ABSTRACT

Title of Document:	DEVELOPMENT OF CARBON NANOTUBE FIELD-EFFECT TRANSISTOR ARRAYS FOR DETECTION OF HER2 OVEREXPRESSION IN BREAST CANCER
	Konrad Hsu Aschenbach, Ph.D., 2011
Directed By:	Professor Romel D. Gomez, Department of Electrical and Computer Engineering

We developed a carbon nanotube biosensor platform that was deployed at the National Cancer Institute and successfully detected the HER2 oncogene in real cancer cells at clinically relevant levels. HER2 is a receptor protein that resides on the surface of certain cancer cells and is associated with higher aggressiveness in breast cancers. Overabundance of HER2 at the chromosomal, cell surface, and intermediate gene expression levels can all indicate a dangerous HER2 status.

At the present, testing for HER2 status requires labor-intensive laboratory procedures using expensive reagents. Cost remains the major barrier to widespread screening. We propose an integrated electronic testing platform based on direct label-free gene detection. The system would integrate the various labor-intensive processes that are usually performed by skilled laboratory technicians. The heart of the system is an array of carbon nanotube field-effect transistors that can detect unlabelled nucleic acids via their intrinsic electric charges.

We developed a scalable fabrication technique for carbon nanotube biosensor arrays, hardware and software for data acquisition and analysis, theoretical models for detection mechanism, and protocols for immobilization of peptide nucleic acid probes and hybridization of nucleic acids extracted from cells. We demonstrated detection of HER2 from real cell lines which express cancer genes, thereby lowering the technological barrier towards commercialization of a low-cost gene expression biosensor. The system is suitable for lab-on-a-chip integration, which could bring rapid, low-cost cancer diagnoses into the clinical setting.

DEVELOPMENT OF CARBON NANOTUBE FIELD-EFFECT TRANSISTOR ARRAYS FOR DETECTION OF HER2 OVEREXPRESSION IN BREAST CANCER

By

Konrad Hsu Aschenbach

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Romel D. Gomez, Chair
Professor Neil Goldsman
Professor John Melngailis
Professor Christopher C. Davis
Professor Lourdes Salamanca-Riba

Dedication

To my brother, mother, and father,

who've been with me since the beginning,

to all the wonderful teachers I've had, and

to all the people who've helped me in

some way or another.

Acknowledgements

I first must express my gratitude to my advisor Professor Gomez for being a great professional mentor as well as occasionally taking the role of father figure. Starting with my undergraduate days, he welcomed me into his laboratory with open arms and provided many opportunities to contribute to graduate level research. When I returned as a Ph.D. student, he warmly welcomed me back to tackle this challenging biosensor project. Along the way, he has never ceased to be friendly and supportive, especially during the intense periods of challenge as should be expected in an endeavor of this magnitude. He continues to be a source of inspiration and encouragement and his support has been vital to my professional and personal development.

I also thank the dissertation committee members Prof. Goldsman, Prof. Melngailis, Prof. Davis, and Prof. Salamanca-Riba for their time and effort to evaluate and improve my dissertation.

I gratefully thank Prof. Julius Goldhar, as well as Profs. Goldsman, Melngailis, and Gomez for supporting my successful application to the Department of ECE Distinguished Dissertation Fellowship.

The research presented in this dissertation is a multidisciplinary effort involving people from many fields, including chemistry, molecular biology, medicine, microelectronics, semiconductor physics, and electrical engineering. As the senior graduate student leading this project, I humbly thank the many who have contributed to the process culminating in this dissertation.

I am indebted to our collaborators at the National Cancer Institute (NCI) Oncogenomics Section, headed by Dr. Javed Khan, M.D., who graciously made himself,

his laboratory and staff available for our research. I worked closely with Dr. Belhu Metaferia to develop the probe attachment chemistry, and with Dr. Young Song and Dr. Jun Wei to develop the hybridization protocols. I also gratefully acknowledge the many others at the NCI who were not directly involved in this work but who welcomed our presence nonetheless.

Many thanks go to our friendly collaborators headed by Dr. Stephanie Getty at the NASA Goddard Space Flight Center.

As all of our chips were fabricated at the Laboratory for Physical Sciences, I acknowledge Dr. Charles Krafft as well as Dr. Cliff Hull for their sponsorship. I extend my gratitude to the LPS cleanroom staff headed by Toby Olver and including Steve Brown, Sean Flannery, Warren Berk, and Scott Horst. Thanks to Dan Hinkel for sharing the scanning electron microscope and his expertise on it. Thanks to Russell Frizell (despite his best efforts), Les Lorenz, and J.B. Dottelis. *Danke schön* to Dr. Michael Dreyer for converting the original DOS-based measurement software to work in a graphical computing environment.

I am ingratiated to Prof. Michael Fuhrer and his former students Dr. Daniel Lenski and Dr. Enrique Cobas in the Department of Physics for sharing the expertise and equipment for growing carbon nanotubes. Thanks go to Dr. Akin Akturk and Dr. Zeynep Dilli in the ECE Department for their help with device modeling, simulations and layout. I thank the good fellows Tom Loughran, John Abrahams, and Jonathan Hummel at the Fablab for their support with atomic layer deposition.

There core nucleus of the project consisted of faculty researchers and graduate students on campus. Thank you (感謝) to Mrs. Min Du for her hard work in the

cleanroom fabricating hundreds of biosensor chips, freeing me to attend to other duties; Jennifer Evangelista Oliva for assuming the tedious task of culturing cancer cells as well as staying late to run biosensor hybridizations; Sowmya Subramanian for measuring countless transistors at the probe station; Dr. Jookyung Lee, Dr. Seokhwan Chung, and Jin-seock Ma for providing iron seed coatings for carbon nanotubes to grow on the aforementioned chips. I am also indebted to my colleague Dr. Herman Pandana, with whom I worked closely to start this project from scratch.

Thanks to Dr. Zoraida Aguilar for sharing her expertise in chemistry, Dr. Ronald de los Reyes for his expertise in materials science, and to our former visiting scholars, Kristine-Ann Buela and John Fadul, for teaching us molecular biology.

Thanks to our former graduate students William Chao-Wei Chen, Jui-Ping Chiang, Harita Tenneti, Saeed Esmaili, and former undergraduates Suchit Bhattarai, Jeremy Tsang, Naveed Haghani, Arman Arkilic, and Arman Fallahkhair.

Many thanks to the alumni of our group not mentioned above, namely: Dr. Sylvia Florez, Dr. Hyuncheol Koo, Dr. Li Gan, and Marcia Golub-Miller.

Thanks to Dr. Jonghee Lee, Dr. Hui Wang, Dr. Taesoon Kwon, Dr. Hanhee Paik, Dr. Dong Hun Park, Dr. Junghwan Kim, Anita Roychowdhury, Filiz Yesilkoy, Vitaley Zaretskey, and Nathan Siwak for their friendship and support in and out of the lab.

Finally, my apologies to the many others to whom I owe my gratitude but, by a slip of my memory, do not appear on this list. I completely understand if you wish to return the favor.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

This dissertation describes the work I have done during my PhD studies on the development of carbon nanotube transistor arrays for the label-free detection of the HER2 breast cancer gene. Each year in the United States, nearly 200,000 women are diagnosed with breast cancer and 40,000 will die from it. The "War on Cancer" in 1971 initiated a search for new cancer diagnostics, therapies, preventative measures, and early detection. Yet today, cost remains the major barrier to widespread screening.

We believe that an integrated electronic testing platform based on direct label-free gene detection may be the answer to reducing cost. The system would integrate the various labor-intensive processes that are usually performed by a skilled technician in a laboratory. The heart of the system would be an array of transistors that can detect unlabelled nucleic acids via their intrinsic electric charges. We probed for the HER2 breast cancer gene as a model biomarker on which to prove the detection system. HER2 status has already been shown to be effective in determining if a patient will respond to the drug Herceptin®, which blocks HER2 from taking disastrous effect.

We embarked on an engineering project to provide a system to be used by molecular biologists at the National Cancer Institute to determine the HER2 status of breast cancer cell lines. We constructed three dependable experimental platforms that provided results repeatedly, an improvement compared to our previous prototype design. My colleague Herman Pandana and I began this research completely from scratch, with no existing fabrication process and no laboratory experience in molecular biology. We discovered that the original goals of the endeavor were far more than one or even two graduate students could accomplish in the course of one Ph.D.

In the initial phase of the project, we made the following contributions:

- Demonstrated a new approach of making carbon nanotube field-effect transistors (CNTFETs) with an oxide surface intended for bio-sensing applications,
- Fabricated and implemented reproducible CNTFETs having intrinsic ambipolar transport characteristics,
- Developed protocols for attachment of synthetic DNA probes onto the CNTFETs,
- Demonstrated detection of specific synthetic DNA genes in the range of 10 – 100 nM using CNTFETs.

Using this initial work as a foundation, and with the help of a number of graduate students and faculty, I was able to make a number of contributions:

- Developed scalable fabrication technique for carbon nanotube biosensor arrays,
- Developed hardware and software for data acquisition and analysis,
- Developed theoretical models for detection mechanism,
- Developed protocols for immobilization of peptide nucleic acid probes and hybridization of nucleic acids extracted from cells,
- Demonstrated proof-of-concept with synthetic DNA targets,
- Demonstrated selectivity, sensitivity, and system reliability,
- Demonstrated biological validation of technique,
- Detected the HER2 oncogene in real cell lines that express cancer genes.

In summary, it is very gratifying to see the project come to fruition: we were able to use our system to discriminate between HER2 positive and negative cell lines. My project involved physics, chemistry, device fabrication, biology, computer hardware and software and theoretical modeling. I have integrated these diverse fields, and subsequently lowered the technological barrier towards commercialization of a rapid, low-cost nucleic acid detection system that will be suitable for lab-on-a-chip integration. My hope is for my work to have a positive impact on patients' lives.

The organization of this dissertation is as follows: We begin with Chapter 2 by introducing the carbon nanotube (CNT) and the fabrication process of CNT field-effect transistors. We also present experimental results focusing on biosensor chip production. In Chapter 3, we will review molecular biology and relevant assays in breast cancer detection. Chapter 4 describes the basic principles of electrochemistry and their application to charge detection via the CNTFET. Our experiments at the National Cancer Institute to detect HER2 from cancer cells are detailed in Chapter 5. Finally, Chapter 6 outlines proposed items for future work.

This dissertation attempts to provide all the details necessary for someone to reconstruct the system and reproduce the results. Our ultimate motivation is to create enough momentum such that investors would be willing to fund a continuation, if not multifold expansion, of the project.

Chapter 2: Carbon Nanotubes, Biosensor Fabrication, & Electrical Testing

2.1 Chapter Purpose

The purpose of this chapter is to impress upon the reader that a large number of CNTFET biosensors were produced from a custom design and went through a comprehensive testing process. We provide the reader with a basic atomic and electronic structure of carbon nanotubes and illustrate in detail how we fabricate biosensors based on nanotube transistors. We review the physics of carbon nanotubes to explain their intrinsically high mobility, which allows for high-sensitivity transistors. In sensing applications, extraordinarily high sensitivity of CNTs unfortunately comes with poor selectivity. By coating the CNTs with oxide, we purposefully compromise on sensitivity to bring improvements in selectivity and reliability.

We designed and fabricated arrays of metal-gated transistor arrays using standard lithographic techniques and carbon nanotube "mats". Over 200 biosensors were fabricated and samples were tested using an infrastructure of custom instrumentation and software. We first review performance at transistor level measurements, then present statistical results showing improvements in the quality of production of CNTFET arrays. We gauge the stability of the biosensor and provide possible explanations for our observations.

2.2 Principles of Carbon Nanotubes

In 1991, S. Iijima reported the discovery of a new form of carbon: graphene tubes ranging from 4 to 30 nm in diameter.[1] These carbon nanotubes (CNTs) exhibit a range of electronic properties, ranging from metallic to semiconducting, that depend on their diameter and chirality (wrapping angle).

2.2.1 Atomic & Electronic Structure

A carbon nanotube (CNT) can be modeled as a cylinder constructed by slicing a strip out of a sheet of graphene and joining the long sides together (Figure 1). As a consequence of quantum mechanics, the atomic and electronic structure of the resulting CNT significantly depends on the orientation of tube axis (**T**) with respect to the hexagonal graphene lattice, as well as the diameter of the tube.

The chirality of the tube is described by the wrapping vector **C**, which is orthogonal to the tube axis and defines the relative location of two equivalent locations in the lattice. The chirality of the tube is specified by a pair of integers (n, m), such that $\mathbf{C} = n\mathbf{a}_1 + m\mathbf{a}_2$, where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the unit vectors of graphene. The diameter of the tube is $d = |\mathbf{C}|/\pi$. If $n = m$, the chirality is called "armchair" and the wrapping angle is $\varphi = 0°$. If $m = 0$, the chirality is called "zigzag" (the atoms along **C** appear to zigzag), and $\varphi = 30°$. Otherwise, the CNT is simply "chiral", where $0 < \varphi < 30°$. The wrapping angle can also be specified relative to the zigzag direction by θ, where $\varphi = 30° - \theta$.

Figure 1 shows an atomically resolved scanning tunneling microscope (STM) image of a chiral nanotube. The STM is able to resolve atomic-scale features by maintaining a conductive tip within quantum tunneling distances of a surface.[2] For this

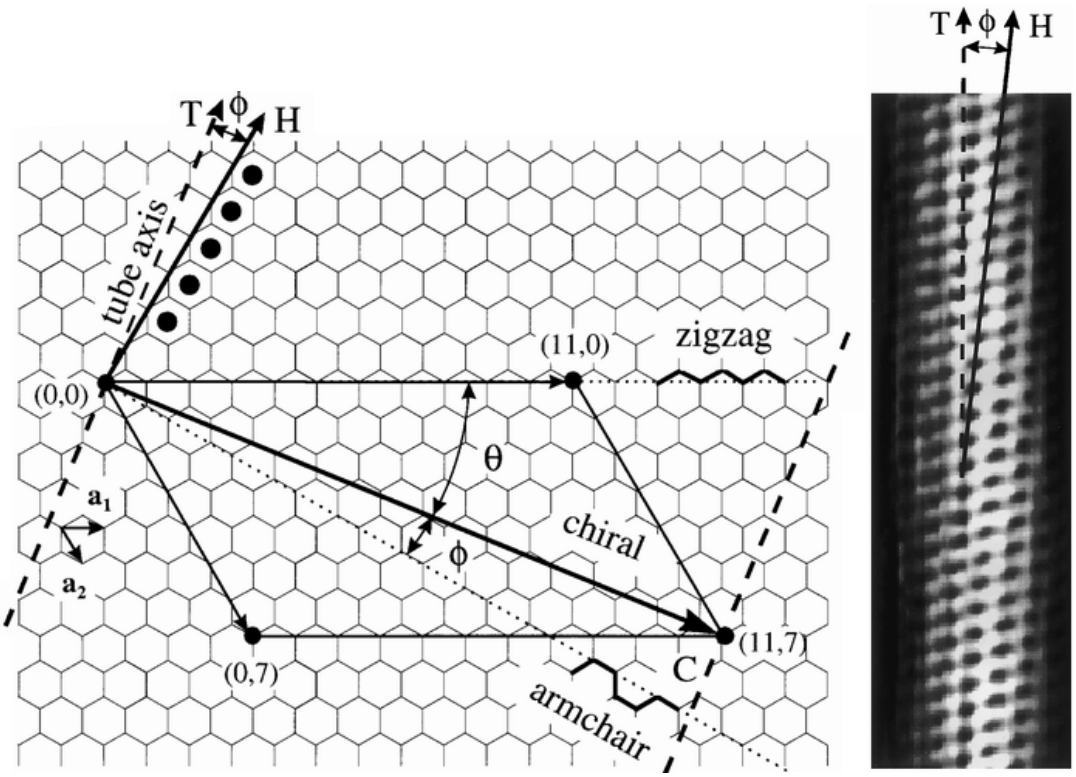


Figure 1. Atomic and electronic structure of a carbon nanotube. (left) Atomic lattice structure of carbon nanotube based on a graphene sheet. The unit vectors of graphene are $\mathbf{a_1}$ and $\mathbf{a_2}$. The chirality is denoted by the wrapping vector **C**, which is orthogonal to the tube axis **T**. The chirality is called "armchair" when the wrapping angle is $\varphi = 0°$. The solid vector **H** is perpendicular to the armchair direction and specifies the direction of nearest-neighbor hexagon rows indicated by the black dots. (right) Scanning tunneling microscopy (STM) image of a carbon nanotube with $(m, n) = (11, 7)$, chiral angle $\varphi = 7°$ and diameter $d = 1.3$ nm. Source: Wildöer *et al*.[3]

single-walled carbon nanotube, $(m, n) = (11, 7)$, the chiral angle is $\varphi = 7°$ and the diameter is $d = 1.3$ nm.

The electronic structure[3,4] of the CNT can be either metallic or semiconducting depending on tube diameter and wrapping angle due to periodic boundary conditions: $\mathbf{C}\cdot\mathbf{k} = 2\pi q$, where **k** is the wavevector, and q is an integer. Only a limited set of allowed

wavevectors can be substituted into the energy dispersion relation for a graphene sheet. q labels the 1-D modes. Slight changes in these parameters determine whether a tube is metallic or semiconducting.

If $2n+m = 3q$ (where q is an integer), then the nanotube is metallic, otherwise the nanotube is semiconducting with an energy bandgap of $E_g = 6a_{CC}/d$, where a_{CC} is the carbon-carbon nearest neighbor distance (142 pm).

All armchair tubes ($n = m$, or $\varphi = 0°$) are metallic. For chiral ($0° < \varphi < 30°$) and zigzag ($\varphi = 30°$) tubes, the following applies. If $n - m = 3l$, where l is an integer, the tube is metallic. Otherwise, it is semiconducting with an energy bandgap of $E_g \sim 0.5$ eV. Given a random distribution of chiralities, 1/3 of nanotubes are metallic, while the remaining 2/3 are semiconducting.

2.2.2 High Mobility

The first carbon nanotube field-effect transistor (CNTFET) was reported in 1998 by Tans *et al*.[5] The basic device consists of source and drain wires contacting a single-walled carbon nanotube laying on a heavily doped silicon substrate capped by a thermal oxide (Figure 2). The substrate was used as a common "back-gate" for all the CNTFETs residing on it. Application of a voltage to the gate electrode can switch the nanotube from a conducting to an insulating state. As shown in Figure 3, the operational characteristics of a CNTFET can be qualitatively described by the semiclassical band-bending models currently used for traditional semiconductor devices.

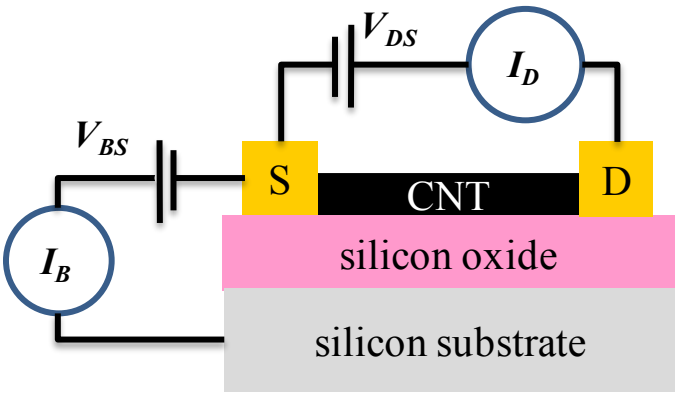


Figure 2. Schematic cross section of a basic CNTFET device. A single- or multi-walled nanotube bridges the gap between the source (S) and drain (D) electrodes. The doped silicon (Si) substrate is used as back gate and is capped with silicon oxide (SiO_2). The drain current I_D is monitored while the drain-source voltage V_{DS} is applied across the channel. The back-gate current I_B is monitored while the back-source voltage V_{BS} is applied to the substrate.

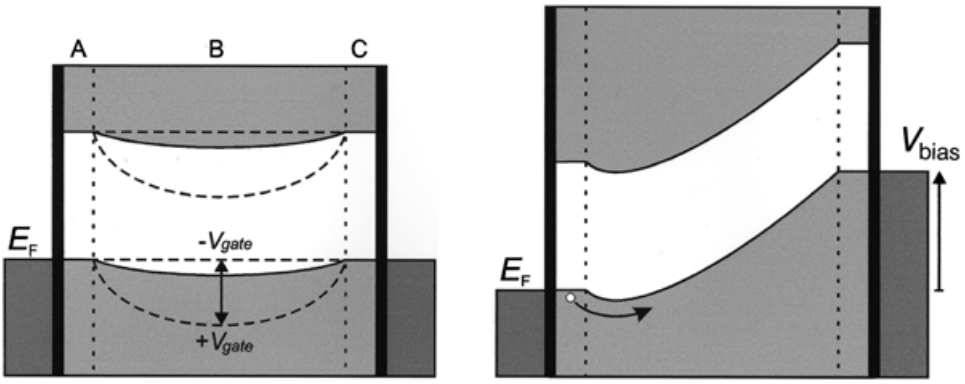


Figure 3. (left) Band diagram of the CNTFET. A CNT with a gap of ~0.6 eV is connected to leads with Fermi energy E_F by tunneling contacts, indicated by black vertical bars. At points A and C, the valence-band edge is pinned to the Fermi energy of the leads. Due to a difference in work function between the CNT and electrodes, the bands bend toward lower energy between the electrodes (B). For positive V_{gate} the bands bend more strongly, leading to an insulating state. For negative V_{gate} the bands flatten, leading to a conducting state. (right) The drain-source bias voltage (V_{bias}) results in a suppression of the barrier. Source: Tans *et al.*[5]

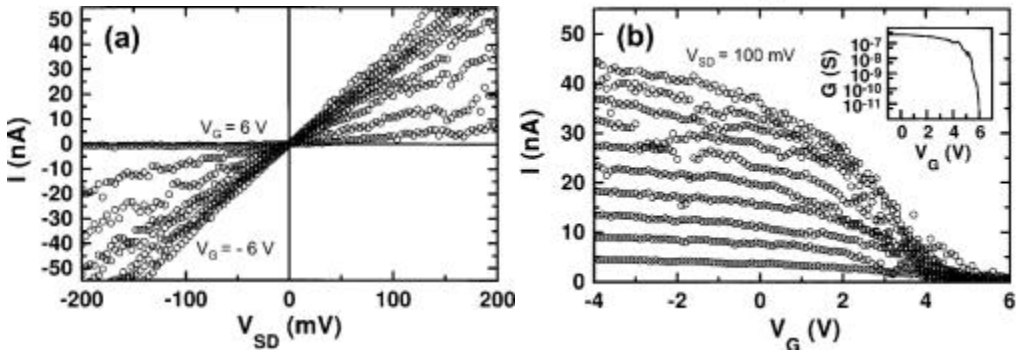


Figure 4. Output and transfer characteristics of a CNTFET. (a) Output characteristic (I_D–V_{SD}) curves measured for V_G = –6, 0, 1, 2, 3, 4, 5, and 6 V, showing ohmic, p-channel behavior. (b) Transfer characteristic (I_D–V_G) curves for V_{SD} = 10–100 mV in steps of 10 mV, showing p-channel operation. The inset shows that the gate modulates the conductance by 5 orders of magnitude. (V_{SD} = 10 mV). Source: Martel *et al*.[6]

Also in 1998, Martel *et al*.[6] analyzed the performance of CNTFETs based on individual single- and multi-wall CNTs. Figure 4 contains the output characteristics and transfer characteristics of a device consisting of one single-walled CNT (diameter 1.6 nm). The output characteristic reveals essentially ohmic behavior, like a resistor whose value is controlled by the gate voltage. The conductance of the device increases with decreasing gate voltage, like a p-channel MOSFET. The p-channel behavior is further affirmed by the transfer characteristic, which shows increases of the drain current as the gate voltage moves from positive to negative. Both the output and transfer characteristics imply that the transport through the CNTFET is dominated by holes.

Dürkop *et al*.[7] studied CNTFETs with channel lengths exceeding 300 µm and found that carrier transport is diffusive and dominated by channel resistance. They demonstrated extremely high mobilities on the order of 10 $m^2/V{\cdot}s$, which is about 200 times that of a typical silicon MOSFET.[8] Such high mobilities even at such long channel lengths are a consequence of the quantum mechanical behavior of charge carriers within

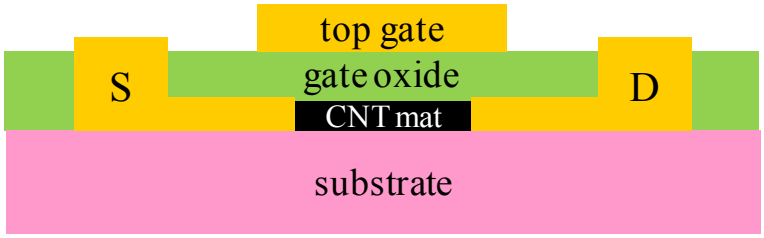


Figure 5. CNTFET with CNT mat covered by a layer of gate oxide for biosensor application.

the one dimensional (1D) physical structure of the nanotube, and suggest the application of carbon nanotubes in chemical/biochemical sensors.

2.2.3 Proposed Use in Biosensors

We propose a CNTFET design based on carbon nanotubes encapsulated in a protective top oxide layer (Figure 5). We choose to use CNTs for their high intrinsic mobility, nanometer-scale geometry, and uncomplicated fabrication requirements. Because a CNT is cylindrical, the spatial confinement focuses the electric field in the vicinity of the CNT and thus creates larger effects for any given quantity of gate charge. Both the high mobility and cylindrical geometry suggest high charge sensitivity. In addition, transistor fabrication using CNTs is far less complicated than in CMOS because it does not require deliberate doping of the channel and the contacts. It is also less restrictive on the choice of starting substrate. Dense, random networks of CNTs ("mats") can be grown "in-house" on silicon, silicon oxide and other metal-oxides, as well organic polymers though nanoimprinting.[9]

The feasibility of using transistors based on CNT mats for detection of unlabelled DNA was successfully demonstrated by Star *et al*.[13] and others,[14] who exploit the fact that single-walled CNTs are more sensitive to DNA charge due to their high surface-to-volume ratio. Since all atoms of the device are surface atoms, the surface to volume ratio

is extremely high so that the effect on a single atom on the surface has far greater influence on the overall transduction of the CNT. Unfortunately, this sensitivity is very non-specific. For example, Kong *et al*.[10] and Collins *et al*.[11] demonstrated that CNTs are extremely sensitive to adsorbed gases such as oxygen, NO_2 and NH_3. Collins *et al*.[12] suggested that surface fluctuations are the origin of excess drift (1/*f* noise) in CNTFETs. Furthermore, designs with exposed CNTs, while extremely sensitive, can suffer from unacceptable levels of false positive readings and very rapid aging in the corrosive environment of biological buffers.

Our answer to these concerns is to completely bury the CNTs and source/drain contacts under a conformal oxide. This prevents their interaction with contaminants from the buffer as well as non-specific interactions with nucleic acids. The channel is thereby extremely robust and shows the ambipolar transfer characteristics of high performance CNT transistors.

Furthermore, our design uses probes that are directly immobilized on the gate rather than substrate backgate used in the design of Maehashi *et al*.[15] One advantage is that the gate voltage used in our case is on order of 1 V as opposed to ~10 V. Such large voltages are well above the oxidation potential of the electrolyte and could easily cause undesirable signals coming from electrochemical processes unrelated to hybridization events. These design points are crucial in the optimal operation of CNT based biosensors.

2.3 Carbon Nanotube Field-Effect Transistor

The earliest CNTFETs were simply gated using the substrate, which was silicon doped to be conductive. While convenient, this method alone provides no way to individually control transistors sharing the same substrate. Top-gates address this need,

and commonly take the form of a dedicated microfabricated electrode atop an oxide layer. As an extreme example, electrolyte can gate a nanotube around the entire circumference and length, without any oxide. We develop a general circuit model for the CNTFET based on the MOSFET, and review the basic design considerations specific to using CNTs.

2.3.1 Transistor Circuit Model

The equations that describe the drain current of a CNTFET are similar to those of a conventional MOSFET but are modified to account for the different geometry of CNTs. Since we will be fabricating devices from dense CNT mats, we replace the concept of channel width W with the number of parallel CNT channels N_{CNT}. We realize that N_{CNT} is not well-defined for a channel comprised of a random CNT network. In this case N_{CNT} represents the *equivalent* or *effective* number of parallel channels. The next task is to replace C_{OX}, the MOS oxide capacitance per unit area, with the gate capacitance per unit length of a CNT, C_G. The solution is to replace $C_{OX}W$ with $N_{CNT}\,C_G$, which maintains proper dimensions. Also, we recognize that CNTFETs are intrinsically ambipolar,[16] but will model them as n-channel devices for simplicity.

The CNTFET operates in three different regimes of operation (deep ohmic, ohmic, and saturation) that are distinguished by the voltages applied to the device. For the saturation ($0 < (V_{GS}-V_{TH}) < V_{DS}$) and ohmic ($0 < V_{DS} < (V_{GS}-V_{TH})$) regions, the drain current for an n-channel CNTFET is:

$$I_D = \begin{cases} \frac{1}{2}\mu\frac{N_{CNT}C_G}{L}(V_{GS}-V_{TH})^2 & \text{for } 0 < (V_{GS}-V_{TH}) < V_{DS} \\ \mu\frac{N_{CNT}C_G}{L}\left[(V_{GS}-V_{TH})V_{DS}-\frac{1}{2}V_{DS}{}^2\right] & \text{for } 0 < V_{DS} < (V_{GS}-V_{TH}) \end{cases} \quad (1)$$

where μ is the mobility, C_G is the gate capacitance per unit length of the CNT, L is the length of the transistor, V_{GS} is the voltage drop from gate to source, V_{TH} is the threshold voltage, and V_{DS} is the voltage drop from drain to source.

In our back-gating experiments, the drain voltage V_{DS} is biased to only 0.1 V, whereas magnitude of the gate voltage is usually anywhere from 10 V to 40 V. In this "deep ohmic" region ($0 < V_{DS} << (V_{GS}-V_{TH})$), the drain current is approximated as:

$$I_D = \mu\frac{N_{CNT}C_G}{L}(V_{GS}-V_{TH})V_{DS} \quad \text{for } 0 < V_{DS} << (V_{GS}-V_{TH}) \quad (2)$$

This linear relationship between the drain current and voltage justifies describing the channel under low bias as ohmic. In the output characteristic, the deep ohmic behavior is near the origin.

The output and transfer characteristics of an idealized CNTFET are shown in Figure 6. These curves have the exact same shape as those of the ideal nMOSFET, with the exception that the ambipolar nature of the CNTFET is depicted by mirroring the transfer characteristic across the line $(V_{GS} - V_{TH}) = 0$. In a similar vein, the output characteristic for any given $V_{GS} = V_{GS,k}$ is the same for $V_{GS} = -V_{GS,k}$ (assuming $V_{TH} = 0$).

The transconductance ($g_m \equiv \partial I_D/\partial V_{GS}$) of the CNTFET can be calculated for any of the three bias conditions but is easily recognizable as the slope of the transfer

characteristic. In practice, we calculate g_m by fitting a line to the transfer characteristic in the deep ohmic region (which occurs at large $|V_{GS}|$), where the slope is given by

$$g_m = \frac{\mu N_{CNT} C_G V_{DS}}{L} \tag{3}$$

This expression is used to calculate the field-effect mobility μ_{FE} from measured transfer characteristics:

$$\mu_{FE} = \frac{g_m L}{N_{CNT} C_G V_{DS}} \tag{4}$$

The mobility can also be calculated from the output characteristic as:

$$\mu_{FE} = \frac{g_{ds} L}{N_{CNT} C_G \left(V_{GS} - V_{TH}\right)} \tag{5}$$

where $g_{ds} \equiv \partial I_D / \partial V_{DS}$, which is the small signal output conductance.

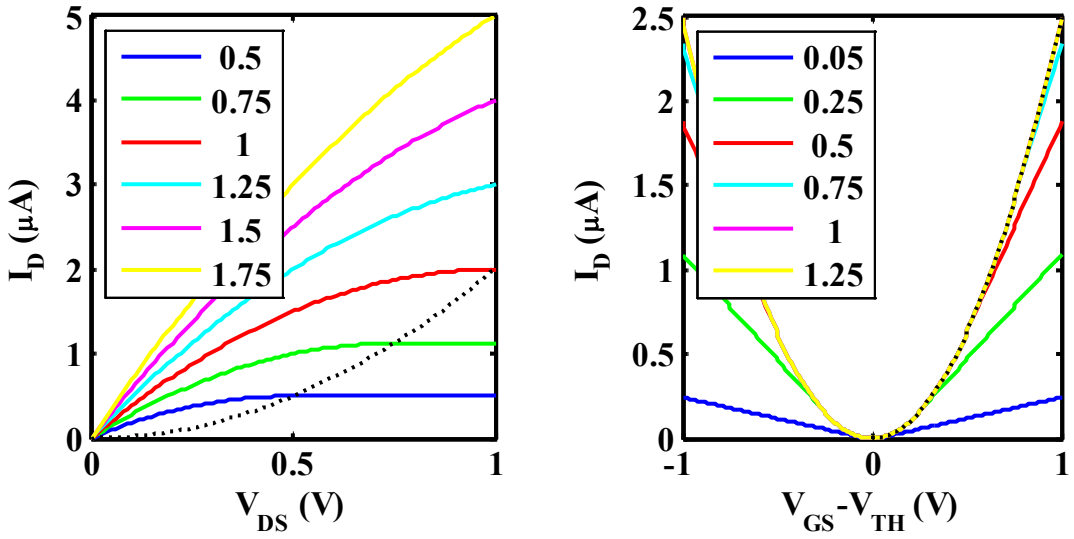


Figure 6. Output and transfer characteristics of a model n-channel CNTFET (N_{CNT} = 1, μ = 10 m^2/V/s, C_G = 250 pF/m, L = 500 μm). (left) Output characteristic (I_D–V_{DS}) curves for V_{GS}–V_{TH} = 0.5–1.75 V in steps of 0.25 V. The dotted curve indicates the boundary between the ohmic region $V_{DS} < (V_{GS}–V_{TH})$ and the saturation region $V_{DS} > (V_{GS}–V_{TH})$. (right) Transfer characteristic (I_D–($V_{GS}–V_{TH}$)) curves for V_{DS} = 0.05–1.25 V in steps of 0.25 V. The dotted curve indicates the extreme of the saturation condition ($V_{GS}–V_{TH}$) < V_{DS}. The ohmic region is the straight part of each line branching to the right.

