9]. Using PDS we demonstrate highly sensitive detection of sub-nanogram quantities of adsorbed explosives such as pentaerythritol tetranitrate (PETN), cyclotrimethylene trinitramine (RDX), and trinitrotoluene (TNT). The sensitivity of the technology presented here can be improved further by optimizing the bimaterial effect. This method has the potential for detection of sub-monolayer amounts of molecules with high selectivity and sensitivity.

The bimaterial effect on cantilevers has been well investigated. A bimaterial cantilever undergoes bending due to differential expansion of constituent layers of the cantilevers. Extremely small changes in temperature can be detected sensitively

by using mechanical deflection of a bimaterial cantilever. The deflection of the microcantilever beam due to this bimaterial effect can be derived using beam equations:

$$z = -\frac{3}{4}(\alpha_1 - \alpha_2)\frac{t_1 + t_2}{t_2^2 K}\frac{t^3}{(\lambda_1 t_1 + \lambda_2 t_2)w}P, \tag{6.1}$$

where z is the deflection of the cantilever tip, α_1 and α_2 are the coefficients of thermal expansion for the two layers, l is the length of the cantilever, t_1 and t_2 are the layer thicknesses, λ_1 and λ_2 are the thermal conductivities of the two layers, w is the width of the cantilever, and P is the total power absorbed by the cantilever. The subscripts 1 and 2 represent the gold and silicon substrates, respectively. The parameter K is given by the expression

$$K = 4 + 6\left(\frac{t_1}{t_2}\right) + 4\left(\frac{t_1}{t_2}\right)^2 + \frac{E_1}{E_2}\left(\frac{t_1}{t_2}\right)^3 + \frac{E_2}{E_1}\left(\frac{t_2}{t_1}\right). \tag{6.2}$$

In the photothermal deflection spectroscopy (PDS), the microcantilever with adsorbed chemical species is sequentially exposed to different infrared radiation wavelengths from a monochromator. Generally, the IR light is chopped at a frequency chosen to improve the sensitivity of detection. The microcantilever deflection as a function of IR wavelength resembles the IR absorption spectrum of the adsorbate molecules. Microcantilever-based photothermal spectroscopy, therefore, combines the extremely high sensitivity of a cantilever beam and the selectivity of molecular vibration spectroscopy. The sensitivity of detection is limited by the thermal sensitivity of the cantilever, which can be optimized by properly designing the bimaterial aspect of the cantilever.

Adsorbed Mass Detection

The technique can be further quantified by monitoring the resonance frequency of the cantilever for measuring the adsorbed mass. By monitoring the resonance frequency shift before and after explosive vapor adsorption, it is possible to calculate the amount of chemical adsorbed on the cantilever. The mass of the adsorbates, σm, on the cantilever can be calculated from

$$\delta m = m\left(\frac{\nu_1^2 - \nu_2^2}{\nu_2^2}\right). \tag{6.3}$$

Here ν_1 and ν_2 are the initial and final resonant frequencies of the microcantilever during mass adsorption and m is the cantilever mass. One advantage of the microcantilever technique is that both the bending and resonance frequencies can be measured simultaneously.

Experimental Techniques

The experiments described here were carried out using commercially available silicon microcantilevers with dimensions of 350 μm length, 35 μm width, and 1 μm thickness (MikroMasch, Oregon). The silicon microcantilevers were made bimaterial by depositing a 600 nm layer of gold with 10 nm of chromium as an adhesion layer using an e-beam evaporator. We have used an optical beam deflection method for monitoring the bending of the cantilever. In the optical beam deflection method, a focused beam from a visible laser diode is reflected off the cantilever into a position-sensitive detector (PSD). The output voltage from the PSD is directly proportional to the cantilever deflection. Cantilever displacement can be monitored with sub-nanometer resolution using this method.

Explosive vapor was deposited on the cantilever using a custom-built explosive vapor generator (Idaho National Laboratory, INL) that can deliver precise concentrations of explosive vapors. In brief, the explosive vapors were produced by the vapor generator, by flowing nitrogen through a chamber containing thickly packed glass wool with adsorbed explosives kept a constant temperature. The glass wool sample was prepared by injecting explosives dissolved in an acetonitrile solution into the thickly packed glass wool. Initially dry nitrogen was passed through the glass wool chamber kept at higher temperature to remove the solvents. Separate vapor generators were used for generating PETN, RDX, and TNT vapor streams to avoid cross-contamination. The glass wool chamber temperature was kept constant using two thermoelectric elements. The vapor pressure of the explosive inside the chamber is related to the temperature of the chamber. The TNT vapor was generated at 60°C, while the RDX and PETN vapors were generated by heating the chamber to 75°C. To avoid condensation in the delivery lines, the delivery lines were also heated to the same temperature as the chamber. The nitrogen flow was kept at 200 sccm for all of the experiments.

We have used the IR source from a Foxboro Miran 1A-CVF spectrometer as the source for monochromatic IR light. In this arrangement an interference filter wheel selects the IR wavelengths in sequential fashion. The photothermal deflection spectrum of the adsorbate on the cantilever were obtained by exposing the cantilever beam to monochromatic infrared radiation from an IR source and the interference filter wheel in the wavelength window of 2.5–14.5 μm. The experimental setup is shown in Fig. 6.1. The IR energy from the source was focused onto the cantilever using a concave mirror. The bending of the cantilever beam was monitored using an optical beam deflection arrangement and a PSD. The PSD signal was fed to an in-house fabricated amplifier circuit for amplification.

To increase the sensitivity, a lock-in amplifier was used to obtain the cantilever signal generated by the pulsed infrared light. The resonance frequency of the cantilever was determined using a spectrum analyzer. The experiments were carried out with cantilevers of known resonance frequency and mass after depositing explosive vapors on them using the vapor generator. In general due to extremely low vapor pressures of the explosives, they condense on the cantilever. The analytes such as volatile organic compounds have very high vapor pressure and do not condense

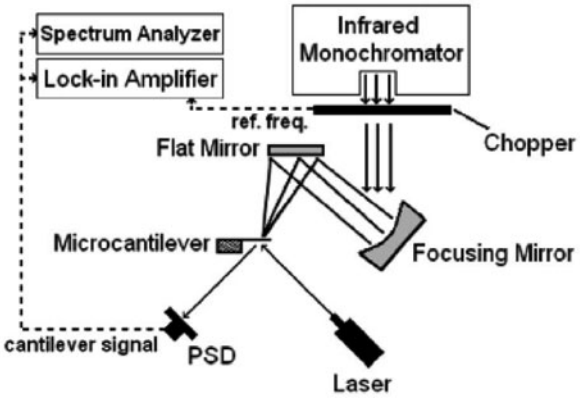


Fig. 6.1 Schematic of the experimental setup used for photothermal deflection spectroscopy (PDS). The infrared light from the monochromator is focused on the cantilever and chopped at 80 Hz. The wavelength is scanned between 2.5 and 14.5 μm. The resulting deflection data is collected with the lock-in amplifier. The spectrum analyzer is used to measure the resonance frequency to calculate the mass of explosives adsorbed on the cantilever

on the cantilever. The mass adsorption results in resonance frequency shift enabling determination of adsorbed mass. The cantilevers with adsorbed explosives were illuminated with a focused IR beam from the monochromator. The wavelength of the IR light was then scanned between 2.5 and 14.5 μm wavelengths while recording the lock-in amplifier signal due to cantilever bending.

Photothermal Deflection Spectroscopy of Adsorbed Explosives

Figure 6.2 shows the PDS spectra of PETN, RDX, and TNT on a silicon cantilever. The spectra shown in Fig. 6.2 were obtained by dividing the IR profile of each explosive by the baseline of the bare silicon cantilever taken prior to explosive adsorption. The resulting spectra were then normalized for qualitative comparison. The observed peaks are in excellent agreement with published IR absorption spectra of the bulk explosives [18].

The intensity of the photothermal signal depends on the bimaterial cantilever thermal sensitivity as well as the amount of adsorbed mass. The heat is generated by the non-radiative de-excitation of the adsorbed molecules. The thermal sensitivity of the cantilever depends on the properties of the constituent elements of the bimaterials such as the heat capacity and thermal conductivity of the cantilever beam and bimaterial layer. Also, the larger the difference in the coefficient of thermal expansion between the cantilever material and the bimaterial layer material, the greater the bending of the device for a given temperature change. The observed photothermal deflection spectra match very well with the known vibrational modes of the explosives.

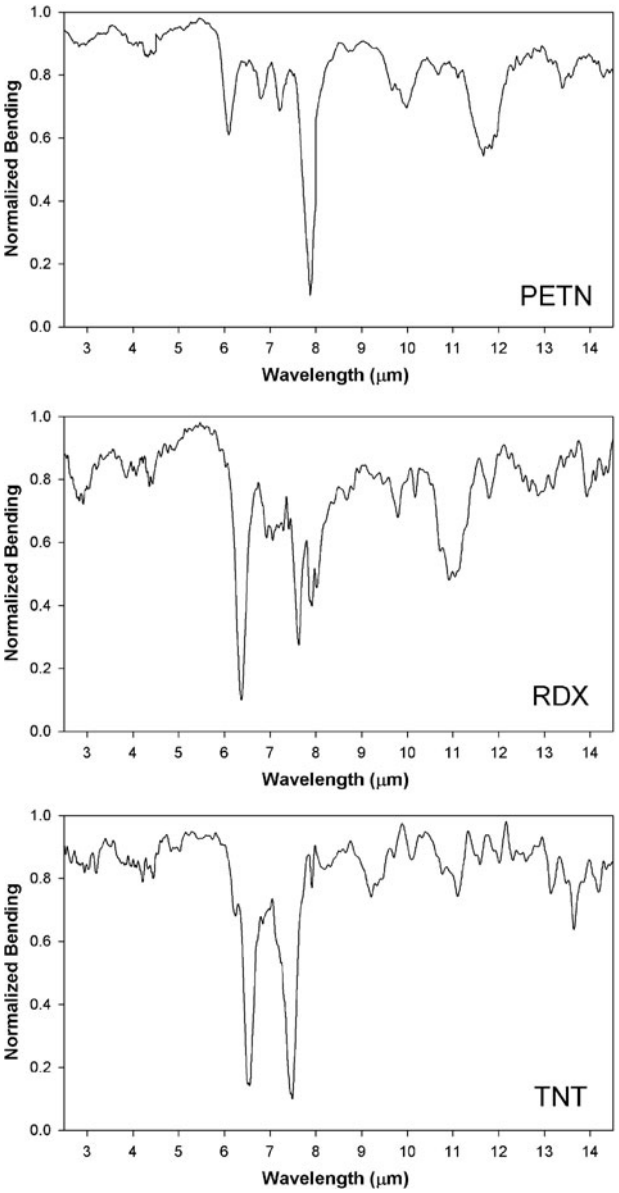


Fig. 6.2 Photothermal deflection spectra of vapor phase explosives adsorbed on the cantilever. In the PETN spectrum (*top plot*) the peak near 6 μm is caused by the asymmetric stretching of the O–NO_2 bonds, while the peak near 8 μm is from the symmetric stretching of the same bonds. Similarly, in the RDX spectrum (*center plot*) the peaks seen at 6.4 and 7.6 μm are caused by the asymmetric and symmetric stretching of the N–NO_2 bonds respectively, and in the TNT spectrum (*bottom plot*) the asymmetric and symmetric stretchings of the C–NO_2 bonds cause the peaks at 6.5 and 7.4 μm, respectively

The observed cantilever deflection peaks can be explained using the conventional peaks observed with infrared absorption spectroscopy. The peak at around 7.4–8 μm, found in all three explosive spectra, is caused by the symmetric stretching vibration of the NO_2 (nitro) group bond [19–22]. The peak near 6–6.6 μm is caused by the asymmetric stretching vibration of the same bond [19–22]. The slight shift in these two peaks is due to the nitro group being bound to a different atom in each explosive compound. In PETN the nitro groups are bound to oxygen (O–NO_2). This bond has a symmetric stretch vibration at wavenumber 1285 cm^{-1} (7.78 μm) and an asymmetric stretch vibration at wavenumber 1658 cm^{-1} (6.03 μm) [19–21]. In RDX the nitro groups are bound to nitrogen (N–NO_2). This bond has a symmetric stretch vibration at wavenumber 1310 cm^{-1} (7.63 μm) and an asymmetric stretch vibration at wavenumber 1570 cm^{-1} (6.37 μm) [19]. In TNT they are bound to carbon (C–NO_2). This bond has a symmetric stretch vibration at wavenumber 1350 cm^{-1} (7.41 μm) and an asymmetric stretch vibration at wavenumber 1530 cm^{-1} (6.54 μm) [19, 20]. The relative intensities of the peaks were slightly different for the photothermal spectra compared to the conventional IR spectra. Also some peaks that are not very prominent in conventional IR spectra appear to have higher intensity in the photothermal spectra. This may be directly related to the efficiency of non-radiative decay of these excited states.

It should be noted that the amount of adsorbate on the cantilever is in the range of monolayers on a surface and not bulk. Therefore, comparing the spectra from monolayers of materials with bulk spectra can be of concern. However, we did not observe any significant shift in the deflection peaks, as compared to standard spectra (bulk), due to substrate effects. One possibility is that the shifts due to surface effects may be too small to be detected by our technique and/or may be also masked by thermal broadening. It was, however, observed that photothermal deflection intensities did not match with the bulk spectra, most probably due to surface effects. The extent of bending is directly proportional to the adsorbed material, the power of IR radiation, the absorption mode, and the thermal sensitivity of the cantilever.

The thermal sensitivity of the cantilever can be further improved by controlling the thicknesses of the cantilever and the metal layer as well as the thermal expansion coefficient and Young's moduli of constituent elements of the bimaterial beam. For example, by selecting different metals and optimizing the thickness of the coating it is possible to make a bimaterial cantilever very sensitive to thermal changes. It is also possible to fabricate cantilevers with an optimized spring constant for increased thermal bending or to pattern the cantilever surface for increased adsorption. In general most of the heating caused by non-radiative decay is lost to the massive base of the cantilever. Only a small fraction of the heat is lost to the ambient air due to the large difference in thermal conductivity between the silicon and the air. Therefore, it is possible to increase the sensitivity by making the cantilever smaller as well as by optimizing the cantilever shape to restrict the heat flow from cantilever into the base of the cantilever. The sensitivity of photothermal deflection spectroscopy may be further increased by increasing the power of the illuminating IR source.

Conclusion

Despite its many advantages, the cantilever-based sensor platform for small molecules detection remains unattractive for practical applications due to the selectivity challenge when mixtures of vapors are present. Photothermal deflection spectroscopy, which combines the extremely high detection sensitivity of a cantilever and the selectivity of the infrared spectroscopy, offers an elegant way of overcoming the selectivity challenge. In this chapter we have described a photothermal technique that can add excellent selectivity to the cantilever sensor platform. The deflections of an adsorbate-covered bimaterial cantilever as a function of illuminating mid-infrared wavelength closely resemble the traditional infrared absorption spectrum of the adsorbed sample. The sensitivity of the technique was shown to be 500 pg. However, this limit of detection can be improved by optimization of the bimaterial cantilever. This method can be used in parallel with approaches based on chemically selective interfaces in a cantilever array providing orthogonal signals for pattern recognition. This technique has obvious advantages such as improved selectivity, sensitivity, low power operation, and fast detection and regeneration.

Acknowledgments This research was supported in part by U.S. Department of Homeland Security and the Office of Naval Research. ORNL is managed by UT-Battelle, LLC for the U.S. Department of Energy under contract no. DE-AC05-00OR22725.

References

1. Lavrik, N.V., Sepaniak, M.J., Datskos, P.G.: Rev. Sci. Instrum. **75**, 2229 (2004)
2. Thundat, T., Oden, P.I., Warmack, R.J.: Microscale Thermophys. Eng. **1**, 185 (1997)
3. Thundat, T., Warmack, R.J., Chen, G.Y., Allison, D.P.: Appl. Phys. Lett. **64**, 2894 (1994)
4. Chen, G.Y., Thundat, T., Wachter, E.A., Warmack, R.J.: J. Appl. Phys. **77**, 3618 (1995)
5. Butt, H.J., Colloid interface Sci., **180**, 251 (1996)
6. Fritz, J., Baller, M.K., Lang, H.P., Rothuizen, H., Vettiger, P., Meyer, E., Güntherodt, H.-J., Gerber, Ch., Gimzeski, J.K.: Science **288**, 316 (2000)
7. Berger, R., Delamarche, E., Lang, H.P., Gerber, C., Gimzewski, J.K., Mayer, E., Guntherdot, H.J.: Science **276**, 2021 (1997)
8. Boisen, A., Thaysen, J., Jensenius, H., Hansen, O.: Ultramicroscopy **82**, 11 (2000)
9. Pinnaduwage, L.A., Boiadjiev, V., Hawk, J.E., Thundat, T.: Appl. Phys. Lett. **83**, 1471 (2003)
10. Barnes, J.R., Stephenson, R.J., Welland, M.E., Gerber, Ch. Gimzewski, J.K.: Nature **372**, 79 (1994)
11. Krause, A., Van Neste, C., Senesac, L., Thundat, T., Finot, E.: J. Appl. Phys. **703**, 094906 (2008)
12. Li, G., Burggraf, L.W., Baker, W.P.: Appl. Phys. Lett. **76**, 1122 (2000)
13. Datskos, P.G., Rajic, S., Sepaniak, M.J., Lavrik, N., Tipple, C.A. Senesac, L.R., Datskou, I.: J. Vac. Sci. Technol. B **19**, 1173 (2001)
14. Datskos, P.G., Sepaniak, M.J., Tipple, C.A., Lavrik, N.: Sens. Actuators, B **76**, 393 (2001)
15. Arakawa, E.T., Lavrik, N.V., Rajic, S., Datskos, P.G.: Ultramicroscopy **97**, 459 (2003)
16. Arakawa, E.T., Lavrik, N.V., Datskos, P.G.: Appl. Opt. **42**, 1757 (2003)
17. Wig, A., Arakawa, E.T., Passian, A., Ferrell, T.L., Thundat, T.: Sens. Actuators, B **114**, 206 (2006)
18. Pristera, F., Halik, M., Castelli, A., Fredericks, W.: Anal. Chem. **32**, 495 (1960)

19. Lewis, I.R., Daniel, N.W. Jr., Griffiths, P.R.: Appl. Spectrosc. **51**, 1854 (1997)
20. Makashir, P.S., Kurian, E.M.: J. Therm. Anal. Calorim. **55**, 173 (1999)
21. Makashir, P.S., Kurian, E.M.: Propellants, Explos. Pyrotech. **24**, 260 (1999)
22. Beal, R.W., Brill, T.B., Appl. Spectrosc. **59**, 1194 (2005)

Chapter 7
Nanoelectronics for DNA Sensing

Predrag S. Krstić

Introduction

Deoxyribonucleic acid (DNA) is a molecule which carries the inheritable genetic information, a blueprint for each living organism. Double-stranded DNA is a combination of two antiparallel polymers, strands built of mutually weakly bonded monomers, the DNA nucleotides. Each of nucleotides is made of the sugar–phosphate backbone to which is attached a DNA basis. There are only four types of the base molecules, adenine (**A**), guanine (**G**), cytosine (**C**), and thymine (**T**) (Fig. 7.1).

A functional sequence of the nucleotides in a segment of DNA, gene, carries the genetic information. A human DNA, human genome, may contain a few billion bases but only tens of thousands of genes. The pairing of DNA nucleotides between the two antiparallel polymers, which creates the well-known double-stranded helical structure, follows particular rules, for example, A pairs to T, G to C. Therefore, DNA sequencing consists of determining the order of bases on a single strand (Fig. 7.2).

DNA sequencing is one of the leading precursors of the personalized medicine, i.e., use of individual genetic code to prevent, diagnose, and treat diseases. Besides medical diagnostic, it has a large importance in biotechnology and in forensic sciences. The next generation of genome sequencing technology is expected to bring the cost (<$1000) and time (<24 h) for sequencing a human genome into a range that is affordable for diagnosis and prognosis related to individual patients.

The development of the DNA sequencing technology is a collective effort in nanoscience which integrates biology, chemistry, physics, and engineering. Particularly interesting is the prospective of physics-based methods, since they are intrinsically fast and can operate on a single DNA molecule. These methods are the main subject of this chapter. The basic idea of the physical methods [1] is electrical detection of the DNA sequences by reading either the transversal electron tunneling

P.S. Krstić (✉)
Physics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6372, USA
e-mail: krsticp@ornl.gov

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_7,

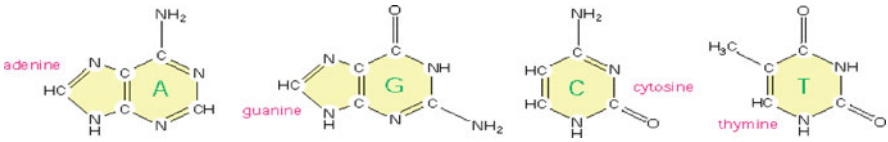


Fig. 7.1 DNA bases: A, G, C, T

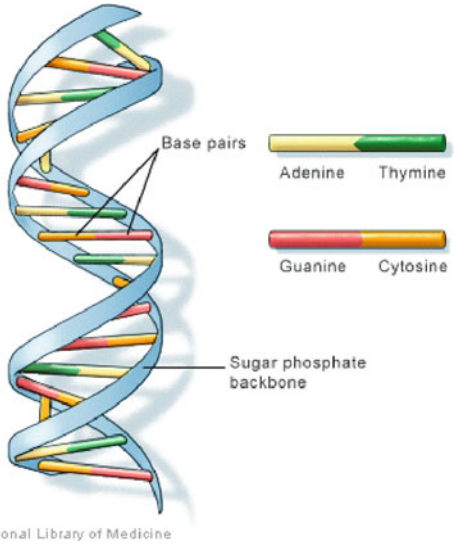


Fig. 7.2 A segment of the DNA double helix (U.S. National Library of Medicine)

current between the electrodes or the ionic current through a pore while a DNA is translocated through the synthetic nanopore or a nanogap (schematically, as in Fig. 7.3). The measurement of the conductance of a single molecule is an attractive concept for molecular detection because single-molecule conductance may be governed by its intrinsic electronic properties. The main challenges in this approach are twofold: (1) electric signal-to-noise ratio that distinguishes between the DNA bases and (2) full control and localization of a DNA while it is translocated through the nanoelectrode gap or nanopore.

However, a DNA experiences significant deformation and stretching when it is pulled through the gap. Furthermore, an across-the-gap E-field of 100 mV nm^{-1} is sufficient to attract and "glue" DNA to a panel electrode surface and prevent control of the translocation. In addition to these difficulties, the thermal fluctuations, Brownian motion, and transversal and electrophoretic fields cause strong fluctuations of instantaneous mutual position of the DNA bases and the electrodes, i.e., a large noise and poor signal-to-noise ratio in the transversal nonresonant tunneling conductance. Concerning improvements of the DNA translocation control, there are a number of approaches [1], for example, translocation through a Paul nanotrap [2] or through a narrow carbon nanotube (CNT) [3], in aqueous/electrolyte environment.

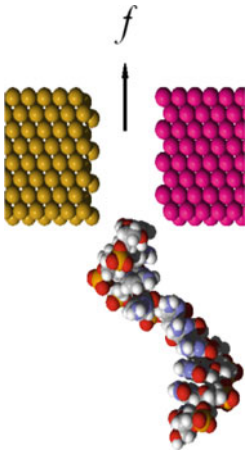


Fig. 7.3 A DNA translocation through a nanogap

The $1000 genome sequence obviously means also a fast sequence, and the race to also achieve this price tag and speed is one of the most active fields of research, even faster developing than computation power. Thus, in 2006 the price of a human genome sequencing was $22 million, which was a 150-fold decrease in cost to the 2002 first complete reading. However, in 2007, the sequencing cost was reduced to less than $1 million (3000 fold), while in 2009 the price was about $100,000 (i.e., 30000 fold less than 2002). If $1000 genome read is achieved in 2015 as expected, that would be 3-millions-fold reduction in price. Of course, this fast development is supported by strategic investment from public and private sectors to invent technology toward $1000 genome, led by the National Human Genome Research Institute (NHGRI) of the National Institute of Health (NIH).

In 2006, the X Prize Foundation established the Archon X Prize for Genomics (http://genomics.xprize.org) which will award $10 million to "the first Team that can build a device and use it to sequence 100 human genomes within 10 days or less, with an accuracy of no more than one error in every 100,000 bases sequenced, with sequences accurately covering at least 98% of the genome, and at a recurring cost of no more than $10,000 per genome."

Challenges

Currently there is considerable interest in studying the properties of macromolecules and biomolecules. There are several approaches to DNA sequencing [1, and references therein]. One method that was proposed early and is still considered to have a great potential uses nanopores for confinement of a DNA [4–11]. However, it has been realized that repeatable measurements of the base-specific signature of each nucleotide depends critically on the relative geometry of the bases to the pore and/or electrodes during the DNA sequencing [12, 13]. For example, it is found that the

variation in the conductance due to the geometry of the base relative to the electrode can easily override the difference between different types of nucleotides [14, 15]. Therefore, control of the orientation and position of the nucleotide as it threads the nanopore becomes a primary concern for the new generation of the DNA sequencing techniques [16–18]. The novel idea of localization of DNA molecules within a Paul-type quadrupole trap [19] offers increased electrical detection efficiency for heteropolymers confined within a nanopore regardless of detection scheme [20, 21]. The negatively charged DNA can be stabilized by a combined static and radio-frequency (RF) quadrupole trapping electric fields in a linear Paul trap, which control its translocation through the device. Each nucleotide on the DNA is expected to maintain its specific orientation or a stable average geometry.

The application of a synthetic nanopore sequencing device faces several key challenges that can be grouped into three categories: (1) control of the DNA translocation rate, (2) control of the nucleobase localization and orientation, and (3) control of the measured signal-to-noise ratio at the nucleobase-distinguishable level. These challenges must be addressed simultaneously since one needs to slow down the DNA dynamics sufficiently for each nucleobase to be read out by the electrodes.

Here are the theoretical considerations and predictions concerning these questions.

Electron Transport Through the DNA Nucleotides

Most of the theoretical considerations, developments, results, and conclusions shown in this section are fully described in [14, 15, 22–32].

Charge transport through biological molecules ultimately depends on the electronic and chemical structure of these molecules, altered by the presence of the electrodes. In [14, 15] we have explored whether the unique conductance signatures of the DNA bases can be used to sequence DNA, translocated through a nanogap between gold nanoelectrodes. The rationale was that the electronic and chemical structures of the four DNA bases (A, C, G, T) are intrinsically different (Fig. 7.1), and should, in principle, produce distinguishable electron transport properties [32]. Assuming existence of a practical way for stable threading of a single-stranded (ssDNA) polymer through the nm gap, the DNA sequence could be read from conductivity measurements as a function of the position of the polymer in the gap.

A fundamental element of such a method is the conductance through the DNA bases. The electronic transmission of a molecule is critically dependent on the contact between electrode and molecule [33], which largely explains the orders of magnitude disagreements between theory and early, groundbreaking scanning tunnel microscopy (STM) experiments in molecular electronics [34]. In the case of nucleotides, both sugar and phosphate groups must be considered as well, for they form part of an ssDNA. The previous theoretical work [15] focused on the basic mechanism—transverse transport of electrons through a nucleotide molecule placed in the nanogap between two nanoelectrodes, in vacuum. In order to study

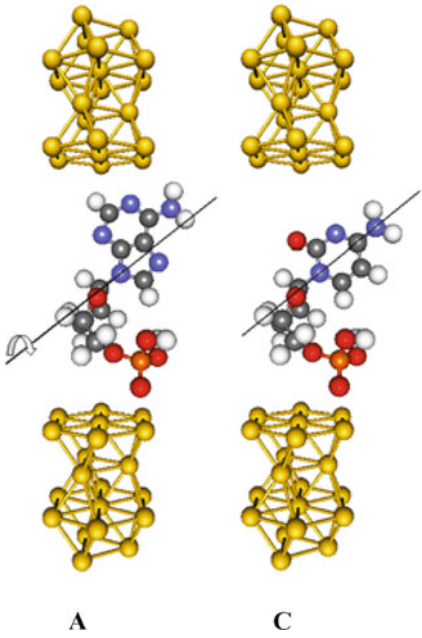


Fig. 7.4 The two DNA nucleotides, A and C, in the 1.5 nm gap between the gold electrodes. Glycosidic bond axis of rotation is shown, here 0°.

the effect of geometry on the conductance as well as to represent more reliably an average effect of the different conformations that an ssDNA segment would randomly assume in a possible experiment, the interelectrode–molecule geometry was varied, for example by varying the angle of rotation around the glycosidic bond axis (Fig. 7.4).

The calculations demonstrate that dominant transverse mechanism of electrical conduction is through the exponential tails of the electrode states near the Fermi energy, which at low voltages can be largely independent of the orbitals localized at the molecule, at energies far away from the Fermi energy of the gold electrodes.

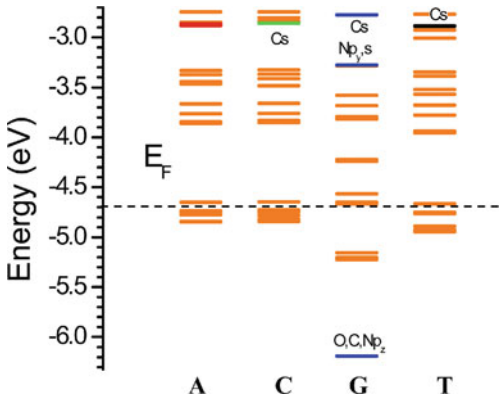


Fig. 7.5 Molecular Eigen-energy levels of the DNA nucleotides coupled to the Au electrodes (as in Fig. 7.4), at 10° glycosidic angle. Orange: Au-centered orbitals [15]. Other colors: orbitals centered at a nucleotide.

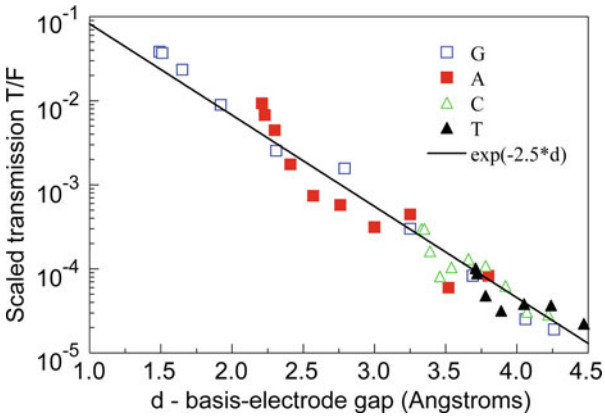


Fig. 7.6 Transmission in terms of geometrical factors: $T = F(\theta)e^{-kd}$, where F is a function of only geometrical factors defining the dimension of the nucleotides, and d the closest distance of a base atom and the gold electrode [14]

As long as the Fermi energy is not quasi-resonant with the levels of the molecule, which seems to be the case with the DNA nucleotide between the gold electrodes (Fig. 7.5), the dominant mechanism of the electrical conduction through a gap with a nucleotide is nonresonant tunneling. This indicates a conclusion with an overarching importance for the transverse-conductance measurements of the ssDNA, and for the ssDNA sequencing using such measurements. Namely in the absence of a resonance, the geometrical characteristics of the gap and the nucleotide, and the position of the nucleotide relative to the electrodes are far more important in the determination of the conductance than difference in electronic structure of the DNA bases (Fig. 7.6). As a result, it would be difficult to achieve single-base resolution due to the geometrical uncertainty (e.g., size of the molecule and the relative orientation between the molecule and the electrodes). Additional work was obviously required to understand the extent to which this distinguishability survives the refinement of the model systems toward more realistic representations of the ssDNA molecule in an aqueous environment, with averaging over a range of possible conformations of ssDNA, and with consideration of background Faraday currents, as well as of other uncertainties affecting the observed signal-to-noise ratios.

