SILICON NANOWIRES AS BIOLOGICAL SENSORS AND HIGHLY EFFICIENT THERMOELECTRIC MATERIALS

Thesis by
Yuri Leonid Bunimovich

In Partial Fulfillment of the Requirements for the
Degree of

Doctor of Philosophy

California Institute of Technology
Pasadena, California
2007
(Defended May 17, 2007)
© 2007

Yuri Leonid Bunimovich

All Rights Reserved
Acknowledgments

I owe much to my family for their unwavering support. My beautiful wife, Olya, is my constant source of inspiration. My parents and my sister always encourage me and elevate my spirit. My grandparents, who have gone and who would have really liked to see me graduate, still give me strength.

I have been very fortunate to be surrounded by kind and talented people at Caltech. I would like to thank my advisor, Jim Heath, for giving me unprecedented freedom and support to pursue various projects in the last five years. It has been a pleasure to interact with my friends and colleagues at the laboratory. Specifically, I have worked closely and productively with Young Shik Shin and Michael Amori on biological sensors and Akram Boukai on thermoelectrics. Working with them has been fun and extremely rewarding.

I would like to thank Caltech and UCLA Medical School for providing me with a wonderful learning environment over the past seven years. UCLA MSTP program has been extremely flexible and accommodating. Our collaborators at Caltech, UCLA and Institute for Systems Biology are fantastic scientists who have tremendously enriched our work. Bill Goddard and his group elevated our understanding of silicon thermoelectrics to a different level. Leroy Hood, Adrian Ozinsky and other truly wonderful biologists at ISB have been very open and sincere in sharing ideas.

Finally, I want to give a shout to my friends from MSTP and medical school, those who have already graduated and those who will graduate soon. Best of luck to all of you guys, we are almost done!
Abstract

Silicon nanowires are of significant interest because of their novel properties which afford new functions. Here, we study silicon nanowires fabricated via a well established top down approach called superlattice nanowire pattern transfer (SNAP). In the first part of the thesis, nanowires are utilized for biological sensing of DNA and proteins in an electrolyte solution. Important electronic and surface properties are considered as means to optimize the device sensitivity. The removal of silicon-oxide interface is shown to improve the limit of detection by two orders of magnitude. The sensitivity can be further improved by the reduction of the doping level to $10^{17}$ cm$^{-3}$. In this way, sub-femtomolar concentration of oligonucleotides in physiological conditions can be detected. While the Debye screening is circumvented by the electrostatic adsorption of primary DNA on the amine-terminated monolayer, the detection of proteins is limited by the size of the antibodies. In low ionic strength solution, $\sim 10\mu$M, human IL2 cytokine is detectable at 1 to 10pM concentrations. Furthermore, a model is developed which allows the determination of kinetic parameters and absolute analyte concentrations from the real-time resistance of the nanowires. This model is consistent with Langmuir model, and could, in principle, be used to determine the amount of low abundance biological molecules at concentrations below those detectable with other label-free methods, such as surface plasmon resonance technique. In addition, a novel electrochemical technique is developed which allows the spatially-selective functionalization of silicon nanowires and the construction of a small library of proteins. In the second part, the discovery of highly efficient thermoelectric materials based on silicon nanowires is discussed. A relatively simple, scalable, and single component system of silicon nanowires with figure of merit of $\sim 1$ at room temperature is
developed. ZT can be tuned at various temperatures to exceed unity by varying nanowire size and/or impurity doping level. Such enhancement in ZT compared to the bulk value is achieved by significantly perturbing the phonon-mediated heat transport in a nanowire. Decreased thermal conductivities and longer lifetimes of long-wavelength phonons in a nanowire are major reasons for an increased thermoelectric efficiency of these structures.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>1. Introduction</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2. Fabrication of Ultra-High Density Silicon Nanowire Arrays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>Supperlattice Nanowire Pattern Transfer (SNAP)</td>
<td>8</td>
</tr>
<tr>
<td>2.3</td>
<td>Diffusion Doping of Silicon Thin Films</td>
<td>13</td>
</tr>
<tr>
<td>2.4</td>
<td>Two-Dimensional Nanowire Circuits</td>
<td>17</td>
</tr>
<tr>
<td>2.5</td>
<td>Discussion</td>
<td>25</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>3. Silicon Nanowires for Real-Time, Label-Free Biological Sensing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>29</td>
</tr>
<tr>
<td>3.2</td>
<td>Experimental Methods</td>
<td>33</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Nanowire Sensor Fabrication</td>
<td>33</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Surface Functionalization and Characterization</td>
<td>38</td>
</tr>
<tr>
<td>3.2.3</td>
<td>SPR and Electronic Measurements</td>
<td>41</td>
</tr>
<tr>
<td>3.3</td>
<td>Single-Stranded DNA Sensing</td>
<td>43</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Nanowire Surface Passivation</td>
<td>43</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Nanowire Impurity Doping Level</td>
<td>52</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>3.4</td>
<td>Protein Sensing</td>
<td>56</td>
</tr>
<tr>
<td>3.5</td>
<td>Quantitative Analysis of Nanowire Response</td>
<td>67</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Analysis of DNA-Sensing Experiments</td>
<td>67</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Analysis of IL2-Sensing Experiments</td>
<td>73</td>
</tr>
<tr>
<td>3.6</td>
<td>Discussion</td>
<td>82</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>85</td>
</tr>
</tbody>
</table>

4. Electrochemically Programmed, Spatially Selective Functionalization of Silicon Nanowires

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>90</td>
</tr>
<tr>
<td>4.2</td>
<td>Experimental Methods</td>
<td>92</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Materials</td>
<td>92</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Organic Synthesis</td>
<td>93</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Surface Functionalization</td>
<td>95</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Electrochemistry, XPS, and Optical Microscopy</td>
<td>97</td>
</tr>
<tr>
<td>4.3</td>
<td>Functionalization of Hydrogen-Terminated Si(111) and Si(100) Surfaces with Hydroquinone</td>
<td>98</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Characterization of Electroactive Organic Monolayers</td>
<td>98</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Electrochemical Oxidation of Silicon: Organic Monolayer Density and Surface Orientation</td>
<td>103</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Diels-Alder Reaction and Michael Addition on Silicon</td>
<td>108</td>
</tr>
<tr>
<td>4.4</td>
<td>Selective Functionalization of Silicon Micro- and Nanowires</td>
<td>113</td>
</tr>
<tr>
<td>4.5</td>
<td>Discussion</td>
<td>116</td>
</tr>
</tbody>
</table>
4.5.1 Mixed Monolayers and Biofouling Minimization .......... 116

4.5.2 Reductive Functionalization ................................................. 119

Bibliography ................................................................................................... 122

5. Silicon Nanowires as Highly Efficient Thermoelectric Materials

5.1 Introduction ................................................................................................. 128

5.2 Experimental Methods ................................................................. 134

5.2.1 Thermoelectric Device Fabrication ...................................... 134

5.2.2 Electronic Measurements ................................................................. 139

5.3 Temperature Dependence of Silicon Nanowire

Thermoelectric Properties ................................................................. 148

5.3.1 Electrical Conductivity ................................................................. 148

5.3.2 Thermal Conductivity ................................................................. 153

5.3.3 Thermopower and Phonon Drag ............................................. 156

5.4 Discussion ................................................................................................. 161

Bibliography ................................................................................................... 165
Chapter 1

Introduction

When an object is described, its size often is the starting point of the description. This is not surprising, since the importance of physical scale is cross-disciplinary. Viruses are the smallest living organisms and are able to live inside other cells undetected. Diffusion rate of a molecule or a subatomic particle is related to its size and in turn may control the pressure or electrical conductivity. Size is intimately related to function. Dramatically changing the size of an object almost always alters its function by causing new physical and chemical properties to emerge. Accessing a new scale, therefore, often means obtaining new functions. While the worlds of biochemistry and organic chemistry revolve around the nanoscale, inorganic solid-state materials have only relatively recently been able to tap into these dimensions. While tremendous progress in the synthesis and research of nanomaterials has been made, relatively few truly useful properties have been discovered and even fewer successfully applied. This is, in part, because nanoscience is largely cross-disciplinary, and major challenges occur at the interfaces of various fields, such as material science and surface chemistry.

This thesis is a reflection of the merging of different fields to obtain and harness useful properties afforded by the nanostructured materials, in particular silicon. High aspect ratio silicon nanowires (SiNWs) comprise a central theme of this work, and each chapter tackles a particular property characteristic to these materials. In chapter two, I introduce a new fabrication technique, called silicon nanowire pattern transfer (SNAP), for making silicon nanowire arrays of ultrahigh density. This method is characterized by
its tremendous versatility and robustness. It affords unprecedented control over most of the important physical and electronic parameters of a one-dimensional semiconductor nanomaterial, such as diameter, length and doping concentration. SNAP is a perfect platform on which to build systematic studies not only of silicon nanowires, but nanowires of a broad range of materials.

Chapter three addresses the use of SiNWs for the sensing of biological molecules in an electrolyte solution. High surface-to-volume ratio of SiNWs causes the surface charges to exert a significant effect on the electrical conductivity of the nanowires. The NWs can essentially act as solution field-effect transistors (FETs) whose current is modulated by the surface binding of charges biological species. Since the binding event may be directly electrically detected without secondary signal amplification, which is often required for optical detection, label-free real-time sensing is possible. Few sensing techniques are real-time and label-free; therefore, the superior sensitivity and extreme scalability of SiNWs makes these sensors unique for many applications, such as the detection of secreted proteins from single cells, for example. The doping level and surface chemistry are two of several important issues addressed in this chapter for proper engineering of SiNWs to detect oligonucleotides and proteins. While femtomolar sensitivity of DNA can be achieved in high ionic strength (physiological) solution, the detection of proteins is limited by the size of antibodies, which are by far the most widely used high-affinity protein capture agents. The Debye screening by the ions, therefore, prevents the sensing of proteins in the physiologically relevant solution. Alternatives to antibodies, such as aptamers, small molecules or peptides may eventually circumvent this limitation. Finally, I address the method of quantititating the response of SiNW sensors.
A model is developed which may be used to extract kinetic parameters as well as an absolute analyte concentration. This feature, unique to a label-free and real-time method of detection, significantly elevates the possible usefulness of SiNW-based sensors.

Chapter four discusses a novel, spatially selective method of functionalizing silicon micro- and nanostructures based on electrochemical programming. A monolayer of hydroquinone molecules is grown on a hydrogen-terminated silicon surface via a UV-catalyzed hydrosilylation method. An electrochemical oxidation of hydroquinone to benzoquinone allows selective attachments of secondary molecules to the benzoquinone via Diels-Alder or Michael addition. Spatial encoding can be carried out at a resolution far beyond the optical diffraction limit, since this method is only limited by the ability to electrically address individual silicon nanostructures. The motivation is that, in principle, individual SiNWs produced via SNAP method and separated by as few as 35 nm may be individually encoded. This would allow a construction of hundreds of electrical sensors in a 100 $\mu$m$^2$ area for the purpose of parallel multiparameter analysis. However, the method is by no means limited to this application, and, in fact, could be easily extended to other Si-based micro- and nanodevices or even to other electrode materials. I address in some detail the ways to minimize electrochemical oxidation of silicon, either through controlling the molecular packing density or utilizing the electrodes with different crystal orientations. Finally, an alternative approach is presented, based on electrochemical reduction, which circumvents the electrochemical oxidation of silicon electrodes altogether.

In chapter five, I diverge from the applications of SiNWs to biosensing, and concentrate on the fundamentals of their thermoelectric properties. The motivation
behind this work is to create a highly efficient thermoelectric device out of nanostructured materials. SiNWs are chosen because of our extensive expertise in their fabrication, and because of all the advantages afforded by SNAP which are described in chapter one. Also, heavily doped semiconductor materials are the usual choice for thermoelectric applications because of their relatively high electrical conductivities and Seebeck coefficients and phonon-dominated thermal conductivities. Systematic analysis of the effects of size and doping on the figure of merit (ZT) of SiNWs is presented. Phonon transport can be markedly altered by one-dimensional confinement, leading to a precipitous drop in thermal conductivity of SiNWs compared to bulk silicon and a corresponding increase in the lifetimes of long-wavelength phonons. In addition, the electrical conductivity of NWs can be maintained at a relatively high level, with the bulk-like thermopower. These factors lead to a figure of merit which approaches 1.0 at room temperature and can be tuned across different temperatures. In addition, phonon drag is observed in heavily, but not degenerately, doped 20 nm wide NWs, leading to a significant increase in thermopower of these materials at temperature in the range of 150 K to 250 K. The highest ZT reported here, 1.2, was measured in 20 nm wide SiNWs at 200 K. This study, therefore, clearly demonstrates that the field of thermoelectrics is no longer limited by the discovery of new highly efficient and exotic materials. Instead, modified phonon transport properties which emerge from nanostructuring silicon, or, in principle, many other materials as well, lead to a dramatic enhancement of the thermoelectric figure of merit. ZT values above 3 may be possible in nanowires, which would make thermoelectrics competitive with conventional compressor-based refrigeration.
Chapter 2

Fabrication of Ultra-High Density Silicon Nanowire Arrays

2.1 Introduction

Nanofabrication patterning of molecules as well as metals and semiconductor structures has received increasing attention. A major challenge in the development of robust nanopatterning techniques involves the reduction of feature size and the increase in pattern density. High-density nanoscale circuits have the potential to be more efficient and faster than the conventional electronic circuits. In addition, new scientific possibilities arise when the patterning techniques begin to approach the scales and densities of macromolecules, likely offering a feasible organic-inorganic interface to biological systems. In particular, nanowire (NW) circuit fabrication has become a very active field of research. To date, NWs have been used to fabricate modular circuit elements such as field-effect transistors (FETs), bipolar junction transistors, p-n diodes, logic gates, lasers and LEDs, molecular memory, and nanoscale electro-mechanical resonators. In addition to NWs, the electronic, thermal and optical properties of single-walled carbon nanotubes (SWNTs) have been extensively studied. In this chapter, I concentrate the discussion on the development of a robust and versatile technique to fabricate high-density silicon nanowire arrays. Significant
advantages afforded by such method lead to a variety of applications which are also briefly addressed here.

Silicon is an attractive material for multiple electronic applications. The physical and mechanical properties of silicon have been extensively studied and are well characterized. In addition, fabrication protocols for etching, patterning and electrically addressing silicon devices are highly developed. Finally, silicon-based devices are easily integrated into a conventional CMOS technology, thus providing an opportunity for large-scale manufacturing and commercialization. These reasons have sparked an extensive research into the fabrication techniques of silicon NWs (SiNWs) and their incorporation into functional devices. One such technique, which, at this time, is the most widely used method of SiNW fabrication for electronic applications, is the nanoparticle-catalyzed vapor-liquid-solid (VLS) growth mechanism.\textsuperscript{16-18} Briefly, a gold nanocluster is heated in the presence of vapor-phase silicon (SiH\textsubscript{4} in H\textsubscript{2}) to 362 °C, a temperature of Au-Si eutectic, resulting in the formation of Au-Si alloy in the form of a liquid droplet. The droplet supersaturates in Si under a continuous flow of vapor Si, resulting in a precipitation of solid silicon. SiNW continues to grow from the solid-liquid interface as long as the droplet is supersaturated. This process may be tuned to produce single-crystal SiNWs with well-controlled diameter and growth orientation.\textsuperscript{19} However, this process also has major drawbacks, which limit its usefulness as a large-scale fabrication tool. One significant problem is the lack of precise control over the doping level. The VLS doping is done \textit{in situ} by adding dopant precursor to the reaction, yielding NWs with doping levels and electrical properties which are largely unknown prior to their integration into a device. Additionally, while the VLS technique allows the
fabrication of many NWs at once, subsequent methods are required to align NWs into parallel arrays or crossbar circuits. One such method which has been the most successful in aligning VLS grown NWs employs a Langmuir-Blodgett (LB) trough. Aside from being impractical for scaling up and manufacturing, LB method produces NWs with large alignment fluctuations and poor end-to-end registry of individual NWs. Interconnection and integration of such SiNW arrays into CMOS compatible circuitry is difficult.

As described above, a major challenge in nanoelectronics is to be able to reproducibly fabricate high-density NW arrays with precise control over the diameter, length, pitch and doping level. Superlattice nanowire pattern transfer (SNAP) technique presented here, combined with the spin-on doping method (SOD), allows such level of control over multiple physical and electronic parameters of SiNWs, and may be extended to the fabrication of NWs from a wide range of materials. The other challenge is higher-order control of the device architecture, with arbitrary complex interconnections within an array of nanostructures. While certain ordered arrangements, such as nanowire crossbar structures, are useful from the point of view of electronic circuit construction, many other applications require at least some level of nonperiodic complexity. Examples include routing networks between nanowire field-effect transistors, or specified defects within photonic crystals. An extension of SNAP technique may be used to construct arbitrarily complex, two-dimensional nanowire structures templated on a single crystal substrate at sublithographic dimensions. As a proof-of-principle demonstration, this method is used to construct routing networks for Si nanowire-based complementary
symmetry logic applications with electrical contacts which are superior to those fabricated with other methods.

### 2.2 Superlattice Nanowire Pattern Transfer (SNAP)

Superlattice nanowire pattern transfer (SNAP)\textsuperscript{13} is a technique for fabricating aligned nanowires of a wide range of materials through a one-step deposition process without subsequent etching or liftoff, which is often necessary for other methods such as e-beam lithography (EBL).\textsuperscript{22} These fully formed NWs may be transferred to any surface. Further processing steps can be utilized to convert these nanowires into an identical pattern out of a thin-film material, such as silicon on insulator (SOI). SNAP technique has been developed to simultaneously address issues of NW size, pitch and alignment. NWs can be fabricated from a thin film of arbitrary material, as long as dry etching of such material is possible. Therefore, the electronic properties of the NWs produced via SNAP may be manipulated by a corresponding control of the properties of the starting thin film. For example, by combining SNAP technique with spin-on doping (SOD) method, a quantitative control of the NW doping level is possible.

Figure 2.1 outlines the SNAP fabrication steps. SNAP uses molecular beam epitaxy (MBE) to create a physical template for NW patterning. Such template is a GaAs/Al\textsubscript{x}Ga\textsubscript{1-x}As superlattice structure consisting of the alternating gallium-arsenide and aluminum-gallium-arsenide layers grown on top of a (100) GaAs substrate. The Al mole fraction may be varied between 0.5 and 0.8. The atomic-level control over the thickness and composition of each layer translates into atomic-level control over the resulting metal or semiconductor nanowires. The fabrication protocol begins by dicing a
Figure 2.1: SNAP fabrication flow-chart. A) A small piece of the GaAs/AlGaAs superlattice is selectively etched, B, forming a comblike structure (inset: SEM image). C) Platinum is deposited at an angle along the ridges of the comb. D) The superlattice template is adhered to an epoxy-coated SOI substrate. E) The superlattice is etched away, releasing the Pt nanowires. F) Dry etch transfers the Pt NW pattern into the underlying silicon film.

portion of the superlattice wafer into 2 mm by 5 mm pieces, or masters (Figure 2.1A). Each master has one side, \{110\} or \{001\} plane, which is atomically flat. With that side facing up, the masters are placed into a Teflon holder and sonicated in methanol for \(\sim10\) seconds. The flat edge is swabbed until all microscopic particles are removed.

The GaAs layers are selectively etched in a solution of NH\(_4\)OH/H\(_2\)O\(_2\)/H\(_2\)O, producing a comblike structure shown in Figures 2.1B and 2.2A. The thickness of the AlGaAs layer determines the NW width, while the GaAs thickness translates into the distance between the NWs. A metal, such as platinum, is then deposited at an angle along the ridges of the AlGaAs layers on the atomically flat edge of the master (Figure 2.1C). The angle is chosen so that no metal is deposited on the etched GaAs layers, and is, therefore, dictated by the pitch of the NWs and the depth of the grooves in the superlattice comb structure. In general, NW pitches between 30 nm and 60 nm will
require a deposition angle between 45° and 15° with respect to the horizontal axis. The thickness of the deposited metal depends on the desired width of the NWs, and is generally 10 nm if the NW width is larger than 10 nm and equals to the NW width otherwise.

\[ \text{Figure 2.2: A) Etched superlattice with comblike Al}_{0.5}\text{Ga}_{0.5}\text{As layers. B) Fifteen nanometers wide SiNWs. C) 20 nm wide Pt nanowires on SOI substrate. D) Twenty nanometers wide SiNWs remaining after pattern transfer of C and the removal of Pt nanowires. The scale bars and the NW pitch of B, C, and D are 50 nm and 30 nm, respectively.} \]

The metal-coated edge of the master is gently positioned on top of the epoxy-coated SOI surface which has been rigorously cleaned after the doping process described below (Figure 2.1D). The epoxy is heat curable, and consists of the mixture of Epoxy Bond 110 (Allied High Tech Products, Ranch Dominguez, CA, 10 drops part A to 1 drop part B) and 0.15 g of 6% PMMA in 20 mL of chlorobenzene. The epoxy is spun on the surface at 7000 RPM to form a film of approximately 10 nm thickness. Because the
master can be aligned on the surface at an arbitrary angle, SiNWs with crystal orientations between (100) and (110) along the length may be produced. The master-SOI substrate assembly is placed on the hot place at 135 °C for 45 minutes, after which it is left in the solution of 1:5:50 (v/v/v) 30% H₂O₂:H₃PO₄:H₂O for ~5 hours, until all AlGaAs layers are etched, releasing the Pt NWs. The remaining master is detached from the substrate and the epoxy is removed in O₂ plasma. The remaining Pt nanowires (Figure 2.2C) are L shaped due to angular deposition onto the etched superlattice.

Reactive ion etching (RIE) may be used to transfer the pattern defined by the resulting Pt nanowires into an identical pattern in the silicon epilayer of SOI substrate. This is done with anisotropic etch in CF₄/He plasma (20/30 sccm, 5 mTorr, 40 W), which etches silicon at a rate of ~10 nm/min. Subsequently, Pt nanowires are removed in hot aqua regia (3:1 v/v HCl:HNO₃ at 120 °C for 20 minutes). As Figure 2.2 (B,D) demonstrates, the resulting SiNWs are straight, perfectly parallel, isolated from each other and have relatively smooth side walls. Since these wires are etched from the epilayer of the SOI substrate, they are also electrically isolated from one another. The procedure to convert the Si nanowires into functioning devices for biological sensing and thermoelectric applications is described in details in the following chapters. It is critical to emphasize the versatility afforded by the SNAP technique. First, various metal NWs may be fabricated on any substrate; so far, the metals which have successfully yielded NWs are gold, chromium, aluminum, titanium, niobium, platinum and nickel. Second, NWs from any thin film material may be fabricated, provided that anisotropic etch for that material exists. Third, physical parameters such as the number of parallel NWs, their
sizes, pitches and lengths are precisely controlled. Finally, the electronic properties of SiNWs can be tightly modulated with quantitative doping of the initial thin silicon film.

**Figure 2.3:** A) An array of 1400 SiNWs 20 nm wide with 35 nm pitch. B) An array of 200 SiNWs 10 nm wide with 60 nm pitch. The inset scale bars are 100 nm.

The size of the array containing SiNWs is only limited by the MBE growth time, and could, in principle, be increased to thousands of NWs. Figure 2.3A demonstrates the largest array produced so far with SNAP technique, containing 1400 SiNWs with 20 nm width. The diameter of the NWs may also, in principle, be significantly decreased to about 1 nm or less, owing to atomic-level control of the MBE process. In practice, however, it has been rather difficult to reproducibly obtain SiNWs thinner than ~7 nm, although it certainly is not by any means the limit of the technique. Ten-nanometers-
wide NWs (Figure 2.3B) have been extensively studied in context of, for example, silicon thermoelectric properties (chapter five).

2.3 Diffusion Doping of Silicon Thin Films

As mentioned above, a major advantage of SNAP technique is that the starting material is a thin film, the electrical properties of which are much easier to control through doping than those of NWs grown via VLS technique. A large SOI surface can be uniformly doped and used to produce multiple SiNW arrays, all of which will have an identical concentration of impurities. This allows us to carry out systematic studies of the effects of doping concentration on, for example, the electrical and thermoelectrical properties of NWs. These studies generally require many NW samples whose doping levels are quantified and are the same. For this purpose we have utilized spin-on doping (SOD) technique, which is based on the thermally mediated impurity diffusion. SOD method has several very important advantages over other thin-film doping techniques, such as ion implantation. Ion implantation, which uses high energy ion flux, results in the lattice damage of the SOI substrate and the reduction the electrical conductivity of the NWs. The diffusion doping process, on the other hand, does not lead to the damage of thin silicon film. The other practical advantage of SOD technique is that the doping may be tuned over a large range of concentrations by simply varying the temperature. This may be readily done on small wafer pieces, introducing the flexibility which is lacking when large wafers have to be sent out for ion implantation. In addition, patterned diffusion doping of the thin-film substrate is necessary to create more complicated
circuits such as complementary logic gates, where n-type and p-type FETs must exist on the same device in close proximity.\(^8\)

The SOD procedure consists of four steps: wafer cleaning, application of the dopant, annealing and removal of the dopant. Wafers have to be thoroughly cleaned to remove any diffusion barriers on the surface which can result in non-uniform doping of the substrate. RCA clean is used: 5:1:1 \(\text{H}_2\text{O}/\text{H}_2\text{O}_2/\text{NH}_4\text{OH}\) for 10 min at 80\(^\circ\text{C}\), followed by the removal of oxide in dilute (50:1 \(\text{H}_2\text{O}/\text{HF}\)) hydrofluoric acid, followed by 6:1:1 \(\text{H}_2\text{O}/\text{H}_2\text{O}_2/\text{HCl}\) for 10 min at 80 \(\circ\text{C}\). The wafers are rinsed with \(\text{H}_2\text{O}\) and a thin film of dopants is spun on the surface. Most spin-on dopants consist of the desired species (such as phosphorus or boron) incorporated into a silica or organic polymer matrix dissolved in an organic solvent. For all the applications in this thesis, Boron A organic film (Filmtronics, Bulter, PA) was used to produce p-type substrates, while phosphorosilica (Emulsitone, Whippany, NJ) films generated n-type devices. Thin films of dopant were

\[\begin{align*}
\text{Figure 2.4:} & \quad \text{Impurity concentrations (cm}^{-3}\text{)} \text{ of p-type (circles) and n-type (triangles) substrates obtained from annealing at different temperatures for 3 min. After the annealing, p-type substrates were either thermally oxidized (grey circles) or simply washed with BOE (black circles).}
\end{align*}\]
baked on a hot plate for 10 min at 200 °C prior to the annealing. Rapid thermal annealer (RTA) was used to precisely control the temperature and time of the annealing, which was carried out under nitrogen. Usually, the temperature was varied and the time was maintained at 3 min to achieve different doping concentrations. Figure 2.4 summarizes the annealing conditions and the impurity concentrations they yield. After the annealing, phosphorosilica film was removed simply by immersion in BOE for a few seconds, resulting in a hydrophobic surface. Boron A film, however, could not usually be completely removed with BOE, as indicated by the hydrophilicity of the substrate. For applications such as biological sensing, where a hydrophobic surface is required at this step to assure the presence of exposed oxide later for subsequent surface functionalization, the substrates were thermally oxidized in RTP for a minute in pure oxygen at the same temperature as the previous annealing step. After the oxidation and BOE treatment, hydrophobic surface could be reliably obtained. The thermal oxidation, however, somewhat lowered the impurity doping (Figure 2.4), probably because of the oxidation and removal of the top few layers of the silicon.