2.3.2 Transistor Design

The design of the CNTFET-based biomolecular sensor will be similar to that of the MOSFET-based ion-sensitive FET (ISFET) used in pH sensors and the like.[17] The engineer must choose transistor width and length, as well as oxide material and thickness within the design requirements of (i) sensitivity, (ii) selectivity and reduction of false positives, (iii) device manufacturability, (iv) compatibility with conventional molecular biology protocols, (v) signal response calibration and (vi) modeling. In the case of the CNTFET, the unit element of width is the tube. The diameters of the tubes are not well-

controlled, but are on the order of 1 nm for single-walled nanotubes.[18] The overall width of the sensor can be increased by connecting tubes in parallel. Since the CNT is of a 1D nature, the length of the sensor has little impact on conductance for lengths greater than a few microns. In fact, defects and Schottky barrier modulation at the contacts[19] may play a greater role in sensitivity. Assuming the contacts have already been engineered, the oxide thickness is the remaining major design parameter. The choice of oxide thickness will be a compromise between a desire of thick oxide to minimize electrolyte current leakage and the desire of thin oxide to maximize the gate coupling capacitance. The following derivation will show that the cylindrical nature of the CNT reduces the requirement for extremely thin oxides.

A toy model for the nanotube-oxide-gate structure is the standard electrostatic textbook situation of a conducting tube (cylinder) in embedded in a dielectric in proximity to a conducting plane. We will assume that all three objects are infinite in one dimension so that the problem can be solved in only two dimensions. The problem reduces to strategically positioning a line of charge a certain distance from the plane that satisfies the boundary conditions of the tube and plane. To aid in problem solving, we use the method of images; a fictitious image charge of equal and opposite magnitude to the first is placed within the plane. The original line charge and its image are an equal distance from the surface of the plane, which is the line of symmetry.

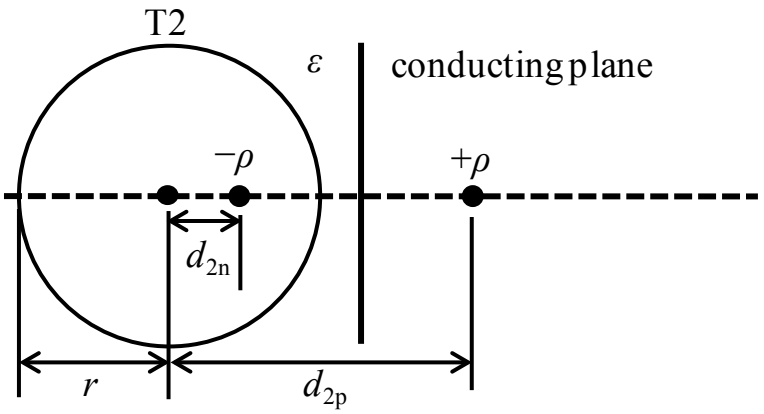


Figure 7. Cross-section of conducting tube T2 embedded in dielectric (ε) near a conducting plane. The negative and positive line charges, $-\rho$ and $+\rho$, are of equal magnitude and placed equidistant from the planar surface (vertical line). The negative charge $-\rho$ is a distance d_{2n} to the right of the tube axis. The positive image charge $+\rho$ is located a distance d_{2p} to the right of the tube axis. Axes of the tubes and line charges all lie on the horizontal axis shown in dotted line.

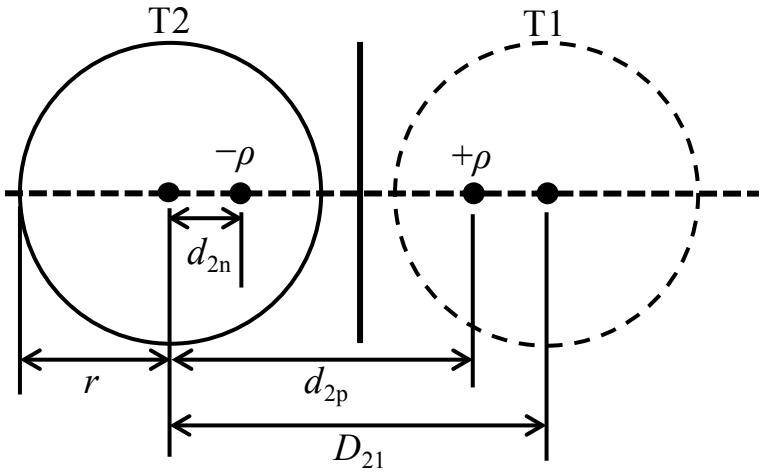


Figure 8. Cross-section of electrostatic model updated to include an equipotential image tube T1 with its axis located distance D_{21} to the right of the axis of tube T2. The same setup models a two-wire transmission line.

First, we must find the location of the equipotential cylinder around one of the two line charges. As shown in Figure 7, a negative line charge $-\rho$ is located a distance d_{2n} to the right of the axis of the equipotential tube T2 of radius r, such that $d_{2n} < r$. Then a positive image charge $+\rho$ is placed at a distance d_{2p} to the right of the tube axis. From geometrical arguments,[20] it turns out that $d_{2n} = r^2/d_{2p}$ makes the tube surface (2) equipotential.

Next, we bring in a fictitious image tube T1 and place its axis at d_{2n} to the right of the image charge $+\rho$, as shown in Figure 8. By symmetry, both tubes T1 and T2 are equipotentials, and so is the surface of the conducting plane. The distance between the

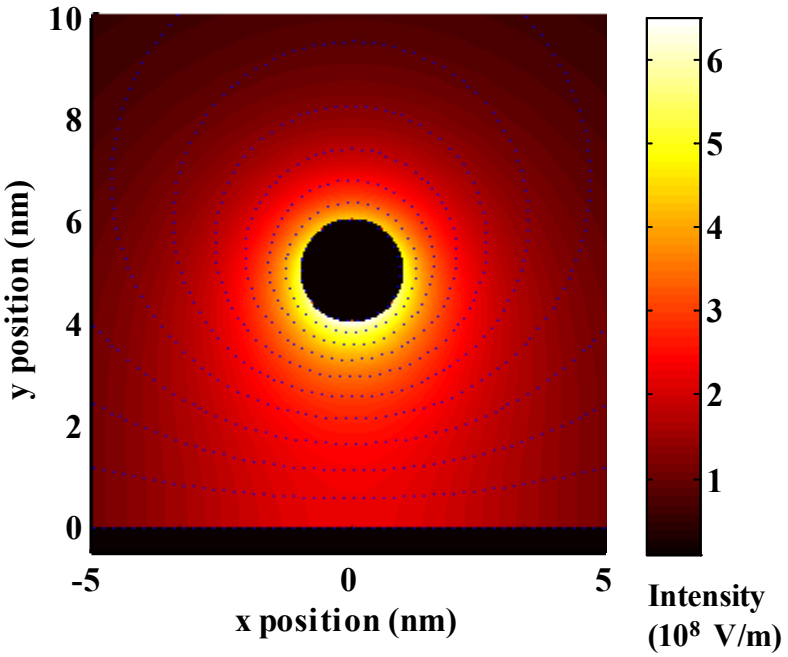


Figure 9. Calculated electric field intensity and potential contour lines (dotted) of a 2 nm diameter tube (r = 1 nm) at 1 V potential, embedded in dielectric and located 5 nm from a conducting surface ($y < 0$).

two tube axes is $D_{21} = d_{2p} + d_{2n} = d_{2p} + r^2/d_{2p}$. Given the radius of the tube and the distance of the tube from the conducting plane, one solves a quadratic equation to determine the positions of the tubes and corresponding line charges.

The electric field intensity and electric potential in the dielectric is calculated by the superposition of the fields and potentials from the two line charges. We plot the field and potential modeling a 2 nm diameter CNT embedded in dielectric at a distance of 5 nm (D_{21} = 10 nm) from a conducting plane in Figure 9.

To calculate the capacitance of the tube to the conducting surface, we need to find the voltage of tube T1 with respect to the conducting plane. Let x denote horizontal position, and let the vertical line of symmetry be the plane where $x = 0$. Taking the potential at any point on $x = 0$ to be zero, the voltage of tube T1 is given by

$$V_1 = \frac{\rho}{2\pi\varepsilon} \ln r/d_{2p}, \tag{6}$$

where ε is the dielectric permittivity. By symmetry, the voltage of tube T2 is $V_2 = -V_1$. The capacitance per unit length between the tube and the conducting plane is given by the charge ρ divided by the voltage V_1, or

$$C = \frac{2\pi\varepsilon}{\operatorname{arcosh}(h/r)}, \tag{7}$$

where $h = D_{21}/2$ is the height of the tube axis from the conducting plane. Using the identity $\operatorname{arcosh}(x) = \ln(x+(x^2-1)^{1/2})$, we see that for $h/r >> 1$, we can approximate the denominator as $\operatorname{arcosh}(h/r) \approx \ln(2h/r)$. Since $\ln(x)$ is a slow function when $x >> 1$, it follows that the capacitance varies little when $h/r > \sim 6$.

Figure 10 plots the gate capacitance as a function of height for a 2 nm diameter tube (r = 1 nm) embedded in aluminum oxide (Al_2O_3). This model is applicable to the

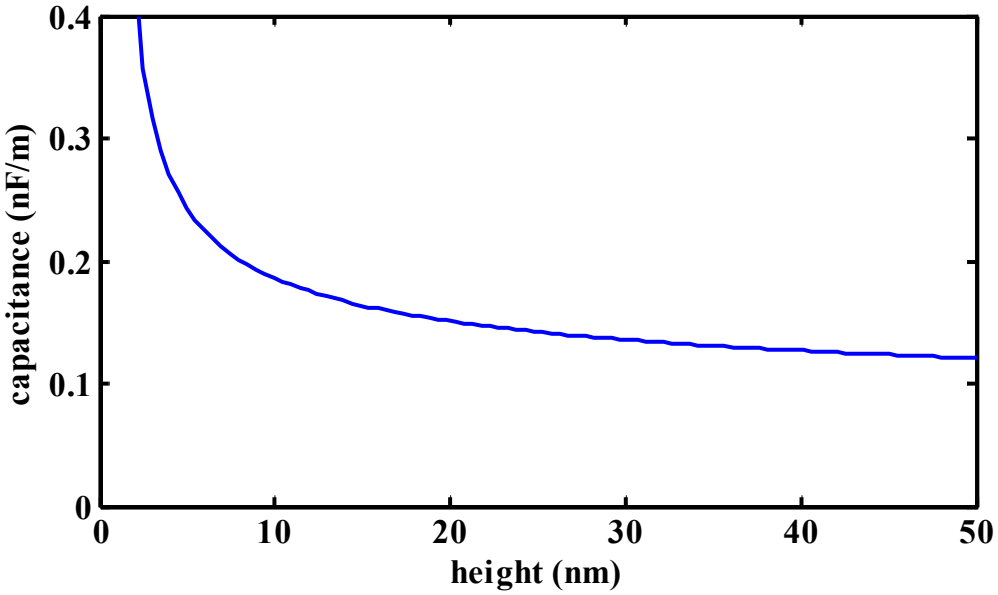


Figure 10. Capacitance per unit length of a 2 nm diameter tube (r = 1 nm) embedded in aluminum oxide ($\varepsilon = 10\varepsilon_0$) as a function of height h from conducting plane. The capacitance varies little at sufficient distances ($h > 6r$).

top-gated CNTFET of Figure 5. We see that the capacitance is indeed a slow function of height when $h > 6$ nm.

Note that we cannot blindly apply (7) to actual CNTFETs because the dielectric above the nanotube is often different than the dielectric below it. In a back-gated bare CNTFET (Figure 2), the CNT is resting atop silicon oxide but the dielectric medium above it is air. Therefore (7) overestimates the electrostatic gate capacitance. For example, the capacitance of a nanotube of diameter 3.9 nm (radius r = 1.95 nm) embedded in SiO_2 ($\varepsilon_{ox} = 3.7\ \varepsilon_0$) at a height h = 500 nm above a conducting plane is C_G = 33 pF/m. All else equal, an electrostatic simulation[7] for a CNT sitting atop a 500 nm thick SiO_2 yields a lower back-gate capacitance of C_G = 19 pF/m.

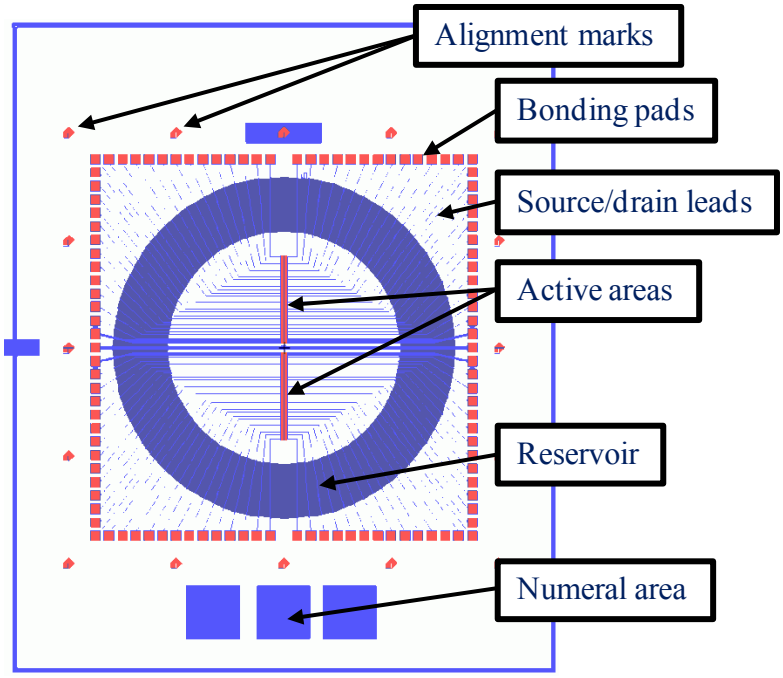


Figure 11. Overlay of all masks for chip fabrication, plus cross section of tubular reservoir. Array of 104 CNTFETs is vertically aligned along red middle strip. Their drain leads connect to bonding pads, the red squares around perimeter. Design is separated into top and bottom halves, where each half contains 50 accessible transistors and one T-shaped common source electrode. An on-chip gate electrode running horizontally across the middle is not currently used. Three boxes on bottom of chip represent three-digit chip ID. Chip size is 9 x 12 mm.

For use as a charge sensor, the top-gate capacitance (from CNT to electrolyte) should be much higher than the back-gate capacitance (from CNT to silicon substrate). According to Figure 10, the top-gate C_G is about 100 pF/m when the oxide thickness h = 50 nm, which is a practical oxide thickness. Thus, we can expect the top-gate to be only five times as effective as the back-gate. We might be tempted to reduce the top-gate oxide thickness to achieve higher capacitance. Unfortunately, Figure 10 shows that major gains in C_G would only be achieved when oxide is less than 6 nm thick. From experience, we found that even 25 nm oxide yielded many failures due to pinholes and defects. These

problems are amplified because the oxide does not grow directly on CNTs. Ultimately, the judicious choice of a high-κ dielectric in combination with an appropriate top-gate oxide thickness provides the optimal compromise between top-gate coupling and adequate protection.

2.3.3 CNTFET Array Layout

The fabrication of the first CNTFETs involved growing randomly oriented CNTs, locating suitable CNTs via scanning electron microscopy, and using the SEM to expose a pattern on a resist designed for electron beam lithography. Using this time-consuming process, Dürkop *et al.*[7] made CNTFETs with single CNTs spanning channel lengths exceeding 300 μm. The fact that transistors of these lengths were functional paved the way for the massively parallel fabrication of similar devices using standard photolithography.

Rather than using single-CNT channels, we developed arrays of transistors based on high-density CNT networks ("mats") between the source and drain. The mask layout for fabricating an array of CNTFETs using standard photolithography is shown in Figure 11. A spider-like formation of numerous leads connects a central array of transistors to bonding pads around the square perimeter. The layout contains two sub-arrays containing 50 individually accessible transistors which share a T-shaped common source electrode. The mask will be shown in closer detail in the fabrication section.

As shown in Figure 12, the design uses a plastic tube to contain the liquid bathing a large number of transistors. Adding more transistors allows for increased data throughput and the possibility of multiple molecular probes per chip. The liquid containment tube provides a stable liquid environment over all the transistors during

chemical exposure protocols, as well as during the top-gating experiments. In this design, the bonding pads were placed on a perimeter large enough to accommodate attachment of a plastic tube up to 5 mm outer diameter, which we refer to as the "reservoir". Larger liquid volume means evaporation does not shorten the duration of experiment.

2.4 **Biosensor Fabrication Process**

The biosensor fabrication process starts with a clean silicon wafer and produces a finished product comprised of a functional CNTFET array. Fingertip-sized chips are diced from a wafer and dusted with a fine layer of pure iron (Fe). The iron serves as the catalyst for the subsequent growth of carbon nanotubes using high temperature chemical vapor deposition (CVD). With CNTs covering nearly the entire chip, the first step in the microfabrication process removes large patches of CNTs to reduce the coverage to only the active transistor areas. Next, contacts are made from the nanotubes to the perimeter bonding pads, oxide is grown over the nanotubes, and the bonding pads are exposed and

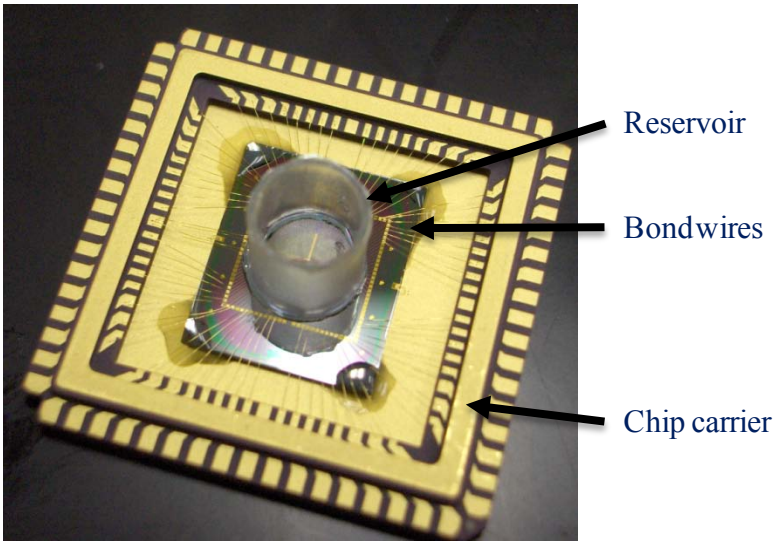


Figure 12. Finished biosensor chip mounted onto leadless chip carrier (LCC), wire bonded, and with reservoir attached.

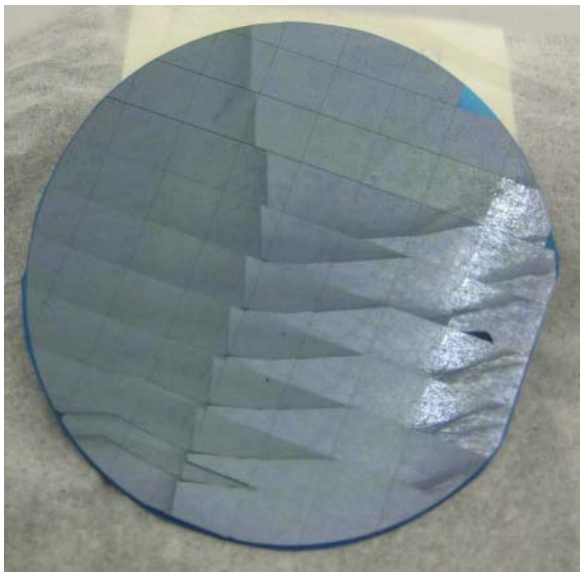


Figure 13. Photograph of diced 100 mm silicon wafer. The rectangular chips (10 mm by 12 mm) still have a protective polymer coating and are attached to the dicing tape. Each wafer yields about 50 useable chips.

built up. The final step is the fabrication of the metal gates, after which the biosensor is ready for attachment to a chip carrier, wirebonding, and the attachment of the tubular reservoir. A detailed discussion of each step is presented below.

2.4.1 Wafer Dicing

The starting material for fabrication is a 4" (100 mm) wafer of silicon capped with 500 nm of thermal oxide. The substrate (Silicon Quest International) is 525 μm thick and highly doped p-type with boron (B) to a resistivity of 1 to 10 Ω·cm. The wafer is mounted onto an adhesive tape to support the numerous individual chips after dicing.

Figure 13 shows a wafer diced into about fifty 10 mm x 12 mm chips on a Disco DAD 321 Automatic Dicing Saw at the Laboratory for Physical Sciences (LPS). This particular size was required due to the relatively small diameter of the quartz tube furnace used to grow the nanotubes. Each chip is then individually cleaned with streams of

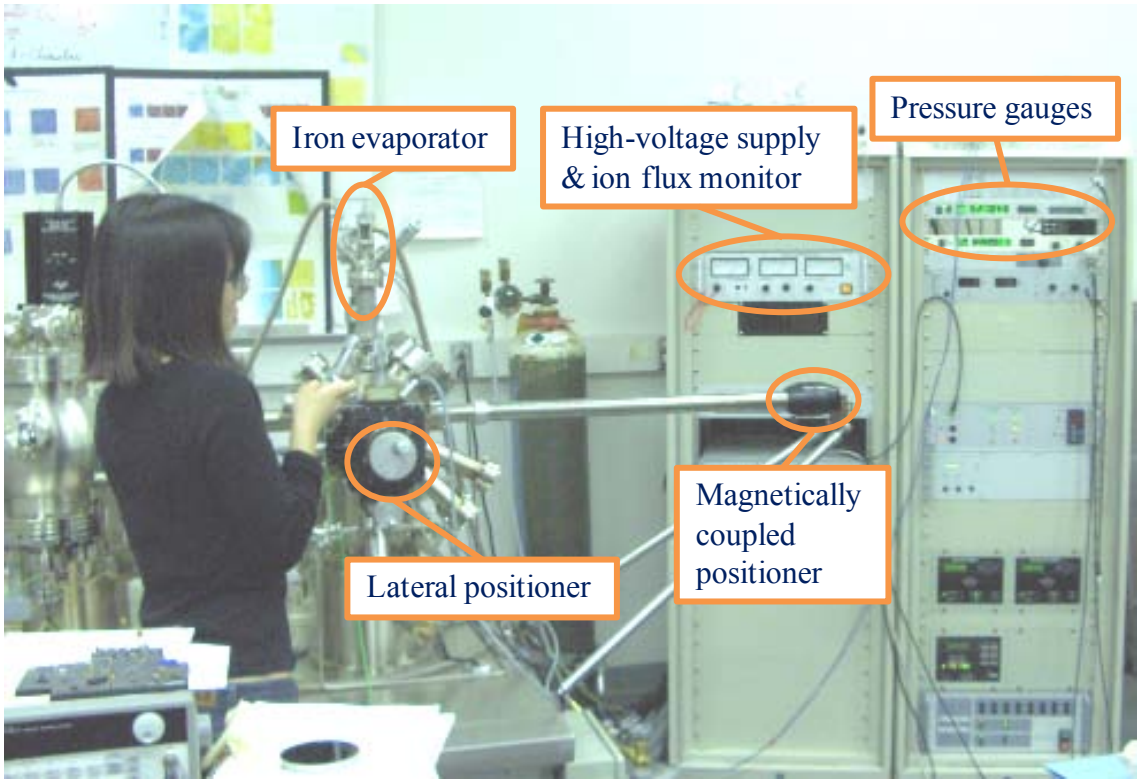


Figure 14. Iron deposition using the ultra-high vacuum electron beam evaporator in the Nanomagnetics Laboratory of the LPS.

acetone, methanol, isopropanol, and deionized water (DI). The final cleaning is done by oxygen plasma ashing to remove organic residues. Then identification numbers are marked on the chips using a diamond tip scribe.

2.4.2 UHV Iron Deposition

High-purity iron (Fe) nanoparticles for the catalyzed nucleation of carbon nanotubes were deposited onto each chip by an electron beam evaporator in an ultra-high-vacuum (UHV) system at the LPS. The deposition of one monolayer of iron is typically performed at an absolute pressure of around 500 nPa and is complete within about 1 minute. At this pressure, the formation of one monolayer of background

contaminants is more than 10 minutes, allowing for the deposition of a high purity iron film.

The UHV system, shown in Figure 14, is isolated from the ambient via a fast-entry load lock which only allows a handful of chips to be processed per cycle. Deposition was performed by personnel trained in operation of the complex vacuum system and specialized manipulators required for UHV. The iron evaporator itself requires a skilled technician to perform proper start-up and operation procedures such as manual adjustment of the high-voltage supply to maintain the proper ion flux density.

2.4.3 CNTCVD

The CNTs are grown via a high-temperature chemical vapor deposition (CVD)

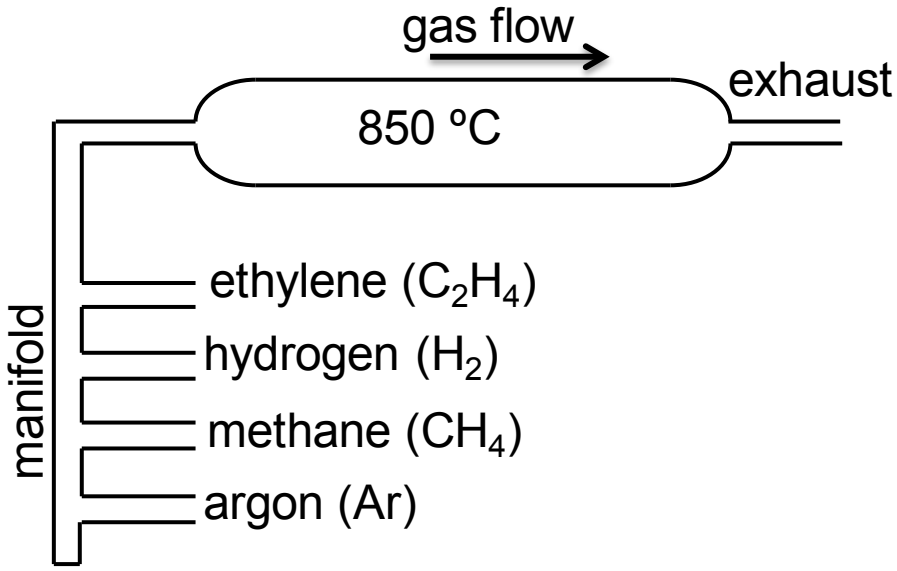


Figure 15. Schematic of carbon nanotube chemical vapor deposition (CNTCVD). The inlet gases include argon for purging, hydrogen for reduction of iron, and methane and ethylene to provide carbon for nanotube growth. The temperature of CNT growth is 850 °C. The cooled exhaust gases are vented to a fume hood.

process in the Fuhrer lab in the Physics Department. This involves loading the iron-coated chips into a tube furnace and growing the nanotubes at 850 °C for 10 min using a mixture of H_2, CH_4, and C_2H_4 using the flow rates and following the method described by Kim *et al*.[21] As shown in Figure 15, argon (Ar) is manifolded with the three CNT growth gases for the purpose of purging the system to avoid hazardous buildup of flammable gases. The exhaust is vented to a fume hood. The process steps are as shown in Table 1.

The following sequence outlines the events happening during the birth of nanotubes. During the heating phase, the sub-monolayer iron film reorders into nanoparticles given the significant kinetic energy of atoms at 845 °C. During the growth phase, the hydrogen gas first reduces the iron islands to elemental iron. Then the metallic particles absorb carbon from the flow of hydrocarbon gases until the point of saturation. Once a catalyst particle reaches saturation, the carbon nanotube rapidly grows out from it

Table 1. CNTCVD process steps used at UM Physics.

1. **Flush:** Open (CCW) argon (Ar) rotameter to 20 mm (729 sccm), at room temperature (RT), for 5 mins.
2. **Purge lines**: Close Ar Swagelok (CW). Open rotameters (CCW) of hydrogen (H_2) 93 mm (1900 sccm), methane (CH_4) 23 mm (1240 sccm), ethylene (C_2H_4) 5 mm (18 sccm), RT. Purge for 5 mins.
3. **Heat up:** Set Ar flow rate to 20 mm and ramp from RT to 850 °C (about 45 mins). To reduce overshoot, take smaller steps: 650, 750, 825, 845 °C.
4. **Grow CNTs**: Close Ar Swagelok, open H_2, CH_4, C_2H_4 Swagelok valves and confirm H_2 93 mm, CH_4 23 mm, C_2H_4 5 mm. Hold at 845 °C for 10 mins. Purpose of hydrogen gas is to reduce iron oxide (Fe_2O_3) islands to the elemental state (Fe).
5. **Cool:** Turn off H_2, CH_4, C_2H_4 Swagelok valves (CW). Immediately turn on Ar 20 mm. Set temperature to 30 °C. This usually takes about 200 mins but can be accelerated by partially opening the heater module.
6. **Vent**: T<200 °C, turn off the main cylinder valves (CW) of H_2, CH_4, C_2H_4, then Ar. Open the Swageloks (CCW) and rotameters (CCW) to vent out the regulators and supply lines. After venting, decrease the regulator controls (fully CCW) and close the outlet valves (CW). Close the Swageloks (CW) and gently close the rotameters (CW).

until stopped by collision or other environmental disturbance. The carbon nanotubes generated by this process have been shown to be mostly single-walled and about two-thirds are semiconducting.[22] The density of catalyst particles determines the density of the CNT mat.

Successful growth of CNTs is associated with direct visual observation of a dark circle corresponding to the UHV iron deposition spot, as shown in Figure 16a. Since this preliminary observation could also be due to a layer of amorphous carbon soot or buckyballs, we must use more advanced imaging methods to confirm the success of CNTCVD.

Atomic force microscopy (AFM) is a variant of scanning probe microscopy where a sharp tip is rastered over a surface and atomic force interactions are measured or used

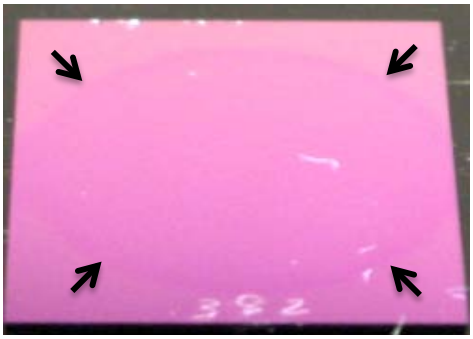
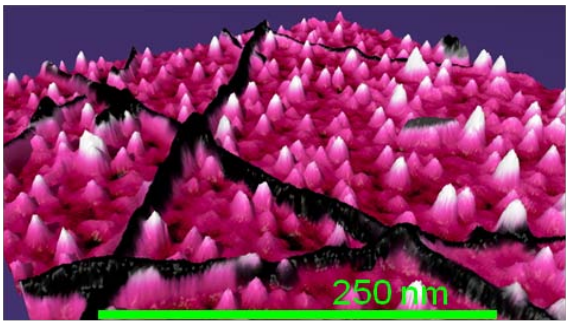
(a)  (b) 

Figure 16. (a) Photograph of SiO_2/Si chip #382 after CNTCVD taken at proper viewing angle for contrast of nanotube film against substrate. The arrows indicate the edges of the dense CNT network, corresponding to the boundaries of iron film. (b) Atomic force microscopy (AFM) 3D rendering of carbon nanotubes (black) on silicon oxide substrate (purple). The white peaks indicate iron nanoparticles or possibly undeveloped carbon buckyballs. One can observe overlap of nanotubes in foreground and midground. It is believed that the growth of individual nanotubes is very rapid once nucleated on an iron nanoparticle, and growth stops when the sprouting end of the tube hits an obstacle.

as feedback signal.[23] This method was used to image the height topography of the nanotubes as shown in Figure 16b, which suggests the CNTs are forming a random network. If one traces the path of one tube, it is apparent that it may overlap or be overlapped by other tubes, implying the sequential growth of CNTs. Furthermore, the junctions between the tubes may have consequences on the mobility and transistor behavior of the overall network.

Scanning electron microscopy (SEM) is a quick, non-contact non-destructive method that we routinely used to confirm the growth of dense networks of CNTs. Figure 17 shows how CNTs grown from our sub-monolayer iron catalyst form a very high-density randomly ordered network suitable for the large-scale fabrication of transistors. For comparison, the CNTs grown from ferric nitrate ($FeNO_3$) catalyst solution[24] are very sparsely distributed and too short to enable large-scale fabrication.

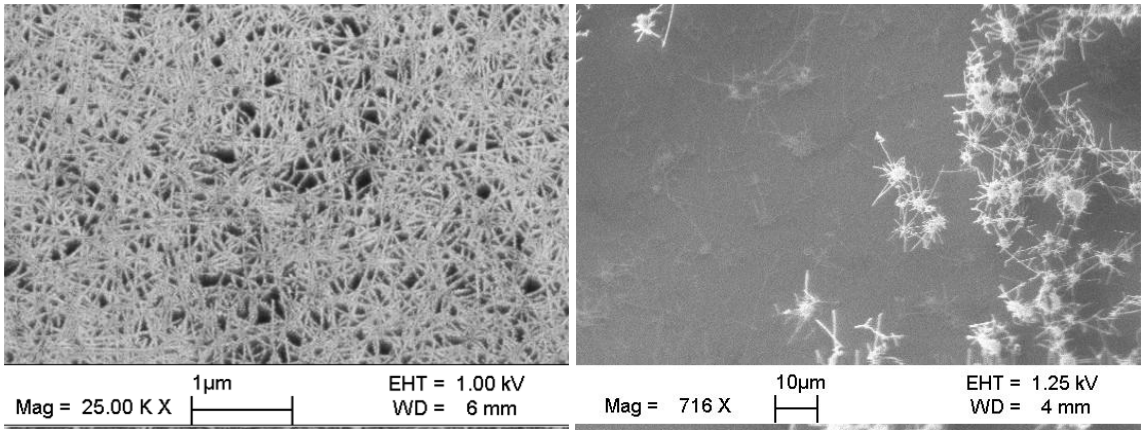


Figure 17. Scanning electron micrographs (SEM) of CNTs (light) on oxide substrates (dark). The contrast of the CNTs against the substrate, as explained by Britlinger *et al*.[25], is due mainly to differing electron charging and discharging effects. (left) High-density randomly ordered CNT network ("mat") grown from UHV iron and suitable for mass fabrication of transistors. (right) Low-density randomly nucleated islands of CNTs grown from $FeNO_3$ catalyst solution.