DNA Translocation Through Nanoelectrode Gaps

Cummings, Zhao, and coworkers applied extensive atomistic simulations of molecular dynamics of translocation of DNA segments through a metallic nanogap [25–28]. In their studies of DNA translocation they simulated the properties such as the structural transformation, energetics, as well as dynamics of the DNA molecule translocating through nanopores. They also studied how a DNA molecule behaves when an external force field exists. Their simulations showed that the DNA molecule

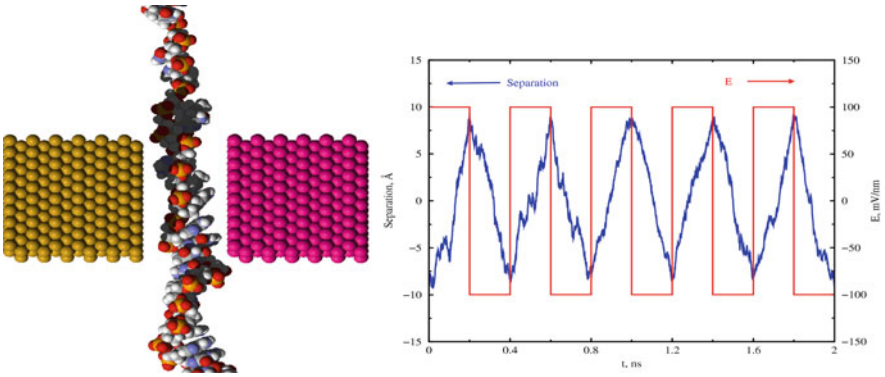


Fig. 7.7 DNA segments respond to reverse of E-field direction almost instantaneously

can be driven through a nanogap as narrow as 1.5 nm [25] and the motion of the DNA can be controlled by applying appropriate driving fields [26]. These simulation studies were performed in aqueous solution at atomistic level. They observed that the solvent plays an important role in the interaction between DNA and nanopores. These simulations demonstrated that molecular dynamics can simulate the behavior of a DNA molecule in aqueous solution with physically meaningful results.

DNA oscillates in a gap when oscillating electrophoretic field is applied. Interestingly, longer DNA is easier to control than shorter one. Once DNA is in the gap, the translocation process is very fast, less than ns per base. However, ssDNA experiences significant deformation and stretching when it is pulled through the gap. Furthermore, an across-the-gap E-field of 100 mV nm^{-1} is enough to attract and "glue" the DNA to the panel surface, as discussed in detail in [25–28]. In addition to these difficulties in translocation, the thermal fluctuations, Brownian motion, transversal and electrophoretic field cause strong fluctuations of instantaneous mutual position of the DNA bases and the electrodes, i.e., a large noise and poor signal-to-noise ratio in the transversal nonresonant tunneling conductance.

Search for Solutions

Quasi-resonant Tunneling

The measurement of the conductance of a single molecule is an attractive concept for molecular detection because single-molecule conductance may be governed by its intrinsic electronic structure. One of the main drawbacks of this concept is that the relevant mechanism for conductance is typically a nonresonant charge tunneling, which is not strongly dependent on the details of electronic structure, but rather on geometric conformations of the molecule in the gap/nanopore and nanoelectrodes. Therefore, the result is often measurements with poor sensitivity

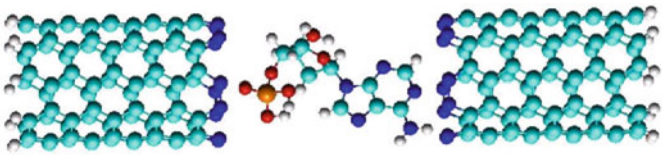


Fig. 7.8 Schematic of the prototype probe

and low signal-to-noise ratio. A possible solution is the application of chemically doped, i.e., functionalized, electrodes to increase coupling between the electrode and the molecule to a degree that electron tunneling through the gap is promoted into the resonant or quasi-resonant regime. This would enhance the measured signal by many orders of magnitude. Additional tuning capability can be generated by using a third gate to apply an electric field to shift the energy levels. These concepts have been recently discussed by Meunier and Krstic [23], leading to a theoretical prediction of a strongly enhanced current through DNA nucleotides in the gap between nitrogen-doped single-walled carbon nanotube (SWCNT) electrodes (Fig. 7.8). Theoretical predictions show a large increase of tunneling conductance of the DNA nucleotides noncovalently coupled to the nitrogen-functionalized electrodes [24]. The main appeal for using a CNT tip lies in the potential to tune its properties by chemical modification that consists of incorporating noncarbon atoms into the carbon network. This concept may be generalized by choosing the electrode functionalization to target specific types of molecules to be detected [31]. However, in order for these ideas to grow to a scientifically sound principle, experimental validation of the theoretical concepts will be needed.

The feature of chemical "activation" of the electrodes to enhance coupling with the chosen molecules is a powerful new idea that has much broader appeal than just for DNA sequencing. Overall increase of the nucleotide conductance and signal-to-noise ratio in all cases is expected through a transition from a regime of nonresonant to resonant tunneling, due to H-bond coupling of the nitrogen at the CNT with HN and HO terminal groups in a nucleotide. For example, while the H-doped or C-caped CNTs have a tunneling current through the molecule in the range of sub-picoamperes (pA), a device based on resonant tunneling would be capable of achieving tunneling currents in the nanoampere (nA) range (Fig. 7.9). The increase in the signal-to-noise ratio with the increased tunneling current promises to overcome the shortcomings of existing approaches and enable accurate measurements on single molecule. As a consequence of the hydrogen bonding between N at the electrode and N or O at the DNA, the energy levels of the molecular orbitals localized at the molecule are moved to an almost resonance with the Fermi energy of the electrodes (Fig. 7.10). This provides a vehicle for quasi-resonant tunneling of electron from one electrode to the other one.

Although results from theory and simulations of attainable base contrast using tunneling measurements with doped CNT described above are highly encouraging, successful operation of a nanopore detector for sequencing depends on a variety

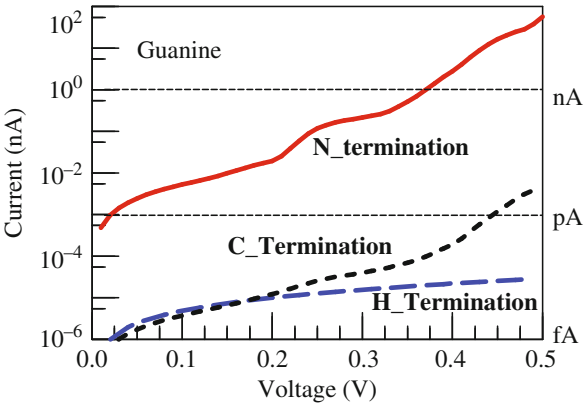


Fig. 7.9 Comparison of the *I–V* characteristics of guanine with N-, C-, and H-terminated CNT leads [24]

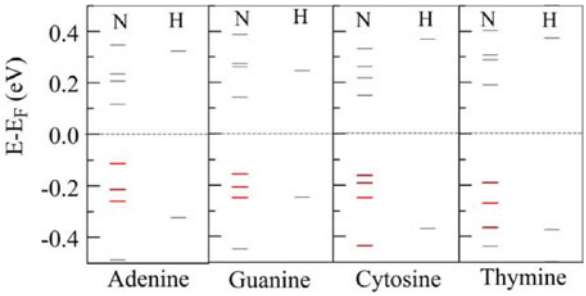


Fig. 7.10 Molecular orbital energies of the system if the type as in Fig. 7.8, for the cases of N and H functionalization of the CNT tip (*black* localized at C, *red* localized at DNA)

of chemical and electrical interactions between DNA molecules, electrical probes, the fluidic environment, and the nanopore itself. The experimental validation of the theoretical discovery is necessary for establishing the principle.

Detecting single chemical and biological molecules without using receptors or chemical interfaces is a paradigm shift in chemical sensing [35, 36]. This approach may also pave the way for developing receptor-free sensing concepts that are purely electrical. It may be possible to develop arrays of nanotubes where each tube is functionalized with moieties that can produce resonant tunneling for specific analytes.

Paul Trap for DNA Translocation

The sensitivity of quantum transport on the interface geometry is not limited to the transversal DNA measurements; it is a universal effect that makes measurement in a single molecule particularly difficult [35, 36]. An attractive idea, emerging from theoretical simulations [20], is to develop a nanofilter-trap which would provide opportunity for isolation and full control of geometry during DNA translocation,

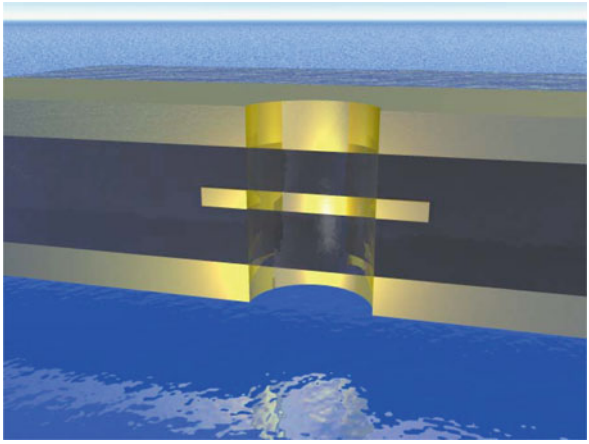


Fig. 7.11 Typical fabrication configuration of a 3D Paul trap (courtesy of Mark Reed)

as the predecessor for the sequencing. A quadrupole micro Paul trap [19] was fabricated and tested in the aqueous environment [37]. Independent of a sensing approach in a future DNA sequencing device, the Paul trap is an attractive approach to localization and stabilization of the single molecular ions, ssDNA in particular. Figure 7.11 shows a typical three-electrode configuration for a 3D Paul trap. This is capable of confining a molecule in about a center of the trap. More attractive for the DNA sequencing tasks is a 2D (linear) Paul trap, which confines an ion around the axis of the trap, enabling its translocation along the axis.

The main advantage of the Paul-type nanotrap is relaxation of critical dimension control. Critical dimension control becomes very problematic at dimensions below 10 nm due to resist resolution and pattern transfer limits of electron-beam lithography. This approach avoids the fabrication difficulties since the electrostatic trapping volume (~nm) is significantly smaller than the fabricated dimensions (20–100 nm range in future, currently in the micrometer range). Large-scale molecular dynamics simulations indicate that this approach to enhanced molecular localization and control is feasible. The fabrication scheme is relatively simple, utilizing a self-aligned approach that positions the Paul trap electrodes automatically in the center of the physical nanopore through which the ssDNA will translocate.

A Paul ion trap exists in both linear and 3D varieties and refers to a quadrupole trap that uses DC and radiofrequency (RF) oscillating AC electric fields to trap ions. The 2D version is a critical component of a quadrupole mass spectrometer. The invention of the quadrupole ion trap is attributed to Wolfgang Paul who shared the Nobel Prize in physics in 1989 for this work [19]. The 3D trap generally consists of two hyperbolic metal electrodes with their foci facing each other and a hyperbolic ring electrode halfway between the other two electrodes. The ions are trapped in the space between these three electrodes by AC and DC electric fields. An excellent review of the functions and applications of the conventional Paul ion traps is given by Leibfried et al. [37].

The resultant 3D Paul trap potential alternates between a focusing potential well in the axial direction with a defocusing inverted radial potential, and a focusing potential well in the radial direction with a defocusing axial potential. By focusing particles in one direction and then in the other, the dynamic Paul trap potential creates a stable equilibrium whereas the static quadrupole potential was capable of only unstable equilibrium (Fig. 7.12).

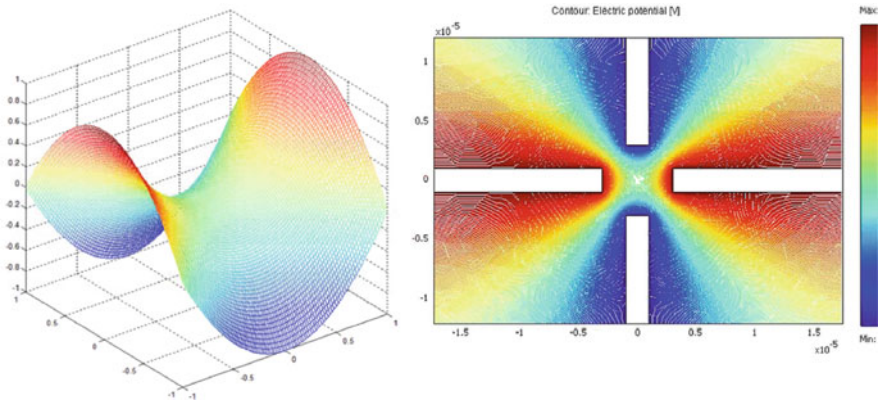


Fig. 7.12 At 1 μm spacing, a saddle potential with a trap depth of ∼1 V is achieved

More recent applications of the Paul traps are for quantum information processing [38–39], for coherent quantum-state manipulation of trapped atomic ion [40], for laser sideband cooling of the motion [41], for formation of ordered structures of trapped ions [42–46], etc.

Though Paul trap has been generally used for trapping of single ions in vacuum, recent molecular dynamics simulations have demonstrated that a nanoscale quadrupole Paul trap is capable of effectively confining ions in an aqueous environment [20]. Such environment might be crucial for supporting the trap functions for biomolecular ions like DNA.

If the applied electric bias is $\Phi_0 = 2(V_{dc} - V_{ac}\cos\Omega t)$, the resulting electric potential in the space between the electrodes is given by $\phi(x, y, t) = \Phi_0 \frac{x^2 - y^2}{2R_0^2}$. The motion for a spherical charged particle of mass M, radius r, and charge Q can be described by Langevin equation. For the motion in x direction with neglecting the thermal fluctuation force, it takes the parametric dimensionless form:

$$\frac{d^2x}{d\tau^2} + b\frac{dx}{d\tau} + (a - 2q\cos 2\tau)x = 0, \tag{7.1}$$

where $\tau = \frac{\Omega t}{2}$ is a dimensionless scaled time, $a = \frac{4QU}{MR_0^2\Omega^2}$ is the dimensionless DC voltage, $q = \frac{2QV}{MR_0^2\Omega^2}$ is the dimensionless AC voltage (Fig. 7.13(a)), and $b = \frac{2\xi}{M\Omega}$ is the dimensionless damping coefficient.

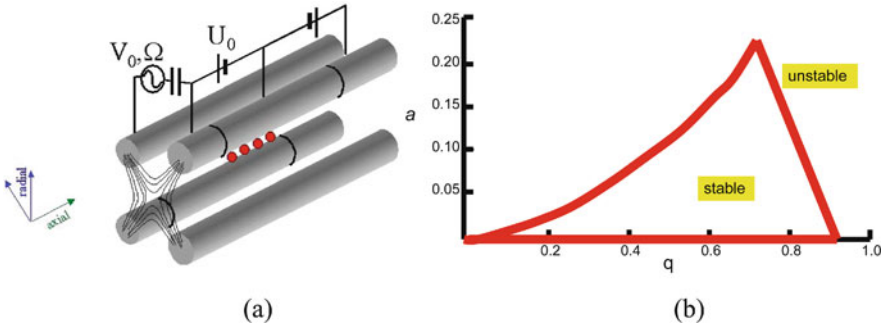


Fig. 7.13 (**a**) Schematic sketch of a 2D Paul trap and (**b**) the stability diagram

The motion in the y-direction can be described with the same expression of Eq. (7.1), except replacing parameters a and q with $-a$ and $-q$, respectively. ξ is the Stokes drag coefficient and can be well approximated by $\xi = 6\pi\eta r$ for small Reynolds number, where η is the dynamic viscosity of medium (at room temperature, water has a dynamic viscosity of 1×10^{-3} Pa·s)

As is well known for a Paul trap in vacuum environment, stable trapping will only occur when parameters are within certain regions in the $q-a$ diagram (Fig. 7.13(b)). For a nanoscale linear Paul trap, the dimension r_0 could range from a few nanometers to hundreds of nm. As a result, to maintain stability, the other parameters have to adjust resulting in large values for the frequency ranging from hundreds of MHz to few GHz. The voltages are limited by the breakdown strength of the medium. In our modeling of the trap, we have considered potentials varying from 1 to 10 V, frequencies ranging from 400 MHz to 10 GHz, and radius of the trap from 5 to 100 nm. If a viscous medium is present (e.g., air or water), the stable trapping region will be not only shifted but also extended [47].

The solution of Eq. (7.1) can be obtained from Floquet's theorem if a transformation is made to change the equation into the canonical form of Mathieu equation [48]. In the stability region of Eq. (7.1), the position of the particle should damp to the center of the device and eventually settle at rest when time $t \to \infty$. This is, however, not the case in the presence of the Brownian motion, which adds the fluctuations in the position of the trapped particle.

When $q << 1$ and $a << 1,\ b = 0$, the first-order solution to Eq. (7.1) is

$$\begin{aligned} x(t) &= x_0 \cos(\Omega t + \varphi)\left[1 + \tfrac{q}{2}\cos(\omega t)\right]. \\ \Omega &= \omega\sqrt{a^2 + q^2/2} \end{aligned} \tag{7.2}$$

Ω is the secular frequency and φ is a phase determined by the initial conditions of position and velocity of the particle. The motion corresponding to $\cos(\omega t)$ term

is the micromotion, driven by the applied ac field. Higher order solutions in powers of q and a contain integer multiple of the frequencies in Eq. (7.2).

When using Paul trap for localization and control of a DNA molecule, several challenging questions arise:

(1) Can Paul trap function at nanoscale?
(2) Can a DNA segment (generically line charge) be trapped in a Paul trap?
(3) Can Paul trap confine an ion in an aqueous environment?
(4) Does electrolyte concentration in the aqueous Paul trap screen the electrode potentials, disabling the confinement functions?

The answers to these questions came through both theoretical and experimental considerations. The nanometer size of the trap would bring new, not observed in the macrotraps, phenomena that may influence the trap confinement functionality. Thus, van der Waals forces from the electrodes, nonhomogeneities and fluctuations of the dielectric constant, temperature, viscosity, and Brownian motion can contribute to the nondesirable effects. The MD simulations of Zhao et al. [20] have shown that these effects do not act destructively to the confinement functions of the trap, and stabilization was reached in the nanosecond time for a wide set of the ion initial conditions in the configuration space. Also, using molecular dynamics simulations, Joseph et al. [2] have studied the dynamics of a filamented structure with line-distributed charges such as a charged linear polymer and a single-stranded DNA in a nanoscale 2D Paul trap, in vacuum. A line charge can be effectively trapped, with stability parameters similar in values to the single charged particle of the same Q/M.

Figure 7.15 shows oscillations of a center of mass of an ssDNA, carrying various bases: G, A, T, and C. These oscillations are shifted both in phase and in frequency. The power spectrum density (PSD) in Figs. 7.14(b) and 7.15(b) is defined by

$$S(\omega) = |F_T(\omega)| = \left| \int_{-T/2}^{T/2} r(t) e^{-j\omega t} dt \right|. \tag{7.3}$$

The stability regions in the $q-a$ diagram for various damping factors b can be numerically determined by using Hasegawa and Uehara's method [47] and are shown by the black regions in Fig. 7.16. Figure 7.16(a) shows a case in vacuum. However, with increase of viscosity, the range of stability increases, as shown in the (q, a) diagram in Fig. 7.16(b), and (q, k) diagram in Fig. 7.16(c, d). We note that fluctuations significantly influence the range of stability as seen in Fig. 7.16(c, d),

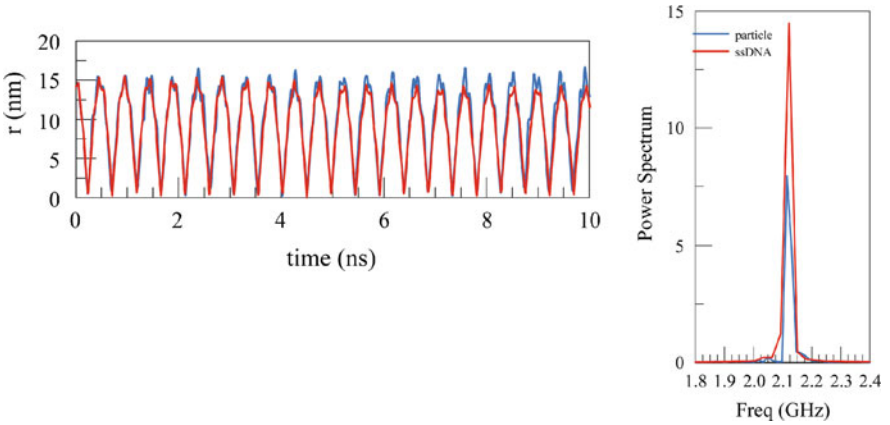


Fig. 7.14 (**a**) The r trajectories of a 60 guanine-base ssDNA as a function of time compared with the trajectory of a single particle with the same q-factor 0.2938. The parameters for the trap are $V_{ac} = 10$ V, $r_0 = 100$ nm, $f = 10$ GHz. (**b**) Comparison of the center of mass power spectrum (Fourier transform) for the ssDNA with that of a particle with the same q-factor. The overlapping peaks show that the secular oscillation of a single particle is similar to the ssDNA even though the charges are distributed

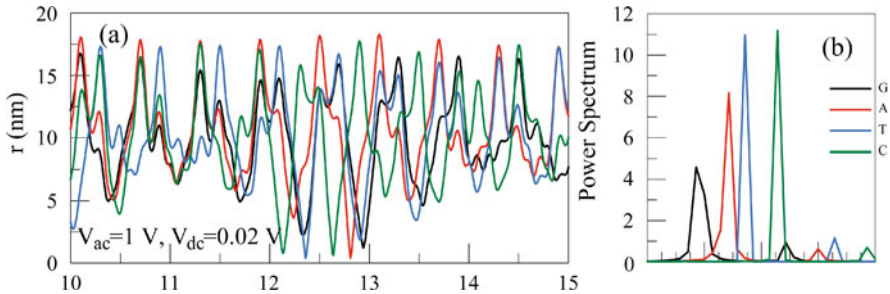


Fig. 7.15 (**a**) Trajectories of the center of mass of ssDNAs with different bases for $V_{dc} = 0.02$ V, $V_{ac} = 1$ V, $r_0 = 50$ nm, $f = 5$ GHz and (**b**) PSD of a 60-base ssDNA for $V_{dc} = 0.02$ V, $V_{ac} = 1$ V, $r_0 = 50$ nm, $f = 5$ GHz

where a phenomenological parameter c is introduced to estimate the effect of thermal fluctuations ($c = 1$ in the absence of fluctuations). The parameter k is equal to the dimensionless b, defined below Eq. (7.1).

In conclusion, the feasibility of a Paul-type trap working in an aqueous solution was demonstrated. The trap utilizes the strong alternating confining force and dynamically traps charged particles in the center of the planar device. This technique opens the possibility to spatially control a bio-object in aqueous solution and can lead to lab-on-a-chip systems controlling single molecules. Further investigations such as the impact of variation of the solution's ion composition and pH on the

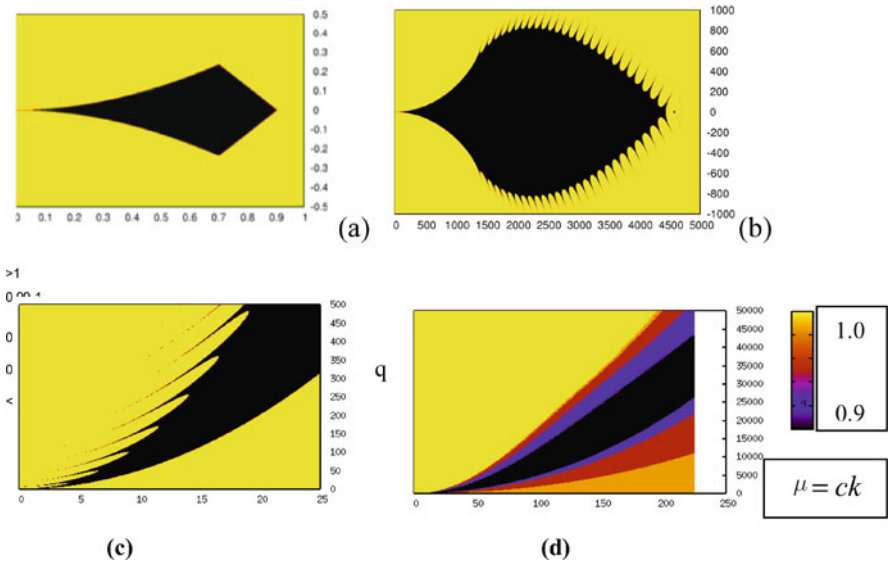


Fig. 7.16 Stability and reduction of the stability region due to random fluctuations. (**a**), (**b**) are a–q diagrams for $b = 0$ and $b = 92$, respectively; (**c**), (**d**) are b–q diagrams for $a = 0$

trapping performance are needed for a better understanding of the Paul trap for biomolecular applications.

Acknowledgments This research was supported by the U.S. National Human Genome Research Institute of the National Institutes of Health under grant no. 1R21HG004764-01. The author acknowledges support by U.S. Department of Energy (DOE) at ORNL managed by a UT-Battelle for the U.S. DOE under contract no. DEAC05-00OR22725, by the U.S. DOE, and U.S. NHGRI under grant no. 1 R21 HG003578-01. A portion of this research was conducted at the Center for Nanophase Materials Sciences, which is sponsored at ORNL by the Division of Scientific User Facilities, U.S. DOE.

References

1. Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S.B., Krstic, P.S., Lindsay, S., Ling, X.S., Mastrangelo, C.H., Meller, A., Oliver, J.S., Pershin, Y.V., Ramsey, J.M., Riehn, R., Soni, G.V., Tabard-Cossa, V., Wanunu, M., Wiggin, M., Schloss, J.A.: The Potential and Challenges of Nanopore Sequencing. Nature Biotechnol. **26**, 1146 (2008)
2. S. Joseph, W. Guan, M.A. Reed, and P.S. Krstic: A long DNA segment in a linear nanoscale Paul trap. Nanotechnol. **21**, 015103 (2010)
3. Haitao Liu, Jin He, Jinyao Tang, Hao Liu, Pei Pang, Di Cao, Predrag Krstic, Sony Joseph, Stuart Lindsay, Colin Nuckolls: Translocation of Single-Stranded DNA Through Single-Walled Carbon Nanotubes. Science **327**, 64 (2010)
4. Chan, E.Y.: Advances in sequencing technology. Mutat. Res. Fund. Mol. Mech. Mut. **573** 13–40 (2005)

5. Fredlake, C. P., Hert, D. G., Mardis, E. R., Barron, A. E.: What is the future of electrophoresis in large-scale genomic sequencing? Electrophoresis **27**, 3689–702 (2006)
6. Healy, K.: Nanopore-based single-molecule DNA analysis. Nanomedicine **2**, 459–81 (2007)
7. Kricka, L. J., Park, J. Y., Li, S. F. Y. Fortina, P.: Miniaturized detection technology in molecular diagnostics. Exp. Rev. Mol. Diagn. **5**, 549–59 (2005)
8. Nakane, J. J., Akeson, M., Marziali, A.: Nanopore sensors for nucleic acid analysis. J. Phys.: Condens. Matter **15**, R1365–R1393 (2003)
9. Rhee, M., Burns, M. A.: Nanopore sequencing technology: Research trends and applications. Trends Biotechnol. **24**, 580–6 (2006)
10. Rhee, M., Burns, M. A.: Nanopore sequencing technology: Nanopore preparations. Trends. Biotechnol. **25**, 174–81 (2007)
11. Ryan, D., Rahimi, M., Lund, J., Mehta, R., Parviz, B. A.: Toward nanoscale genome sequencing. Trends Biotechnol. **25**, 385–9 (2007)
12. Lagerqvist, J., Zwolak, M., Di Ventra, M.: Comment on Characterization of the tunneling conductance across DNA bases. Phys. Rev. E **76**, 013901 (2007)
13. Tabard-Cossa, V., Trivedi, D., Wiggin, M., Jetha, N. N., Marziali, A.: Noise analysis and reduction in solid-state nanopores. Nanotechnology **18** 305505 (2007)
14. Zhang, X.-G., Krstic, P.S., Zikic, R., Wells, J. C., Fuentes-Cabrera, M.: First-principles transversal DNA conductance deconstructed. Biophys. J. **91**, L04 (2006)
15. Zikic, R., Krstic, P.S., Zhang, X.-G., Fuentes-Cabrera, M., Wells, J., Zhao, X.C.: Characterization of the tunneling conductance across DNA bases. Phys. Rev. E **74**, 011919 (2006)
16. Trepagnier, E. H., Radenovic, A., Sivak, D., Geissler, P., Liphardt, J.: Controlling DNA capture and propagation through artificial nanopores. Nano Lett. **7**, 2824–30 (2007)
17. Tsai, Y.S., Chen, C.M.: Driven polymer transport through a nanopore controlled by a rotating electric field: off-lattice computer simulations. J. Chem. Phys. **126**, 144910 (2007)
18. Chen, C.M., Peng, E.H.: Nanopore sequencing of polynucleotides assisted by a rotating electric field. Appl. Phys. Lett. **82** 1308–10 (2003)
19. Paul, W.: Electromagnetic traps for charged and neutral particles. Rev. Mod. Phys. **62**, 531 (1990)
20. Zhao, X., Krstic, P.S.: A molecular dynamics simulation study on trapping ions in a nanoscale Paul trap. Nanotechnology **19**, 195702(2008)
21. Arnott, D., Henzel, W.J., Stults, J.T.: Rapid identification of comigrating gel-isolated proteins by ion trap mass spectrometry. Electrophoresis **19**, 968–80 (1998)
22. Zikic, R., Krstic, P.S., Zhang, X.-G., Fuentes-Cabrera, M., Wells, J., Zhao, X.C.: Comment on Characterization of the tunneling conductance across DNA bases. Reply, Phys. Rev. E **76**, 013902 (2007)
23. Meunier, V., Krstic, P.S.: Enhancement of the Transverse Conductance in DNA nucleotides. J. Chem. Phys. **128**, 041103 (2008)
24. Zhao, X.C., Payne, C.M., Cummings, P.T., Lee, J.W.: Single stranded DNA molecules translocation through nanoelectrode gaps. Nanotechnology **18**, 424018 (2007)
25. Zhao, X.C., Payne, C.M., Cummings, P.T.: Controlled translocation of DNA segments through nanoelectrode gaps from molecular dynamics. J. Phys. Chem. C 112, 8 (2008)
26. Payne, C.M., Zhao, X.C., Vlcek, L., Cummings, P.T.: Molecular dynamics simulation of ss-DNA translocation through a copper nanoelectrode gap. J. Phys. Chem. B112, 1712 (2008)
27. Payne, C.M., Zhao, X.C., Cummings, P.T.: Molecular simulation of DNA transport in solution. Mol. Simulat. **33**, 399 (2007)
28. Predrag Krstic, Erica Forzani, Nongjian Tao, Anatoli Korkin, Design and function of molecular and bioelectronics devices. Editorial Nanotechnol **18**, 420201 (2007)
29. Miguel Fuentes-Cabrera, Vincent Meunier, Bobby G Sumpter, Benzo homologated nucleobases in a nanotube-electrode set-up for DNA sequencing. Nanotechnology **18**, 424019 (2007)