The electrical resistivity of the thin film was measured with 4-point technique by applying a 10 mA DC current between the outer pins and measuring the voltage drop across the inner pins. The resistivity was converted to the approximate impurity concentration by using the reference tables generated for bulk silicon. The doping profile as a function of depth into the silicon epilayer was determined experimentally (Figure 2.5, blue triangles and red squares) by thinning a 50 nm Si epilayer in 10 nm steps via CF₄ etching, each time measuring the electrical resistivity of the film. Diffusion
Figure 2.5: Dopant density vs. depth for 50 nm SOI substrates with and without dopant drive-in. Si epilayers that were diffusion doped normally as described in the text (triangles and squares) show a rapid decrease in dopant density with depth as expected from calculations (dashed line). SOI subjected to an additional dopant drive-in as described in the text (circles) show a more homogeneous dopant density with depth.\textsuperscript{1}

doping produces an impurity gradient which falls rapidly away from the surface. The model of the infinitely thick Si epilayer (Figure 2.5, dashed line) agrees well with the experimental data.\textsuperscript{1} The inhomogeneous doping profile shown in Figure 2.5 has important implications for a variety of applications, as will be described later in this chapter. However, there are cases when a homogeneous doping is required. An additional annealing step is then used. The dopant film is removed as described above and \textasciitilde250 nm of SiO\textsubscript{2} is deposited (PECVD) on the substrate. Subsequently, the wafer is annealed at 1000 °C for 10-15 minutes to produce a homogeneous doping profile shown in Figure 2.5 (green circles).\textsuperscript{1}
2.4 Two-Dimensional Nanowire Circuits

As described above, high-density nanocircuits with more complex architectures such as a crossbar are usually fabricated either using EBL or, in the case of VSI SiNWs or SWNTs, utilizing Langmuir-Blodgett trough technique or an AC electric field, respectively. All these methods are serial and generally produce functional devices with low yield, making them sufficient only for one-off demonstrations. Being a top-down approach, SNAP may be easily extended to fabricate more complicated two-dimensional (2D) circuits for a variety of applications. For example, the production of 2D SiNW logic and routing circuits at sublithographic dimensions involves combining SNAP process with other patterning techniques such as e-beam lithography (EBL). The unprecedented dimension and density that can be achieved by the SNAP method for the production of aligned NW arrays is complemented by the arbitrary complexity that can be achieved via the incorporation of traditional patterning methods.

In a typical process, as demonstrated in Figure 2.6A, an array of Pt NWs is first deposited by the SNAP technique onto an SOI substrate as described above. The substrate is then further patterned using electron beam lithography to generate arbitrarily complex structures on top of the first set of Pt NWs. After the evaporation and subsequent liftoff of 20 Å Ti (as a sticking layer) and 100 Å Pt, a 2D structure consisting of Pt NWs and perpendicular Pt bars is obtained. Pt is then used as an etch mask to transfer the pattern into the underlying SOI epilayer via a directional dry etching with CF$_4$ gas. The entire process is concluded with the removal of Pt NWs and bars in aqua regia (Figure 2.6A). The resulting 2D Si structure consists of NWs connected by arbitrary routing bars, all formed from a single-crystalline layer of Si. Figure 2.6B shows an array
Figure 2.6: Two-dimensional SiNW arrays with monolithic connections. (A) Schematic illustrations of the fabrication process. Starting with Pt NWs fabricated by the SNAP method, e-beam or photolithography is used to create patterns on top of the Pt NWs (i). After metal evaporation and liftoff, a 2D Pt NW structure is formed (ii). This pattern is transferred into an SOI epilayer and the Pt mask is removed (iii). (B) An array of 400 SiNWs with monolithic connection bars. (C) An integrated inverter with SiNWs connected by monolithic bars (bottom inset). The Si/metal contacts (top inset) are between the monolithic Si bars (green) and metal leads (red). Scale bars of the insets are 500 nm.

of 400 20 nm wide SiNWs with a pitch of 33 nm. The NWs are connected by Si bars with various widths, ranging from 50 to 500 nm. Alternatively, instead of connecting them all at once, a subset of NWs can be selectively connected, as shown in Figure 2.6C. This structure can be used to fabricate, among other things, integrated electronic logic gates. The false color-coded green structures in Figure 2.6C are current routing bars,
which are single crystalline extensions of the SiNWs. They negate the need for direct electrical contacts to the nanowires. The device depicted in Fig 2.6C is a complementary symmetry inverter, or NOT gate.

The monolithic NW structure can be utilized to form high-quality and reliable electrical contacts which are crucial for any application such as high performance NW field effect transistors (FETs), thermoelectrics (chapter 5) or SiNW biosensors (Chapter 3). Despite enormous efforts by various researchers, obtaining reliable electrical contacts to SiNWs ~10-20 nm wide using metal pads remains a challenge. Because of the high surface-to-volume ratio, NWs are more sensitive to surface states than their bulk counterparts. Electrical contacts to the NWs, therefore, are significantly influenced by the surface states and their quality depends on particular steps in the fabrication process, such as contact annealing. Often, these necessary processing steps vary significantly for n-type and p-type Si FETs. Single-crystalline, 2D SiNW circuits may represent a universal solution to this challenging problem. First, the electrical contacts to the NWs are established through relatively large Si pads, rather than through the NWs themselves (Figure 2.6C). These types of metal/Si contacts are less sensitive to the surface states and have been extensively studied and optimized for conventional MOSFET fabrication. Second, the doping levels within contact regions and those of NWs may be separately optimized, as described later in the text. Finally, the routing Si bars can have an arbitrarily complex architecture (Figure 2.6C).

As described in the previous section, diffusion doping of SOI epilayer results in an impurity gradient, with highest doping concentration at the silicon surface (Figure 2.5). Depending on the doping conditions, the impurity density can decrease by a factor of 10
to 100 at a distance of ~10 nm below the top NW surface. Such sharp gradient can assist in separately modulating the NW and contact doping levels. High impurity concentration, in general, means that the Fermi level of the NWs is close to or in the conduction (n-type) or the valence band (p-type). As a result, the changes in the gate potential are inadequate to significantly alter the Fermi level, and the channel resistance of an FET does not change appreciably, yielding a poor gate modulation. Heavily doping NWs, on the other hand, results in a short depletion width within Si and a narrow Schottky barrier between metal and Si. Charge transfer through the contact region is, therefore, highly efficient. Such contact is less sensitive to the surface states due to the reduced depletion width. Overall, a heavily doped semiconductor forms good electrical contact but does not respond well to gate modulation. Conversely, the Fermi level of a lightly doped semiconductor is typically in the midgap, slightly deviating toward the conduction (n-type) or valence band (p-type). Small changes in the gate potential cause large shifts of the Fermi level and an efficient gate modulation. However, a lower doping level also means a longer depletion width, making the FET body very sensitive to surface and interface states. In general, it is difficult to establish good electrical contacts to lightly doped silicon.

An ideal FET structure exhibits different doping profiles in different regions: the source-drain (S-D) channel is lightly doped, while the S and D contact regions are heavily doped. Conventional FET structures are fabricated with such inhomogeneous doping profiles, which are achieved using ion implantation methods. To avoid damaging the NWs, alternative approaches are required to achieve similar doping profiles. In meeting this challenge, various approaches have been reported. These include the
formation of metal silicides at the NW S and D contacts\textsuperscript{28} or synthesizing NWs with axially heterogeneous doping profiles.\textsuperscript{29,30} While promising initial results exist, reliability in forming high quality contacts has not been demonstrated yet. Our approach to create inhomogeneous doping profiles along the FET channel involves a two-step process. FETs are first fabricated with the entire body, including the channel and semiconductor/metal interface, doped heavily (i.e. >10\textsuperscript{19} cm\textsuperscript{-3}). After the deposition of metal contacts, anisotropic etching with CF\textsubscript{4} gas removes the top (heavily doped) layers of Si. Metal contacts protect the S/D silicon regions, which remain heavily doped.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{selective_NW_channel_etching.png}
\caption{Selective NW channel etching to form lightly doped devices with heavily doped contacts. (A) As-made NW FETs with heavily doped channel. Device structure is schematically shown as the inset. The current is not modulated by the sweeping of the gate voltage (V\textsubscript{GS}). (B) After the selective etching, the doping level is greatly reduced in the channel region and the FETs can be fully depleted. The contact regions are still protected by the metal, as shown in the inset.}
\end{figure}

The effect of selective CF\textsubscript{4} etching described above is clearly demonstrated in Figure 2.7. Initially, the devices are highly doped and the source/drain (S/D) conductance changes little when the bottom gate is swept from -20 to 20 V (Figure 2.7A). After the etching, the reduced channel doping level yields a much better gate modulation: the S/D conductance changes by more than 3 orders of magnitude in the same range of
the gate voltage (Figure 2.7B). Importantly, this method is compatible with both p- and n-type FETs. The development of a single process which can be applied to both p- and n-type NW FETs is enabling for power-efficient complementary symmetry NW logic applications.\(^8\) Alternatively, one must rely on different annealing methods or other phenomenological treatments to get reliable contacts to NW FETs.\(^{28}\)

---

**Figure 2.8:** SiNW FETs with monolithic and heavily doped contacts. Electrical characterization of p-type 20 nm wide NWs: (A) I-V curves at different gate potentials (from -20V to 10V, 5V increments) and (B) current as a function of \(V_{GS}\). (C, D) Identical electrical characterization of n-type FETs. The gate range in (C) is from 20V to -10V, with -5V increments.

Combining the selective channel etching and the monolithic structures yields high quality electrical contacts (Figure 2.8). For both p- and n-FETs, the linear I-V behavior
in the low-bias regions at all measured gate potentials indicates that the contacts are ohmic. Corresponding current versus gate voltage traces reveal that the FET conductance can be modulated with high efficacy. Furthermore, this method is highly reliable and reproducible. For more than 100 devices, less than twenty percent variation in the on-current was observed from device to device.

In addition to constructing reliable, NW-based complementary symmetry logic circuits, the 2D patterning methods can be extended towards the production of a host of other novel and potentially useful structures as well. A first example is shown in Figure 2.9. Ultra-dense arrays of SiNWs can be suspended, thus allowing for the measurements of thermal properties of nanowires\(^{31}\) or for creating high-frequency NW resonators.\(^{12}\)
Figure 2.10: Fabrication of ultra-high density arrays of NW crossbars and holes in Si film. (A) NW crossbars are fabricated with two SNAP steps, which produce (B) two sets of NW arrays orthogonal to each other. Subsequent anisotropic dry etching (A) transfers the crossbar pattern into thin silicon epilayer, producing (C) an array of holes with ~10 nm in diameter.

When an array of aligned, closely spaced NWs is suspended over a trench, various processing steps, coupled with the van der Waals interactions between the NWs, inevitably cause NWs to collapse into bundles (Figure 2.9B). The integrity of the suspended NWs can be maintained by fixing them on either end with an in-plane monolithic structure of the type described above (Figure 2.9C). Nevertheless, it is worth pointing out that this method does have its limitations. When the trench is too wide, proximity of the NWs still may cause them to collapse together, even in the presence of the monolithic bars.
Multiple SNAP transfers can also be utilized in sequence to generate different three-dimensional circuits. For example, two orthogonal NW arrays may be constructed for applications in ultra-high density molecular memory based on the crossbar architecture (Figure 2.10B). Crossbar devices of this type may be further used to fabricate arrays of nanometer diameter holes in silicon with nanometer spacing (Figure 2.10). The nanohole arrays can be tuned in a number of ways. First, their diameter can be adjusted by using NW superlattice with a different pitch. Second, the spacing between holes can be tuned by using NWs of different widths. Third, their shape can be changed by aligning the two sets of NW arrays at different angles. The nanohole arrays are of some interest and potentially useful in photonic applications.

2.5 Discussion

As demonstrated in this chapter, Superlattice Nanowire Pattern Transfer (SNAP) is a very versatile technique for the fabrication of high-density nanowires from a wide range of materials, including silicon. Important advantages of SNAP over other SiNW fabrication methods, such as VLS technique, allow for systematic studies of the fundamental properties of nanowires which go beyond one-off device demonstrations. SNAP allows precise control over the number, width, spacing and length of the nanowires. In addition, perfectly parallel arrays of NWs are fabricated, eliminating the necessity for complicated alignment techniques such as those utilizing Langmuir-Blodgett trough. Complicated circuits, therefore, may be readily fabricated with high precision and yield. In addition, the top-down nature of SNAP means that the impurity levels of NWs can be carefully controlled by standard thin-film diffusion doping methods. As demonstrated here and in the following chapters, it is critical to be able to
systematically and quantitatively manipulate the dopant concentrations for optimizing the electrical properties of SiNW-based FETs, thermoelectrics and biological sensors. Using the traditional lithography in combination with SNAP yields 2D NW circuitry of arbitrary complexity. Reproducibly obtaining high quality contacts to SiNWs is a major challenge. However, by integrating an additional processing step, very high quality metal/Si contacts to 10 nm wide SNAP NWs can be fabricated. Robustness and versatility of the technique, ease of fabrication and the quality of the produced nanomaterial are central issues in nanoelectronics. SNAP exhibits superior performance in each of these categories.
Bibliography


Chapter 3

Silicon Nanowires for Real-Time, Label-Free Biological Sensing

3.1 Introduction

Over the past few years a number of new biomolecular sensors have been reported.\(^1\)\(^-\)\(^5\) The development of these devices is, in part, driven by the emerging needs of both systems biology\(^6,\)\(^7\) and personalized and predictive medicine\(^8\)–both of which are increasingly requiring quantitative, rapid, and multiparameter measurement capabilities on ever smaller amounts of tissues, cells, serum, etc. To meet these needs, many groups have focused their attention on developing real-time, highly sensitive and potentially scalable tools for detecting nucleic acids and proteins. One-dimensional nanostructures such as nanotubes,\(^9\)\(^-\)\(^12\) semiconductor,\(^13\)\(^-\)\(^15\) metal oxide nanowires (NWs),\(^16\) and conducting polymer nanofilaments\(^17\) have all been demonstrated as capable of the label-free detection of small molecules, nucleic acids, and proteins.

Silicon nanowire (SiNW) biosensors are promising label-free, electronic-based detectors of biomolecules.\(^2\) However, significant scientific challenges remain before SiNW sensors can be viewed as a realistic technology. One challenge relates to the use of these devices in biologically relevant media, which is typically a 0.14M electrolyte. NW sensors detect the local change in charge density (and the accompanying change in
local chemical potential) that characterizes a target/capture agent binding event. That changing chemical potential is detected as a ‘gating’ voltage by the NW, and so, at a given voltage, affects the source (S) \(\rightarrow\) drain (D) current value, or \(I_{SD}\). However, that change is screened (via Debye screening) from the NW by the solution in which the sensing takes place.\(^{18, 19}\) Debye screening is a function of electrolyte concentration, and in a 0.14M electrolyte (which represents physiological environments such as serum) the screening length is about 1 nm.\(^{20}\) Because of this, all reports on SiNW sensors for proteins or DNA have been carried out in low ionic strength solutions.\(^{14, 15, 21}\) In this chapter, we demonstrate that a single-stranded complementary oligonucleotide is able to significantly change the conductance of a group of 20 nm diameter SiNWs (p-doped at \(\sim 10^{19} \text{ cm}^{-3}\)) in 0.165M solution by hybridizing to a primary DNA strand that has been electrostatically adsorbed onto an amine-terminated organic monolayer atop the NWs. This intimate contact of the primary strand with the amine groups of the NW surface brings the binding event close enough to the NW to be electronically detected. In addition, within a 0.165M ionic strength solution the DNA hybridization is more efficient.\(^{10, 22}\) However, we further demonstrate that the sensing of proteins in physiological conditions is fundamentally limited by the size of the antibodies, which, at the moment, remain the most widely used high affinity probes for most proteins. This problem may be circumvented by utilizing alternative probes, such as aptamers.\(^{12}\) Thus, if the antibodies are used as the probes for electrical label-free protein detection, the experiments are usually limited to low ionic strength solutions.

A second challenge involves demonstrating reproducible and high-throughput nanofabrication methods that can produce nearly identical NW sensors time and time again, and that allow for multiple measurements to be executed in parallel. Based on
electrostatic considerations, it is well known that nanowires are more sensitive to surface charges than planar ion-sensitive field-effect transistors (ISFET) or chemical field-effect transistor (CHEMFET). Such dimensional arguments\textsuperscript{23} imply that nontraditional methods must be utilized to fabricate the NWs.\textsuperscript{24, 25} While biological sensing with silicon produced by wet etch\textsuperscript{26} or dry oxidation\textsuperscript{27} was reported, to date, most reports of NW sensors have utilized semiconductor NWs grown as bulk materials\textsuperscript{28} using the vapor-liquid-solid (VLS) technique.\textsuperscript{29} This method produces high quality NWs, but they are characterized by a distribution of lengths and diameters, and they also must be assembled into the appropriate device structure (or the device structure must be constructed around the nanowire\textsuperscript{30}). In this study we utilize the Superlattice Nanowire Pattern Transfer (SNAP) method\textsuperscript{31} (chapter 2) to produce highly aligned array of 400 SiNWs, each 20 nm wide and \textasciitilde2 millimeters long. Standard semiconductor processing techniques are utilized to control the NW doping level,\textsuperscript{32} to section the NWs into several individual sensor arrays, to establish electrical contacts to the NW sensors, and to integrate each array into a microfluidic channel. Such integration is rather challenging in itself;\textsuperscript{33} however, it is extremely important for obtaining low-noise, reproducible measurements. The resulting NWs exhibit excellent, controllable, and reproducible electrical characteristics from device to device and across fabrication runs. The sensor platforms may also be fabricated in reasonably high throughput. A key advantage which is provided by the top-down approach of SNAP versus the bottom-up VLS technique is the precise control of doping level of the nanowires. We utilize diffusion doping technique (chapter 2) to create nanowires with well characterized doping levels ranging from \textasciitilde10\textsuperscript{17} to \textasciitilde10\textsuperscript{20} cm\textsuperscript{-3}. We demonstrate that the doping level profoundly affects the limit
of detection of both DNA and protein; therefore, nanowires can be tuned to a specific dynamic range window with an appropriate concentration of impurities.

A third challenge involves the SiNW surface. The effectiveness of SiNWs for biomolecular sensing arises in part because of their high surface-to-volume ratio. The native (1-2 nm thick) surface oxide on a SiNW may limit sensor performance due to the presence of interfacial electronic states. In addition, the oxide surface of SiNWs acts as a dielectric which can screen the NW from the chemical event to be sensed. Covalent alkyl passivation of Si(111) surfaces can render those surfaces resistant to oxidation in air and under oxidative potentials (chapter 4). Recently, methyl-passivated SiNWs were shown to exhibit improved field-effect transistor characteristics. More complex molecules, such as amine terminated alkyl groups, can be covalently attached to H-terminated Si surfaces (including SiNWs) via UV-initiated radical chemistry. Such chemistry has been used for a covalent attachment of DNA to VLS grown SiNWs. DNA may also be immobilized on amine-terminated surfaces via electrostatic interactions. In this work, we explore how the characteristics of SiNW sensors vary as the nature of the inorganic/organic interface is varied. We find that SiNW sensors in which the native oxide provides the interface for organic functionalization are significantly inferior in terms of both sensitivity and dynamic range when compared with SiNW sensors that are directly passivated with an alkyl monolayer.

A final challenge is actually an opportunity that is provided by the intrinsic nature of a label-free, real-time sensor. The standard such sensing technique is surface plasmon resonance (SPR). SPR is utilized to determine the $\kappa_{on}$ and $\kappa_{off}$ rates, and hence the equilibrium binding affinities, of complementary DNA strands or antibody-protein pairs.
The capture agent (single-stranded DNA or an antibody) is typically surface bound, and so the key experimental variables are the analyte (complementary DNA strand or a protein) concentration and time. If $\kappa_{on}$ and $\kappa_{off}$ are both known, then SPR can be utilized to quantitate the analyte concentration. Very few biomolecular sensing techniques are quantitative. In this work, dope the NW sensors so that their sensing dynamic range is optimized to match that of SPR for the detection of DNA hybridization or protein binding to an antibody. We demonstrate the equivalence of these two methods, and thus demonstrate the potential use of SiNW sensors for quantitating analyte concentrations. SiNW sensors can be optimized for significantly higher sensitivity than SPR by an appropriate surface modification and doping, and thus can potentially be utilized to quantitate the concentrations of specific biomolecules at very low concentrations. That would constitute a unique application of these devices.

3.2 Experimental Methods

3.2.1 Nanowire Sensor Fabrication

The SiNW arrays were fabricated as previously described, and all fabrication was done within a class 1000 or class 100 clean room environment. A typical NW sensor device employed in this work for DNA sensing is shown in Figure 3.1. The starting material for the SNAP process was an intrinsic, 320 Å thick silicon-on-insulator (SOI) substrate with (100) orientation (Ibis Technology, Danvers, MA) and with a 1500 Å buried oxide. Cleaned substrates were coated with either p-type (Boron A, Filmtronics, Bulte, PA) or n-type (Phosphorosilica, Emulsitone, Whippany, NJ) spin-on-dopants (SODs). SODs were thermally diffused into the SOI film (chapter 2). We reproducibly controlled the resulting substrate doping concentration, as quantified by 4-point
resistivity measurements on the SOI film, by varying the diffusion temperature. The exact doping conditions, including the temperatures corresponding to particular doping levels, are described in details in chapter 2. The p-type substrates were thermally oxidized in O₂ for 1 minute at 850 °C, which was necessary to remove the organic SOD residue. The SOD films were removed by brief immersion in piranha (70% H₂SO₄, 30% H₂O₂), followed by a water rinse, and immersion in buffered oxide etchant (BOE; General Chemical, Parcippany, NJ).

Figure 3.1: A diagram (A) and an SEM image (B) of a single device section containing three groups of ~10 SiNWs in a microfluidics channel. The wafer is covered with Si₃N₄ except for an exposed active region with SiNWs (A, inset; B). B, inset: High resolution SEM image of 20 nm SiNWs.
The SNAP method for NW array fabrication translates the atomic control achievable over the individual layer thicknesses within an MBE-grown GaAs/Al\(_x\)Ga\(_{(1-x)}\)As superlattice into an identical level of control over NW width, length and spacing. This method has been described in some detail elsewhere,\(^{31,46}\) and will not be described here (refer to chapter 2 for details). We utilized the SNAP process to produce a 2 mm long array of 400 SiNWs, each of 20 nm width and patterned at 35 nm pitch (Figure 3.1B, inset).

The SiNWs were sectioned into \(~30\) \(\mu\)m long segments using e-beam lithography (EBL) and SF\(_6\) RIE etching, producing groups of \(~10\) SiNWs with a diameter of 20 nm. Six identical sections from a single imprint, each containing 3 NW segments, were produced. One such section is shown in Figure 3.1. When fully integrated with the microfluidics channels, this allowed for six separate measurements, with three independent NW segments per measurement. Source (S) and drain (D) electrical contacts, \(~500\) nm wide and separated by 10-15 \(\mu\)m, were patterned using electron beam lithography (EBL) on each section of SiNWs. Prior to metallization, the native oxide of the SiNWs over the contacts was removed with BOE to promote the formation of ohmic contacts. Finally, 400 Å Ti and 500 Å Pt were evaporated to form S/D contacts. Immediately after the liftoff, the devices were annealed in 95\% N\(_2\), 5\% H\(_2\) at 475\(^\circ\)C for 5 minutes. This step greatly improves the characteristics of SNAP SiNW FETs. To provide room for a 1 cm by 1.5 cm PDMS chip with microchannels for analyte delivery to each section of the SiNWs (Figures 3.1A and 3.2), the electrical contacts were extended to the edges of the substrate using standard photolithography techniques followed by evaporation of 200 Å Ti and 1500 Å Au. To eliminate parasitic current between metal contacts in solution, approximately 70 nm of Si\(_3\)N\(_4\) was deposited using
plasma-enhanced chemical vapor deposition (PECVD) everywhere on the chip except in 5 μm by 20 μm regions over the NWs and the outer tips of the Au contacts. Briefly, 100 nm of chromium was deposited over an active region of the NWs. PECVD was used to deposit Si₃N₄ film at 300°C (900mT, 20W, 13.5MHz) from N₂ (1960 sccm), NH₃ (55 sccm) and SiH₄ (40 sccm) gases. The nitride film was selectively etched with CHF₃/O₂ plasma over the protected NW region using PMMA as a mask, followed by the removal of chromium with CR-7C (Cyantek Corp., Fremont, CA).

