Self-assembly of one-dimensional nanostructures at silicon surfaces


Department of Physics, University of Wisconsin Madison, 1150 University Avenue, Madison, WI 53706-1390, USA

Abstract

Nanostructures at surfaces and interfaces are a fertile testing ground for bringing the idea of ‘tailored solids’ towards reality. Electronic properties can be controlled systematically by confinement or by interface effects. The presence of a single crystal substrate allows for the self-assembly of highly regular nanoobjects, such as stripes and strings of dots with sizes of about 10 nm. Using silicon as substrate facilitates the electronic integration of nanodevices into micro-electronics. We speculate how such structures might evolve into future devices, such as data storage arrays with densities of Terabits/cm² and self-assembled, highly parallel data processors for pattern recognition. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: A. Nanostructures; A. Quantum wells; A. Surfaces and interfaces; B. Nanofabrications; C. Scanning tunnelling microscopy

PACS: 81.05.Ys; 81.05.Zx; 73.20.At; 79.60.Jv

1. Characteristics of nanostructures

In recent years there has been a veritable explosion of ideas for potential applications of nanostructured materials. In fact, nanostructures are seen as the key for many applications in data storage, communications and biotechnology. Examples are highly parallel data processors [1], high-density storage media [2–5], and bio-sensors [6].

Novel computing paradigms proposed today are in most cases dependent on the availability of massively parallel circuits [1,7]. One hopes to create these in an efficient way by self-assembly of nanostructures. In chemically self-assembled computer arrays one would use molecular switches [7–9] that are addressed by orthogonal grids of nanowires. Organizing such systems as cellular automata [10–12] requires both very precise control of the size and shape of the elements and their relative positioning on the nanometer scale.

The density limit in magnetic data storage is largely determined by the inhomogeneity of the magnetic particles that make up the storage medium [3–5]. Overcoming variations in particle size, shape, magnetic switching, and arrangement currently requires the use of a few hundred particles for each recorded bit. The error limits are extremely stringent (less than one error in $10^8$ read/write cycles, which can be further reduced to one error in $10^{12}$ cycles by error-correcting codes). The size of the individual particles in today’s media approaches the superparamagnetic limit already, where thermal fluctuations flip the magnetization. For further improvements one has to use fewer, but more homogeneous particles. The ultimate limit is a single particle per bit, which would increase the storage density by more than two orders of magnitude.

Biosensors and biomaterials represent an attempt to combine the sensing and processing techniques of microelectronics with the ability to detect or amplify biological signals. Many important biological structures have sizes in the nanometer range (DNA, viruses, proteins, antibodies [13–16]). It is important for the functioning of such biostructures to present them a surface that is functionalized and patterned on the same scale. For example, most cells require a membrane to grow on, and structural features on the 100 nm scale have been found on such membranes [17]. For tasks ranging from drug discovery to DNA sequencing one wants to selectively deposit biologically active materials with sub-micron precision in order to make assays for rapid, highly parallel analysis.

 Trying to explain the broad appeal of nanostructures eventually boils down to two observations:

(A) Nanostructuring creates new materials with tailored properties.
(B) Self-assembly produces large quantities of structured materials. In order to tailor the electronic structure of a material one has to go down to the single-digit nanometer regime. For example, silicon keeps its bulk properties down to the smallest commercial sub-micron devices, but changes its fundamental band gap when confined to dimensions in the single-digit nanometer regime [18,19]. It is interesting to note that several arguments lead to such a length scale, as shown in Fig. 1. Consider the following electronic devices and require that they be operable at room temperature, that is in the presence of thermal energy fluctuations of about $kT \approx 25$ meV:

1. **Quantum wells.** The quantum-mechanical energy difference between the lowest two quantum well states is $E_{QM} = \frac{3\hbar^2}{8ml^2}$ for an infinite potential well of width $l$. Requiring $E_{QM} \approx kT$ for preventing thermal excitations gives a width $l \approx 7$ nm.

2. **Single electron devices.** The Coulomb energy for a single electron on a sphere of radius $r$ in a surrounding medium with dielectric constant $\varepsilon$ is $E_C = \frac{e^2}{\varepsilon r}$. Using $\varepsilon$ of silicon and requiring again $E_C \approx kT$ one obtains a diameter $2r \approx 9$ nm.