30. Krstic, P.S., Wells, J.C., Miguel Fuentes-Cabrera, Dong Hu, Lee, J.W.: Toward Electronic Conductance Characterization of DNA Nucleotide Bases. Solid State Phenomena **121–123**, 1387 (2007)
31. Krstic, P.S., Meunier, V.: Sensing Single Molecules Between Doped Carbon Nanotubes, Invention disclosure, Docket #: 1300001981, DOE S#: S-111, 599 (August 2007)
32. Lee, J.W., Thundat, T.G., Greenbaum, E.: DNA and RNA sequencing by nanoscale reading through programmable electrophoresis and nanoelectrode-gated tunneling and dielectric detection, US Patent No. 6905586 (2005)
33. Cui, X.D., Primak, A., Zarate, X., Tomfohr, J., Sankey, O.F., Moore, A.L., Moore, T.A., Gust, D., Harris, G., Lindsay, S.M.: Reproducible measurement of single-molecule conductivity. Science **294**, 571 (2001)
34. Reed, M.A., Zhao, C., Muller, C.J., Burgin, T.P., Tour, J.M.: Conductance of a molecular junction. Science **278**, 252 (1997)
35. Roy, S., Vedala, H., Datta Roy, A., Do-h Kim, Doud, M., Mathee, K., Shin, H-k, Shimamoto, N., Prasad, V., Choi, W.: Direct Electrical Measurements on Single-Molecule Genomic DNA Using Single-Walled Carbon Nanotubes. Nano Letter **8**, 26 (2008)
36. Tao, N.J.: Electron transport in molecular junctions. Nature Nanotechnol. **1**, 173 (2006)
37. Leibfried, D., Blatt, R., Monroe, C., Wineland, D.: Quantum dynamics of single trapped ions. Rev. Mod. Phys. **75**, 281–324 (2003)
38. Vant, K., Chiaverini, J., Lybarger, W., Berkeland, D.J.: Photoionization of strontium for trapped-ion quantum information processing arXiv:quant-ph 0607055 (2006)
39. Seidelin, S., Chiaverini, J., Reichle, R., Bollinger, J.J., Leibfried, D., Britton, J., Wesenberg, J.H., Blakestad, R.B., Epstein, R.J., Hume, D.B., Itano, W.M., Jost, J.D., Langer, C., Ozeri, R., Shiga, N., Wineland, D.J.: Microfabricated surface-electrode ion trap for scalable quantum information processing. Phys. Rev. Lett. **96**, 253003 (2006)
40. Wineland, D.J., Monroe, C., Itano, W.M., Leibfried, D., King, B.E., Meekhof, D.M.: Experimental issues in coherent quantum-state manipulation of trapped atomic ions J. Res Natl. Inst. Stand. Technol **103**, 259–328 (1998)
41. Abich, K., Keil, A., Reiss, D., Wunderlich, C., Neuhauser, W., Toschek, P.E.: Thermally activated hopping of two ions trapped in a bistable potential well. J. Optics B-Quant Semiclassical Optics **6**, S18–S23 (2004)
42. Schiffer, J.P.: Order in confined ions. J. Phys. B-Atomic Mol. Optical Phys. **36**, 511–23 (2003)
43. Shi, L., Zhu, X.W., Feng, M., Fang, X.M.: Ordered structures of a few ions. Paul trap Commun. Theor. Phys. **31**, 491–6 (1999)
44. Itano, W.M., Bergquist, J.C., Bollinger, J.J., Wineland, D.J.: Cooling methods in ion traps. Phys. Scripta **T59**, 106–20 (1995)
45. Walther, H.: From a single-ion to a mesoscopic system—crystallization of ions. Paul traps Physica Scripta **T59**, 360–8 (1995)
46. Edwards, C.S., Gill, P., Klein, H.A., Levick, A.P., Rowley, W.R.C.: Laser-cooling effects in few-ion clouds of YB+. Appl. Phys. B-Lasers Optics **59**, 179–85 (1994)
47. Hasegawa, T., Uekara, K.: Appl. Phys. B **1995**, 61, 159–163
48. Chicone, C.: Ordinary Differential Equations with Applications, 2nd edn, Texts in Applied Mathematics, NewYork: Springer Verlag (2006)

Chapter 8
Nanostructured Electrode Materials for Lithium-Ion Batteries

A. Manthiram and T. Muraliganth

Abstract Electrochemical energy storage systems are becoming increasingly important with respect to their use in portable electronic devices, medical implant devices, hybrid electric and electric vehicles, and storage of solar and wind energies. Lithium-ion batteries are appealing for these needs because of their high energy density, wide range of operating temperatures, and long shelf and cycle life. However, further breakthroughs in electrode materials or improvements in existing electrode materials are critical to realize the full potential of the lithium-ion technology. Nanostructured materials present an attractive opportunity in this regard as they could offer several advantages such as fast reaction kinetics, high power density, good cycling stability with facile strain relaxation compared to their bulk counterparts. In this chapter, we present an overview of the latest progress on the nanostructured anode and cathode materials for lithium-ion batteries.

Introduction

Energy storage and production is one of the greatest challenges facing human kind in the twenty-first century. Based on moderate economic and population growth, the global energy consumption is anticipated to triple by the year 2010. The rapid depletion of fossil fuel reserves, increasing fuel costs, increasing greenhouse gas emission, and global climate changes are driving the development of sustainable clean energy technologies like fuel cells, solar cells, high energy density batteries, and supercapacitors. While fuel cells and solar cells are energy conversion devices, batteries and supercapacitors are energy storage devices.

Nearly one-third of the total energy consumption in the United States is by the transportation sector, and internal combustion engine based automobiles are a significant contributor of green house gas emission. Hybrid electric vehicles (HEV) and plug-in hybrid electric vehicles (PHEV) are the most viable near-term option

A. Manthiram (✉)
Materials Science and Engineering Program, The University of Texas, Austin, TX 78712, USA
e-mail: rmanth@mail.utexas.edu

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_8,

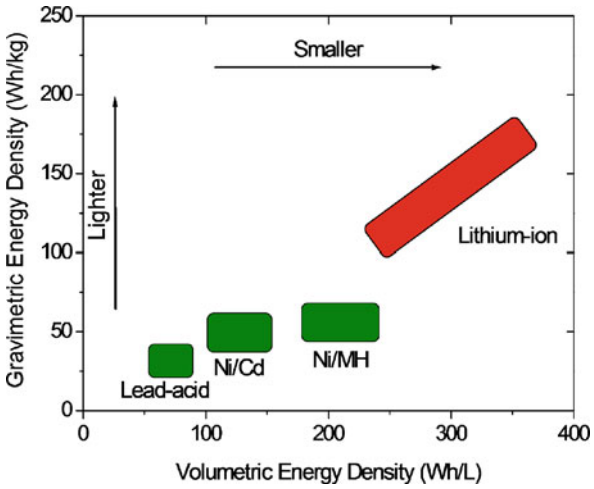


Fig. 8.1 Comparison of the energy densities of different battery systems

for transportation. Energy storage is also critical for the efficient utilization of the electricity produced from solar and wind energies. Lithium-ion batteries are appealing for automobiles (HEV and PHEV) and the storage of solar and wind energy as they provide higher energy density compared to other rechargeable battery systems such as lead acid, nickel–cadmium, and nickel metal hydride batteries as seen in Fig. 8.1 [1]. However, cost and safety are the major issues with respect to large-scale applications like automobiles and solar and wind energy storage, and they need to be adequately addressed.

Lithium-Ion Batteries

Rechargeable lithium batteries involve a reversible insertion/extraction of lithium ions into/from a host electrode material during the charge/discharge process. The lithium insertion/extraction process occurring with a flow of ions through the electrolyte is accompanied by a reduction/oxidation (redox) reaction of the host matrix assisted with a flow of electrons through the external circuit (Fig. 8.2).

The open-circuit voltage V_{oc} of such a lithium cell is given by the difference in the lithium chemical potential between the cathode ($\mu_{Li(c)}$) and the anode ($\mu_{Li(a)}$) as

$$V_{oc} = (\mu_{Li(c)} - \mu_{Li(a)})/F, \tag{8.1}$$

where F is the Faraday constant. A schematic energy diagram of a cell at open circuit is given in Fig. 8.3. The cell voltage is determined by the energies involved in both the electron transfer and Li^+ transfer. While the energy involved in electron transfer is related to the work functions of the cathode and the anode, the energy involved in Li^+ transfer is determined by the crystal structure and the coordination geometry of the site into/from which the Li^+ ions are inserted/extracted [2]. Thermodynamic stability considerations require the redox energies of the cathode (E_c) and anode

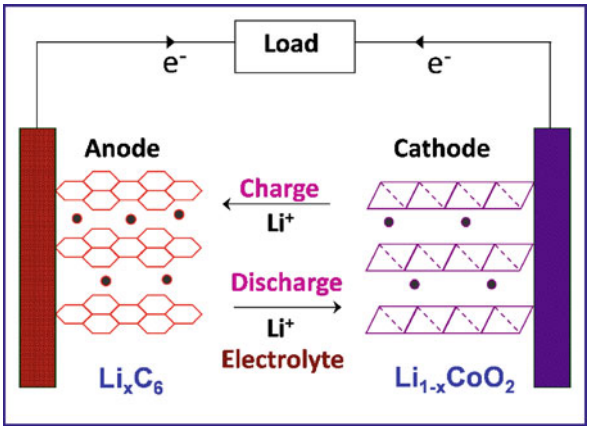


Fig. 8.2 Illustration of the charge–discharge process involved in a lithium-ion cell consisting of graphite as an anode and layered $LiCoO_2$ as a cathode

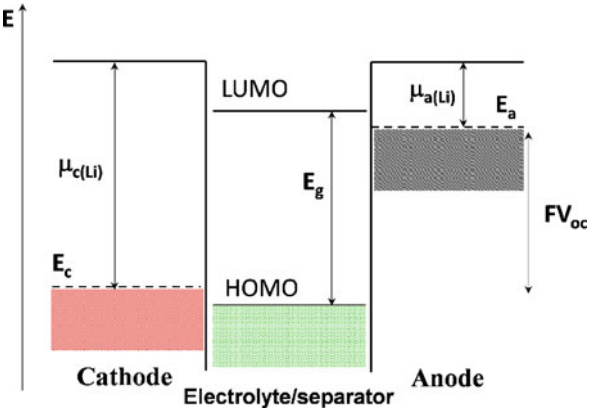


Fig. 8.3 Schematic energy diagram of a lithium cell at open circuit. HOMO and LUMO refer, respectively, to the highest occupied molecular orbital and lowest unoccupied molecular orbital in the electrolyte

(E_a) to lie within the band gap E_g of the electrolyte, as shown in Fig. 8.3, so that no unwanted reduction or oxidation of the electrolyte occurs during the charge–discharge process. With this strategy, the anode and cathode insertion hosts should have, respectively, the lowest and highest voltages vs. metallic lithium in order to maximize the cell voltage.

The concept of rechargeable lithium batteries was first illustrated with a transition metal sulfide TiS_2 as the cathode, metallic lithium as the anode, and a nonaqueous electrolyte [3]. Following the initial demonstration, several other sulfides and chalcogenides were pursued during the 1970s and 1980s as cathodes [4]. However, most of them exhibit a low cell voltage of <2.5 V vs. lithium anode. This limitation in cell voltage is due to an overlap of the higher valent M^{n+}:d band with the top of the nonmetal:p band—for example, with the top of the S:3p band as shown in Fig. 8.4. Such an overlap results in an introduction of holes into (or removal of electrons from) the S^{2-}:3p band and the formation of molecular ions such as S_2^{2-},

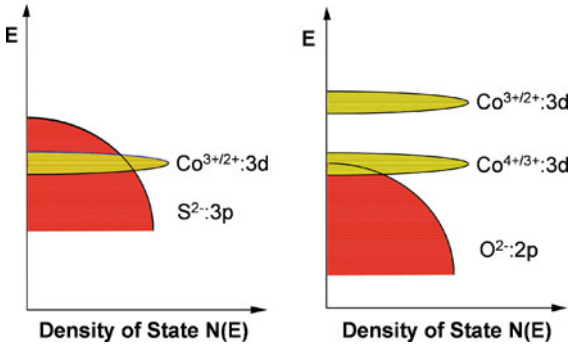


Fig. 8.4 Relative energies of metal:d (for example, Co:3d) and nonmetal:p in (**a**) a sulfide and (**b**) an oxide

which in turn leads to an inaccessibility of higher oxidation states for M^{n+} in a sulfide $Li_xM_yS_z$. The stabilization of higher oxidation state is essential to maximize the cell voltage. Recognizing this difficulty with chalcogenides, Goodenough's group at the University of Oxford focused on oxide cathodes during the 1980s [5–7]. The location of the top of the O^{2-}:2p band much below the top of the S^{2-}:3p band and a larger raising of the M^{n+}:d energies in an oxide compared to that in a sulfide due to a larger Madelung energy (Fig. 8.4) make the higher valent states accessible in oxides. For example, while Co^{3+} can be readily stabilized in an oxide, it is difficult to stabilize Co^{3+} in a sulfide since the $Co^{2+/3+}$ redox couple lies within the S^{2-}:3p band as seen in Fig. 8.4. Accordingly, several transition metal oxide hosts crystallizing in a variety of structures (two-dimensional layered and three-dimensional framework structures) have been pursued during the past two decades. Among them $LiCoO_2$, $LiNiO_2$, and $LiMn_2O_4$ oxides having a high electrode potential of 4 V vs. metallic lithium have become attractive cathodes for lithium-ion cells. On the other hand, graphite with a low electrode potential of <0.3 V vs. metallic lithium has become an attractive anode.

Although the initial efforts in rechargeable lithium batteries employed lithium metal as the anode, the strategy failed to attain commercial success due to safety limitations [8, 9]. The inherent instability of lithium metal and the dendrite formation during charge–discharge cycling eventually forced the use of intercalation compounds as anodes. This led to the commercialization of the lithium-ion battery technology by Sony in 1990 with $LiCoO_2$ as the cathode and graphite as the anode. However, the cost and safety issues and the performance limitations associated with the $LiCoO_2$ cathode and carbon anode have prompted the search for new electrode materials as well as improvements in existing electrode materials.

Nanostructured materials have attracted a lot of interest in the past two decades because of their unusual electrical, mechanical, and optical properties compared to their bulk counterparts [10–15]. Recently, nanostructured materials are also gaining increasing popularity for energy storage applications [16–21]. This chapter focuses on providing an overview of some of the recent advances in nanostructured anode and cathode materials for lithium-ion batteries.

Nanostructured Electrode Materials for Lithium-Ion Batteries

Nanomaterials offer advantages and disadvantages as electrode materials for lithium-ion batteries. Some of the advantages are given below:

- The smaller particle size increases the rate of lithium insertion/extraction because of the short diffusion length for lithium-ion transport within the particle, resulting in enhanced rate capability.
- The smaller particle size enhances the electron transport in the electrode, resulting in enhanced rate capability.
- The high surface area leads to enhanced utilization of the active materials, resulting in higher capacity.
- The smaller particle size aids a better accommodation of the strain during lithium insertion/extraction, resulting in improved cycle life.
- The smaller particle size enables reactions that could not otherwise occur with micrometer-sized particles, resulting in new lithium insertion/extraction mechanisms and improved electrochemical properties and performances.

Some of the disadvantages are given below:

- Complexities involved in the synthesis methods employed could increase the processing cost, resulting in higher manufacturing cost.
- The high surface area may lead to enhanced side reactions with the electrolyte, resulting in high irreversible capacity loss and capacity fade during cycling.
- The smaller particle size and high surface-to-volume ratio could lead to low packing density, resulting in low volumetric energy density.

Nanostructured Cathodes

Nanostructured Layered Oxide Cathodes

Oxides with the general formula $LiMO_2$ (M = V, Cr, Co, and Ni) crystallize in a layered structure in which Li^+ and M^{3+} ions occupy the alternate (111) planes of the rock salt structure to give a layered sequence of -O-Li-O-M-O- along the *c* axis. The Li^+ and M^{3+} ions occupy the octahedral interstitial sites of the cubic close-packed oxygen array as shown in Fig. 8.5. The structure with strongly (covalently) bonded MO_2 layers allows a reversible extraction/insertion of lithium ions from/into the lithium planes. The lithium-ion movement between MO_2 layers provides fast two-dimensional lithium-ion diffusion [22], and the edge-shared MO_6 octahedral arrangement with a direct M–M interaction provides good electronic conductivity. $LiCoO_2$ is the most commonly used transition metal oxide cathode in commercial lithium-ion batteries. It has been used because of its high operating voltage (4 V), ease of synthesis, and good cycle life. However, only 50% (~140 mAh g^{-1}) of the theoretical capacity of $LiCoO_2$ can be utilized in practical lithium-ion

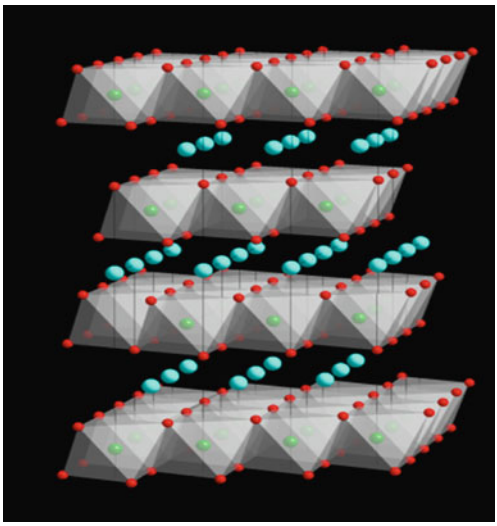


Fig. 8.5 Crystal structure of layered $LiCoO_2$

cells due to structural and chemical instabilities at deep charge with $(1-x) < 0.5$ in $Li_{1-x}CoO_2$ as well as safety concerns [23, 24]. Moreover the element cobalt is toxic and expensive. In this regard, $LiNiO_2$ provides an important advantage compared to $LiCoO_2$ since Ni is less expensive and less toxic than Co. However, it suffers from a few problems: (i) difficulty to synthesize $LiNiO_2$ with all Ni^{3+} and as a perfectly ordered phase without a mixing of Li^+ and Ni^{3+} ions in the lithium plane [25–27], (ii) Jahn–Teller distortion (tetragonal structural distortion) associated with the low-spin Ni^{3+}:d^7 ($t_{2g}^6 e_g^1$) ion [28, 29], and (iii) exothermic release of oxygen at elevated temperatures and safety concerns in the charged state [30, 31]. As a result, $LiNiO_2$ is not a promising material for lithium-ion cells. However, some of these difficulties have been overcome by a partial substitution of Co for Ni forming $LiNi_{1-y}Co_yO_2$ [32]. For example $LiNi_{0.85}Co_{0.15}O_2$ has been found to deliver a high capacity of ~180 mAh g^{-1} with good cyclability.

In general, $LiCoO_2$ and $LiNi_{1-y}Co_yO_2$ are prepared by solid-state reactions at high temperatures (800–900°C) [33]. The high-temperature method usually results in larger particles and often in irregular morphology and a broad particle size distribution. A number of synthetic routes such as sol–gel, co-precipitation, and emulsion methods have been pursued over the years to synthesize nanostructured layered oxides [34–38]. Decreasing the particle size to nanometer range significantly shortens the lithium-ion diffusion distance and improves the rate performance of the layered oxide cathodes. However, the large surface area of the nanoparticles incurs undesirable side reactions with the electrolyte, especially at higher operating voltages. This effect could be minimized by having nano-sized primary particles agglomerating into micron-sized secondary particles.

In this regard, template-assisted methods with the use of porous anodic aluminum oxide have been used to prepare $LiCoO_2$ nanotubes with lengths up to several

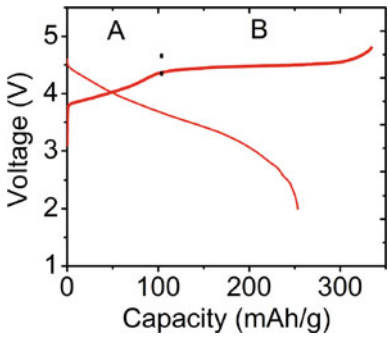


Fig. 8.6 First charge–discharge profiles of solid solutions between layered $Li[Li_{1/3}Mn_{2/3}]O_2$ and $Li[Ni_{1-y-z}Mn_yCo_z]O_2$

micrometers and diameter around 200 nm [39]. Since the rate-determining step in $LiCoO_2$ electrodes is the solid-state lithium-ion diffusion, decreasing the particle size can significantly improve the kinetics because of shorter lithium diffusion length. Accordingly, $LiCoO_2$ with a nanotube architecture has been found to exhibit improved rate performance compared to the conventional $LiCoO_2$ [40].

Recently, solid solutions between $Li[Li_{1/3}Mn_{2/3}]O_2$ (commonly known as Li_2MnO_3) and $LiMO_2$ (M = $Mn_{0.5}Ni_{0.5}$, Co, Ni, and Cr) have become appealing as they exhibit a high reversible capacity of around 250 mAh g^{-1} with a lower cost and better safety compared to the $LiCoO_2$ cathode [41–45]. This capacity is nearly two times higher than that found with the $LiCoO_2$ cathode. However, the $Li_2MnO_3 - LiMO_2$ solid solutions have a few drawbacks. First, they lose oxygen irreversibly from the lattice during first charge as indicated by a voltage plateau at 4.5 V following an initial sloping region corresponding to the oxidation of the transition metal ions to 4+ oxidation state as seen in Fig. 8.6. This leads to a large irreversible capacity loss in the first cycle. Second, these oxides suffer from poor rate capability, impeding their high-power applications.

Decreasing the lithium-ion diffusion distance and having a high surface area nanostructure could enhance their rate performance. Especially, one-dimensional nanowire morphology could be beneficial compared to nanoparticles. Accordingly, $Li[Ni_{0.25}Li_{0.15}Mn_{0.6}]O_2$ nanowires with a diameter of around 30 nm and a high aspect ratio of around 100 have been synthesized via a template-free hydrothermal method. Interestingly, the $Li[Ni_{0.25}Li_{0.15}Mn_{0.6}]O_2$ sample thus prepared has been found to deliver a stable cycle life and superior rate performance with a high discharge capacity of around 260 mAh g^{-1} at 7C rate [46].

Surface-Modified (Nanocoating) Layered Oxide Cathodes

As pointed out in the earlier section, only 50% of the theoretical capacity of layered $LiCoO_2$ cathode could be utilized in practical lithium-ion cells due to the chemical instability in contact with the electrolyte for $(1 - x) < 0.5$ in $Li_{1-x}CoO_2$ [23, 24]. One way to suppress the chemical reactivity is to coat or modify the surface of the cathode with other inert oxides. In fact, surface modification of the layered

$LiCoO_2$ cathode with nanostructured oxides like Al_2O_3, TiO_2, ZrO_2, SiO_2, MgO, ZnO, and MPO_4 (M = Al and Fe) has been found to increase the reversible capacity of $LiCoO_2$ from ~140 mAh g^{-1} to about 200 mAh g^{-1}, which corresponds to a reversible extraction of ~0.7 lithium per formula of $LiCoO_2$ [47–54]. The surface modification suppresses the reaction of the cathode surface with the electrolyte and thereby decreases the impedance growth of the cathode and improves the capacity retention. This clearly demonstrates that the limitation in practical capacity of $LiCoO_2$ is primarily due to the chemical instability at deep discharge and not due to the structural (order–disorder) transition at $(1 - x) = 0.5$. However, the long-term performance of these nano-oxide-coated cathodes will rely on the robustness of the coating. In addition to the improvement in electrochemical properties, nanocoating of $AlPO_4$ on $LiCoO_2$ has also been found to improve the thermal stability and safety of $LiCoO_2$ cathodes [55].

As indicated in the previous section, solid solutions between layered Li_2MnO_3 and $Li[Ni_{1-y-z}Mn_yCo_z]O_2$ exhibit a huge irreversible loss during first cycle. The large irreversible capacity loss is observed to be due to the extraction of lithium as "Li_2O" in the plateau region as shown in Fig. 8.6 and an elimination of the oxygen vacancies formed to give an ideal composition "MO_2" at the end of the first charge, resulting in less number of lithium sites available for lithium insertion/extraction during subsequent cycles [43, 56]. However, a careful analysis of the first charge–discharge capacity values in our laboratory with a number of compositions suggests that part of the oxygen vacancies should be retained in the lattice to account for the high discharge capacity values observed in the first cycle [57]. More importantly, it was found that the irreversible capacity loss in the first cycle can be reduced significantly by coating these layered oxide solid solutions with nanostructured Al_2O_3 and $AlPO_4$ [44, 57]. Figure 8.7 shows the TEM images of $AlPO_4$-coated layered oxide, in which the thickness of the $AlPO_4$ coating is around 5 nm.

Figure 8.8a and b compares the first charge–discharge profiles and the corresponding cyclability data of a series of solid solutions between layered $Li[Li_{1/3}Mn_{2/3}]O_2$ and $Li[Ni_{1/3}Mn_{1/3}Co_{1/3}]O_2$ before and after surface

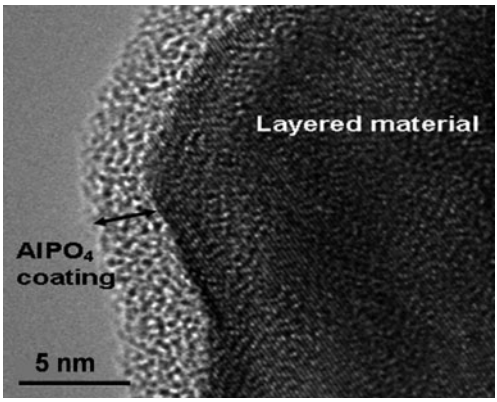


Fig. 8.7 TEM image of 4 wt.% nano-$AlPO_4$-modified $Li[Li_{0.2}Mn_{0.54}Ni_{0.13}Co_{0.13}]O_2$ cathode

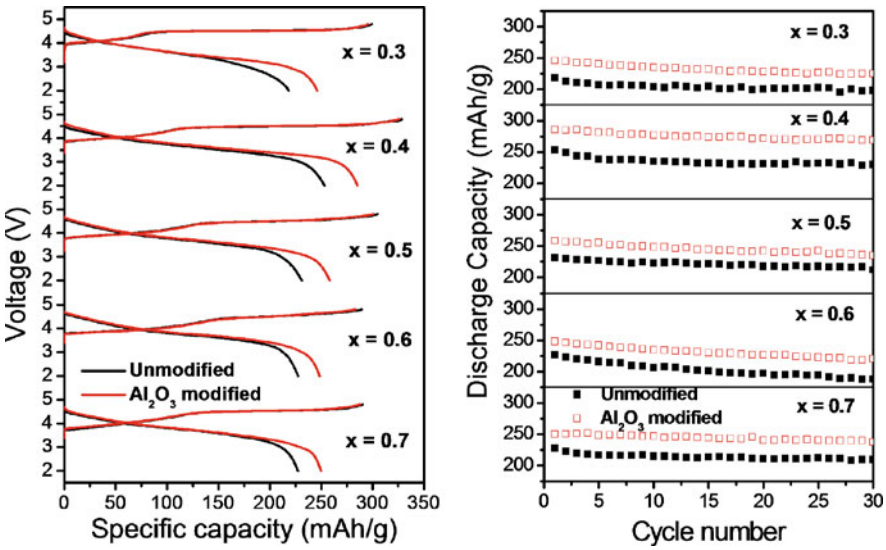


Fig. 8.8 (**a**) First charge–discharge profiles of the layered $(1 - x)\,Li[Li_{1/3}Mn_{2/3}]O_2 - x\,Li[Ni_{1/3}Mn_{1/3}Co_{1/3}]O_2$ solid solutions before and after surface modification with 3wt.% nanostructured Al_2O_3, followed by heating at 400°C (**b**) Cyclability of the layered $(1 - x)\,Li[Li_{1/3}Mn_{2/3}]O_2 - x\,Li[Ni_{1/3}Mn_{1/3}Co_{1.3}]O_2$ solid solutions before and after surface modification with 3 wt.% nanostructured Al_2O_3, followed by heating at 400°C

modification with nanostructured Al_2O_3 [44]. The surface-modified samples exhibit lower irreversible capacity loss and higher discharge capacity values than the pristine layered oxide samples. This improvement in surface-modified samples has been explained on the basis of the retention of more oxygen vacancies in the layered lattice after the first charge compared to that in the unmodified samples [57–59]. The bonding of the nano-oxides to the surface of the layered oxide lattice suppresses the diffusion of oxygen vacancies and their elimination. Remarkably, the surface-modified $(1 - x)Li[Li_{1/3}Mn_{2/3}]O_{2-x}Li[Ni_{1/3}Mn_{1/3}Co_{1/3}]O_2$ composition with $x = 0.4$ exhibits a high discharge capacity of ~280 mAh g^{-1}, which is two times higher than that of $LiCoO_2$. Moreover, the surface-modified cathodes have been found to exhibit higher rate capability than the unmodified samples despite the electronically insulating nature of the coating materials like Al_2O_3 and $AlPO_4$ [59]. This is believed to be due to the suppression of the formation of thick solid-electrolyte interfacial (SEI) layers as the coating material minimizes the direct reaction of the cathode surface with the electrolyte at the high charging voltages.

However, these layered oxide solid solutions have to be charged up to about 4.8 V, so more stable, compatible electrolytes need to be developed to fully exploit their potential as high energy density cathodes. Moreover, oxygen is lost irreversibly from the lattice during first charge, and it may have to be vented appropriately during cell manufacturing. Also, the long-term cyclability of these high-capacity cathodes needs to be fully assessed.

Nanostructured Spinel Oxide Cathodes

Oxides with the general formula $LiMn_2O_4$ (M = Ti, V, and Mn) crystallize in the normal spinel structure (Fig. 8.9) in which the Li^+ and the $M^{3+/4+}$ ions occupy, respectively, the 8a tetrahedral and 16d octahedral sites of the cubic close-packed oxygen array. A strong edge-shared octahedral $[M_2]O_4$ array permits reversible extraction of the Li^+ ions from the tetrahedral sites without collapsing the three-dimensional $[M_2]O_4$ spinel framework. While an edge-shared MO_6 octahedral arrangement with direct M–M interaction provides good electrical conductivity, the interconnected interstitial (lithium) sites in the three-dimensional spinel framework provide good lithium-ion conductivity.

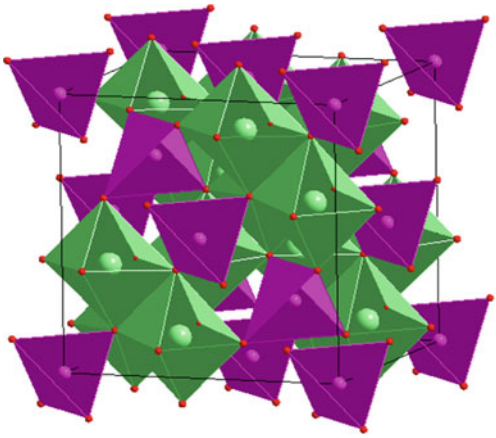


Fig. 8.9 Crystal structure of spinel $LiMn_2O_4$

As a result, spinel $LiMn_2O_4$ has become an attractive cathode. Moreover, the element Mn is inexpensive and environmentally benign compared to Co and Ni involved in the layered oxide cathodes. However, $LiMn_2O_4$ delivers only a limited capacity of around 120 mAh g^{-1} at an operating voltage of 4 V. Moreover, $LiMn_2O_4$ tends to exhibit capacity fade particularly at elevated temperatures (55°C). Several factors such as Jahn–Teller distortion occurring on the surface of the particles under conditions of nonequilibrium cycling [60, 61], manganese dissolution into the electrolyte [62, 63], formation of two cubic phases in the 4 V region, loss of crystallinity [64], and development of micro-strain [65] during cycling have been suggested to be the source of capacity fade. Several strategies have been pursued to overcome the capacity fade of $LiMn_2O_4$. The most significant of them is cationic substitution to give $LiMn_{2-y}M_yO_4$ (M = Li, Cr, Co, Ni, and Cu) to suppress the difficulties of Jahn–Teller distortion and manganese dissolution [66].