**Microfluidics fabrication.** The soft lithography microfluidics chips were fabricated as described by others.⁴⁷ We observed that manual introduction/changing of solutions caused serious noise, capacitive currents and baseline shifts in real-time recordings. Thus, for low noise, stable real-time electronic measurements, we found it necessary to automate fluid injection and solution switching by using PDMS multilayer, integrated elastomeric microfluidics chips of the type developed by the Quake and Scherer groups.⁴⁸ The size of the wafer containing SiNWs did not permit the inclusion of all necessary flow and control lines necessary for the fluidic handling chip, and so that was fabricated as a separate chip. Such PDMS chip was fabricated using a standard photolithography: mixed PDMS (Dow Corning, Midland, MI) was applied over a premade photoresist (Shipley SPR 220-7) molding on silicon wafer and incompletely cured at 80 °C for 30 minutes. The chip containing microchannels was cut out of the PDMS layer and 0.5 mm diameter holes were punctured to serve as microchannel inlets and outlets. The fluidic chip and the device containing SiNWs were then brought into contact, with the 100μm wide microchannels aligned over the individual nanowire sections. The assembled device was cured to completion overnight at 80 °C.
Figure 3.2: A) Fabrication and assembly of the two-layer PDMS chip for solution injection (top) with the sensing device composed of SOI wafer and a single-layer PDMS chip with six separate microchannels (bottom). B) A photograph of a sensing device with PDMS chip containing six microchannels. The chip is fixed onto a chip carrier and the gold pads on the device are wirebonded to the gold pads of the chip carrier. C) The sensing device with the PDMS chip on a chip carrier (B) is fixed into a chip carrier socket, and the PDMS chip for solution injection (A, top) is shown connected to one of the six microchannels of the sensing device, as outlined in the schematic A. The tubing with different solutions is feeding into the solution injection chip.

To automate an injection/changing of analyte solutions, we also introduced a second PDMS chip which can sequentially inject four different solutions into one of six
microchannels on silicon wafer. Such sample injection chip is composed of two layers, control layer and flow layer (Figure 3.2A). For the fabrication of the flow layer, mixed PDMS was spin coated on a photoresist mold at 2500 rpm for 50 sec and incompletely cured at 80 °C for 30 minutes. Control layer was fabricated by applying mixed PDMS over a photoresist mold directly and incompletely curing at 80 °C, followed by the puncturing of holes for inlets and outlets. The two layers were aligned together and the inlets/outlets for the flow layer were created. After two hours at 80 °C, the two-layer PDMS chip was bonded to a glass slide via an O₂ plasma treatment. By utilizing such sample injection chip, we were able to control the injection and solution-changing processes without disturbing the measurement, while maintaining the sensing device in an electrically isolated chamber at all times. The valves of the sample injection chip were actuated with the Labview program by means of the BOB3 Microfluidic Valve Manifold Controller and solenoid cluster manifolds (Fluidigm, San Francisco, CA). By introducing a waste outlet into the sample injection chip, we were able to remove any bubbles arising from switching between different solutions, which also helped in maintaining a stable baseline reading.

3.2.2 Surface Functionalization and Characterization

*Synthesis of tert-Butyl allylcarbamate.* To a solution of allylamine (2.27 g, 39.8 mmol) in THF (20 mL) was added N,N-diisopropylethylamine (13 mL, 80.0 mmol) followed by di-tert-butyl dicarbonate (8.7 g, 39.9 mmol). After 1 hr, the organic solvent was evaporated under reduced pressure, and the residue was purified by silica gel chromatography (Hex:EtOAc = 9:1) to give 6.6 g (93%) of a product as a clear oil. ^1^H NMR 300 MHz (CDCl₃) δ 5.82 (m, 1H), 5.12 (m, 2H), 3.74 (bm, 2H), 1.45 (s, 9H).
Surface functionalization. The two procedures used to functionalize SiNWs with and without oxide layer are shown in Schemes 3.1 and 3.2, respectively.

### Scheme 3.1: Functionalization of Si(100) oxide surface with amine.

Both procedures resulted in an amine terminated organic monolayer atop SiNWs. For the oxide surface functionalization, cleaned SiNWs were treated with 2% (v/v) 3-aminopropyldimethylethoxysilane (Gelest, Morrisville, PA) in toluene for 2 hrs. The wafers were then rinsed in toluene and methanol and incubated at 100 °C for 1 hr.

The procedure described previously\(^{38, 44}\) (chapter 4) was used to functionalize hydrogen terminated SiNWs with tert-Butyl allylcarbamate (scheme 3.2). SiNWs were immersed in 2% HF solution for 3 seconds, washed with Millipore water and blow-dried under N\(_2\) stream. The wafer was immediately placed in a custom made quartz container which was then pumped down to ~10\(^{-5}\) Torr, followed by an argon purge. Under positive argon pressure, a mixture of 1:2 tert-Butyl allylcarbamate:methanol (v/v) was applied to the wafer, completely covering the SiNWs. The wafer was illuminated with UV (254 nm, 9 mW/cm\(^2\) at 10 cm) for 3 hours. SiNWs were then rinsed in methylene chloride and methanol. The deprotection of t-Boc amine was carried out in a solution of TFA in
methanol (1:4 v/v) for 4 hours, followed by extensive methanol washing. In the case of antibody attachment, the amine terminated surfaces were reacted with water-soluble homobifunctional N-hydroxysuccinimide ester (NHS ester), followed by the introduction of 50 mg/ml of monoclonal anti human IL-2 or anti human TNFα antibodies. The unreacted amines were quenched with ethanolamine (100mM in 1×PBS).

**X-ray photoelectron spectroscopy.** X-ray photoelectron spectroscopy (XPS) was utilized to quantify the amount of oxide on Si(100) wafers after surface treatments outlined in Schemes 3.1 and 3.2. XPS was also used to follow the attachment of antibodies to silicon surfaces. All XPS measurements were performed in an ultrahigh vacuum chamber of an M-probe surface spectrometer that has been previously described.\textsuperscript{49} Experiments were performed at room temperature, with 1486.6 eV X-ray from the Al Kα line and a 35° incident angle measured from the sample surface. ESCA-2000 software was used to collect the data. An approach described elsewhere\textsuperscript{36, 49} was used to fit the Si 2p peaks and quantify the amount of surface SiO\textsubscript{x}, assuming that the oxide layer was very thin. Any peak between 100 eV and 104 eV was assigned to Si\textsuperscript{−}−Si\textsuperscript{4+} and fitted as described in the literature.\textsuperscript{50} SiO\textsubscript{x}:Si 2p peak ratio was divided by a normalization constant of 0.17 for Si(100) surfaces.

**Contact angle measurements.** The sessile contact angle of water on the functionalized Si(100) surface was used to check the fidelity of surface chemistry as described in Schemes 3.1 and 3.2. Contact angle measurements were obtained with an NRL C.A. Goniometer Model #100-00 (Rame-Hart, Netcong, NJ) at room temperature. All measurements were repeated three times and averaged to obtain the contact angle θ for the surface.
3.2.3 SPR and Electronic Measurements

**Surface plasmon resonance (SPR).** All SPR experiments were performed on the Biacore 3000 with carboxylic acid terminated Biacore CM5 chips. The active flow cells were first primed in 1x SSC (15 mM NaCitrate, 150 mM NaCl, pH 7.5). To generate an amine surface, the carboxylic acid groups were converted to succinimide esters by flowing EDC/NHS prior to exposure of a 1 mg/ml solution of polylysine (Sigma-Aldrich, St. Louis, MO). Single stranded DNA (5’TGGACGCATTGCACAT3’, Midland Certified, Midland, TX) was electrostatically absorbed unto the polylysine matrix. Complementary DNA was then immediately introduced and allowed to hybridize to the active surface. The flow cell was regenerated with two 1 minute pulses of 50 mM NaOH, after which ssDNA was reabsorbed electrostatically before another cDNA pulse was introduced for hybridization. The antibodies in acetate buffer (pH 5.5) were attached directly immediately following the surface treatment with EDC/NHS, and the remaining esters were reacted with ethanolamine. The antigen was introduced at various concentrations in 1x PBS buffer at the flow rate of 30 μl/min. Between each addition, the surface was regenerated with glycine/HCl buffer (pH 3.0).

**Electronic measurements.** The 4-point resistivity of silicon film as well as SiNW resistances and solution gating were measured with Keithley 2400 Source Meter (Keithley Instruments, Cleveland, OH). The sensing experiments were performed with SR830 DSP lock-in Amplifier (Stanford Research Systems, Sunnyvale, CA). Figure 3.3 shows the experimental setup for the electronic measurements. A 50 mVrms at 13 Hz voltage source \( V_{SD} \) was applied to one terminal of the nanowire, with the amplifier input operating in the current-measure mode. For the DNA-sensing experiments,
platinum wire was inserted into the microchannel and used as a solution gate, while it was kept at a ground potential throughout the real-time measurements to reduce the noise in the system (Figure 3.1A). In the case of protein sensing, the handle of the wafer (backside Si) was held at a ground potential instead of the platinum electrode in solution. The devices were functionalized and assembled as described above. Single stranded 10 μM DNA (same as in SPR experiments) in 1x SSC buffer was flowed through the microchannel for 1 hr and allowed to electrostatically adsorb to the amine terminated surface of SiNWs. The nonbound DNA was washed thoroughly with 1x SSC buffer. Complementary DNA (5’ATGTGCAATGCGTCCA3’, Midland Certified, Midland, TX) of varying concentrations in 1x SSC buffer was sequentially injected from the injection PDMS chip into the microchannel containing SiNWs at a flow rate of 2.0
μl/min as the resistance of the NWs was recorded in real time. Noncomplementary DNA (noncomp. DNA) (5’CATGCATGATGTCACG3’) was used as a control. In general, a different SiNW sensor was utilized for each of the measurements described here. Similar procedure was followed for protein sensing. The surface functionalized with capture antibodies was subjected to the 10 μM PBS solution containing various antigen concentrations (1 pM to 100 nM). After the introduction of a particular concentration, the surface was completely regenerated with 10 μM PBS, followed by the introduction of the next antigen concentration in the same microchannel.

3.3 Single-Stranded DNA Sensing

3.3.1 Nanowire Surface Passivation

We used contact angle measurements to follow the functionalization processes of various surfaces. Table 3.1 presents the data for both Schemes 3.1 and 3.2 (See Experimental Methods Section). The procedure in Scheme 3.1 generates a large increase in contact angle. Similarly, large changes in contact angles are observed for photochemically treated Si surface before and after t-Boc deprotection. The resulting contact angle of ~60° is observed for surfaces prepared by Scheme 3.1 and 3.2, arguing for an existence of chemically similar, amine terminated monolayers on these surfaces.
Table 3.1: Measured contact angles for various Si(100) surfaces.

<table>
<thead>
<tr>
<th>Si(100) surface</th>
<th>contact angle (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>With nonfunctionalized oxide</td>
<td>11 ± 1</td>
</tr>
<tr>
<td>Scheme 1: amine terminated</td>
<td>61 ± 1</td>
</tr>
<tr>
<td>Scheme 2: t-Boc protected</td>
<td>81 ± 1</td>
</tr>
<tr>
<td>Scheme 2: deprotected, amine terminated</td>
<td>60 ± 1</td>
</tr>
</tbody>
</table>

Quantifying the amount of oxide on the SOI NWs is extremely challenging. Therefore, we used Si(100) bulk surfaces to approximate the amount of surface oxide remaining after photochemical functionalization. Figure 3.4A shows XPS scan in the Si/SiO_x region. The Si(100) surface with native oxide exhibited approximately 1.9 equivalent monolayers of SiO_x. In contrast, the Si(100) surface treated according to Scheme 3.2 contained 0.08 equivalent monolayers of SiO_x prior to TFA deprotection and 0.3 monolayers of SiO_x after the deprotection step and a 10 hr exposure to 1x SSC buffer.
Figure 3.4: A) XPS of Si 2p region of Si(100) surface functionalized as in Scheme 3.2 before (dark grey) and after (light grey) TFA deprotection and 10 hr in 1x SSC buffer. Nonfunctionalized Si(100) surface with native oxide (black). Inset: N 1s region of nonfunctionalized Si(100) surface (black), Si(100) functionalized by Scheme 3.1 (light grey) and Scheme 3.2 (dark grey). B) Current-Voltage (IV) graphs of SiNWs functionalized by Scheme 3.1 in solutions of varying pH. Inset: solution gated ($V_{SG}$) n-type hydroxyl terminated SiNW in solutions of varying pH.
The roughness of a SiNW surface may cause a more extensive oxidation than the one observed on the bulk surface, but the data in Figure 3.4A does demonstrate a significant reduction in oxide thickness after photochemical treatment. Furthermore, we used XPS to determine the presence of amine terminated monolayer on bulk Si(100) surfaces post functionalization with two different schemes. Figure 3.4A inset demonstrates the XPS scans of N 1s region. Nitrogen peak is clearly visible for surfaces functionalized by Schemes 3.1 and 3.2, while no peak is present for the nonfunctionalized Si.

Scheme 3.1 functionalized SiNWs demonstrate sensitivity to pH which is different than for native oxide-passivated NWs. The isoelectric point of silica is ~2, implying that for hydroxyl terminated, non-functionalized SiNWs at low pH, the SiOH groups are largely protonated. At high pH, negative charges on SiO− should deplete carriers in the n-type SiNWs, causing a decrease in \( I_{DS} \) (Figure 3.4B, inset). Above pH 4 the conductance is no longer modulated by increasing the pH, as most of the hydroxyl groups are deprotonated. When the surface is functionalized with amine (\( pK_a \approx 9-10 \)), the opposite trend is expected. At low pH, the amine is protonated, causing carrier depletion or increased resistance in p-type SiNW. This trend is observed in Figure 3.4B, where the sharpest transition in resistance occurs between pH 9 and 10. The observation of the correct pH effects on the resistance of the SiNWs serves as a confirmation of the presence of surface functional groups, amine in this case.

As shown in Figure 3.5, oxide covered SiNWs (~8e18 cm\(^{-3}\) doping level) in 1x SSC buffer (0.165M, pH 7.2) respond weakly to the applied solution gate voltage, \( V_{SG} \), showing no significant on-off current transition between 0.8 and -0.8 Volts.
In contrast, directly passivated SiNWs (Scheme 3.2) exhibit on-off current ratios of \(~10^2\). Figure 3.5 strongly suggests that directly passivated SiNWs exhibit an enhanced response to surface charges and should therefore serve as superior NW sensors compared with similarly functionalized, but oxide-passivated SiNWs.

![Graph showing solution gating of SiNWs functionalized by Scheme 3.1 (grey) and by Scheme 3.2 (black) (\(V_{SD}\) was 50 mV). (Right inset) IV curves of SiNWs in air with (black) and without (grey) oxide. (Left inset) Resistances in air of SiNWs functionalized by Scheme 3.1 (left) and Scheme 3.2 (right).]

The Scheme 3.2 procedure does involve an HF etch step, which can be potentially detrimental to the device conductance. We, therefore, checked the conductivity of SiNWs before and after photochemical treatment. Lightly doped SiNWs provide for superior FET properties,\(^{52}\) and, in fact, we have reported that lightly doped (10\(^{17}\) cm\(^{-3}\)) p- or n-type SiNWs are more sensitive biomolecular sensors than the highly doped nanowires (10\(^{19}\) cm\(^{-3}\)) (section 3.3.2).\(^{53}\) Our doping process preferentially dopes the top few nanometers of the SiNWs, as described in chapter 2.\(^{54}\) Thus, if the HF etching of the
Si surface was extensive enough, we could expect an enhancement in SiNW current modulation by $V_{SG}$ to be entirely due to the decrease in carrier concentration and not the removal of surface oxide. The insets in Figure 3.5 demonstrate that the NW resistance increased only, on average, by a factor of 2 following the HF treatment. This relatively negligible resistance increase indicates that the major reason that the SiNWs prepared by Scheme 3.2 exhibit an improved solution FET performance originates from the elimination of oxide via direct silicon passivation. This result is consistent with the recent demonstration that, for VLS grown SiNWs, direct methylation of the SiNW surface leads to a $10^3$ to $10^4$ fold enhancement in the on-off conductance of the FETs made from those nanowires.\textsuperscript{39}

**Figure 3.6:** Real-time response of SiNWs functionalized as in Scheme 3.1 to the addition of (a) 10µM ssDNA and (b) 100nM complementary DNA. Right top inset: Real-time SiNW response to the sequential addition of (a) 0.165M SSC, (b) 0.0165M SSC, and (c) 0.00165M SSC buffers. Left inset: SPR measurement demonstrating the addition of 10µM ssDNA to poly-L-lysine coated CM5 sensor chip. $V_{SD} = 50$ mV.
Figure 3.6 shows SiNW real-time detection of the electrostatic adsorption of 10μM single-stranded DNA (ssDNA), followed by the hybridization in 1x SSC buffer of 100nM complementary DNA strand. As expected, the resistance of p-type SiNWs is decreased with the addition of negative surface charges. The metal contacts to NWs have been covered with Si₃N₄ layer, and there is no background conductance through the solution. We have observed an insignificant change in the resistance of the highly doped NWs upon switching from dry environment to buffer solution (data not shown). Moreover, as Figure 3.6 (right inset) shows, changing the ionic strength of the solution does not affect the resistance. This behavior is particular to highly doped wires. The low doped wires do show a change in conductance as the ionic strength of the solution is varied. In addition, the automated solution injection (section 3.2.1) removes any baseline shifts or transient changes in the resistance. SPR was also utilized in parallel to SiNWs in order to validate the surface chemistry and to obtain kinetic parameters such as $\kappa_{\text{on}}$, $\kappa_{\text{off}}$ and affinity constant $K_A$ for this particular DNA pair. Poly-L-lysine was covalently attached to the SPR sensor chips, mimicking the amine terminated monolayer of SiNWs. Figure 3.6 (left inset) shows the SPR response to the electrostatic adsorption of primary DNA from a 10μM DNA solution. The surface density of adsorbed DNA was estimated as $2.5 \times 10^{13}$ cm$^{-2}$, using the conversion factor of 1000RU = 100 ng cm$^{-2}$ from the literature. The surface density is approximately an order of magnitude higher than the average surface density of $10^{12}$ cm$^{-2}$ obtained when localizing biotinylated DNA on a streptavidin covered surface. Such high surface density of primary DNA is expected because the poly-L-lysine treated surface is positively charged. It is likely that the amine-terminated SiNW surface has less surface charges than the poly-L-lysine covered surface and thus contains fewer sites for electrostatic adsorption of oligonucleotides.
Figure 3.7 demonstrates real-time label free detection of ssDNA by SiNWs and by SPR. In either case, the primary DNA strand was electrostatically immobilized on the sensor surface. Known DNA concentrations were injected after a stable reading with 1x SSC buffer was obtained, and the flow was maintained throughout the experiment. Different concentrations were detected with different groups of SiNWs. We observed that the hybridization on SiNWs is essentially irreversible on the relevant time scales when the analyte DNA was being washed away with the buffer solution. Such behavior is in contrast to SPR measurements, where the slow reversal of hybridization was observed (Figure 3.7C). The performance of the NWs functionalized according to Scheme 3.1 (Figure 3.7A) was compared to SiNW sensors prepared according to Scheme 3.2 (Figure 3.7B). The SPR experiments, although carried out on Au substrates, also utilized primary ssDNA that was electrostatically adsorbed onto an amine terminated surface. The intention here was to find experimental conditions that could serve to validate the NW experiments by obtaining kinetic parameters for these particular DNA strands under specific experimental conditions. Control experiments with noncomplementary DNA yielded no response for either SiNWs or SPR measurements (black traces of Figures 3.7A and 3.7C). These negative controls were also independently validated via fluorescent detection in microfluidic channels on two different Si surfaces (Figures 3.7A and B, insets). Figure 3.7D demonstrates that the NW response $(\Delta R/R_o)$ varies as $\log[\text{DNA}]$. Such a logarithmic dependence has been previously reported.\textsuperscript{21}
Figure 3.7: Concentration-dependent, real-time sensing of complementary DNA by SiNWs and by SPR in 0.165M electrolyte. A) Real-time responses of SiNWs that were surface functionalized according to Scheme 3.1 and coated with electrostatically adsorbed primary DNA. The black trace represents exposure of the SiNW sensors to 100nM non-complementary ssDNA. Each curve represents measurements from a different set of NWs. Inset: Fluorescence image of Si(100) surface (with overlaying PDMS microfluidics chip) treated as in Scheme 3.1 followed by 10μM primary DNA addition and addition of (microchannel a) 100nM noncomplementary fluorescent DNA and (microchannel b) 100nM complementary fluorescent DNA. PDMS chip was removed before the image was collected. B) As in A, except the SiNWs were functionalized according to Scheme 3.2. Inset: Same as in A inset, but Si(100) surface was treated as in Scheme 3.2. C) SPR measurement of the hybridization of complementary DNA to electrostatically adsorbed primary DNA on a poly-L-lysine surface. D) Normalized SiNW responses for Scheme 3.1 (black dots) and Scheme 3.2 (red dots) surface preparations, as a function of the log of DNA concentration. For all measurements, $V_{SD} = 50mV$. 

\[ \text{RU (x10^2)} \]

\[ \text{Time (min)} \]

\[ \text{DNA} \]

\[ \text{noncomp. DNA} \]

\[ \text{500nM 1xSSC buffer} \]

\[ \text{250nM} \]

\[ \text{125nM} \]

\[ \text{63nM} \]

\[ \text{32nM} \]
As demonstrated in Figure 3.7D, the dynamic range of SiNWs is increased by 100 after the removal of oxide and UV-initiated chemical passivation; the limit of detection (LOD) increased from 1nM to 10pM.

### 3.3.2 Nanowire Impurity Doping Level

The way a semiconductor material such as silicon responds to surface potential or surface charges strongly depends on the impurity concentration (doping level) inside the material. Highly doped devices exhibit short Debye lengths and short depletion regions. The Debye length, $L_D$, is inversely proportional to the square root of the impurity concentration, $N$. $L_D$ is a measure of how effectively the electric field is screened inside a material. Short $L_D$ means that the screening is highly effective. For example, $L_D$ of silicon doped at $\sim 10^{17}$ is approximately 10 nm, and that of $\sim 10^{19}$ doped silicon is only 1 nm. Our silicon nanowires are doped non-uniformly: the top 10 nanometers of the wire are doped the heaviest, with the effective doping decreasing by an order of magnitude approximately 10 nanometers from the top surface of the wire. Therefore, we would expect that the nanowires doped at $\sim 10^{17}$ will be the most sensitive to surface charge fluctuations due to molecular binding events in solution. Microscale EIS (electrolyte-insulator-silicon) FET devices were shown to be sensitive to the surface binding of DNA only when the sensor was doped at extremely low level of $\sim 10^{15}$. The doping level is, therefore, an extremely important parameter to optimize, and, unlike the VLS grown wires, where the doping level is very hard to control precisely, SNAP method combined with SOD (spin-on doping, chapter 2) allows us to vary the doping of nanowires in a highly predictable and reproducible fashion.
In Figure 3.8A we present typical I-V curves for lightly doped p- and n-type SiNWs used for biological sensing. The linearity of the graphs suggests that the electrical contacts are ohmic. Moreover, the contact resistances are negligible compared with $10^7$ to $10^8$ Ω resistances of lightly doped SiNWs. Slight variations in measured resistances are expected due to a variable number of SiNWs contacted by the S/D contacts (between 5 and 10). We are currently working towards an individually
addressable arrays of SiNW sensors. The resistance obtained from the slopes of I-V graphs may be used to estimate the doping concentrations, assuming mobilities of 100 cm$^2$/Vs and 300 cm$^2$/Vs for lightly doped (100) p-type and n-type silicon. The resulting values of $\sim 3.5e17$ cm$^{-3}$ for n-type and $\sim 5.6e17$ cm$^{-3}$ for p-type agree well with the doping concentration of $\sim 3e17$ cm$^{-3}$ measured by the 4-point method on the initial 300 Å SOI film.

The ability of SiNWs to sense charged species on their surfaces is directly linked to their performance as FETs in aqueous solution. Figure 8B demonstrates that the drain-source current ($I_{DS}$) can be gated by the voltages applied to the Pt electrode patterned near the sensor with the solution acting as the gate dielectric. The gating profiles are consistent with p and n-type FETs: $I_{DS}$ of n-type (p-type) wires increases with the application of positive (negative) gate bias. Lightly doped SiNWs yield larger $dI_{DS}/dV_{SG}$ values, suggesting that the lightly doped NWs should be more sensitive biosensors. The sensing response of SiNWs has largely been attributed to the modulation of the conductance of the 1D NWs through a change in surface charges (probably a combination of charges on biomolecules/solvating ions and exclusion of water/ions by the biomolecules).

We investigated how the sensitivity of SiNWs to single stranded DNA changes as the doping concentration is varied (Figure 3.9). As expected, lightly doped p- and n-type NWs showed increased levels of sensitivity: nanowires with impurity concentrations in the range of $10^{17}$ cm$^{-3}$ exhibited the best performance. Wires doped at $\sim 10^{19}$ cm$^{-3}$ and higher exhibited no detectable response to complementary DNA below the concentration
Figure 3.9: A) Real-time sensing results from n-type (black: top x-axis) and p-type (grey: bottom x-axis) SiNW sensors, both doped at ~ 3e17 cm\(^{-3}\), demonstrating sensitivity in 100attoM (10\(^{-16}\)M) range in 1x SSC buffer. The various points indicated are (n-type) n1=220attoM, n2=22femtoM, n3=2.2pM, n4=220pM complementary DNA; (p-type) pC=22nM non-complementary DNA, p1=220attoM, p2= 22femtoM, p3=2.2pM complementary DNA. B) Both p- and n-type nanosensors exhibit sensitivity over a broad dynamic range. The sensitivities, however, depend on doping concentrations: n-type ~3e17 cm\(^{-3}\) (black circles), ~3e18 cm\(^{-3}\) (black triangles), ~1e19 cm\(^{-3}\) (black squares); p-type ~3e18 cm\(^{-3}\) (grey triangles), ~3e17 cm\(^{-3}\) (grey circles).
of 1nM, consistent with the data from Figure 3.7. The lower limit of detection of single-stranded 16-mer DNA for our NW sensors is at a concentration of approximately 1fM, consistent with other reported SiNW experiments.\textsuperscript{14,27} The dynamic range, defined as the range over which $\Delta R/R_0$ versus $\log[\text{DNA}]$ is linear, is approximately $10^6$. As was mentioned previously, due to the electrostatic adsorption of primary DNA strand to the sensor surface, 1fM sensitivity is obtained in the high ionic strength (0.165M) solution. Therefore, proper engineering of the sensor, controlling impurity concentration and surface passivation, leads to optimized device performance. However, in the following section on protein sensing we will demonstrate that optimizing surface chemistry and the electronic properties of the sensor is not enough to obtain sensitivities which are significantly superior to optical detection. This is due to the fundamentally limiting effect of screening in high ionic strength solutions. In such environments, the size of the probe and the distance from the surface of the sensor at which the binding event takes place dominate the sensitivity of the electronic detection.