3. **Magnetic data storage.** The superparamagnetic limit for the minimum magnetic particle size is determined by the energy barrier for reversing the magnetization. In the case of pure shape anisotropy of a needle-shaped particle the barrier is of the order $E_M = \frac{1}{2}M^2V$ [5] for a volume $V$ and a magnetization $M$. In iron, this energy becomes larger than $kT$ for an average particle size $V^{1/3} = a \approx 3$ nm.

Surprisingly, three different types of energies with different power laws in $l$, $r$, $a$ lead to comparable length scales. Quantum mechanics, electrostatics, and magnetism conspire to produce a common length scale in the single-digit nanometer regime.

The methods for fabricating structures of that size can be classified into two types, i.e. writing and self-assembly. Writing is best for creating an individual device, whereas self-assembly is well suited for producing macroscopic amounts of tailored materials or highly parallel arrays of
devices. In the following, we will focus on self-assembly. Nanostructures with one-dimensional order will be emphasized, such as arrays of wires and strings of dots. Stepped silicon surfaces serve as atomically precise templates. For the subsequent self-assembly of wires and dots we exploit principles borrowed from microlithography, such as the use of an inert mask and the selective deposition between masked areas. We follow this path in Section 2, starting with silicon templates, applying CaF$_2$ masks, and selectively depositing molecules and metals. Section 3 looks ahead towards structures near the atomic limit. They allow us to demonstrate exotic electronic structures that occur in one-dimensional systems. We will allude to the dream of the ultimate data storage medium that stores a bit in a single atom. In Section 4 we speculate how such one-dimensional nanostructures might evolve into future devices, such as structured data storage media and highly parallel processors for pattern recognition.

![Fig. 2. Schematic of one-dimensional arrays on a stepped surface (top) and STM topographs of actual structures fabricated with CaF$_2$ on stepped silicon (bottom). (A) Stepped silicon surface, serving as a template (detail in Fig. 3). (B) CaF$_2$ stripes on top of stepped silicon, which has been coated by a CaF$_2$ layer ([23], detail in Fig. 4). (C) Linear array of CaF$_2$ dots on stepped silicon [23,25].](image)

![Fig. 3. Step arrays on Si(111)7×7 with step spacings from 6 to 80 nm, depending on the miscut angle. Kink densities as low as one in 20 000 can be achieved. The reason for such perfection is the 7×7 reconstruction which divides the surface into 2.3 nm wide units (d). The x-derivative of the tip height is displayed in these STM images, simulating illumination from the left and producing dark lines at step edges. (a) 10° miscut towards [112], triple steps [29]. (b) 1° miscut towards [112], single steps [28]. (c) 6° miscut towards [112], bunched steps [28]. (d) Close-up of (b) showing the grooves at the edges of the 7×7 unit cell that determine the location of step edges [28].](image)
2. Emulating silicon microlithography on the nanoscale

Silicon-based microlithography is a highly developed technology. Any attempt to extend patterning techniques into the nanometer regime will benefit from paying attention to silicon lithography techniques. A silicon substrate will also facilitate the electronic connection of nanodevices with the macroscopic world, one of the biggest hurdles for using nanostructures. In the following we outline a program for silicon-based nanopatterning, with the extra twist of using stepped silicon substrates to impose one-dimensional order. We retain the concept of a mask combined with selective etching and deposition around the mask, which is characteristic of microlithography. Instead of a photoresist we utilize calcium fluoride as mask, an inert material that grows epitaxially on silicon with an atomically precise interface [20–22]. Typical masks consist of stripes or strings of dots, as shown in Fig. 2 [23–25].

Regular arrays of steps can be produced on silicon surfaces with periodicities ranging from 2 to 80 nm [26–30]. Examples are given in Fig. 3. On Si(111)7 × 7 the steps are much straighter than on metal surfaces (compare Ref. [31]). The kink density is suppressed to values as low as one in 20 000 edge atoms [27,28]. The reason for such straight steps is the large-scale surface reconstruction that is characteristic of many semiconductor surfaces, and of Si(111)7 × 7 in particular. At this surface, there are stacking faults that introduce natural grooves every seven atomic rows (2.3 nm). These can be seen in Fig. 3 as vertical lines connecting the corner holes of the 7 × 7 unit cell. The resulting grid of grooves forms a template for step arrays and allows us to control the step spacing with atomic precision in intervals of 2.3 nm. The top of a step edge always coincides with one of these grooves. In order to form a kink one has to add 14 atomic rows, thereby extending the terrace by seven rows two layers deep (because each step consists of two Si(111) layers). Such a process is highly suppressed.