2.4.4 Microelectronic Fabrication Process

With carbon nanotubes already covering the silicon oxide substrate, the next steps are to build the transistor source and drain contacts using standard cleanroom processing techniques. While the microelectronic fabrication process is similar to that of a silicon MOSFET, there are some unique considerations when working with CNTs. Since all chemical bonds of the graphitic carbon atoms are filled, there are no dangling chemical bonds on the CNT as silicon has. A positive consequence is that CNTFETs may incorporate a broader choice of oxide material than silicon devices, which are plagued by detrimental oxide interfacial states. This same reason is responsible for why CNTs do not adhere well to a many materials.

Since the bonding pads will be ultrasonically wire-bonded to a chip carrier, we need excellent adhesion of the bonding pad films to the substrate. Furthermore, we need the source and drain metal leads to maintain adhesion to the substrate under varying temperatures and moisture levels. To increase adhesion of both the bonding pads and metal leads, we re-designed the process to clear out CNTs from the area underneath these areas as much as possible.

We also experienced an adhesion issue when we started to use gold to decrease the resistance of the metal leads. This resistance was responsible for increasing the apparent threshold voltage of transistors due to significant voltage drop from the transistor source to the source bonding pad. The use of gold reduced this effect, but caused decreased adhesion with the ALD aluminum oxide. This is the reason for using a triple layer (Ti/Au/Ti) for the metal wires.

The biosensors were almost entirely fabricated at facilities within the University of Maryland, College Park. Most process steps were performed in the cleanroom of the Laboratory for Physical Sciences (LPS). The equipment used included the Zuss MJB-3 Mask Aligner, Plasmatherm 790 reactive-ion etcher (RIE), and CHA industries electron beam evaporator. Other process steps were performed in the Department of Physics and the FabLab Class 1000 research cleanroom. The details of the microfabrication steps will be discussed in the following.

2.4.4.1 Alignment Marks

Figure 18 shows the mask for alignments used in subsequent processing steps. Upon their entrance into the cleanroom environment, the chips are rinsed with acetone, methanol (MeOH), and isopropanol (IPA), and blown dry in nitrogen (N_2). Ultrasonic cleaning, piranha etch, and other harsh cleaning methods were not used because it would remove the CNTs.

Standard photolithography was used to transfer patterns in a contact mask to a photoresist. After the chips were spin-coated with Futurrex NR9-1500PY negative photoresist, a pre-exposure bake was done on a 150 °C hotplate for 60 s. Then the pattern in the mask was transferred by exposure for 16 s with a 350 nm (i-line) wavelength contact aligner. The post-exposure bake is 100 °C on a hotplate for 60 s. Pattern development was done by immersion in RD6 for 15 s. The masks shown are clear field (dark feature), so the negative resist will developed away only in the dark areas (indicated in red).

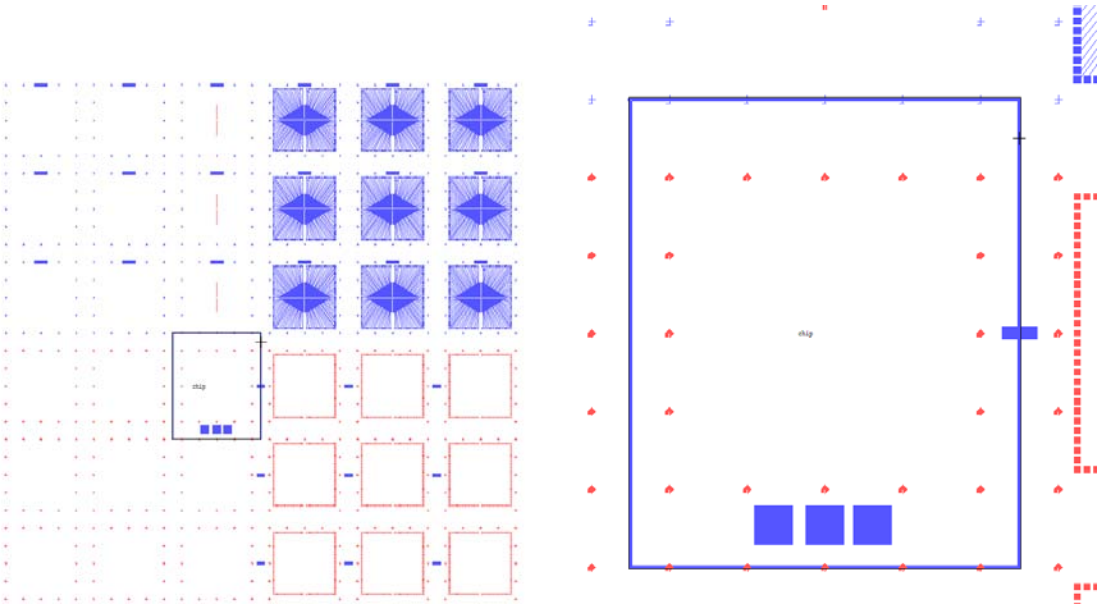


Figure 18. (left) Single contact mask containing patterns for entire fabrication process: (clockwise quadrants starting from upper right) metal wires, bonding pad etch/deposition, alignment marks, and field ash. The rectangular outline containing three small boxes representing etched numbers indicates the position of a chip ready for alignment marks. (right) Zoomed view of mask for alignment marks (red) and outline indicating proper chip position (blue). The alignment marks have 2 mm spacing.

To ensure good adhesion of the metal film, the chip is exposed to oxygen plasma for 20 s to remove CNTs. Then 5 nm of titanium (Ti) was evaporated, followed by lift-off in acetone, allowing for direct adhesion of the Ti film on the area colored with red, while CNT mats still remain elsewhere. This first metal layer serves as the alignment reference for subsequent pattern transfers.

2.4.4.2 Unwanted CNT Removal

The CNTs in the inactive areas (field) are removed prior to the deposition of any metal wires or bonding pads. The CNTs remaining in the active area are in the channel gap (under the gate) and a 10 µm overlap with the source and drain leads. Figure 19

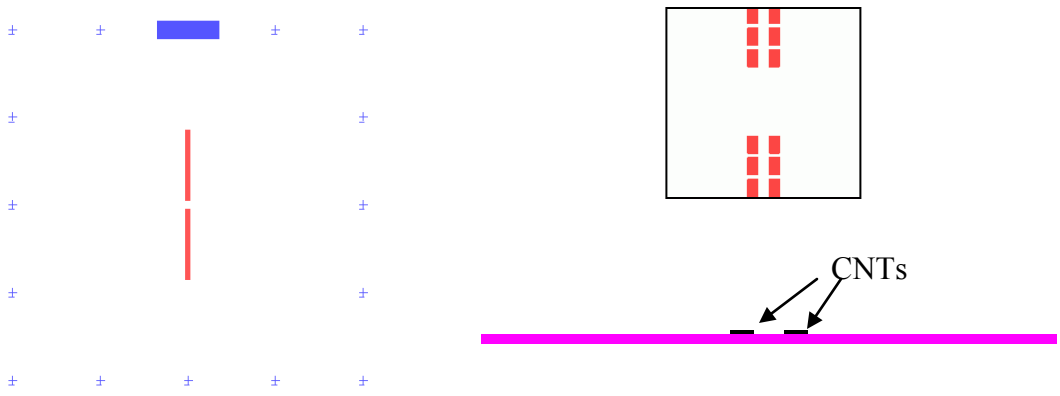


Figure 19. (left) Mask for ashing of CNTs in the field. (right top) Zoomed view of mask showing twelve CNTFET active areas. (right bottom) Cartoon illustration of cross-section of remaining CNTs on SiO_2 substrate to be used for two CNTFETs. Alignment marks and bulk substrate are not shown.

shows the mask pattern used for this step. The goal is to remove all CNTs except for where the mask is opaque.

To begin this step, we spin OIR 908-35, a positive photoresist and bake it at 90 °C on a hotplate for 60 s. Then the pattern in the mask was transferred by exposure for 13 s on the contact aligner. There is no post-exposure bake for this step. Pattern development was done by immersion in OPD4262 for 60 s. The positive resist will be developed away only in the open areas of the mask. As before, CNT removal is accomplished by exposing the chip to oxygen plasma. CNTs remain only in the areas protected by photoresist. The end result is a chip with alignment marks and CNTs in the active areas.

2.4.4.3 Metal Contacts & Leads

After clearing out unwanted CNTs, we are ready to deposit the metal wires. These include the source and drain contacts to the CNTs, as well as wires that lead from the CNTFETs proper to the bonding pads. NR9-1500PY coating, pre-bake, exposure, post-bake, and development, as described before, are performed to transfer the patterns in the

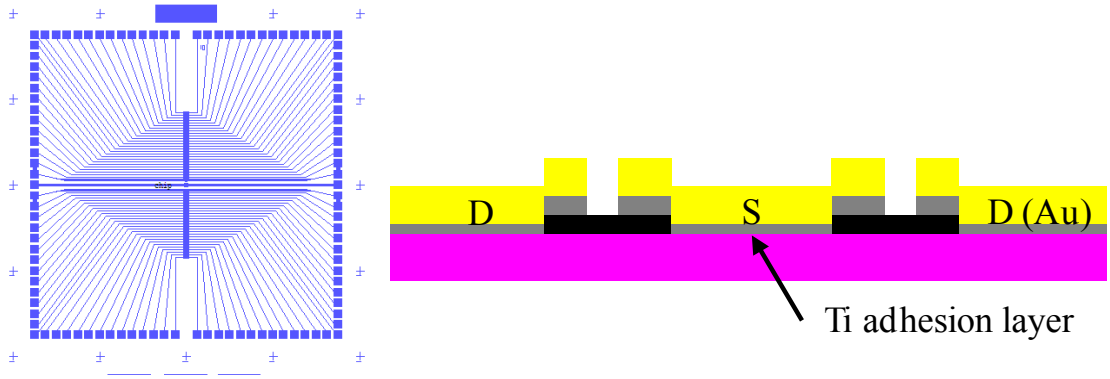


Figure 20. (left) Mask for the source, drain, and corresponding leads of the metal layer. (right) Zoomed-in cross-section of two transistors after metal (Au with Ti adhesion layer) for drain electrodes (left and right) and common source electrode (middle) are deposited onto substrate and CNTs. At this stage, the CNTFETs are occasionally tested in back-gate measurements.

mask shown in Figure 20. Then the appropriate metal is evaporated and lift-off is performed in acetone. At this stage, the transistor is a bare CNTFET; back-gate testing is described later in Section 2.5.1.1.

Based on the literature on annealing CNTFETs,[16] the metal layer of our early chips was a single layer of titanium (Ti) only 50 nm thick. Given that titanium is 19 times less conductive than gold, and the wires from the bonding pads are long and skinny, we suffered a series resistance on the order of 10's of kΩ. Upon realizing that we did not need to pursue annealing but needed to decrease the source and drain wire resistance, we changed the recipe to 100 nm of gold (Au) on 10 nm of titanium (Ti). This resulted in decreasing the wire resistance to about 100 Ω and improved the average transconductance from 0.6 µS to 2.5 µS.

2.4.4.4 Atomic Layer Deposition (ALD)

As shown in Figure 21, approximately 50 nm of aluminum oxide (Al_2O_3) is deposited via atomic layer deposition (ALD) performed either on campus or at

Figure 21. Illustration of cross-section of transistor after atomic layer deposition (ALD) of top-gate oxide. The oxide layer (green) is depicted as unidirectional, but actual ALD films are conformal and anisotropic.

Cambridge Nanotech, Inc (Boston, MA). The principle of ALD is based on sequential pulsing of precursor vapors which individually react with the sample surface to generate one layer per pulse. The resultant oxide coating is highly conformal and of quality suitable for transistor gates. Other oxides such as thermally evaporated silicon dioxide layers were unreliable; pinholes in the insulator lead to cracking and degradation under conditions of electrolysis.

ALD films of aluminum oxide were deposited using alternating doses of trimethylaluminum ($Al(CH_3)_3$ or TMA) and water (H_2O). During the TMA pulse, methyl groups are released as the aluminum binds with oxygen from hydroxyl groups on the surface. The methane by-product is subsequently pumped out, along with any excess TMA. During the water pulse, the water vapor converts methyl groups of the bound TMA into fresh hydroxyl groups available for binding during the next TMA pulse. In this manner, the oxide film is built layer-by-layer.

On our sensors, the ALD grows on the silicon oxide substrate; the CNTs do not participate in the ALD reactions. Eventually the oxide film grows to envelop the CNTs. Ultimately the biosensor will be exposed to a wet buffer electrolyte full of ions. One of the most important reasons for choosing Al_2O_3 is that it is a good barrier against mobile alkali ions, which have been the cause of device instabilities in SiO_2 films.[26] For

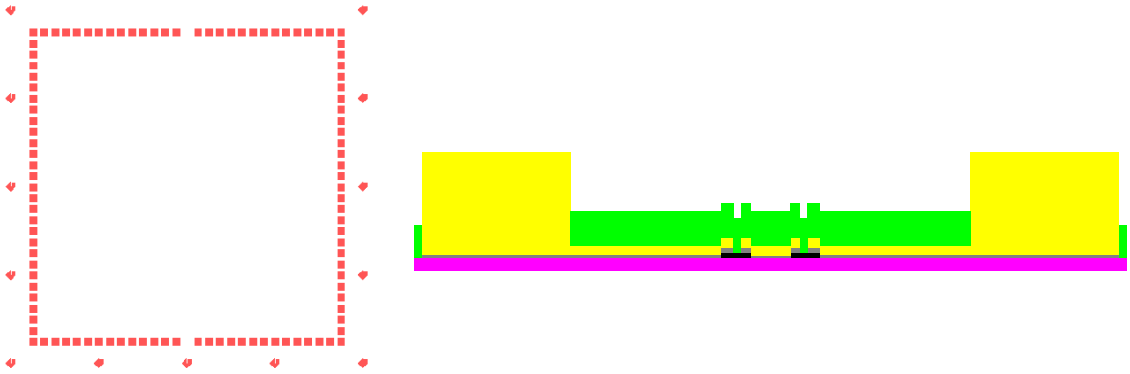


Figure 22. (left) Mask for bonding pads. (right) Cross-section of a CNTFET for which bonding pads have been deposited.

example, potassium, a common component of biological buffers, is mobile in silicon oxide.[27] Another advantage to using Al_2O_3 is that its high dielectric constant (high-κ) allows for a more sensitive device.

2.4.4.5 Bonding Pad Oxide Etch & Deposition

With ALD oxide covering the entire chip surface, the next step is to regain access to the bonding pad areas of the metal wire layer. NR9-1500PY photoresist was patterned to leave open features corresponding to the bonding pads (red squares) in Figure 22. To reduce oxide degradation due to thermal stresses, the pre-exposure bake and post-exposure bake were done at a reduced temperature of 110 °C for 5 minutes.

The top oxide covering the bonding pad areas is removed via fluorine-based reactive ion etching (40 mTorr, 18 sccm CHF_3, 2 sccm O_2, 175 W). This recipe normally achieves an etch rate of 50 nm SiO_2/min. However, since aluminum fluoride is not volatile, it only achieves 2.5 nm Al_2O_3/min; the etching may be dominated more by sputtering than chemical reaction. The etch is run for approximately 25 minutes, such that openings (windows) have been etched completely down to the underlying gold.

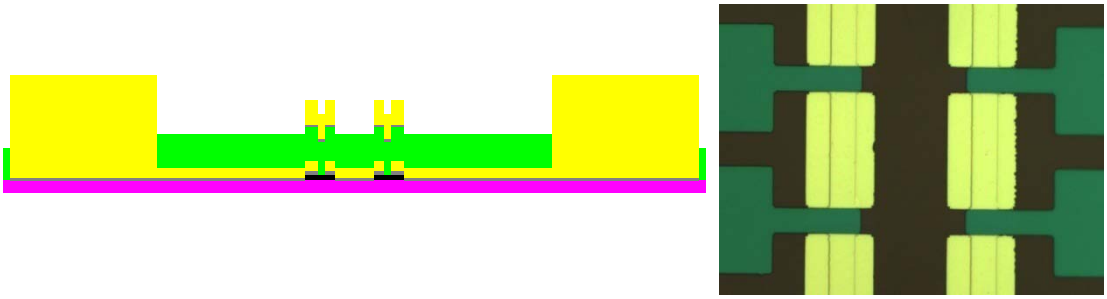


Figure 23. (left) Cross-section of finished CNTFET including gold attachment surfaces (GAS). Mask for field ash (FA) is reused for this step. (right) Microscopic view of six CNTFETs, including six gold gates, six drain wires, and one common source wire in middle.

After etching windows to the bonding pads, the same photoresist serves as a lift-off mask. Without removing the NR9 resist, 300 nm of gold is deposited via electron-beam evaporation. This relatively thick layer is needed for strong wire bonds. Liftoff is done in acetone as before.

2.4.4.6 Metal Gates

The next cleanroom process is to create the CNTFET gates, which also serve as surfaces on which to attach probe nucleic acids. The pattern for the gold attachment surfaces (GAS) is created using the same photomask used to remove unwanted CNTs (Figure 19). NR9-1500PY photoresist is patterned to leave open features corresponding to the CNT areas. Any remaining resist residue ("scum") is removed through exposure to oxygen plasma for 15 s. The resist pattern is now ready for the metal deposition and subsequent lift-off in acetone.

Since gold does not adhere well to aluminum oxide, we must first deposit an adhesion layer such as chromium (Cr) or titanium (Ti). Early in the development of metal-gated transistors, we found that titanium gave us V_{TH} too far negative, and that chromium placed the threshold voltages closer to zero. However, as we explain in detail

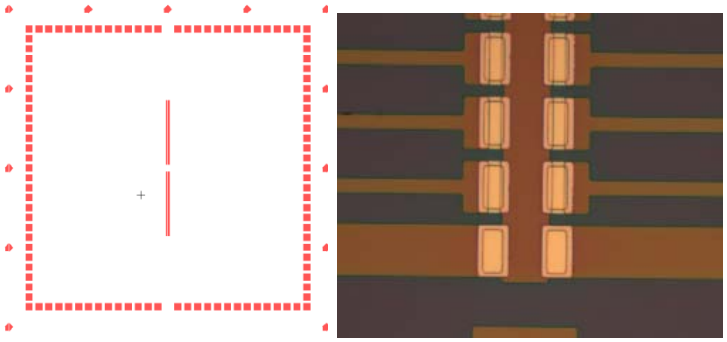


Figure 24. (left) Mask for protective coating. (right) Close-up of gate areas over which the protective coating has been patterned. The coating remains on the field but has been etched away over the metal gates. The coating is patterned such that the openings are smaller than the metal gates and bonding pads.

in Section 2.5.4.2, we ultimately switched to palladium (Pd) after we started using ALD oxide from an external vendor. The current practice is to deposit 7 nm Pd and 100 nm Au by e-beam evaporation. As shown in Figure 23, the resulting structure is a floating-gate CNTFET; testing is discussed later in Section 2.5.1.2.

2.4.4.7 Protective Coating

The final cleanroom process is the application of a protective coating that prevents etching of the aluminum oxide surrounding and underneath the metal gates. We discovered through side experiments that aluminum oxide, while relatively impervious to ion penetration, has the weakness of a slow but significant etch rate in the hybridization buffer. Overnight exposures to buffer caused the film to change color and became hydrophilic. The color change indicates a reduction in thickness, while the hydrophilic nature implies an abundance of fresh surface hydroxyl groups due to etching.

The mask for this step, shown in Figure 24, is a combination of the patterns for the metal gates and the bonding pad windows. First, Microchem PMGI SF8 is spun to a thickness of ~600 nm, then baked at 180 °C for 60 s on a hotplate. Next, the film is

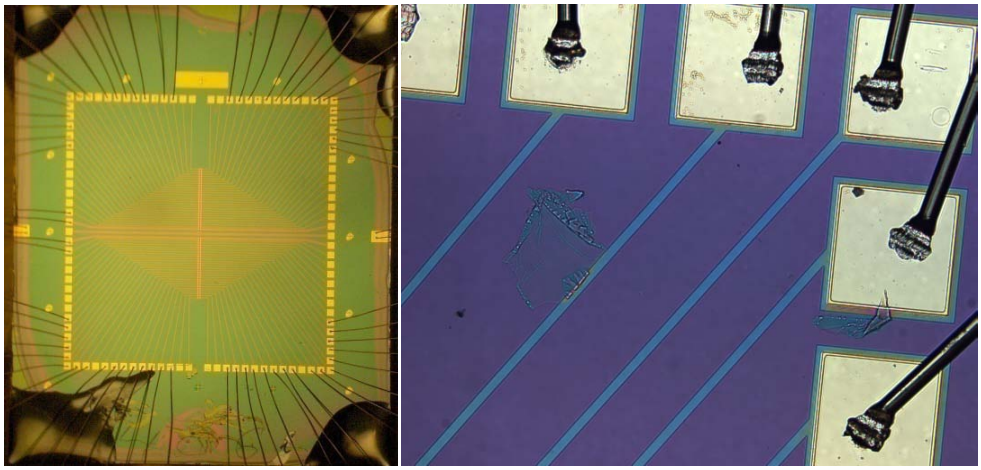


Figure 25. (left) Photograph of chip after mounting on the LCC068 and wirebonding. (right) close-up of wire bonds created by a wedge bonding tool with ultrasonic agitation. The bond wires are 25 μm in diameter.

thinned to ~150 nm in RD6 developer. After rinsing and drying, NR9-1500PY photoresist is spun and patterned to leave open features over the metal gates and bonding pads. The exposed PMGI SF8 is removed via oxygen plasma for about 4 minutes. At this point the chips are ready for final electrical testing on a probe station using needles to make direct connections to the source and drain, and a fine gold wire to contact individual gates.

2.4.5 Wirebonding & Reservoir Attachment

Upon passing the electrical quality control testing, the chip will be mounted to a chip carrier and wire bonded. The chip is mounted to the 68-pin ceramic leadless chip carrier (LCC) using room temperature vulcanizing (RTV) silicone sealant. As shown in Figure 25, the connections from the LCC to the chip are wirebonded using a West•Bond 747677E-79 wire bonder at LPS.

Next a liquid containment reservoir is attached using silicone sealant (Figure 11). Reservoirs were made from a section of a certified nuclease-free MBP pipette tip cut on one side using a hot-knife. Other variations of the reservoir have been used (such as sections of thick-walled polypropylene tubing, and thin-walled PCR tubes).

The biosensor can now be used in molecular biology experiments. Wet measurements can be done easily on the automated measurement platform. Liquid-gated chip performance is tested before any probe molecules are immobilized.

2.5 CNTFET Electrical Testing

Throughout the biosensor fabrication process, we perform a series of measurements to confirm the quality of the transistors. This is important because the downstream processes and experiments consume time and effort that must be carefully spent. The electrical measurements consist of mainly of the transfer characteristics and occasionally the output characteristics as measured on a probe station. Here we present data on representative transistors and extract parameters such as the transconductance and field-effect mobility. Although it is difficult to correlate the mobility of our devices with the intrinsic mobility of individual carbon nanotubes in the literature, we extracted a value still significantly higher than found in silicon MOSFETs. A more important parameter to us is the threshold voltage, which we characterize for entire CNTFET arrays during fabrication to ensure viability of the finished chip in the final application as a biosensor. Finally we review the improvements made in device fabrication at the batch level.

2.5.1 Transfer Characteristics

2.5.1.1 Bare CNTFET, Back-Gated

After a few lithography steps, the chip is mid-way to fabrication completion and is tested to confirm proper transistor operation. At this point, the carbon nanotube active areas have been defined and the non-active areas are cleared free of CNTs. The source and drain contacts, along with the wires leading to the bonding pads, have also been formed. Thus, we are able to measure "back-gate" characteristics using the basic setup shown in Figure 26 to scan the back-gate voltage V_{BS} while maintaining the drain bias V_{DS} = 0.1 V. While ohmic contact to the silicon substrate is not a trivial matter, we use a salt solution to conduct the small charging currents required.

The highly consistent transfer characteristics in Figure 26 indicate that the intrinsic device parameters such as number of tubes (N_{CNT}), back-gate mobility ($\mu_{FE(\text{back})}$), and gate length (L) are fairly constant across the chip. Most of the FETs have a minimum I_D around 1 μA and maximum I_D around 3 μA, which corresponds with 1/3 of nanotubes being metallic and 2/3 being semiconducting. The large counter-clockwise hysteresis is expected because the exposed CNTs are subject to shielding and polarization effects of water molecules adsorbed to the CNT and silicon oxide under it.[28] While the large hysteresis might impede biosensor designs based on exposed CNTs, the p-type transistor action at this stage is enough to strongly forecast good oxide-covered transistors at the end of the fabrication process.

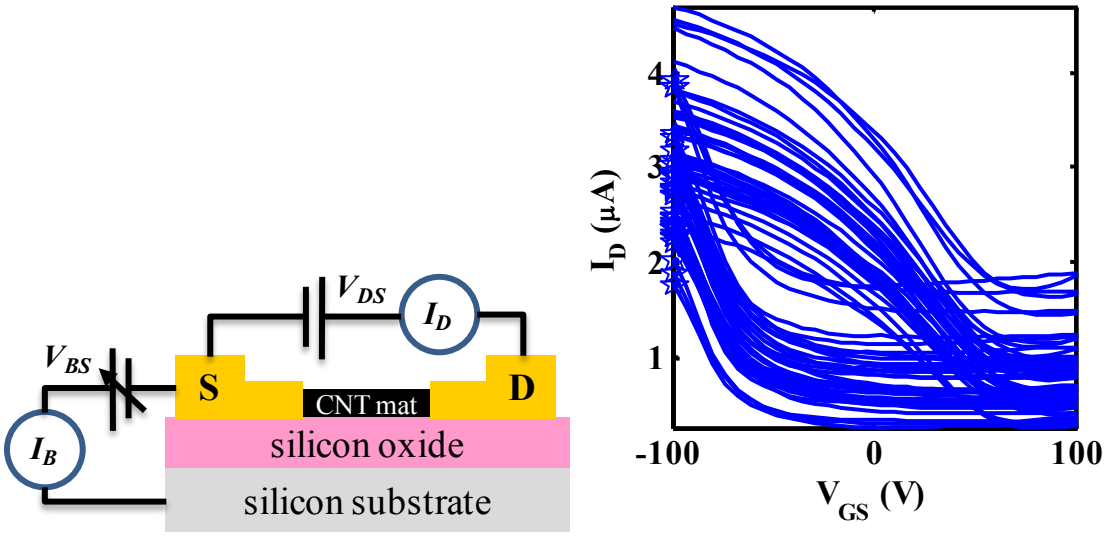


Figure 26. (left) Cartoon illustration of back-gate operation of a bare CNTFET in post metal wire (MW) stage. The back-gate voltage V_{BS} is applied via the substrate and is global to all CNTFETs fabricated on the same substrate. (right) Transfer characteristics of several bare CNTFETs from one chip (V_{DS} = 0.1 V). The counter-clockwise hysteresis loop starts at the stars on the left. The ohmic back-gate transconductance is $|g_{mb}| \approx 0.05$ μS and occurs near $V_{BS} = -100$ V. The gate current (not shown) was capacitive and less than 2 nA.

2.5.1.2 Metal-Gate CNTFET, Top-Gated

On the finished biosensor, each of the CNTFETs has its own gate electrode. As is not possible in a global back-gating configuration, we can conveniently modulate the gate voltage of a single transistor and measure the output and transfer characteristics at the probe station (Figure 27). This allows us to characterize the threshold voltage and assess any yield problems (such as gate leakage) on a per transistor basis.

Our paramount quality concerns boil down to two transistor parameters: the threshold V_{TH}, and the gate current I_G. For automated tracking of the threshold voltage, we need a parabolic I_D–V_{GS} curve with transconductance high enough so that noise and

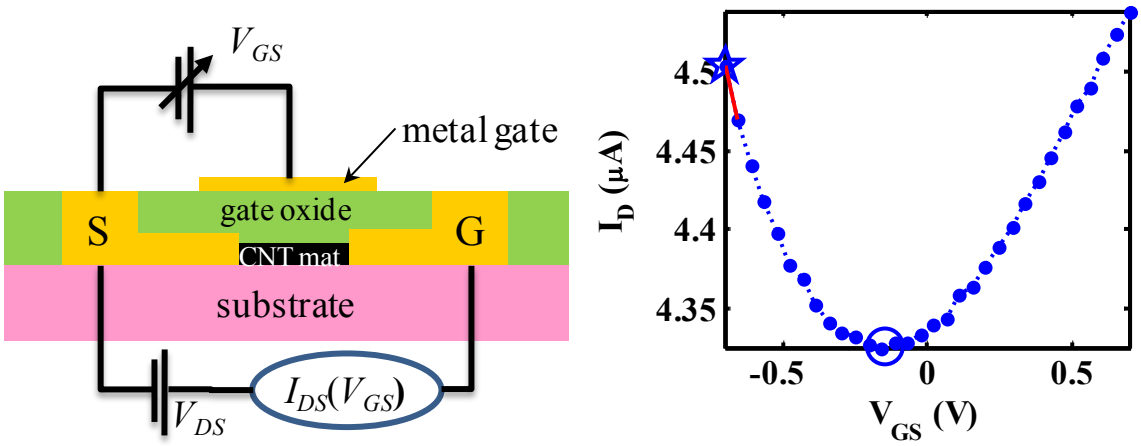


Figure 27. (left) Schematic of direct top-gated CNTFET with metal gate. As a simplification, we omit the back-gate substrate, which we short to the source. (right) Direct top-gate transfer characteristic of a CNTFET. Device has 30 nm of Al_2O_3 top oxide and metal-gate of 30 nm Au with 1 nm Ti adhesion layer. The ohmic top-gate transconductance (red mark) is $|g_m| = 0.8$ μS. The circle marks the threshold location. Direct contact to the GAS was made with a fine (25 μm diameter) wire.

hysteresis do not interfere. We also need $|V_{TH}| < 1$ V or so to avoid electrolytic reactions that occur at occasional oxide pinholes. Finally, the leakage current of the gate should be $I_G < 1$ nA over the intended V_{GS} scan range.

The direct top-gate transfer characteristic of a representative good transistor is shown in Figure 27. Compared to the bare CNTFET of Figure 26, this device is ambipolar, has practically no hysteresis, and has a V_{TH} readily identified by our custom threshold-tracking software. Furthermore, the top-gate transconductance (red line) is nearly 1 μS, which is 20 times the average back-gate transconductance of Figure 26. This is nearly four times better than the top/back ratio we predicted in Section 2.3.2! We attribute the improvement to a combination of increased modulation of the Schottky

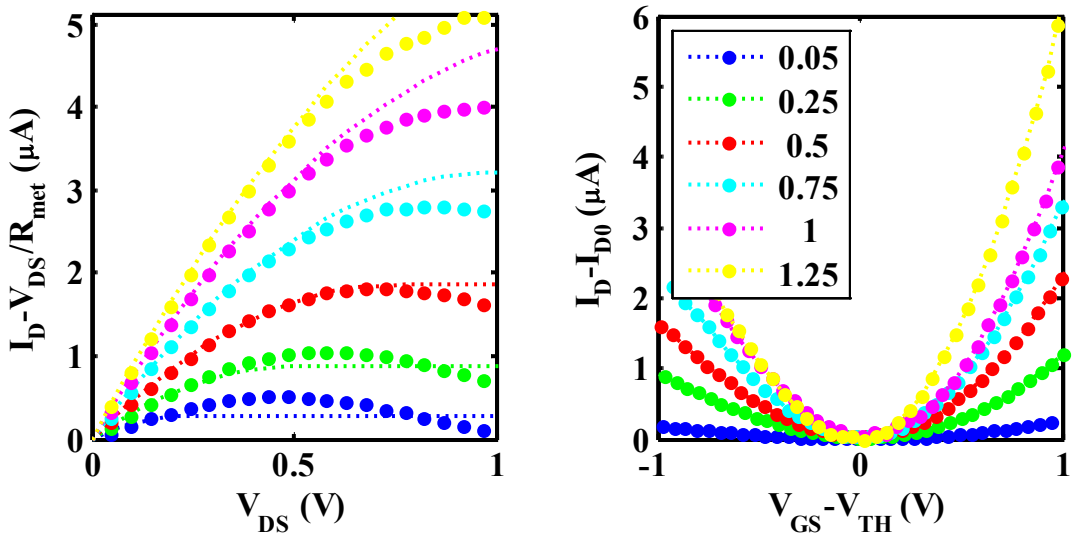


Figure 28. Measured output and transfer characteristics of a CNTFET, compensated for metallic component (~35 kΩ). (left) Output characteristics (I_D–V_{DS}), measured (dots) and from calculation (dotted lines) for V_{GS} = 0.5–1.75 V in steps of 0.25 V. Calculations used (N_{CNT} μ C_G/L) = 5 μA/V^2 (extracted from deep ohmic region) and $V_{TH} \approx$ 250 mV. The drooping in drain current at higher V_{DS} may be due to pulling of V_{TH} to higher value. (right) Transfer characteristic (I_D–(V_{GS} −V_{TH})) curves for V_{DS} = 0.05–1.25 V in steps of 0.25 V. The V_{TH} shifted from −58 mV to 487 mV under increasing V_{DS}. Variations in threshold (V_{TH}) and metallic/offset current (I_{D0}(V_{DS})) have been compensated for.

barriers at the CNT-metal contacts with the high-κ dielectric, plus reduced contact resistance due to annealing during the top-oxide deposition.

2.5.2 Output Characteristic

So far we have been primarily interested in transfer characteristics (I_D–V_{GS}), which allow us to easily track the device threshold and measure transconductance. However, we can also extract important parameters from the output characteristic (I_D–V_{DS}) such as output resistance and mobility.

We measured six I_D–V_{DS} curves and six I_D–V_{GS} curves of a single top-gated CNTFET. Both sets of data must be compensated for a significant metallic component that would other render the plots unrecognizable. For example, the output characteristic with the metallic background included look basically like a diagonal line (resistive). The adjusted output and transfer characteristics are plotted in Figure 28. The output characteristic also includes curves calculated from (1) and (2) in Section 2.3.1, assuming $(N_{CNT}\mu\ C_G/L) = 5\ \mu A/V^2$.