### 3.4 Protein Sensing

Robust label-free detection of proteins below the concentration of ~10pM is of considerable importance in rapid clinical evaluation, cancer marker detection, disease staging, etc. The real-time nature of electronic label-free detection also offers additional benefits such as characterization of new affinity probes, drug screening, and could, therefore, be potentially useful in basic research as well as in clinical practice. For these reasons, we have extended the above study to the detection of proteins. Such an endeavor, however, faces a fundamental challenge, owning to the significant charge screening in the solution of high ionic strength. The extent of such screening may be
characterized by the Debye length, $1/\kappa$\textsuperscript{20}, which describes a distance from a point charge at which the potential due to that charge drops off to $\sim e^{-1}$ of its value. Scheme 3.3 demonstrates the relevance of Debye screening to the electrical detection of biomolecules in solution. Here, we assume that the antibodies, which serve as capture probes for proteins, are approximately 10 nm long, and are randomly oriented on the surface of the nanowire. The change in charge density, $\Delta \sigma$, due to the equilibrium binding of proteins is smeared over a distance $b$,

\begin{equation}
\Delta \Psi = \frac{\Delta \sigma}{b} \left[ \frac{e^{ - \kappa d} - e^{ - \kappa (d+b)}}{\varepsilon_0 \varepsilon_w \kappa^2} \right]
\end{equation}

which is a distance $d$ away from the NW surface. The change in chemical potential per area at the surface of the nanowire may be described by the Debye-Hückel equation,\textsuperscript{18,20}
charge density. Surface potential and the distance from the surface at which the binding takes place are intimately coupled. If the screening length is much smaller than $d$, $\kappa^{-1} << d$, then the potential due to protein binding will be completely screened from the surface of the nanowire. Therefore, the condition $\kappa^{-1} \geq d$ must be met in order to detect charged species in the solution a distance $d$ away from the surface. In the case of DNA sensing (section 3.3), the capture probe single-stranded DNA was electrostatically adsorbed on the NW surface, and the hybridization was taking place very close, ~1 nm from the surface, allowing us to carry out sensing in high ionic strength conditions of 0.165M. The antibodies, however, are fairly large biomolecules (Scheme 3.3). At 25 °C the Debye length of aqueous solution is

$$\frac{1}{\kappa} = \frac{0.304}{\sqrt{[NaCl]}} \text{nm}$$

(3.2)

for 1:1 electrolytes such as NaCl.

<table>
<thead>
<tr>
<th>[NaCl]</th>
<th>1/\kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>100mM</td>
<td>1nm</td>
</tr>
<tr>
<td>1mM</td>
<td>10nm</td>
</tr>
<tr>
<td>10\muM</td>
<td>100nm</td>
</tr>
</tbody>
</table>

As Table 3.2 demonstrates, the size of the antibodies dictates the ionic strength of the solution in which the electronic detection may take place. This is a serious limitation if the physiological medium such as serum (0.14M) must be used to detect low abundance
proteins, without the possibility of appropriate dilutions. To circumvent this problem, new high-affinity probes such as aptamers, small molecules and short peptides must be developed, all of which are significantly smaller than the antibodies.

For protein sensing, devices were modified (original devices as in Figure 3.1) to include a control channel for each measurement (Figure 3.10), which contained nonfunctionalized SiNWs. It is expected that such bare SiNWs provide a measure of nonspecific protein binding to the surface; therefore, a differential measurement taking biofouling and random drift into account is more accurate.
Figure 3.11: XPS of carbon 1s region, followed through the process of nanowire functionalization with antibodies. Starting with bare silicon-on-insulator (SOI) substrate (dark grey), the wafer is functionalized with 3-aminopropyltrimethoxysilane (APMES) as described in experimental methods section (light grey). Bifunctional cross-linker (NHS ester) is then attached to the surface amines (grey), followed by the covalent attachment of antibodies (black).

SiNW surfaces for protein sensing were functionalized in a similar manner to the experiments with DNA. Native oxide of silicon was functionalized with amine-terminating monolayer. A bifunctional cross-linker, with an NHS ester on either end, was coupled to the primary amines on the surface, followed by the coupling of antibodies to the other end of the linker. This chemistry may potentially involve any of the primary amines of the antibody, and, therefore, probably results in a random orientation of the antibodies on the surface (Scheme 3.3). Since the surface area of a nanowire is rather small, $\sim 10^{-13}$ m$^2$, this may result in the broadening of the distribution of the responses from the identical nanowires. Measuring the response of a larger array of nanowires, therefore, may lead to a smaller variance in the response.
Figure 3.12: A) SiNW (~10^{18} \text{ cm}^{-3}) response to the application of solution gate voltage in 10\mu M PBS solution. B) Current-voltage traces of SiNW (~10^{18} \text{ cm}^{-3}) a) with surface amine in air, b) with surface amine in 10\mu M PBS solution, c) after antibody attachment in 10\mu M PBS solution.

XPS scans of the carbon 1s region from the silicon surface treated as described above are presented in Figure 3.11. Clear emergence of the C=O and C-O/C-N bonds is visible, suggesting that the antibodies are successfully attached to the surface. Low doped nanowires (~10^{18} \text{ cm}^{-3}) were used for this study based on the observation of the
influence of doping level on the sensitivity of SiNWs (Figure 3.9). This doping level was chosen as the one which reproducibly yields wires with resistances of ~1 MΩ, ohmic contacts and good solution transconductance behavior (Figure 3.12A). Wires with higher doping are not as sensitive to proteins and those with lower doping yield fewer functional devices.

The attachment of antibodies was further verified by monitoring the resistance of the nanowires during the functionalization process. Figure 3.12B shows IV curves of the nanowires in air, in solution, and after the attachment of the antibodies. The resistance of the amine-terminated nanowires drops significantly after immersion in solution and further yet after the attachment of the antibodies. The pH of the 10μM PBS is approximately 6.0, which may account for the increase in the resistance in the solution. It is also possible that pH 6.0 is below the isoelectric point of the particular antibody used here, and the excess positive charges on the antibody surface lead to the further increase in the resistance.

Real-time detection of proteins with SiNWs is demonstrated in Figures 3.13 and 3.14. In each case, the same microchannel (the same SiNWs) was used for introducing antigen at various concentrations. After the saturation of the signal, phosphate buffer was used to remove bound interleukin-2 (IL2). This process was repeated several times with different IL2 concentrations, and as Figures 3.13A and 3.14 demonstrate, the antigen-antibody binding is fully reversible. SiNWs can, therefore, similarly to the SPR chip, be reused multiple times for protein detection. The data in Figure 3.13 was collected in
Figure 3.13: A) Real-time sensing of human IL2 with SiNWs doped at $10^{-18}$ cm$^{-3}$. The nanowires were functionalized with 3-aminopropyltriethoxysilane (APTES), followed by the attachment of anti human IL2 antibodies. Human IL2 in 1.5μM PBS at different concentrations (1nM, 10nM and 100nM) was introduced, each time followed by the removal of bound IL2 with 1.5μM PBS. B) Normalized resistance of nanowires ($R_0$ is an initial resistance) as a function of time and IL2 concentration. C) Normalized change in resistance $\{\Delta R/R_0 = (R_{sat} - R_0)/R_0\}$ as a function of protein concentration.

1.5μM PBS using the chip which was functionalized with 3-aminopropyltriethoxysilane (APTES). APTES forms multilayers on the surface due to intermolecular polymerization. After the data in Figure 3.13 was collected, the chip was cleaned in organic solvents and briefly in O$_2$ plasma. The surface of SiNWs was then functionalized with 3-aminopropyldimethylethoxysilane (APMES), which forms a monolayer on silicon oxide surface. Subsequent protein sensing was carried out in 10μM PBS. As evident from
Figure 3.14: A) Real-time, differential sensing of human IL2 in 10μM PBS. The solutions containing IL2 at different concentrations (100nM, 10nM, 1nM and 100pM) were flown sequentially over an active region of SiNWs which were functionalized with antihuman IL2 antibodies (black curve), with the addition of 10mM PBS after each IL2 concentration to wash away bound proteins. Simultaneously, the same solutions were also introduced into a separate channel (Figure 3.10) containing SiNWs without antibodies on the surface (grey curve). B) SPR of IL2-antiIL2 antibody interaction at different concentrations of IL2. The flow rate was 30μl/min.
comparing Figures 3.13 and 3.14, while the same device was used for protein sensing, the changes in resistance corresponding to the same concentrations of IL2 are markedly different. The reason for this difference is difficult to pinpoint exactly. In Figure 3.13, the functionalization with APTES may have resulted in a higher density of surface amines, which translated to a higher surface density of anti-IL2 antibodies, and, therefore, to a larger saturation signal. Also, longer Debye screening length, corresponding to the detection in 1.5μM PBS (Figure 3.13), versus 10μM PBS (Figure 3.14), may have also contributed to higher signals. Finally, O₂ treatment may have oxidized the surface, leading to a drop in sensitivity (chapter 4). Regardless of the exact reason, it is evident that the sensing devices may be reused multiple times for protein detection. Figure 3.14 demonstrates that the increase in resistance is specific to antibody-

![Figure 3.15: Real-time detection of 10nM TNFα with n-type 20 nm SiNWs (~10^{18} cm^{-3}) (black curve) functionalized with antiTNFα antibodies. Grey curve is the control experiment demonstrating no change in resistance upon introduction of 10nM IL6.](image)
IL2 binding. When the antibodies are absent from the SiNW surface, no changes in the resistance are observed, meaning that the nonspecific binding of the antigen is below the detection threshold. In addition, the response of the nanowires to the binding of IL2 is consistent with the majority surface charge of the protein. The isoelectric point of recombinant human IL-2 is between pH 6.5 and 7.5. Therefore in dilute PBS, pH ~6.0, there should be a prevalence of positive surface charges on the protein, leading to an increase in resistance of p-type silicon nanowires (Figures 3.13 and 3.14). In contrast, tumor necrosis factor alpha (TNFα) cytokine, with the isoelectric point of between pH 4 and 5, should have majority negative charges at pH 6.0. Figure 3.15 demonstrates the detection of TNFα with n-type SiNWs, which exhibit an increase in resistance, consistent with the lower isoelectric point of the cytokine.

The above protein sensing experiments demonstrate that it is possible to engineer silicon nanowires to detect protein concentrations below the detection limit of other label-free methods, such as SPR (Figure 3.14B). The detection of two important proteins, IL2 and TNFα, which are crucial cytokines of the immune system, is demonstrated. Potentially, this method may be utilized to detect low levels of cancer markers in the serum. For those applications especially, quantitative information about the fluctuation of protein concentrations in blood is of utmost importance. As shown above, a relative change in resistance of SiNWs for a given concentration of a protein may differ from device to device. The real-time and label-free nature of this method, however, affords a powerful opportunity to obtain quantitative kinetic information about protein-antibody interaction, and ultimately to convert that information into an absolute
protein concentration. In the next section, we will discuss a method of obtaining kinetic parameters from the kind of sensing data presented here.

3.5 Quantitative Analysis of Nanowire Response

3.5.1 Analysis of DNA-Sensing Experiments

SiNW sensors can potentially be utilized to quantitate analyte concentration and binding constants. In order to explore this possibility, the SiNW sensing response must be compared with other label-free, real-time methods such as SPR. It is also critical to design experimental parameters for both sensing modalities that are as similar as possible, as was described above. In this section, we first discuss the use of electrostatically adsorbed primary DNA for detecting complementary DNA analyte. We then discuss the development of a self-consistent model that allows for the direct comparison of SPR measurements with nanowire sensing data. Finally, we test that model by utilizing the nanowire sensing data to calculate 16-mer DNA binding constants and analyte concentrations.

Previous studies have demonstrated that the Langmuir model can be applied for parameterization of the hybridization processes of short oligonucleotides.\textsuperscript{22, 56} We used the Langmuir model to calculate kinetic parameters from the SPR hybridization measurements (Figure 3.7C) and obtained $k_{on}=1 \times 10^5$, $k_{off}=2 \times 10^{-2}$, $K_A=5 \times 10^6$ (Table 3.3). This $K_A$ value is between 10 and 100 times smaller than that reported for similar length DNA measured with a quartz crystal microbalance, SPR,\textsuperscript{22} and surface plasmon diffraction sensors (SPDS).\textsuperscript{56} The average primary DNA surface coverage in those studies was $\sim5 \times 10^{12}$ molecules/cm$^2$. As stated above, the electrostatically adsorbed DNA
coverage in our SPR experiments was approximately 10 times higher, at $2.5 \times 10^{13} \text{ cm}^{-2}$. This difference in coverage likely arises from the differing methods of DNA immobilization; while in our system the DNA is electrostatically adsorbed, other studies utilized a streptavidin-biotinylated DNA linkage for surface immobilization.\textsuperscript{22, 56} High surface coverage of primary DNA significantly reduces the efficiency of hybridization.\textsuperscript{56, 59} In addition, the hybridized duplex of electrostatically adsorbed and covalently bound DNA may be structurally and energetically different. It has been proposed that a preferred structural isomer of an oligonucleotide pair on a positively charged surface is a highly asymmetrical and unwound duplex.\textsuperscript{60} We believe that such non-helical nature of DNA duplex, together with steric effects of highly packed surface play major roles in the reduced affinity for the 16-mer pair used in this study.

We now turn toward developing a model for using SiNW sensors to quantitate complementary DNA pair binding constants, and, if those numbers are known, to determine the solution concentration of the analyte. A discussion of the kinetics of a surface binding assay, as measured within flowing microfluidics environments, is required. Zimmermann and co-workers modeled the kinetics of surface immunoassays in microfluidics environments.\textsuperscript{61} Their model was based on four differential equations: the two Navier-Stokes partial differential equations, the Convection-Diffusion equation, and the ordinary differential equation resulting from the Langmuir binding model (i.e., the binding/hybridization equilibrium). A key result was that in the limit of high analyte flow speeds ($>0.5 \text{ mm/sec}$) (which is the case for all the experiments in this work) the process is reaction limited. Therefore, the amount of analyte that is captured and detected can be described by the ordinary differential equation resulting from the Langmuir binding model:
\[
\frac{d\Theta_t}{dt} = k_{on} C (\Theta_{max} - \Theta_t) - k_{off} \Theta_t 
\]  

(3.3)

Here, \( \Theta_t \) = surface density of bound analyte molecules; \( k_{on} \) = rate constant for association; \( k_{off} \) = rate constant for dissociation; \( C \) = solution concentration of analyte (a constant under flowing conditions); \( \Theta_{max} \) = maximum number of binding sites available per surface area.

Equation (3.3) can be solved analytically to yield:

\[
\Theta_t = \frac{k_{on} \Theta_{max} C}{k_{on} C + k_{off}} \left(1 - e^{-(k_{on} C + k_{off}) t}\right) 
\]  

(3.4)

The challenge is to translate from the resistance change of a SiNW sensor to the analyte concentration, \( C \). However, the exact relationship between a measured resistance change and the surface density of bound analyte molecules is not intuitively clear. Here we attempt to determine the nature of that relationship. We demonstrated above (Figure 3.7D) that the cumulative change in SiNW sensor resistance arising from the binding of a charged analyte (ssDNA) at a concentration-dependent saturation was linearly proportional to the \( \log[DNA] \), similar to what has been reported for VLS SiNW detection of prostate specific antigen (PSA).\(^{21}\) In mathematical terms, this means that as we approach saturation for a given concentration:

\[
\frac{\Delta R}{R_0} = \alpha \ln C 
\]  

(3.5)

where \( \alpha \) is a constant, \( \Delta R = R - R_0 \), \( R \) is resistance at time \( t \), and \( R_0 \) is the resistance at \( t=0 \).

At saturation levels equation (3.4) reduces to

\[
\Theta_t = \frac{k_{on} \Theta_{max} C}{k_{on} C + k_{off}} = \frac{K_A \Theta_{max} C}{K_A C + 1} 
\]  

(where the
binding affinity $K_d = \frac{k_{on}}{k_{off}}$. In the limit where $K_d C \ll 1$ (which is usually the case with values of $C \leq 10^{-9}$ and values of $K_d < 10^8$), this reduces to $\Theta_t = K_d \Theta_{max} C$.

Therefore, at saturation, and with $K_d C \ll 1$, $\Theta_t$ scales linearly with $C$. From our previous discussion, this implies that at saturation $\frac{\Delta R}{R_0}$ scales logarithmically with $\Theta_t$ (or equivalently that $\Theta_t$ is an exponential function of $\frac{\Delta R}{R_0}$ at saturation). In estimating the relationship between resistance changes at all times (not just at saturation) and the surface density of bound analyte molecules at all corresponding times, we start by assuming the same functional relationship that we experimentally observe at saturation. We also impose two boundary conditions: (1) when the measured resistance reaches its
saturation level we would expect the maximum number of binding events to have taken place and for that number to be consistent with the prediction from the Langmuir binding model (equation 3.4), and (2) when the measured resistance is unchanged from its starting level we expect zero binding events (again consistent with the Langmuir model at time=0). Based on these assumptions and boundary conditions we can thus estimate that the surface density of bound analyte molecules as a function of resistance change has the form:

\[
\Theta_t = \frac{k_{on} \Theta_{max} C}{k_{on} C + k_{off}} \left(1 - e^{-\frac{\Delta R}{R_{max} - R}}\right) \quad (R_{max} = R \text{ at saturation}) \quad (3.6)
\]

It is important to note that eq. (3.6) is only a guess based on the relationship observed experimentally. In the next section, we will attempt to derive a relationship which is consistent with the fundamentals of solid state physics. The validity of eq. (3.6) can be tested by considering the following expression that is derived from eq. (3.6) and comparing it to the same expression derived from eq. (3.4):

\[
\frac{\Theta_t}{k_{on} \Theta_{max} C} = \frac{1 - e^{-\frac{\Delta R}{R_{max} - R}}}{k_{on} C + k_{off}} = \left[1 - e^{-\frac{(k_{on} + k_{off}) t}{k_{on} C + k_{off}}}\right] \quad (3.7)
\]

Note that eq. (3.7) is expressing the fraction of bound analyte molecules at time \( t \) relative to the level at saturation in terms of \( \Delta R \) (first term in brackets) and in terms of binding constants (second term in brackets). Time appears explicitly in the second term in brackets, while it is implicit in the first term in brackets (i.e., at a given time \( t \) there is a given \( R \) and \( \Delta R \)). If we plot the first term in brackets in eq. (3.7) (the term containing
ΔR) against the second term in brackets (using \( \kappa_{on} \) and \( \kappa_{off} \) values from an SPR analysis), we find that the two curves are qualitatively similar (Figure 3.16).

A second test of eq. (3.6) is to utilize it to extract binding kinetics. As we can infer from eq. (3.7), if eq. (3.6) is equivalent to the Langmuir binding model (eq. 3.4), then:

\[
\frac{\Delta R}{R_{max} - R} = \left( k_{on} C + k_{off} \right) \times t
\]

(3.8)

We can thus extract \( \kappa_{on} \) and \( \kappa_{off} \) values from measured resistance data.

<table>
<thead>
<tr>
<th>SiNWs - concentration pair:</th>
<th>SPR (this work) (poly-L-lysine surface, 16-mer DNA)</th>
<th>SPDS (ref. 56) (avidin-biotin linkage, 15-mer DNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 nM 100 nM 1 nM 100 nM 1 nM 10 nM</td>
<td>100 nM (actual); 68(52) nM calculated</td>
<td>10 nM (actual); 14(9) nM calculated</td>
</tr>
<tr>
<td>( \kappa_{on} (M^{-1} s^{-1}) )</td>
<td>( 3.5(3.4) \times 10^5 ) ( 4.2(2.4) \times 10^5 ) ( 6.2(9.6) \times 10^5 )</td>
<td>( 1.01 \times 10^5 ) ( 6.58 \times 10^4 )</td>
</tr>
<tr>
<td>( \kappa_{off} (s^{-1}) )</td>
<td>( 3.1(0.5) \times 10^2 ) ( 2.4(0.8) \times 10^2 ) ( 2.4(0.9) \times 10^2 )</td>
<td>( 2.01 \times 10^2 ) ( 1.32 \times 10^4 )</td>
</tr>
<tr>
<td>( K_A (M^{-1}) )</td>
<td>( 1.1 \times 10^7 ) ( 1.8 \times 10^7 ) ( 2.6 \times 10^7 )</td>
<td>( 5.02 \times 10^6 ) ( 4.98 \times 10^8 )</td>
</tr>
</tbody>
</table>

Table 3.3: Kinetic Parameters estimated from SiNW biosensors for the hybridization of 16-mer DNA and corresponding comparisons with analogous SPR and SPDS (surface plasmon diffraction sensor). The calculated concentrations (bottom row) were estimated with eq. (3.6), by using the pair of SiNW measurements that did not include the concentration to be determined. For example, the 1nM and 100nM measurements were used to determine the concentration at 10nM. Standard deviations are given in parentheses.

To extract \( \kappa_{on} \) and \( \kappa_{off} \) values from the resistance versus time data, we used eq. (3.8) to create a series of two equations with two unknowns (one equation from each concentration) which we solved to get the implied \( \kappa_{on} \) and \( \kappa_{off} \). For each concentration in
the pair, we chose to use all data points starting at a time where our model (the first term in brackets in eq. 3.7) indicated a value of 0.63 (i.e., a time equal to the characteristic time of this experimental function) and ending 150 seconds later (time close to saturation, i.e., a value of 1 for eq. 3.7). We chose this part of the data because the assumptions underlying the model indicate that values close to saturation are the ones where our model fits real data the best. For each concentration pair, therefore, we had 150 pairs of equations, each yielding a value for $\kappa_{on}$ and $\kappa_{off}$. To extract the implied concentration values from the resistance versus time data, we used eq. (3.8), this time using $\kappa_{on}$ and $\kappa_{off}$ values obtained from a concentration pair that did not contain the concentration we were trying to estimate. Again, we chose 150 data points from the same portion of the graph used to extract $\kappa_{on}$ and $\kappa_{off}$ values. Each data point yielded one equation in one unknown, which we solved to get the implied concentrations. We then calculated the average implied concentration and the standard deviation for all 150 data points. The $\kappa_{on}$, $\kappa_{off}$, and $K_A$ values, along with the analyte concentrations, are summarized in Table 3.3. The $\kappa_{on}$ constants determined from the SiNW experiments are 3 to 5 times larger than $\kappa_{on}$ obtained with SPR experiments. The nanowire-measured $\kappa_{off}$ values, however, are consistently quite close to those measured with SPR. As stated above, the variation in $\kappa_{on}$ values may be a reflection of steric affects that arise from the unusually high surface density of primary DNA adsorbed onto the poly-L-lysine surfaces that were used for the SPR experiments.\textsuperscript{56, 59}

### 3.5.2 Analysis of IL2-Sensing Experiments

In the previous section, eq. (3.6) was proposed based on the observation of linear relationship between the log of analyte concentration and $\Delta R/R_0$, as well as the
appropriate boundary conditions. It was demonstrated that kinetic parameters and analyte concentrations which are quite close to those measured with SPR could be derived from eq. (3.8). Here, we attempt to ground the derivation of kinetic parameters and analyte concentrations, obtained from the silicon nanowire sensing data, on the fundamentals of solid state physics. We assume, as in section 3.5.1, that the sensing process is reaction limited due to a high flow rate of solution containing the analyte, and that the nanowire response must, therefore, be consistent with the solution (eq. 3.4) to an ordinary differential equation (3.3). Let us define $\Theta_{\text{sat}}$ as a number of analyte molecules captured at saturation on the surface of a sensor:

$$
\Theta_{\text{sat}} = \frac{k_{\text{on}} \Theta_{\text{max}} C}{k_{\text{on}} C + k_{\text{off}}}
$$  \hfill (3.9)

Further, let $\Theta_{f}$ be a fraction of the saturation value of molecules captured at time $t$:

$$
\Theta_{f} = \frac{\Theta_{t}}{\Theta_{\text{sat}}} = \left(1 - e^{-\left(k_{\text{on}} C + k_{\text{off}} \right) t}\right)
$$  \hfill (3.10)

We now turn to some fundamental equations describing the electronic properties of semiconductors.\(^{57}\) The resistivity ($\rho$) of p-type material may be described as follows:

$$
\rho = \frac{1}{q \mu p}
$$  \hfill (3.11)

where $q$ is hole charge, $\mu$ is the mobility of holes and $p$ is the density of holes, given by:

$$
p = n_i e^\frac{(E_g - E_F)}{kT}
$$  \hfill (3.12)
where \( n_i \) is intrinsic carrier concentration, \( E_i \) is the intrinsic energy, \( E_F \) is Fermi energy, \( k \) is Boltzmann constant and \( T \) is the temperature. The resistance is given by \( R = \rho L/A = \rho G \), where \( L \) is the length of the material and \( A \) is the area of the cross section.