Decreasing the particle size to nanometer level can enhance the power performance of $LiMn_2O_4$ cathodes further. As a result, a variety of synthetic approaches such as sol–gel [67], solution phase [68], mechanochemical [69], spray pyrolysis [70], and templating methods [71] have been pursued to synthesize nano-sized

$LiMn_2O_4$. Recently, single-crystalline $LiMn_2O_4$ nanorods obtained by using β-MnO_2 nanorods synthesized by hydrothermal reaction have been shown to exhibit high power performance [72].

However, the application of nanometer-sized spinel particles for practical lithium-ion batteries is not favorable as the high interfacial contact area between the electrode and the electrolyte will aggravate the dissolution of manganese from the spinel lattice into the electrolyte and increase the capacity fade further.

Nano-oxide-Coated Spinel Cathodes

As pointed out in the previous section, the major issue with the $LiMn_2O_4$ spinel cathode is the Mn dissolution from the lattice in contact with the electrolyte. Consequently, coating of the $LiMn_2O_4$ spinel cathode with nanostructured oxides like Al_2O_3, TiO_2, ZrO_2, SiO_2, MgO, and ZnO has been found to suppress the Mn dissolution from the spinel lattice in contact with the electrolyte and improve the capacity retention [73–75].

Another drawback with the spinel $LiMn_2O_4$ cathode is the lower energy density compared to the layered oxide cathodes. In this regard, the $LiMn_{1.5}Ni_{0.5}O_4$ spinel cathode is appealing as it offers a discharge capacity of around 130 mAh g^{-1} at a higher voltage of ~4.7 V vs. lithium. However, the spinel $LiMn_{1.5}Ni_{0.5}O_4$ encounters the formation of NiO impurity during synthesis and the ordering between Mn^{4+} and Ni^{2+} leads to inferior performance compared to the disordered phase [76]. It has been found that the formation of the NiO impurity phase and ordering can be suppressed by appropriate cation doping as in $LiMn_{1.5}Ni_{0.42}Zn_{0.08}O_4$ and $LiMn_{1.42}Ni_{0.42}Co_{0.16}O_4$[77].

One major concern with the spinel $LiMn_{1.5}Ni_{0.5}O_4$ cathode is the chemical stability in contact with the electrolyte at the higher operating voltage of 4.7 V. To overcome this difficulty, surface modification of $LiMn_{1.42}Ni_{0.42}Co_{0.16}O_4$ cathodes with oxides like $AlPO_4$, ZnO, Al_2O_3, Bi_2O_3 have been carried out (Fig. 8.10) [78]. The surface-modified cathodes exhibit better cyclability and rate capability compared to the pristine unmodified samples as shown in Figs. 8.11 and 8.12. The surface coating not only acts as a protection shell between the active cathode material surface and the electrolyte, but also offers fast lithium-ion and electron diffusion channels compared to the SEI layer formed by a reaction of the cathode surface with the electrolyte, resulting in enhanced cycle life and rate performance. X-ray photoelectron spectroscopic (XPS) analysis has shown that the surface modification indeed suppresses the formation of thick SEI layers and thereby improves the rate capability [78]. Moreover, the surface modification helps to maintain the high rate capability as the cathodes are cycled compared to the unmodified cathode, resulting in better rate capability retention during long-term cycling.

Nanostructured Polyanion-Containing Cathodes

A major drawback with cathodes containing highly oxidized redox couples like $Co^{3+/4+}$ and $Ni^{3+/4+}$ is the chemical instability at deep charge and the associated

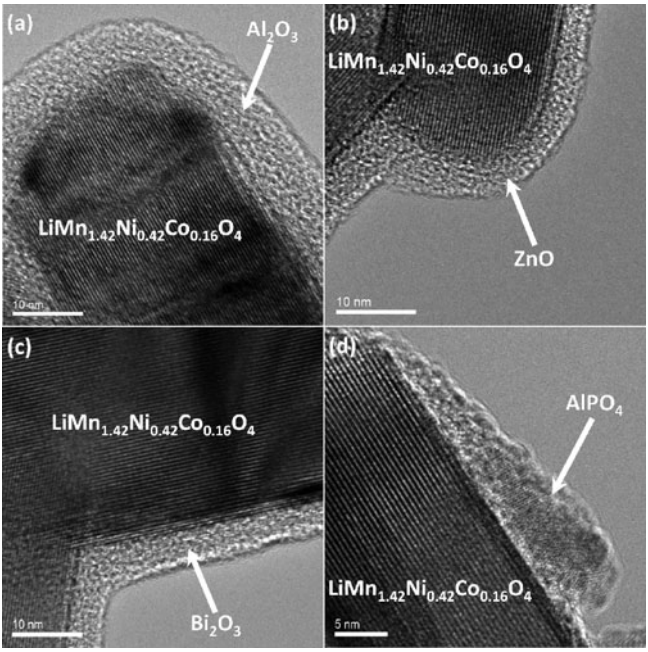


Fig 8.10 High-resolution TEM images of 2 wt.% (**a**) Al_2O_3-, (**b**) ZnO-, (**c**) Bi_2O_3-, and $AlPO_4$-coated $LiMn_{1.42}Ni_{0.42}Co_{0.16}O_4$

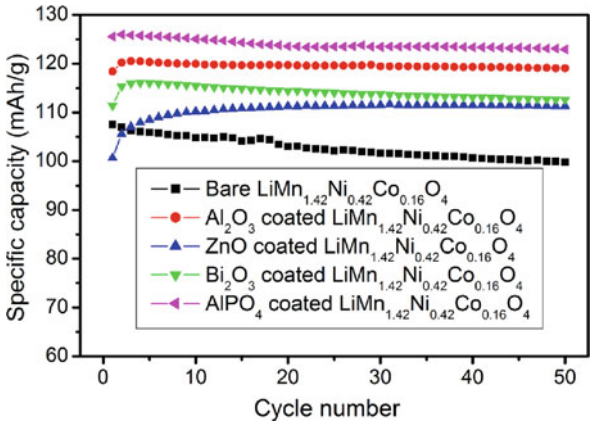


Fig. 8.11 Cycling performances of the bare and 2 wt.% Al_2O_3-, ZnO-, Bi_2O_3-, and $AlPO_4$-coated $LiMn_{1.42}Ni_{0.42}Co_{0.16}O_4$

safety problems. Recognizing this, oxides like $Fe_2(XO_4)_3$ that contain the polyanion $(XO_4)^{2-}$ (X = S, Mo, and W) and crystallizing in the NASICON-related three-dimensional framework structures were shown in the 1980s to exhibit flat discharge voltage profiles at 3.0 or 3.6 V [79, 80]. In these structures, the FeO_6 octahedra share corners with SO_4 tetrahedra with a Fe-O-S-O-Fe linkage and lithium ions

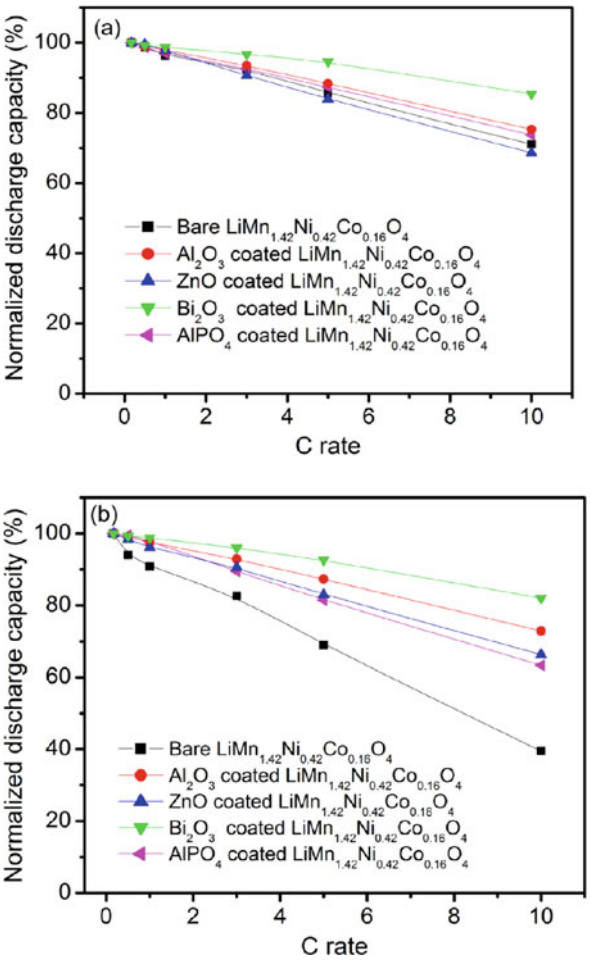


Fig. 8.12 Comparison of the rate capabilities and rate capability retentions of $LiMn_{1.42}Ni_{0.42}Co_{0.16}O_4$ before and after coating with 2 wt.% Al_2O_3, ZnO, Bi_2O_3, and $AlPO_4$: (**a**) normalized discharge capacity at 3rd cycle. (**b**) normalized discharge capacity at 50th cycle

could be inserted into the interstitial voids of the framework. Although the lower valent $Fe^{2+/3+}$ couple in a simple oxide like Fe_2O_3 would be expected to offer a lower discharge voltage of <3 V, a higher voltage of 3.6 V is observed with the $Fe^{2+/3+}$ couple in $Fe_2(SO_4)_3$ due to inductive effect caused by the countercation S^{6+}. A stronger S–O covalent bonding weakens the л-bond Fe–O covalence through inductive effect, which results in a lowering of the $Fe^{2+/3+}$ redox couple and an increase in the cell voltage. However, a poor electronic conductivity associated with the Fe-O-X-O-Fe (X = S or P) linkage leads to poor rate capability.

Nanostructured Phospho-olivine Cathodes

Following the initial concept of using polyanions [78, 79], several phosphates have been investigated in recent years [81–83]. Among them, $LiFePO_4$ crystallizing in

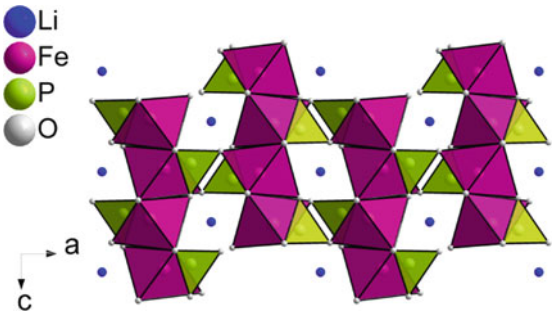


Fig. 8.13 Structure of olivine $LiFePO_4$

the olivine structure (Fig. 8.13) with FeO_6 octahedra and PO_4 tetrahedra has been shown to be a promising material exhibiting a flat discharge voltage of ~3.45 V, with a theoretical capacity of 170 mAh g^{-1}. Unlike in the case of layered $LiMO_2$ (M = Co, Ni, or Mn) oxides, the presence of covalently bonded PO_4 units as well as the operation of $Fe^{2+/3+}$ couple rather than $M^{3+/4+}$ couples leads to good structural and chemical stabilities, resulting in good safety features. Moreover, iron is inexpensive and environmentally benign. However, the initial work was able to extract only <0.7 lithium ions from $LiFePO_4$ even at very low current densities, which corresponds to a reversible capacity of <120 mAh g^{-1}. As the lithium extraction/insertion occurred by a two-phase mechanism with $LiFePO_4$ and $FePO_4$ as end members without much solid solubility, the limitation in capacity was attributed to the diffusion-limited transfer of lithium across the two-phase interface. Thus, the major drawback with $LiFePO_4$ is its poor lithium-ion conductivity resulting from one-dimensional diffusion of Li^+ ions along the chains (*b*-axis) formed by edge-shared LiO_6 octahedra and poor electronic conductivity (~10^{-9} S cm^{-1}).

Tremendous efforts have been made in recent years to overcome these problems by cationic doping, decreasing the particle size through various synthesis methods, and coating with electronically conducting agents [84–89]. Particularly, nano-sized $LiFePO_4$ particles have been shown to exhibit excellent performance with high rate capability due to a shortening of both the electron and lithium-ion diffusion path lengths within the particles. In this regard, dimensionally modulated nanostructures such as nanorods, nanowires, and nanosheets are appealing as they can efficiently transport charge carriers while maintaining a large surface-to-volume ratio, enhancing the contact with the electrolyte and the reaction kinetics.

Among the various synthesis approaches pursued in the past few years, solution-based methods have been particularly successful for $LiFePO_4$ with respect to controlling the chemical composition, tailoring the crystallite size, and particle morphologies. However, these methods require either long reaction times (5–24 h) or further post-heat treatment processing at temperatures as high as 700°C in reducing atmospheres to achieve phase pure samples and a high degree of crystallinity. In this regard, microwave-assisted synthesis approaches are extremely appealing as they can shorten the reaction time from several hours to a few minutes with enormous energy savings and cleanliness. Consequently, our group has demonstrated

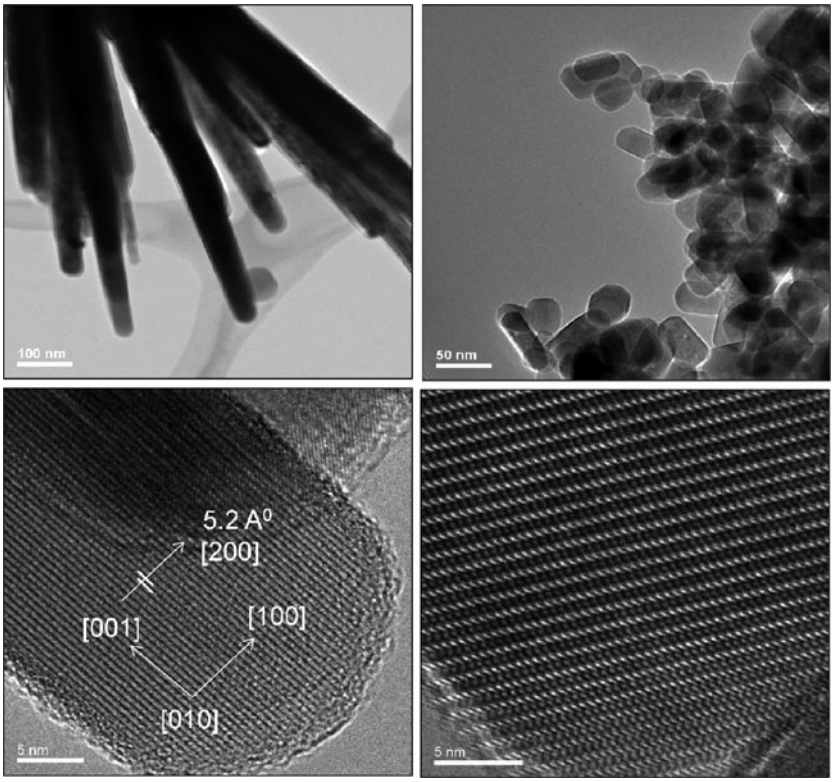


Fig. 8.14 TEM images of (**a**) large $LiFePO_4$ nanorods, (**b**) small $LiFePO_4$ nanorods, (**c**) and (**d**) high-resolution TEM images of the $LiFePO_4$ nanorods

the synthesis of $LiFePO_4$ with nanorod morphology by a microwave-solvothermal (MW-ST) method within 15 min at temperatures as low as 300°C [90–92]. The TEM images shown in Fig. 8.14 reveal well-defined crystalline nanorod morphology with controlled size. The width and length of the nanorods depend on the synthesis conditions (e.g., reactant concentration), which could help to tune the rate capability and volumetric energy density.

The MW-ST method has also been extended to synthesize $LiMnPO_4$, $LiCoPO_4$, and $LiNiPO_4$, which provides an opportunity to realize higher energy density because of the higher operating voltages of, respectively, 4.1, 4.8, and 5.1 V [93]. In addition, from the diffraction information provided by the FFTs, we could further confirm that the long axis and the width of the nanorods correspond, respectively, to the *c*-crystallographic axis [001] and the *a*-crystallographic axis [100], while the *b*-crystallographic axis of the orthorhombic olivine structure (lithium diffusion direction) is parallel to the electron beam direction [010]. Recent computational models and experimental methods on $LiMPO_4$ have shown that in the orthorhombic olivine structure, the lowest Li^+ migration energy is found for the pathway

along the [010] channel, indicating one-dimensional lithium-ion mobility along the *b*-axis during the charge–discharge process [94–96]. Consequently, there is enormous interest in synthesizing nanostructured phospho-olivines with the *b*-axis along the shortest dimension of the crystallites. The dimensionally modulated $LiMPO_4$ nanorods offered by the MW-ST approach thus exhibits a unique and favorable morphology since the *b*-axis is one of the two short dimensions of the nanostructures, resulting in a shorter lithium-ion diffusion path length.

While decreasing the particle size to nanometer level has been able to reduce the diffusion length of lithium ions and overcome the lithium-ion transport limitations in $LiFePO_4$, the pristine $LiFePO_4$ nanorods still suffer from poor electronic conductivity. In this context, electronically conducting interconnects with better wetting properties will be appealing. For example, addition of conducting polymers and nanonetworking with multiwalled carbon nanotubes (Figs. 8.15 and 8.16) have been found to offer significantly improved electrochemical performances [90, 92].

Figure 8.17 compares the discharge capacities at various C-rates for the long and short $LiFePO_4$ nanorods before and after networking with MWCNT. With both the pristine and MWCNT-networked samples, the shorter $LiFePO_4$ nanorods exhibit

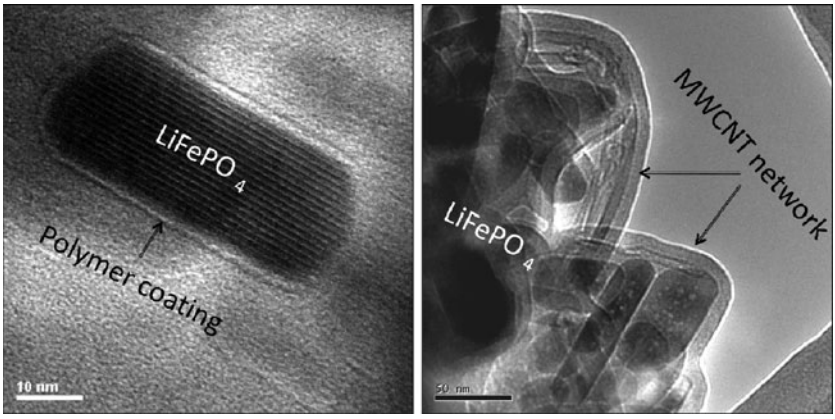


Fig. 8.15 TEM image of (**a**) $LiFePO_4$-PEDOT and (**b**) $LiFePO_4$-MWCNT nanocomposites

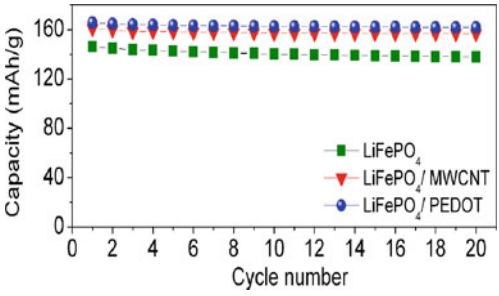


Fig. 8.16 Cyclability of pristine $LiFePO_4$ prepared by the MW-ST method, after networking it with MWCNT, and after encapsulating it with *p*-TSA-doped PEDOT

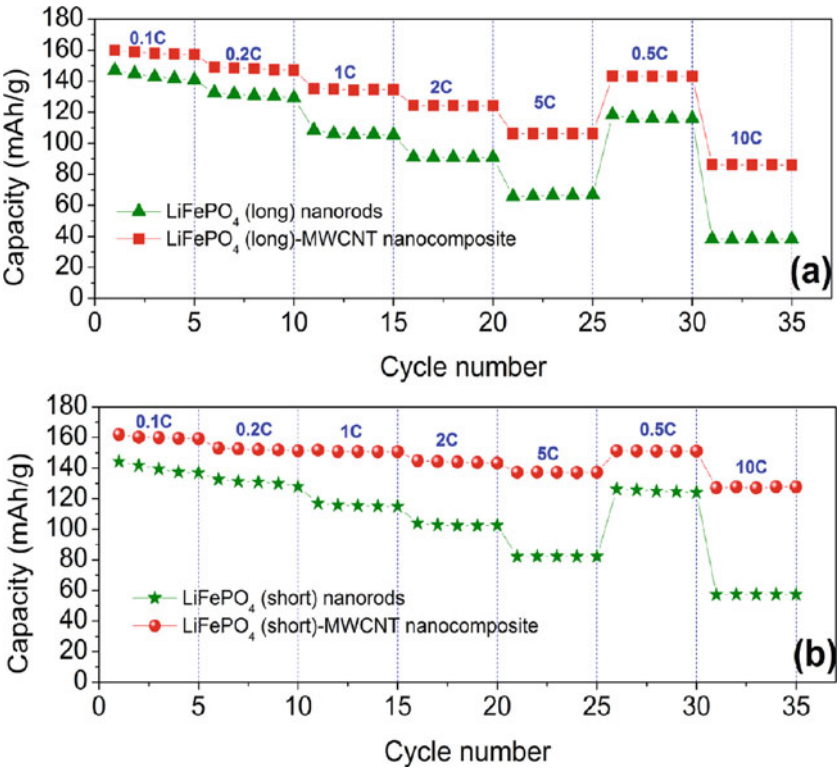


Fig. 8.17 Cyclability of the (**a**) long $LiFePO_4$ nanorods and (**b**) short $LiFePO_4$ nanorods at different discharge rates from C/10 to 10C before and after networking with MWCNT. The charging rate was kept constant at C/10 for all the samples

higher discharge capacity at a given C-rate than the long nanorods due to a faster lithium-ion diffusion arising from a shorter diffusion length. With both the long and short nanorods, the $LiFePO_4$-MWCNT nanocomposites exhibit higher capacity at a given C-rate than the pristine $LiFePO_4$ due to the enhancement in electronic conductivity. At 10C rate, the $LiFePO_4$-MWCNT nanocomposite with the short $LiFePO_4$ nanorods shows a discharge capacity of close to 130 mAh g^{-1} while that with the long $LiFePO_4$ nanorods exhibits a discharge capacity of only 90 mAh g^{-1}. This size–property correlation clearly showcases the beneficial effect of decreasing particle size on the electrochemical rate performance of $LiFePO_4$. The superior rate performance of the $LiFePO_4$-MWCNT nanocomposite is due to the synergistic effect of small lithium diffusion path length in the short nanorods and the highly conductive matrix provided by the carbon nanotubes.

Although the initial work by Goodenough's group [83] revealed a two-phase reaction mechanism with $LiFePO_4$ and $FePO_4$ as end members, subsequent investigations have indicated several interesting observations. For example, the miscibility gap between the two phases has been found to decrease with increasing temperature,

and the occurrence of a single-phase solid solution Li_xFePO_4 with $0 \leq x \leq 1$ has been reported at 450°C. Similarly, the miscibility gap has been found to decrease with decreasing particle size, and the complete solid solubility between $LiFePO_4$ and $FePO_4$ at room temperature has been reported for 40 nm size particles. Thus, what was originally found to be a two-phase reaction mechanism with micrometer-sized particles has turned into a single-phase reaction mechanism with nano-sized particles. This is a clear demonstration of how nanoparticles can behave entirely different from micrometer-sized counterparts. Defect chemistry with the existence of cationic vacancies in the samples prepared by the low-temperature approaches has been suggested to be partly the reason for the contrasting behavior of the nano-sized particles.

$LiMnPO_4$ is of particular interest to the battery community because of the ideal location of the $Mn^{2+/3+}$ couple at 4.1 V vs. Li/Li^+, which is compatible with the presently available commercial electrolytes. However, due to the extremely low intrinsic electronic conductivity (~10^{-14} S cm^{-1}) compared to that of $LiFePO_4$ (~10^{-9} S cm^{-1}) [97, 98], $LiMnPO_4$ exhibits inferior electrochemical performance. Optimizing the synthesis process and carbon coating have recently shown promising electrochemical performances for $LiMnPO_4$ nanoparticles [99, 100]. Realization of near theoretical capacity in materials like $LiCoPO_4$ that has a higher operating voltage of ~4.8 V can enhance the energy density significantly, appealing for next-generation lithium-ion cells. However, $LiCoPO_4$ offers a lower discharge capacity of ~120 mAh g^{-1} [93], and the lower capacity value and the capacity fade on cycling could be due to the lack of compatible electrolytes to operate at the high voltage of 4.8 V, and development of more stable electrolyte compositions has the potential to improve the performance of $LiCoPO_4$ further. With a theoretical voltage of 5.1 V vs. Li/Li^+, $LiNiPO_4$ poses even a tougher challenge on the electrolyte oxidation issue, so we were not able to carry out the electrochemical tests on the synthesized $LiNiPO_4$ nanorods with the available conventional electrolytes (1 M $LiPF_6$ in 1:1 diethyl carbonate/ethylene carbonate).

Other Polyanion-Containing Cathodes

Polyanion-containing compounds other than $LiMPO_4$ such as $Li_3M_2(PO_4)_3$ (M = V, Fe, or Ti), $LiVPO_4F$, and $LiVOPO_4$ have also attracted a great deal of interest in recent years because of their high thermal stability and attractive electrochemical properties [101–105]. The main drawbacks of these materials are their poor electronic conductivity and the consequent slow reaction kinetics. So various synthetic routes such as sol–gel and hydrothermal methods have been employed to prepare nanostructured materials, and coating with electronically conductive agents has been carried out to improve their electrochemical performance.

While only one lithium per formula unit could be reversibly extracted/inserted from/into the $LiMPO_4$ hosts, a new class of cathode materials based on silicates such as Li_2MSiO_4 (M = Mn, Fe, Ni, and Co) offers the possibility of reversibly extracting/inserting two lithium per formula unit and two times higher capacity than $LiFePO_4$. Various synthetic routes such as hydrothermal and mechanochemical

methods have been employed to synthesize nanoparticles of Li_2MSiO_4 [106, 107]. However, although close to two lithium per formula unit could be extracted from Li_2MnSiO_4, it shows poor cycle life. On the other hand, only one lithium per formula could be extracted from Li_2FeSiO_4.

Nanostructured Anodes

Nanostructured Carbon-Based Anodes

Carbon-based anodes are attractive because of their abundance, high chemical stability, and wide electrochemical potential windows [108]. Graphite with a low atomic weight, redox energy close to that of lithium, and a high theoretical capacity of ~372 mAh g^{-1} is the most commonly used anode in commercial lithium-ion cells. The lithium insertion/extraction reaction can be written as

$$6C + Li^+ + e^- <==> LiC_6. \tag{8.2}$$

The insertion of lithium at very low potentials (<100 mV) results in electrolyte reduction with subsequent formation of solid-electrolyte interfacial (SEI) layer at the surface of graphite anode during first charge. The SEI layer prevents further reaction of graphite anode during cycling and also inhibits exfoliation of graphite layers. However, the low operating potential may potentially plate lithium on the graphite anodes, resulting in serious safety issues and poor cycle performance of the cell. SEI layer formation also causes irreversible capacity loss during first cycle. The fraction of lithium irreversibly trapped requires an additional mass of cathode materials.

Anodes based on hard carbon are also used to overcome the graphite exfoliation problem during lithium insertion/extraction. The hard carbons are typically obtained by thermal decomposition of phenolic and epoxy resins and products from petroleum pitch. The hard carbons exhibit a high irreversible capacity and a sloping voltage profile between 0 and 1 V unlike graphite, which shows a nearly flat discharge profile between 0 and 0.3 V. One important advantage of hard carbon anode compared to the graphite anode is that it works better with spinel manganese cathodes. Most of the commercial cells employing spinel manganese oxide cathode use hard carbon anodes.

Carbon Nanotubes and Graphene Nanosheets

Carbon nanotubes and graphene nanosheets having a high electro-active surface area, high conductivity, and good mechanical properties are attractive anode candidates for lithium-ion batteries. The nanostructured architecture along with their high surface area will be particularly beneficial for achieving high power performance.

Recently, graphene nanosheets (GNS) have been shown to exhibit high capacity values [109–112]. Graphene nanosheets are single layer of carbon atoms tightly packed into two-dimensional (2D) honeycomb sp^2 carbon lattice. Figure 8.18

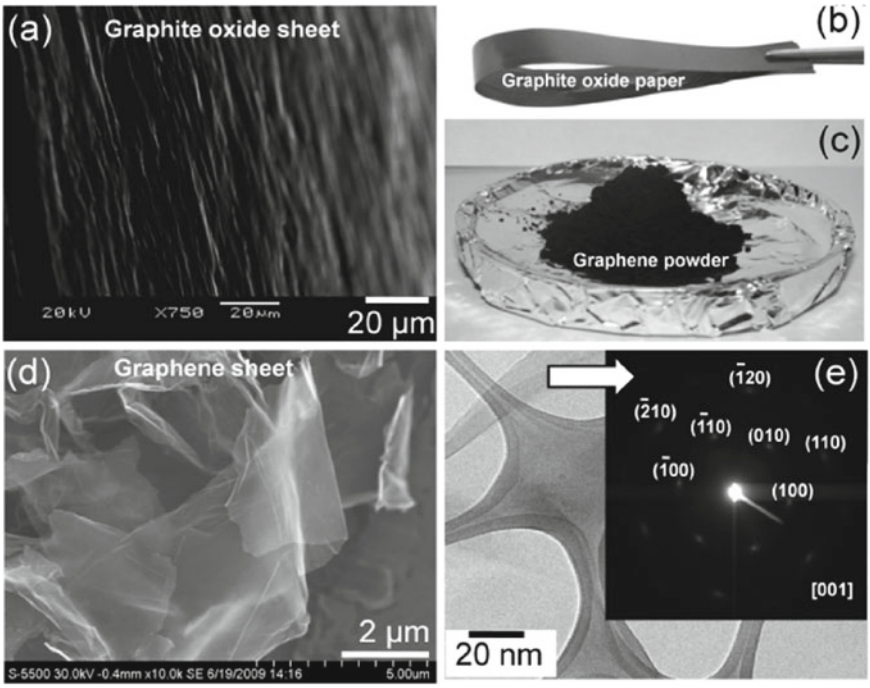


Fig. 8.18 SEM image of the cross-section of the GO paper, (**b**) photographic image of GO paper, (**c**) bulk quantity of grapheme powder produced by the MW-ST process, (**d**) ultrahigh resolution FE-SEM image of large (~15 × 5 μm size), single paper-like GNS, and (**e**) bright field TEM image of optically transparent GNS, with the *inset* showing the electron diffraction pattern of GNS

shows the SEM and TEM images of graphene sheets synthesized by a microwave-solvothermal (MW-ST) reduction of graphite oxide (GO) in our laboratory. Figure 8.19 compares the first charge–discharge profiles at C/20 rate of commercial graphite and GNS synthesized by the MW-ST process. The GNS offers much higher reversible capacity (~820 mAh g^{-1}) than the graphite anode (~340 mAh g^{-1}). The high capacity can be attributed to the absorption of lithium on both sides of GNS, resulting in two layers of lithium per layer of graphene sheet to give Li_2C_6, and nanoporous carbons have been shown to operate with similar mechanism [113, 114]. This is quite different from a conventional graphite anode, which involves the insertion of one layer of lithium per carbon layer to give LiC_6. However, GNS exhibits a huge irreversible capacity loss (685 mAh g^{-1}) in the first cycle due to electrolyte decomposition and formation of SEI layer. Also, the GNS exhibits some capacity fade. Nevertheless, the cyclability of graphene has been improved by decorating GNS with low cost, conducting polymer polyaniline (PANI). The GNS/PANI nanocomposite has shown reduced irreversible capacity loss with a significant improvement in capacity retention due to the modification of the surface and pore structure by conducting polymer [112].