The measured I_D–V_{DS} curves are generally in good agreement with their calculated counterparts. However, one may notice a drooping in drain current at higher V_{DS}, which may be due to pulling of V_{TH} to higher values as semi-mobile charges move to shield the higher electric field across the channel. The measured transfer characteristics (I_D–($V_{GS} - V_{TH}$)) follow a highly parabolic trajectory as usual, but according to the model, should follow a more linear path at higher gate voltages. In both cases, the threshold voltage was not completely constant; in the transfer characteristic, the V_{TH} shifted from −58 mV to 487 mV under increasing V_{DS}.

While it may be possible to apply a more rigorous analysis to explain the deviations from theory using first principles, this would be very difficult given several unknowns such as the number of CNTs comprising the channel, the nature of their complicated interconnections, and the Schottky barriers at the interface of the CNT network to the metallic contacts. Rather, we resort to extracting a few key parameters.

The slope of the output characteristics in the deep ohmic region is the output conductance (inverse of the output resistance), which can be used with (5) to provide an

estimate of the field-effect mobility. Noting that g_{ds} is a function of V_{GS}, we measure an average value of $k_{ds} \equiv g_{ds}/(V_{GS}-V_{TH}) = 5\ \mu A/V^2$. This value is used to plot the calculated output characteristics in dotted line and was also used in the model of Figure 6.

The slopes of the transfer characteristics in the deep ohmic region can be used with (4) to provide another estimate of the mobility. We extract an average value of $k_m \equiv g_m/V_{DS} = 11\ \mu A/V^2$. We will use k_m and k_{ds} and g_m from the previous back-gate and top-gate measurements to calculate mobility estimates.

2.5.3 Calculation of Field-Effect Mobility

The mobility of individual carbon nanotubes has been thoroughly investigated by Dürkop *et al*.,[7] who extracted field-effect mobilities on the order of 10 $m^2/V{\cdot}s$. Since it is generally difficult to determine the mobility of the CNT channel itself due to contact effects, Dürkop *et al*. used very long channels ($L \approx 300\ \mu m$) to reduce the contribution of the contacts. Since our channel lengths are much shorter ($L = 10\ \mu m$), the contacts are more likely to reduce the apparent mobility of the devices. With this caveat in mind, we estimate the effective field-effect mobility from measurements on our CNTFETs.

The estimates of μ_{FE} depend on assumptions for the values of the numbers of nanotubes and the gate capacitance. We take $N_{CNT} = 1$ and $C_{G(back)} = 20$ pF/m and $C_{G(top)} = 100$ pF/m. The drain bias V_{DS} is always 100 mV. Using these values, we compute the field-effect mobilities, summarized in Table 2. The back-gate transfer characteristic of the bare (unpassivated) CNTFET yielded the lowest figure of 0.25 $m^2/V{\cdot}s$, while the highest figure of 1.1 $m^2/V{\cdot}s$ was extracted from the direct top-gated transfer characteristic of the finished CNTFET.

Table 2. Summary of field-effect mobility estimates extracted from measured transfer and output characteristics.

Measurement	g_m (μS)	C_G (pF/m)	$k_{ds}\,k_m$ (μA/V^2)	μ_{FE} (m^2/V·s)
Back-gate (Figure 26)	0.05	20	0.5	0.25
Top-gate (Figure 27)	0.8	100	8.0	0.8
Output char. (Figure 28)	varies	100	5	0.5
Transfer char. (Figure 28)	varies	100	11	1.1

While we hope for the gold standard of 10 m^2/V·s, we realize that these estimates are already optimistic because the effective number of CNTs per transistor is likely much greater than one. We are also reminded that the effective mobility of our devices is likely to be degraded due to contact effects (i.e. Schottky barriers at the metal-CNT interfaces). However, our mobilities, on the order of 1 m^2/V·s, are still 20 times better than typical silicon devices.

2.5.4 Fabrication Process Quality Control & Improvements

The goal of this project is to use our biosensors in a molecular biology experiment, rather than to explore the details of various device parameters. Under this framework, the fabrication and engineering activities behaved more like a company than a research group. We were responsible for producing over 200 biosensors with a certain level of quality and consistency amidst fluctuating process parameters as found in a research environment. In fact, it was not so much a matter of *maintaining* quality as it was a matter of constantly *improving* the quality of our devices.

In the following, we assess the performance of our biosensors at the chip level based on a few key parameters: chip threshold (VTH), chip transistor yield (Y), and chip transconductance (Gm). The chip VTH and Gm are simply the average of g_m and V_{TH} over all the CNTFETs on a chip (or a subset of the 104 FETs on a chip). The yield of a chip is defined as the ratio of the number of "good" FETs to the total number of FETs, where we define the criteria for "good" later. A "good" chip generally has easily distinguishable thresholds within a V_{GS} range that can be accessed without degrading the chip due to electrolysis.

2.5.4.1 Advantage of Metal Gate vs. Oxide Gate

As alluded to earlier in Section 2.4.4.6, one of the major challenges for us was controlling the threshold voltage of the CNTFETs. The transistor gate oxide and its interface to the outside world turn out to have a major impact on V_{TH}. For convenience, we define the *shift* in chip threshold as $\Delta VTH \equiv VTH_k - VTH_{k-1}$, where k is refers to the current trial, and $k-1$ is the previous trial.

As shown in Figure 29, the VTH of a gateless chip (oxide-covered but before gate metallization) is about 1200 mV. (In this special case, the top-gate voltage was applied via an electrolyte even before a reservoir was attached.) The direct-gated VTH of the finished chip (after gate metallization), is nearly 1400 mV lower. Conversion of the finished chip from direct to liquid gating produces a ΔVTH of just 140 mV. Given this miniscule shift, one might conclude that the metallization causes a large negative ΔVTH.

However, we recall that the threshold voltage of a MOSFET is:

$$V_{TH} = \varphi_M - \varphi_S + 2\varphi_F + \frac{\sigma_{etc}}{C_{OX}} \tag{8}$$

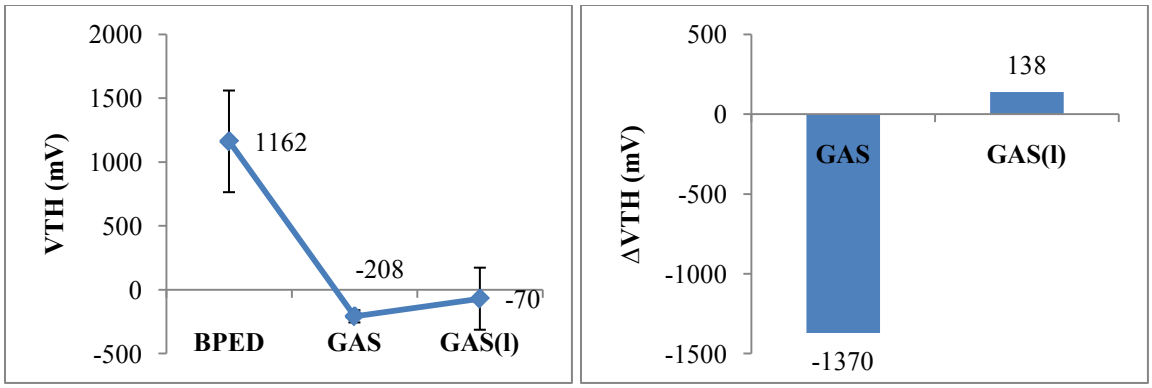


Figure 29. (left) VTH (over 6 CNTFETs) measured after completion of the following fabrication stages: BPED (Bonding Pad-oxide Etch and Deposition), GAS (gold attachment surface), and GAS (liquid gated). Error bars indicate one standard deviation. (right) ΔVTH shifts due to GAS (vs. BPED) and due to liquid gating vs. direct gating. The shift of −1370 mV is likely due to the work function of the metal gate, since liquid gating causes a shift of just +138 mV with respect to direct gating.

where φ_M and φ_S are the respective work-functions of the metal and semiconductor, φ_F is the semiconductor surface fermi level, and σ_{etc} is the sum various charges in the oxide and at material interfaces.[29] The gateless FET is bathed in electrolyte containing mobile ions of relatively low work function, suggesting that metallization of the gate should cause a threshold *increase*. Unfortunately this was not observed. Moreover, the similarity of VTH of the finished chip under direct and liquid gating suggests that the metal gate somehow stabilizes the chip. By process of elimination, our attention is focused on σ_{etc}, specifically the charges at the interface of the oxide and electrolyte. As will be seen in detail in Chapter 4, surface charges comprised of ionized amphoteric sites on the oxide surface indeed generate a significant potential that shifts the V_{TH} in the direction opposite to their polarity.

Relative to the metal-gate transistor, the gateless transistor has a quite positive threshold, implying negative surface charge. This is quite expected because aluminum oxide tends to form negatively charged hydroxyl (-OH) surface groups when exposed to water. The variability of this process with respect to ionic concentration and pH is a vulnerability that supports the argument for protecting the gate oxide with metal gates.

Furthermore, the relatively higher spreads of thresholds in both liquid-gated experiments suggests that the electrolyte and/or Ag/AgCl electrode may play a role in increasing the noise of the system.

2.5.4.2 Controlling Threshold Voltages

Problems with threshold voltage rose again when we needed to switch the vendor providing the ALD oxide growth. As shown in Figure 30, the standard recipe gave us VTH fairly close to zero, which is good for avoiding electrochemical degradation. Due to material supply issues, we were forced to send three batches of chips to an external vendor to perform ALD "for-hire". The fabrication process was then completed on a handful of chips that yielded a batch average VTH ≈ −2000 mV, which would require large V_{GS} scan ranges that would cause the loss of many chips.

We again refer to (8) to solve this problem. The only term we have control over is φ_M. In past experience, chromium and titanium, with work functions of 4.5 eV and 4.33 eV, respectively, yielded thresholds less than zero. In fact, this is how we made many unusable transistors with thresholds from −2 to −4 V. Only a handful of elements (Pt, Pd, Ir, Au, and Ni) are available to us with a work function above 5 eV, but most of these were also known to have poor adhesion to oxide. Ultimately, palladium (Pd,

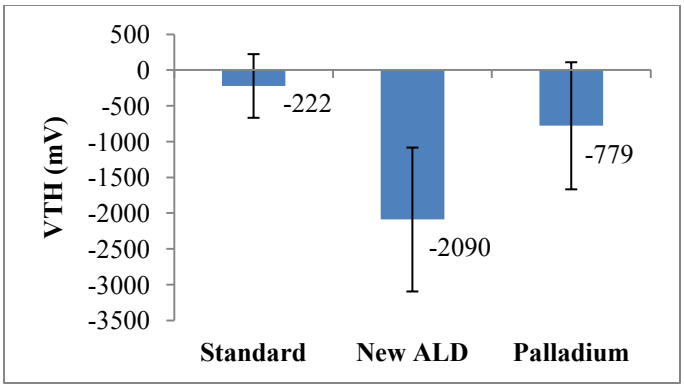


Figure 30. VTH of three incremental versions of chips: standard recipe, new vendor for ALD oxide, and change in metal gate adhesion material. The new oxide caused a large shift in VTH. Change the adhesion layer from Cr to Pd brought the VTH back to a reasonable value. The mean values and error bars are taken over these 15, 6, and 10 chips, respectively.

5.41 eV) proved to move the threshold voltages closer to zero (VTH ≈ −780 mV) and provide sufficient adhesion.

2.5.4.3 Improving Transistor Yield

Another major problem we faced was a low yield of working transistors per chip. A working transistor ("pass") basically has a clear threshold voltage with low gate leakage. A transistor that fails may have a transfer characteristic with undesirable properties, such as high gate current, more than one minima in the transfer function, or a threshold voltage that is outside of the acceptable range (window). A MATLAB script was written to automatically categorize each transistor as "pass" or "fail". If the percentage of good transistors on a chip is more than about 75%, we deem the chip worthy of use in molecular biology experiments.

By far the major symptom of transistor failure was gate current leakage, which tends to be super-imposed onto the transfer characteristic when some portion of the gate current leaks into the drain current. An investigation was launched to figure out why

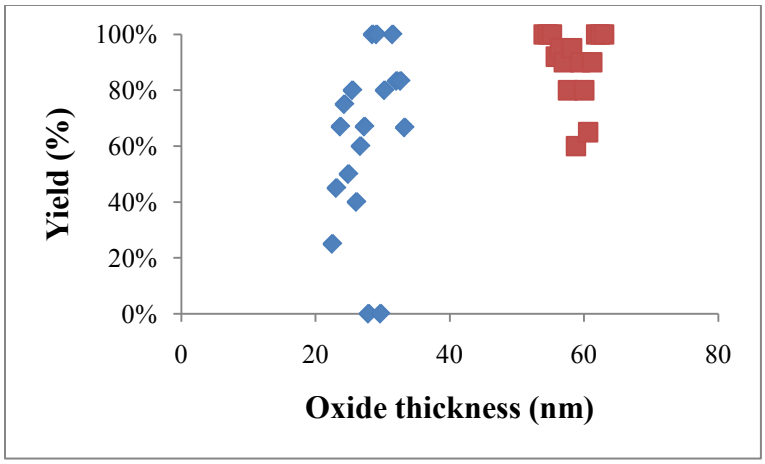


Figure 31. Transistor yield vs. ALD oxide thickness for 16 chips using the standard thickness (blue) and 22 chips with double thickness (red). Doubling the oxide thickness improved the transistor yield from 62% to 91%.

some chips were not yielding any useful number of good transistors. As shown in Figure 31, the percentage of transistors on a chip with acceptable gate leakage was originally 62%, with large variation.

Figure 32 contains two SEM images of our biosensors. The first image is an almost "side-view" of a metal layer coated with ALD oxide. Asperities exist on both the substrate and metal layer with roughness similar in magnitude to layer thicknesses. The other image, a top view of the active area of the CNTFET, shows that significant roughness under the gate area is imparted by the presence of CNTs and other presumably organic particles in the transistor active area (continually protected from O_2 plasma cleaning). It follows that the multitude of asperities may hinder the development of a thin, uniform conformal oxide.

This realization prompted the doubling of ALD oxide thickness to reduce the chance of gate leakage. As seen in Figure 31, this action resulted in an improvement of yield to an average of 91% with full yield for many chips.

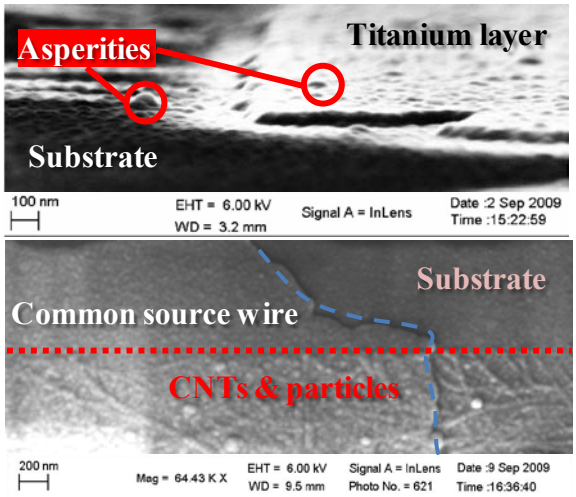


Figure 32. SEM images revealing source of surface roughness possibly responsible for loss of gate isolation. (top) Asperities can be seen on substrate and metal layer on this chip, which has a 30 nm oxide layer over a patterned 50 nm metal layer. Roughness of surface is similar in comparison to layer thicknesses. (bottom) Significant roughness under the gate is imparted by the presence of CNTs and other presumably organic particles in the transistor active area, which was protected from O_2 plasma cleaning. It follows that the multitude of asperities may hinder the development of a thin, uniform conformal oxide.

2.5.4.4 Increasing Transconductance

Based on the literature on annealing CNTFETs,[16] the metal layer of our early chips was a single layer of titanium (Ti) 50 nm thick. Upon realizing that we did not need to pursue annealing but needed to decrease the source and drain wire resistance, we changed the recipe to 100 nm of gold (Au) on 10 nm of titanium (Ti). The gold, with its low sheet resistance, takes the role of the main conductor, whereas the titanium promotes adhesion to the substrate and forms intimate contact with the CNTs. As shown in Figure 33, this resulted in an increase of the average transconductance from 0.6 µS to 2.5 µS.

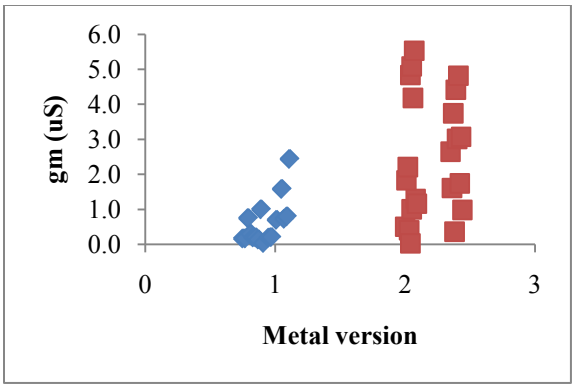


Figure 33. Transconductance of chips from metal layer version 1 (Ti only) to version 2 (Au/Ti). Average transconductance was improved from 0.6 µS to 2.5 µS.

2.6 Threshold Stability

It is important to establish the stability of our transistors so that changes in the threshold voltage can be assumed to come from nucleic acid hybridization. We assess stability in two different experimental configurations: dry direct gating and liquid-gating. Direct-gating proves stability of the transistor itself, whereas liquid-gating further tests the stability of the electrodes and the transistor immersed in electrolyte.

2.6.1 Direct-Gating

To characterize individual transistor stability, we repeatedly scan the direct-gated transfer characteristic of one transistor at a probe station. The gold top gate of the transistor is contacted by a thin, flexible gold wire so as to not damage the gate or underlying oxide. The drain and back-gate voltages are held constant (V_{DS} = 100 mV, V_{BS} = 0 V). After visually confirming placement of the probes, the V_{GS} scanning is started.

Figure 34 shows the threshold voltage of a representative CNTFET over 100 consecutive scans over a period of eight minutes (a sampling interval of 5.2 s). The V_{TH} experiences a significant amount of instantaneous fluctuation but drifts downward from

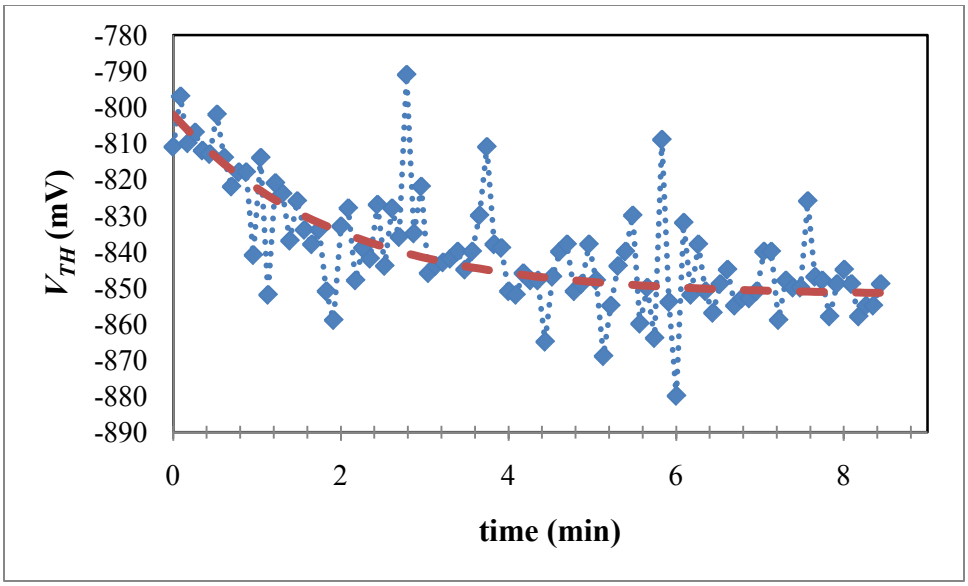


Figure 34. Threshold voltage of a direct-gated CNTFET over 9 minutes of repeated scanning (100 scans). An exponential settling function (red dashed) with a time constant of 1.9 minutes is fitted via a least-squares error method. The RMS noise level is ~11 mV. The V_{GS} scan range is (−2, 0) V.

−810 mV and settles at a level of −850 mV. The drift follows an exponential settling pattern over the long-term with a time constant of 1.9 minutes. The RMS noise level is ~11 mV; however, a few large shifts of about ±50 mV exist.

We postulate that the settling behavior is due to relaxation of the top-gate capacitance through a finite gate resistance, which abides by a time constant given by $\tau_{OX} = R_{OX}C_{OX}WL$, where R_{OX} is the gate resistance, C_{OX} is the gate capacitance per unit area (as defined before), W is the gate width (50 µm), and L is the gate length (10 µm). Given the measured time constant and gate capacitance of $C_{OX}WL$ = 740 fF, we estimate the gate resistance is R_{OX} = 160 TΩ. According to data from the ALD oxide company, the oxide resistance should be 2000 TΩ. The higher figure achieved by the company may be due to the testing method (liquid Hg probe), or better substrate/growth conditions. We recall that ALD oxide does not nucleate directly on CNTs.

Fluctuations in the threshold voltage of CNTFETs are especially expected given the small dimensions and high surface to volume ratio of CNTs. Long-term tracking of a DC parameter such as threshold voltage is also prone to 1/*f* noise, which increases at lower frequencies. Collins *et al*.[12] found that a wide variety of single- and multiwalled carbon nanotube devices (including individual SWNTs, two-dimensional networks, and three-dimensional mats) all exhibit a relatively large amount of 1/*f* noise compared to devices made of other materials. This is not surprising given that the current in nanotubes is transmitted through surface atoms and is easily perturbed by local charge fluctuations. Snow *et al*.[30] investigated 1/*f* noise in two-dimensional carbon nanotube networks like the ones used in this work and found that increasing the device size can reduce the 1/*f* noise. They also postulated that the gate field might also produce charge fluctuations in the gate dielectric.

Rather than perform a full noise analysis to ascertain the physical noise sources in our transistors, we conclude that to reduce the effect of 1/*f* noise, we should avoid the use of excessively long measurement durations in an attempt to average out thermal noise. For example, it should be sufficient to measure the V_{TH} a few times before hybridization and take the mean. Do likewise after hybridization. Since the hybridization signal is the shift between two threshold voltages, and given the noise in the settled V_{TH}, it follows that the uncertainty in any hybridization signal is about 10 mV at the minimum.

Like 1/*f* noise, the observed settling behavior may be due to charges within the oxide dielectric which might move under influence of the gate potential. In this experiment, the V_{GS} was biased at 0 mV before the scanning was initiated. During the scans, the V_{GS} goes from −2000 mV to 0 mV, applying an average of −1000 mV to the

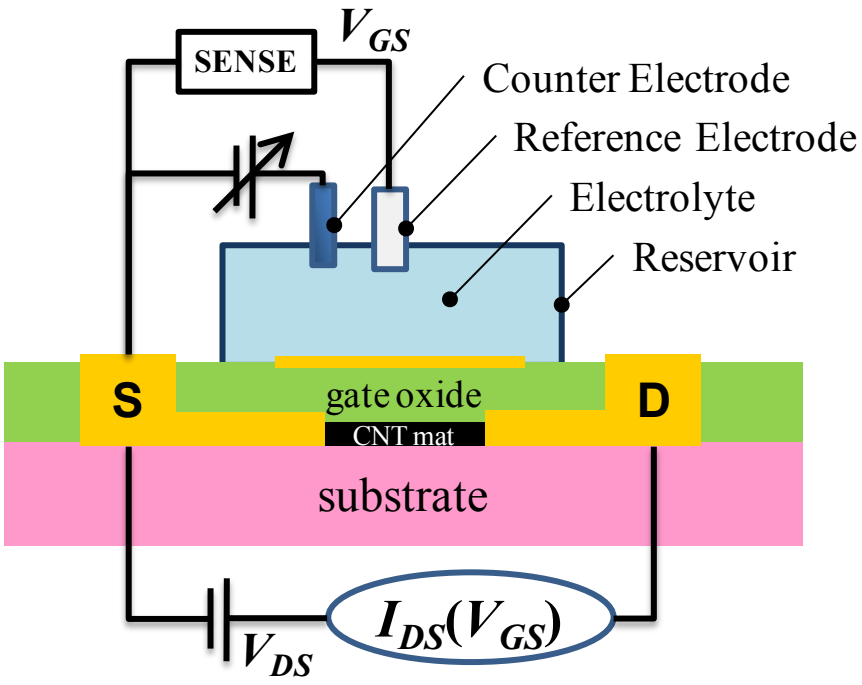


Figure 35. Schematic of liquid-gated CNTFET after reservoir attachment. As a simplification, we omit the conductive back-gate substrate, which we short to the source.

gate. This change in average V_{GS} level may have attracted positive charges or repelled negative charges underneath the metal gate, causing a gradual negative shift in threshold. To avoid this effect, it is imperative to use a symmetrical V_{GS} range.

2.6.2 Liquid-Gating

Next we are interested in assessing the stability of the liquid-gated CNTFET (Figure 35), which takes into account the interaction of the electrolyte with the reference electrode, counter electrode, and metal gate. The threshold voltage of a CNTFET over 280 successive liquid-gated scans is shown in Figure 36. The V_{TH} began in the vicinity of −470 mV and drifted upward by 80 mV to a final level of −390 mV, following an exponential settling function with a time constant of 25 minutes. The RMS level of the noise is ~25 mV.

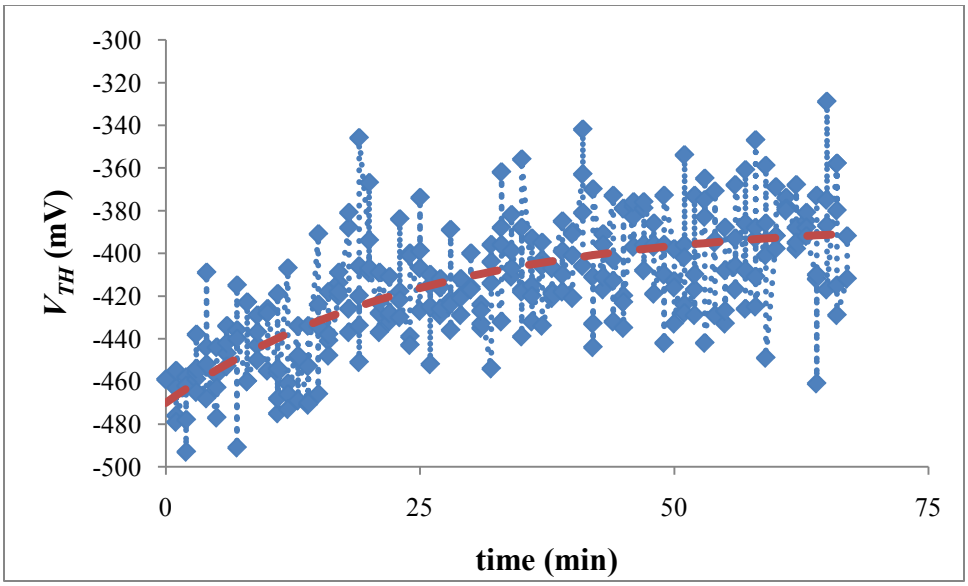


Figure 36. Threshold of a liquid-gated CNTFET over 67 minutes of repeated scanning (280 scans). An exponential settling function (red dashed) with a time constant of 25 minutes is fitted via a least-squares error method. The RMS noise level is ~25 mV. The V_{GS} scan range is (−1.0, +1.2) V.

As in direct top-gating, we again assume that the settling behavior is due to relaxation of a capacitance through a finite resistance. However, we must take into account additional resistances and capacitances introduced by liquid gating. As is explained in Chapter 4, the capacitance at the electrolyte-gate interface is relatively large compared to the oxide capacitance. This could easily increase the capacitance term in the RC time constant by a factor of 50! Thus, we attribute the longer time constant in liquid-gating to the electrolyte.

The noise level in V_{TH} in this liquid-gated setup is about twice the level of the dry direct-gated setup. Hassibi *et al.*[31] modeled the noise processes in the ISFET and showed that a main source of noise is the electrolyte thermal noise (both in the bulk and the diffusion layer). Thus, we also attribute the increased noise to the presence of the electrolyte.

2.7 Conclusion

We designed, fabricated and tested more than 200 biosensor chips, each containing an array of over 100 CNTFETs. The high mobility of carbon nanotubes makes them an attractive candidate for high-sensitivity field-effect detection of the intrinsic electric charge of DNA and RNA. However, the extraordinarily high sensitivity of CNTs is accompanied by poor selectivity. By coating the CNTs with ALD oxide, we strike a balance between selectivity and sensitivity.

Before employing the biosensors in molecular biology experiments, we electrically test them for two important parameters: threshold voltage and gate current. The thresholds must be easily distinguishable and within a range that can be accessed without degrading the chip due to electrolysis. The gate current must be low so as to ensure that the metal gate surface is electrically isolated from the rest of the transistor, and that the transistor itself is protected from destructive electrolysis.

We realized the advantages of metal gates as opposed to the typical oxide-only gate of an ISFET. The metal gate protects the underlying oxide from chemical etching by the hybridization buffer, as well as allows for control of threshold voltages by careful choice of the adhesion layer material. Throughout the production of the CNTFETs, we improved the transistor yield to over 90%, which is excellent for a research process.

We characterized the threshold stability of the biosensor under direct gating and liquid gating conditions. Our devices have low hysteresis and a settling behavior as expected considering the interaction of the electrolyte and gold gate electrodes.

In practice, all devices exhibit some hysteresis and drift. This may be due to semi-mobile or slow charges. For example, it is known that the aluminum oxide and its

interfaces to other materials trap such charges. These may also explain why the threshold voltage depends on the gate voltage range. Once the range has been chosen for a series of measurements, it will be kept constant.

After the biosensors have been fabricated and shown to be electrically operational, it is time to functionalize them with the appropriate probe that will target the disease of interest. The next chapter will provide an introduction to the disease of breast cancer and the molecular biology behind testing for it.

Chapter 3: Introduction to HER2, Existing Testing Methods, & Proposed Biosensor

3.1 Chapter Purpose

This chapter uses HER2 as a framework to educate the reader about breast cancer and molecular biology. We briefly overview established HER2 testing methods and present our vision of an integrated lab-on-a-chip system to incorporate the various processes required to extract and analyze HER2 mRNA from a sample of cancer cells. We review the literature on biosensor research, focusing on the state-of-the-art in direct label-free RNA detection.

We then detail our collaboration with the National Cancer Institute (NCI) to design and synthesize high-performance probes for HER2. One HER2 probe was selected for the biosensor, implemented in peptide nucleic acid (PNA) technology, and required the development and validation of new surface attachment and hybridization protocols. Finally we demonstrate validation of the probe with established fluorescent methods.

3.2 Breast Cancer & HER2

In 2009, more than 192,000 women were diagnosed with breast cancer in the United States.[32] The National Cancer Institute (NCI) estimates that this year, 207,090 more women will be diagnosed, and that nearly 40,000 will die from it.[33]

For some fraction of breast cancer patients, a possible treatment is the antibody Herceptin® (trastuzumab), which costs nearly $100,000 a year. Insurance is usually

hesistant to cover the cost. This treatment is only effective in cases where the cancer cells have an overabundance of the HER2 protein, causing them to be more aggressive than other types of breast cancer. The aggressive nature of HER2 and the high cost of Herceptin® bolster the need for widespread, routine laboratory testing to determine HER2 status.

We embarked on an engineering project with the goal of lowering the barrier to HER2 testing via commercialization of an integrated testing system. We provided a prototype system that was used at NCI to determine the HER2 status of breast cancer cell lines. A commercialized platform must be able to provide valid diagnostic results repeatedly, as opposed to a one-off prototype with sporadic performance.

There are numerous challenges in the adoption of a new technology in a field of study that evolved from empirical observation and inferences. Many methods involve or have evolved from direct observables such as color change, or fluorescent light emission. Furthermore, given the complexity of biological systems, validation based on agreement with first principles calculations, as is often acceptable in engineering, is often not possible or even necessary in molecular biology. The legitimacy of an experiment is based on sound formulation, inclusion of the proper controls, and comparison with existing data. We see this methodology as we review established and possible alternative testing methods for HER2.

3.3 HER2 & Testing Methods

3.3.1 Background

HER2 is a receptor protein that resides on the surface of certain cancer cells and is associated with higher aggressiveness in breast cancers. All normal cells contain the genetic instructions for the production of HER2, which can help promote normal growth of the cell. However, in about 25% of breast cancer patients, the cancer cells have an excessively high number of HER2 receptors.[34] These HER2-positive (HER2+) cells are more likely to grow and reproduce uncontrollably when stimulated by growth factors that are naturally occurring in the body.

Research into HER2 has found that overexpression of this receptor is associated with increased disease recurrence and worse prognosis. HER2 has been identified as a marker of advanced stage breast cancer.[35] Even women with apparently benign (as seen under a microscope) breast tissue samples that ultimately test HER2 positive appear to be at increased breast cancer risk.[36] Because of its prognostic role as well as its ability to predict response to trastuzumab (trade name Herceptin), HER2 status tests are recommended for all invasive breast cancers.[37]

3.3.2 HER2 Status Testing

Testing to determine HER2 status is performed on samples taken from an initial biopsy or breast cancer surgery. A core biopsy is the use of a needle to remove a tissue sample from a mass or lump. The tissue can then be examined under a microscope and submitted for laboratory testing. The basic histology is performed by examining a thin slice (section) of tissue under a light microscope. At this point, a limited diagnosis of

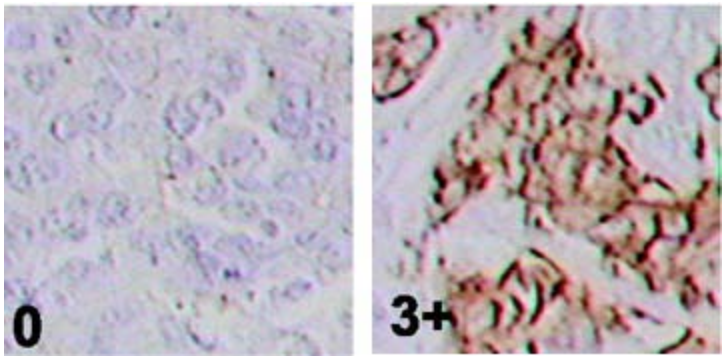


Figure 37. Light microscopy images comparing of HER2 expression in primary breast tumors via IHC. (left) HER2 expression scored as 0. (right) HER2 expression scored as 3+ (=positive). Expression of the HER2 protein was determined on 5-μm paraffin sections by a two-step immunohistochemical staining. Following incubation with the primary rabbit antibody (DAKO) to human HER2 protein, a secondary goat anti-rabbit immunoglobulin conjugated with horseradish peroxidase was applied for visualization. Source: Vinatzer *et al*. [38]

cancer can be made, such as benign, malignant, or indeterminate. To determine HER2 status, the tissue must be sent to a laboratory to undergo immunohistochemistry (IHC) and/or fluorescence in-situ hybridization (FISH).