\( E_i \) can be rewritten as \( E_i = -qV_G = -qA\Theta_t S \), where \( V_G \) is gate voltage, \( S \) is total charge on a bound analyte molecule and \( A \) is some proportionality constant relating chemical potential and total surface charge (similar to equation 3.1). Rewriting the expression for resistance \( R \) by combining equations (3.11) and (3.12), we obtain (\( G = L/A \)):

\[
R = \frac{G}{q\mu n_i e^{-E_i/kT} e^{-qA\Theta_t S}}
\]  

(3.13)

Solving for \( \Theta_t \) as a function of \( R \) yields:

\[
\Theta_t = \frac{1}{qAS} \left[ \ln \left( \frac{q\mu n_i}{Ge^{E_i/kT}} \right) + \ln \left( R \right) \right]
\]  

(3.14)

We invoke the boundary condition similar to those discussed in section 3.5.1. The number of captured molecules at saturation, when \( R = R_{\text{sat}} \), equals \( \Theta_{\text{sat}} \).

\[
\Theta_{\text{sat}} = \frac{1}{qAS} \left[ \ln \left( \frac{q\mu n_i}{Ge^{E_i/kT}} \right) + \ln \left( R_{\text{sat}} \right) \right]
\]  

(3.15)

At time zero (\( t = 0 \)) there are zero captured molecules (\( \Theta_t = 0 \)) and \( R = R_0 \).

\[
0 = \frac{1}{qAS} \left[ \ln \left( \frac{q\mu n_i}{Ge^{E_i/kT}} \right) + \ln \left( R_0 \right) \right]
\]  

(3.16)

Therefore,
\ln \left( \frac{q \mu n_i}{Ge^{E_i/kT}} \right) = -\ln(R_0) \quad (3.17)

Finally, we can represent the fraction of molecules captured simply in terms of the nanowire resistance:

\[ \Theta_f = \frac{\Theta_t}{\Theta_{sat}} = \frac{1}{qAS} \left( -\ln(R_0) + \ln(R) \right) = \frac{\ln \left( \frac{R}{R_0} \right)}{\ln \left( \frac{R_{sat}}{R_0} \right)} \quad (3.18) \]

Equations (3.10) (Langmuir model) and (3.18) (solid state physics) both represent the fraction of captured molecules and can be equated.

\[ \Theta_f = \frac{\ln \left( \frac{R}{R_0} \right)}{\ln \left( \frac{R_{sat}}{R_0} \right)} \left( 1 - e^{-\left(\kappa_{sat} + \kappa_{off}\right)t} \right) \quad (3.19) \]

Therefore, if the kinetic parameters \( \kappa_{\text{on}} \) and \( \kappa_{\text{off}} \) are known, the concentration can be obtained from the resistance value at any time \( t \).

\[ C = -\frac{1}{\kappa_{\text{on}}t} \ln \left( 1 - \frac{\ln \left( \frac{R}{R_0} \right)}{\ln \left( \frac{R_{sat}}{R_0} \right)} + \kappa_{\text{off}}t \right) \quad (3.20) \]
Figure 3.17: Graphical representation of equation 3.19 ($\Theta_t$) using the kinetic parameters obtained from the SPR experiments (1x PBS) ($\kappa_{on}=4.71e5~M^{-1}s^{-1}$ and $\kappa_{off}=1.92e-3~s^{-1}$) (grey curves) and SiNW resistance versus time data (10μM PBS) (black curves) for four concentrations of IL2. Time zero is the time when pure 10μM PBS was switched to a solution containing indicated concentrations of IL2.

We now attempt to use the above derivation, particularly eq. (3.19), to extract kinetic parameters and analyte concentrations from the SiNW sensing data shown in Figure 3.14A. We first qualitatively compare Langmuir model (right side of eq. 3.19) with the SiNW data (left side of eq. 3.19) (Figure 3.17). Kinetic parameters obtained from SPR experiments (Figure 3.14B), $\kappa_{on}=4.71e5~M^{-1}s^{-1}$ and $\kappa_{off}=1.92e-3~s^{-1}$, were used. As Figure 3.17 depicts, there are significant deviations between the Langmuir 1:1 model (grey curves) and SiNW data (black curves). Of course this comparison is purely
qualitative, and merely serves as a guide. Several major differences exist between the SPR and SiNW experiments, which may have resulted in the discrepancies observed in Figure 3.17. First, the SPR experiments were carried out in 1x PBS (0.14M), while nanowire experiments were carried out in 10μM PBS. We attempted to perform SPR experiments in 10μM PBS; however, the data was highly irregular and irreproducible. Therefore, it is probable that the kinetic parameters are affected by the salt concentration, and may vary between the two sets of experiments. The second important different has to do with the surface area of the sensor. The surface area of a SiNW sensor is \( \sim 10^{-12} \text{ m}^2 \) (ten nanowires), while that of the SPR sensor chip is significantly larger. If we assume the same probe surface density for the two experiments, the number of antibodies available for analyte capture in the case of SPR is much higher. In contrast, fewer surface probes on the surface of SiNWs may lead to “noisier” signal, with larger signal variance.

We can further extract kinetic parameters from the electrical measurements alone, also using eq. (3.19). The decaying part of the curve in Figure 3.14A (when only buffer is flowing and the surface is initially saturated with the analyte) should follow the relation:

\[
\Theta_f = \ln\left(\frac{R}{R_0}\right) = e^{-\kappa_{off} t}
\]

(3.21)

Thus, it is possible to determine the \( \kappa_{off} \) from the decaying part of the sensing data, and then use the obtained value to determine \( \kappa_{on}C \) from eq. (3.19). If the concentration (C) or
\( \kappa_{\text{on}} \) is known, the other parameter may be readily obtained. Taking the log of both sides of eq. (3.21) yields:

\[
\ln \left( \frac{R}{R_0} \right) = -\kappa_{\text{off}} t
\]

\( \ln \left( \frac{R_{\text{sat}}}{R_0} \right) \)
Since we have a large sample of resistance (R) values versus time, we can utilize linear regression analysis to obtain 95% confidence interval for the $\kappa_{\text{off}}$ value for each IL2 concentration. Figure 3.18 shows eq. (3.22) plotted using the SiNW experimental data obtained for different concentrations of IL2. The $\kappa_{\text{off}}$ values indicated are quite close to the off rate obtained with SPR, $1.92\times 10^{-3}$ s$^{-1}$. Table 3.4 summarizes the results of linear regression of the data in Figure 3.18, including the $R^2$ values and the 95% confidence intervals of $\kappa_{\text{off}}$ values for different analyte concentrations. We now use the off-rate values obtained above to determine the $\kappa_{\text{on}}$ and the concentration of the analyte.

Equation (3.19) can be rearranged to yield:

$$-\kappa_{\text{on}} Ct = \ln \left(1 - \frac{R}{R_0} \right) + \kappa_{\text{off}} t = \Gamma(t)$$  \hspace{1cm} (3.23)

Using the appropriate $\kappa_{\text{off}}$ values, $\Gamma(t)$, defined in eq. (3.23), may be graphed versus the negative values of $\kappa_{\text{on}} Ct$ (Figure 3.19). Again, regression analysis yields the values of $\kappa_{\text{on}}$ if we assume that we know exactly the concentration of the analyte. Table 3.4 summarizes the $\kappa_{\text{on}}$ values obtained in this manner as well as the apparent concentration of analyte assuming the $\kappa_{\text{on}}$ value from the SPR experiment. In the range of 1nM to 10nM, the estimated kinetic parameters and apparent analyte concentrations are quite close to real concentrations and the SPR kinetic parameters. However, obvious discrepancies are observed for the 100nM and 100pM IL2 concentrations. Figure 3.17 suggests that serious discrepancies, particularly at these concentrations, should occur, but the origin of such deviations is not obvious. While the off-rates are quite close to the
SPR value at all concentrations, the forward rate, $\kappa_{on}$, is significantly slower for the 100nM and significantly higher for 100pM than what is predicted with the Langmuir model and observed by the SPR (Table 3.4). It is possible that at high analyte concentrations in dilute buffer (10μM PBS), protein-protein interactions become significant and compete with antibody-protein interactions, causing the $\kappa_{on}$ to be underestimated. At very low concentrations, inherently “noisy” measurements due to small surface area of the sensor and the low abundance of the analyte may lead to

**Figure 3.19:** Equation (3.23) plotted using the data obtained from the SiNW experiment while flowing 10μM PBS solution containing different concentrations of IL2. The parts of the data outlined with grey bars were used for linear regression analysis, which yielded slopes ($\kappa_{on}$C) that were divided by appropriate known concentrations to obtain $\kappa_{on}$ values indicated on the graphs. Time zero is the time when the 10μM PBS was switched to 10μM PBS containing IL2.
Table 3.4: Kinetic parameters calculated from linear regression analysis of the SiNW sensing data for four analyte concentrations (equations 3.22 and 3.23). The 95% confidence intervals are included. Concentrations of the analyte (C) are obtained by using the SPR $\kappa_{\text{on}}$ value, while the SiNW $\kappa_{\text{on}}$ values are obtained by using the known analyte concentrations. For comparison, the kinetic parameters obtained from the SPR experiment are also presented.

<table>
<thead>
<tr>
<th>nM</th>
<th>$K_{\text{off}} \times 10^{-3}$ (s$^{-1}$)</th>
<th>95% CI</th>
<th>$R^2$</th>
<th>$K_{\text{on}} \times 10^5$ (M$^{-1}$s$^{-1}$)</th>
<th>95% CI</th>
<th>$R^2$</th>
<th>C (nM)</th>
<th>95% CI</th>
<th>$K_A \times 10^8$ (M$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>2.54</td>
<td>2.49</td>
<td>2.59</td>
<td>0.98</td>
<td>0.12</td>
<td>0.11</td>
<td>0.13</td>
<td>0.82</td>
<td>2.52</td>
</tr>
<tr>
<td>10</td>
<td>2.66</td>
<td>2.56</td>
<td>2.75</td>
<td>0.94</td>
<td>2.32</td>
<td>2.22</td>
<td>2.41</td>
<td>0.94</td>
<td>4.92</td>
</tr>
<tr>
<td>1</td>
<td>2.04</td>
<td>1.94</td>
<td>2.14</td>
<td>0.92</td>
<td>3.48</td>
<td>2.76</td>
<td>4.19</td>
<td>0.33</td>
<td>0.74</td>
</tr>
<tr>
<td>0.1</td>
<td>2.35</td>
<td>2.22</td>
<td>2.47</td>
<td>0.91</td>
<td>60.5</td>
<td>45.2</td>
<td>75.8</td>
<td>0.24</td>
<td>1.29</td>
</tr>
<tr>
<td>SPR</td>
<td>1.9</td>
<td></td>
<td></td>
<td>4.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

significant deviations from the Langmuir model. It is, however, very encouraging that the method derived above yields meaningful values of the kinetic parameters and accurate analyte concentrations in the range between 100nM and 100pM. Further theoretical and experimental work is needed to determine the reasons for the deviations of SiNW sensing data from the Langmuir binding model.

3.6 Discussion

Real-time and label-free detection of DNA 16-mers and proteins with SiNWs was performed. Primary DNA was electrostatically adsorbed onto an amine-terminated SiNW surface and hybridized to the complementary strand in a microfluidics channel under flow. Electrostatic adsorption of single-stranded DNA to poly-L-lysine coated surface has previously been electronically detected at nanomolar concentrations with capacitive methods on lowly doped Si electrodes in 0.015M solution.$^{58}$ The ability to
detect DNA under physiological conditions and at a detection limit of 1fM, as demonstrated in this work, is of significance as it may allow the direct use of biological samples such as serum or tissue culture media. It is likely that because the primary DNA is electrostatically bound and hybridization occurs very close to NW surface, Debye screening does not prevent SiNW based detection. We attempted to carry out hybridization in pure water; however, no changes in SiNW resistance were observed (data not shown), presumably because DNA-DNA repulsion is significant in that case.

SiNWs with significantly reduced oxide coverage exhibited enhanced solution FET characteristics (Figure 3.5) when compared to SiNWs characterized by a native SiO$_2$ surface passivation. Oxide covered, highly doped SiNWs were designed to exhibit a similar dynamic range of DNA detection as the best near-infrared imaging SPR technique$^{62}$-$10nM for 18-mer, corresponding to $\sim10^{11}$ molecules/cm$^2$. When identical nanowires were functionalized by the UV-initiated radical chemistry method, resulting in near-elimination of the Si-SiO$_2$ interface, the limit of detection was increased by two orders of magnitude, with an accompanying increase in the dynamic range. This result highlights the importance of controlling surface chemistry of SiNWs for their optimization as biological sensors. In the future, surface chemistries yielding higher coverage than UV-initiated alkylation may be utilized to passivate and electrochemically convert SiNWs into arrays for multiparameter analysis.$^{38,63}$

Sensing of an important cytokine, interleukin-2, has also been performed. Protein detection is significantly limited by the size of the capture agent. Using antibodies poses a limitation on the ionic strength of the buffer containing the analyte. To circumvent this limitation, aptamers and small peptides must be developed as alternative high-affinity protein probes. However, a combination of an appropriate doping level and surface
chemistry will undoubtedly allow the detection down to a subpicomolar regime, which is more than sufficient for most relevant clinical applications.

Finally, a model that is consistent with both the standard Langmuir binding model and with the fundamentals of semiconductor physics is developed. Kinetic parameters and analyte concentrations that are consistent with SPR values may be extracted from the silicon nanowire experiments. The potential for SiNW sensors to quantitate the concentrations of low-abundance biomolecules within physiological relevant environments is an intriguing one, and we are currently vigorously pursuing this possibility. The most useful application of our model would be in extracting otherwise unknown concentration values once $\kappa_{on}$ and $\kappa_{off}$ values are known. As demonstrated here, SiNW sensors can be used for label-free biomolecule detection at concentrations significantly below the limits of detection for SPR. The robustness of the fabrication technique (SNAP) employed here, which yields nanowire sensors that exhibit reproducible and highly tunable behavior, holds a promise for the future integration of this technology within the clinical setting.
Bibliography


Chapter 4
Electrochemically Programmed, Spatially Selective Functionalization of Silicon Nanowires

4.1 Introduction

The spatially selective biofunctionalization of surfaces has proven to be an enabling capability, beginning with the early work of Fodor and co-workers on utilizing photolabile surface groups to construct DNA libraries, to the use of inkjet technologies for the construction of protein chips. The dip-pen lithography methods from Mirkin’s group represent the current limit of patterning density for protein chips. Chip-based array methods are largely predicated upon the optical detection of the target/probe binding events, which imposes the optical diffraction limit on pixel density. However, electronically transduced detectors, such as chemically gated silicon nanowires (SiNW), may circumvent this limitation. Methods for fabricating ultra-high density circuits of silicon nanowires with excellent conductivity and field-effect transistor properties required for biosensors have been established. To construct an array of sensors, NWs must be functionalized with different receptor probes, such as antibodies or aptamers, against their designated molecular targets. Here we describe an electrochemical approach that, while possibly having multiple applications, should be applicable toward the selective biopassivation of silicon nanowire sensor arrays, and is spatially limited only by the ability to electronically address the individual sensor elements.
Recent advances in alkylation of H-terminated Si surfaces has made it possible to bypass the necessity of chemical modification of the native oxide of silicon.\textsuperscript{11-13} Hydrosilylation produces cleaner, more stable and more reproducible monolayers than silane-based SAMs.\textsuperscript{14} The benefits of utilizing organic monolayers formed on H-terminated Si for biosensing are multifold. First, the removal of the SiO\textsubscript{x} tunneling barrier brings the target/probe pair 1-2 nm closer to the conducting surface,\textsuperscript{15} which, as we have previously reported,\textsuperscript{5} translates into an increase in the sensitivity of the device. Second, the electrical properties of sufficiently small diameter silicon NWs are dominated by the surface characteristics. Removal of what is often an electrically imperfect Si-SiO\textsubscript{2} interface and a disordered oxide film with a high density of trap sites\textsuperscript{16} is desirable. Finally, the difficulty in controlling the smoothness of the SiO\textsubscript{2} layer results in rough and grainy surfaces upon the growth of the siloxane-anchored monolayers.\textsuperscript{17,18}

Alkyl monolayers grown on appropriately prepared H-terminated Si(111) can reflect the atomic flatness of the underlying substrate.\textsuperscript{19,20} Multiple studies have focused on the methods of chemical passivation of silicon via a formation of Si-C bond. Hydrosilylation has been accomplished with radical initiators,\textsuperscript{21,22} through thermally induced\textsuperscript{14,22-26} or photochemical methods,\textsuperscript{27-31} or utilizing Lewis acid catalysts.\textsuperscript{32,33} Alkylmagnesium reagents have been successfully employed for the alkylation of H-terminated Si surfaces,\textsuperscript{34,35} and halogenated surfaces have been alkylated with alkylmagnesium and alkyllithium reagents.\textsuperscript{28,36-40} Furthermore, electrochemical methods of H-terminated silicon functionalization have been explored.\textsuperscript{41} Several research groups have demonstrated an ability to utilize these functionalization methods to nonspecifically attach DNA to silicon surfaces.\textsuperscript{42-44}
Electroactive monolayers have attracted attention due to the growing interest in selective molecular and cellular immobilization on surfaces. Electrochemical activation of hydroquinone\textsuperscript{45, 46} and hydroquinone esters on gold,\textsuperscript{47} as well as oxidation of thiols on Si-SiO\textsubscript{2} surfaces\textsuperscript{48} have been demonstrated for such applications. Notably, Mrksich’s group has accomplished a selective attachment of proteins and cells to monolayers on gold via a Diels-Alder reaction between 1,4-benzoquinone and cyclopentadiene.\textsuperscript{45, 46} Here, we extend this chemistry of hydroquinone terminated monolayers\textsuperscript{49} to H-terminated silicon surfaces. We also demonstrate an alternative method of molecular attachment via Michael addition of thiol-terminated molecules to p-benzoquinone.\textsuperscript{50, 51} In addition, we discuss strategies to reduce biofouling by utilizing mixed monolayers consisting of electroactive molecules and oligo(ethylene glycol).\textsuperscript{52,53} Finally, we discuss an alternative electrochemical functionalization strategy which utilizes reductively driven electroactive molecule that has been “clicked” to the acetylenylated silicon surface.\textsuperscript{54} Such strategy avoids deleterious oxidative damage of the silicon surface, which is detrimental to solution sensing applications.

4.2 Experimental Methods

4.2.1 Materials

Single-polished Si(111): n-type, 550 μm thick, resistivity 0.005-0.02 Ω·cm (Montco Silicon Technologies), single-polished Si(100): n-type, 500 μM thick, 0.005-0.01 Ω·cm (Wacker-Chemitronic, Germany), and silicon-on-insulator (SOI) (100): n-type, 9-18 Ω·cm, 50 nm device thickness, 145 nm buried oxide thickness (Ibis Technology Corp.), substrates were used. All reagents used were of highest purity. Streptavidin-AlexaFluor and
streptavidin-gold conjugates, together with Silver-Enhancement Kit were purchased from Molecular Probes. Dulbecco’s Phosphate Buffered Saline (DPBS), (2.7mM KCl, 1.5mM KH$_2$PO$_4$, 137mM NaCl, 8mM Na$_2$HPO$_4$) pH 7.4, was from Sigma. Molecule D or 2-(Biotinamido)ethylamido-3,3’-dithiodipropionic-acid-N-hydroxysuccinimide ester was purchased from Sigma. Dichloromethane (DCM) was distilled from calcium hydride, tetrahydrofuran (THF) was distilled from sodium/ketone. All other chemicals were bought from Sigma Aldrich and used as received. Molecule C was a generous gift from Dr. C. J. Yu.

4.2.2 Organic Synthesis

Scheme 4.1 demonstrates the molecules used to functionalize silicon surfaces.

Scheme 4.1: Molecules used in this study.
**Molecule A.** 2-(4-(tetrahydro-2H-pyran-2-yl oxy)phenoxy)-tetrahydro-2H-pyran was synthesized according to literature methods. To a solution of THP protected hydroquinone (4.26 g, 15 mmol) in THF, was added t-butyllithium (15 mL of 1.7M solution in pentane) dropwise at -78 °C. The reaction mixture was stirred at room temperature for 2 hours, followed by the addition of 12 mL of 11-bromo-1-undecene, and then stirred overnight at 40 °C. The reaction mixture was diluted with ethyl acetate, washed with saturated NH₄Cl and brine, then dried over MgSO₄. Column chromatography with 1:1 hexane/dichloromethane gave mg (60% yield) of molecule A as a colorless oil. ¹H NMR (CDCl₃, 300MHz) δ 6.99 (m, 1H), 6.84 (m, 2H), 5.81 (m, 1H), 5.29 (s, 2H), 4.95 (m, 2H), 3.91 (m, 2H), 3.59 (m, 2H), 2.58 (t, 2H), 1.62-2.02 (m, 28H)

**Molecule B.** 11-(2,5-dimethoxyphenyl)-1-undecene was synthesized by following the procedure outlined in the literature. To a solution of 2.76 g (0.02 moles) of p-dimethoxybenzene in 20 mL THF was added dry tetramethylethylenediamine (2.19 mL, 14.5 mmol). The solution was cooled to -78 °C and purged under argon. N-butyllithium in hexanes (1.6M, 12.5 mL) was slowly added and the resulting pale yellow reaction was stirred for 2 hours. Subsequently, solution of 11-bromo-1-undecene (4.3 mL, 19.6 mmol) was added. The colorless reaction mixture was stirred for 12 hours after it was warmed up to room temperature. The mixture was diluted with ethyl acetate, washed with saturated NH₄Cl solution, water and brine and dried over MgSO₄. The concentrated pale yellow organic oil was purified by column chromatography with hexane/dichloromethane (500 mL 1:10, 500 mL 1:5, 500 mL 1:3) to afford 3.2g (55% yield) of product as colorless oil: ¹H NMR (CDCl₃) δ 6.74 (m, 3H, -C₆H₃(OCH₃)₂), 5.82 (m, 1H, -CHCH₂), 4.96 (dq, J = 2.0 and
Molecule E. Previously reported method\textsuperscript{57} was modified to synthesize molecule E as follows. To a solution of 50 mL triethylene glycol (375 mmol) in dry dimethylformamide (DMF) was added 3.06 g NaH (60% in mineral oil, 76.5 mmol). After four hours, 12 mL 11-bromo-1-undecene (54.7 mmol) was added and the reaction was left overnight. DMF was evaporated and the resulting oil was diluted in DCM. After several extraction steps with DCM, the solution was dried over Na\textsubscript{2}SO\textsubscript{4} and concentrated. The product was purified by column chromatography with ethyl acetate/hexanes (1:3) to afford 14.67 g of clear oil (88.7% yield): \textsuperscript{1}H NMR (CDCl\textsubscript{3}) \delta 5.7 (m, 1H, -CHCH\textsubscript{2}), 4.9 (m, 2H, CHCH\textsubscript{2}), 3.5-3.7 (d, 12H, -OCH\textsubscript{2}CH\textsubscript{2}OCH\textsubscript{2}CH\textsubscript{2}OCH\textsubscript{2}CH\textsubscript{2}OH), 3.42 (t, 2H, J=7Hz, CH\textsubscript{2}CH\textsubscript{2}CH\textsubscript{2}O-), 2.8 (br s, 1H, -OH), 2.0 (q, 2H, J=7Hz, -CHCH\textsubscript{2}-), 1.55 (qui, 2H, J=7Hz, -CH\textsubscript{2}CH\textsubscript{2}CH\textsubscript{2}O-), 1.25 (br s, 12H, -CH\textsubscript{2}CH\textsubscript{2}-). To 3 g of the above purified undec-1-en-11-yltri(ethylene glycol) in dry DMF was added 0.23 g NaH (2 mmol) and allowed to react for 2 hours at room temperature. Subsequently, 2.84 g of iodomethane (20 mmol) was added and the reaction was left overnight at 40 °C to produce pale yellow solution. The concentrated clear oil was purified on silica gel with DCM/hexanes (1:1) to afford 0.92 g of clear oil (30% yield).

4.2.3 Surface Functionalization

Si(111) and Si(100) substrates were cleaned in piranha (H\textsubscript{2}SO\textsubscript{4}:H\textsubscript{2}O\textsubscript{2} = 2:1) at 90 °C for 15 minutes and etched in degassed 40% NH4F [Si(111)] for 20 minutes or in 2.5%
HF for 15 seconds, respectively. In the case of SOI substrates, wafers were spin-on doped with phosphorosilica film to concentrations of $10^{18} \text{ cm}^{-3}$ as measured with 4-point probe (chapter 2). The top 50 nm Si layer was patterned using conventional photolithography on AZ5214 photoresist and etched in SF$_6$ by reactive ion etching (RIE). In the case of nanowires, scanning electron microscopy (SEM) was utilized to pattern the PMMA e-beam resist. After acetone liftoff, the substrate was cleaned in ALEG (J.T. Baker Microelectronics) at 90 °C for 15 minutes and etched with 2.5% HF for 2 seconds. Immediately after etching, neat degassed (by freeze, pump, thaw cycles) Molecules A or B (Scheme 4.1) were spotted onto the chip. The samples were immediately placed into custom-made quartz chamber, which was then attached to the vacuum line and pumped down (Figure 4.1). Functionalization was carried out for 2 hours with UV (254 nm, 9 mW/cm$^2$) illumination under $\sim$10$^{-5}$ Torr. Chips were rinsed with CH$_2$Cl$_2$, deprotected with either 1% TSA in methanol (Molecule A) or with BBr$_3$ in anhydrous DCM (Molecule B). After deprotection, a 15° to 20° drop in contact angle was observed.
To selectively attach molecules to the silicon wires, functionalized device was fixed into a custom-made Teflon cell so that only the center of the chip was in contact with solution. Individual wires were then connected to the outside metal pins with a thin gold wire using silicon-indium-gold contacts. Wires were oxidized at 700mV (vs. Ag/AgCl) for 5 seconds. Immediately after the oxidation the devices were left in the solution of 10mM Molecule C (1:1 DPBS:MeOH) for 30 minutes. Subsequently, the chips were sonicated in methanol for 10 minutes and streptavidin-AlexaFluor (10nM in DPBS) or streptavidin-Au (10pM with 0.05% Tween20 in DPBS) conjugate was introduced for 5 to 10 minutes. Devices were sonicated in 0.05% Tween20/DPBS for 20 to 30 minutes. In the case of Nanogold streptavidin detection, the nano-particles were amplified with silver enhancement reagents for 20 minutes.