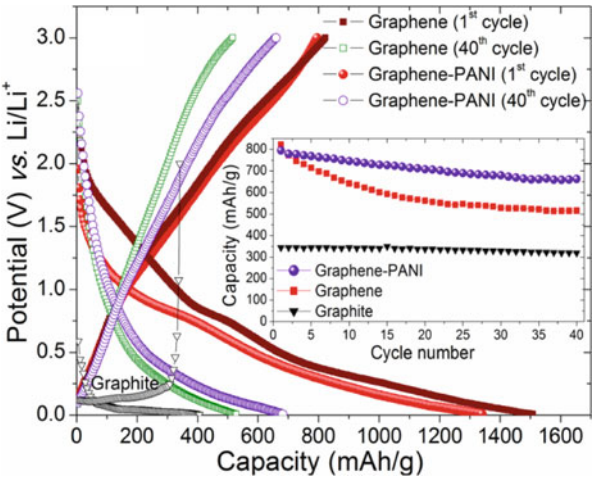


Fig. 8.19 Comparison of the charge–discharge profiles and cyclability (*inset*) in lithium cells of commercial graphite, GNS, and grapheme-PANI nanocomposite at a constant discharge rate of C/15

Multiwalled carbon nanotubes have been found to deliver capacities up to 800 mAh g^{-1} [109]. In addition, flexible, freestanding carbon nanotube papers obtained by CVD process have also been demonstrated to exhibit a high capacity of 572 mAh g^{-1} and are attractive candidates for application in flexible batteries [110]. Accommodation of extra lithium into nanometer-sized cavities of CNTs and storage of additional lithium at the edges and surfaces are the main reason for the high capacity of these nanoarchitectures.

Although carbon nanotubes and graphene exhibit two times higher lithium storage capacity than graphite, the sloping voltage profile during the charge–discharge is undesirable. Moreover, these materials invariably suffer from large SEI layer formation and high irreversible capacity loss because of their high surface area. Also, a huge voltage difference occurs between the lithium insertion and extraction processes (hysteresis loss), resulting in a decrease in the efficiency of the energy storage reaction.

Nanostructured Titanium Oxide Anodes

Titanium oxides have attracted a lot of attention as an alternative, low-cost, nontoxic anode material, with a discharge–charge voltage well above the lithium plating region. In particular, $Li_4Ti_5O_{12}$ with the spinel structure exhibits excellent lithium insertion/extraction properties with little volume change (zero strain material), which is desirable for long-term cycling [115–117]. The three-dimensional network of channels for lithium-ion diffusion in the spinel lattice makes it attractive for high-rate applications. It can be cycled over the composition range $0 \leq x \leq 3$ in

$Li_{4+x}Ti_5O_{12}$ with a flat charge–discharge voltage of about 1.5 V vs. Li^+/Li. The higher operating voltage eliminates the SEI layer formation and lithium plating problems, rendering the material much safer than graphite. Also, it is not poisoned by dissolved manganese released by the $LiMn_2O_4$ cathode. In addition, the $Ti^{3+/4+}$: t_{2g} band lies well above the top of the O^{2-}:2p band and this makes it chemically stable as well.

Significant efforts have been devoted to synthesize nanoparticles of $Li_4Ti_5O_{12}$, which facilitate facile lithium-ion transport due to shorter lithium diffusion distance. For example, $Li_4Ti_5O_{12}$ nanowires synthesized from lithium acetate and TiO_2 nanowires have been shown to offer superior rate performance and cycle life [118]. However, the main drawback with $Li_4Ti_5O_{12}$ is that it exhibits a much lower capacity of around 175 mAh g^{-1} at a much higher voltage of 1.5 V vs. Li^+/Li, resulting in a significant reduction in energy density of the lithium-ion cells. Recently, nanocoating of carbon on $Li_4Ti_5O_{12}$ has also been pursued to improve its rate performance [119].

TiO_2 is also an attractive anode because of the ease of synthesis and its commercial availability. Although not much lithium could be inserted into the conventional micron-sized particles of TiO_2, decreasing the particle size to the nanometer scale increases the degree of lithium insertion. Accordingly, various nanostructures such as nanoparticles, nanowires, and nanotubes of different forms of TiO_2 (rutile, anatase, and TiO_2(B)) have been studied to improve the capacity and rate performance [120–123]. The lithium insertion/extraction into/from TiO_2 occurs at around 1.7 V with <4% volume change and a theoretical capacity of ~335 mAh g^{-1} as

$$TiO_2 + x\,Li^+ + x\,e^- <==> Li_xTiO_2 (0 \leq x \geq 1). \tag{8.3}$$

The low volume change offers good structural stability and long-term cycling performance.

Nanostructured Alloy Anodes

Metals or nonmetals that undergo alloying reactions and store lithium are among the most appealing anode candidates for lithium-ion batteries as they offer much higher capacities than carbon. Lithium is known to form alloys (Li_xM) with various metals and nonmetals such as Mg, Al, Si, Ge, Sn, Sb, Bi, and Zn [124–127]. When polarized to a sufficient negative potential, these metals or nonmetals undergo reversible alloying reaction with lithium as follows:

$$xLi^+ + xe^- + M <==> Li_xM. \tag{8.4}$$

The theoretical capacities of some alloy anodes are 10 times higher than that of the commercially used graphite anodes. Another important feature of alloying reaction is that it operates at a higher operating potential (0.3–1 V) than graphite anodes. This will eliminate or minimize the safety hazards associated with the SEI layer

formation and lithium plating. Also, they do not suffer from solvent co-intercalation, unlike the graphite anode.

Unfortunately, the accommodation of such a large amount of lithium results in a huge volume expansion–contraction (100–300%) during the discharge–charge process. The huge volume change leads to rapid deterioration (cracking) and eventual pulverization of the active materials, resulting in a failure of the alloy anodes after a few charge–discharge cycles. Significant research effort has been devoted to overcome this problem, and one of the most attractive approaches is to reduce the particle size into the nanometer range.

Cycling performance of the alloy anodes could be improved significantly by employing an active–inactive nanocomposite strategy. While the active components undergo lithium storage reaction and expand/contract, the inactive components act as a "matrix" buffering the volume expansion. However, for successful application of this concept, the following criteria should be satisfied: (i) the active materials should be in the nanometer size and finely dispersed in the matrix to avoid cracking during volume expansion and (ii) the inactive matrix should support a facile transport of lithium ions and electrons.

Nanostructured Tin- and Antimony-Based Alloys

Tin (Sn) and antimony (Sb) alloys are particularly attractive because they offer high theoretical capacities of, respectively, 990 mAh g^{-1} ($Li_{4.4}Sn$) and 660 mAh g^{-1} (Li_3Sb) with an operating voltage well above that of metallic lithium. However, the reaction of Sn and Sb with lithium is accompanied by a large volume change of around 150%, which results in cracking of the electrode and capacity fading. To overcome this problem nanocrystalline intermetallic compounds with electrochemically inactive and active elements such as SnCo, SnNi, SnSb, Cu_2Sb, CoSb, CrSb, InSb, and Zn_4Sb_3 have been studied [128–133]. These nanocrystalline intermetallics have shown considerable improvement in cycle life over pure metals, and the nano-sized powders lead to enhanced reaction kinetics.

The alloy anodes have been synthesized by different synthetic routes such as sol–gel, electrodeposition, and ball-milling. High-energy ball-milling process is advantageous because of its ease of operation and ability to scale up. Recently, our group has synthesized Sb-MO_x-C (M = Al, Ti, or Mo) nanocomposites by a mechanochemical reduction of Sb_2O_3 with Al, Ti, or Mo in the presence of acetylene black [134]. Figure 8.20 shows the STEM images of the M = Al and Ti samples. The reduced antimony, MO_x, and carbon are homogenously dispersed in the nanocomposite. HR-TEM and FFT images (Fig. 8.20) confirm the presence of crystalline antimony particles of 15–25 nm, well surrounded by the amorphous oxide MO_x and carbon phases. The MO_x-C ceramic matrix acts as a buffer to alleviate the volume expansion. The Sb-MO_x-C nanocomposites exhibit charge capacities of around 500 mAh g^{-1} with an average operating voltage of 1.0 V. Figure 8.21 compares the cyclability of the three Sb-MO_x-C nanocomposites at a constant current density of 100 mAh g^{-1}. While pure antimony exhibits continuous capacity fade, the nanocomposites exhibit good cyclability.

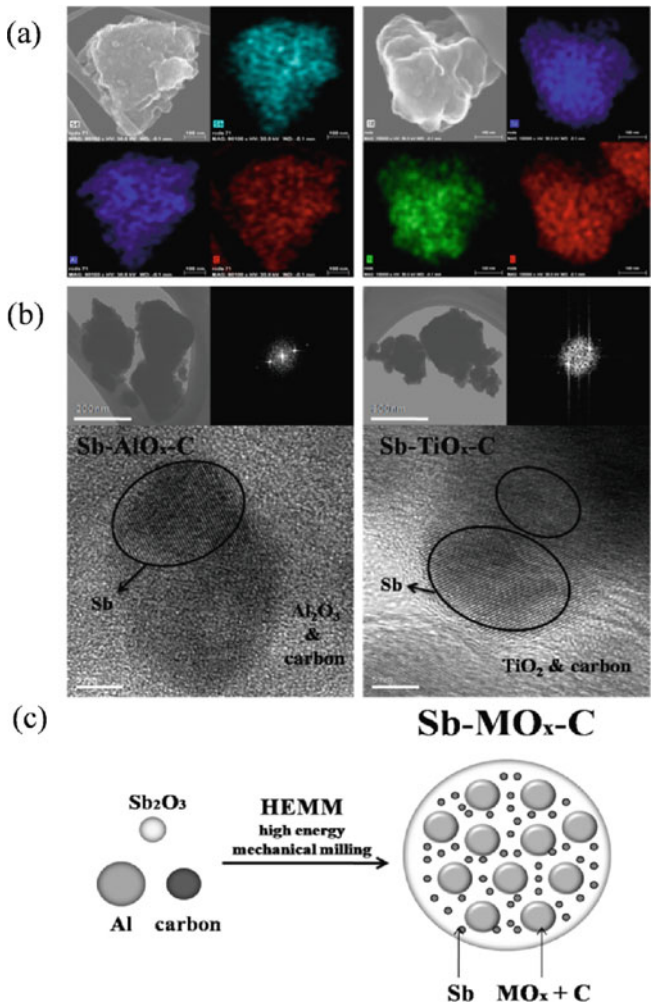


Fig. 8.20 (**a**) STEM images, (**b**) HRTEM images, and (c) schematic description of the Sb-Mo$_x$-C (M = Al and Ti) nanocomposites. Also, shown are the fast Fourier transform images over selected regions

Materials based on SnO_2 deliver a high reversible capacity of 782 mAh g^{-1}. During first discharge, SnO_2 reacts with lithium to form Li_2O and Sn, which then undergoes reversible reaction with lithium to form $Li_{4.4}Sn$. However, the practical application of SnO_2 is hindered by poor cyclability arising from a large volume change during the electrochemical reaction. Extensive work has been focused on improving the electrochemical performance of tin oxide by designing different nanostructures such as nanowires, nanotubes, mesoporous structures, ordered nanostructure/carbon composites, and nanoparticles [135–137].

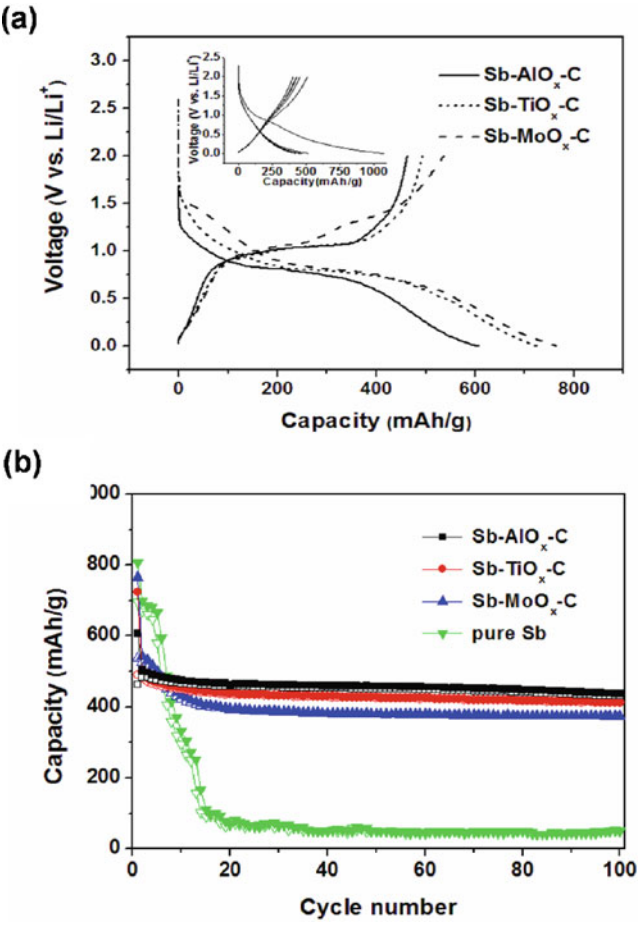


Fig. 8.21 Comparison of the (**a**) discharge–charge profiles and (**b**) cycling performances of the Sb-Mo_x-C (M = Al, Ti, and Mo) nanocomposites

Silicon Nanowires

Silicon is an attractive anode material for lithium-ion batteries because of its low discharge potential and high theoretical capacity of ~4,200 mAh g^{-1}, which is 10 times higher than that of the currently used graphite anodes. However, silicon suffers from a huge volume change of up to 400% during the reaction with lithium. As a result, bulk films and micrometer-sized silicon particles have been shown to suffer from severe capacity fade during cycling [138–140].

Recognizing that the huge volume expansion causes pulverization and loss of contact between the active materials and current collectors, a nanoarchitecture approach has been pursued to overcome this problem. Accordingly, silicon nanowires have been grown directly on stainless steel substrate using a

vapor–liquid–solid growth method [141]. The main advantages of this nanowire architecture are as follows: (i) the small nanowire diameter allows for better accommodation of large volume change without fracture during cycling, (ii) each silicon nanowire is electrically connected to current collector so that all nanowires take part in the electrochemical reaction, and (iii) the nanowires have direct one-dimensional electron pathways resulting in efficient charge transport. The silicon nanowires thus obtained show a first charge and discharge capacity of 4,277 mAh g^{-1} and 3,124 mAh g^{-1} at C/20 rate in lithium cells. They have also shown good cycle life and high rate performance. The improved capacity and cycle life of silicon nanowire architecture compared to the micron-sized or nano-sized particles demonstrate the advantage of this specific nanowire design. Similar strategy has also been found to show improved performance with germanium nanowire anodes with a high capacity of ~1,140 mAh g^{-1} [142]. Yet another approach that has been pursued is nanocomposites such as Si/C, Si/TiB_2, or Si/Sn [143], which have shown improved electrochemical performances compared to Si.

Nanostructured Metal Oxide Anodes

Electrodes based on classical intercalation/de-intercalation are limited in capacity because of the limited number of lithium ions that can be accommodated within the insertion host. It has been shown that lithium can react with a range of transition metal oxides, sulfides, nitrides, and phosphides by a conversion reaction, forming composites consisting of nanometal particles and Li_mX (X = O, S, N, or P) matrix [144–148]. The reaction with lithium can be written as

$$M_xX_y + ne^- + nLi^+ <==> xM^0 + yLi_mX\,(m = n/y). \tag{8.5}$$

Such reactions show good reversibility, providing almost four times higher capacity than graphite anodes. Moreover, the possibility of using low-cost transition metals such as iron and manganese make them attractive. Furthermore, the conversion reaction provides an opportunity to control the redox potential by tuning the electronegativity of the anion.

Among the transition metal oxides undergoing conversion reaction with lithium, magnetite (Fe_3O_4) is attractive because of its high capacity, eco-friendliness, natural abundance of Fe, and high electronic conductivity. However, its application in practical lithium-ion batteries is hindered due to its poor reaction kinetics and poor cycling stability resulting from volume expansion occurring during cycling. In this context, employing nanowire architectures could enhance the cycling performance. Moreover, the one-dimensional nanowire morphology is beneficial for achieving high rate performance since it facilitates better electron and lithium-ion transport than electrodes comprising small nanoparticles in which the electrons and lithium ions have to move through particles and are limited by the interparticle contacts [149–151].

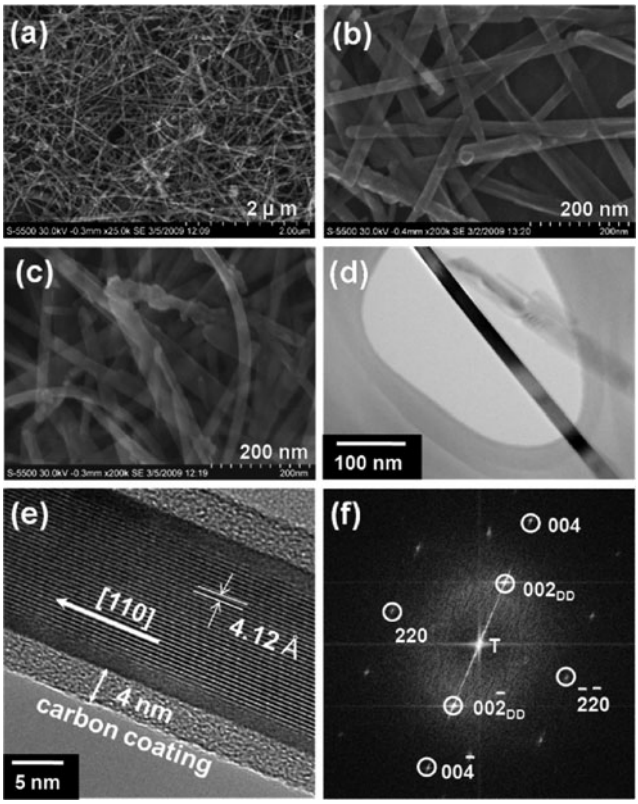


Fig. 8.22 Morphological characterization of Fe_3O_4 nanowires: (**a**) low- and (**b**) high-magnification FE-SEM images of as-synthesized Fe_3O_4 nanowires, (**c**) FE-SEM image of carbon-decorated Fe_3O_4 nanowires, (**d**) TEM and (**e**) HRTEM images showing crystalline Fe_3O_4 nanowires surrounded by amorphous carbon, and (**f**) FFT image of single-crystaline Fe_3O_4 nanowire. The subscript DD refers to double diffraction

Our group has demonstrated the synthesis of carbon-decorated, single crystalline magnetite nanowires by a microwave-hydrothermal approach, employing PEG-400 as a soft template (Fig. 8.22) [152]. Figure 8.23 shows the first charge–discharge profiles of as-synthesized and carbon-decorated Fe_3O_4 nanowires. The as-synthesized Fe_3O_4 nanowires exhibit a reversible capacity of ~922 mAh g^{-1}, corresponding to the extraction/insertion of eight Li^+ ions. However, as seen in Fig. 8.23, the as-synthesized sample exhibits gradual capacity fade during cycling, retaining only 50% of their initial capacity after 50 cycles.

Interestingly, Fe_3O_4 nanowires after decorating with carbon give a capacity of ~830 mAh g^{-1} without much capacity loss during 50 cycles. The enhanced performance of the carbon-coated sample is due to the presence of the carbon buffer layer around the Fe_3O_4 nanowires, preventing direct contact among adjacent nanowires and thereby minimizing aggregation of nanowires during cycling.

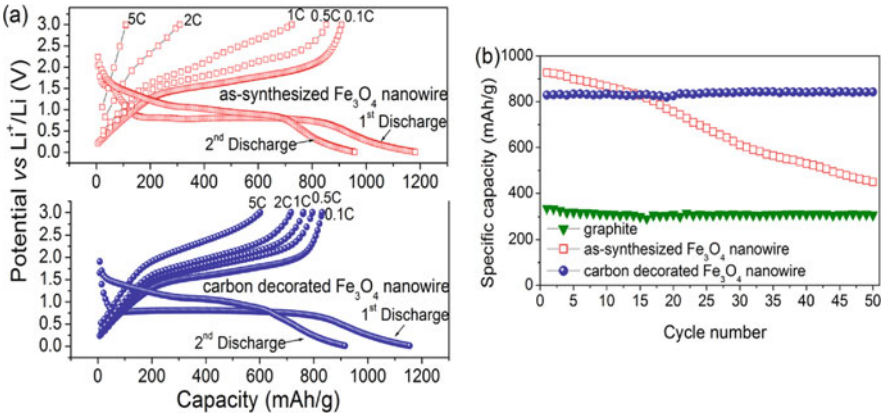


Fig. 8.23 (**a**) Galvanostatic charge–discharge curves and (**b**) cycling performances of the Fe_3O_4 nanowires before and after carbon decoration. For a comparison, data for natural graphite electrode is also shown

The carbon nanocoating also provides an elastic, inactive buffer matrix that can absorb massive volume expansion and contraction during the charge–discharge process. More importantly, the carbon-decorated Fe_3O_4 nanowires exhibit improved rate performance compared to the as-synthesized Fe_3O_4 nanowires. For example, as seen in Fig. 8.23, the carbon-decorated Fe_3O_4 nanowires can still deliver a high capacity of around 600 mAh g^{-1} at a high charge rate of 5C. In other words, 73% of the initial capacity could be retained at a high rate of 5C. In contrast, only 12% of the initial capacity could be retained with the as-synthesized Fe_3O_4 nanowires at such a high charge rate of 5C. Nanowires of Co_3O_4 synthesized by several methods have also been shown to offer high discharge capacities with stable cycle life. Major drawbacks with the conversion reaction anodes are their poor kinetics and a large difference between the charge and discharge voltages.

Conclusions

This chapter presented an overview of the use of nanostructured materials as cathode and anode materials for lithium-ion batteries. They offer the important advantages of high capacity and rate capability, arising from a large surface-to-volume ratio and short lithium ion and electron diffusion path lengths. In addition, they provide a facile accommodation of volume expansion–contraction occurring during cycling. The smaller particles also often lead to new reaction pathways with lithium and formation and stabilization of kinetically stable phases. However, the larger surface-to-volume ratio can aggravate the reaction with the electrolyte and increase the solid-electrolyte interfacial layer formation when the charge–discharge voltage is >4.3 or <1 V vs. Li/Li^+. Thus, the nanostructure is potentially advantageous only with certain electrode materials as has been demonstrated, for example,

with $LiFePO_4$ (3.45 V) cathode and $Li_4Ti_5O_{12}$ (1.5 V) anode. The smaller particle size also leads to low tap densities, resulting in a reduced volumetric energy density. Furthermore, the nanomaterials could increase the processing and manufacturing cost in some cases. Nevertheless, nanomaterials and nanoarchitectures have the potential to offer breakthrough electrode materials for future energy storage devices.

Acknowledgment Financial support by the Welch Foundation grant F-1254 and by the Office of Vehicle Technologies of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 is gratefully acknowledged.

References

1. Lithium Battery Energy Storage (LIBES) Publications, Technological Research Association, Tokyo (1994)
2. Aydinol, M.K., Ceder, G.J.: J. Electrochem. Soc. **144**, 3832 (1997)
3. Whittingham, M.S.: Science **192**, 1126 (1976)
4. Whittingham, M.S., Jacobson, A.J.: Intercalation Chemistry. Academic, New York, NY (1982)
5. Mizushima, K., Jones, P.C., Wiesman, P.J., Goodenough, J.B.: Mater. Res. Bull. **15**, 783 (1980)
6. Goodenough, J.B., Mizushima, K., Takeda, T.: Jap. J. Appl. Phys. **19**, 305 (1983)
7. Thackeray, M.M., David, F.W.I., Bruce, P.G., Goodenough, J.B.: Mat. Res. Bull. **18**, 461 (1983)
8. Megahed, S., Scrosati, B.: Interface **4**(4), 34 (1995)
9. Takehare, Z., Kanamura, K.: Electrochim. Acta **38**, 1169 (1993)
10. Sun, Y., Xia, Y.: Science **298**, 2176 (2002)
11. Rolison, D.R.: Science **299**, 1698 (2003)
12. Duan, X., Niu, C., Sahi, V., Chen, J., Parce, J.W., Empedocles, S., Goldman, J.L.: Nature **425**, 274 (2003)
13. Xia, Y., Yang, P., Sun, Y., Wu, Y., Mayers, B., Gates, B., Yin, Y., Kim, F., Yan, H.: Adv. Mater. **15**, 353 (2003)
14. Li, Y., Qian, F., Xiang, J., Lieber, C.M.: Mater. Today **9**, 18 (2006)
15. Shui, J., Li, M. J.C.: Nano Lett. **9**, 1307 (2009)
16. Arico, A.S., Bruce, P., Scrosati, B., Tarascon, J.–M., Shalkwijk, W.V.: Nat. Mater. **4**, 366 (2005)
17. Bruce, P.G., Scrosati, B., Tarascon, J.-M.: Angew. Chem. Int. Ed. **47**, 2930 (2008)
18. Manthiram, A., Vadivel Murugan, A., Sarkar, A., Muraliganth, T.: Energy Environ. Sci. **1**, 621 (2008)
19. Wang, Y., Cao, G.: Adv. Mater. **20**, 2251 (2008)
20. Chen, J., Cheng, F.: Acc. Chem. Res. **42**, 173 (2009)
21. Bazito, C.F.F., Torresi, R.M., J. Braz. Chem. Soc. **17**, 627 (2006)
22. Whittingham, M.S.: Chem. Rev. **104**, 4271 (2004)
23. Chebiam, R.V., Prado, F., Manthiram, A.: Chem. Mater. **13**, 2951 (2001)
24. Venkatraman, S., Shin, Y., Manthiram, A.: Electrochem. Solid State Lett. **6**, A9 (2003)
25. Dutta, G., Manthiram, A., Goodenough, J.B.: J. Solid State Chem. **96**, 123 (1992)
26. Hirano, A., Kanno, R., Kawamoto, Y., Takeda, Y., Yamamura, K., Takeda, M., Ohyama, K., Ohashi, M., Yamaguchi, Y.: Solid State Ionics **78**, 123 (1995)
27. Kanno, R., Kubo, H., Kawamoto, Y., Kamiyama, T., Izumi, F., Takano, Y.: J. Solid State Chem. **140**, 145 (1998)
28. Nakai, I., Nagagome, T.: Electrochem. Solid State Lett. **1**, 259 (1998)

29. Nakai, I., Takahashi, K., Shiraishi, Y., Nakagome, T., Nishikawa, F.: J. Solid State Chem. **140**, 145 (1998)
30. Dahn, J.R., Fuller, E.W., Obrovac, M., Von Sacken, U.: Solid State Ionics **69**, 265 (1994)
31. Zhang, Z., Fouchard, D., Rea, J.R.: J. Power Sources **70**, A49 (2001)
32. Li, W., Curie, J.,:J. Electrochem. Soc. **144**, 1773 (1997)
33. Liu, H., Wu, Y.P., Rahm, E., Holze, R., Wu, H.Q.: J. Solid State Chem. **8**, 450 (2004)
34. Treuil, N., Labrugere, C., Menetrier, M., Portier, J., Campet, G., Deshayes, A., Frison, J.-C., Hwang, S.-J., Song, S.-W., Choi, J.-H.: J. Phys. Chem. B **103**, 1200 (1999)
35. Im, D., Manthiram, A.: J. Electrochem. Soc. **149**, A1001 (2002)
36. Lu, C.H., Wang, H.C.: J. Mater. Chem. **13**, 428 (2003)
37. Liu, Z., Wang, A.-L., Liu, X., Wu, M., Li, D., Zeng, Z.: J. Solid State Chem. **177**, 1585 (2004)
38. Hong, Y.-S., Park, Y.J., Ryu, K.S., Chang, S.H., Kim, M.G.: J. Mater. Chem. **14**, 1424 (2004)
39. Cheng, F., Tao, Z., Liang, J., Chen, J.: Chem. Mater. **20**, 667 (2008)
40. Li, X., Cheng, F., Guo, B., Chen, J.: J. Phys. Chem. B **109**, 14017 (2005)
41. Numata, K., Sakaki, C., Yamanaka, S.: Solid State Ionics **117**, 257 (1999)
42. Lu, Z., Beaulieu, L.Y., Donaberger, R.A., Thomas, C.L., Dahn, J.R.: J. Electrochem. Soc. **149**, A778 (2002)
43. Armstrong, A.R., Holzapfel, M., Novak, P., Johnson, C.S., Kang, S., Thackeray, M.M., Bruce, P.G.: J. Am. Chem. Soc. **128**, 8694 (2006)
44. Wu, Y., Manthiram, A., Electrochem. Solid State Lett. **9**, A221 (2006)
45. Arunkumar, T.A., Wu, Y., Manthiram, A.: Chem. Mater. **19**, 3067 (2007)
46. Kim, M.G., Jo, M., Hong, Y.-S., Cho, J.: Chem. Commun. **218** (2009)
47. Kim, J., Noh, M., Cho, J., Kim, H.: J. Electrochem. Soc. **152**, A1142 (2005)
48. Kim, B., Kim, C., Kim, T., Ahn, D., Park, B.: J. Electrochem. Soc. **153**, A1773 (2006)
49. Kim, Y.J., Cho, J., Kim, T.-J. Park, B.: J. Electrochem. Soc. **150**, A1723 (2003)
50. Liu, L., Wang, Z., Li, H., Chen, L., Huang, X.: Solid State Ionics **152**, 341 (2002)
51. Cho, J., Kim, C.-S., Yoo, S.-I.: Electrochem. Solid State Lett. **3**, 362 (2000)
52. Kannan, A.M., Rabenberg, L., Manthiram, A.: Electrochem. Solid State Lett. **6**, A16 (2003)
53. Wang, Z., Huang, X., Chen, L.: J. Electrochem. Soc. **150**, A199 (2003)
54. Fang, T., Duh, J., Sheen, S.: J. Electrochem. Soc. **152**, A1701 (2005)
55. Cho, J., Kim, Y.-W., Kim, B., Lee, J.-G., Park, B.: Angew. Chem. Int. Ed. **42**, 1618 (2003)
56. Thackeray, M.M., Kang, S.-H., Johnson, C.S., Vaughey, J.T., Benedek, R., Hackney, S.A.: J. Mater. Chem. **17**, 3112 (2007)
57. Y. Wu, Murugan, A.V., Manthiram, A.: J. Electrochem. Soc. **155**, A635 (2008)
58. Wu, Y., Manthiram, A.: Solid State Ionics **180**, 50 (2009)
59. Wang, Q.Y., Liu, J., Murugan, A.V., Manthiram, A.: J. Mat. Chem. **19**, 4965 (2009)
60. Ohzuku, T., Kitagawa, M., Hirai, T.: J. Electrochem. Soc. **137**, 769 (1990)
61. Thackeray, M.M., Y. Shao-Horn, Kahaian, A.J., Kepler, K.D., Skinner, E., Vaughey, J.T., Hackney, S.A.: Electrochem. Solid State Lett. **1**, 7 (1998)
62. Jang, D.H., Shin, Y.J., Oh, S.M.: J. Electrochem. Soc. **143**, 2204 (1996)
63. Inoue, T., Sano, M., J. Electrochem. Soc. **145**, 3704 (1998)
64. Xia, Y., Yoshio, M.: J. Electrochem. Soc. **143**, 825 (1996)
65. Kannan, A.M., Manthiram, A.: Electrochem. Solid State Lett. **5**, A167 (2002)
66. Shin, Y., Manthiram, A.: Electrochem. Solid State Lett. **6**, A34 (2002)
67. Curtis, C.J., Wang, J., Schulz, D.L.: J. Electrochem. Soc. **151**, A590 (2004)
68. Nieto, S., Majumder, S.B., Katiyar, R.S.: J. Power Sources **136**, 88 (2004)
69. Choi, H.J., Lee, K.M., Kim, G.H., Lee, J.G.: J. Am. Ceram. Soc. **84**, 242 (2001)
70. Taniguchi, I., Lim, C.K., Song, D., Wakihara, M.: Solid State Ionics, **146**, 239 (2002)
71. Nishizawa, M., Mukai, K., Kuwabata, S., Martin, C.R., Yoneyama, H.: J. Electrochem. Soc. **144**, 1923 (1997)
72. Hosono, E., Kudo, T., Honma, I., Matsuda, H., Zhou, H.: Nano Lett. **9**(3), 1045 (2009)
73. Sun, Y., Hong, K., Prakash, J.: J. Electrochem. Soc. **150**, A970 (2003)