The step arrays in Fig. 3 are prepared from vicinal Si(111) wafers with a miscut of 1–10° after a special annealing...
sequence [27,28]. A flash to 1250°C produces a clean, oxygen- and carbon-free surface, and a long post-anneal below 850°C creates a single-domain $7 \times 7$ reconstruction which is necessary for the formation of a regular grid of grooves and the resulting suppression of kinks. Step spacing and step height can be controlled by the miscut angle. Cuts towards the [1$\bar{1}$2] azimuth produce single steps 0.31 nm high (Fig. 3B), those towards the opposite [112] azimuth create bunched steps (Fig. 3C). At the highest miscut angles towards [1$\bar{1}$2] one observes triple steps, such as for the 10° cut shown in Fig. 3A. As described above, the $7 \times 7$ terraces are divided neatly into quanta of 2.3 nm width by the stacking faults that are built into the $7 \times 7$ structure. The dense step structure in Fig. 3A represents a single such quantum, combined with a triple step. After adsorption of 0.2 monolayer of Au the single steps are restored and the terrace width goes down further from 6 to 2 nm, forming a Si(5 5 7)-Au structure (not shown; compare Ref. [30] and Section

Fig. 5. Surface state on the striped Si(111)5 × 2-Au structure (compare Fig. 6). Its character changes continuously from one-dimensional to two-dimensional [43]. (a) Angle-resolved photoemission spectra of clean Si(111)7 × 7 and Au-covered Si(111)5 × 2, showing a strong surface state in each case. (b)–(d) Constant energy contours of the surface state in momentum space for Si(111)5 × 2-Au. The horizontal contour at the top of the band in (d) is one-dimensional, because it does not depend on the emission angle perpendicular to the stripes $\theta \perp$. It can trigger a Peierls distortion that doubles the period along the stripes and creates a new zone boundary (ZB$^{\perp 2}$). In (c) and (b) the contours become gradually more two-dimensional, extending diagonally in both emission directions.
3). A perfectly regular step spacing is achieved for the narrow terraces (Fig. 3A), where the interaction between steps is largest. While the steps are still extremely straight for wider terraces on 1° miscut samples, there is some variation in their spacing (Fig. 3B).

As mask material we have chosen calcium fluoride (Figs. 2 and 4). It is lattice-matched with silicon to 0.7% at room temperature, does not intermix at the interface, and is chemically inert. Two types of atomically sharp interfaces exist [20–22], a fluorine-terminated interface when annealing below 650°C, and a calcium-terminated interface after heating above 750°C. The low-temperature interface contains a full F–Ca–F triple layer. At about 700°C the F-layer adjacent to Si desorbs and allows the formation of a Si–Ca bond at the interface which takes up the second valence of Ca. This Si–Ca–F layer has an optical band gap of only 2.4 eV [20–22], as opposed to the 12 eV gap of CaF_2. The difference in the band gap can be detected via scanning tunneling spectroscopy [24] by tunneling into the respective conduction band minima of CaF_1 and CaF_2.

Images of the tunneling current at the corresponding bias voltages identify CaF_2, CaF_1, and Si with 1 nm resolution. An array of CaF_2 stripes is obtained after depositing between 1 and 2 monolayers of CaF_2 and forming the Ca-terminated interface at 830°C. The result is a complete CaF_1 overcoat on the Si(111) surface with CaF_2 stripes growing on top of the steps (Figs. 2B and 4). The stripes are continuous and completely separated from each other. These are the requirements for producing continuous and insulated wires. CaF_2 stripes keep avoiding each other even after decreasing their separation to 1 nm with a coverage close to two monolayers. That can be explained by the fact that the Ca-terminated interface rotates the lattice by 180° about the Si(111) surface normal, which is equivalent to a reversal of the stacking from ABC to CBA [20–23]. As a result, calcium fluoride stripes on adjacent terraces are laterally mismatched and avoid each other. By growing on top of a step edge the CaF_2 in the outer layer avoids making contact with the CaF_1 layer next to it on the upper terrace.