4.2.4 Electrochemistry, XPS, and Optical Microscopy

Electrochemistry of monolayers on bare silicon was performed in custom made cells with VersaStat II (EG&G Instruments). DPBS (see materials section) was used as the electrolyte, with Pt counter and Ag/AgCl (BAS Instruments) reference electrodes. For the determination of molecular coverage, the area under the cathodic peak was converted to the number of molecules through a stoichiometric ratio of 2 electrons to 1 electroactive molecule, divided by electrode surface area and normalized to Si atom surface density (7.8e14 for Si(111) and 6.8e14 for Si(100)). XPS was performed in a UHV chamber that has been described elsewhere.\textsuperscript{58} Experiments were performed at room temperature, with 1486.6 eV X-ray from the Al K\(\alpha\) line and a 35° incident angle measured from the sample surface. ESCA-2000 software was used to collect the data. An approach described elsewhere\textsuperscript{58, 59} was used to fit the Si 2p peaks and quantify the amount of surface SiO\(_x\),
assuming that the oxide layer was very thin. Any peak between 100 eV and 104 eV was assigned to Si$^{+}$-Si$^{4+}$ and fitted as described in the literature.\textsuperscript{60} SiO$_x$:Si 2p peak ratio was divided by a normalization constant of 0.21 in the case of Si(111) and 0.17 in the case of Si(100).\textsuperscript{59} Fluorescence spectroscopy was performed on Nikon Eclipse e800 microscope with a D-eclipse C1 confocal system equipped with 488nm, 546nm and 633nm sources.

### 4.3 Functionalization of Hydrogen-Terminated Si(111) and Si(100) Surfaces with Hydroquinone

#### 4.3.1 Characterization of Electroactive Organic Monolayers

Scheme 4.2: UV-catalyzed functionalization of hydrogen-terminated silicon with hydroquinone containing molecules. I) Functionalization is carried out in vacuum, under UV irradiation, while the surface is covered with neat electroactive molecule. II-III) Deprotection of R group (II) and reversible electrochemical activation (III), which converts hydroquinone (HQ) into p-benzoquinone (Q). IV-V) Two attachment strategies, selective to p-benzoquinone: addition of cyclopentadiene (IV) and thiolated molecule (V) with arbitrary functional group R$^*$. Scheme 4.2 shows the overall strategy of spatially-selective immobilization via a Diels-Alder reaction (Step IV) or Michael addition (Step V). The initial step in organic monolayer formation on H-terminated silicon involves a wet etch of native oxide. We employed an established wet etch (40% NH$_4$F) for preparing atomically flat Si(111).\textsuperscript{61}
However, the flattest Si(100) surface that can be generated by a wet etch is substantially rougher, with (111) facets, and is thus more prone to oxidation.\textsuperscript{62-64} Our SNAP method for NW fabrication can be carried out most easily using silicon-on-insulator (SOI) substrates,\textsuperscript{6} which are commercially available mostly in the (100) orientation. Bonded Si(111) wafers are also available, although they require substantial additional processing steps before being suitable for NW fabrication. Thus, the Si(111) surface is more ideal, while the Si(100) surface is more practical.

The quality of the formed monolayer on silicon is critical in determining the interfacial electrical properties and the susceptibility of silicon to oxidation in air and in aqueous solution under oxidative potentials.\textsuperscript{65, 66} In particular, higher packing density

---

**Figure 4.2:** Survey scan XP spectra of Si(111) (left) and Si(100) (right) surfaces (a-b) cleaned and etched as described, (c-d) functionalized with neat Molecule B and (e-f) neat Molecule A at $10^{-5}$ Torr under UV for 2 hours.
leads to a more stable silicon electrode. The size of the molecule in the chemisorbed monolayer dictates the susceptibility of the surface to oxidation by limiting the packing density. We explored two hydroquinone hydroxyl protecting groups, THP (Molecule A) and CH$_3$ (Molecule B) (section 4.2.2). THP is removed under milder conditions, and is compatible with using tri-ethylene glycol (TEG, Molecule E) in conjunction with the electroactive molecule (Step II in Scheme 4.2). As a co-component of a monolayer, TEG helps prevent the non-selective binding of cells and proteins (Heath group, unpublished data).

In Figure 4.2 we present representative X-ray photoelectron spectroscopy (XPS) survey scans for the photochemical functionalization of Si(111) and Si(100) with Molecules A and B. All of native oxide has been successfully removed via the wet etch, as evidenced by the absence of O 1s peak in Figure 4.2 (a, b) for (111) and (100), respectively. Furthermore, no Si 2p peaks were observed on high resolution XPS scans between 100 BeV and 104 BeV, which would be expected if traces of SiO$_2$ remained. No

![Figure 4.3: Tapping mode atomic force micrographs of Si(111) (left) and Si(100) (right) functionalized with Molecule B.](image)
adventitious C 1s peaks at 285 BeV were observed for either (100) or (111) immediately after the wet etch.

As has been shown previously, atomic force microscopy (AFM) is a useful tool for an assessment of surface stability. Figure 4.3 presents AFM images of Si(111) and Si(100) surfaces functionalized with Molecule B. Atomically flat terraces with monoatomic steps, which resemble those of a well-etched, H-terminated surface, are evident on the (111) substrate. The topography of these surfaces does not change after storage in air or application of small positive potentials (<1V) in aqueous electrolytes. The root mean square (rms) roughness of these surfaces was measured after zooming in on a 0.15μm by 0.15μm area on a single terrace. As expected, the hydroquinone functionalized Si(100) surface is rougher (0.129±0.006nm) than the Si(111) surface (0.065±0.004nm). It is likely that the molecular films on (100) are more disordered than those on (111).

Molecule B was synthesized in order to (1) study the effects of molecular size on monolayer stability, and (2) assess the quality of monolayer on silicon prepared by thermally and photochemically induced methods. Molecule A cannot be reacted with silicon surface under thermally induction, since the tetrahydropyran protected group is thermally unstable; however, Molecule B is stable under thermal or photochemical functionalization. Molecule B, however, cannot be used in conjunction with Molecule E since harsh BBr₃ deprotection of methyl groups destroys Molecule E. Figure 4.4A demonstrates the cyclic voltammetry on Si(111) surface functionalized with Molecule B under thermal conditions, 80 °C for 2 hours. We often found that thermal functionalization affords more stable surfaces with better coverage than photochemical functionalization.
Moreover, Figure 4.4A demonstrates that the redox reaction of hydroquinone coupled to silicon surface is diffusion limited.\textsuperscript{68} The reaction can be written as follows:

\[ Q + 2H^+ + 2e^- \leftrightarrow HQ \]

The linear scaling of peak cathodic current with the scan rate indicates that likely the hydrogen diffusion is the limiting rate of the reaction, not the electron transport through the electrode electrolyte interface.

**Figure 4.4:** A) Cyclic voltammetry of a Si(111) surface thermally functionalized with Molecule B. The peak cathodic current (\( \iota_{\text{peak}} \)) is linearly proportional to the scan rate, \( \nu \). B) CV of S(111) surface photochemically functionalized with Molecule A before (black) and after (grey) tetrahydropyran (THP) deprotection. Scan rate 30 mV/s.
4.3.2 Electrochemical Oxidation of Silicon: Organic Monolayer Density and Surface Orientation

We use cyclic voltammetry (CV) to determine the molecular coverage following the deprotection of hydroquinone (Step III of Scheme 4.2). Figure 4.5 illustrates this approach for a monolayer of Molecule B on Si(111) and Si(100), and molecular coverage together with contact angle measurements are presented in Table 4.1. Coverage was obtained by integrating the cathodic peak after all of the surface molecules were converted to the benzoquinone form on a first oxidation sweep to at least 0.6V and then cycled back to the hydroquinone form. A complete conversion to benzoquinone was achieved at potentials above 0.6V because the integrated current plateaus at that value. To correct for non-Faradaic processes, all samples were cycled once at the beginning of the measurement from an open circuit potential to -700mV. The charge obtained from the integration of that cycle was then subtracted from the integrated cathodic peak. The deprotection of the hydroquinone hydroxyl groups is accompanied by a decrease in the static contact angle of 13° to 20°. This decrease in contact angle was reached within 1.5 and 2 hours during the deprotection step for Molecules A and B, respectively, indicating that the deprotection was complete by this time. Complete deprotection eliminates the possibility of underestimating the molecular coverage. Figure 4.4B demonstrates the CV of Si (111) surface functionalized with Molecule A before and after THP deprotection. As expected, the emergence of the reduction peak can only be observed after the hydroxyl deprotection.
Figure 4.5: Measurements of the coverage of Molecule B on Si(111) and Si(100) by cyclic voltammetry (CV). A) Si(111) (top) and Si(100) (bottom): CV scans with increasing peak anodic potential at a scan rate of 50 mV/sec. B) Coverage of electroactive molecules on Si(111) (■) and Si(100) (○) obtained by integrating the reduction peak. The electrolyte was Dulbecco’s Phosphate Buffered Saline (pH 7.4).
No difference in contact angles was detected between monolayers of Molecules A and B for the same surface. However, comparing (111) and (100) reveals that while the contact angle is similar for Molecules A and B prior to deprotection, both monolayers exhibit a 5° to 7° higher contact angle on (100) after deprotection.

<table>
<thead>
<tr>
<th>Hydroquinone protecting group</th>
<th>Static contact angle /° (before and after –OH deprotection)</th>
<th>Coverage (# molecules/Å²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Si(111)</td>
<td>Si(100)</td>
</tr>
<tr>
<td>CH₃ (Mol. B)</td>
<td>73.7±0.8</td>
<td>54.6±1.4</td>
</tr>
<tr>
<td>THP (Mol. A)</td>
<td>72.0±1.3</td>
<td>53.2±1.2</td>
</tr>
</tbody>
</table>

Table 4.1: Monolayer characterization, separated by substrate and by hydroquinone protecting group. The coverage data has been normalized to per silicon atom, taking into consideration the difference in surface densities of Si(111) and Si(100).

As expected, the coverage of Molecules A and B on Si(100) is lower than on Si(111). Table 4.1 also demonstrates that molecular size plays less of a role on the packing density on Si(111): molecules A and B exhibit similar coverage on (111) but quite different coverage on (100).

Due to lower surface density of atoms, Si(100) is more susceptible to oxidize in an electrolyte solution. Figure 4.6 presents CV scans of an electrochemical oxidation of freshly prepared, hydrogen terminated Si(100) and Si(111) surfaces. The electrode areas are identical, so the current density of Si(100) oxidation is larger, corresponding to a more extensive oxidation. Moreover, Si(100) functionalized with Molecule A is progressively oxidized by cycling between two forms of the molecule, as evidenced by the linear drop in the peak cathodic current in Figure 4.5. On Si(111) surface, such drop in current is not
Figure 4.6: Cyclic voltammetry of freshly prepared, hydrogen-terminated A) Si(111) and B) Si(100) surfaces in DPBS, pH 7.4. Scan rate 30 mV/s. C) CV scans of Si(100) surface functionalized with Molecule A. Inset: peak cathodic current density as a function of time after consecutive oxidation scans. Scan rate 30 mV/s.

evident, arguing that the molecule itself is stable after multiple CV scans. Figure 4.7 trace is the silicon oxidation of H-terminated and organically functionalized (111) and (100) substrates through various processes, including the formation of the molecular monolayer, incubation in air at 20% relative humidity, deprotection of the monolayer to form the hydroquinone, and electrochemical oxidation to form the benzoquinone. The method used to quantify the overlying oxide is presented in greater details in section 4.2.4.
Figure 4.7: A) The fractional monolayer coverage of SiO$_x$ on either H-terminated Si(111) and Si(100) or functionalized with either Molecule A (THP protected) or Molecule B (CH$_3$ protected) through various chemical steps: (a) immediately after functionalization, followed by (b) 135 hours in air at 20% relative humidity, followed by (c) deprotection for 2 hours (in either 1% TSA in MeOH for Mol. A or BBr$_3$ in DCM for Mol. B), followed by (d) chronoamperometry at 700mV (vs. Ag/AgCl) for 5 sec in DPBS (pH 7.4). B) High resolution XP spectra of the Si 2p region of Si(100) that was (a) freshly etched and (b) freshly functionalized with Mol. B, (c) exposed to ambient laboratory air for 135 hours, and (d) electrochemically activated at 700mV (vs. Ag/AgCl) for 5 sec in DPBS (pH 7.4).

*Calculated by dividing SiO$_x$:Si 2p peak area ratio by either 0.21 for (111) or 0.17 for (100), different normalization constants calculated as described elsewhere.$^{59}$

We assumed that the organic monolayer causes minimal attenuation of the photoelectrons, although it is possible that the data presented in Figure 4.7 is slightly underestimated because of this assumption. H-terminated and functionalized Si(100) consistently demonstrate higher degree of oxidation than Si(111), consistent with the literature.$^{62-64}$ In
addition, substrates with monolayers of Molecule A exhibit consistently higher levels of oxide than those functionalized with Molecule B. Our data indicate that the functionalized (100) surfaces are oxidized the same amount as H-terminated Si when stored at room temperature and 20% relative humidity, while Si(111) show extreme chemical stability. We have observed that after eight months of storage, less than half of an equivalent SiO$_2$ monolayer exists on functionalized Si(111) samples. Curiously, Si(100) functionalized with Molecule A showed more oxide after incubation in air than H-terminated Si(100). This may have resulted from trace water present in the Molecule A solution. For both (100) and (111) surfaces, a negligible change was observed in the amount of oxide grown during the deprotection steps of either Molecule A or B.

### 4.3.3 Diels-Alder Reaction and Michael Addition on Silicon

Keeping the immobilized biological molecules in a functional state is absolutely necessary for the development of any sensor array, including NW arrays. One corresponding constraint for this chemistry is that the biological probe attachment step and sensing must be done in an aqueous electrolyte. Alkyl monolayers formed onto H-terminated Si surfaces have been shown to reduce anodic oxidation in aqueous media, however, the quality of the monolayer again plays a critical role here. Figure 4.7 clearly demonstrates that a short pulse of a small positive potential, although sufficient to oxidize all surface hydroquinone molecules (Figure 4.8A), only slightly oxidizes the functionalized Si(111), compared with the non-functionalized substrate. The difference is less pronounced for the case of Si(100). We have observed a direct correlation between the packing density and the amount of anodic oxidation of functionalized silicon in phosphate buffer at pH 7.4. Consistent with the packing density data from Table 4.1, molecular size
plays an insignificant role in anodic oxidation of Si(111), while Si(100) functionalized with Molecule A is oxidized more extensively than that functionalized with Molecule B.

Figure 4.8: A) CV scans of n-Si(100) functionalized with deprotected Molecule A. Dashed line–starting with benzoquinone form and scanning negative of open circuit potential (OCP). Solid line–starting with hydroquinone and scanning positive of OCP. The data was collected at a 50 mV/sec scan rate in DPBS, pH 7.4. B) CV of a deprotected 50% Molecule B–50% decene monolayer on n-Si(111) in the presence of 9mM 2, 4-cyclopentadiene-1-acetic acid; taken at 30 mV/sec in aqueous DPBS (pH 7.4):methanol (20:1) solution. The inset shows the decrease in peak cathodic current density measured at -640mV (vs. Ag/AgCl) over the time of the scan.

The oxidation and reduction peaks of hydroquinone on medium doped (~$10^{18}$ cm$^{-3}$) n-type silicon are observed at approximately 350mV and -500mV (vs. Ag/AgCl) on an
initial CV scan. CV measurements on very similar molecules formed as SAMs on Au surfaces reveal the oxidation peak at approximately 300mV and the corresponding reduction peak at -150mV. The redox potentials of hydroquinone/benzoquinone on the functionalized Si electrodes, however, depend upon the quality of the organic monolayer and shift to larger overpotentials as oxidation of the substrate adds to the surface impedance. Figures 4.5A and 4.6C demonstrate a small decrease in current density with consecutive CV scans and a shift of the anodic peak to higher overpotentials. Consistent with the data reported elsewhere, the open circuit potential (OCP) of alkylated silicon electrodes used in this study was shifted to between -100 and -200mV (Figure 4.8A). Silicon oxidation in an aqueous electrolyte causes an accumulation of electrons on the Si surface, shifting the OCP to negative overpotentials.

The effects of monolayer composition and surface functional groups on the kinetics of Diels-Alder interfacial reactions have been clearly demonstrated for the case of Au substrates. Figure 4.8A shows that the monolayer is completely oxidized and reduced at 800mV and -900mV, respectively. Furthermore, both reduced and oxidized forms of the molecule are quite stable. The relaxation rate of p-benzoquinone is slower than either the Michael addition or Diels-Alder reaction (data not shown). Therefore, it is sufficient to pulse the monolayer for a few seconds without applying either subsequent pulses or holding the Si electrode at anodic potentials. An attractive aspect of using the hydroquinone and Diels-Alder chemistry for spatially selective biomolecular attachment is that the redox reaction is electrochemically reversible, which allows for an accurate determination of reaction rates using CV. Figure 4.8B shows a CV trace for a mixed monolayer in the presence of 9mM 2,4-cyclopentadiene-1-acetic acid. To increase the
packing density and to minimize anodic oxidation, mixed monolayers of electroactive molecule combined with 1-decene were used. In general, using short unsubstituted alkenes in conjunction with hydroquinone leads to denser packing of the monolayer, which renders the underlying substrate less prone to oxidation. The rate of this Diels-Alder reaction could be obtained by fitting the decrease in peak cathodic current to the following equation\textsuperscript{45}

\[ I_t = I_r + (I_o - I_r) \exp^{-\kappa t} \]

Here, $\kappa = 0.005 \text{ s}^{-1}$, which is about five times faster than the rate of addition of cyclopentadiene of equal concentration reported on Au electrodes.\textsuperscript{45} The cathodic current was evaluated to determine the reaction rates because the current associated with the oxidation of the Si surface is convoluted into the anodic peak, as is clearly visible in Figures 4.5 and 4.6C.

We also demonstrated Michael addition of thiolated molecules to 1,4-benzoquinone as an alternative coupling chemistry (Step V of Scheme 4.2) that is not available for hydroquinones bound to Au surfaces. This chemistry is advantageous because the native cysteines of antibodies or other proteins could potentially be utilized,\textsuperscript{51, 69} and thiol-terminated nucleic acids are commercially available. The high-resolution XPS data presented in Figure 4.9 (A, B) reveal clear S 2p and N 1s peaks of an oxidized monolayer on Si(100) which was reacted with thiol-terminated biotin (reduced Molecule D) or cysteine, respectively. A reduced hydroquinone monolayer that was taken through the same chemical exposures exhibited no evidence of sulfur or nitrogen presence. Other small, thiol-containing molecules such as L-cysteine (Figure 4.9B) and glutathione were also reacted with the oxidized monolayer.\textsuperscript{50, 70} Figure 4.10 demonstrates the cyclic voltammetry of hydroquinone on silicon before and after addition of thiol-containing
Figure 4.9: High resolution XP spectra of (left column) sulfur 2p peak at 163.8 eV and (right column) nitrogen 1s peak at 400 eV. Si(100) samples functionalized with Molecule A were either (A, a and c; B, b and d) oxidized at 700mV for 5 sec or (A, b and d; B, a and c) reduced at -800mV for 5 sec in the presence of A) 60μM Molecule D and 0.6mM TCEP or B) 60μM L-cysteine in DPBS, pH 7.4. molecules. In each case, the oxidation and reduction peaks were diminished and shifted to higher overpotentials. The covalent attachment of molecules was validated with XPS (Figure 4.9).
4.4 Selective Functionalization of Silicon Micro- and Nanowires

The Scheme 4.2 strategy could be used to selectively functionalize micro- and nanostructures patterned on silicon-on-insulator (SOI) substrates and to subsequently immobilize proteins onto selected electrodes (Figures 4.11 and 4.12). Microwires were patterned using conventional photolithography, and nanowires were patterned using electron beam lithography. Following the steps outlined in Scheme 4.2, a potential of 700mV (vs. Ag/AgCl) was applied to specific wires. All of the wires were immersed into the solution containing either Molecule C (Figures 4.11A and 4.12) or reduced Molecule D (Figure 4.11C), followed by the introduction of a fluorescent dye-streptavidin conjugate (Figure 4.11) or gold nanoparticles functionalized with streptavidin (Figure 4.12). Figure 4.11A demonstrate that this approach can be used to construct a protein library of two elements.
After a particular protein was coupled to one silicon electrode, biofunctionalization of an additional electrode on the same device did not affect the previously functionalized electrode. It is thus apparently possible to saturate all the active sites on a given silicon electrode and to prevent cross-functionalization between two electrically isolated electrodes. However, it was necessary to clean the surfaces by sonication in 0.05% Tween 20 solution after every addition of protein to remove nonselectively bound proteins, and this step may explain the dark patches appearing on the electrodes in Figure 4.11. The sites vacated by leaving proteins may not be suitable for further molecular attachment. It was possible to assess the degree of nonspecific adsorption on a given electrode by recording the current density during the oxidation. While the current densities upon the oxidation of

Figure 4.11: Optical micrographs revealing the spatially selective functionalization of Si(100) electrodes (B) using Diels-Alder cycloaddition chemistry (A) and Michael addition (C). The SOI substrates were photolithographically patterned with eight 20 μm wide wires. A) After oxidation of the middle wire, the device was treated with DPBS solution containing 10mM Molecule C. A 10nM Streptavidin-AlexaFluor 488 was introduced for 10 min and the device was washed with 0.05% Tween 20. The top wire was then treated in an identical manner; however, Streptavidin-AlexaFluor 568 was used instead. The bottom wire was maintained in the reduced state throughout the experiment. Two fluorescent images were collected using appropriate excitation wavelengths and filters, and then merged into a single image. Background noise arising from the scattering of two different laser sources was subtracted from the image. C) A fluorescence micrograph of spatially selective functionalized using Michael addition. Two out of eight wires were oxidized at 700mV and the device was incubated with 60μM Molecule D/0.6mM TCEP in DPBS. Streptavidin-AlexaFluor 488 was then introduced, and was found to bind only to the activated wires.
the first set of electrodes on a device did not vary, activation of the subsequent electrodes after the device had already been exposed to protein consistently exhibited lower current densities. Furthermore, the electrode edges almost always exhibited more non-specific fluorescent signal than did the central portions. This may arise from sharper field gradients at those edges.

**Figure 4.12:** Electron micrographs revealing the spatially selective immobilization of streptavidin-gold nanoparticle conjugates via Diels-Alder reaction. A) An SOI device with ten micrometer wide wires. Four wires are shown, two oxidized (O) at 700mV and two reduced (R) at -800mV. After incubation of the device with Molecule C, gold nanoparticles conjugated with streptavidin were introduced in DPBS solution with 0.05% Tween 20. Silver amplification is only visible on the oxidized wires. B) An electron micrograph of a different SOI device with four 100 nm wires. The third wire from the left was activated and the above procedure was repeated; however, streptavidin conjugated gold nanoparticles were added without Tween 20.

We extended this approach to the selective biopassivation of 100 nm wide, 50 nm high nanowires (patterned using e-beam lithography) that were spaced too closely together (0.3 μm pitch) to easily resolve using fluorescence microscopy. Therefore, we utilized the well-established Ag⁺ amplification scheme\textsuperscript{71} to assess fidelity of the biopassivation on those nanostructures, and the results are shown in Figure 4.12. In this scheme, streptavidin conjugated gold nanoparticles are bonded to the electroactive site, and a silver layer is grown onto the Au particles. The nanowires exhibited more nonspecific binding than the
microwires. As noted above, the edges of the microwires exhibited more non-specific binding, and this affect may be exacerbated for the case of nanowires. In addition, as Figure 4.12 demonstrates, it was possible to essentially eliminate the nonspecific binding on the microwires by including Tween 20 detergent with the streptavidin. Therefore, even without the use of TEG or other chemical approaches designed to prevent non-specific binding, this method can be used for spatially selective immobilization of proteins.

4.5 Discussion

4.5.1 Mixed Monolayers and Biofouling Minimization

We describe an electrochemical method to selectively functionalize silicon micro- and nanoelectrodes utilizing Diels-Alder and Michael additions. Such route of selective molecular immobilization should be especially valuable for the biopassivation of ultra dense arrays of electronically addressable silicon nanowires\textsuperscript{72} (chapter 2) or other nanostructures such as nanomechanical sensors.\textsuperscript{73, 74} This method avoids the spatial limitations of alternative techniques such as inkjet spotting, or the alignment and registry limitations of dip-pen lithography. While the selectivity of this process, as shown in

![Figure 4.13: Tapping mode atomic force micrographs of Si(111) surface A) hydrogen terminated, B) functionalized with Molecule E, or C) functionalized with Molecule A, after exposure to 90nM streptavidin solution.](image)
Figures 4.11 and 4.12, is evident, it can be substantially improved beyond the results presented here. For example, mixed monolayers incorporating TEG units will minimize nonspecific adsorption of proteins and improve the selectivity of this approach. Our preliminary results on such mixed monolayers demonstrate that the high packing density of TEG terminated alkanes (Molecule E) yields more stable silicon electrodes than the ones functionalized only with hydroquinone-terminated alkanes. Figure 4.13 demonstrates the resistance of such surfaces to protein biofouling. Biofouling does not only cause significant noise in electronic sensing, but will most likely be a hindrance in building a large element library with the electrochemical method described above. However, mixed monolayers composed of the electroactive molecules and tri(ethylene glycol) will most likely decrease the problem of biofouling, while further stabilizing the
The cyclic voltammetry on Si(111) surface with such mixed monolayer is shown in Figure 4.14. As is evident from the presence of a reversible reduction peak, the hydroquinone moiety remains functional at low densities, which can be tuned to yield sufficient selective attachment of probe molecules while minimizing the non-specific binding. Furthermore, electrode stability may be substantially improved by modulating the doping levels, or utilizing p-type Si rather than the highly doped n-type substrates used in this study.\textsuperscript{62, 63} Extending this approach to nanowire circuits (Figure 4.15), however, is going to present a number of challenges, not the least of which will involve analytical methods, since most of the techniques utilized here are not easily translated to high density nanowire circuits.