74. Han, J., Myung, S., Sun, Y.: J. Electrochem. Soc. **153**, A1290 (2006)
75. Thackeray, M.M., Johnson, C.S., Kim, J.S., Lauzze, K.C., Vaughey, J.T., Dietz, N., Abraham, D., Hackney, S.A., Zeltner, W., Anderson, M.A.: Electrochem. Commun. **5**, 752, (2003)
76. Kim, J.-H., Myung, S.-T., Yoon, C.S., Oh, I.-H., Suna, Y.-K.: J. Electrochem. Soc. **151**, A1911 (2004)
77. Arun Kumar, T.A., Manthiram, A.: Electrochem. Solid State Lett. **8**, A403 (2005)
78. Jun, L., Manthiram, A.: Chem. Mater. **21**, 1695 (2009)
79. Manthiram, A., Goodenough, J.B.: J. Power Sources **26**, 403 (1989)
80. Manthiram, A., Goodenough, J.B.: J. Solid State Chem. **71**, 349 (1987)
81. Padhi, A.K., Nanjundaswamy, K.S., Masquelier, C., Okada, S., Goodenough, J.B.: J. Electrochem. Soc. **144**, 1609 (1997)
82. Padhi, A.K., Nanjundaswamy, K.S., Masquelier, C., Goodenough, J.B.: J. Electrochem. Soc. **144**, 2581 (1997)
83. Padhi, A.K., Nanjundaswamy, K.S., Goodenough, J.B.: J. Electrochem. Soc. **144**, 1188 (1997)
84. Chung, S.-Y., Bloking, J.T., Chiang, Y.-M.: Nat. Mater. **1**, 123 (2002)
85. Wang, Y., Wang, J., Yang, J., Nuli, Y.: Adv. Funct. Mater. **16**, 2135 (2006)
86. Dominko, R., Gaberscek, M., Drofenik, J., Bele, M., Pejovnik, S., Jamnik, J.: J. Power Sources **119**, 770 (2003)
87. Ellis, B., Ka, W.H., Makahnou, M. W.R., Nazar, L.F.: J. Mater. Chem. **17**, 3248 (2007)
88. Wang, C.S., Hong, J.: Electrochem. Solid State Lett. **10**, A65 (2003)
89. Doeff, M.M., Hu, Y., McLarnon, F., Kostecki, R.: Electrochem. Solid State Lett. **6**, A207 (2003)
90. Vadivel Murugan, A., Muraliganth, T., Manthiram, A.: Electrochem. Commun. **10**, 903 (2008)
91. Vadivel Murugan, A., Muraliganth, T., Manthiram, A.: J. Phys. Chem C **112**, 14665 (2008)
92. Muraliganth, T., Vadivel Murugan, A., Manthiram, A.: J. Mater. Chem. **18**, 5661 (2008)
93. Vadivel Murugan, A., Muraliganth, T., Manthiram, A., Inorg. Chem. **48**, 946 (2009)
94. Islam, M.S., Driscoll, D.J., C.A. Fisher, J., Slater, P.R.: Chem. Mater. **17**, 5085 (2005)
95. Morgan, D., Ven der ven, A., Ceder, G.: Electrochem. Solid State Lett. **7**, A30 (2007)
96. Nishimura, S.-I., Kobayashi, G., Ohoyama, K., Kanno, R., Yashima, M., Yamada, A.: Nat. Mater. **7**, 707 (2008)
97. Delacourt, C., Laffont, L., Bouchet, R., Wurn, C., Leriche, J.-B., Morcrette, M., Tarascon, J.-M., Masquelier, C.J.: J. Electrochem. Soc. **152**, A913 (2005)
98. Sauvage, F., Baudrin, E., Gengembre, L., Tarascon, J.-M.: Solid State Ionics **176**, 1869 (2005)
99. Delacourt, C., Poizot, P., Morcrette, M., Tarascon, J.-M., Masquelier, C.: Chem. Mater. **16**, 93 (2009)
100. Kwon, N.-H., Drezen, T., Exnar, I., Teerlinck, I., Isono, M., Graetzel, M.: Electrochem. Solid State Lett. **9**, A277 (2006)
101. Saidi, M.Y., Barker, J., Huang, H., Sowyer, J.L., Adamson, G.: J. Power Sources **119**, 266 (2003)
102. Yin, S.C., Grond, H., Srobel, P., Huang, H., Nazar, L.F.: J. Am. Chem. Soc. **125**, 326 (2003)
103. Hung, H., Yin, S.C., Kerr, T., Taylor, N., Nazar, L.F.: Adv. Mater. **14**, 1525 (2002)
104. Barker, J., Saidi, M.Y., Swoyer, J.L.: J. Electrochem. Soc. **150**, A1394 (2003)
105. Ren, M.M., Zhou, Z., Gao, X.P., Liu, L., Peng, W.X.: J. Phys. Chem. C **112**, 13043 (2008)
106. Dominko, R., Bele, M., M. Gaberšček, Meden, A., Remskar, M., Jamnik, J.: Electrochem. Commun. **8**, 217 (2006)
107. Gong, Z.L., Li, Y.X., He, G.N., Li, J., Yang, Y.: Electrochem. Solid State Lett. **11**(5), A60 (2008)
108. Imanishi, N., Takeda, Y., Yamamoto, O.: Lithium Ion Batteries: Fundamentals and Performance, p. 98. Wiley-VCH, Weinheim (1998)

109. Frackowiak, E., Bēguin, F.: Carbon **40**, 1775 (2002)
110. Chen, J., Minett, A.I., Liu, Y., Lynam, C., Sherrell, P., Wang, C., Wallace, G.C.: Adv. Mater. **20**, 566 (2008)
111. Yoo, E.J., Kim, J., Hosono, E., Zhou, H.-S., Kudo, T., Honma, I.: Nano Lett. **8**, 2277 (2008)
112. Vadivel Murugan, A., Muraliganth, T., Manthiram, A.: Chem. Mater. **21**, 5004 (2010)
113. Yata, S., Kinoshita, H., Komori, M., Ando, N., Kashiwamura, T., Harada, T., Tanaka, K., Yamabe, T.: Synth. Met. **62**, 153 (1994)
114. Xue, J.S., Dahn, J.R.: J. Electrochem. Soc. **142**, 3668 (1995)
115. Ferg, E., Gummow, R.J., Dekock, A., Thackeray, M.M.: J. Electrochem. Soc. **141**, L147 (1994)
116. Jansen, A.N., Kahalan, A.J., Kepler, K.D., Nelson, P.A., Amine, K., Dees, D.W., Vissers, D.R., Thackeray, M.M.: J. Power Sources **81**, 902 (1999)
117. Reale, P., Panero, S., Scrosati, B., Garche, J., WohlfahrtMehrens, M., Wachtler, M.: J. Electrochem. Soc. **151**, A2138 (2004)
118. Kim, J., Cho, J.: Electrochem. Solid State Lett. **10**(3), A81 (2007)
119. Wang, C.J., Gao, J., Fu, J.I., Zhao, N.H., Wu, Y.P., Takamura, T.: J. Power Sources **174**, 1109 (2007)
120. Deng, D., Kim, M.G., Leem, J.Y., Cho, J.: Energy Environ. Sci. **2**, 818 (2009)
121. Wei, M., Konishi, Y., Zhou, H., Sugihara, H., Arakawa, H.: Sold State Commun. **133**, 493 (2005)
122. Bavykin, D.V., Friedrich, J.M., Walsh, F.C.: Adv. Mater. **18**, 1 (2006)
123. Xu, J., Jia, C., Cao, B., Zhang, W.F.: Electrochim. Acta **52**, 8044 (2007)
124. Johnson, C.E., Fosterm, M.S.: J. Electrochem. Soc. **116**, 1612 (1969)
125. Lai, S.: J. Electrochem. Soc. **123**, 1196 (1976)
126. Muller, W., Schafer, H., Naturforsch Z., 28b, 246 (1973)
127. Weppner, W., Huggins, R.A.: J. Electrochem. Soc. **124**, 1569 (1977)
128. Winter, M., Besenhard, J.O.: Electrochim. Acta **45**, 31 (1999)
129. Fransson, L., Vaughey, J.T., Benedek, R., Edstrom, K., Thomas, J.O., Thackeray, M.M.: Electrochem. Commun. **3**, 317 (2001)
130. Alcantara, R., Fernandez-Madrigal, F.J., Lavela, P., Tirado, J.L., Jumasb, J.C., Oliver-Fourcade, J.: J. Mater. Chem. **9**, 2517 (1999)
131. Yang, J., Takeda, Y., Imanish, N., Yamamoto, O.J.: J. Electrochem. Soc. **146**, 4009 (1999)
132. Li, H., Shi, L., Lu, W., Huang, X., Chen, L.: J. Electrochem. Soc. **148**, A91 (2001)
133. Hewitt, K.C., Beaulieu, L.Y., Dahn, J.R.: J. Electrochem. Soc. **148**, A402 (2001)
134. Yoon, S., Manthiram, A.: Chem. Mater. **21**, 3898 (2009)
135. Meduri, P., Pendyala, C., Kumar, V., Sumanasekera, G.U., Sunkara, M.K.: Nano Lett. **9**, 612 (2009)
136. Chen, G., Wang, Z., Xia, D.: Chem. Mater. **20**, 6951 (2008)
137. Lou, X.W., Li, C.M., Archer, L.A.: Adv. Mater. **21**, 1 (2009)
138. Kasavajjula, U., Wang, C., Appleby, A.J.: J. Power Sources, **163**, 1003 (2007)
139. Green, M., Fielder, E., Scrosati, B., Wachtler, M., Moreno, J.S.: Electrochem. Solid State Lett. **6**, A75 (2003)
140. Ryu, J.H., Kim, J.W., Sung, Y.-E., Oh, S.M.: Electrochem. Solid State Lett. **7**, A306 (2004)
141. Chan, C.K., Peng, H., Liu, G., McIlwrath, K., Zhang, X.F., Huggins, R.A., Cui, Y.: Nat. Nanotech. **3**, 31 (2008)
142. Chan, C.K., Zhang, X.F., Cui, Y.: Nano Lett. **8**, 307 (2008)
143. Kim, H., Cho, J.: Nano Lett. **8**(11), 3688 (2008)
144. Poizot, P., Laruelle, S., Grugeon, S., Dupont, L., Tarascon, J.-M.: C.R. Acad. Sci. Ser. IIc **3**, 681 (2000)
145. Pereira, N., Dupont, L., Tarascon, J.-M., Klein, L.C., Amatucci, G.G.: J. Electrochem. Soc. **150**, A1273 (2003)
146. Li, H., Ritcher, G., Maier, J.: Adv. Mater. **15**, 736 (2003)

147. Silva, D.C.C., Crosneir, O., Ouvrard, G., Greeedan, J., Safasefat, A., Nazar, L.: Electrochem. Solid State Lett. **6**, A162 (2003)
148. Pralong, V., Souza, D.C.S., Leung, K.T., Nazar, L.: Electrochem. Commun. **4**, 516 (2002)
149. Tarascon, P.L., Mitra, S., Poizot, P., Simon, P., Tarascon, J.-M.: Nature Mater. **5**, 567 (2006)
150. Nam, K.T., Kim, D.-W., Yoo, P.J., Chiang, C.-Y., Meethong, N., Hammond, P.T., Chiang, Y.-M., Belcher, A.M.: Science **312**, 885 (2006)
151. Li, Y., Tan, B., Wu, Y.: Nano Lett. **8**, 265 (2008)
152. Muraliganth, T., Vadivel Murugan, A., Manthiram, A.: Chem. Commun. 7360 (2010)

Chapter 9
Synthetic Models of Copper Proteins for Biofuel Cell Applications

Dominic F. Gervasio

Abstract Recently there has been a renaissance in biological science and engineering. Powerful new technologies have emerged whose development crosses the classic borders between biology, chemistry, and engineering disciplines.

A number of practical applications have resulted, including sensors, drug dosing and making, combinatorial diagnostics, and artificial organs. Using such devices in vivo is highly desired, and this is motivating the development of a means for generating electrical power that is compatible with biological environments. For in vivo applications, providing electrical power by using a fuel cell is especially attractive, because a fuel cell could indefinitely supply power to an implanted device by scavenging reactants from the sera in living organisms. Attractive as such "biofuel cells" may be, serious challenges remain. One challenge is finding catalysts to activate electrode reactions for power generation from chemical species found in sera, while maintaining stability in complex physiological fluids. Perhaps the most important of these reactant species is O_2, because it is usually available and is apt to be required for the cathode in a biofuel cell.

Introduction

Recently there has been a renaissance in biological science and engineering. Powerful new technologies have emerged whose development crosses the classic borders between biology, chemistry, and engineering disciplines. A number of practical applications have resulted, including sensors, drug dosing and making, combinatorial diagnostics, and artificial organs. Using such devices in vivo is highly desired, and this is motivating the development of a means for generating electrical power that is compatible with biological environments. For in vivo applications,

D.F. Gervasio (✉)
Department of Chemical and Environmental Engineering, University of Arizona, Tucson, AZ 85721, USA
e-mail: Don.Gervasio@asu.edu

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9_9,

providing electrical power by using a fuel cell is especially attractive, because a fuel cell could indefinitely supply power to an implanted device by scavenging reactants from the sera in living organisms. Attractive as such "biofuel cells" may be, serious challenges remain. One challenge is finding catalysts to activate electrode reactions for power generation from chemical species found in sera, while maintaining stability in complex physiological fluids. Perhaps the most important of these reactant species is O_2, because it is usually available and is apt to be required for the cathode in a biofuel cell.

Iron-based hemoglobin, the heme center itself, and heme-like molecules are often considered the most relevant catalysts for promoting oxygen chemistry in biological systems. The heme-like iron-porphyrins, phthalocyanins, and related compounds have been extensively investigated as air cathode catalysts. The intention of this review is to highlight synthetic analogues of another class of enzymes as catalytically active for oxygen chemistry as hemoglobin, namely the copper proteins, such as the laccases, oxidases, tyrosinase, and hemocyanin. These Cu enzymes and their synthetic analogues have promise and unique advantages as catalysts for making biofuel cell air cathodes.

The Two-Copper-Ion Fragment and Oxygen

A structural unit often found in copper proteins that catalyzes reactions of molecular oxygen is an antiferromagnetically coupled, two-copper-ion fragment. This two-copper-ion fragment has been associated with the active sites of enzymes when they perform many remarkable phenomena, such as (1) the reversible binding of molecular oxygen, O_2, during oxygen transport (e.g., by the enzyme called hemocyanin); (2) the catalysis of the electron transfers which lead to the oxidation of small molecules and the reduction of O_2 to H_2O (e.g., by the blue oxidases, such as ascorbate oxidase, ceruloplasmin, and the laccases); and (3) the catalytic insertion of an oxygen atom, O, into organic molecules, in which the source of O is O_2 (e.g., by tyrosinase). Simple copper ions in aqueous solution are not nearly so active as catalysts for these reactions as are the copper proteins. The identity and geometry of the ligating residues, which the protein offers to the two copper ions to form such active Cu-based sites in the enzymes, have long been under intense investigation by many research groups.

This review is intended to summarize the salient features of the binuclear copper active sites and to describe approaches for synthesizing new copper coordination complexes. The goal of the syntheses is to make small copper-containing molecules, the so-called "Cu enzymes mimics," which model the structure and spectroscopy and perhaps most importantly effectively catalyze reactions of oxygen with small molecules in the same way these reactions are catalyzed by the binuclear copper active sites found in nature. The natural binuclear copper active sites have the copper ions coordinated with atoms from moieties from a protein strand. Oxygen can come

from phenol, nitrogen from imidazole, etc. A plausible approach to biomimicry is to effectively replicate the active sites found in natural systems by making copper complexes with the same donor atoms and coordination to the copper as in the enzymes. A copper donor atom may originate from the same moiety found in the protein, such as N from imidazole, but an atom coordinated to copper could just as well come from non-natural synthetic moieties. For example, one useful synthetic moiety is pyrazole, an isomer of imidazole. Pyrazole is an unsaturated, "soft," chelating group, which stabilizes lower oxidation states [28] leading to high redox potentials. Neutral pyrazole is a unidentate ligand. The pyrazole anion (also called the pyrazolate anion) can act as a bridging unit between two copper ions. Binuclear copper-ion sites in enzymes have been structurally characterized by X-ray diffraction, and the resulting structures show that the inter-copper separation is typically 3–4 Å. Similarly, X-ray diffraction studies show the bridging pyrazolate group promotes Cu–Cu separations of 3.6–3.9 Å.

Proteins with Coupled Binuclear Copper Centers

Copper ions are found in the active sites of numerous enzymes and proteins [60, 34]. Spectroscopic and magnetic studies of native proteins containing binuclear copper centers and their derivatives suggest that there are only two significantly distinct types of binuclear copper sites, namely, the Type 3 active site and the active site found in hemocyanin or tyrosinase [72]. Proposed active site structures and how they interact with O_2 are shown in Fig. 9.1.

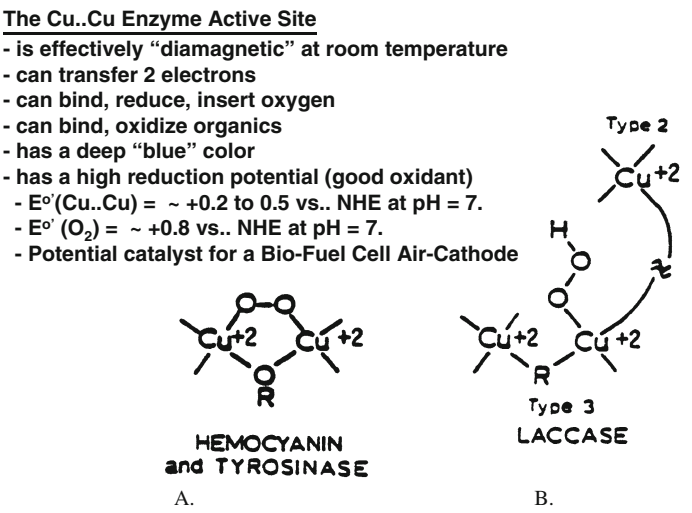


Fig. 9.1 Proposed Active Site Structures for the Cu_2–O_2 Adducts in A Hemocyanin (or Tyrosinase) and B. the Type 3 centers

Hemocyanin and Tyrosinase

Hemocyanin [58] and tyrosinase [94] cycle between the cupric, Cu(II), and cuprous, Cu(I), oxidation states during biological functions. The binuclear cupric ions are so strongly antiferromagnetically coupled that the Cu(II)–Cu(II) site in these proteins is effectively diamagnetic under anaerobic conditions at room temperature [42]. Figure 9.2 summarizes the electronic and magnetic qualities of such a Cu(II)–Cu(II) site [41, 85]. The set of 10 molecular orbitals for the Cu dimer is shown in Fig. 9.2a(i). These molecular orbitals are the in and out of phase combinations of the primarily d-atomic orbitals from the two monomeric copper-ion fragments which make up the dimer. The eight lowest energy dimer orbitals are filled by 16 of the 18 electrons available in the Cu(II)–Cu(II) oxidation state. Figure 9.2a (ii) shows the possible electronic configurations for the filling of the remaining

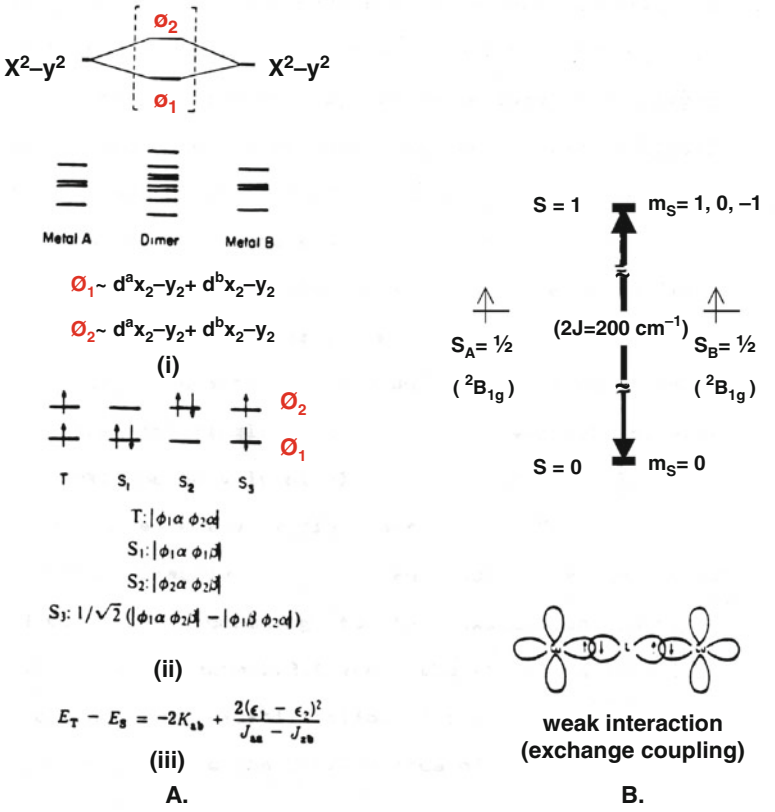


Fig. 9.2 Electronic Description of the Antiferromagnetically Coupled Cu(II)–Cu(II) Center. A. (i) The 10 d-like Dimer Molecular Orbitals; (*ii*) Possible Electronic Configurations; (*iii*) expression for $E_T - E_S(2J)$. **b** Diagram for Comparing the Total Energy of the S = 1/2 States of the Monomeric fragments to the S = 0 and the S = 1 States of the Dimer

highest two dimmer orbitals and the subsequent spin-orbitals. Only the state corresponding to the T, S_1, and S_2 configurations need to be considered, because only these can possibly have appreciable populations up to 300 K. Figure 9.2a(iii) shows a semi-quantitative expression [41] for the energy difference between the triplet state (T for which $S = 1$) and the singlet state (S_1 and S_2 for which $S = 0$). This energy difference, $E_T - E_S$, is called $2J$ and is equal to $-2K_{ab} + 2[(\in_1 - \in_2)^2]/[J_{aa} - J_{ab}]$.

K_{ab}, the kinetic exchange energy between $Cu(II)_a$ and $Cu(II)_b$, is virtually constant and equal to ~50 cm^{-1}. J_{aa} and J_{ab} are the coulombic and exchange integrals, respectively, and their difference $J_{aa} - J_{ab}$ is also virtually constant and equal to ~40,000 cm^{-1}. $E_T - E_S$ is only strongly dependent on the square of the difference of the energy, $_1$, of molecular orbital phi_1 (**Ø$_1$**, refer to Fig. 9.2a(i)) and the energy, $_2$, of molecular orbital phi_2 (**Ø$_2$**). The occupancy for the Boltzmann distribution of the triplet and singlet states at room temperature, 300 K, is almost exclusively singlet (i.e., the Cu(II)–Cu(II) center is effectively diamagnetic) when $E_T - E_S$ is ~700 cm^{-1}. This effective diamagnetism requires the orbital energy difference, $_1 - _2$, to be ~4,000 cm^{-1} according to the expression in Fig. 9.2a(iii). Figure 9.2b shows the stabilization of $S = 0$ state and the destabilization of the $S = 1$ states of the dimer relative to the $S = 1/2$ states of the two monomeric fragments for an antiferromagnetic interaction.

Anaerobic samples of hemocyanin or tyrosinase in the oxidized state, Cu(II)–Cu(II), do not show the 330 nm absorption band characteristically associated with proteins with a Type 3 copper center under similar conditions. The Cu(I)–Cu(I) fragment in the reduced form of hemocyanin or tyrosinase does, however, react reversibly with molecular oxygen, O_2, to readily form a peroxide-bridged binuclear cupric fragment [42], which exhibits two characteristic new

$$[Cu(I)\ldots Cu(I)]^{+2} + O_2 \rightarrow [Cu(II)\text{ - O - O - }Cu(II)]^{+2}$$

charge transfer transitions, one at ~350 nm with an absorptivity equal to 20,000 $M^{-1}cm^{-1}$ and the other at ~570 nm with an absorptivity of ~1000 $M^{-1}cm^{-1}$. Figure 9.3 shows the pi (π) antibonding orbital of O_2 overlapping the cuprous dimer orbitals in hemocyanin or tyrosinase [54] and the subsequent molecular orbital diagram [30]. The intensity enhancement of the 750 cm^{-1} Raman band seen when the excitation radiation is near 350 or 570 nm is attributed to a resonance Raman effect and confirms the presence of the bridging peroxide [30, 54].

Although the entire biological function of hemocyanin is limited to reversibly binding O_2, tyrosinase can also function as a mono-oxygenase as well as a two-electron oxidase [42]. When the tyrosinase Cu(I)–Cu(I) site binds O_2 to make $[Cu(II)\text{-O-O-}Cu(II)]^{+2}$, this peroxide-bridged cupric fragment is a source of activated oxygen. The substrate, phenol, formally receives an oxygen atom from the $[Cu(II)\text{-O-O-}Cu(II)]^{+2}$ fragment (e.g., a bound phenolate is converted to catecholate). The oxene-depleted fragment, $[Cu(II)\text{-}O^{-2}\ldots Cu(II)]^{+2}$, accepts two protons and liberates water. The water-depleted cupric fragment which remains, $[Cu(II)\ldots Cu(II)]^{+2}$, accepts two electrons from the substrate (transforming catechol

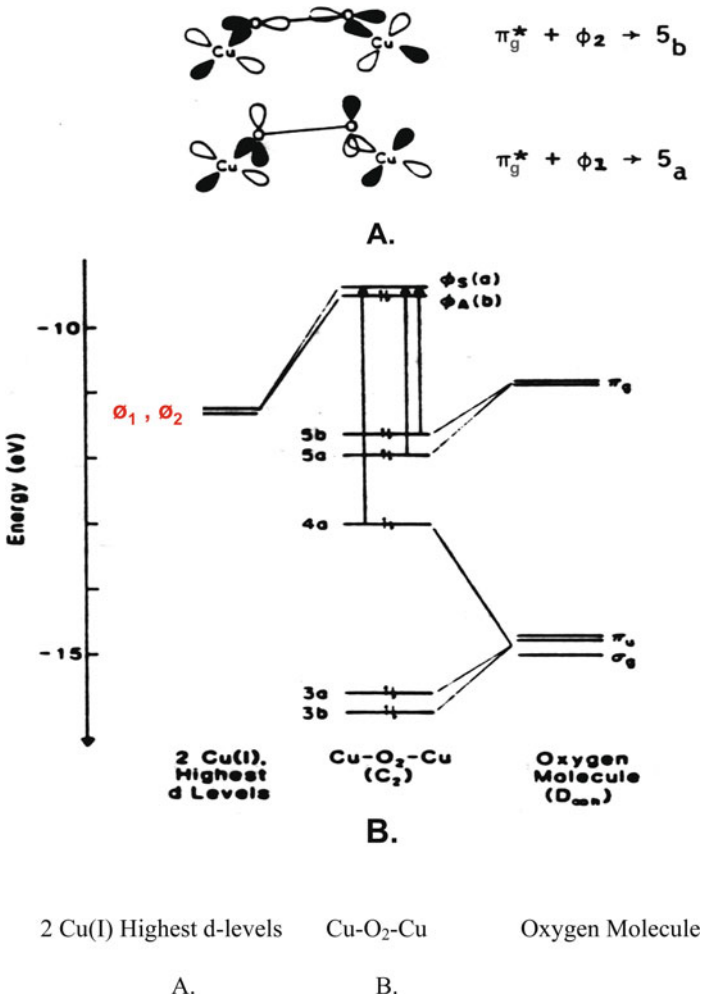


Fig. 9.3 Electronic Description of the Peroxide-Bridged Binuclear Cupric Fragment. A. Overlap of the Cu(I)–Cu(I) orbitals with π Orbitals of Molecular Oxygen, O_2. B. The Subsequent Molecular Orbital Diagram

to quinone). This sequence of events regenerates the Cu(I)...Cu(I) fragment in the active site, which is then ready to undergo another catalytic cycle. This proposed scheme for tyrosinase activity is summarized in Fig. 9.4 [85].

Type 3 Centers

Copper-ion active sites in enzymes and proteins have been classified as belonging to the categories called the Type 1, Type 2, and Type 3 copper centers [63]. In the

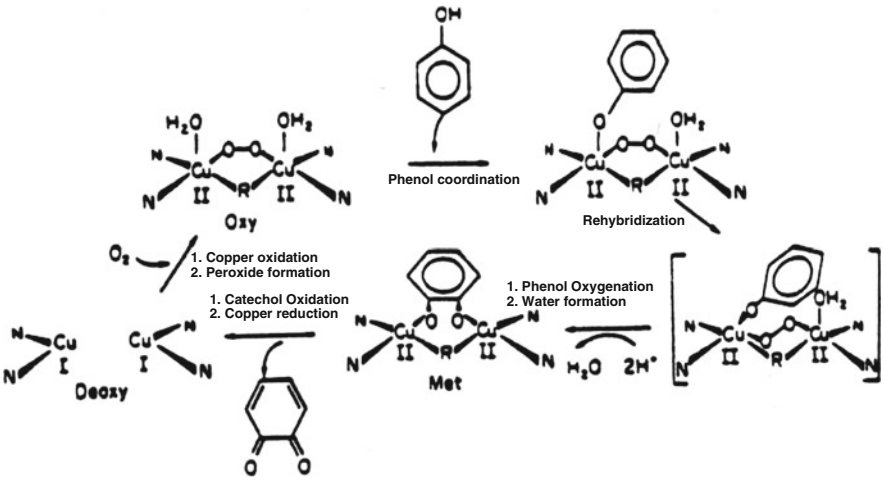


Fig. 9.4 Proposed Scheme for Tyrosinase Activity

native enzymes these three centers can be found either together or individually. Both the Type 1 and Type 2 centers denote sites containing only one copper ion. The Type 3 center was initially described as a copper cation site which (i) can accept two electrons, with the reduction potential for site being ~400–800 mV versus the normal hydrogen electrode at pH of ~7; (ii) is effectively diamagnetic at room temperature and shows no electron paramagnetic resonance signal at 77 K while the site is in its oxidized state; and (iii) absorbs 330 nm ultraviolet radiation with an absorption coefficient of ~2,000–4,000 M^{-1} cm^{-1}. Sensitive, variable temperature, magnetic susceptibility measurements confirmed that the oxidized state of this site has two Cu(II) ions which are antiferromagnetically coupled so strongly that this active site should be effectively diamagnetic at room temperature [81]. The essential electronic and magnetic properties for this binuclear copper center are consistent with the description given in Fig. 9.2.