Strings of CaF_2 dots are obtained by depositing 1/5 of a monolayer of CaF_2 onto stepped Si(111), as shown in Fig. 2C [23,25]. This growth mode occurs at lower temperatures (650°C) where the interface is F-terminated. The 180° lattice rotation is absent for this interface, and the CaF_2 attaches itself to the bottom of the steps, not the top. The dots are fairly homogeneous in size (about 7 × 10 nm^2, a single CaF_2 layer high). These arrays are denser than most dot arrays fabricated so far [32–34] and have the advantage of being neatly lined up along the step edges. Such an arrangement looks promising for data storage media with linear readout schemes, such as a scanning probe or a shift register (see Section 4). Above a critical coverage of 1/2 monolayer the CaF_2 dots convert into stripes. Their aspect ratio changes by two orders of magnitude within a few tenths of a monolayer. The drastic change of the growth mode is explained by the influence of the Si substrate steps. Elastic interaction between the steps and island edges parallel to the steps is repulsive and stronger than the effective repulsion between the islands. At higher coverage an island can minimize the elastic energy increase by growing along a step rather than towards the adjacent step, which effectively keeps the width of the islands constant while increasing their length and aspect ratio. The step repulsion is in fact so strong that nearly all the islands have merged into continuous stripes before they continue to grow towards the adjacent step.

Both of the growth modes of CaF_2 on stepped Si(111) in Fig. 2 defy conventional wisdom, which would favor either step flow or two-dimensional Stranski–Krastanov growth for a material with low surface energy, such as CaF_2 [5]. In both cases the CaF_2 would be attached to the bottom of
the steps, not the top. If dots were to occur (such as in Stranski–Krastanov growth) they would never convert into stripes at higher coverage. Such unexpected growth modes create a rich repertoire of options for growing new types of nanostructures by self-assembly.

The last step in the emulation of microlithography is the selective deposition of a material into areas not covered by the mask. First results are shown in Fig. 4 for the selective deposition of organic molecules (A) and of Fe nanowires (B). Various organic molecules adsorb selectively onto the more reactive CaF₂ trenches [35,36]. Some of these molecules can be used as precursors for chemical vapor deposition of metals, such as ferrocene for the laser-assisted deposition of Fe [37]. At room temperature, the molecules adsorb everywhere (not shown), but at elevated temperature (100–300°C) they accumulate in the CaF₂ trenches which allow them to form a stronger bond. This selectivity is explained by the band gap of CaF₂ (2.4 eV) which matches the gap between the occupied and unoccupied molecular orbitals (typically 3 eV). CaF₂, on the other hand, has a very large band gap of 12 eV which makes it chemically inert. The ferrocene molecule consists of a Fe atom sandwiched between two cyclopentadienyl rings (C₅H₅). Exposing selectively adsorbed ferrocene molecules to ultraviolet photons from a nitrogen laser dissociates the organic ligands from the Fe without leaving other organic fragments. The Fe atoms remaining in the CaF₂ trenches act as adsorption sites for additional ferrocene and keep the selective deposition process going. Similar metallocene molecules can be used to deposit the other ferromagnetic transition metals Co and Ni with a minimum amount of carbon and oxygen incorporated. Such structures are of interest for magnetoelectronics [5], as explained in Section 4.

3. Low-dimensional electronic states

One of the primary goals of fabricating nanostructures by self-assembly is the creation of new materials with tailored electronic properties. That can be accomplished by confining electrons to structures comparable to their wavelength, such as in a quantum well. In fact, two-dimensional character is an essential ingredient in the few nanostructures that have been incorporated into electronic devices already. One of them is the quantum well laser, another the latest generation of reading heads for hard disks. The former utilizes a quantum well to keep electrons and holes confined to the same region in space, in order to enhance their radiative recombination rate. The latter uses the spin-dependent reflection of electrons at an interface between magnetic and non-magnetic metals. This spin filter effect generates a giant magnetoresistance (GMR) that is able to sense the weak magnetic field emanating from a stored bit [5].