\textbf{Figure 4.15:} Electron micrograph of ultra-high density, single-crystal silicon nanowires, 15 nm wide, with 35 nm pitch. Top right: zoomed-in image of the nanowires.
4.5.2 Reductive Functionalization

As described above, electrochemical conversion of hydroquinone to $p$-benzoquinone causes appreciable oxidation of silicon surfaces, and may play an adverse role in nanowire electrical sensing. Therefore, our group has embarked on finding alternative strategies of spatially selective, electrochemical functionalization of silicon. In particular, an electrochemical method which relies on the reductive conversion of the electroactive molecule (Scheme 4.3) to its active form will be particularly beneficial in avoiding oxidative damage to the nanowire surface. Scheme 4.3 outlines the steps in electrochemically converting a benzoquinone molecule to a primary surface amine via an intramolecular cyclization and liberation of a lactone group. The electrochemical reduction is irreversible, as demonstrated in Figure 4.16A. Surface amine is a general group which can be utilized for subsequent functionalization with a variety of molecules and biological probes. We employed this chemistry to demonstrate a spatially selective coupling of gold nanoparticles functionalized with N-hydroxysuccinimidyI ester (NHS ester) to the surface amines generated via a reductive electrochemical deprotection (Figure 4.16B).
Recent developments in spatially selective electrochemical functionalization indicate that this field is of interest due to its relevance in a wide variety of applications.\textsuperscript{75-80} This chapter addresses some important issues pertaining to the passivation of hydrogen terminated silicon with electroactive molecules, with emphasis on obtaining the most stable surfaces which are resistant to oxidation in air and in aqueous media. Parameters such as molecular size, monolayer density, and crystal orientation of silicon surface affect the silicon electrode stability in a profound way. We further demonstrate that individually addressed silicon micro- and nanowires may be utilized to create an electrochemically encoded protein library of an arbitrary density.\textsuperscript{81} Direct surface passivation of silicon affords several important advantages. First, reactions involving immobilized electroactive molecules may only be carried out at low voltages (-1V < voltage < 1V) if the native oxide

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.16}
\caption{A) Two CV scans with Si(100) electrode functionalized with benzoquinone molecule (Scheme 4.3). A large reduction peak is evident on the first scan, corresponding to the reduction of benzoquinone to hydroquinone. Intramolecular cyclization results in the irreversible liberation of lactone, evident by the absence of reduction peak on the second scan. B) Si(100) microelectrodes were functionalized as in Scheme 4.3. The middle electrode was subjected to irreversible reduction, followed by the coupling of gold nanoparticles functionalized with NHS ester and silver amplification (experimental methods).}
\end{figure}
of silicon is replaced with a short organic monolayer. Second, directly passivated silicon nanowires are much more effective biological sensors in high ionic strength solution.\textsuperscript{72}
Bibliography


Chapter 5

Silicon Nanowires as Highly Efficient Thermoelectric Materials

5.1 Introduction

Thermoelectrics is an old field, which is now experiencing a tremendous acceleration due to the development of new nanomaterials.\(^1\)\(^-\)\(^6\) The principle of the field of thermoelectrics rests on two effects, discovered in the nineteenth century, which describe the conversion between thermal and electrical gradients in materials.\(^7\)\(^,\)\(^8\) In the Seebeck effect, the temperature difference across a material creates a voltage difference between the cold and hot ends due to the diffusion of thermally excited charged carriers down the temperature gradient. This effect is used for power generation. The Peltier effect, conversely, describes the development of a thermal gradient in the material in which an electrical current is present. The Peltier effect is used in the refrigeration/cooling applications. The major advantages of the thermoelectrics over conventional power generation and refrigeration methods are the absence of moving parts and toxic gases, robustness and the potential of power generation from waste heat. Scalability makes thermoelectric devices attractive for the thermal management of integrated circuits, which is the biggest challenge in microelectronics today. Widespread applications of thermoelectric devices, however, still remain elusive, owning to their poor efficiency in comparison with conventional power generators and compressor-based
refrigeration. Therefore, the limited applications where thermoelectrics found practical use to date are those where the convenience and reliability outweigh the economy. For example, portable beverage storage, computer CPU or an infrared detector cooling are accomplished with thermoelectric refrigerators with bismuth telluride alloys. Also, NASA’s deep space probes use thermoelectric generators as power sources. Clearly, thermoelectrics is a very promising field; however, high efficiency devices are still a major bottleneck.

The efficiency of a thermoelectric material is described by the nondimensional figure of merit $ZT$,

$$ ZT = \frac{S^2 \sigma T}{\kappa} \quad (5.1) $$

where $\sigma$ is the electrical conductivity, $\kappa$ is the thermal conductivity, $T$ is the temperature and $S$ is the thermoelectric power (a.k.a. Seebeck coefficient) defined as the thermoelectric voltage ($V_p$) produced per degree temperature difference:

$$ S = \frac{dV_p}{dT} \quad (5.2) $$

The central issue in the thermoelectrics research is to increase $ZT$. From (5.1), this may be accomplished by increasing the Seebeck coefficient of the material ($S$), minimizing joule heating losses by maximizing $\sigma$, and minimizing $\kappa$ to reduce heat leakage. However, the major challenge in maximizing $ZT$ arises because the three material parameters comprising $ZT$ are not mutually exclusive. The best thermoelectric materials were summarized as phonon-glass electron crystal. Traditionally, the best $ZT$ materials have been heavily doped semiconductors. Insulators have low thermal conductivity, but
also very poor electrical conductivity. Metals, where the thermal diffusion of electrons and holes largely cancels the thermal voltage, have relatively low Seebeck coefficients. In addition, the thermal conductivity of a metal, which is dominated by the electrons at room temperature, is in most cases proportional to the electrical conductivity, as dictated by the Wiedmann-Franz law. Metals, therefore, generally are poor thermoelectric materials. In semiconductors, the thermal conductivity has contributions from both electrons ($\kappa_e$) and phonons ($\kappa_{ph}$), with the majority usually coming from phonons. Therefore, phonon contribution to thermal conductivity can be reduced without causing significant reduction in the electrical conductivity. Lightly doped bulk semiconducting materials, in general, have large Seebeck coefficients, but poor electrical conductivities. While heavily doped semiconductors in bulk do exhibit lower thermopower, it can be enhanced, as we show in this work, by nanostructuring the material to a 1D phonon confinement.

Traditionally, the way to reduce $\kappa$ without affecting $S$ and $\sigma$ in bulk materials was to use semiconductors of high atomic weight, such as Bi$_2$Te$_3$ and its alloys with Sb, Sn and Pb. Thermal conductivity is reduced in the materials of high atomic weight due to the decrease in the speed of sound. However, such approach fails to produce bulk materials with $ZT>3$ required for a widespread application. The development of a better bulk material for thermoelectric applications has, at least for now, reached a deadend. Recent developments in nanomaterial science, nevertheless, hold a significant promise of improving the thermoelectric figure of merit beyond the current limits. Over the past decade, effects of quantum confinement on the thermoelectric properties have received increasing attention. Dresselhaus' work was the first to inspire the study of low-
dimensional structures as a means of improving the electronic performance of thermoelectric materials. The thermopower of a metal or a degenerately doped semiconductor is proportional to the derivative of the log of the density of states (DOS) n(E) with respect to energy, evaluated at the Fermi energy (E_F):

\[
S(T) \propto \left( \frac{\partial \ln n(E)}{\partial E} \right)_{E_F}
\]

By nanostructuring semiconductors with sizes comparable to the electron wavelength, sharp peaks in the electronic density of states are produced. It has been hypothesized that by matching the location of E_F with the peak in the DOS, the thermopower may be tremendously enhanced. Moreover, the mobility and thus the electrical conductivity would also be increased, leading to a further enhancement of ZT. So far, metallic SWNT remains the only 1D system where the behavior described by the Mott formula has been clearly verified experimentally.

It was realized early on that a potential improvement in ZT may rely less on quantum confinement of electrons and more on the phonon dynamics and transport. The size of the material may be tuned to be smaller than the mean free path of the phonons and larger than that of electrons and holes, thus reducing the thermal conductivity without significantly altering the electrical transport. However, while the charge transport in thermoelectricity is almost monoenergetic (E=E_F±kT), phonon-mediated heat transport is broadband. The lowest thermal conductivity in crystalline solids is that of an alloy, and is referred to as the ‘alloy limit,’ which is dominated by the scattering of short-wavelength phonons. By using nanostructures such as superlattices,
nanowires,\textsuperscript{16} and nanoparticles,\textsuperscript{17} it is possible to beat the alloy limit by means of the scattering of mid- and long-wavelength phonons in addition to short-wavelength phonons. At room temperature, the dominant heat-carrying phonons typically have mean free path of 10-100nm.\textsuperscript{6} Thus, the nanostructures which effectively reduce the thermal conductivity have the dimensions of that order. Reports of enhanced ZTs of nanostructures have provided the first set of experimental evidence of the importance of controlling phonon dynamics in thermoelectric materials. Venkatasubramanian et al. have quoted ZT of \(~2.4\) at room temperature of thin-film superlattices of Bi\textsubscript{2}Te\textsubscript{3} and Sb\textsubscript{2}Te\textsubscript{3}.\textsuperscript{15} In addition, PbSeTe/PbTe quantum dot superlattices with ZT \(~1.3\) to \(1.6\) were reported.\textsuperscript{17} More recently, cubic AgPb\textsubscript{m}SbTe\textsubscript{2+m} bulk nanocomposites with ZT \(~2.2\) at 800K were demonstrated.\textsuperscript{18}

Besides the superlattice and quantum dot structures, nanowires are also capable of significantly reducing the phonon-mediated thermal conductivity compared with the bulk materials. Specifically, silicon nanowires (SiNWs)\textsuperscript{16} were shown to exhibit size dependent reduction of thermal conductivity compared with thin-film\textsuperscript{19, 20} and bulk\textsuperscript{9} single-crystal silicon. The majority of the drop in thermal conductivity in silicon nanowires was attributed to boundary scattering. Being the workhorse of the microelectronic industry, silicon is extremely well studied and micro/nanofabrication techniques for silicon are well established. This makes silicon a desirable candidate material for thermoelectric applications. To date, however, no reports on the thermoelectric figure of merit of SiNWs have been published. This is not entirely surprising, since the simultaneous measurement of thermopower, electrical and thermal conductivities on the same nanowire as a function of temperature is very difficult. The
challenges are numerous, including precise control of the doping levels, diameter and length of the nanowires, as well as the high quality electrical contacts to the medium and highly doped SiNWs. In addition, obtaining statistically significant values by measuring many identical nanowires is required. Current VLS technique for growing SiNWs,\textsuperscript{21} while capable of creating high quality wires, falls short of the level of control over multiple parameters which is required for carrying out the comprehensive study into the thermoelectric properties of SiNWs. The SNAP method,\textsuperscript{22} described in previous chapters of this thesis, is an ideal tool for such an endeavor.

In this work, we study the effects of doping, diameter and temperature on the figure of merit (ZT) of silicon nanowires. We demonstrate that, as previously shown,\textsuperscript{16} the thermal conductivity of SiNWs is greatly reduced compared to the bulk material. This reduction occurs without a significant compromise of the electrical conductivity of heavily doped SiNWs, which remains quite high and comparable to the bulk. We further show that, while the thermopower of the heavily doped (∼10\textsuperscript{20} cm\textsuperscript{-3}) nanowires is proportional to the temperature (similar to the bulk), the thermopower of the wires doped a bit lower, at a level of ∼3e19 cm\textsuperscript{-3} exhibit evidence of phonon drag\textsuperscript{7} below room temperature. Phonon drag in bulk semiconductor materials has only been observed at very low temperatures\textsuperscript{23} and significantly lower doping levels.\textsuperscript{9} We believe that the existence of phonon drag in nanowires is due to a combination of low thermal conductivity and increased relaxation times of long-wavelength phonons. Furthermore, we hypothesize that this may be generic to every 1D phonon system. Thus, the separate contributions of high thermopower (due to phonon drag), high electrical conductivity (due to high doping level) and low thermal conductivity (due to phonon boundary
scattering) together yield a significant enhancement of the figure of merit of SiNWs compared with the bulk silicon materials. The highest value of ZT observed is ~1.2 at 200K and 0.9 at 300K. We believe that the ZT above 3 at room temperature may be possible in silicon nanowires, and we discuss strategies to improve the efficiency beyond what is reported here.

5.2 Experimental Methods

5.2.1 Thermoelectric Device Fabrication

The starting substrate for fabrication was SOI Smart Cut wafer that has been thinned down to between 20 and 35 nm thick device layer by thermal oxidation followed by BOE wet etch. The wafers were p-type doped by the spin-on diffusion doping (SOD) method described in detail in chapter 1. Afterward, the remaining thin polymer film was removed by alternating washes with acetone, water and BOE. Platinum nanowires were deposited onto the wafer as described in chapter 1. For 7 to 10 nm wide wires, prior to pattern transfer to silicon epilayer, e-beam lithography (EBL) and Ti/Pt (5 nm/30 nm) evaporation were used to define two large pads approximately 1 μm apart for in-plane silicon contacts (chapter 1). For the thermopower measurements, each device consisted of approximately 10 nanowires, 3 μm long. Sometimes, particularly in the case of highly doped 10 nm wires, more NW (40-100) were used in parallel to obtain reliable conductivity and thermal voltages. For the thermal conductivity measurements, usually 100 to 150 NWs were used to increase signal-to-noise ratio. While the thermal voltages are quoted here for a group of wires, the thermal conductivities and the electrical conductivities are per wire for every device. After sectioning the lattice into an appropriate number of NWs, electrical
heaters were defined by EBL and the evaporation of 10 nm Ti and 100 nm Pt. Next, four metal contacts were fabricated on top of the SiNWs by EBL and the evaporation of 10 nm Ti and 180 nm Pt. Immediately prior to metallization, the device was briefly treated in O\textsubscript{2} plasma (20 mTorr, 20 sccm, 10W) for 30 seconds, followed by the immersion into BOE solution for 4 seconds, water rinse and N\textsubscript{2} blow-drying. This step is necessary to obtain good quality electrical contacts between silicon and titanium. After the acetone liftoff, the device was annealed in forming gas (95% N\textsubscript{2}, 5% H\textsubscript{2}) for five minutes at 475 °C. Large gold pads (10 nm Ti, 250 nm Au) for wirebonding to the chip carrier were defined by photolithography. Scanning electron microscopy (SEM) images of the representative devices used to measure thermopower and electrical conductivity of 10 nm and 20 nm SiNWs are shown in Figure 5.1.
Figure 5.1: A) SEM of a device for measuring thermopower and electrical conductivity. Temperature gradient is created with either one of the two heaters (i) and measured with two 4-point resistive thermometers (iii). Thermal voltage is measured also between the two thermometers (iii). Nanowire resistance is measured with 4-point method, where the current is applied between the outer two leads (ii) and the voltage is measured between the inner two leads (iii). The heaters and all the leads are made of platinum, with a thin titanium adhesion layer. B) Zoomed-in image of A. C) Twenty-one 10 nm SiNWs connected in parallel. D) Eighteen 10 nm SiNWs in parallel with monolithic Si-metal contacts. Inset shows 10 nm SiNWs with monolithic contacts before the deposition of metal contacts. E) and F) are images of representative 20 nm and 10 nm SiNWs, respectively, used in this study. The pitch of the 10 nm SiNW lattice is 50 nm while that of 20 nm SiNW lattice is 35 nm.
For the thermal conductivity measurements, the nanowires and heaters were suspended on platinum leads over an approximately 600 μm by 600 μm area. Polymethylmethacrylate (PMMA 950, 6%) was spun on top of the wafer containing the device (Figure 5.1A). Scheme 5.1 demonstrates the fabrication steps used to suspend the nanowires. A large window was opened in PMMA with EBL that included all of the area from the nanowires to large gold pads, but excluded a ~10 μm by 20 μm area containing the heaters and nanowires, which remained protected by the PMMA. A dry etch consisting of CF₄/He (20 sccm/30 sccm, 10 mTorr, 40W) was used to etch through ~150 nm of oxide (SOI box layer), exposing the silicon handle. Immediately after etch, the device was placed in a custom-made vacuum chamber and silicon was isotropically etched with XeF₂ (2.5 Torr) for 1.5 minutes, releasing the nanowires and heaters from the underlying silicon handle. The PMMA mask remaining on the device was etched in O₂ plasma (20 mTorr, 20 sccm, 55W) for 5 minutes. The device was wire bonded to a custom made chip carrier made from a copper piece with which the wafer made an intimate contact (Figure 5.2B).
The chip carrier was attached to a custom made chip carrier socket such that the copper piece of the chip carrier made intimate contact with the gold-plated copper holder of the cryostat (Figure 5.2 A,C). Before the measurement, the cryostat chamber (Janis VPF-475) was pumped down to \(~10^{-7}\) Torr with Pfeiffer turbo pump TSH-071E.

**Figure 5.2:** A) Cold finger liquid N\(_2\) cryostat with gold-plated copper holder containing chip carrier socket which houses the chip carrier made from a copper piece. Silicon wafer with a fabricated device is fixed on top of the chip carrier. B) High resolution picture of a device wirebonded to a chip carrier. C) High resolution image of A, showing a copper chip carrier making intimate contact with the gold-plated cryostat holder. The temperature of the cryostat holder is controlled by PID temperature controller via a heater and a silicon diode thermocouple.
5.2.2 Electronic Measurements

A single suspended device pictured in Figure 5.3 can be used for the measurement of every parameter comprising the figure of merit (ZT), namely the thermopower, electrical conductivity and thermal conductivity.

**Thermopower and electrical conductivity.** A method described elsewhere\(^\text{24}\) was adapted for the measurements of thermopower of SiNWs. Devices were fabricated as described in

![Figure 5.3: Suspended 10 nm SiNW device for the measurement of thermopower, thermal conductivity and electrical conductivity. Temperature gradient is set up by joule heating of the heater. Current is applied to one of the two heaters (red arrows, \(H_{I1}-H_{I2}\)), and the 4-point resistance of the heater is measured by determining the voltage drop (\(H_{V1}-H_{V2}\)). Thermal voltage (\(V_p\)) is measured between the two thermometers (\(T_{LI1,2}T_{LV1,2}-T_{RI1,2}T_{RV1,2}\)). The temperature difference between the two thermometers is measured by determining the 4-point resistances of both thermometers (left thermometer–voltage (\(T_{LV1}-T_{LV2}\)) and current (\(T_{LI1}-T_{LI2}\))) as a function of heater power and cryostat temperature. Electrical conductivity is measured as a 4-point, where the current is applied through the SiNWs between the outer leads (\(A_1A_2-B_1B_2\)) and the voltage is recorded between the inner leads (\(T_{LI1,2}T_{LV1,2}-T_{RI1,2}T_{RV1,2}\)).
Figure 5.4: A) Experimental setup and instrumentation used for electronic measurements of thermoelectric properties of SiNWs. Keithley 707A switching matrix (a) is used to select between instruments and device connections. The 4-point resistances of the two Pt thermometers (right and left) are recorded with SR830 lockin amplifiers (two per thermometer), (b) and (c) respectively. The Keithley 2400 source-meter (d) is used to apply DC current to the heater for generating a temperature difference. Keithley 2182A nanovoltmeter (e) is used to either record thermal voltage between the two thermometers or to measure 4-point voltage of the suspended heater for thermal conductivity measurements. B) Break-out box (a) is used to connect all the leads on the device to the switching matrix. The device is kept in the cold finger liquid N2 cryostat (b), and the temperature is controlled via a heater and a silicon diode sensor of a PID temperature controller (e). High vacuum (c) is maintained throughout the measurements, and the pressure is recorded with the ion gauge controller (d). C) All of the instrumentation and the temperature controller are controlled with the Labview program. The picture shows the simultaneous recording of the resistances of the two thermometers as the current through the e-beam fabricated heater is increased in a stepwise fashion.
section 5.2.1 (Figure 5.3). The experimental setup is shown in Figure 5.4. Joule heating \( P=I^2R \) from a single heater was used to locally heat the substrate and create a temperature gradient along the nanowires. To obtain thermopower, \( S=dV_p/dT \), thermal voltage and temperature difference between identical points on the nanowires (inner metal leads, Figure 5.3) were independently obtained as a function of 2-point heater power. Keithley 2400 source-meter was used to apply DC voltage and measure DC current through two leads of the heater (Figure 5.3, \( H_{11}-H_{12} \)), while Keithley 2182A nanovoltmeter was used to measure the thermal voltage (an average of 6 measurements). For every device, we performed a test to make sure the recorded voltage is due to thermal gradient, and not to leakages, etc.

While keeping the polarity of the two voltmeter leads the same, we measured the thermal

![Graph](https://via.placeholder.com/150)

**Figure 5.5:** Thermal voltage of 20 nm SiNWs as a linear function of heater power. Both the right (dark grey) and left (black, light grey) heaters were used to confirm the correct sign reversal of the thermal voltage. Because the heaters are identical and are approximately equidistant from the NWs, the temperature gradient and thus thermal voltages are identical for both heaters for the same power dissipated. Reversing the current in a heater does not alter the sign or the magnitude of the recorded thermal voltage. Error bars represent one standard deviation \( (\sigma) \) of six measurements.
voltage twice, first using one heater, then the other. As expected (Figure 5.5), the thermal voltage switches signs as the temperature gradient is reversed. Moreover, the sign of the thermal voltage was always consistent with p-type material where the holes are the majority charge carriers (Figure 5.5). In addition, since the joule heating is independent of the direction of the current, reversing the current in the heater does not alter the sign or the magnitude of the thermal voltage (Figure 5.5).

To convert thermal voltage \( V_P \) as a function of heater power \( W_H \) to thermal voltage as a function of temperature difference \( dT \), two additional measurements were carried out. With the first one, the resistances of two thermometers \( R_{LT-\text{left}}, R_{RT-\text{right}} \) were simultaneously measured as a function of heater power (either left or right heater) (Figure 5.6A). The DC current to the heater was supplied with Keithley 2400 source-meter (Figure 5.3, \( H_{I1}-H_{I2} \)) and the voltage was measured with both Keithley 2400 (for 2-point resistance, \( H_{I1}-H_{I2} \)) and the 2182A nanovoltmeter (for 4-point resistance, \( H_{V1}-H_{V2} \)). The second measurement provided the relationship between the resistances of the two thermometers and the cryostat holder temperature \( dT \), which was controlled with a Lake Shore 331 temperature controller (Figure 5.6B). The resistances of the thermometers in both measurements were obtained with four SR830 DSP lockin amplifiers, two for each thermometer. One lockin amplifier was used to measure the current (Figure 5.3, \( T_{LI1}-T_{LI2} \) or \( T_{RI1}-T_{RI2} \)) while the other was used to measure the voltage drop in a short section around the NWs (Figure 5.3, \( T_{LV1}-T_{LV2} \) or \( T_{RV1}-T_{RV2} \)). All measurements yielded linear functions, which were used to calculate the thermopower as follows (assume left heater was used):
Figure 5.6: A) Simultaneous real-time measurement of resistances of both thermometers (black and red, left y-axis) as a heater power (blue, right y-axis) is increased stepwise. The left thermometer is more responsive because it is closer to the left heater. For each heater power, 80 resistance points were collected, and last twenty of those were averaged. Inset: Linear relationship between the thermometer resistance and heater power, $dR_T/dW_{LH}$. B) Simultaneous real-time measurements of resistances of both thermometers (black and red, left y-axis) as the cryostat temperature (blue, right y-axis) is increased stepwise by 1K. Both thermometers respond identically to the global temperature change. For each temperature, ~800 resistance points were collected and last 150 of those were averaged. Inset: Linear relationship between the thermometer resistance and global temperature, $dR_T/dT$. For convenience here, all resistances are normalized by the initial resistance, $R_0$. Error bars are normalized standard deviations of the thermometer resistances.
\[ S = \frac{dV_p}{dT} = \frac{dV_p}{dW_{LH}} \times \left( \frac{dR_{LT}}{dT} - \frac{dR_{RT}}{dT} \right) \]

Figure 5.6 demonstrates an example measurement for determining the temperature gradient along the NWs as a function of heater power (the term in parenthesis in equation 5.4).

The electrical conductivity of SiNWs was measured with a 4-point method, sourcing DC current through NWs between the outer metal leads (Figure 5.3, A1A2-B1B2) with Keithley 2400 source-meter and measuring the voltage drop between the inner leads (Figure 5.3, T_{LI1,2}T_{LV1,2}-T_{RI1,2}T_{RV1,2}) with Keithley 2182A nanovoltmeter. The voltage was averaged five times for each current. Monolithic in-plane silicon/metal contacts were used.

**Figure 5.7:** Current versus voltage (4-point) graphs of 10 nm wires doped at \(-5\text{e}19 \text{ cm}^{-3}\). The resistance was significantly decreased when monolithic in-plane contacts (grey dots, lower inset) were used, as opposed to contacting nanowires directly (black dots, upper inset). The current was measured between the outer leads and the voltage was measured between the inner two leads. The current is normalized per wire per length for direct comparison. The scale bars in the insets are 500 nm.
for 10 nm SiNWs to minimize the contact resistance (Figures 5.1 C,D and 5.7). This was especially important for wires doped at \( <10^{20} \text{ cm}^{-3} \). The monolithic contacts for 20 nm wires in the doping regime used in this study were not necessary due to negligible contact resistance.