Enzymes with a Type 3 center (e.g., the blue oxidases: ascorbate oxidase, ceruloplasmin, and the laccases) catalyze the oxidation of substrates with molecular oxygen, O_2, being the ultimate oxidant [63, 43]. As the substrate is oxidized, the Cu(II)–Cu(II) fragment can accept two electrons [24] to form a Cu(I)–Cu(I) fragment [76]. This Cu(I)–Cu(I) fragment in the fully reduced enzyme is reoxidized by O_2 thereby forming H_2O [43, 25]. This is summarized below:

$$\mathrm{Cu(II)-Cu(II)} + 2\mathrm{HA} \rightarrow \mathrm{Cu(I)-Cu(I)} + 2\mathrm{A} + 2\mathrm{H}^+$$
$$\mathrm{Cu(I)-Cu(I)} + 1/2\ \mathrm{O_2} + 2\mathrm{H}^+ \rightarrow \mathrm{Cu(II)-Cu(II)} + \mathrm{H_2O}$$

Molecular oxygen, O_2, is incapable of reoxidizing the Cu(I)−Cu(I) fragment found in the Type 3 center if a Type 2 center is not present [85]. Therefore the Type 2 center is associated with the working Type 3 site.

The origin of the 330 nm absorption band remains debatable. Spectrophotometric titration studies [85, 76, 62, 7] associate this 330 nm band with the cupric oxidation state of the Type 3 site. The 330 nm absorption was thought to arise from a thioether S to Cu(II) charge transfer transition [8, 10]; however, X-ray absorption (EXAFS) studies of the Type 3 center in fungal laccase reject the presence of any sulfur donor atoms in the copper inner-coordination sphere. A histidine to Cu(II) charge transfer transition could account for this absorption band [82] or the band may be associated with the Type 1 or 2 sites.

It is remarkable that the apparently simpler binuclear copper-ion active site in tyrosinase can accommodate reversible oxygen binding, oxygenation, and oxidation catalysis, while only oxidation catalysis is associated with the apparently much more complicated multicopper active site structures found in the blue oxidases.

Modeling Enzymes

The active sites in copper proteins are not rigid protein-independent centers like the heme group. The copper ions are coordinated to proximal ligating residues, which are from the amino acids making up the polymeric protein strand, which is uniquely folded in the catalytically active form of the enzyme. For this reason, it is unrealistic to expect to synthesize a copper complex, which is a detailed model for a copper protein active site. Parallel studies of synthetic models and the enzymes should, however, help to clarify the structure and physical properties of such enzyme active sites. Perhaps more importantly, model studies may lead to simple synthetic compounds, which effectively reproduce the desirable redox and O_2 chemistries exhibited by copper proteins. Specific attempts to make compounds, which mimic enzymes, follow. Table 9.1 summarizes the electrochemistry of the enzymes and synthetic mimics relevant to oxygen reduction.

Robson [39] reported the syntheses of novel binuclear cupric complexes (see Fig. 9.5), which he regarded as potential geometrical models for nitrogen-fixing systems. Although these molecules were never found suitable for catalyzing the nitrogen reduction reaction, this report led others to use 2-hydroxy-5-methylisophthalaldehyde as a building block for the synthesis of new complexes for mimicking binuclear copper proteins.

Table 9.1 Redox properties of binuclear copper proteins and mimics

Compound	Figure	E_o' (vs NHE)	References
Laccase	1	+0.8 V	[1–4]
Grzybowski′s complex	6	−0.5 to −0.7	[22]
Gagnés complex	7	−0.75 to −1.1	[24]
Lindvedt′s complex	8	−0.8 to −1.1	[27–29]
Lehn′s bincleating cryptanc	9	+0.46	[32,33]
Murakami′s complex	18	−0.34	[89]
Gervasio′s complex	21	+0.1	[89,90]

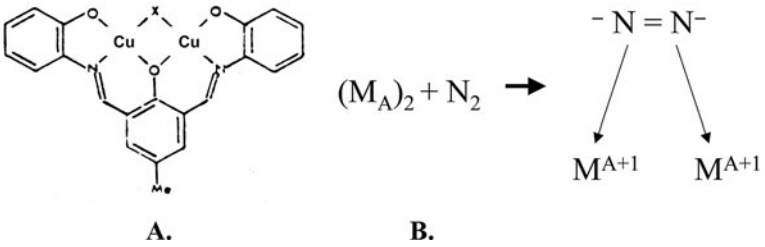


Fig. 9.5 (A) Robson's Binuclear Cupric Complex and (B) Proposed Scheme for Interacting The Complex With N_2

Urbach and Grzybowski [39] performed reactions between aminoalkylpyridines and aminoalkylimidazoles and 2-hydroxy-5-methylisophthalaldehyde to make new binucleating ligands. Pyridine and imidazole groups were used with the intention of stabilizing the Cu(I) oxidation state of the complexes. See Fig. 9.6.

The electroreduction of the hydroxide-bridged cupric complexes on a platinum electrode in 0.1 N tetraethylammonium perchlorate and acetonitrile electrolyte solution is irreversible and occurs at ~−0.5 to −0.7 V versus the saturated calomel electrode (SCE). Ascorbate, which can be oxidized at −0.18 V versus the saturated calomel electrode [61], could not chemically reduce these cupric complexes.

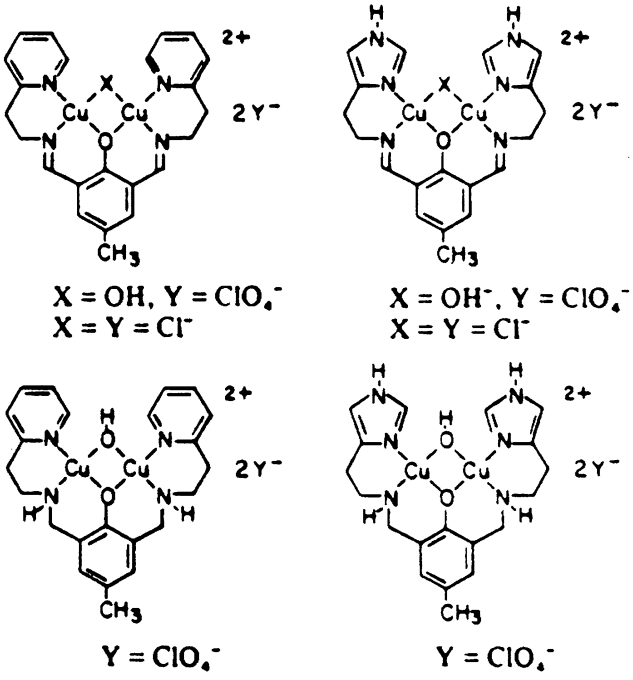


Fig. 9.6 Grzybowski's Binuclear Cupric Complexes

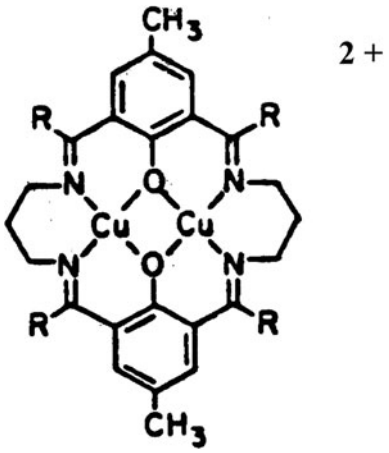


Fig. 9.7 Gagné's Binuclear Cupric Macrocycle Derived from 2-hydroxy-5-methylisophthalaldehyde and 1,3-propanediamine (R = H)

Okawa et al. [34], Addison [72], Gagné [32], and later Hendrickson [57] reported a variety of macrocyclic complexes, which are made by condensing 2-hydroxy-*S*-methyl-isophthalaldehyde with diamines (see Fig. 9.7). Gagné showed such macrocyclic complexes exhibit two, fairly reversible, one-electron reductions at ~−0.75 and −1.1 V versus the normal hydrogen electrode (NHE = −0.26 V versus SCE). The separation of successive one-electron reductions permitted mixed valence species to be isolated and studied. However, any compound, which has such negative reduction potentials, is not well disposed for promoting the catalysis of oxidation reactions, which use oxygen as the ultimate oxidant [93].

Lintvedt [27] prepared binuclear cupric complexes using a series of ligands based on substituted 1,3,5-hexanetrione (see Fig. 9.8) and their diamine Schiff-base derivatives. All of these complexes exhibit very low room temperature magnetic moments (0.7 Bohr magnetons or less) [97]. In the absence of sodium or lithium ions, solutions of these binuclear triketonate complexes display only one cyclic voltammetry wave between −0.8 and −1.1 V versus the saturated calomel electrode. This wave is attributed to a reversible, one-electron transfer on the basis of wave symmetry and peak current calculations. When excess Na^+ or Li^+ is added,

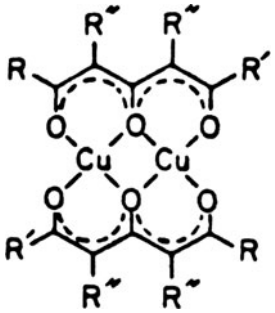


Fig. 9.8 Lintvedt's Dimeric Triketonate

the wave shifts anodically by 0.3 V and changes form. Peak current calculations and the $Ep_c - Ep_c^{1/2}$ diagnostic indicate a reversible, sequential transfer of two electrons at or very near the same potential when excess Ka or Li is present [56]. The magnetic and charge transfer equivalent are reminiscent of the enzyme active sites, but the redox potentials at which these triketonate complexes charge transfer are much lower than the enzyme active sites and too low for catalyzing reactions of molecular oxygen.

Lehn and coworkers reported the elegant syntheses of a variety of macrocyclic and macropolycyclic ligands, which encapsulate two copper ions [4]. See, e.g., Fig. 9.9.

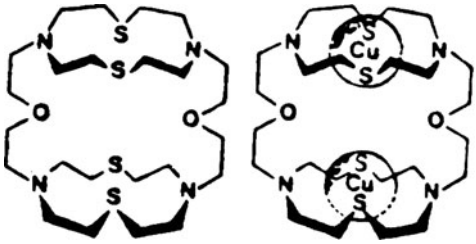


Fig. 9.9 Lehn's Binucleating Cryptand

Following the lead of Rorabacher [46], who used thioether-containing macrocycles to make monomeric cupric complexes with high positive reduction potentials, Lehn incorporated thioether sulfur atoms into his ligands. The rotating disk and cyclic voltammetries have been studied for one compound, 1,7,13,19-tetra aza-4,16-dioxa-10,22,27,32-tetrathia-tricyclo-17,5.5,5 tetra tri acontane dicopper(II) (shown in Fig. 9.9), dissolved in 0.1 M aqueous potassium chloride using platinum as the working electrode [38]. A two-electron transfer occurs at 200 mV versus SCE. The Cu–Cu distance is 5.62 Å , and the X-band epr spectrum of aqueous solutions of this complex exhibits four bands separated by ~85–90 Gauss in the $\Delta M_S = 1$ region [59]. From these data, it is concluded that this complex transfers two electrons at one potential, because there is very little or no interaction between the metal-ion sites.

Other binuclear species are based on similar macrocycles and show strong magnetic interactions [20]. An exogenous bridge (e.g., OH^-, N_3^-), not the macrocyclic ligand, promotes the strong coupling of the copper ions [21, 13]. (See Fig. 9.10 for two examples of such complexes.)

In the absence of the exogenous bridge, OH^-, Osborn's substituted *para*-xylyl compound has displayed partly reversible O_2 binding in propylene carbonate and in the solid state [12]. (Refer to Structure B, Fig. 9.10.)

Binucleating ligands, which offer only two or three donor atoms per copper, seem to favor the cuprous oxidation state. Gagné found [35] that Grzybowski's complex [39] (see Fig. 9.6), which exhibits one irreversible electron transfer step at −0.6 V, versus SCE, will exhibit two nearly reversible one-electron transfers at ~0 V and −0.25 V versus SCE, respectively, if the bridging hydroxide is replaced by a bridging pyrazolate anion. These two waves are ~0.5 V more positive than the two

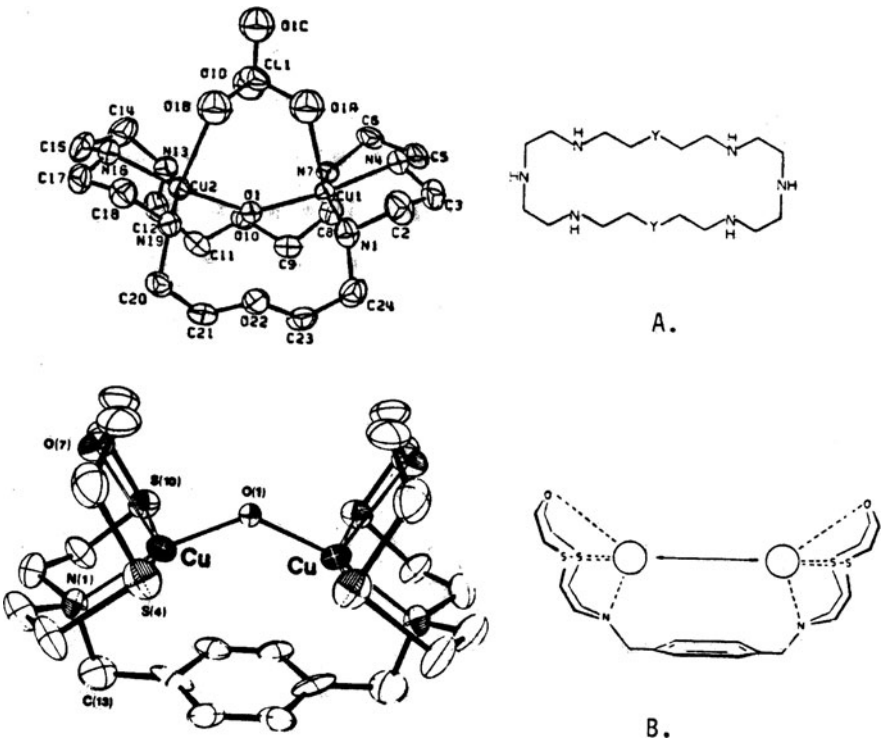


Fig. 9.10 Complexes using Lehn's Macrocycles. A. Lippard's Hydroxide-Bridged Cupric Complex. B. Osborn's Hydroxide-Bridged "Ear Muff" Complex

quasi-reversible waves observed for Gagné's macrocyclic complex (see Fig. 9.7) [32]. The cuprous derivative of this pyrazolate adduct of Grzybowski's complex was synthesized, and crystals were isolated which were suitable for an X-ray structure determination (see Fig. 9.11) [35]. In Gagné's macrocyclic complex (Fig. 9.7) the copper atoms are constrained to be roughly square-planar. This geometry stabilizes the cupric oxidation state. In the cuprous pyrazolate-bridged complex, the terminal pyridines need not and do not bind to the Cu atoms, and so the stabilization of

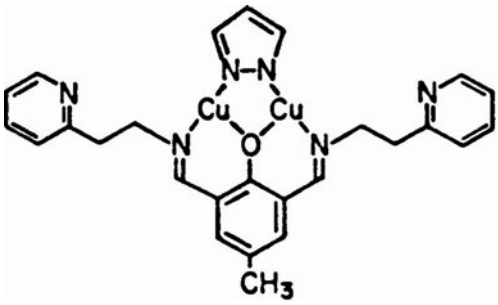


Fig. 9.11 Gagné's Cuprous Pyrazolate-bRidged Derivative of Grzybowski's Schiff-Base Complex

the cupric state is lost. This variable coordination effect could explain the ~+0.5 V shift in the Cu(II)/Cu(I) couple of the pyrazolate-bridged complex compared to the macrocyclic complex.

Nelson reports a binucleating macrocyclic ligand, which offers two donors per cuprous ion. In addition, each cuprous ion coordinates to an acetonitrile donor yielding trigonal coordination about each metal ion (see Fig. 9.12) [66]. The Cu–Cu distance in this complex is 3.35 Å. The cupric form of this complex is spontaneously reduced to a cuprous species when heated in acetonitrile. If suitable substrates (e.g., azobenzene, catecholate, ascorbate) are present, metal reduction is accompanied by substrate oxidation. The oxidation of 3,5-di-tert-butyl catechol is catalytic in the presence of O_2 when carried out in tetrahydrofuran. The observed stoichiometry in this catalytic reaction is 0.5 mole of O_2 consumed per 3,5-di-tert-butylquinone formed. Furthermore, in the autoxidation of the di-cuprous complex itself, a stoichiometry of 0.5 O_2 per 1 Cu(I)...Cu(I) is found. This stoichiometry suggests that each O_2 is associated with four Cu(I) ions (i.e., two binuclear complexes) during O_2 reduction. No electrochemical data were reported, but the spontaneous reduction to a di-cuprous complex in acetonitrile (MeCN) suggests that this solution species has a highly positive reduction potential.

Karlin and coworkers have developed a model system which mimics the dioxygen reactivity found in the copper mono-oxygenases (see Fig. 9.13) [48, 47].

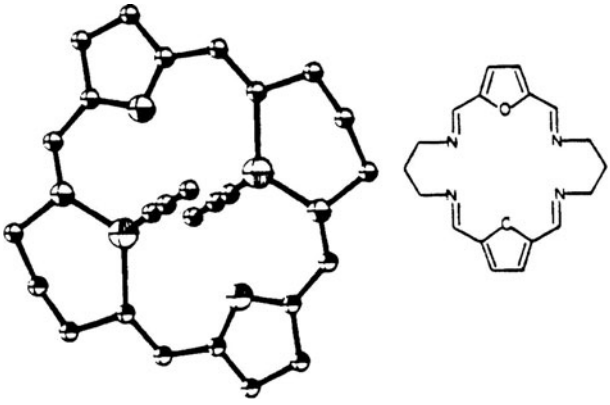


Fig. 9.12 Nelson's Cuprous Macrocycle with tWo N-Bound Acetonitrile Donors

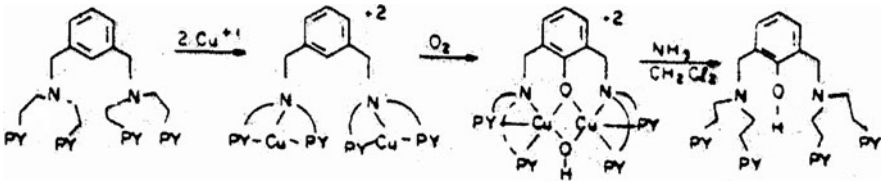


Fig. 9.13 Karlin's Scheme for Hydroxylating a Substituted meta-Xylyl Ligand

A substituted *meta*-xylyl ligand offers three donor atoms to each Cu(X) in the binuclear complex. When this cuprous complex is reacted with O_2, the aromatic ring is oxygenated in the position between the side arms to yield a binuclear cupric complex containing a new binucleating ligand with a phenolate-bridging group and an exogenous hydroxide bridge. The binuclear cupric complex formed as a product has a Cu–Cu separation of 3.1 Å. Removal of the cupric ions allows the substituted phenol ligand to be isolated. The binuclear cuprous complex of the new phenol ligand has been synthesized by the addition of cuprous ion to the isolated ligand. An X-ray study (see Fig. 9.14) of this binuclear cuprous complex shows no exogenous bridge and a Cu–Cu separation of 3.65 Å, which is very similar to the binuclear site in hemocyanin and tyrosinase[2]. When an orange dichloromethane solution of the binuclear cuprous complex is exposed to O_2 at $-70°C$, an intense violet color develops, due to a strong absorption at 505 nm. Manometric measurements show that one O_2 molecule is consumed per binuclear cuprous complex. Raman scattering using an excitation radiation of energy near the 505 nm electronic absorption shows resonance Raman enhancement of the intensity of 803 cm^{-1} vibrational band. This

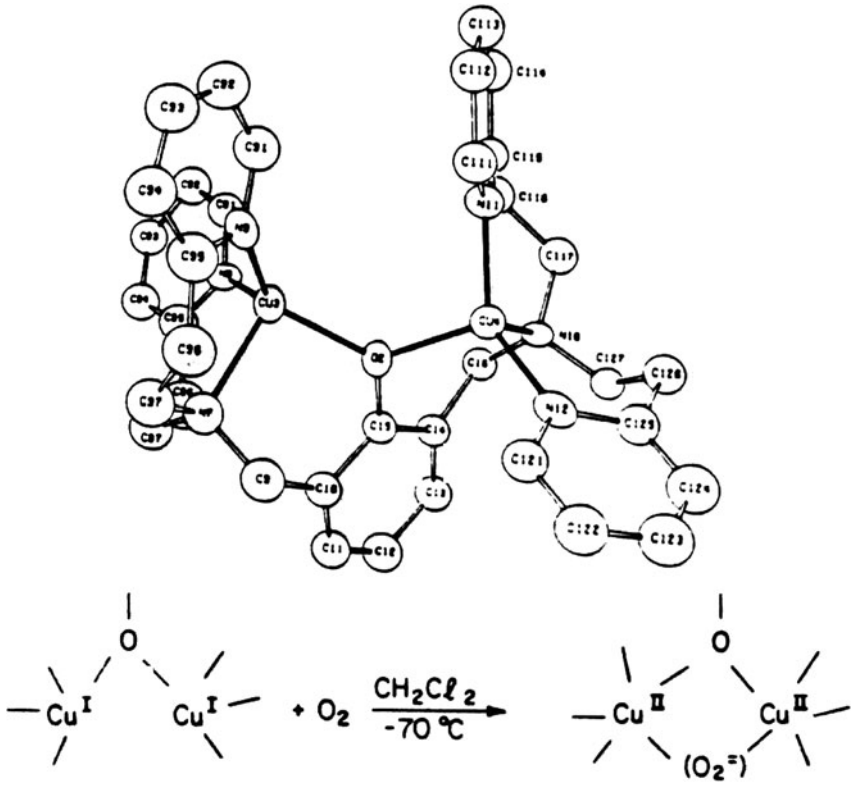


Fig. 9.14 Karlin's Phenolate-Bridged Binuclear Cuprous Complex and Scheme for Producing the pEroxide-Bridged Binuclear Cupric Fragment

vibrational band occurs at an energy only slightly higher than the range observed for the O–O stretch in metal peroxide complexes and strongly suggests formation of a peroxide-bridged binuclear cupric fragment in this solution species (see Fig. 9.14).

Both mononuclear and binuclear copper complexes with highly positive reduction potentials have been synthesized for mimicking enzyme sites and reactivity. Powers and Nikles [71, 70] synthesized a series of tetradentate ligands with thioether and amine or pyridine donors. This series of ligands allowed systematic variations in the chelate ring sizes produced in the copper complexes. One result of this study is that it showed thioether S bound to copper ions stabilizes the cuprous oxidation state [69] in agreement with other studies [46, 73]. Molecular orbital calculations performed as part of this study show that a pi-bonding interaction between the Cu d-orbitals and the empty thioether S d-orbitals is a feasible explanation for the stabilization of the cuprous state.

Durfee reported the synthesis and study of thioether containing analogues of Grzybowski's binuclear copper complexes [23]. Compared to Grzybowski's compounds, Durfee's compounds cycle between the Cu(II) and Cu(I) states more reversibly and at more positive potentials since the cuprous state has enhanced stability in the presence of the thioether donors (see Fig. 9.15).

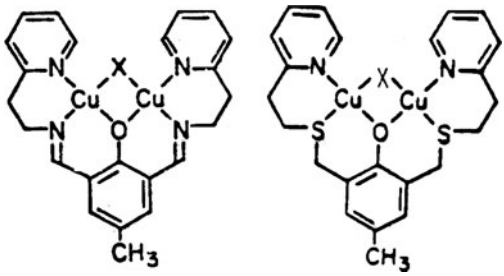


Fig. 9.15 Comparison of Grzybowski's Complex (left) to Durfee's Complex (right)

Pyrazole

The goal of mimicking a binuclear Cu enzyme active site is to replicate the structure, electronic properties, and functionality of the site. An atom coordinated to Cu in the mimic could just as well come from a naturally or not naturally occurring moiety. For example, pyrazole is an *isomer* of the biologically relevant chelate, imidazole. Pyrazole is an unsaturated, "soft," chelating group, which stabilizes lower oxidation states [28], a property leading to high redox potentials needed for effective O_2 reduction catalysis. Neutral pyrazole is a unidentate ligand (see the structure labeled I in Fig. 9.16) [28]. The pyrazolate anion can function as a 1,2 bridge between metal ions to space binuclear Cu ions ~3–4 Å apart, the same Cu–Cu separation found in the active site of Cu proteins (see structures II and III in Fig. 9.16) [28, 91, 90]. The planar bridging structure, II, in Fig. 9.16 is typical for copper-ion dimers.

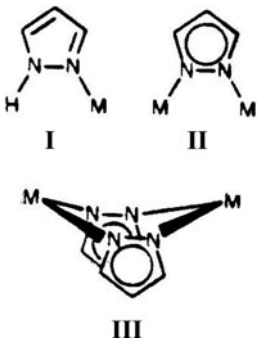


Fig. 9.16 Modes of Pyrazole Coordination

The binuclear complex, $[Cu_2(pz)_2(acac)_2]$, in which pz = pyrazolate and acac = acetylacetonate, shows moderate antiferromagnetic coupling (viz., $2J = \sim 70\ cm^{-1}$) [5]. This magnetic behavior has been rationalized in terms of a "superexchange" (or ligand–orbital) coupling of the singly occupied d_{x2-y2} orbital in each copper ion. The structure which is implicit to this mechanism has two coplanar pyrazolate anions bridging the two cupric ions, and each bridge is like mode II in Fig. 9.16.

One of the first reported structures of a copper dimer bridged by pyrazolate alone was a compound with a planar $Cu_2(pz)_2$ core. The Cu(II)–Cu(II) separation is 3.99 Å (see Fig. 9.17) [17].

The features outlined above make pyrazole complexes attractive for modeling the binuclear copper enzymes. Coupled binuclear copper complexes, which

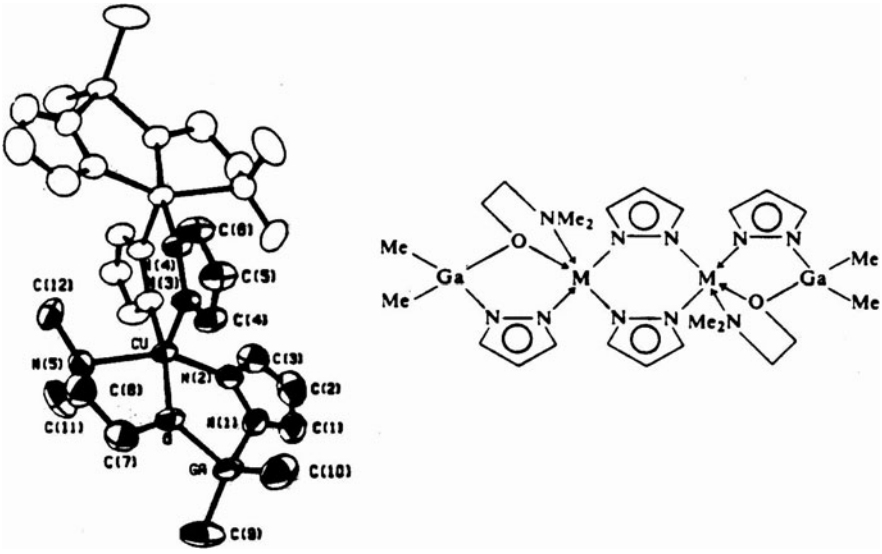


Fig. 9.17 A Cupric Dimer Bridged by Two Pyrazolate Anions

have multidentate ligands based upon a 1,2 bridging pyrazolate anion core, were first synthesized in the Urbach laboratory. The first such multidentate ligands had amide-substituted pyrazoles derived from 3-pyrazole carboxylic acid or 3,5-pyrazole dicarboxylic acid. The stylized structures of typical complexes in this series are shown in Fig. 9.18.

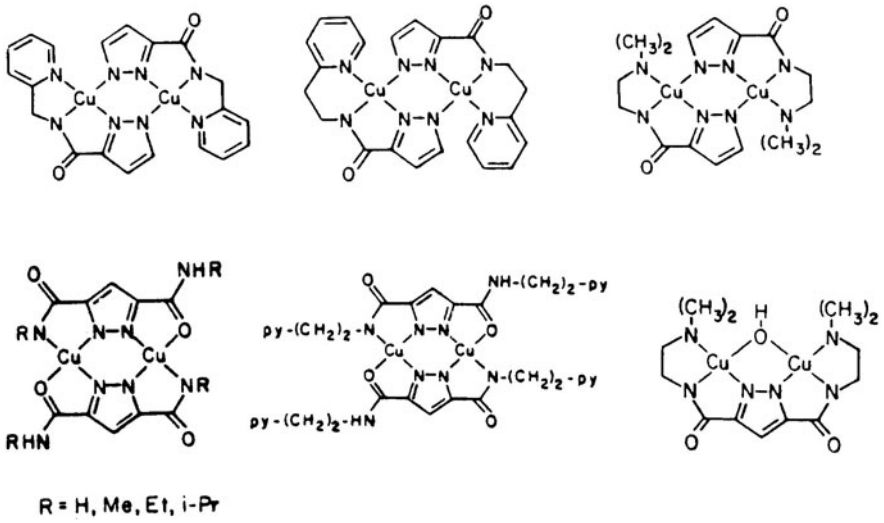


Fig. 9.18 Typical Cupric Complexes of The Amide-Pyrazole Type Ligands

One complex, $[Cu_2(pzdma)OH(H_2O)\ (H_2O)ax]$, has been studied in detail. When two equivalents of *N*,*N*-dimethylethylenediamine were reacted with methyl 3,5-pyrazoledicarboxylate, a diamide of pyrazole results. This diamide was further reacted with two equivalents of hydroxide and two equivalents of cupric ion to yield the complex, $[Cu_2(pzdma)OH(H_2O)(H_2O)ax]$, which has been formulated as $Cu_2N_6O_5C_{13}M_{26}$. The structure of this complex has been given by the single-crystal, X-ray diffraction method for the data from two-theta, 2Θ, spanning 2°–40° and isotropically refined to an error of fit of 12.6. The structure revealed under these conditions reliably shows a Cu–Cu distance of ~4.44 Å (see Fig. 9.19).