3.3.2.1 Immunohistochemistry (IHC)

As shown in Figure 37, immunohistochemistry can be used to reveal how much HER2 proteins are present, or expressed, on the surfaces of cells. It is basically a staining technique that is able to hone in on HER2 by exploiting the principle that antibodies bind specifically to matching proteins. In the simplest scheme, the antibody probe contains a binding site that matches the HER2 receptor, as well as a reporter dye molecule for contrast visualization. For increased sensitivity, a stronger signal can be produced by the use of unlabeled *primary antibodies* that are detected by labeled *secondary antibodies* that bind to it.

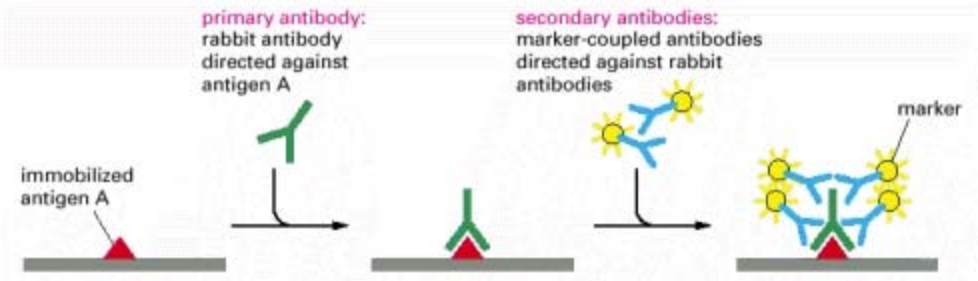


Figure 38. Schematic of indirect immuno-cytochemistry. The immobilized antigen A corresponds to HER2 receptors on the cell surface. The primary antibody is derived from rabbits but is engineered to have a component specific to human HER2 protein. The secondary antibody is anti-rabbit (only binds with rabbit antibodies) and is conjugated with a marker, or label. This detection method is very sensitive because many molecules of the secondary antibody are able to bind to a single primary antibody. Source: MBoC.[39]

Figure 38 illustrates the principle of signal amplification using primary and secondary antibodies. The HER2 receptors (antigen A) on the cell surfaces bind with primary antibodies engineered from rabbits to specifically probe for HER2. The secondary antibody is chosen because it binds only with rabbit antibodies, and was engineered to contain a marker, or label, for the purpose of visualization. This indirect detection scheme is very sensitive because many molecules of the secondary antibody are able to bind to a single primary antibody. The FDA-approved IHC assays HercepTest® (DAKO) and Pathway™ (Roche) are based on the principle of indirect immunohistochemistry.

3.3.2.2 Fluorescence *in situ* Hybridization (FISH)

Fluorescence *in situ* hybridization (FISH) checks whether the cancer cells have a normal level of HER2 gene copies in the cell nuclei. If the number of chromosomal copies of HER2 is higher than normal, this is called gene amplification. Figure 39 contains fluorescent images from a HER2 normal and HER2 amplified tissue specimen.

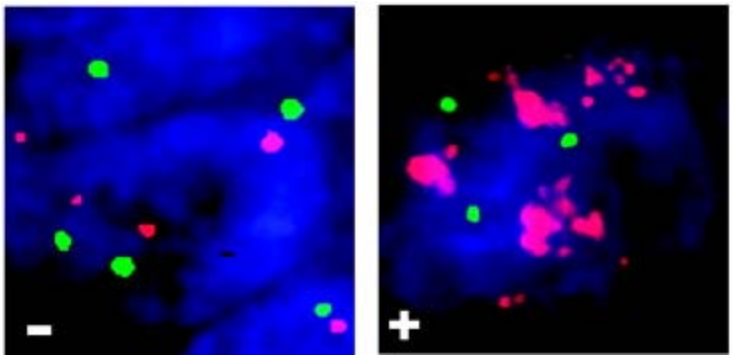


Figure 39. Fluorescent images from FISH comparing HER2 gene amplification in primary breast tumors in patients showing HER2 signals (red) and centromere 17 signals (green). (left) A normal (HER2−) nonamplified specimen. (right) A HER2 gene amplified (HER2+) cancerous specimen. Tumors were scored as harboring a HER2 amplification when the ratio of the average HER2 copy number to that of centromere 17 was ≥2.0, as required by the U.S. Food and Drug Administration for the approval of this method. Source: Vinatzer *et al*.[38]

The HER2 signals are red, while the reference signals are green. HER2 gene amplification is determined from the ratio of the number of signals from HER2 to that of centromere 17 is ≥2.0, as required by the U.S. Food and Drug Administration.

The principle of FISH, illustrated in Figure 40, is the specific hybridization of a fluorescent probe nucleic acid sequence with its complementary sequence within chromosomal DNA. The HER2 probe spans the entire HER2 gene sequence, as determined from a gene sequencing database, and is attached to a fluorophore (i.e., SpectrumOrange™, Texas Red®). The prepared sample is incubated with the HER2 probe. If the probe is complementary to DNA in the chromosomes of the cell, it will hybridize with it. Probe that has not hybridized with genes in the cell will be washed away. Then the hybridization signals are observed under fluorescent microscope.

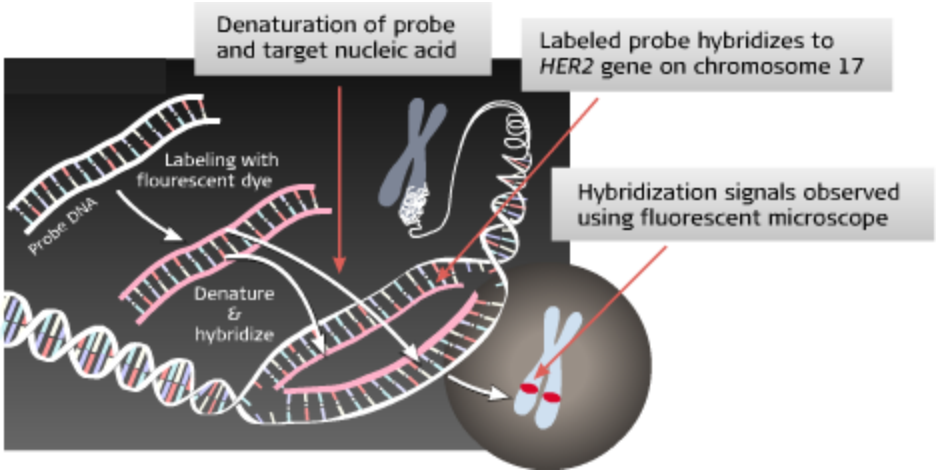


Figure 40. Principle of FISH. Probe DNA sequence, labeled with fluorescent dye, is introduced into a cell. The HER2 probe specifically hybridizes with its complementary sequence in chromosome 17. After unbound excess probe is washed away, the remaining hybridization signals are observed under fluorescent microscope. Source: Dako.[40]

Table 3. Comparison of FISH and IHC.

	FISH	**IHC**
Overall cost	High ($800)	Low
Time to result	>4 days	>3 days
Amount of training required	High	Low
Internal control	Yes	No
Interpretation	Objective/quantitative	Subjective/qualitative

3.3.2.3 Comparison of IHC & FISH

A comparison of FISH and IHC is shown in Table 3. FISH is currently the gold standard because it directly measures HER2 gene amplification at the DNA level. It is also expensive ($800/test, billed to patient), due to the labor and reagent costs. Time of skilled labor is consumed in preparation, overnight hybridization, and signal counting under a microscope.

Compared to FISH, IHC is relatively lower in cost, time, and skill requirements. Its weakness is that it relies on antibodies, so is prone to false negative and false positive results due to inappropriate tissue handling, and may even lead to false positive results due to staining artifacts.[41]

Both FISH and IHC require tissue samples to be taken from the patient, fixed, and embedded in wax. The technique of core needle biopsy uses hollow needles to extract cylindrical samples. The needles are up to 1.6 mm in diameter (14 gauge) and 20 cm long. The needle is inserted into the person while using ultrasound or another technique to identify the approximate location of the suspected tumor. Typically three to six cores will be extracted in one session, frozen, and sent to a laboratory for analysis.

At the laboratory, these samples are dehydrated, fixed in formalin for 18 to 24 hours, dehydrated in ethanol and xylene, and embedded in paraffin wax. The formalin-fixed paraffin-embedded (FFPE) tissue sample must then be cut into thin slices of about 5 μm thickness. Finally, the sections must be dried and frozen. This labor-intensive process takes at least two days.

Neither FISH or IHC is solidly quantitative. For FISH, a specialist must observe about two-hundred tumor cells under a microscope, counting the red (HER2) and green (CEN-17) signals. Tumors are scored as harboring HER2 amplification when the ratio of the average HER2 copy number to that of centromere 17 is $\geq$2.0 (required by the U.S. FDA). In the case of IHC, the criteria are even more subjective. Four scoring levels exist (0 to 3+) that correspond to the intensity of the stain. If the tissue scores 0 to 1+, it is called “HER2 negative.” If it scores 2+ or 3+, it’s called “HER2 positive.”

We have seen two approaches, FISH and IHC, which gauge the presence of HER2 at the structure level and the gene level, respectively. Like a blueprint, DNA encodes genetic information about the development and function of living organisms, and is usually found in the well-known double helix form. Proteins perform many roles in the structure, signaling, and enzymatic reactions of cells. Their specific sequences of amino acids are indeed generated from corresponding sequences in DNA.

There is another nucleic acid that acts as an intermediate between DNA and proteins. Ribonucleic acid (RNA) serves as a template to convey the information in DNA for the synthesis of proteins. The overall process, illustrated in Figure 41, is summarized in the central dogma of molecular biology, which states that genetic information in each cell of the body is stored in deoxyribonucleic acid (DNA), transcribed to ribonucleic acid (RNA), and finally translated into protein. The specific sequences of amino acids uniquely identify a protein and indeed are generated from corresponding sequences in DNA.

The amount of specific proteins in a given cell depends upon the expression levels of the corresponding messenger RNA (mRNA) that has made its way out of the cell nucleus. mRNA is a processed version of the “blueprint” code from DNA and serves as a template for translation into protein. Thus, an alternative approach to FISH and IHC could be to measure the expression of HER2 at the mRNA level.

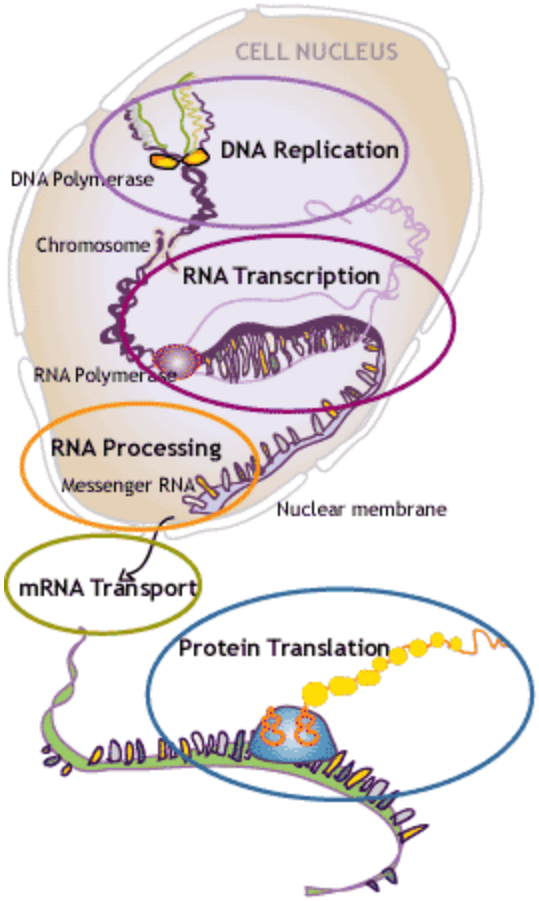


Figure 41. Schematic of relationships of DNA, RNA, and proteins in the cell. DNA in the cell nucleus is organized into chromosomes. The corresponding genes are transcribed into RNA. RNA is processed (non-coding parts are removed) into mRNA, which is moved out of the nucleus in the form of mRNA. mRNA contains the code for protein translation. Source: NobelPrize.org.[42]

3.3.3 Potential Alternative Testing Methods

3.3.3.1 Q-RT-PCR

As a potential alternative for assessment of HER2 status, Vinatzer *et al*. explored quantitative real-time reverse transcription-polymerase chain reaction (RT-PCR) as a fast and inexpensive method for producing quantitative results.[38] First, they extracted total RNA from homogenized tumor tissue and reverse transcribed it into complementary DNA (cDNA), which was then amplified by polymerase chain reaction (PCR). Reverse transcription is required because the replication enzyme (DNA polymerase) can act only on DNA.

PCR is able to amplify target nucleic acid sequences a billionfold (Figure 42). Two primers (probes) with sequences that flank the sequence of interest are synthesized chemically. Upon mixing with the sample of DNA, PCR proceeds in cycles of heating and cooling. Every cycle doubles the amount of DNA as compared to the previous cycle. In practice, 20 to 30 cycles are able to produce enough copies of the gene to be detected by other methods such as fluorescent hybridization. Variations on the standard PCR have evolved since its inception.

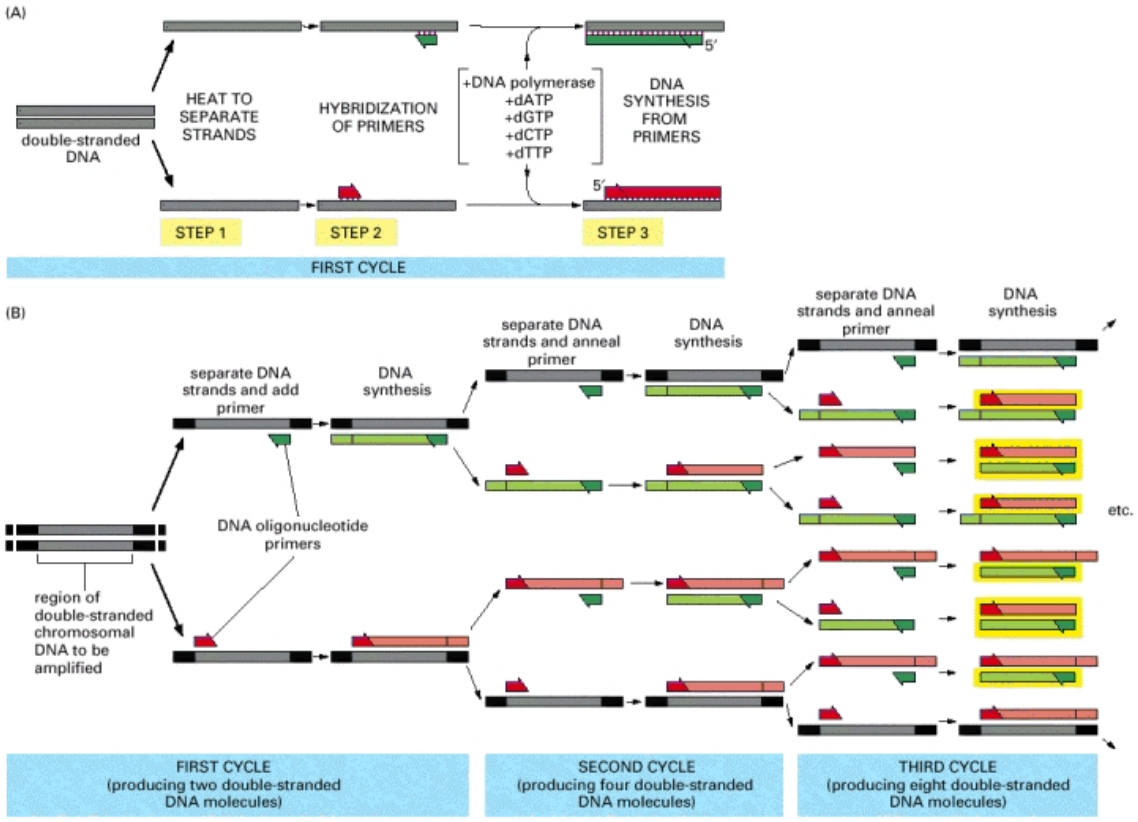


Figure 42. Amplification of DNA using the basic PCR technique. (A) Cycle 1, Step 1: Denaturation of the double-stranded DNA by heating. Step 2: Hybridization with primers targeting region of interest. Step 3: DNA synthesis activated by DNA polymerase. Monomers (dATP, dGTP, dCTP, dTTP) extend each primer to form two new daughter strands. (B) Cycle 2: Repeat of the thermal cycling yields 4 double-stranded DNA (dsDNA). Cycle 3 yields 8 dsDNA. Further cycling produces exponentially increasing copies of dsDNA. Source: MBoC.[39]

Quantitative PCR is basically the same as standard PCR but detects the amplified DNA as the reaction takes place. Fluorescent signals are detected in real-time as fluorescent-labeled gene-specific DNA probes hybridize with their complementary targets. The experiment will often target more than one gene; reference targets serve to normalize the target of interest. For example, Vinatzer *et al*.[38] normalized HER2 levels to

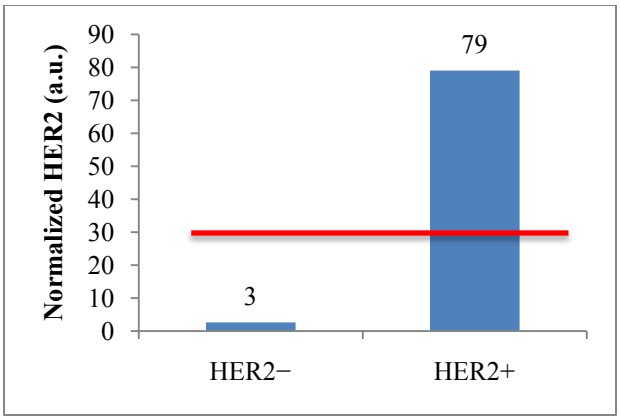


Figure 43. HER2 mRNA expression levels determined with quantitative RT-PCR in primary breast tumors of two patients: low relative expression of HER2 (HER2−) and high relative expression of HER2 (HER2+), as determined by comparison to a statistically determined threshold of 30 (red line). Results are in agreement with HER2 status via IHC and FISH. Source: Vinatzer *et al.*[38]

those of the β-actin protein to compensate for variability in the amount of cDNA starting material.

Figure 43 shows the normalized levels of HER2 obtained by quantitative RT-PCR from two different cancer patients. Statistical analysis was used to determine a cutoff level of 30 above which the sample would be scored as HER2 positive. This way, HER2 status can be determined automatically without the need for a specialist to examine multiple images and make subjective judgments. Overall, this method yielded comparable correlations of HER2 status with the overall survival of the patients. Thus, it could be a promising complement or alternative to current methods for HER2 testing, particularly in laboratories lacking FISH or IHC technology.

3.3.3.2 Fluorescent Microarray

DNA microarray technology, the workhorse of genomic medicine, provides a means to simultaneously monitor the RNA expression of thousands of genes. Schena *et al*. published its first use for gene expression profiling in 1995, utilizing high-speed robotic printing of DNA on glass.[43] The principle behind its operation is straightforward. An array of single-stranded probe DNAs are spotted on a chip and followed by a blocking process to prevent further attachment of nucleic acids in the non-spotted areas. Each element of the array contains a specific nucleotide sequence (chemically synthesized) that probes the sample solution for complementary sequences.

The probe DNA preferentially captures a sequence that is complementary to it, and the binding is highly specific since it is stabilized by Watson-Crick[55] base pairing in which adenine (A) only binds with uracil (U) for RNA or thymine (T) for DNA, and guanine (G) with cytosine (C). For example, the sequence 5'-GATTACA-3' will preferentially bind with its complementary RNA sequence, 3'-CUAAUGU-5' or DNA sequence 3'-CTAATGT-5'. The identity of any nucleotide fragment that hybridizes on the chip is thus known simply by detecting its location in the array.

Figure 44 depicts a cancer analysis based on a *comparative* gene expression microarray, which includes cDNA from healthy cells as a control measure. First, genes expressed as RNA are reverse transcribed and amplified using PCR. The resulting cDNA from the cells under investigation are tagged with a fluorescent red, while the cDNA from healthy control cells are tagged fluorescent green. A mixture of both of these labeled cDNA is then dropped onto all the elements of the microarray. If a certain gene is expressed more strongly of the cancerous cells, the corresponding spot in the array is

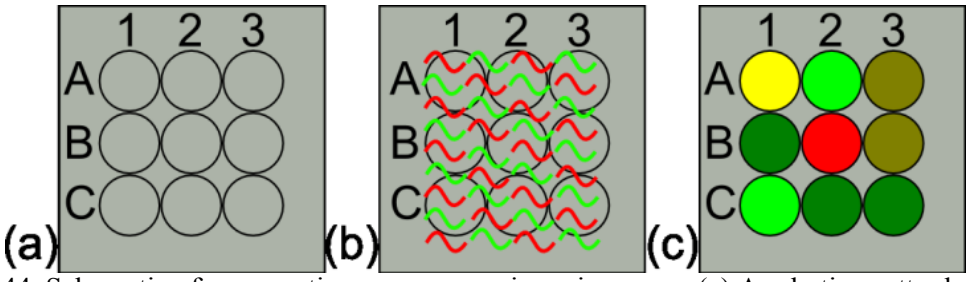


Figure 44. Schematic of comparative gene expression microarray. (a) A robotic spotter has deposited DNA probe sequences at each array position. These do not fluoresce because they are unlabeled. (b) The entire array is exposed to equal doses of cDNA from control (green) and cancerous (red) cells. (c) After a stringency wash, the remaining fluorescence signal indicates how DNA is expressed differently in cancerous cells. If a certain gene is expressed more strongly into RNA of the cancerous cells, the corresponding spot in the array is redder. Conversely, the spot is greener if the gene is expressed higher in the reference cells. Yellow indicates high expression in both control and cancerous cells.

redder. Conversely, the spot is greener if the gene is not expressed so much as in the reference cells. Thus one can use DNA microarrays to generate a profile of the genes that are selectively switched on or off to cause cancer.

As in the case of Q-RT-PCR, the mRNA of interest is likely at such low levels that it cannot be detected directly. In this case, RNA from the cells would first be reverse transcribed into cDNA, amplified by PCR, and tagged with fluorophore. In reverse transcription, the double-stranded cDNA contains the same sequence as the original RNA with U replaced by T. The positions in the array that fluoresce are recorded by a scanning-laser microscope and cross-referenced back to the known probe sequences.

A commercialized version of the microarray concept is the Affymetrix GeneChip which has an array of over 500,000 known probe DNA sequences synthesized on a 16 mm square glass surface via photolithography techniques. The Affymetrix Genome-

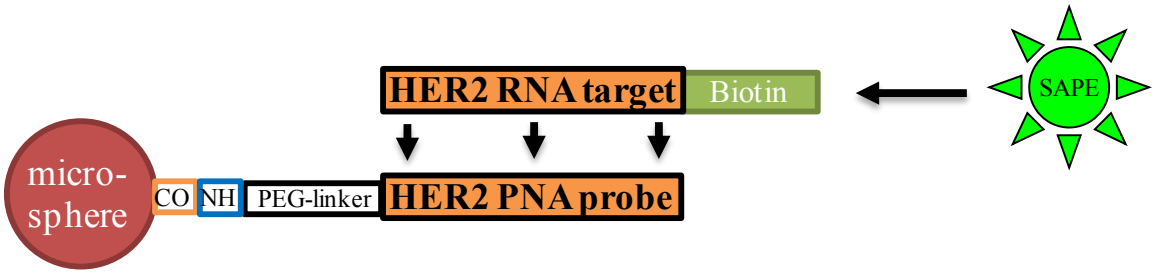


Figure 45. Schematic cartoon of Luminex principle of operation. A microsphere is functionalized with PNA probe via carboxyl-amine reaction using EDC. During hybridization, the HER2 RNA target (labeled with Biotin) forms a duplex with the HER2 PNA probe. As a final step, the biotin-labeled end of the RNA is stained with streptavidin-phycoerythrin (SAPE). Streptavidin has a high affinity for Biotin, while phycoerythrin is a fluorophore. During measurement, a laser scanner excites and reads the fluorescent signature of the bead and the intensity of the SAPE.

Wide Human SNP Array costs $250 per sample and is designed to search for single nucleotide polymorphisms (SNPs), changes in just one nucleotide that can significantly affect the risk of cancer or other disease.

The Luminex XMAP platform could be seen as an evolution of the microarray concept. In place of a monolithic glass substrate, the system color-codes polystyrene microspheres into 100 distinct sets, each functionalized with a unique probe (Figure 45). This allows multiplexing of several probes in one reaction tube. During the measurement phase, the probe on any bead is identified via correspondence with the fluorescent signature.

For the purpose of determining HER2 status, it is likely that less than 100 unique probes would be needed. In concept, the microarray idea could be scaled back to match the requirements at hand. Furthermore, auxiliary processes in the laboratory could be integrated into a single miniaturized system.

3.4 **Proposed Biosensor**

3.4.1 Lab-on-a-chip Integration

Q-RT-PCR and microarray are two existing lab methods to indirectly detect the overexpression of HER2 at the mRNA level. Our vision of determining a patient's HER2 status at the point-of-care calls for the integration of a series of steps that is usually performed manually. The process flow common to both methods is generalized in Figure 46, and involves a series of steps to extract RNA from tissue, amplify the genes of interest via PCR, and detect them via fluorescence. Many of these steps are labor-intensive and must be done in a laboratory with high-cost equipment, reagents, and highly trained personnel. To get this system out of the laboratory and into the field, we need to reduce the number of process steps, thereby also avoiding the cost of reagents and equipment.

Since DNA polymerase can act only on DNA templates, the mRNA must first be reverse-transcribed. If we can directly detect the HER2 mRNA, we could avoid the labor and costs of both reverse transcription and amplification. Direct detection of mRNA is also biologically relevant because mRNA reveals more or less the exact sequence of amino acids in the protein that will be synthesized from it. It follows that therapy is possible by blocking bad mRNA from being translated to disease-causing proteins.

By using CNTFET arrays to directly detect RNA via its intrinsic charge, we can also do away with fluorescent labeling. The remaining steps can be integrated and miniaturized through the use of microfluidic systems.[44] This resulting system could be highly automated, battery-powered, and contained in a package about the size of a mobile phone.

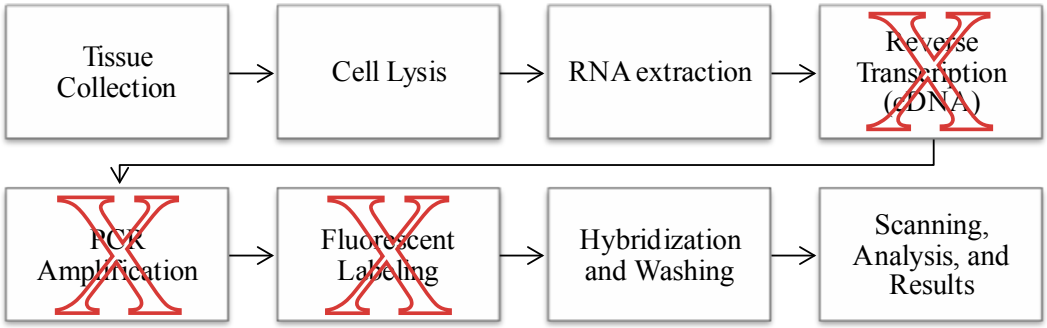


Figure 46. Gene expression assay process flow. After tissue collection, the cells are lysed, the RNA is extracted, and the RNA is reverse transcribed into cDNA. PCR amplification multiplies the copy number of genes of interest. On a microarray, the genes are labeled, hybridized to probes attached to glass, and washed. Finally, results are compiled from fluorescence image data. In Q-RT-PCR, fluorescent labeling is integrated with PCR primers to allow real-time measurement of gene levels. Reverse transcription, PCR amplification, and fluorescent labeling could be eliminated in a direct, label-free detection scheme.

We recognize a number of advantages to our envisioned method. Point-of-care testing means tissue will not need to be fixed or embedded in wax, which is known to cause unreliability in IHC. Avoiding PCR means avoiding unfair bias in replication due to the difference of hydrogen bonds between A-T (2 bonds) and G-C (3 bonds).

Despite its success, conventional microarray technology will not be the optimal platform for our application. Fluorophores are expensive ($300 for 1 µmole of oligonucleotides) and difficult to attach to RNA. The array-scanner itself, which detects the color intensity is priced more appropriately for research labs than clinics (the baseline Affymetrix GeneChip fluorescent array scanner costs over $200,000).

The new challenge we face is in developing a cost-effective RNA detection method with enough sensitivity to obviate the need for PCR, and enough selectivity to make reliable and repeatable diagnoses.

3.4.2 Field-Effect Detection of Nucleic Acids

At biologically relevant pH levels, DNA and RNA bear intrinsic electrical charges that could be detected by field-effect transistors. As shown in Figure 47, each nucleotide (nt) consists of a base (G, T, C, or A), a sugar ring, and a phosphate group. In RNA, the sugar ring is ribose, whereas the sugar ring in DNA is deoxyribose (ribose lacking an oxygen atom). Multiple nucleotides are linked together covalently via the oxygen atoms of the phosphate groups. Three of the oxygen atoms in a phosphate participate in two bonds. But one of the oxygen atoms participates in a single bond, leaving an unshared electron. This is the origin of the negatively charged character of the sugar-phosphate backbone.

The quest to detect charged molecules in aqueous solution via the field-effect transistor dates back to 1970, when Bergveld introduced the ion-sensitive field-effect transistor (ISFET).[45] The ISFET combined the principles of glass reference electrodes and the MOSFET to measure ion concentration. The introduction of silicon as a substrate for making such measurements opened the possibility of integrating the sensor with conventional CMOS circuitry such as amplifiers and signal conditioning. This approach has been so well-developed that they are the basis for Honeywell's line of Durafet Non-Glass pH Sensors.

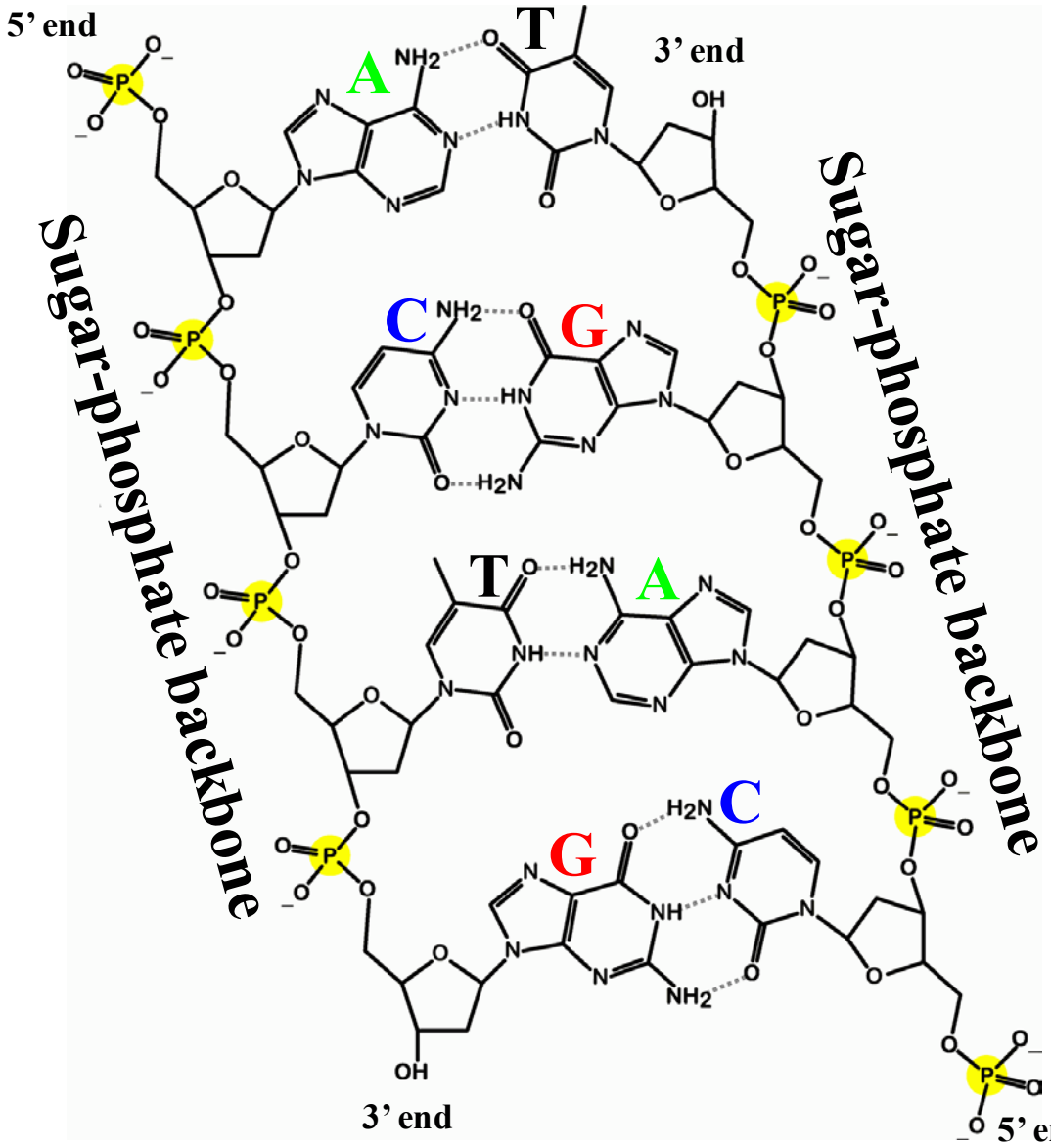


Figure 47. Chemical structure of the DNA molecule. Each base (G, T, C, or A) is covalently attached to one sugar (deoxyribose) and a phosphate group (PO^{3-}), comprising a nucleotide. The nucleotides are linked together covalently through the sugar-phosphate backbone, formed by the 3′-hydroxyl (-OH) group of one sugar and the 5′-phosphate (P) of the next. The phosphate group is the origin of the intrinsic charge of DNA and RNA. Based on: Wikipedia.org.[46]

A review of silicon-based chemical sensors making use of the field-effect concept was given in 2003 by Bergveld.[17] More recently, a summary of DNA or gene field-effect transistors, or "GenFETs", was compiled by Poghossian *et al*.[47] The first attempt in the direct detection of a hybridization event via a GenFET was reported by Souteyrand *et al*.[48] in 1997 in which they observed a shift in flat-band voltage of a FET (10 nm gate oxide) about 100 mV for the hybridization of ~2 μg/mL of 1000-mer poly(dA) to immobilized 1000-mer poly(dT). Fritz *et al*. in 2002 reached a hybridization detection sensitivity of 2 nM of 12-mer (8 ng/mL) DNA via differential measurements at 1 kHz on a capacitive sensor with 2 nm of thermally grown oxide.[49] Perkins *et al*. observed threshold shifts on the order of mV in response to the hybridization of 1 μM 15-mer DNA on depletion mode floating gate FETs,[50] while Pouthas *et al*. built a linear FET (10 nm oxide) array and observed a 10 mV response to the deposition (but not hybridization) of 5 μM of 20-mer single stranded DNA.[51] They also observed a 50 mV shift in response to the deposition of 20 μg/mL of double stranded DNA. Refinements by Ingebrandt *et al*.[52] allowed them to detect a signal amplitude of 4 mV using a chip-to-chip differential sensor readout for target concentrations of 1-3 μM.