**Thermal conductivity.** The method described elsewhere\(^{25}\) was adapted for the measurements of thermal conductivity of SiNWs, with a few major differences. The method in reference 25 is tailored for nanowires and nanotubes which are fabricated separately (bottom up) from the rest of the device. Once the suspended heaters are fabricated, the nanostructure is then introduced between the two suspended \( \text{Si}_3\text{N}_4 \) membranes. Therefore, the 1D nanostructure is completely decoupled from the substrate except at the two points of attachment to the suspended membranes. Our SiNW fabrication procedure is top down. Therefore, the nanowires are fabricated attached to the underlying oxide, and should ideally be released from the substrate at some point prior to the measurements. We have found it very challenging to suspend both the two heaters and SiNWs. The only way to suspend the nanowires in our case is with the wet etch such as BOE. If the NWs are suspended first, all of the subsequent processing steps tend to cause the NWs to collapse. If the NWs are suspended as a last step, the rest of the suspended structure collapses as a result. We decided to circumvent this problem by only suspending one oxide island containing the SiNWs and both heaters (Figure 5.3), and performing two measurements, one with the wires and one without the wires. Isotropic etch with \( \text{XeF}_2 \) is specific to silicon, and does not etch oxide. Therefore, we first measured the thermal conductance of SiNWs plus the underlying oxide, and, after the selective etch of NWs, we repeated the measurement, obtaining the thermal conductance of oxide alone. Subtracting
the two values yields the thermal conductance of SiNWs. It is important to mention that, other than the XeF₂ etch, the samples were not perturbed between the measurements. In fact, the sample was etched in the XeF₂ chamber while remaining wire bonded to the chip carrier.

**Figure 5.8:** A) SEM of SiNW device suspended over ~600 μm by 600 μm area. B) Zoomed-in SEM of a suspended oxide island containing two heaters and two thermometers. The SiNWs are evident between the thermometers. C) The device from B, but after a selective SiNW etch with XeF₂. Scale bars in B,C are 10 μm.

Figure 5.8 demonstrates a typical suspended device used for the measurements of thermal conductivity of SiNWs before (Figure 5.8B) and after (Figure 5.8C) a XeF₂ etch. The isotropic dry etching of silicon nanowires is selective, and does not alter the device.
The second difference between the method described by Shi et al. and our fabrication process is that the method described here allows the measurement of the full figure of merit, including thermopower, electrical and thermal conductivities, on a single device. Therefore, instead of utilizing the heaters as resistive thermometers, we fabricated the thermometers separately from the heaters. This also allows us to more accurately measure the temperature gradient along the nanowires, since we cannot assume that the oxide underneath the heater is isothermal. The temperature change at the two thermometers ($\Delta T_h$: thermometer closer to the heater, $\Delta T_s$: thermometer further away from the heater) generated by a small increase in heater power (usually $\sim$2-5 $\mu$W 2-point) is measured with SR830 lockin amplifiers as described above. For each thermometer, one amplifier supplies a 150-250nA, 913 Hz sinusoidal excitation current, while the other amplifier measures the voltage drop (on the order of 10-50$\mu$V) across a short section of the thermometer whose width is the width of the nanowire array (~5 $\mu$m). The joule heat ($Q_h$) produced by the heater on the oxide membrane is measured with 4-point method by sourcing a small DC current (50-100$\mu$A). The heat dissipated in both platinum leads carrying the DC current ($Q_L$) is measured by subtracting $Q_h$ from the total (2-point) joule heat dissipated by the heater. According to the derivation performed by Shi et al., the thermal conductance of the material between the two thermometers is

$$G_S = \left( \frac{Q_h + Q_L}{2} \right) \left( \frac{\Delta T_s}{\Delta T_h + \Delta T_s} \right) \left( \frac{\Delta T_s}{\Delta T_h - \Delta T_s} \right)$$  \hspace{1cm} (5.5)$$

$$G_{NW} = G_{S(oxide+NW)} - G_{S(oxide)}$$  \hspace{1cm} (5.6)$$
where $G_{NW}$ is thermal conductance of silicon nanowires. Then, the thermal conductivity of silicon nanowires ($\kappa_{NW}$) can be calculated in the following way:

$$\kappa_{NW} = G_{NW} \times \frac{L_{NW}}{nA_{NW}}$$  \hspace{1cm} (5.7)

where $n$, $L_{NW}$ and $A_{NW}$ are number, length and cross-sectional area of nanowires, respectively.

### 5.3 Temperature Dependence of Silicon Nanowire Thermoelectric Properties

#### 5.3.1 Electrical Conductivity

As mentioned above, the electrical conductivity of a thermoelectric device should be maximized for an optimum figure of merit. In semiconductors, the phonons carry a significant portion of heat, and, therefore, by reducing the diameter of the wires we are reducing the thermal conductivity of the material. If the doping is kept high enough, the electrical conductivity may be optimized without affecting the thermal conductivity. The Seebeck coefficient, however, is expected to decrease with the increase in carrier concentration. First, the electronic, or diffusion, component of the thermopower ($S_d$) for moderately doped semiconductors is affected by the hole concentration through the following relationship:\textsuperscript{23}

$$S_d \propto \ln \frac{n_0}{n}$$  \hspace{1cm} (5.8)
Figure 5.9: A) Four-point resistance of p-type (~5e18 cm⁻³) 20 nm SiNWs as a function of gate voltage ($V_G$). Inset: SEM picture of the device demonstrating top gate Pt electrode separated from the NWs by 10 nm thick Al₂O₃ dielectric. B) Modulation of thermopower, $\Delta V_P/\Delta T$, with the top gate voltage. Each thermal voltage point was obtained by averaging six measurements of $V_P$, and the standard deviations are represented by the error bars. The Seebeck coefficients are obtained from the linear regression analyses, and represent the straight lines. Thermopower was measured for four values of the gate voltage: -1V (green squares), 0V (red triangles), +1V (black circles) and +1.5V (blue diamonds).
where $n_0$ is effective density of states in conduction band and $n$ is the density of carriers ($\text{cm}^{-3}$). In addition, in bulk silicon\textsuperscript{9} and germanium\textsuperscript{26} the thermopower due to phonon drag ($S_{\text{ph}}$) has been shown to decrease significantly as the number of impurity atoms and charge carriers increased. Doping level is an important electronic parameter which is most easily tunable in our system. At the beginning of the project, with a large parameter space mainly unexplored, determining the maximum power factor, $S^2\sigma$, as a function of carrier concentration in SiNWs was crucial. We have studied the effect on the thermopower of carrier modulation inside SiNWs with a top gate. Ten nanometers of $\text{Al}_2\text{O}_3$ dielectric were deposited on top of the nanowires (p-type, $\sim5\times10^{18} \text{ cm}^{-3}$) and a platinum gate electrode was fabricated as shown in Figure 5.9A. Between the gate voltages of -1V and 1V, the resistance of the NWs increased by two orders of magnitude (Figure 5.9A). However, the thermopower was only increased 3.4 times, as indicated in Figure 5.9B. A small enhancement of thermopower, as suggested by equation 5.8 and Figure 5.9, argues that it is probably advantageous to maximize the electrical conductivity of the nanowires rather than attempt to enhance $S_d$. We also expected that possible advantages of phonon drag, evident in very low doped bulk silicon material\textsuperscript{9} (a 2.4 times increase of $S_{\text{total}}$ at room temperature–1 mV/K from 0.4 mV/K–in n-type silicon doped at $10^{16} \text{ cm}^{-3}$ versus $10^{19} \text{ cm}^{-3}$) would be completely offset by a dramatic reduction of the electrical conductivity of low doped SiNWs. Therefore, we chose to concentrate on the relatively high impurity concentration range between $\sim10^{19}$ and $10^{20} \text{ cm}^{-3}$. Figure 5.10 demonstrates the temperature dependence of typical electrical conductivities of 10 nm, 20 nm and 500 nm wide p-type nanowires, with approximate impurity concentrations between $5\times10^{19}$ and $2\times10^{20} \text{ cm}^{-3}$. In calculating the electrical conductivities, we assume that the NWs are effectively 10 nm thick, since the
Figure 5.10: Electrical conductivity (per nanowire) versus temperature of typical nanowires with effective cross-sectional areas of 10 nm x 10 nm (squares), 20 nm x 10 nm (circles) and 500 nm x 10 nm (triangles). The doping level of each NW is indicated.

diffusion doping method (chapter 1) yields the majority of impurities in the top ten nanometers of the film. Degenerately doped (>10^{20} \text{ cm}^{-3}) 10 \text{ nm} and 20 \text{ nm} wide SiNWs, like the bulklike 500 \text{ nm} wide wires, exhibit metallic dependence of electrical conductivities on temperature, namely $\sigma \propto T^{-1}$. The 10 \text{ nm} wide SiNWs doped below 10^{20} \text{ cm}^{-3}, on the other hand, generally behave as semiconductors; their electrical conductivities vary as $\sim e^{-a/T}$. The metallic behavior in 10 \text{ nm} wires is observed only when the SOI film is diffusion doped at $\geq 1000 \degree \text{C}$. Clearly, large discrepancies between 10 \text{ nm} and 20 \text{ nm} wide NWs (equally doped, i.e., same doping temperature) in the values and the temperature dependences of their electrical conductivities argue for carrier scattering at the boundaries and possibly local fluctuations in the number of dopant atoms. Such fluctuations, or variance in the average number (N) of the dopant atoms, are expected to play an increasing role as the width of the wire (and N) is reduced, since the variance scales as $N^{-1/2}$. Our
attempts to fabricate wires with widths smaller than 10 nm (~5 to 7 nm) resulted in a further reduction of the electrical conductivities. Such drop in $\sigma$ could not simply be accounted for by the decrease in the cross-sectional area of the nanowires. Therefore, generating nanowires with widths $\leq 10$ nm and with bulk electrical conductivities is a major challenge. As will be demonstrated later in the text, this is the primary reason why at present 10 nm SiNWs doped below $10^{20}$ cm$^{-3}$ consistently yield lower figures of merit compared to 20 nm NWs. As discussed above, low electrical conductivities of SiNWs doped at the level $<10^{19}$ cause these materials to exhibit a low figure of merit. A way to enhance the electrical conductivity of low doped NWs without affecting the Seebeck coefficient would certainly be beneficial in improving their efficiency. Recently, novel and significant linear electro-optic effect was observed in strained silicon.$^{27}$ Depositing Si$_3$N$_4$ layer on top of silicon compressively strains nitride layer and expands the underlying silicon. Such strain effect leads to the breaking of crystal symmetry of silicon. Jacobsen et al. demonstrated that in properly strained silicon, the bulk refractive index ($n$) varies linearly as a function of external applied electric field ($E$).$^{27}$ Thus, strained silicon is a novel material with new electrical and optical properties. We have applied this concept to SiNWs. The expectation is that lifting the degeneracy in the band structure of silicon strained with Si$_3$N$_4$ layer leads to an increase in electrical conductivity, with a negligible effect on the thermopower. Figure 5.11 demonstrates the electrical conductivities of 10 nm and 20 nm NWs (~5e18 cm$^{-3}$ doping level) before and after the deposition of 0.7 $\mu$m PECVD silicon nitride layer on top of the wires. The electrical conductivity in both cases increases as the silicon is strained, with a more dramatic change observed for the 10 nm SiNWs. Initial attempts to measure thermopower after Si$_3$N$_4$ deposition suggest that there
are no significant effects on the S (data not shown). While more work must be done to understand fully the effects of strain on silicon thermoelectric properties, preliminary data clearly suggest that the electrical conductivity of SiNWs can be significantly improved under a proper strain.

### 5.3.2 Thermal Conductivity

As described in the introduction, the thermal conductivity should be minimized for an optimal ZT. Low thermal conductivity helps reduce the heat leakage down the temperature gradient. In semiconductors, heat is primarily carried by crystal vibrations (or phonons). Typically, at room temperature the dominant heat-carrying phonons have mean free paths of 10 to 100 nm. Therefore, by reducing the diameter of a wire down to these...
dimensions, it is expected that the thermal conductivity decreases dramatically. Model of phonon transport in silicon quantum wires yielded a reduction of the lattice thermal conductivity due to spatial confinement of acoustic phonons. The model took into account the modification of phonon dispersions due to spatial confinement and the change in nonequilibrium phonon distribution due to boundary scattering, leading to the prediction of reduced phonon group velocities and an order of magnitude reduction of the thermal conductivity of 20 nm cylindrical wires versus the bulk. In addition, molecular dynamics simulations have demonstrated two orders of magnitude reduction in thermal conductivity of nanometer diameter NWs compared with the bulk Si thermal conductivity. Finally, the first experimental evidence of the thermal conductivity in SiNWs came from Majumdar et al. As predicted by the theory, the thermal conductivity of 1D silicon nanowires was dramatically reduced. Since the drop in thermal conductivity strongly
correlated with the diameter of the wires, the authors concluded that phonon-boundary scattering was the primary reason for the reduction.

We have carried out the measurements of the thermal conductivity of the SiNWs fabricated with the SNAP method. These nanowires have a rectangular cross section and, due to an RIE etch step in their fabrication, somewhat rougher sidewalls compared with the VLS SiNWs. As explained in the methods section of this chapter, releasing these NWs from the oxide substrate while maintaining the integrity of the suspended heaters presented a major challenge. We, therefore, obtained the thermal conductivities of nanowires on the oxide and of the oxide alone (Figure 5.8). The difference in the two values was attributed to the thermal conductivity of the NWs. The selective removal of the NWs with XeF$_2$ was confirmed once by the drop in the electrical conductivity (Figure 5.12) and again with the

\textbf{Figure 5.13:} Measured temperature dependence of thermal conductivities of heavily doped (~1e20 cm$^{-3}$) SiNWs of different widths: 520 nm (black circles), 25 nm (grey circles), 20 nm (black triangles) and 15 nm (grey triangles).
SEM (Figure 5.8) after the measurement was completed. The thermal conductivities of wires of different widths are shown in Figure 5.13. At room temperature, the thermal conductivity of 520 nm wide (35 nm thick) wires was 112 Wm$^{-1}$K$^{-1}$, similar to the bulk value. Reducing the width of the wire had a profound effect on the thermal conductivity, which decreased to 0.28 Wm$^{-1}$K$^{-1}$ for 10 nm wide (20 nm thick) SiNW. This strongly corroborates the previous finding$^{16}$ of the dominance of phonon boundary scattering on the heat transport in 1D silicon nanowires. It is likely that the surface roughness of SiNWs fabricated with SNAP method significantly contributes to the reduction of NW thermal conductivity.

5.3.3 Thermopower and Phonon Drag

The thermopower of bulk silicon strongly depends on the impurity concentration.$^9$ It is difficult to achieve a Seebeck coefficient above $\sim$400 to 500 $\mu$V/K at room temperature for heavily doped bulk ($\sim$10$^{19}$ cm$^{-3}$) silicon. In this regime, the thermopower is purely due to carrier diffusion ($S_e$), and for a degenerately doped semiconductor, as for metals, is given by the Mott formula$^7$:

$$S_e = \frac{\pi^2 k_B^2 T}{3e} \left( \frac{d \ln \sigma(E)}{dE} \right)_{E_F} \approx \left( \frac{283 \mu V}{K} \right) \left( \frac{kT}{E_F} \right)$$

(5.9)

where the derivative of the logarithm of the conductivity $\sigma(\varepsilon)$ is the reciprocal of the energy scale over which it varies (the Fermi energy $E_F$ for metals). Assuming hole doping occurs in the heavier Si valence band (mass 0.49) leads to $E_F=0.076$ eV=880K and $k_F=0.1$Å$^{-1}$ for $n=3\times10^{19}$ cm$^{-3}$. This leads to an electronic term $S_e(T)=aT$ where $a=0.34 \mu$V/K$^2$. 
Figure 5.14 presents measured Seebeck coefficients versus temperature for SiNWs of different widths and doping levels. All of the degenerately doped (≥10^{20} \text{ cm}^{-3}) SiNWs exhibit $S \propto T$ relation, indicative of the dominance of the carrier diffusion on the thermopower. However, the Seebeck coefficients of SiNWs with impurity concentrations less than 10^{20} \text{ cm}^{-3} show very different temperature dependence. As the temperature is decreased from 300K, the thermopower increases, reaching a peak between 150K to 200K, and then continues to decrease below 150K. This behavior is observed in the 10 nm and 20 nm wide NWs, but not in the 500 nm samples. The inflection of thermopower is indicative of the phonon drag contribution ($S_{\text{ph}}$) to the thermopower.
Phonon drag is observed in metals at low T because phonon lifetimes are long enough to make a measurable addition to the heat flux from electronic transport. At sufficiently low T (<20K), the phonon scattering length saturates to the dimensions of the sample leading to $S_{ph} \sim T^3$ from the phonon specific heat ($\sim T^3$). At high temperatures ($kT \gg \Theta_{Debye}$), the specific heat becomes constant and the total number of phonons available for phonon-phonon scattering scales as $\sim T$ leading to $S_{ph} \sim 1/T$.

For p-type Si, the holes are near the valence band maximum at $k=0$ ($\Gamma$ point). This leads to a Fermi surface with Fermi wavevector $k_f$. The largest momentum (shortest wavelength) phonon modes participating in phonon drag by being excited due to the holes scattering with phonons are longitudinal acoustic with wavevector, $k_{LA} = 2k_f = 0.2 \ \text{Å}^{-1}$ (assuming an impurity doping level of $3 \times 10^{19}$ cm$^{-3}$). The wavelength is $\lambda_{ph} = 2\pi/k = 31 \ \text{Å}$. The speed of sound in the (100) direction for Si is $c_L = 8.43 \times 10^5 \ \text{cm/s}$, leading $\omega_{LA} = c_L k = 1.7 \times 10^{13} \ \text{sec}^{-1}$ with energy $\hbar \omega_{LA} = 0.011 \text{ eV} = 129 \text{K}$. Only phonon wavelengths larger than 31 Å can participate in phonon drag. In metals, $k_f$ is on the order of the reciprocal lattice vector. Phonon wavelengths as short as a few lattice spacings participate in phonon drag. Short wavelength phonon lifetimes are less than long wavelength phonons leading to a small phonon drag contribution in metals.

Normal phonon-phonon scattering conserves crystal momentum and hence cannot dissipate heat. Thus Umklapp processes determine the rate of heat dissipation of the phonons. Umklapp scattering requires at least one phonon to have momentum of the size of the reciprocal lattice vector. The single parameter that sets the energy scale for Umklapp scattering is the Debye energy, $\Theta_{Debye}$. Thus the number of Umklapp phonons available to
dissipate the long wavelength longitudinal phonons dragged by the holes is given by the Bose-Einstein function

\[
N_U = \frac{1}{e^{\Theta_D/T} - 1}
\]

leading to a scattering rate \(1/\tau_{ph} \sim N_U \sim 1/\exp(\Theta_D/T) - 1\). When \(T >> \Theta_D\), \(1/\tau_{ph} \sim T\). Since \(\Theta_D = 640\,\text{K}\) for Si, the full Bose-Einstein expression must be applied for the temperature range of interest here, \(T < 300\,\text{K}\).

The total thermopower \(S\) is the sum of the electronic term \(S_e\) and the phonon term \(S_{ph}\). The high temperature \((T > 200\,\text{K})\) data of the 20 nm wire with doping \(n = 3 \times 10^{19} \,\text{cm}^{-3}\) fits the expression

\[
S = S_e + S_{ph} = aT + b\left[e^{\Theta_D/T} - 1\right]
\]

where \(a\), \(b\) and \(\Theta_D\) are varied to obtain the best fit (Figure 5.15). The maximum error is found to be 6.1 \(\mu\text{V/K}\) and the rms error is 1.8 \(\mu\text{V/K}\). If the Debye energy is fixed at its experimental value, \(\Theta_D = 640\,\text{K}\), the best fit occurs for \(a = 0.4 \,\mu\text{V/K}^2\) and \(b = 12.2 \,\mu\text{V/K}\). The maximum error and rms error are 8.9 \(\mu\text{V/K}\) and 2.1 \(\mu\text{V/K}\) respectively (data not shown).
Figure 5.15: Thermopower calculation, plotted along with experimental data (red points) from a 400 nm² (20 nm× 20 nm) cross section Si NW p-type doped at 3×10¹⁹ cm⁻³. The black curve is the fitted expression for the total thermopower shown by the red points. The green curve is the phonon contribution $S_{ph}$ and the blue line is the electronic term $S_e$ arising from the fit. The coefficients are $a = 0.337 \mu V/K^2$, $b = 22.1 \mu V/K$ and $\Theta_{Debye} = 534$ K. The fitted electronic coefficient value is almost identical to our estimated value of $a = 0.34 \mu V/K^2$ assuming hole doping of Si with no Boron impurity band. This demonstrates that phonon drag provides a consistent explanation for the observed thermopower. Moreover, the black data points are experimental values for bulk wires with doping $2 \times 10^{20}$ cm⁻³ (plus signs), 10nm wires with doping $7 \times 10^{19}$ cm⁻³ (open diamonds), and 20nm wires with doping $1.3 \times 10^{20}$ cm⁻³ (open triangles) for which only a linear T electronic contribution was observed. These data points are close to the linear electronic contribution from the red data points and demonstrate that the magnitude of the extracted electronic term is reasonable.

As stated above, the phonon drag in bulk heavily doped silicon is very small because of shorter lifetimes of dragged phonons. Only when the bulk silicon is weakly doped can the phonon drag contribute significantly to the thermopower, raising it to >1 mV/K at room temperature.⁹ As Figure 5.14 demonstrates, heavily doped silicon nanowires exhibit phonon drag behavior at relatively high (200-300K) temperatures.
While the absolute thermopower of heavily doped nanowires is not significantly higher than the bulk value, phonon drag induced deviation from the linear $S \propto T$ behavior yields a 4 to 5 times increase in $S$ at $T \sim 200K$, translating into an order of magnitude increase in $ZT$. Importantly, a large electrical conductivity is maintained in these wires because of the high impurity concentration (Figure 5.10).

### 5.4 Discussion

The data presented in this chapter clearly suggest that there are multiple benefits of using nanostructured materials as thermoelectrics. Particularly, while bulk silicon has a poor thermoelectric figure of merit, nanowires made from single-crystal silicon exhibit $ZT$ values that are over two orders of magnitude higher than those of the bulk Si (Figures 5.16 and 5.17). Such dramatic improvement arises in large part due to an ability to independently optimize the Seebeck coefficient, electrical conductivity and thermal conductivity in a nanowire–something that is impossible to accomplish in bulk semiconductors.

We believe that the electronic properties of 10 and 20 nm SiNWs are those of 3D material, and 1D DOS arguments do not apply in this case.\textsuperscript{13} The improvement in the $ZT$, therefore, comes primarily from the modified phonon transport.\textsuperscript{6} The drop in thermal conductivity is a number-one reason for improved efficiency in SiNWs (Figure 5.13). As others have reported,\textsuperscript{16} and as has been verified by this work, the thermal conductivity of SiNWs decreases by over two orders of magnitude at room temperature compared to the bulk Si, evidently due to phonon-boundary scattering. The major challenge is to keep the electrical conductivity of SiNWs, especially those with diameters less than 20 nm, from
Figure 5.16: Figures of merit of 20 nm SiNWs compared with those of 500 nm x 20 nm bulk Si. The impurity concentrations are indicated in the legend. The left column of the legend is plotted on the left y-axis (grey circles, black squares, open squares), while the right legend column is shown on the right y-axis (black and open triangles).

drastically decreasing compared to the bulk (Figure 5.10). For 10 nm wires, this can only be accomplished at this time by degenerately doping the NWs, to a $>10^{20}$ cm$^{-3}$ level, at temperatures $\geq 1000$ °C (Figure 5.17). Any reduction in the impurity concentration below this level, or thinning the degenerately doped NWs via a dry etch, produces a sharp drop in the electrical conductivity and a corresponding decrease in the ZT (Figure 5.17). The Seebeck coefficient of these NWs is approximately the same as of the heavily doped bulk
(Figure 5.14), and varies as $\sim T$. Therefore, as demonstrated in Figure 5.17, the ZT of degenerated doped 10 nm SiNWs approaches one at room temperature, and most likely increases further at higher temperatures, until the thermal conductivity becomes appreciable. Despite higher electrical conductivity, the figure of merit of 20 nm degenerately doped wires is somewhat smaller than 10 nm at room temperature because of higher thermal conductivity. However, when the doping of the wires is reduced to $\sim 5 \times 10^{19}$ cm$^{-3}$, at temperatures between 150K and 300K, phonon drag contribution dramatically increases the thermopower (Figure 5.14), elevating the ZT at 200K above 1.0 (Figure 5.16). The phonon drag has only been observed in the lightly doped bulk silicon. We believe that 1D phonon confinement increases the lifetimes of long-wavelength phonons, presumably favoring phonon-hole scattering. Therefore, we are able to observe the phonon-drag effect in heavily doped SiNWs.

The results described above reaffirm that the field of thermoelectrics will benefit tremendously from the emerging progress in nanotechnology. No longer is the search for an ideal bulk material with the efficiency superior to Bi$_2$Te$_3$ alloys a bottleneck. Thermal and electrical properties of nanomaterials generally make them superior to the bulk materials in terms of the thermoelectric efficiency, as has already been demonstrated by several groups.$^{15,17}$ Clever manipulation of the electronic$^{10}$ and phonon$^6$ transport in the nanomaterial to enhance the thermoelectric figure of merit is a critical new requirement. Here, we demonstrate that a ZT of $\sim 1.0$ over a broad temperature range is possible in p-type 10 nm and 20 nm wide silicon nanowires, a value which is over 100 times higher than that of the bulk silicon. Nevertheless, it is not unreasonable to think that the figure of merit of silicon may be improved even further.
Figure 5.17: A) Measured ZT values of 10 nm SiNWs of different doping. Anisotropic etching of the NWs was done in the RIE using CF$_4$/He gases (20/30 sccm, 10 mTorr, 40W) for ~40 sec at a time, removing approximately top 3 to 4 nm (Si and B atoms). B) Electrical conductivity of the samples shown in B. Doping below $10^{20}$ cm$^{-3}$ (or etching degenerately doped SiNWs) is accompanied by the reduction in the electrical conductivity and the emergence of a semiconductor behavior.
Bibliography