Here we expand the scope to the exotic electronic states that are expected in lower dimensions, and particularly in one-dimensional objects, such as arrays of stripes on a surface. Already in a very naïve, classical picture the electrons in a one-dimensional wire are highly correlated, because they cannot avoid crossing each other’s path when traveling in opposite directions. More detailed, quantum-mechanical analysis shows that the concept of a single electron as a well-defined quasiparticle breaks down and is replaced by collective excitations resembling charge and spin density waves. In certain parts of parameter space, they might be loosely viewed as an electron breaking up into two quasiparticles, one containing its charge degree of freedom (the holon) the other its spin (the spinon) [38–40]. The traditional approach to testing such predictions has relied on three-dimensional crystals containing chain structures, such as chains of transition metal ions spaced by counter-ions or conjugated polymer chains. In order to go beyond this relatively limited class of compounds given to us by Nature there have been recent attempts at designing one-dimensional chain structures at surfaces with the goal of customizing their electronic structure, such as controlling the chain–chain interaction, modifying the band filling, and preventing a Peierls distortion which wants to convert one-dimensional metals into semiconductors by doubling the unit cell.

Semiconductor substrates have played a prominent role because they naturally exhibit long-range reconstructions, due to the tendency of the broken surface bonds to rearrange themselves into stable configurations. The existence of a gap allows for truly one-dimensional states that do not couple to bulk states. On Si(111) there exist several one-dimensional reconstructions of the type $n \times m$, where $n$ is the number of atomic rows between the chains and $m$ is either 1 or 2, the latter for chains that might contain a Peierls distortion. Examples are $5 \times m$-Au [30,40–43], $4 \times m$-In [44,45], and $3 \times m$ Ag, Li, Na, K, Mg [46]. Figs. 5 and 6 show the electronic structure and topography of the Si(111)$5 \times 2$-Au structure. For the analysis of the electronic states we ignore the 1/40 of a monolayer of additional atoms residing on top of the stripes (white dots in Fig. 6). These will become interesting during the discussion of advanced storage media in Section 4. The underlying stripes give rise to an intense surface state, as shown in Fig. 5A [43]. It has a very peculiar band dispersion, which exhibits one-dimensional character at the top of the band (near the Fermi level), but gradually becomes two-dimensional when moving towards the bottom (Fig. 5B). The one-dimensional section at the top of the band is located at the appropriate point in the Brillouin zone (ZB) to trigger a Peierls distortion from $5 \times 1$ to $5 \times 2$ and render the surface semiconducting. It is possible to prevent the transition to a semiconductor at a highly stepped Si(111) surface where each $5 \times m$ stripe is combined with a step. Metallic behavior has been observed [40,47] and spin-charge separation has been proposed [40] to explain the observation of two features in the photoemission spectrum where only one surface state was expected. This is the Si(557)-Au surface that came up in Section 2.
already when searching for the Si(111) surface with the minimum step spacing [29,30]. Thus, the idea of locking the chain atoms to the rigid silicon lattice at a step and preventing the Peierls distortion could lead to new types of electron liquids.

4. Evolution towards devices

The results obtained so far provide a proof of principle that the strategy of emulating silicon microlithography can be extended all the way down to the to the single digit nanometer regime. The next step will be an expansion of the repertoire of techniques. Eventually, one would like to explore what types of applications can take advantage of the special electronic features of self-assembled nanoarrays.

One of the microlithography techniques neglected so far in our work is selective etching. The CaF₂ masks can be expected to be fairly stable against halogen attack, suggesting various approaches of dry etching. Halogen etching of semiconductors has been studied at the atomic level [48], and a variety of other etching methods at the 10–100 nm level are under development.