Based upon the literature, the field-effect detection of nucleic acids is nothing new. However, the direct detection of mRNA will present some unique challenges.

3.4.3 State of the Art in Direct Detection of mRNA

It has already been shown that Q-RT-PCR can be used to detect the HER2 at the mRNA level, but indirectly. If we can avoid the reverse transcription and amplification altogether, we can avoid the associated labor and reagent costs. Direct detection of mRNA might yield better results because mRNA reveals the exact sequence of amino

acids in the protein that will be synthesized from it, whereas cDNA is known to contain biases and artifacts induced by PCR amplification.[53]

We face a number of challenges in direct detection of RNA. RNA is chemically unstable, due to the polar hydroxy group on the ribose sugar. RNA-digesting enzymes are ubiquitous and careful handling is required throughout experiment. The biggest challenge may be specifically detecting the very low levels of HER2 mRNA relative to all the other RNA in the cell. About twenty million strands of RNA are present, including messenger RNA, transfer RNA, ribosomal RNA, and others types. Detecting the HER2 mRNA, which might be present in just 50 strands, is like finding a needle in a haystack. In a real life situation, HER2 would be further diluted. In a biopsy sample taken from a patient, affected (cancer) cells are mixed with healthy cells and cells from other tissues. Also, hybridization could be hindered by the natural tendency of RNA to form loopy secondary structures.

3.4.4 Peptide Nucleic Acid (PNA) Probe Technology

In 1991, Nielsen *et al*. reported on the design and synthesis of peptide nucleic acid (PNA),[54] an artificial DNA mimic that follows Watson-Crick[55] base pairing with DNA and RNA but with high stability and sequence-selectivity. PNA molecules also exhibit excellent chemical stability and resistance to enzymatic digestion.[56]

These desirable traits are based on the fact that its pseudo-peptide backbone, substituted in place of the charged sugar-phosphate backbone, is electrically neutral (Figure 48). Therefore, PNA is highly suitable as a probe molecule since it does not electrostatically repel target RNA and will strongly maintain a duplex formation after hybridization.

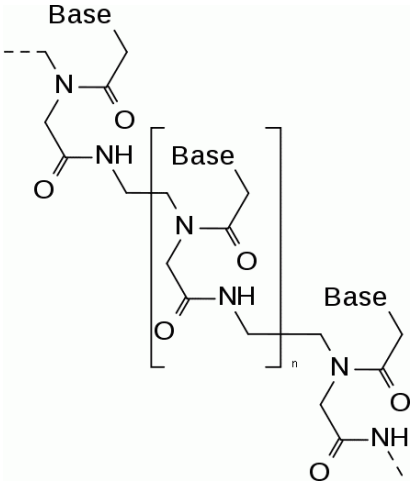


Figure 48. The backbone of PNA is composed of repeating N-(2-aminoethyl)-glycine units linked by peptide bonds. The bases are the same four bases found in DNA (A, C, G, T). Since the backbone of PNA contains no charged phosphate groups, the binding between PNA/RNA strands is stronger than between DNA/RNA strands due to the lack of electrostatic repulsion. Source: Wikipedia.[57]

The high specificity of PNA/RNA duplex formation means that a single base mismatch can be discriminated against better than a similar mismatch in a DNA/RNA duplex. The success of PNA as a probe for RNA is attested by the fact that a vast majority of published studies have focused on PNA oligonucleotides that target mRNA. In fact, it is also possible for PNA to form triplexes with double-stranded chromosomal DNA.[58] Hence, cell penetrating PNAs are being sought for gene therapy or silencing.

The application of PNA as probes for DNA hybridization arrays was demonstrated as early as 1997.[59] More recently, PNA has been used in microarrays by Jacob *et al*. in 2004[60] and 2006.[61] PNA has been used in the direct detection of DNA using nanowire sensors by Hahm and Lieber,[62] and on CNTFETs by Maehashi *et al*.[15]

In our application, a very important aspect of PNA is that it can hybridize with target RNA in the complete absence of salt in the hybridization buffer. This helps us

avoid contaminating the transistor gate oxide with mobile ionic contaminants, as well as allows for higher detection signals due to less screening. It also opens to the door to the possibility of real-time monitoring of the transistor thresholds as the hybridization reaction is occurring, which is not possible with DNA probes, which require high strength buffers that drown out the signal.

Furthermore, in a low salt environment, the target RNA is less likely to maintain secondary folding structures and more likely to be linearized and accessible to hybridization with the probe. This also allows us to avoid the fragmentation of long RNA before hybridization. Longer RNA could potentially increase the measured signal if the more remote tails do form secondary structure or otherwise lay flat on the transistor, bringing more charge within the Debye screening length.

Although PNA is relatively easy to synthesize, it is not widely commercially available for customized use and has seen limited application (i.e., the chromosome 17 probes in the *HER2* FISH pharmDx™ Kit). Since custom PNA probes are not commercially available, our collaborators at the NCI designed and synthesized PNA probes for this and other efforts such as targeted delivery and gene therapy.

3.5 **Results: Design, Synthesis, & Validation of HER2 probe**

In close collaboration with our NCI chemist, we produced multiple HER2 PNA probe sequences and validated their performance on a proven, commercially-available nucleic acid detection system. Bioinformatics specialists at the NCI designed a unique set of probe sequences to specifically capture HER2. The probes were then synthesized and purified using in-house peptide synthesis techniques. After the probes were validated

using an accepted molecular biology platform, we collaborated to develop and validate the probe attachment chemistry relevant to our biosensors.

3.5.1 Design of HER2 Probe Sequences

The design of HER2 probe PNA sequences was based on the existing bioinformatics techniques used to design DNA probes, but with modifications that basically account for the stickiness of PNA. Whereas the HER2 probe DNA in FISH may extend the entire HER2 gene (~41,000 base pairs), a PNA probe needs to be fairly short (16 nt) to avoid spontaneous aggregation due to poorer water solubility.

As a starting point for probe design, our molecular biologist at NCI used the HER2 probe sequences from the Affymetrix gene chip, which has 25-mer DNA probes. (Since the DNA-DNA binding is relatively weak, duplexes shorter than 25-mer may simply melt at room temperature.) Using these 25-mer windows to the HER2 gene, a protein sequence database (PSD) and other bioinformatics methods were used to search for 16-mer PNA sequences within these windows with high melting temperature.[63] Candidate sequences were checked against specific design rules that ensure solubility in aqueous solutions and avoid self-complementary.

Table 4 contains the sequence and relevant calculations for the HER2 PNA probe sequence selected for use on the biosensor. The melting temperature T_m is the temperature at which half the population of PNA-DNA duplexes are denatured. T_m is important because the duplexes melt at lower temperatures when the buffer has less ionic screening. However, higher temperatures may be needed for high specificity. The sequence must also satisfy the following rules on purine content: (a) the purine (adenine, guanine) content must be ≤ 60%, (b) the number of any purines in a row must be ≤ 4, and

Table 4. Sequence and calculation results from Applied Biosystems, Inc. (ABI) for the anti-sense HER2 PNA probe suitable for synthesis. Anti-sense polarity means the probe can detect mRNA directly, as opposed to a sense probe for detecting cDNA from reverse-transcription. The melting temperature (T_m) decreases with lower ionic strength. The probe modification (mod) is attached to the 5' end.

HER2 probe	5'-(mod)-ACT GGA CCC TAG AGT C-3'	
Complementary target	3'-TGA CCT GGG ATC TCA G-5'	
T_m ([1 μM])	76 °C	
T_m ([1 nM])	64 °C	
T_m ([1 pM])	53 °C	
Molecular weight	4369.8 g/mol	
% purine	50.0%	Limit: 60%
Longest purine stretch	4 (AGAG)	Limit: 4 mixed A/G, 3 G in-a-row

(c) the number of any guanines in a row ≤ 3. The HER2 probe sequence shown obeys these rules and is suitable for synthesis.

A chemical drawing of the probe is shown in Figure 49. The modification of this particular version of the probe is for surface attachment via an acrylic-sulfur bond. The 5' acrylic group is coupled to the main PNA sequence via a spacer that increases the solubility of the entire molecule. The acrylic group can be replaced with the appropriate functional group to match the groups on the attachment surface. The subject of attachment chemistry will be discussed next.

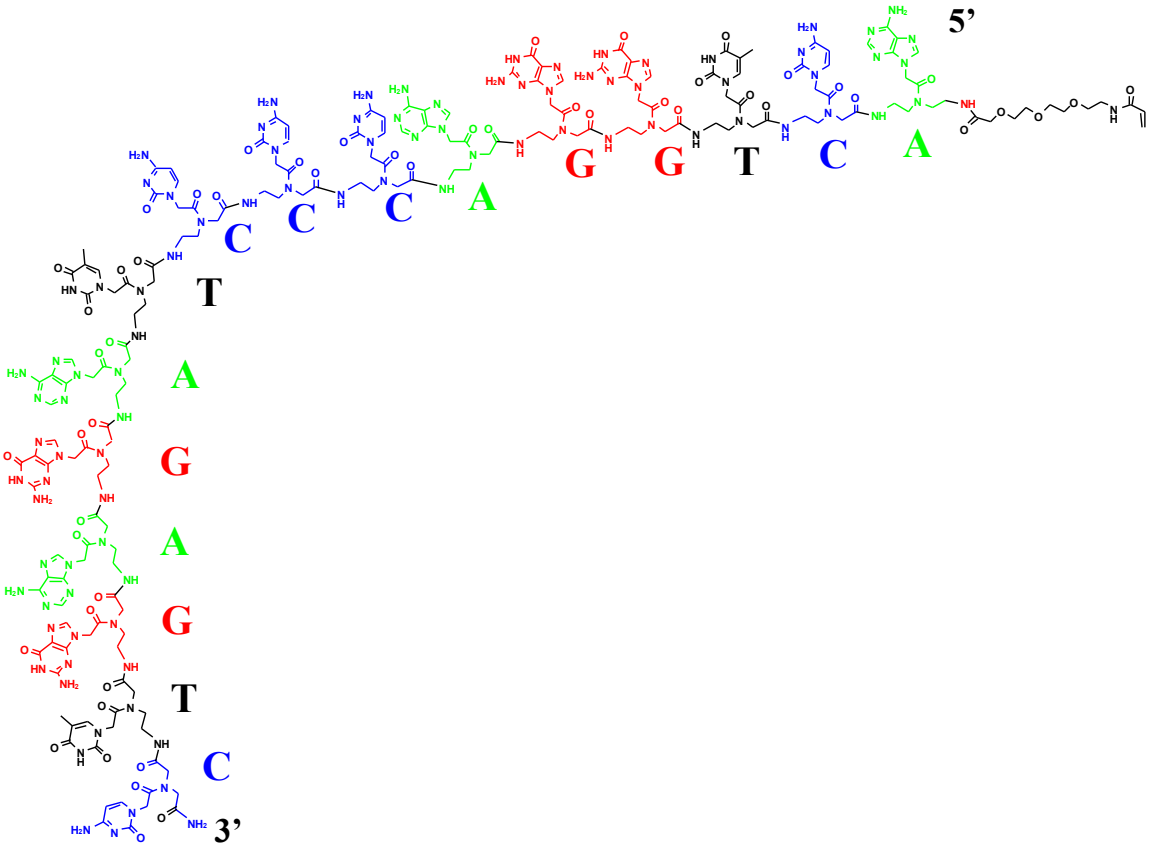


Figure 49. Chemical drawing of acrylic-modified HER2 anti-sense PNA probe. 5' acrylic group is on the upper right. The 16-mer PNA sequence is linked to the acrylic group by an 11-atom PEGylated (polyethylene glycol) spacer. The distance created by spacer increases the accessibility of the acrylic group for immobilization and the nucleotides for hybridization. The use of PEG groups is to increase water solubility.

3.5.2 Attachment Chemistry

A variety of options exist for attaching nucleic acid probes to surfaces. The choice of attachment chemistry involves determining what reactive group will be used to modify the probe and what surface groups must be available for reaction. Amino, thiol, and acrylamide are among the routinely used modifications in conventional microarrays. In

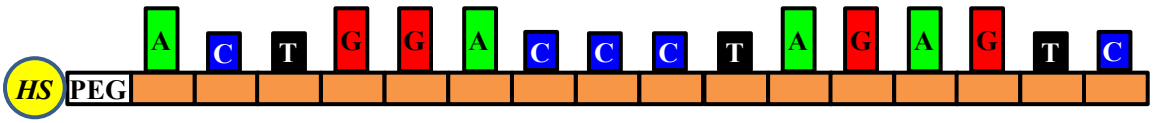


Figure 50. Cartoon drawing of thiol (HS) modification, PEG spacer, and PNA sequence. Thiol group is intended for direct attachment to gold (or other noble metal) via dative bond.

these cases, the issues affecting the decision revolve around fluorescence performance (scattering, background), as well as chemical stability and loading capacity.

In our case, we wanted a simple probe attachment strategy that would also maximize hybridization loading capacity. Given the choice of gold CNTFET gates, it followed that two options existed: functionalization of the gold surface with a self-assembled monolayer (SAM) followed by covalent attachment reaction, or direct attachment of thiol-modified probe to the gold.

The first option was used by Maehashi *et al.*[15] to attach PNA probes to the back gate of CNTFETs. It involves incubating the bare gold surface in a solution of 3-mercaptopropionic acid to form the intermediate SAM, then activation of carboxylic acid groups using an aggressive coupling agent (ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride, or EDC). After activation of the carboxy function with N-hydroxysuccinimide (NHS), the 5'-amino-modified PNA may then be attached.

Ultimately we adopted the simpler option and synthesized PNA with 5'-thiol modifications (Figure 50). Thus we avoided the difficulties we encountered with the moisture sensitivity of EDC and could run the attachment reaction in the mild conditions of DI water. This attachment strategy was used by Steel *et al.*[64,65], who achieved nearly complete target hybridization levels on gold surfaces. They also were able to enhance the accessibility of the probes for hybridization by blocking unreacted surface sites with 6-

mercapto-1-hexanol (MCH). However, we employed the basic sulfur-gold chemistry and will defer the exploration of blocking to a later optimization phase.

3.5.3 Synthesis & Purification

The PNA was synthesized via Fmoc chemistry solid phase peptide synthesis[66] and purified using high-performance liquid chromatography (HPLC). In short, peptides are synthesized by coupling the carboxyl group of one amino acid to the amino group of another. Protection groups based on 9-fluorenylmethyloxycarbonyl (Fmoc) prevent unwanted side reactions on the developing oligomer until which time the next nucleotide should be appended.

The ABI 433A Peptide Synthesizer was used to make the crude probe PNA, which will contain a mixture of the desired product and other unwanted fragments. The Agilent 1200 series HPLC was used to isolate the desired product (Figure 51). As compared to automated DNA synthesis, peptide synthesis is more difficult. Thus, the purification step here is more important than for DNA oligos, since peptide synthesis often does not perform at high efficiency. Matrix-assisted laser desorption/ionization (MALDI) was used to measure the distribution of molecular weights of compounds within the purified product. The MALDI spectrum of successful PNA synthesis shows clear peaks corresponding to the expected molecular weight of the synthesis product.

Figure 51. (left) Applied Biosystems ABI 433A Peptide Synthesizer used to make probe PNA. (right) Agilent 1200 series HPLC used for purification of crude PNA.

3.5.4 Validation

We collaborated with NCI to verify the performance of the exact HER2 PNA sequence that we would be immobilizing on our biosensors. Validation was performed on two different fluorescent-based systems: (1) the commercially available Luminex XMAP platform, and (2) plain gold surfaces. Luminex was used to generate quantitative calibration curves (intensity vs. target concentration) to find the optimal temperature for specificity. The plain gold chips were then used to qualitatively demonstrate successful hybridization and specificity of the same probe sequence but different attachment chemistry.

Table 5. HER2 “sense” probe and synthetic target sequences used in Luminex experiment. The sense polarity probe is designed to detect cDNA from reverse transcription-PCR. Mutated nucleotides are underlined.

Probe	5'-/amine/GAC TCT AGG GTC CAG T-3'
Wild type (wt)	5’-/biotin/ACU GGA CCC UAG AGU C-3’
Mutant 1 (mt1)	5’-/biotin/ACU GGA C<u>U</u>C UAG AGU C-3’
Mutant 2 (mt2)	5’-/biotin/ACU GGA C<u>U</u>C UAG <u>C</u>GU C-3’

3.5.4.1 Validation via Luminex XMAP

The Luminex experiment involves immobilizing probes to the Luminex microspheres (beads), then hybridizing them at 55 °C with synthetic, biotin-modified RNA. After hybridization, the biotin-RNA is stained with streptavidin-phycoerythrin (SAPE) fluorophore; streptavidin binds strongly with biotin. During measurement, a laser scanner reads the fluorescent intensity from the SAPE bound to the RNA. Since hybridization occurs in a 96-well plate, it is convenient to test several concentrations and different targets in parallel.

As shown in Table 5, a sense-polarity probe and three different sequences of biotin-labeled synthetic RNA were used: fully complementary wild-type (wt), single nucleotide mutation (mt1), and two nucleotide mutations (mt2). The sense-polarity probe is intended for capturing HER2 in PCR-amplified cDNA that is reverse transcribed from cellular mRNA. The mutant targets constitute single and double nucleotide polymorphisms and are included to test the specificity of sequence recognition.

Table 6. Summary of Luminex XMAP protocol for attachment of probes to Luminex microspheres, hybridization with synthetic DNA, and fluorescent staining.

Probe Immobilization
1. Dilute one million beads (10^6) in 10 µL MES buffer, pH 4.5.
2. Add 1 µL of NH_2-modified PNA probe (100 pmol).
3. Add EDC (5 µg, 2.5 µL) and 30 min at room temp (repeat once).
4. Wash the conjugated beads with 0.02% Tween20 and 0.1% SDS.
5. Resuspend in 20 µL TE pH 8.0.
6. Count the beads under a hemacytometer (~5,000 beads/µL).
Hybridization
1. Dilute the conjugated bead (5000 beads) in 33 µL 1.5x TMAC per well.
2. Add biotin-labeled target DNA/RNA in 17 µL TE.
3. Denature at 95 °C for 2 min, then hybridize at 55 °C for 1 hour.
4. Wash the bead with 1 mL 0.06x SSC buffer and recover the beads by centrifugation.
5. Add 75 µL SAPE (3 µg/ml) in 1x TMAC and 45 °C 10 min.
6. Measure fluorescence intensity on Luminex-100.

The protocols for probe immobilization and target hybridization are summarized in Table 6. The microspheres are pre-functionalized with carboxylic acid groups (COOH). The attachment chemistry is similar to the EDC/NHS method used by Maehashi *et al*.[15] The 5'-amino-modified PNA is directly mixed with the beads while the acid groups are activated with EDC in acidic conditions (pH 4.5). The conjugated beads are then washed in detergents and resuspended in TE storage buffer.

Hybridization was at 55 °C in 50 µL of buffer based on Tris-EDTA, followed by SAPE staining. No explicit washing step was incorporated because the measurement procedure further dilutes the beads and the unbound SAPE, producing a negligible background.

Calibration curves (fluorescent intensity vs. synthetic RNA concentration) are shown in Figure 52. As expected, each curve shows an increasing signal with increasing

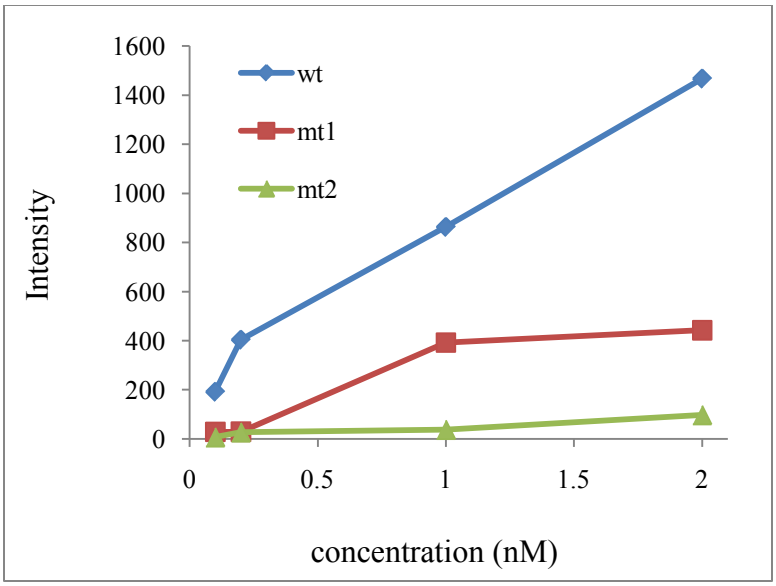


Figure 52. Luminex calibration curves (fluorescent intensity vs. synthetic RNA concentration) of HER2 PNA probe hybridized with synthetic RNA of three different sequences: wild-type (wt), single nt mutation (mt1), and two-nt mutation (mt2). The fluorescence signal originates from streptavidin-phycoerythrin (SAPE), which binds with biotin on the labeled RNA. Hybridization is at 55 °C in 50 µL Tris-EDTA buffer containing trimethylammonium chloride (TMAC). Each curve shows an increasing signal with increasing concentration. The fully complementary wild-type sequence produces the highest signal overall, while increasing the number of mutations caused decreased signals. At low concentrations (<0.25 nM), the PNA probe cannot discriminate between one mutation and two mutations in the target. The lowest concentration used was 0.1 nM. (Data courtesy of Dr. Young Song, NCI. To be published elsewhere).

RNA concentration. The probe successfully discriminates between RNA that is fully complementary and RNA that has point mutations. The fully complementary wild-type sequence produces the highest signal overall, while increasing the number of mutations caused decreased signals. At concentrations below 0.25 nM, the PNA probe cannot discriminate between one mutation and two mutations in the target. Therefore this number is claimed as the sensitivity limit.

Table 7. Summary of protocol for attachment of probes to test surface, hybridization with synthetic DNA, and fluorescent staining.

Probe Immobilization
1. Heat PNA to 55 °C for 2 mins. Agitate.
2. Apply spots of 1 mM PNA solution to chip.
3. Hold overnight in humidified enclosure.
4. Rinse chip with TGB gently, decant, and blow dry.
5. Mount to chip carrier. Wirebond. Mount cups using silicone adhesive.
6. Cure adhesive overnight.
7. Dilute stock DNA in hybridization buffer. Make 25 µL per cup. Vortex and apply.
Hybridization
8. Hybridize in sealed chamber at 55 °C for 3 hours with humidity.
9. Preheat blank hybridization buffer to 55 °C. Rinse cups twice.
10. Remove cups and silicone residue.
11. Mix 2.5 µL SAPE in 245 µL TMAC. Apply under a cover slip, wait 15 mins.
12. Rinse gently with 0.1X SSarc. Remove as much as possible.
13. Image chip on Typhoon 9410 with Cy3 setting.

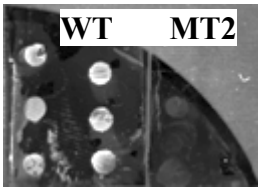
For our application, it may not be necessary to detect single nucleotide mutations. In a future design, several HER2 probes would be used and the determination of HER2 status would be based on signals over the aggregate. Thus the overall selectivity is more important than the selectivity with respect to a single nucleotide.

3.5.4.2 Verification on Gold Test Surfaces

The attachment chemistry and type of surface may affect the hybridization performance of the probe. We performed experiments using the exact probe on surfaces similar to that which comprises the CNTFET metal gates. Table 7 summarizes the protocols for probe attachment and target hybridization. The thiolated PNA probe was attached to the gold surface simply by incubation. The hybridization was carried out with biotinylated synthetic DNA targets at 55 °C. As in Luminex, visualization is

accomplished via a final staining step (SAPE). A flatbed laser scanner was used to image the results.

The fluorescent images in Figure 53 show the successful immobilization and hybridization of the HER2 probe on gold surfaces. First, a piece of gold-coated wafer was split into two chips. One of the chips was hybridized with wild type DNA (WT) and produced bright spots only at the areas where the probe was spotted. The other chip was hybridized with mutant (MT2) and produced only dim signals. A third chip was patterned with PMGI stripes to check if it intereferes with immobilization or hybridization. The process is not hampered by the presence of the PMGI protective layer. Furthermore, absence of signal over the PMGI stripes indicates that probe does not adsorb to the PMGI.

(a) 

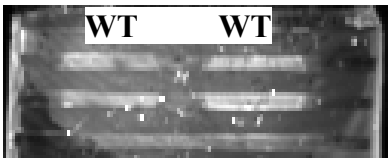
(b) 

Figure 53. Representative fluorescent image of synthetic DNA targets hybridized with thiolated HER2 PNA probe immobilized on gold coated wafer. The targets were labeled with biotin and stained with SAPE, then imaged by Typhoon 9410 Variable Mode Imager. (a) A quarter-circular piece of wafer was split into two chips. The entire left chip was hybridized with wild type DNA (WT) and produced bright spots only at the areas where the probe was spotted. The right chip was hybridized with mutant (MT2) and produced only dim signals. (b) Another gold chip was patterned with PMGI stripes and attached to two plastic tubular cups. The probe was spotted inside the cups in the bright circular areas. Both spots were hybridized to wild type (WT). The absence of signal from where the PMGI is indicates that the probe was not adsorbed to the PMGI.

3.6 **Conclusion**

We envision a lab-on-a-chip integrated system for the detection of cancer. For this iteration of platform development, we specifically focused on the HER2 breast cancer gene and designed, synthesized, and verified custom PNA probes. PNA imparts the advantages of high sensitivity and high selectivity under low ion conditions, traits important to the success of a reliable diagnostic system. We hope that this system, which detects HER2 at the mRNA level, could be used adjacent or in competition with existing HER2 status methods based on fluorescence and or antibody staining. In the next chapter, we will show that the prototype system was able to detect overexpression of HER2 in real breast cancer cells.

Chapter 4: Theory of Operation

4.1 Chapter Purpose

This chapter seeks to develop the theory of operation of the CNTFET biosensor for the purpose of estimating how much DNA or RNA must be captured to produce a measurable signal, as well as what fundamental limitations might govern the maximum signal that could be generated. The double-layer theory, based upon the laws of drift and diffusion, will be used to explain the operation and purpose of all the electrodes that are immersed in the top-gate liquid: the counter electrode, reference electrode, and gate

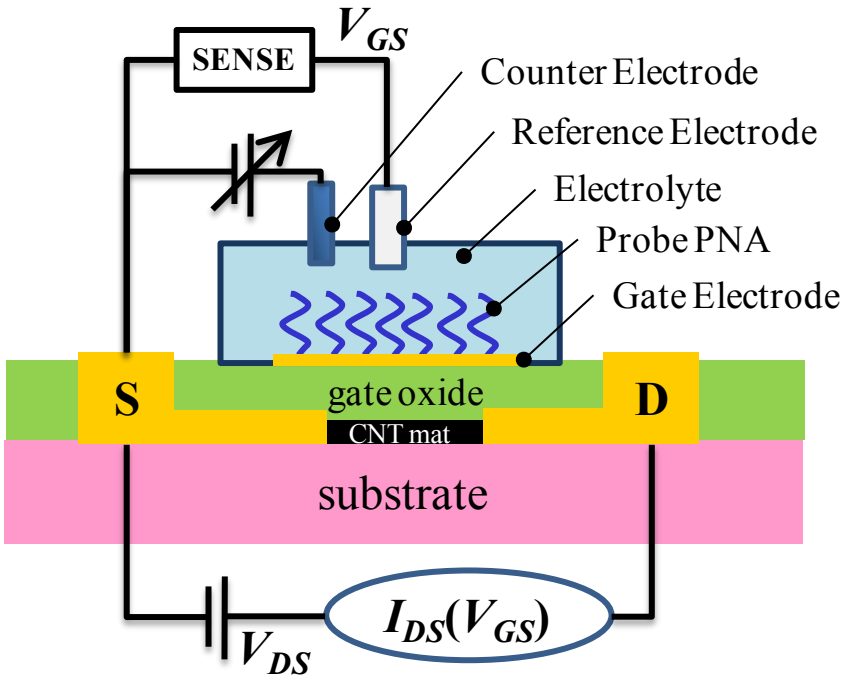


Figure 54. Schematic diagram of a liquid gated CNTFET, including counter, reference, and gate electrodes, all in contact with the top-gate electrolyte. The electrolyte potential is controlled by the counter electrode, measured by the reference electrode, and is capacitively coupled to the gate electrode. The probe PNA is responsible for the specific recognition of DNA or RNA.

electrode (see Figure 54). The double-layer model will be combined with basic electrostatic equations to arrive at the Grahame equation, which gives the relationship between surface charge and measurable voltages in solution. We will see that this relationship is linear for low densities of nucleic acid, but is moderated at higher densities. Finally, we will use the linear relationship to generate a circuit model to gauge the sensitivity of nucleic acid detection.

4.2 Overarching Theory of Electric Charges in Solution

The principle of operation of the CNTFET biosensor is derived from basic theory that describes the distribution of charges when a solid electrode is immersed in a liquid electrolyte. The arrangement of the charges is referred to as a "double layer".

Suppose a negatively charged electrode is immersed in a liquid electrolyte (see Figure 55). The electrolyte consists of dissolved positive and negative ions (for example, Na^+ and Cl^- in pure water). The charges on the electrode will be counterbalanced by a cloud of charge in the liquid called the "diffuse layer". The surface charges and the diffuse layer comprise a "double layer".

Near the surface of the solid, the diffuse layer is mostly positive ions (attracted by the negative solid). However, this dense concentration of positive charge will tend to attract a somewhat lesser dense layer of negative ions from the liquid. This layer of negative charge will then attract a less dense layer of positive ions. Further out into the fluid bulk, this "layering" will continue but with decreased intensity until some distance where the electrolyte is practically neutral (bulk region).

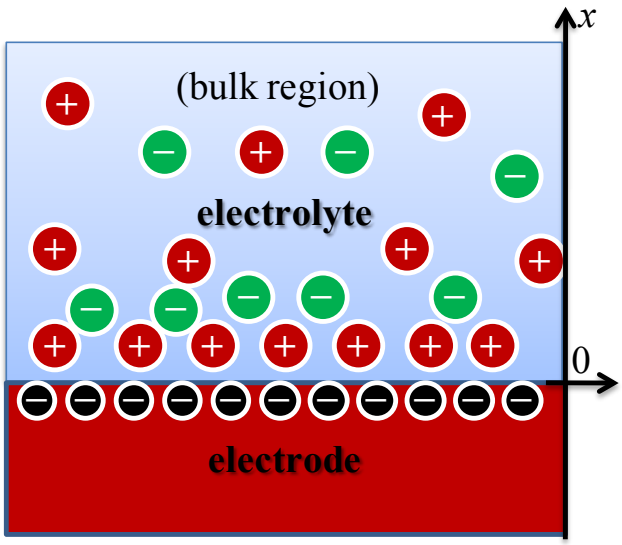


Figure 55. Simplified model of charges in an electrochemical half-cell, consisting of the electrolyte, electrode, and their "double-layer" interface ($x = 0$). The bulk region of the electrolyte contains an equal density of positive and negative ions. The electrons on the electrode surface are due to the ionization of atoms on the metal electrode and their subsequent dissolution. The charges on the electrode will be counterbalanced by a cloud of charge in the liquid called the "diffuse layer". The double-layer is comprised of the surface charges and the diffuse layer.

The structure of the diffuse layer is a consequence of two phenomena: drift and diffusion. Drift current depends on electric field and follows Ohm's Law: $V = IR$, where V is the voltage across resistor R with current I flowing through it. Diffusion is the entropic spreading out of particles due to thermal agitation.

4.2.1 Fick's First Law of Diffusion

The diffusion current is directly proportional to the concentration gradient. This is the basis for Fick's First Law of Diffusion:

$$J_D = -qD\frac{dn}{dx} \tag{9}$$

where J_D is the diffusion current density (A/m^2), q is the charge per particle (C), D is the diffusion constant (m^2/s), n is the local concentration (number density) of the charges (1/m^3), and x is the location (m). The negative sign means that the flux of particles is opposite to any increase in particle density. (The diffusion constant D must be positive) The total current density due to drift and diffusion is $J = J_E + J_D$, where J_E is the drift current density (A/m^2).