Although the amide–pyrazole complexes can be reduced by two electrons to form cuprous complexes, electrochemical characterization showed the greatest deficiency of these complexes as enzyme models. "Hard," negatively charged, amide donors strongly favor the cupric oxidation state over the cuprous state. The reduction potential for the two-electron transfer of $[Cu_2(pzdma)OH(H_2O)(H_2O)ax]$ is centered about ~−0.6 V versus the saturated calomel electrode. This is much too negative for a copper protein model[1,2].

In order to increase the stability of the cuprous oxidation state in pyrazole complexes, the amide amine side groups $[\text{-}CON^{-}\text{-}(CH_2)_x\text{-}N(CH_3)_2]$ were replaced by thioether-pyridine side groups $C\text{-}CH_2\text{-}S\text{-}(CH_2)_x$-pyr or tertiary amine-double pyridine or "tripod" side groups $\{\text{-}CM_2\text{-}N\text{-}[\text{-}(CH_2)_x\text{-pyr}]_2\}$. See Fig. 9.20.

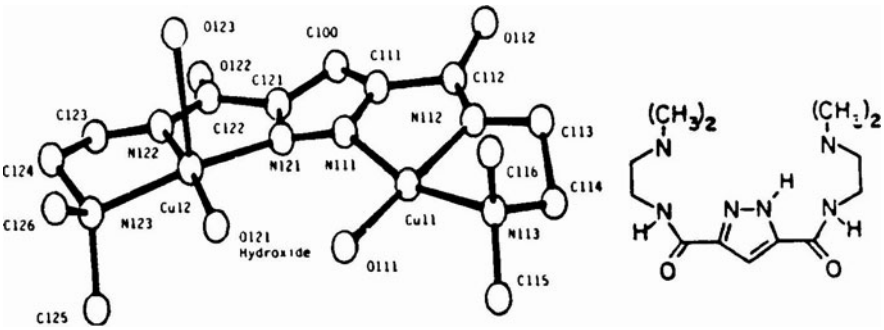


Fig. 9.19 Murakami's Binuclear Cupric Complex

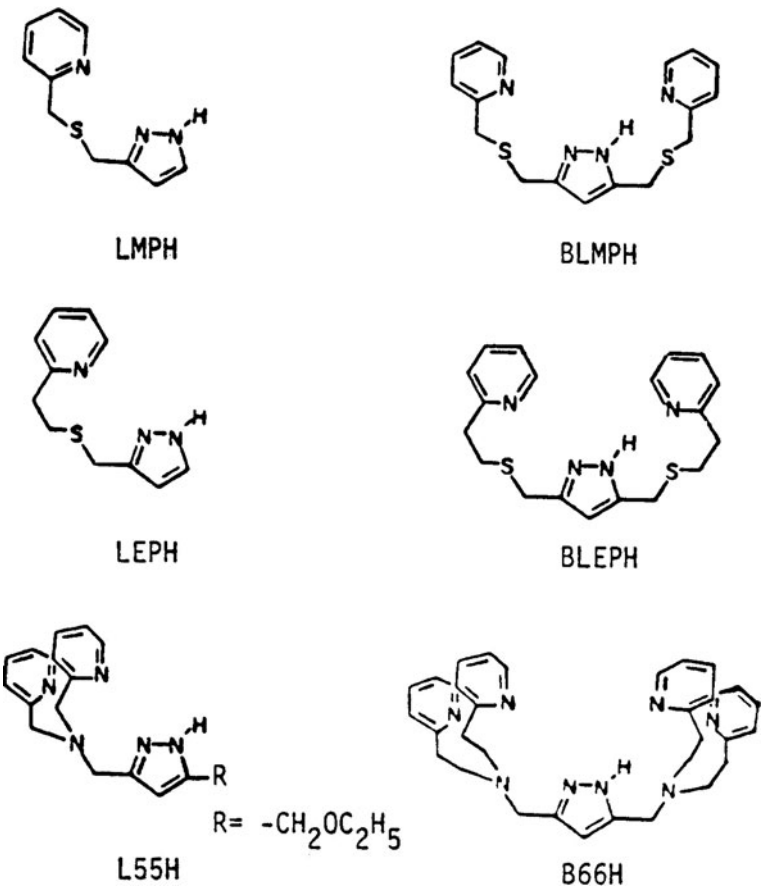


Fig. 9.20 New Pyrazole Macrocyclic Ligands with Pyridine and Thioether sidearms

The abbreviations for the short ligand names are: L, pyrazole with a methylene substituent at position 3; BL, pyrazole with methylene substituents at positions 3 and 5; MP, thiomethylpyridine side arm; EP, thioethylpyridine side arm; 55, amino-bis(methylpyridine) side arm; 66, amino-bis(ethylpyridine) side arm. Stylized structures of the cupric complexes synthesized in the authors' lab are shown in Fig. 9.21.

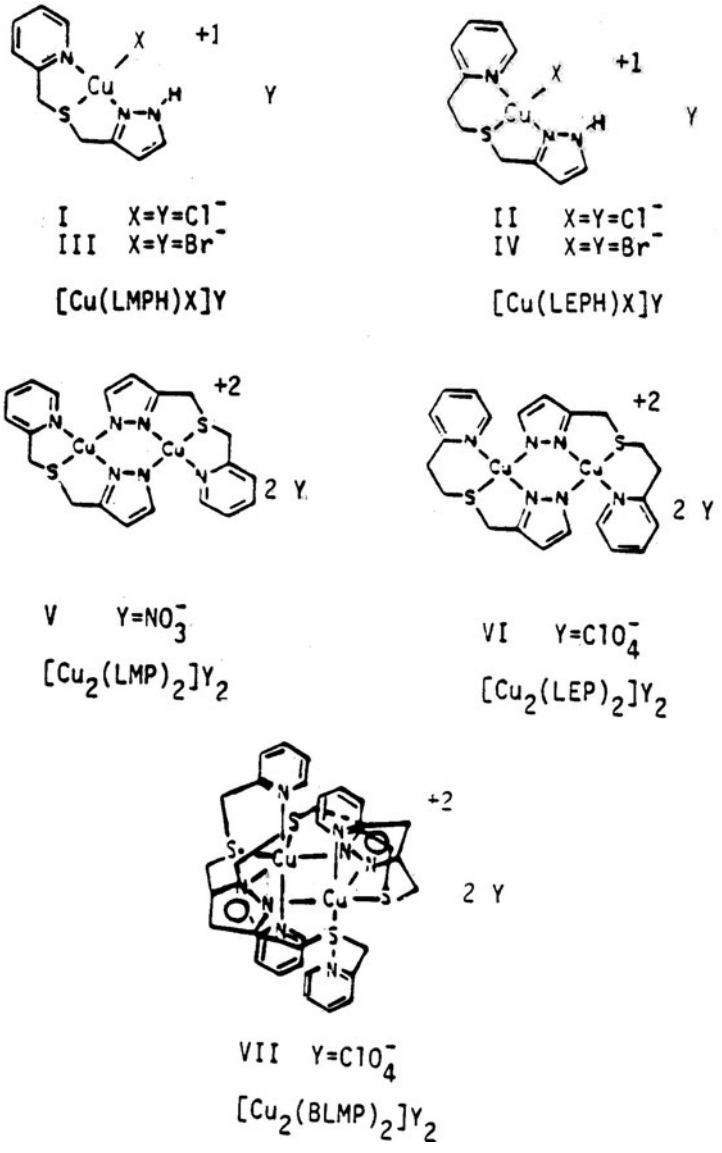
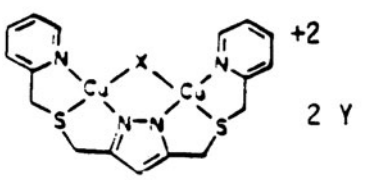
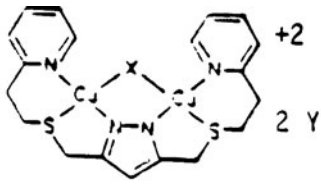


Fig. 9.21 (Continued)



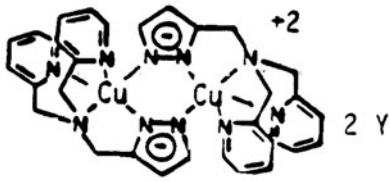
VIII X=Y=Br^-
IX X=OH^- Y=ClO_4^-
X X=pz^- Y= "

$[Cu_2(BLMP)X]Y_2$



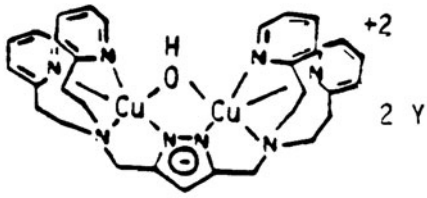
XI X=Y=Cl^-
XII X=Y=Br^-
XIII X=OH^- Y=ClO_4^-
XIV X=Cl^- Y= "
XV X=pz^- Y= "

$[Cu_2(BLEP)X]Y_2$



XVI Y=ClO_4^-

$[Cu_2(L55)_2]Y_2$



XVII X=OH^- Y=ClO_4^-

$[Cu_2(B66)(OH)]Y_2$

Fig. 9.21 Part 1—New Cupric Complexes of Pyrazole Macrocyclic Ligands with Pyridine And Thioether Sidearms Part 2—New Cupric Complexes of Pyrazole Macrocyclic Ligands with Pyridine And Thioether Sidearms

Of these compounds, one compound, namely compound XIII, $[Cu_2(BLEP)(OH)](ClO_4)_2$, was particularly well behaved and showed marked improvement, compared to Murakami's amide amine compound, $Cu_2(pzdma)(OH)$, for functionally modeling a binuclear Cu enzyme active site which catalyzes reactions involving O_2 reduction. Compound XIII is structurally similar to compound XI which was characterized by single-crystal methods and whose structure is given in Fig. 9.22.

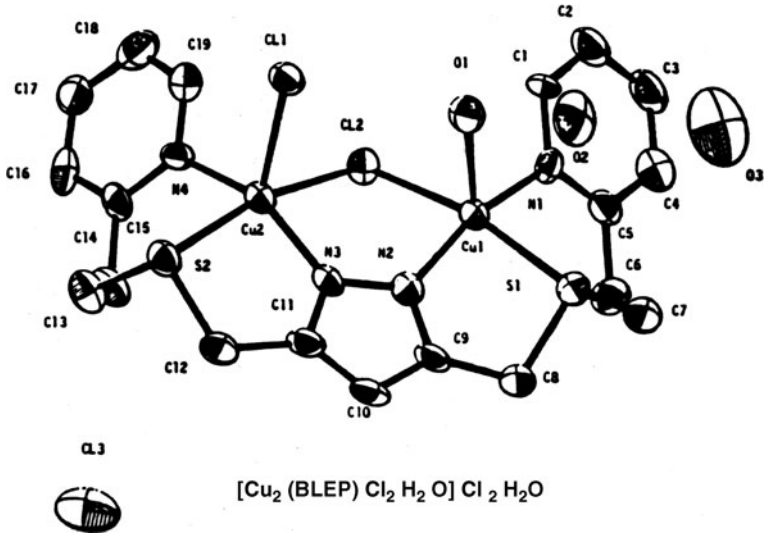


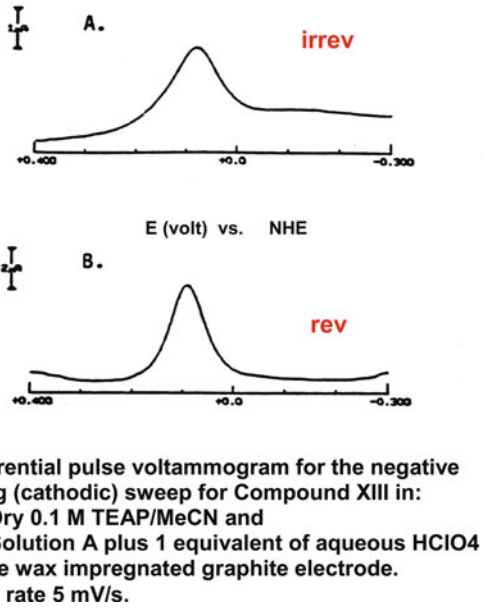
Fig. 9.22 Single-Crystal X-ray structure of compound XI, $[Cu_2(BLEP)Cl_2\ H_2O]Cl{\cdot}2H_2O$

When $[Cu_2(BLEP)(OH)](ClO_4)_2$ was reacted with 1 equivalent of acid, H^+, it was found to form a solution species exhibiting a fairly reversible two-electron transfer process which was found to occur at ~0.1 V versus NHE as shown in Fig. 9.23. This is a much higher, in fact ~0.7 V higher, redox potential, than the ill-behaved two-electron transfer process found for Murakami's amide amino Cu dimer, $Cu_2(pzdma)(OH)$, occurring at ~−0.6 V versus NHE. The positive shift in reduction potential of the acidified $[Cu_2(BLEP)(OH)]$ species compared to the $Cu_2(pzdma)(OH)$ species is attributed to the substitution of "harder" amide and amine N-donor atoms by softer thioether S and pyridine N-donor atoms.

Spectrophotometric titration showed that in $[Cu_2(BLEP)(OH)](ClO_4)_2$, Cu(II)...Cu(II) could catalyze the spontaneous air oxidation of ascorbic acid, HA, to the oxidized form, A, in homogeneous solution as shown:

$$\begin{array}{l} \text{HA oxidation } Cu(II)\ldots Cu(II) + 2HA \rightarrow Cu(I)\ldots Cu(I) + 2A + 2H^+ \\ \underline{O_2 \text{ reduction } Cu(I)\ldots Cu(I) + 1/2\ O_2 \rightarrow Cu(II)\ldots Cu(II) + 2A + H_2O} \\ \text{Net reaction } 2HA + 1/2\ O_2 + 2H^+ \rightarrow 2A + H_2O \end{array}$$

A new goal is to immobilize a synthetic Cu...Cu complex on an electrode, so O2 electroreduction can be heterogeneously catalyzed in a biofuel cell.



Differential pulse voltammogram for the negative going (cathodic) sweep for Compound XIII in:
A. Dry 0.1 M TEAP/MeCN and
B. Solution A plus 1 equivalent of aqueous HClO4
At the wax impregnated graphite electrode.
Scan rate 5 mV/s.

Fig. 9.23 The differential pulse voltammogram of compound XIII, $[Cu_2(BLEP)(OH)](ClO_4)_2$, in acetonitrile before and after the addition of one equivalent of aqueous perchloric acid

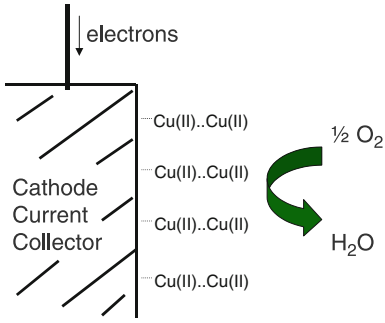


Fig. 9.24 A synthetic Cu. . .Cu complex, such as $[Cu_2(BLEP)](ClO_4)$, on an electrode, so O_2 electro-reduction to water can be heterogeneously catalyzed in a biofuel cell

$[Cu_2(BLEP)(OH)](ClO_4)_2$ is an attractive candidate as an oxygen electroreduction catalyst for an air cathode in a biofuel cell, provided that suitable schemes can be developed for immobilizing $[Cu_2(BLEP)(OH)](ClO_4)_2$ on an inert current collecting support as schematically suggested and illustrated in Fig. 9.24.

Conclusion

Developing materials which can coordinate O_2 and reversibly transfer multiple electrons at a high reduction potential is an approach for making a synthetic biomimetic oxygen reduction catalyst. The copper proteins and their synthetic analogues have been suggested as oxygen cathode catalysts for use in fuel cells running under physiological conditions. The copper enzymes, namely, the "Type 3" active centers found

in the laccases and hemocyanin and tyrosinase, can be potent catalysts for promoting the four-electron electroreduction of molecular oxygen to water, although two-electron reduction to peroxide and stability have been problems in using these enzymes for practical applications. Binuclear copper complexes have been made in a number of laboratories and have been shown to mimic the structure and properties of the binuclear copper-ion active sites found in copper proteins. Accordingly, enzyme-mimicking synthetic Cu complexes offer promise for developing an active yet stable catalyst for an O_2 cathode for biofuel cells operating under physiological conditions.

A number of binuclear cupric complexes have been synthesized using macrocyclic ligands which bridge two copper ions. These synthetic binuclear copper complexes mimic numerous essential features of the active sites of copper enzymes, which catalyze chemical reactions involving reduction of molecular oxygen. Most of the studies of the compounds reviewed here were carried out in homogeneous solution. However, it is a logical extension to use a synthetic Cu complex immobilized on an electrode, as an alternative to using a Cu enzyme immobilized on an electrode, as the catalyst for oxygen electroreduction in a biofuel cell cathode.

The need for power sources for new applications emerging from the recent growth in biotechnology suggests that renewed studies are warranted for exploiting binuclear copper complexes as enzyme active site mimics. Decorating a surface with a copper compound, which mimics a binuclear copper enzyme active site, is suggested as a route for making an air cathode in a biofuel cell operating under physiological conditions.

Acknowledgment D. Gervasio gratefully acknowledges F.L. Urbach, Professor Emeritus of the Chemistry Department at Case Western Reserve University in Cleveland, OH 44106, who supervised much of the work done at the university that is cited in this review.

References

1. Addison, A.W., Borer, L.L., Sinn, E.: Models of copper(II) proteins. Small-molecule models of copper(II) proteins, especially those involving copper-sulfur bonding. Metallothioneins 387–407. (1992)
2. Addison, A.W.: Inorg. Mud. Chem. Lett. **12**, 899 (1976).
3. Addison, A.W.: Spectroscopic and redox trends from model copper complexes. Copper Coord. Chem.: Biochem. Inorg. Perspect. 109–128 (1983).
4. Alberts, A.H., Annunziata, R., Lehn, J.M.: J. Am. Chem. Soc. **99**, 8503 (1977).
5. Barraclough, C.G., Brooks, R.W., Martin, R.L.: Aust. J. Chem. **27**, 1843 (1974).
6. Bauer, L., Gardella, L.A., Jr.: J. Org. Chem. **26**, 82 (1961)
7. Beinert, H.: Coord. Chem. Rev. **23**, 119 (1977).
8. Bennett, L.E.: Progress in Inorganic Chemistry. In: Lippard, S.J. (ed.) vol. 18. Wiley, New York (1973).
9. Bernaducci, B., Schwindinger, W.F., Hughey, J.L. IV, Krogh-Jespersen, K., Schugar, H.J.: J. Am. Chem. Soc. **103**, 1686 (1981).
10. Bosnich, B., Whelan, J., Amundsen, A.: J. Am. Chem. Soc. **99**, 6730 (1977).
11. Brubaker, G.N., Brown, J.N., Yoo, M.K., Kinsey, R.A., Kutchan, T.M., Mottel, E.A.: Inorg. Chem. **18**, 299 (1979).

12. Bulkowski, J.E., Burk, P.L. Ludmann, K.F., Osborn, J.A.: J. Chem. Soc. Chem. Comm. **498** (1977).
13. Burk, P.L., Osborn, J.A., N-T. Youinou, Agnus, Y., Louis, R., Weiss, R.: J. Am. Chem. Soc. **103**, 1273 (1981)
14. Li, C.-Y., Gulliken, J.S., Jr.: J. Chem. Ed. **54**, A217 (1977).
15. Cabral, J.O.: Electrochemical studies of synthetic models of copper proteins. NATO ASI Series, Series C: Math. Phys. Sci. **385**, 151–70 (1993)
16. Chen, H.-W., Fackler, J.P.: Inorg. Chem. **17**, 22 (1978).
17. Chong, K.S., Rettig, S.J., Storr, A., Trotter, J.: Can. J. Chem. **55**, 4166 (1977).
18. Chou, J.-L., Chyn, J.-P., Urbach, F.L., Gervasio, D.F.: Dinuclear copper(II) complexes incorporating a novel pyrazolo-based ligand with S- and N-rich coordination spheres. Polyhedron **19**, 2215–2223 (2000).
19. Churchill, N.R.: Inorg. Chem. **12**, 1213 (1973).
20. Comarmond, J., Plumere, P., Lehn, J.M., Agnus, Y., Louis, R., Weiss, R., Kahn, O., Morgenstern-Badarau, I.: J. Am. Chem. Soc. **104**, 6330 (1982).
21. Coughlin, P.K., Lippard, S.J.: J. Am. Chem. Soc. **103**, 3228 (1981).
22. Desjardins, S.R., Penfield, K.W., Cohen, S.C., Musselman, R.L., Solomon, E.I.: J. Am. Chem. Soc. **105**, 4590 (1983).
23. Durfee, W.S.: Ph. D. Thesis, Case Western Reserve University (1984).
24. Ehrenberg, A., Malmstrom, B.G., Broman, L., Mosbach, R.: J. Mol. Biol. **5**, 450 (1962).
25. Farver, O., Goldberg, M., Pecht, I.: Eur. J. Biochem. **104**, 71 (1980).
26. Feltham, R.D., Hayter, R.G.: J. Chem. Soc. 4587 (1964).
27. Fenton, D.B., Lintvedt, R.L.: J. Am. Chem. Soc. **100**, 6367 (1978).
28. Fieselmann, B.F., Stucky, G.D.: Inorg. Chem. **17**, 2074 (1978).
29. Figgis, B.N., Lewis, J.: Techniques of Inorganic Chemistry. In: Jonassen, H.B., Weissberger, A. (eds.), vol. IV, p. 137. Wiley, New York, N. Y. (1965)
30. Freedman, T.B., Loehr, J.S., Loehr, T.M.: J. Am. Chem. Soc. **98**, 2809 (1976).
31. Gaber, B-P., Miskowski, V.M., Spiro, T.G.: J. Am. Chem. Soc. 96,6868 (1974).
32. Gagné, R.R., Koval, C.A., Smith, T.J.: J. Am. Chem. Soc. **99**, 8367 (1977).
33. Gagné, R.R., Kovel, C.A., Lisensky, G.C.: Inorg. Chem. **19**, 2855 (1980).
34. Gagné, R.R., Kreh, R.P., Dodge, J.: Electrochemical and Spectroelectrochemical Studies of Biological Redox Components. In: Karl N. Kadish (ed.) Advances in Chemistry Series, vol. 201, p. 145. American Chemical Society, Washington, D.C. (1962).
35. Gagné, R.R., Kreh, R.P., Dodge, J.A.: J. Am. Chem. Soc. **101**, 6917 (1979).
36. Gervasio, D.F.: Binuclear Cupric Complexes Bridged by the Pyrazolate Anion. Ph. D. Thesis, Case Western Reserve university, Cleveland, OH (1984).
37. Ghendini, N., DeMunno, G., Denti, G., Mannotti Lanfredi, A.M., Tiripicchio, A.: Inorg. Chim. Acta **57**, 87 (1982).
38. Gisselbrecht, J.P., Gross, M., Alberts, A.H., Lehn, J.M.: Inorg. Chem. **19**, 1386 (1980).
39. Grzybowski, J.J., Merrell, P.H., Urbach, F.L.: Inorg. Chem. **17**, 3078 (1978).
40. Hathaway, B.J., Billings, D.E.: Coord. Chem. Rev. **5**, 143 (1970).
41. Hay, P.J., Thibeault, J.C., Hoffman, R.: J. Am. Chem. Soc. **97**, 4884 (1975).
42. Himmelwright, R.S., Eickman, N.C., LuBien, C.D., Lerch, K., Solomon, E.I.: J. Am. Chem. Soc. **102**, 7339 (1980).
43. Holwerda, R.A., Wherland, S., Gray, H.B.: Annu. Rev. Biophys. Bioeng. **5**, 363 (1976).
44. Jahn, H.A., Teller, E.: Proc. Roy. Soc. **220** (1937).
45. Jones, R.G.: J. Am. Chem. Soc. **71**, 3994 (1949).
46. Jones, T.B., Rorabacher, D.B., Ochyrmowycz, L.A.: J. Am. Chem. Soc. **97**, 7485 (1975).
47. Karlin, K.D., Cruse, R.W., Y. Gultneh, Hayes, J.C., Zubieta, J.: J. Am. Chem. Soc. **106**, 3372 (1904).
48. Karlin, K.D., Feller, D.M., DiPierro, L.T., Simon, R.A., Zubieta, J.: J. Chem. Soc. Chem. Comm. **881** (1981).
49. Karlin, K.D., Hayes, J.C., Juen, S., Hutchinson, J.P., Zubieta, J.: Inorg. Chem. **21**, 4106 (1982).

50. Kim, M.-H., Birke, R.L.: Anal. Chain. **55**, 1735 (1983).
51. Koenig, D.B., Brookhaven Laboratories, Chicago, Illinois [personal communication] (1979).
52. Kroneck, P.M.H.: Redox properties of blue multi-copper oxidases, Multi-Copper Oxidases, pp. 391–407. World Scientific Publishing Co. Pte. Ltd., Singapore (1997). (See also chapter 14 in same book, "Model Compounds for Multi-Copper Oxidases," by Chaudhuri, P.).
53. Langford, C.H., Grey, H.B.: Ligand Substitution Processes. In: Breslow, R., Karplus, N. (eds.) Frontiers in Chemistry. Benjamin, W.A., Inc., New York, N. Y. (1966).
54. Larabee, J.A., Spiro, F.G.: J. Am. Chem. Soc. **102**, 4217 (1980).
55. Libus, W., Strzelecki, H.: Electrochimica. Acta **15**, 703 (1970).
56. Lintvedt, R.L., Ranger, G., Schoenfelner, B.A.: Inorg. Chem. **23**, 688 (1984).
57. Long, R.C., Hendrickson, D.N.: J. Am. Chem. Soc. **105**, 1513 (1983).
58. Lontie, R., Vanquickenborne, L.: Metal Ions in Biological Systems. In: Sigel, H. (ed.) vol. 3, p. 183. Marcel Dekker, New York (1974)
59. Louis, R., Agnus, V., Weiss, R.: J. Am. Chem. Soc. **100**, 3604 (1978)
60. Machonkin, T.E., Solomon, E.I.: The Thermodynamics, Kinetics, and Molecular Mechanism of Intramolecular Electron Transfer in Human Ceruloplasmin. J. Am. Chem. Soc. **122**, 12547–12560 (2000). (The newest value for Eo for ceruloplasmin).
61. Mahler, H.R., Cordes, E.H.: Biological Chemistry, 2nd edn., p. 30. Harper and Row, New York, N. Y. (1971)
62. Malkin, R., Malmstrom, B.G., Vanngard, T.: Eur. J. Biochem. **10**, 324 (1969).
63. Malkin, R., Malmstrom, B.G.: Adv. Enzymol. **33**, 177 (1970).
64. Mann, P., Woffson, M., Germain, G., MULTAN: A Computer Program for the Automatic Solution of Crystal Structures. Department of Physics, University of York, England (1971).
65. Miskowski, V.M., Thich, J.A., Solomon, R., Schugar, H.J.: J. Am. Chem. Soc. **98**, 8344 (1976).
66. Nelson, S.K., Esho, F., Lavery, A.: J. Am. Chem. Soc. **105**, 5693 (1983).
67. Nicholson, R.S., Shain, I.: Anal. Chem. **36**, 706 (1964).
68. Nicholson, R.S.: Anal. Chem. **38**, 1406 (1966).
69. Nikles, D.E., Anderson, A.B., Urbach, F.L.: Copper Coordination Chemistry Biological and Inorganic Perspectives. In: Karlin, K.D., Zubieta, J. (eds.), p. 203–223. Adenine Press, Guilderland, New York (1982)
70. Nikles, D.E., Powers, M.J., Urbach, F.L.: Inorg. Chem. **22**, 3210 (1983).
71. Nikles, D.E., Powers, N.J., Urbach, F.L.: Inorg. Chem. Acta. **37**, L499 (1979).
72. Okawa, H., Honda, N., Kida, S.: Chem. Letters 1027 (1972).
73. Patterson, G.S., Holm, R.H., Bioinorg. Chem. **4**, 257 (1975).
74. Polcyn, D.S., Shain, I.: Anal. Chem. **38**, 370 (1966).
75. Prochaska, H.J., Schwindinger, W.F., Schwartz, N., Burk, M.J., Bernaducci, E.., Lalancette, R.A., Potenza, J.A., Schugar K.J.: J. Am. Chem. Soc. **103**, 3446 (1981).
76. Reinhammar, B.: Biochim. Biophys. Acta, **275**, 245 (1972).
77. Robson, R.: Inorg. Nucl. Chem. Letters **6**, 125 (1970).
78. Romary, J.K., Zacharissen, R.D., Barger, J.O., Schiesser, H.: J. C. S. (C), 2884 (1968).
79. Schonsees, W., Glick, M.: [from a personal communication to Professor Fackler, J.P.] Wayne State University, Detroit, Michigan (1977).
80. Solomon, E.I., Chen, P., Metz, M., Lee, S.-K., Palmer, A.E.: Oxygen binding, activation, and reduction to water by copper proteins. Angewandte Chemie, Inter. Edition **40**, 4570–4590. (2001)
81. Solomon, E.I., Dooley, D.N., Wang, R.-H., Gray, H.B., Cerdanio, N., Mogno, F., Romani, G.L.: J. Am. Chem. Soc. **98**, 201 S (1976).
82. Solomon, E.I., Penfield, K.W., Wilcox, D.E.: Structure Bond. **53**, 1 (1983).
83. Solomon, E.I., Sundaram, U.M., Machonkin, T.E.: Multicopper Oxidases and Oxygenases. Chem. Rev. **96**, 2563–2605 (1996). (Gives table of values of Eo for copper proteins.).
84. Solomon, E.I.: Copper Proteins. In: Spiro, T.G. (ed.) p. 41. Wiley, New York (1981)
85. Solomon, E.I.: Pure Appl. Chem. **35**, 1069 (1983).

86. Sorrell, T.N., Jameson, D.L.: .J. Am. Chem. Soc. **19**, 6013 (1983).
87. Sorrell, T.N., Malochowski, N.R.., Jameson, D.L.: Inorg. Chem. **21**, 3250 (1982).
88. Strouse, C.: Personal Communication, U. C. L. A., Los. Angeles, California (obtained from Doctor Raymond G. Teller of the Standard Oil of Ohio, Warrensville, Ohio) (1980).
89. Thompson, J.S., Marks, T.B., Ibers, J.A.: J. Am. Chem. Soc. **101**, 4180 (1979).
90. Trofimenko, S., Chem. Rev. **72**, 497 (1972).
91. Trofimenko, S.: Acc. Chem. Res. **4**, 17 (1971).
92. Urbach, F.L.: Metal Eons in Biological Systems. In: Sigel, H. (ed.) vol. 13, p. 73. Marcel Dekker, New York (1981)
93. van der Brink, F., Barendrecht, E., Visscher, W.: Recuil Review **99**, 253 (1980).
94. Vanneste, W.H., Zuberbuhler, A.: Molecular Mechanisms of Oxygen Activation, Hayaishi, O. (ed.), p. 371. Academic Press, New York (1974)
95. von Rothenburg, R.: J. Prakt. Chem. **52**(2), 45 (1895); Knorr Ann. **279**, 231 (1894)
96. Williams, R.J.P.: J. Chem. Soc. 1371 (1955).
97. Wishart, J.F., Ceccarelli, C., Lindvedt, R.L., Berg, J.N., Foley, D.P., Frey, T., Hahn, J.E., Hodgson, H.O., Weiss, R.: Inorg. Chem. **22**, 1667 (1983).

Index

A. Korkin et al. (eds.), *Nanotechnology for Electronics, Photonics, and Renewable Energy*, Nanostructure Science and Technology, DOI 10.1007/978-1-4419-7454-9,