Beyond traditional microlithography one may look to biochemistry for new self-assembly techniques. Short snippets of synthetic, single-stranded DNA have been used as ‘rubber bands’ to connect Au nanoclusters to each other [14–16] and to attach a DNA molecule to an electrode or to a scanning probe tip. For connecting two nanoparticles one attaches DNA strands with complementary sections and lets them form a double-strand. We envisage such techniques to become quite useful for arranging mono-disperse nanoclusters into linear arrays at steps. For such a program it is necessary to make silicon surfaces bio-compatible. There are standard methods to achieve that, such as thiol chemistry on gold [14–16] or siloxane chemistry on SiO₂. We have pursued the former method and tested whether it is able to preserve step features at the nanometer level [49,50]. In order to obtain smooth, continuous Au films on stepped Si(111) it is necessary to start with a wetting layer of a reactive transition metal, such as Ti. After depositing >10 monolayers of Au it is possible to attach alkane chains and short DNA segments that are functionalized with a thiol group to bond to the Au [49,50]. These rod-like molecules even exhibit a preferred orientation perpendicular to the Au/ Si surface. Other organic molecules, such as liquid crystals, can be oriented on stepped Si(111) in a direction parallel to the step edges [51].

Among the initial applications could be patterned data storage media for high-density data storage, such as in hard disks. For this purpose one would like to have regular arrays of identical magnetic particles. Linear arrangements, such as particles attached to straight step edges, would facilitate readout by a scanning probe. As mentioned in the introduction, the ultimate goal in magnetic data storage is a single magnetic particle per bit, with a size close to the superparamagnetic limit of 3–10 nm. Current commercial hard disks reach a density of close to 20 Gbits/in.² (using units familiar to the storage industry). The array of CaF₂ dots in Fig. 2C has a density of 2 Tera-dots/in.², with a dot size of 7 × 10 nm². That would be right at the superparamagnetic limit if the dots could be used as masks for producing magnetic particles. Another approach would be to attach perfectly magnetic clusters [3,4] to the step edges, possibly via complementary DNA snippets [14–16]. Going far beyond such extensions of traditional storage media one might speculate about storage at the atomic limit, i.e. using a single atom per bit. Fig. 6 shows a naturally formed chain structure formed by gold on silicon stepped Si(111), which resembles the micrograph of a CD-ROM. The main difference is the scale (nm versus μm), which produces a 10⁶ time higher density of 300 Tera-atoms/in.². The Si(111)5 × 2-Au structure discussed in Section 3 forms the red stripes in Fig. 6, which are reminiscent of a formatted disk with well-defined tracks. The silicon surface self-assembles itself into such a structure and allows extra atoms to occupy 5 × 4 lattice sites randomly (white spots in Fig. 6). These would correspond to the stored bits. They are stable at room temperature and prefer a coverage of about half of the allowed sites, independent of gold coverage. That would correspond to an encoding scheme with equal numbers of bits with 0 and 1.

The big issue with such high storage densities is the speed of writing and reading the information. The proven techniques for writing and reading on the atomic scale involve typically a scanning tunneling microscope, either for moving atoms or for displaying their position. The speed of an STM is many orders of magnitude slower than that of a magnetic reading head, requiring the use of massively parallel tip arrays (>10⁶ tips). Arrays are beginning to be built, with the record standing at about 10³ tips per array [52]. An important feature of our dot arrays is their linear architecture induced by the silicon steps. That facilitates readout via a tip array scanning parallel to the steps. The tips will automatically encounter the stored data along the tracks. Other linear architectures might also be considered, such as electronic readout via a shift register. The problem with readout speed at high storage densities appears to be rather fundamental. The energy that can be stored in a bit decreases with the size of the bit and requires longer time integration to preserve the necessary signal-to-noise ratio.

Attempting to proceed from storage media to data processing, such as in molecular computers, will be difficult. Self-assembled devices are prone to errors. One will have to rely on software correction schemes that route signals around defective hardware, such as those demonstrated with a conventional computer containing defective parts [1]. Therefore, it might be wise to take advantage of the sheer numbers in a vast array of nanoprocessors and to select tasks that require massively parallel processing but do not critically depend on complete error correction. Examples might be found in the area of pattern recognition, such as
recognizing speech, handwriting, and a person’s face. These complex tasks have to be solved quickly, but without requiring every single bit to be processed free of error.

Acknowledgements

This work was supported by NSF under Award Nos. DMR-9815416 and DMR-9632527. Part of it was conducted at the Synchrotron Radiation Center, University of Wisconsin-Madison, which is supported by the NSF under Award No. DMR-0079983.

References

[29] A. Kirakosian et al., submitted for publication.