4.2.2 Ohm's Law & Diffusion

The drift current is directly proportional due to the applied electric field. Thus, the relevant form of Ohm's Law is

$$J_E = \sigma E \tag{10}$$

where J_E is the drift current density (A/m^2), σ is the conductivity (S/m), and E is the electric field (V/m). To combine Fick's Law and Ohm's Law, we seek a way to relate the conductivity to the diffusion constant. In the absence of any "drag" (scattering mechanisms), an electric field accelerates charges to infinite velocities. In reality, charges accelerate only until they have reached a "drift velocity". Thus, a microscopic view of Ohm's Law is expressed as $J_E = qnv$, where v is the drift velocity. The drift velocity is directly proportional to the electric field: $v = \mu E$, where μ is the mobility. (Note: as used here, the mobility is negative for negative charges).

For a system in thermal equilibrium that obeys Boltzmann statistics, the diffusion coefficient is related to the mobility by the Einstein relation: $D = \mu\,(kT/q)$, where k is the

Boltzmann constant (1.38×10^{-23} J/K) and T is temperature (K). We can rewrite the Einstein relation as $\mu = qD/kT$. Thus, Ohm's Law can be expressed as:

$$J_E = nq\mu E = nq\frac{D}{(kT/q)}E \tag{11}$$

Now we can combine drift and diffusion into one equation via superposition:

$$J = J_D + J_E = -qD\frac{dn}{dx} + nq\frac{D}{(kT/q)}E \tag{12}$$

4.2.3 Nernst-Planck Formula

We rewrite (12) to arrive at the Nernst-Planck Formula, which combines the effects of diffusion and drift:

$$J = D\left(-q\frac{dn}{dx} + \frac{nq}{(kT/q)}E\right) \tag{13}$$

In equilibrium, $J = 0$, so it simplifies to:

$$\frac{dn}{dx} = \frac{qE(x)n(x)}{kT}, \tag{14}$$

which is rearranged as:

$$\frac{dn}{dx}\frac{1}{n(x)} = \frac{qE(x)}{kT}. \tag{15}$$

Via the rules of integration, we rewrite:

$$\frac{d}{dx}(\ln n(x)) = \frac{qE(x)}{kT}. \tag{16}$$

By definition, $E(x) = -\frac{d\phi}{dx}$, where φ is the potential. Thus

$$\frac{d}{dx}\left(\ln n(x)\right) = -\frac{q}{kT}\frac{d\phi}{dx} \quad . \tag{17}$$

Now integrate both sides with respect to x:

$$\int_{x'}^{bulk} \frac{d}{dx}\left(\ln n(x)\right)dx = -\frac{q}{kT}\int_{x'}^{bulk} \frac{d\phi}{dx} \quad . \tag{18}$$

In the bulk electrolyte (for large x), the electrolyte ionic density is n_0 and the potential is zero ($\varphi = 0$). Therefore, replacing x' with x, we arrive at the **Nernst relation**:

$$\ln n(x) - \ln n_0 = -\frac{\phi(x)}{kT/q} \quad . \tag{19}$$

4.2.4 Nernst Relations

For simplicity, we will assume that the electrolyte is composed of singly ionized species such as Na^+ ($q = +q_e$) and Cl^- ($q = -q_e$), where q_e is the elementary charge ($+1.609 \times 10^{-19}$ C). We write separate Nernst relations for each polarity of charge:

$$\ln p(x) - \ln n_0 = -\frac{\phi(x)}{U_T} \quad , \tag{20}$$

and

$$\ln n(x) - \ln n_0 = \frac{\phi(x)}{U_T} \quad , \tag{21}$$

where $p(x)$ and $n(x)$ are the number densities of positive and negative ions, respectively, and the "thermal voltage" is $U_T = kT/q_e$ (~26 mV at room temperature). We also note that the bulk electrolyte contains an equal density of positive and negative ions, so it is not necessary to define $p_0 = n_0$.

We can now make the following observations. The density of positive ions is highest (and the density of negative ions is lowest) near the negatively charged electrode. This is expected, since opposite charges attract and like charges repel.

4.2.5 Boltzmann Distributions

To proceed further, we exponentiate the Nernst relations to arrive at the following Boltzmann distributions:

$$\begin{aligned} p(x) &= n_0 \mathrm{e}^{-\phi(x)/U_T} \\ n(x) &= n_0 \mathrm{e}^{+\phi(x)/U_T}, \end{aligned} \tag{22}$$

(these may look familiar in study of semiconductor diodes). This result shows that concentration at any point depends exponentially on the electric potential. Suppose the electrode potential is at −100 mV relative to the electrolyte. Since $\exp(-(-100\text{ mV})/26\text{ mV}) = \sim 47$, the concentration of positive ions at the electrode is about 50 times that of the bulk ($p = 47\ n_0$), whereas the concentration of negative ions is about 2% that of the bulk ($n = n_0/47$).

4.3 Electrodes

Now that we have presented the basic theories concerning the electrochemistry, we will use them to gain an understanding of how the electronic biosensor works. In particular, we will explain the operation and purpose of all the electrodes that are in direct contact with the top-gate liquid: the counter electrode and reference electrode (Figure 56). An electrode and the electrolyte around it form a basic electrochemical half-cell.

4.3.1 Electrochemical Half-cell

A simple half-cell is comprised of a metal electrode immersed in an electrolyte, an aqueous solution of the metal's salt (see Figure 55). We are familiar with half-cells as they comprise the common alkaline battery. The chemical reaction occurring at either pole is either oxidative or reductive. A typical anodic (oxidation) reaction involves a metal atom from the electrode being oxidized and dissolved into the electrolyte. The reaction can be written as: $M \rightleftharpoons M^+ + e^-$, where M is a metal atom, M^+ is the metal ion, and e^- is an electron.

The newly freed electron stripped from the metal atom remains on the electrode,

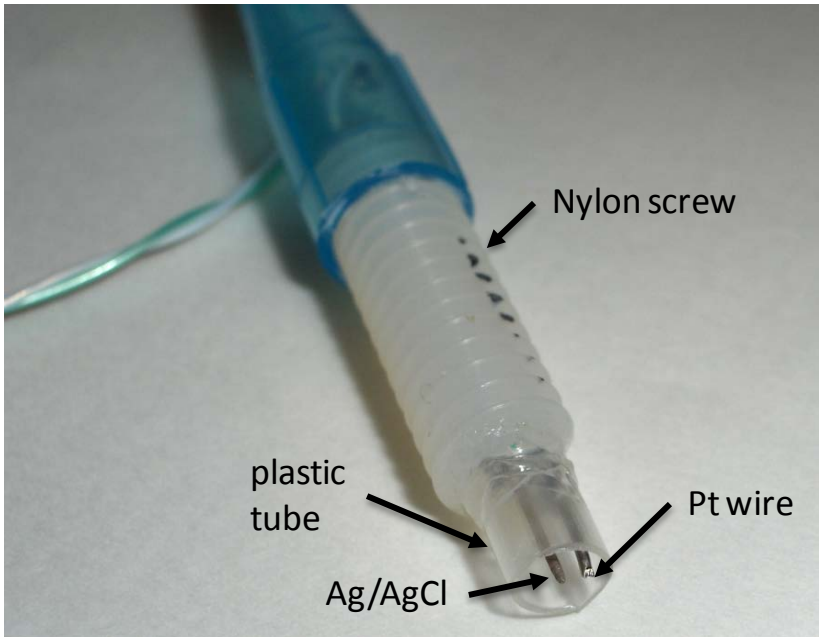


Figure 56. Top-gate electrode combination. A glass capillary tube surrounds the silver-silver chloride reference electrode. The counter electrode is a coiled platinum (Pt) wire. They are both mounted in a nylon screw. A plastic tube serves to hold a small amount of top gate buffer to keep the electrodes continually immersed. The wires travel through the screw and protrude out to the left.

which acquires negative charge. The newly ionized metal atom contributes positive charge to the electrolyte. The reaction proceeds forward until the buildup of electric potential no longer favors the generation of metal ions. In accordance with the theory presented earlier in this chapter, a double layer of charge forms at the interface between the electrode and electrolyte. This equilibrium is reached almost instantly in an isolated half-cell.

The Nernst relations as presented earlier are applicable to the half-cell. If we only consider the negative ions, the relevant equation is:

$$\ln n(x) - \ln n_0 = \frac{\phi(x)}{U_T} \quad . \tag{23}$$

Let us rearrange the Nernst relation as:

$$\phi(x) = U_T \ln\left(\frac{n(x)}{n_0}\right) \quad . \tag{24}$$

The potential of the electrode with respect to the electrolyte is given by:

$$\phi_{electrode} = U_T \ln\left(\frac{n_{electrode}}{n_0}\right) = U_T \ln(n_{electrode}) - U_T \ln(n_0) \quad , \tag{25}$$

where $n_{electrode}$ is the negative ion density just outside the electrode.

Note that negative ions are less dense near the negative electrode ($n_{electrode} < n_0$) due to electrostatic repulsion. As expected, the Nernst relation yields an electrode potential that is negative with respect to the electrolyte ($\varphi_{electrode} < 0$).

In practice, we want to know the potential of the electrode with respect to other electrodes in the electrolyte. Also, $n_{electrode}$ may not be easily determined. Thus, a more practical form of the Nernst equation for the chemical reaction is written as

$$E_N = E^0 - U_T \ln(c_0), \tag{26}$$

where E_N is the electrode potential, E^0 is the standard electrode potential for the reaction, and c_0 is the ionic strength in molar. E_N and E^0 are both taken with reference to a standard electrode (not the electrolyte). Values of E^0 have been tabulated for many electrodes through measurement in a standard half cell, usually with reference to the standard hydrogen electrode. A standard half cell consists of a metal electrode in a 1 M aqueous solution of the metal's salt at standard temperature (298 K, or 25 °C).

Note that the theory presented is developed assuming equilibrium conditions, with zero net current flowing into the electrode. Practical electrode systems can be modeled using two extreme cases: ideal non-polarizable electrodes and ideal polarized electrodes. As shown in Figure 57, a non-polarizable electrode has a well-defined potential given by the Nernst relation and independent of current flow, whereas an ideal polarized electrode permits no current flow at any potential. In reality, a practical electrode will be some combination of the two theoretical extremes.

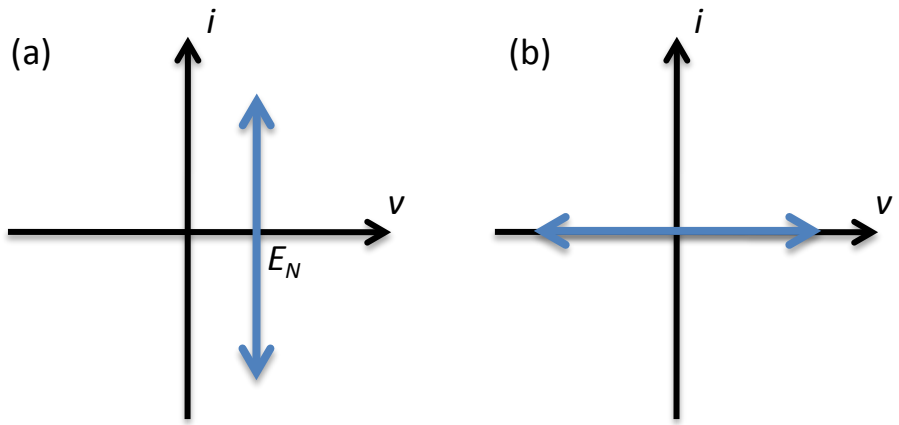


Figure 57. I-V curves of an ideal non-polarizable electrode and ideal polarized electrode. (a) The non-polarizable electrode has a well-defined potential given by the Nernst relation and independent of current flow. (b) The ideal polarized electrode permits no current flow at any potential.

4.3.2 Ag/AgCl Reference Electrode

In the biosensor system, a reference electrode senses the potential of the top-gate electrolyte (TGE) while charge is driven by a platinum (Pt) counter electrode. Thus, the reference electrode should have a stable and well-known potential with respect to the TGE. In addition, this electrode need not conduct significant current because it will be part of a high impedance circuit. Thus we seek an electrode that has the character of an ideal non-polarizable electrode under low current conditions.

The silver-silver chloride electrode (Ag/AgCl) is used commonly as a reference electrode in pH sensors due to its low cost and good performance (see Polk2006).[67] It is composed of a mixture of silver metal (Ag) and its chloride salt (AgCl). Once immersed in solution, the following chemical equilibrium reactions occur: $Ag^0(s) \leftrightarrow Ag^+ + e^-$, and

$AgCl(s) \leftrightarrow Ag^+ + Cl^-$. As in our simple model of the half-cell, some portion of the electrode dissolves while electrons are left behind on the electrode. However, since

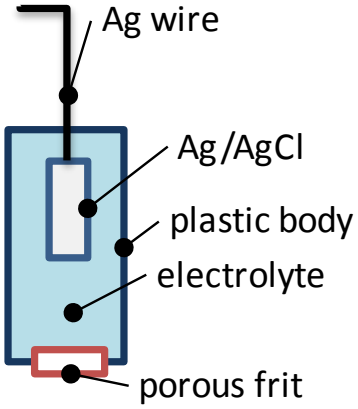


Figure 58. A commercial reference electrode is based on a silver wire coated at one end with Ag/AgCl, housed in a plastic body filled with electrolyte. The electrode assembly interacts with the field environment via a porous frit that is ion permeable.

silver is insoluble, the net balance of the two reactions is such that the introduction of electrons to the electrode releases chloride ions into solution: $AgCl(s) + e^- \leftrightarrow Ag^0(s) + Cl^-$.

In field use, a reference electrode is usually encapsulated in a saturated electrolyte for greater stability. The construction of a commercial Ag/AgCl reference electrode, illustrated in Figure 58, is based on a silver wire coated at one end with Ag/AgCl. The electrolyte interacts with the field environment via a porous frit that is ion permeable.

In our system, the reference electrode is immersed directly into the top-gate electrolyte, which is buffered to provide a stable pH. We can use the Nernstian relation ($E_N = E^0 - U_T \ln(c_0)$) to calculate the electrode potential for a model situation.

The standard electrode potential of the Ag/AgCl electrode against the standard hydrogen electrode (SHE) is $E^0 = 230 \pm 10$ mV. A typical ion concentration in our experiments is $c_0 = [Cl^-] = 1$ mM. Thus, the potential of the reference electrode is $E_N = E^0 + 180$ mV $= 410$ mV with respect to the SHE. If we wish to determine the drop

from the electrode to the electrolyte, we would rather measure than calculate directly. However, it is useful to know the expected variation in the potential drop given changes in ionic strength. This is calculated by substituting $\ln(10)\log(c_0)$ for $\ln(c_0)$ in the Nernstian: $E_N = E^0 - U_T \ln(10)\log(c_0)$. This shows that each decade increase in ionic strength causes a decrease in electrode potential of about 60 mV. Thus the sensitivity of any detector relying on Nernstian potential is limited to –60 mV/decade. For example, pH sensors, which measure the concentration of hydrogen ions, have a maximum sensitivity of 60 mV/pH.

4.3.3 Counter Electrode

In the biosensor, the role of the counter electrode is to modulate the potential of the electrolyte. Since the electrolyte should only be charging the double-layers at the reference electrode and the transistor gates, the counter electrode normally drives small currents.

If there are short-circuits through oxide pinholes or defective transistor gates, the electrode may take the burden of faradaic currents generated by chemical reduction or oxidation of a substance at the electrode. Under these conditions, it is desirable for the electrode to be chemically inert but it is not necessary for the counter electrode to be an ideal non-polarizable or ideal polarized electrode.

The noble metals (gold, platinum, etc) and glassy carbon are common examples of materials used to make counter electrodes. The counter electrode in our system is made of platinum, which will not chemically react with the electrolyte.

We can safely neglect the details of the potential drop from the counter electrode to the electrolyte, since the potential of the reference electrode will simply be a constant offset from the electrolyte potential.

4.3.4 CNTFET Metal Gates

The gates of the transistors in our biosensor are electrically floating films of gold. They are isolated from the rest of the transistor by dielectric and from the bulk electrolyte by the double-layer. Under normal circumstances, no electrochemical reaction should occur on them that would produce faradaic currents. Thus they are best modeled as ideal polarized electrodes. The double-layer model and Nernst relations presented earlier will serve to model the capacitive coupling of the gates to the electrolyte. In addition, more comprehensive models account for the possibility of ions adsorbed directly on the electrode surface. We will move forward to develop these models to describe the theory of operation.

4.4 Biosensor Theory of Operation

In the case of direct field-effect detection, we wish to measure the electrical charge on DNA or RNA (macroions) tethered to the gate of a transistor. We have already modeled how a diffuse layer of counterions balance a layer of surface charge on an electrode. Now we extend the double-layer theory to describe the response of the biosensor to a sheet charge density.

The theory of ion distributions in aqueous solutions was initiated almost a century ago by Gouy[68] and Chapman,[69] who applied the Poisson-Boltzmann theory to calculate the spatial distribution of monovalent ions near a uniformly charged interface. (Prior to

this, Helmholtz made the assumption that any surface charges on an immersed electrode would be neutralized by a sheet of counterions placed at one ionic radius away from the surface. However, this does not occur in nature.) The Gouy-Chapman model proposes that a layer of charge at a solid surface is neutralized by a layer of adsorbed counterions (via van der Waals interaction) as well as a diffuse region of increased counterion concentration in aqueous solution. The effect of excess salt concentration was extended by Debye and Hückel.[70] In 1924, Stern[71] accounted for the finite ionic radius by introducing a layer of ions, now called the Stern layer, located near the charged interface with a different dielectric constant. Most recently, Pouthas *et al.* provided a clear theoretical picture of the operation of a GenFET.[51] The following arguments assume a monovalent 1:1 electrolyte (such as NaCl).

4.4.1 Poisson-Boltzmann Formula

We start with the **Boltzmann distributions** presented earlier in (22):

$$\begin{aligned} p(x) &= n_0 \mathrm{e}^{-\phi(x)/U_T} \\ n(x) &= n_0 \mathrm{e}^{+\phi(x)/U_T} \end{aligned}, \tag{27}$$

Then recall the **Poisson formula** from electrostatics:

$$\frac{d^2\phi}{dx^2} = \frac{-\rho_q}{\varepsilon_w}, \tag{28}$$

where ρ_q is the local charge density, and ε_w is the permittivity of the solution, usually taken to be similar to that of water (~80 ε_0).

The local charge density includes both positive and negative charges:

$$\rho_q(x) = q(p(x) + n(x)) \quad (29)$$

Combining the Poisson formula with the Boltzmann distributions, we arrive at the **Poisson-Boltzmann** formula:

$$\frac{d^2\phi}{dx^2} = \frac{-qn_0}{\varepsilon_w}\left(e^{-\phi(x)/U_T} - e^{+\phi(x)/U_T}\right). \quad (30)$$

Now, if we define the geometry and boundary conditions (potentials), we can solve for the profile of the electric potential, as well as field and charge. Later, we will develop the theory further to solve the electrostatics given a certain density of DNA on the transistor.

4.4.2 Debye Screening Length

We now define λ_D as the "Debye length":

$$\lambda_D \equiv \sqrt{\frac{U_T \varepsilon_w}{2qn_0}}. \quad (31)$$

The **Poisson-Boltzmann** formula (30) can be written as:

$$\frac{d^2\phi}{dx^2} = \frac{-U_T}{2\lambda_D^2}\left(e^{-\phi(x)/U_T} - e^{+\phi(x)/U_T}\right). \quad (32)$$

Now we see that the solutions will involve the Debye length within exponentials. At room temperature (T = 300 K) in a solution of 1:1 monovalent salt, the Debye length is approximated as:

$$\lambda_D \cong \frac{0.31\,\text{nm}}{\sqrt{[c]}}, \quad (33)$$

where [c] is the bulk ionic concentration specified in molarity (mol/L). We will show soon that the Debye length determines the severity of counterion screening.

Suppose a sheet of charge exists at $x = 0$. Via application of Gauss' Law, the electric field due to the sheet charge is $E = \sigma_s/\varepsilon_w$, where σ_s is the surface charge density, and the field points away from the charge if the charge is positive. By definition, the electric field is equal to the negative of the gradient of potential: $E = -d\varphi/dx$. Thus, a solution to the Possion-Boltzmann formula must satisfy the boundary condition $d\varphi/dx = -\sigma_s/\varepsilon_w$ at $x = 0$.

For small enough potentials ($\varphi(x) < U_T$), approximate solutions for the electric field and potentials are given by:

$$E = \frac{\sigma_s}{\varepsilon_w} e^{-x/\lambda_D} \quad \text{and} \quad \phi(x) \cong \frac{\sigma_S \lambda_D}{\varepsilon_w} e^{-x/\lambda_D}, \tag{34}$$

which are valid for a weakly charged surface ($\sigma_s < \varepsilon_w U_T / \lambda_D$) and $x > 0$. Furthermore, the potential drop from the bulk electrolyte to the sheet charge is:

$$\phi_{EC} = \phi(\infty) - \phi(0) \cong -\frac{\sigma_S \lambda_D}{\varepsilon_w}. \tag{35}$$

It is now clear that the Debye length characterizes the spatial decay in electric field strength. This screening effect means that biomolecules do not feel one another until they are very close, but also means that their interactions can be very specific to their detailed shape. This allows for highly specific complementary nucleic acid binding and protein-antibody binding.

We illustrate the impact of Debye screening with the following toy model. A typical buffer for DNA hybridization has an ionic strength of 150 mM. The Debye length in this solution is 0.31 nm/√(150 mM) ≈ 0.8 nm. Now assume a length of hybridized double stranded DNA (dsDNA) is attached perpendicular to a surface. Given that one

base pair (bp) of double stranded DNA (dsDNA) is 0.34 nm long, only the first two bases pairs will contribute to the electric field as the intrinsic charges on the third base pair and beyond would be heavily screened by the electrolyte.

Let us take the liberty to reduce the ionic strength of the buffer to increase the Debye length. Selective recognition requires probes to be at least 16 nucleotides long,[72] or 5.5 nm from end to end. The required Debye length should be longer than this in order to detect the charge of the hybridized DNA; an electrolyte of less than 3 mM ionic strength is required.

Decreasing the ionic strength of the hybridization buffer to 1 mM results in a Debye length of about 10 nm, or about 29 mers. Steel *et al.*[64] measured surface densities of 25-mer DNA immobilized on gold of $(1-10) \times 10^{16}$ m^{-2}. At the lower density of 10^{16} m^{-2}, the DNA are spaced at a square pitch of 10 nm. This equates to a charge density of $\sigma_{DNA} = -25\ q_e \times 10^{16}\ m^{-2} = -40$ mC/m^2. According to our approximation, the potential drop from the bulk electrolyte to the sheet charge is $\varphi_{EC} = (-\sigma_{DNA}\ \lambda_D / \varepsilon_w) = (40\ \text{mC/m}^2)$ (9.8 nm)/$80\varepsilon_0$ = 550 mV. This voltage would be easily measureable but exceeds the bounds of approximation ($U_T \approx 25$ mV). We would need to scale back the DNA loading by 20 times to be at that limit. We now proceed to develop an exact solution to the electrostatics.

4.4.3 Grahame Equation

We continue to develop the model of operation, which is based on the Grahame Equation. For the purposes of the model, we will assume that the CNTFET can be modeled as a planar FET and the DNA is in direct contact with the top gate oxide. Minor adjustments will need to be made for presence of the gold transistor gate. We will also

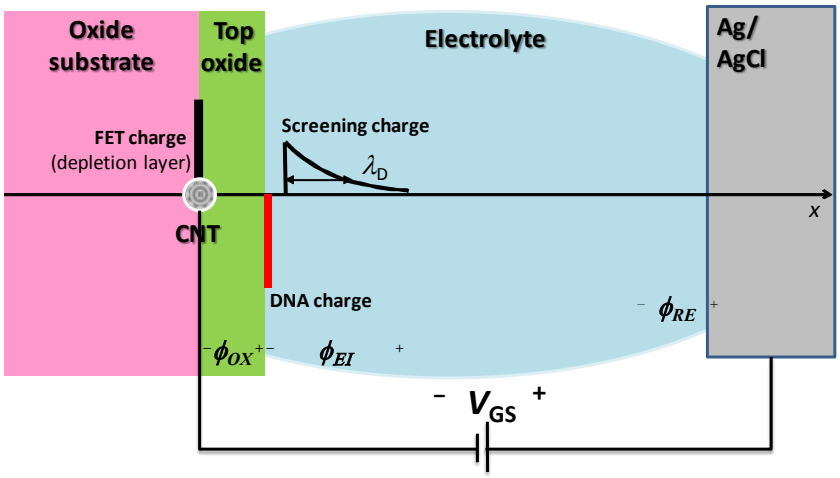


Figure 59. Simplified schematic of CNTFET electrostatic model. The charge on the FET and DNA is screened by the diffuse layer, characterized by the Debye length (λ_D). By Kirchoff's Voltage Law (KVL), $V_{GS} = \varphi_{RE} + \varphi_{EI} + \varphi_{OX}$. The counter electrode is not shown for simplicity.

ignore potential drop from the reference electrode to the electrolyte since this potential is invariant with respect to the hybridization process.

A cartoon model of charge vs. position is shown in Figure 59, which illustrates how the DNA charge is screened partially by the FET and partially by the electrolyte. The diffuse screening charge of the electrolyte is described by the double-layer and characterized by the Debye length. By Kirchoff's Voltage Law (KVL), the gate-source voltage (V_{GS}) is the sum of the potential drops from the Ag/AgCl reference electrode to the source: $V_{GS} = \varphi_{RE} + \varphi_{EI} + \varphi_{OX}$, where φ_{RE} is the electrode-electrolyte drop, φ_{EI} is the electrolyte-gate drop, and φ_{OX} is the gate-source drop. For simplicity, the counter electrode is not shown, and we neglect φ_{RE} because it is invariant with respect to the charge of the DNA. Thus, $V_{GS} = \varphi_{EI} + \varphi_{OX}$.

The electric field in the oxide is $E_{ox} = \varphi_{OX} / t_{ox}$. By Gauss' Law, the potential across the oxide is $\varphi_{OX} = -\sigma_{FET}\, t_{ox} / \varepsilon_{ox}$, where t_{ox} is the thickness of the top oxide and ε_{ox} is its permittivity. The gate voltage V_{GS} is known because it is applied by a voltage

source. Therefore, the remaining task is to find the drop from the electrolyte to the interface (φ_{EI}) as a function of DNA charge density (σ_{DNA}). From this point, we will follow the recent treatment by Pouthas *et al.*[51]

The Gouy-Chapman model proposes that the layer of DNA charge would be neutralized by the diffuse layer as well as a layer of adsorbed counterions (via van der Waals interaction). This model includes a more exact solution to the Poisson-Boltzmann formula:

$$\phi(x) = 2U_T \operatorname{arsinh}\left(-\frac{d\phi}{dx}\frac{\lambda_D}{2U_T} \right). \tag{36}$$

A schematic representation of the charge density, electric field and potential profile of our device, which we treat analogous to a MOSFET, is shown in Figure 60. At the oxide-electrolyte interface, there are two important sheet charges: the surface charge σ_1 is mainly due to the ionization of amphoteric sites on the oxide surface, and the positive compact layer (also known as the Stern layer) charge σ_2 is due to adsorbed cations with loci at the plane of $x = \delta$. On the right of the Stern layer is the sheet of DNA charge σ_{DNA}, which is negative due to phosphate groups on the DNA backbone. The charge on the transistor is σ_{FET}, located at $x = -t_{ox}$.

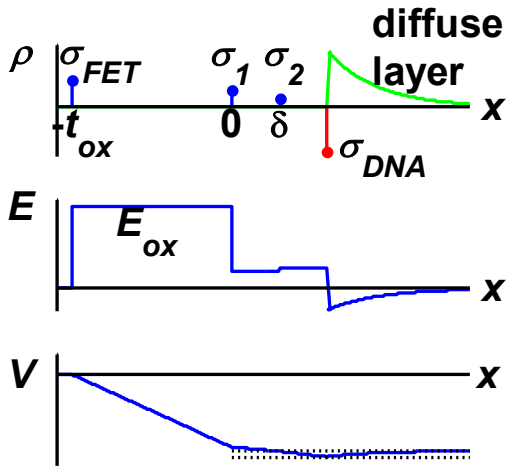


Figure 60. Charge density ρ, electric field E, and potential profile V through FET, oxide, and electrolyte (not to life scale). The charge on the transistor is σ_{FET}, located at $x = -t_{ox}$. The surface charge σ_1 is mainly due to the ionization of amphoteric sites on the oxide surface, while the charge σ_2 is due to adsorbed cations with loci at the plane of $x = \delta$. Immediately to the right is the sheet of DNA charge σ_{DNA}, which is negative due to phosphate groups on the DNA backbone. The diffuse layer balances the net of all of the aforementioned charges. The electric field in the oxide is E_{ox}.

The total charge prior to the diffuse layer is $\sigma_{eff} = \sigma_1 + \sigma_2 + \sigma_{DNA} + \sigma_{FET}$. Then the electric field just to the right of $x = \delta$ is by Gauss' Law: $E = \sigma_{eff}/\varepsilon_w$. Therefore the diffuse layer potential drop from the bulk electrolyte to the interface is given by:

$$\phi_{EI} = -2U_T \operatorname{arsinh}\left(\frac{\sigma_{\text{eff}}\lambda_D}{2U_T\varepsilon_w}\right). \tag{37}$$

This equation, a form of the Grahame equation,[73] shows that potentials of greater magnitude can be expected with increasing charge density and longer Debye lengths. The result of the Grahame equation is plotted as a function of charge density for several ionic concentrations in Figure 61a. At low ionic strength (1 mM), the charge sensitivity is quite

high but only for low charge density. At high ionic strength (1M), the sensitivity is relatively lower but is approximately linear over a larger range. In either case, the arsinh begins to reduce the charge sensitivity once $\varphi_{\mathrm{EI}} > 2U_T \approx 50$ mV. The sensitivity roll-off is especially severe when $\varphi_{\mathrm{EI}} > 200$ mV.

Suppose a feedback system works to maintain the gate voltage V_{GS} such that σ_{FET} is always zero. Equivalently, we can track the threshold voltage over several scans of the transfer characteristic, where we assume the channel conduction is lowest because $\sigma_{FET} = 0$. In either case, $\varphi_{\mathrm{OX}} = 0$, and the gate voltage is simply $V_{GS} = \varphi_{\mathrm{EI}}$. If we further assume σ_1, σ_2 are insignificant, then $\sigma_{eff} = \sigma_{DNA}$; any change in the gate voltage is equal to only the Grahame potential from the DNA hybridization.

Let us estimate how much voltage could be generated by a reasonable DNA loading. Our previous estimate of 550 mV was satisfyingly large but out of the range of approximation. Let us take $\sigma_{DNA} = -40$ mC/m^2 as before, and assume σ_1, σ_2, and σ_{FET} are zero, so $\sigma_{eff} = \sigma_{DNA}$. The Grahame equation predicts a resulting voltage of $\varphi_{\mathrm{EI}} = 160$ mV, which, although much less than our original estimate, would still be easily measurable.

Let us further explore the dampening effect of the hyperbolic sinh and the possibility of linearization. Define $\sigma_0 \equiv 2\ U_T\ \varepsilon_w / \lambda_D$. Figure 61b shows that the arsinh does not have much significance for low charge densities ($|\sigma_{eff}| < \sigma_0$), corresponding to $\varphi_{\mathrm{EI}} < 2U_T \approx 50$ mV. We call the regime the linear range, in which the charge sensitivity is $|\partial\varphi_{\mathrm{EI}} / \partial\sigma_{eff}| = \lambda_D / \varepsilon_w$. This opens the possibility for creating a linear circuit model.

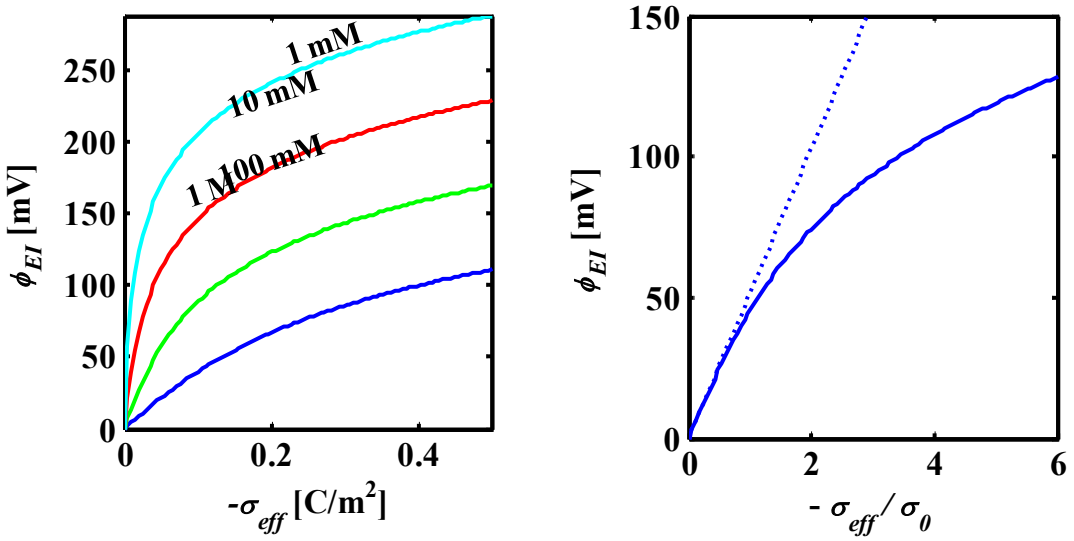


Figure 61. (a) Potential drop from the bulk electrolyte to the charged interface as a function of charge density (σ_{eff}) with ionic concentration as a parameter. At high ionic strength, the charge sensitivity is lower but the linear range is wider. At low ionic strength, the sensitivity is higher but the linear range is smaller. (b) (solid) Potential drop from the bulk electrolyte to the charged interface as a function of normalized charge density ($-\sigma_{eff}/\sigma_0$). (dotted) Plot of $2U_T(-\sigma_{eff}/\sigma_0)$ for reference. For small charge density ($|\sigma_{eff}| < \sigma_0$), the response is approximately linear. For larger charge densities ($|\sigma_{eff}| > \sigma_0$), the charge sensitivity is moderated by the arsinh function.

At higher charge densities ($|\sigma_{eff}| > \sigma_0$), the charge sensitivity is significantly moderated. Using the identity $\text{arsinh}(x) = \ln(x+(x^2+1)^{1/2})$, we see that for $x >> 1$, $\text{arsinh}(x) \approx \ln(2x)$. Thus although this region is less sensitive, there is no fundamental limit restricting the magnitude of the potential at ever larger loading densities.

4.4.4 Circuit Model

We now formally incorporate the CNTFET, electrolyte, and reference electrode into a single electric circuit (Figure 62). The circuit contains the gate-source voltage source (V_{GS}) as well as two capacitors: the Debye capacitance (C_D), and the FET oxide

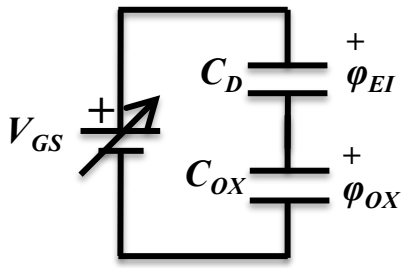


Figure 62. Electrical schematic diagram of biosensor. C_D is the Debye capacitance, and C_{OX} is the oxide capacitance.

capacitance (C_{OX}). By Kirchoff's Voltage Law (KVL), the gate-source voltage (V_{GS}) is the sum of the potential drops from the reference electrode to the source: $V_{GS} = \varphi_{EI} + \varphi_{OX}$, where φ_{EI} is the electrolyte-gate drop, and φ_{OX} is the gate-source drop.

To aid in circuit analysis, we simplify the Grahame equation. First, we neglect σ_1 and σ_2 such that $\sigma_{eff} = \sigma_{DNA} + \sigma_{FET}$. Second, we approximate arsinh to unity for small charge density ($|\sigma_{eff}| < (2\ U_T\ \varepsilon_w / \lambda_D)$). Using this linear approximation, $\varphi_{EI} = -(\sigma_{DNA} + \sigma_{FET})/C_D$, where $C_D \equiv \varepsilon_w/\lambda_D$.

The oxide capacitance is defined as $C_{OX} \equiv \varepsilon_{ox}/t_{ox}$. Then $\varphi_{OX} = -\sigma_{FET}/C_{OX}$, where the negative sign is consistent with Gauss' Law. Combining this with the approximation for the Grahame equation and substituting into the KVL equation, the gate voltage is given by:

$$V_{GS} = -\frac{1}{C_D}\left(\sigma_{DNA} + \sigma_{FET}\right) - \frac{\sigma_{FET}}{C_{OX}} \quad (38)$$

Solving for the FET charge:

$$\sigma_{FET} = \frac{-C_{ox}C_D}{C_{ox}+C_D}\left(V_{GS}+\frac{\sigma_{DNA}}{C_D}\right) = -C_{XD}\left(V_{GS}+\frac{\sigma_{DNA}}{C_D}\right). \quad (39)$$

In the deep ohmic region ($V_{DS} << (V_{GS}-V_{TH})$), the current-voltage relation for a standard MOSFET can be approximated as $I_D = \mu$ (W/ L) V_{DS} C_{OX} ($V_{GS}-V_{TH}$). (In experiment, we apply a V_{DS} of 100 mV and V_{GS} on the order of 1000 mV). The drain current can be decomposed into two parts: the first part is μ (W/ L) V_{DS}, which has units of m^2/s. The second part is C_{OX} ($V_{GS}-V_{TH}$), having units of C/m^2. By observation, we realize that the threshold voltage of the biosensor is:

$$V_{TH} = \frac{-\sigma_{DNA}}{C_D}. \quad (40)$$

Thus, the threshold is directly proportional to the DNA charge, and the sensitivity is simply given by $-1/C_D$. The Debye capacitance is the only capacitance that plays a role in shifting the threshold of the biosensor in response to the capture of nucleic acids. If we want to maximize sensitivity, we should seek to minimize the Debye capacitance by using a buffer with as low ionic strength as possible.

4.4.5 Tradeoff: Sensitivity vs. Linear Range

In the linear region, the sensitivity of the threshold shift in response to DNA charge is given by $|\partial V_{TH} / \partial \sigma_{DNA}| = 1/C_D = \lambda_D / \varepsilon_w = 0.31$ nm/($[c]^{1/2}$ 80 ε_0). From the previous section, we already know that lowering the ionic strength increases the sensitivity, but also decreases the linear input range.

We take the optimal DNA loading as the highest density of DNA such that the charge on the FET and DNA obey $|\sigma_{DNA} + \sigma_{FET}| < \sigma_0$ ($\equiv 2\ U_T\ C_D$). In other words, the sum of the DNA and FET charge is such that the threshold voltage is $V_{TH} = 2\ U_T = 51$ mV.

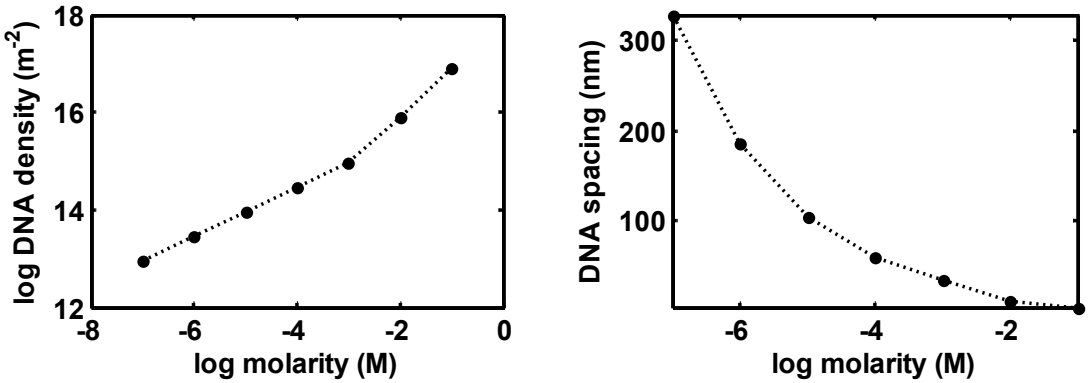


Figure 63. Optimal DNA loading vs. ionic strength of measurement buffer. The optimal loading is taken as the highest possible density of 25-mer DNA still in the linear range. Each decade increase in ionic strength allows for a $\sqrt{10}$ (≈3.16) increase in DNA loading. The extreme of 10^{-7} M ionic concentration represents deionized water.

Suppose a feedback system works to maintain the gate voltage such that σ_{FET} is always zero. Equivalently, we can track the threshold voltage, where we assume the channel conduction is lowest because $\sigma_{FET} = 0$. In either case, $\varphi_{OX} = -\sigma_{FET}/C_{OX} = 0$, $V_{GS} = \varphi_{EI}$, and $\sigma_{eff} = \sigma_{DNA}$; any change in the gate voltage is equal to only the Grahame potential from the DNA hybridization.

Thus the optimal DNA charge density is $\sigma_{DNA} = -2\ U_T\ C_D = 2\ U_T\ \varepsilon_w / \lambda_D \approx -[c]^{1/2}$ 118 mC/m^2, which is basically determined by the buffer ionic strength. The optimal DNA loading as a function of buffer strength is shown in Figure 63, where each decade increase in ionic strength allows for a $\sqrt{10}$ (≈3.16) increase in DNA loading.

We again assume 25-mer DNA with density on the order of 10^{16} m^{-2} (Steel *et al.*[64]). The optimal ionic concentration for linear response is in the regime of 100 mM. In the case that higher sensitivity is more important than linearity, a lower ionic strength buffer may be used. For a 1 mM buffer, the optimal DNA loading is 3.6 mC/m^2, which corresponds to a square pitch of 33 nm and areal density of 10^{15} m^{-2}, which suggests that

if linearity is of paramount importance, we may want to increase the spacing between the probes using spacer molecules.[74] Alternatively, we can choose to tolerate the roll-off in sensitivity at higher loading levels.

4.5 **Conclusion**

We developed the theory of operation of the CNTFET biosensor beginning from the basic laws of drift and diffusion. The double layer model leads to the Nernst relations, providing an understanding of the double-layer model as it applies to electrodes in electrolytes, as well as serving as the basis of the Grahame equation. We found that no fundamental limitations restrict the maximum signal that could be generated by DNA hybridization, but that there is a finite range of DNA densities over which the electrical response is linear. The choice of buffer strength determines the tradeoff between sensitivity and linear input range when the probe density is fixed.

Chapter 5: Biosensor for HER2 Overexpression in Breast Cancer Cells

5.1 Chapter Purpose

This purpose of this chapter is to show that the prototype HER2 detection system was successful in detecting HER2 overexpression in real breast cancer cells. In collaboration with the NCI, we developed protocols for immobilizing the probe, hybridizing the RNA on the biosensor, and determining status of HER2 at the mRNA level. We demonstrated consistent dose-response behavior wherein higher RNA concentration produced larger hybridization signals. Over-expression of HER2 breast cancer gene in HER2 positive cancer cells was gauged relative to two types of experimental controls. The first is detection of the GAPDH house-keeping gene alongside HER2 for the purpose of normalization. The second is to use a HER2 negative cell line for reference. We were able to discriminate between the two cancer cell lines based on the relative HER2 expression level.

The direct detection of label-free mRNA has not been highly documented in the literature. We may be the first to accomplish the discrimination between two cancer cell lines via direct detection of label-free mRNA.

5.2 Detection of Total mRNA from Cancer Cells

Before delving into detection of HER2 mRNA from cancer cells, we focused on detection of generic mRNA as a stepping stone from prior success in using DNA probes to detect synthetic DNA targets. Like synthetic DNA oligomers, most RNA molecules

are single-stranded. However, biological RNA is more challenging to work with for two major reasons. First, naturally occurring RNA comes in a wide variety of sizes (lengths) and shapes. RNA molecules tend to form self-loops and take on complex three-dimensional structures. Second, RNA is generally difficult to work with because it is chemically unstable. The sugar-phosphate backbone of RNA is weaker due to the free hydroxyl group in the ribose ring.

Total RNA includes messenger RNA (mRNA), ribosomal RNA, and transfer RNA. In this experiment, we targeted messenger RNA (mRNA), which are distinguished from other RNA in that they have a 3' polyadenine (poly(A)) tail which protects them from enzymatic degradation and promotes their export from the cell nucleus for translation. If we were able to detect mRNA despite the presence of other RNA in this mixture, the stage would be set for specifically detecting HER2 mRNA.

5.2.1 Methods

The methodology of this preliminary experiment is to generate a basic dose-response (or standard) curve that relates the threshold shift of the biosensor with the concentration of total RNA applied during hybridization. A schematic of a dose-response curve is shown in Figure 64a. The response is linear over a finite range of concentration but saturates above a certain limit, often due to maximum loading of the probe molecule. Below the limit of detection (LOD), the response is "in the noise" and is not significant. Our goal is to experimentally determine what range of RNA concentration is the "sweet spot" (approximately linear regime) of the biosensor, while being above the limit of detection and below saturation.

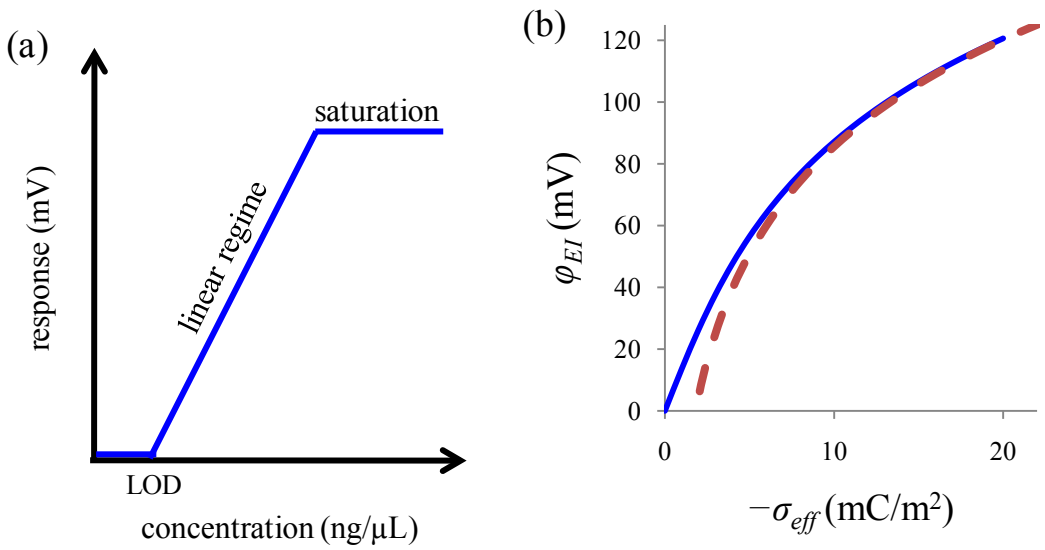


Figure 64. (a) Schematic of a dose-response or standard curve. The response is linear over a finite range of concentration. Below the limit of detection (LOD), the response is "in the noise" and considered zero. Above a certain limit, the response saturates. (b) Potential from electrolyte to gate interface as a function of effective surface charge density. Dashed line is logarithmic approximation (see Chapter 4).

As described in Section 4.4.3, the Grahame equation (37) captures the relationship between the areal density of RNA charge and the measured voltage, which strongly depends upon the Debye screening length of the buffer. For an ionic strength of 1 mM, this relationship is plotted in Figure 64b, which can be used to infer the amount of RNA charge per unit area based on the measured voltage shift in the transfer characteristic. In theory, given the relationship between the *concentration* of RNA (#/m^3) in the hybridization buffer and the *areal density* (#/m^2) achieved due to hybridization, one could predict the voltage response to RNA concentration. However, it is difficult to know the concentration-binding relationship ahead of experiment.

Cellular RNA was harvested from live MCF7 breast cancer cells cultured from cryogenically frozen stock at the Advanced Technology Center (ATC) of the National

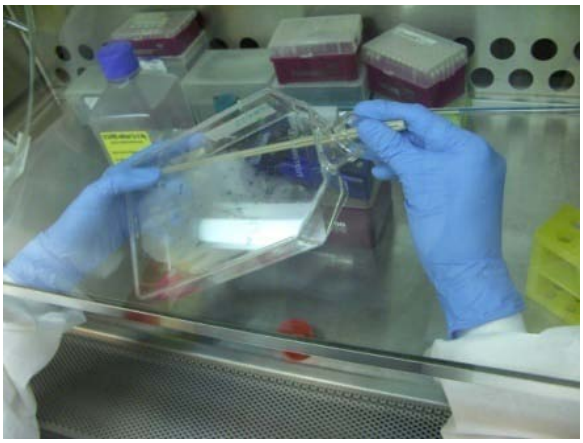
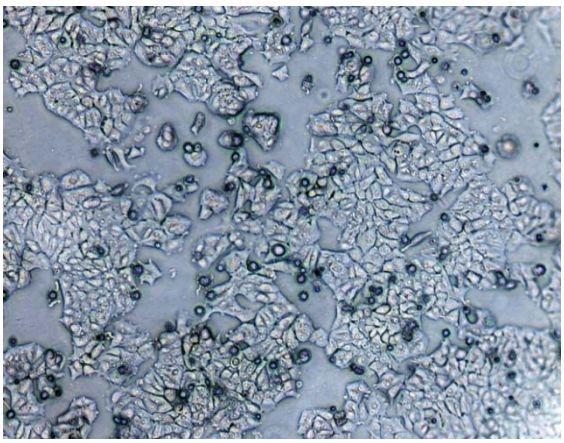



Figure 65. (left) Cancer cells being scraped from bottom of cell culture flask during RNA extraction under laminar flow biological containment hood. One flask contains about 7 million cells. (right) Magnified view of MCF7 cancer cells adhering to the bottom of a cell culture flask before passaging (splitting a dense population of cells into two new groups).

Cancer Institute (NCI). The MCF7 breast cancer cell line was originally isolated from a woman in 1970 and was the first mammary cell line that was capable of lasting longer than a few months. Cells were cultured over several months to provide ample supply of RNA for this project. In cell culture, the cells grow in plastic flasks containing a liquid media that provides appropriate pH and nutrients. The cancer cells reproduce until they are so crowded that they must be either split among two new flasks (see Appendix: How to Passage Cells) or chemically lysed to extract the RNA (see Figure 65). If the cells are too dense, nutrients will be depleted and the cells will die. Total RNA was extracted from the cancer cells via the Qiagen® RNeasy MiniKit protocol, eluted in RNase-free deionized water (Qiagen) to a stock concentration of about 420 ng/µL, and stored at −80 °C until use.

Table 8. Probe sequence and structure of mRNA. UTR = untranslated region.

Probe	5'-/thiol/TTT TTT TTT TTT TTT T-3'
Target	5'-/cap/5'UTR/Coding Sequence/3'UTR/AAA…AAA/-3'

We developed a 16-mer poly-thymine (poly(T)) PNA probe to selectively target mRNA (see Table 8). This probe can easily hybridize to mRNA since it can bind to any consecutive string of 16 adenine bases along the poly(A) tail, which is more than 200 nucleotides long.

To functionalize the biosensor for experiment, the freeze-dried poly(T) PNA probe was diluted in DI water purged with nitrogen and applied to the chip at a concentration of 1 mM and volume of 50 µL. After overnight incubation, the chip was rinsed with 0.1X TE (Quality Biological TE pH 8.0 diluted down to 1 mM with RNase-free water), which was also used as the default top-gate buffer (TGB) during threshold scans. The functionalized chip was then plugged into the measurement platform and scanned repeatedly to gauge stability. To prepare for hybridization, the chip was taken off the platform and excess TGB was removed. The hybridization was performed by pipetting the RNA (in 25 µL of TGB) into the reservoir and incubating for one hour at room temperature in a sealed humid environment. After a manual rinse with TGB, the chip was scanned to acquire the post-hybridization thresholds.

5.2.2 Results

As shown in Figure 66, chip-averaged threshold (VTH) shifts of 72 mV and 99 mV occurred in response to two different dilutions of total RNA: (A) 25 ng/µL and

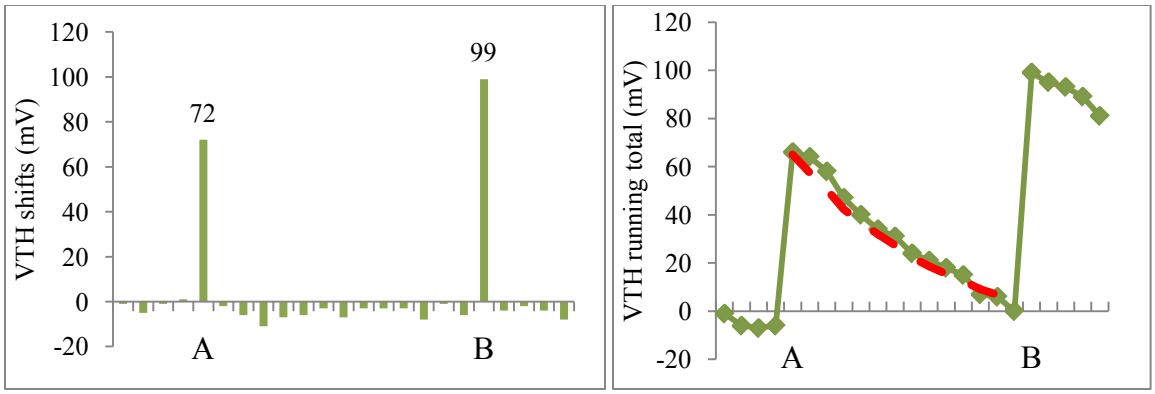


Figure 66. (left) Biosensor threshold (VTH) shift and (right) VTH running total probing for total mRNA in two concentrations of total RNA extracted from breast cancer cells. Hybridizations at 25 ng/µL (prior to scan A) and 50 ng/µL (prior to scan B) produced shifts of 72 mV and 99 mV, respectively. An exponential settling curve with a time constant of 30 mins was fit to the data.

(B) 50 ng/µL. With the exception of these two hybridization signals, the shifts between any two consecutive scans ("repeats") were miniscule in comparison. During the 13 scans between the two hybridizations, the threshold drifted at an average rate of −5 mV/scan. An exponential settling function with a decay time constant of 30 minutes was fit to these 13 scans.

5.2.3 Analysis & Discussion

The dose response data of Figure 67 are increasing as a function of the total RNA concentration, as expected. We also performed a hybridization with 1374 ng/µL of unrelated synthetic oligos (data not shown), which produced a shift of only 17 mV, suggesting that the probe is specifically detecting mRNA.

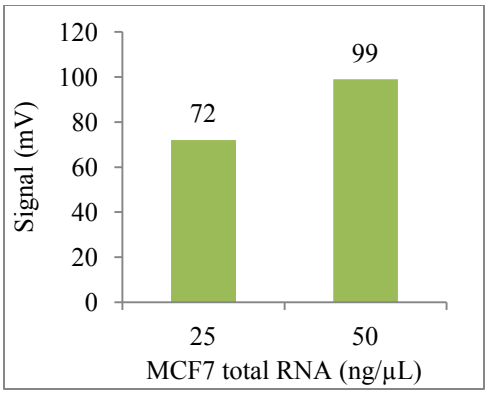


Figure 67. Signal (threshold shifts) of biosensor targeting mRNA from two increasing concentrations of total RNA extracted from the MCF7 breast cancer cell line. The amount of RNA used corresponds to roughly 30,000 cancer cells.

Since doubling the dose of RNA did not produce twice the signal of the first dose, the sensor is not responding linearly. The apparent saturation in response at the higher RNA dose could be due to the probe being loaded by the first dose. In a strict sense of the dose-response, each experiment should start with a fresh chip that has never been hybridized before.

It is also possible that the surface charge of the tethered RNA is so dense that we are seeing the inherent non-linear response of field-effect detection. From Section 4.4.5, we see that any response greater than 50 mV is already outside of the "linear" regime of the Grahame equation. Using the Grahame equation (37) and Figure 64b, we back-calculate that a response of about 100 mV implies an RNA surface charge density of 10 mC/m^2. Assuming that the effective length of each RNA is 25 mers (approximately the Debye length), the RNA density is $\sigma_{RNA}/(25q_e) = 0.25 \times 10^{16}$ m^{-2}, or equivalently a square pitch of 20 nm. A reasonable estimate of the probe density is twice the RNA density (hybridization efficiency of 50%), or 0.5×10^{16} m^{-2}. This is an order of magnitude lower than the densities of ~5×10^{16} m^{-2} measured by Steel *et al.*[64]

The area of probe immobilization, on the gold gates, is 1.5×10^{-7} m^2. With the probe loading at ~10^{16} m^{-2}, the total number of PNA probe molecules is 1.6×10^9. The number of mRNA floating in solution is estimated using an average molecular weight of 500,000 g/mol to be 7.5×10^{10} molecules. The ratio of floating mRNA to probe PNA is ~50, suggesting that we are still far from completely capturing all the target molecules in solution. While our estimate of the probe loading seems low, it is important to increase both probe loading *and* hybridization efficiency. Herne & Tarlov[74] showed that a tradeoff exists; if the probe loading is too high, hybridization efficiency decreases. However, one can create just the right probe density by inserting spacer molecules between adjacent probe molecules to prevent steric hindrance.

The negative drift following each hybridization appears to have a time constant of 30 minutes, which is consistent with the time constant of 25 minutes as we found the liquid gating experiment in Section 2.6.2. Therefore the settling behavior probably originates from the slow relaxation of charges in the electrolyte equilibrating with the transistor gate.

Let us now get a grasp of the clinical relevance of the detection sensitivity. A typical mammalian cell contains about 20 pg total RNA. Using the 5% estimate, the amount of mRNA in one cell is about 1 pg. The lower concentration of total RNA in this experiment contained 630 ng total RNA. Given that mRNA makes up less than 5% of total RNA, our system was detecting less than 32 ng of mRNA. This implies our system was detecting mRNA from less than 32,000 cells. For comparison, a core biopsy of 1 mm in diameter and 2 mm in length could be expected to yield one million cells.

5.3 Detection of HER2 mRNA from Cancer Cells

Given the positive results from the experiments probing for generic total mRNA from real cancer cells, we were ready to probe specifically for HER2 mRNA in two different cell lines. The SKBR3 breast cancer cell line is HER2-positive (HER2+) and overexpresses HER2 mRNA,[75] whereas MCF7 is HER2-negative (HER2−) and expresses HER2 mRNA at a normal level.[76]

Before we move into the details of the experiments, let us briefly examine the challenge we are facing. One cell contains about twenty million strands of RNA from many different genes. Of these, just over a million strands are messenger RNA, amounting to about 1 pg of mass. Now we want to detect HER2 specific mRNA, of which there are as many as 52 copies in the SKBR3 cell or as little as a single copy in the MCF7 cell. All else being equal, we are aiming to measure a signal 27,000 times as small as the previous experiment!

5.3.1 Response to MCF7 (HER2−) & SKBR3 (HER2+) Cell Lines

We assign one chip per cell line so the experiment can proceed on the same measurement platform with the two chips alternating in a pipeline stream. The probe for HER2 was immobilized on both chips at a concentration of 1 mM in a volume of 5 μL. After overnight incubation, the chip was rinsed with TGB gently, then mounted on a chip carrier, wirebonded, and attached to reservoir tubes. The adhesive used to attach the reservoirs was allowed to cure overnight.

The next morning, we ran a few baseline scans to establish stability. Then we removed the excess measurement buffer and applied 25 μL of total RNA at the stock concentration of 420 ng/μL in RNase-free water. This concentration is nearly an order of

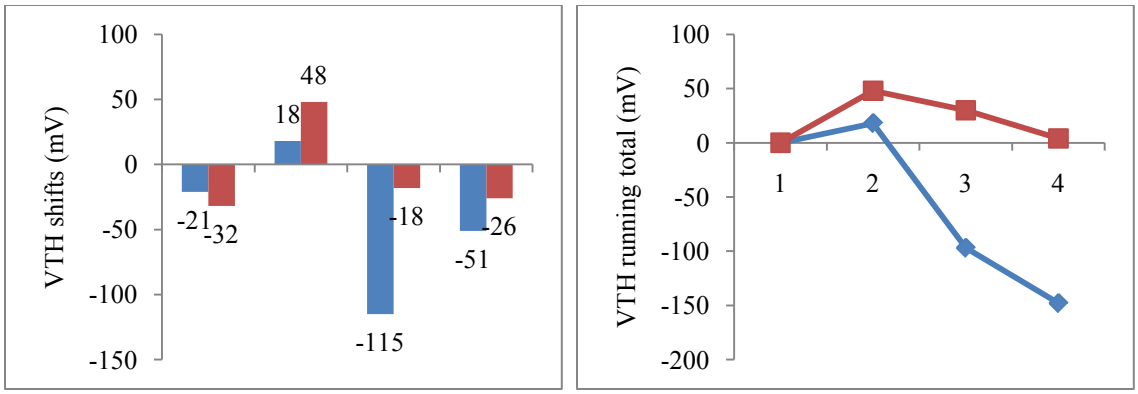


Figure 68. Threshold shifts (left) and running total (right) probing for HER2 mRNA in dose of 420 ng/µL total RNA from MCF7 (blue) and SKBR3 (red). Stability was established with multiple scans prior to the four scans shown. Scan 1 was taken before hybridization. After hybridization, three successive scans were taken (2, 3, 4).

magnitude higher than what we used when probing for total mRNA. We chose to use the highest concentration available to increase the chances of seeing a signal from HER2 mRNA, which, at best, exists at levels less than 40 parts per million (ppm) relative to total mRNA. Hybridization was performed through incubation at room temperature for approximately one hour, after which the chips were rinsed several times with TGB. Finally, three post-hybridization scans were measured.

A windowed portion of the scan data is shown in Figure 68. The hybridization was performed between scan 1 and scan 2: 18 mV for MCF7 and 48 mV for SKBR3. Scans 3 and 4 were taken to gauge system stability. MCF7 chip experiences a large negative shift (−115 mV from scan 2 to 3), followed by a smaller settling shift. The SKBR3 chip also settles but the shifts are within the drift signal of −20 to −30 mV/scan.

The large negative settling shift of MCF7 may be due to rapid dehybridization of RNA, which is notoriously chemically unstable. Another possibility is instability in the

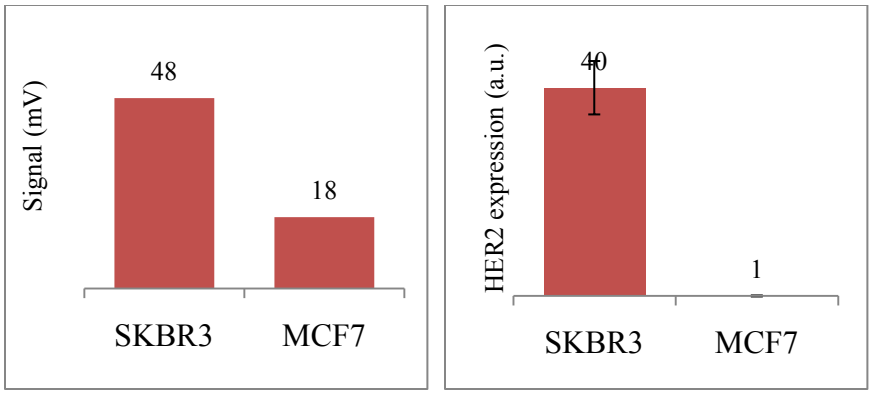


Figure 69. (left) Biosensor response probing for HER2 from total RNA extracted from two breast cancer cell lines: SKBR3, which over-expresses HER2, and MCF7, which expresses HER2 at normal levels. RNA was incubated on each chip at a concentration of 420 ng/μL total RNA. A different chip was used for each cell line. (right) HER2 expression in SKBR3 as measured by quantitative real-time PCR (arbitrary units) was 40 times that of MCF7. We have qualitative agreement with an established method.

transistors (corrosion of oxide and metal wires). The large difference in chip stability is reason to incorporate blank measurements into the experimental protocol.

The signal due to HER2 is the increase in threshold from scan 1 to scan 2, which is summarized in Figure 69. The small response to MCF7 and large response to SKBR3 are consistent with the overexpression of HER2 mRNA only in SKBR3. However, it is also possible that total mRNA from SKBR3 happened to be expressed or extracted at a higher level than from MCF7. Thus, we need a basis on which to normalize the HER2 signal.

Although the hybridization buffer for the first mRNA detection experiment was TGB, in this case, it was simply DI water. This is because we wanted to use the highest concentration of RNA available (the stock). The absence of a proper buffer in molecular biology is expected to significantly affect hybridization. Furthermore, the switch from

Table 9. Mass and concentrations of total RNA and targeted RNA for this and previous experiment.

Type of RNA targeted	Total RNA mass (µg)	Total RNA concentration (ng/µL)	Targeted mRNA mass (pg)	Targeted mRNA concentration (fg/µL)	Signal (mV)
Total mRNA	0.625	25	31000	1,300,000	72
HER2 mRNA in MCF7	11	420	0.7	27	18
HER2 mRNA in SKBR3	11	420	20	780	48

TGB during measurement to DI during hybridization and back to TGB following hybridization could also elicit a response from the biosensor. This response must be accounted for by performing hybridization with samples consisting of only DI water.

A comparison of the amounts of RNA and measured signals in this experiment with those from the previous experiment can be made from Table 9. We increased the dose of total RNA over one order of magnitude since the last experiment. However, the mass and concentration of mRNA targeted in this experiment is more than three orders of magnitude lower in concentration than in the previous experiment. Yet we were able to measure significant signals and match them to the correct cell line. It is remarkable that we were able to measure any signal whatsoever!

On the following day, we tested the two chips in response to a blank hybridization in TGB, followed by a blank in DI water (Figure 70). The blank response of M0 was −4 mV for TGB and +4 mV for DI. The blank response of S0 was 25 mV for TGB and

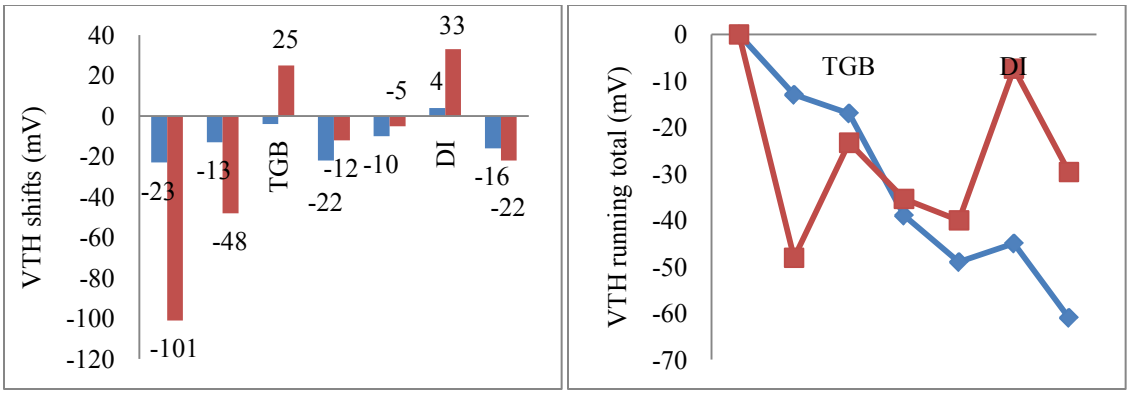


Figure 70. Threshold shifts (left) and running total (right) for biosensors M0 (blue) and S0 (red) used previously with MCF7 and SKBR3, in experiment to characterize dependence of blank hybridizations on hybridization solvent.

33 mV for DI. In both cases, the DI response is nearly 10 mV greater than the preceding TGB response. The general response of M0 over the 7 scans is a gradual decrease, while the response of S0 seems to decrease initially but recover somewhat.

The signals of these two chips due to blank incubations with TGB and DI are summarized in Figure 71. The fact that these two chips had such very different blank responses even under the same buffer highlights the need to incorporate regular blanks into the experimental plan. It also points to the use of TGB for both measurement and hybridization.

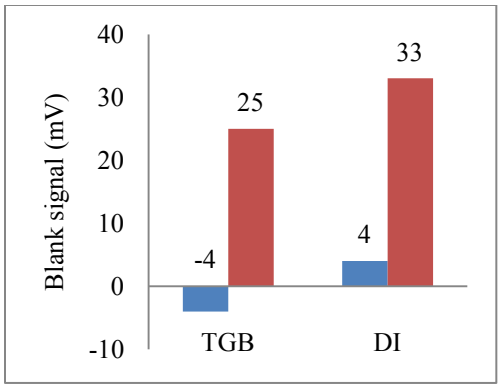


Figure 71. Signals due to blank incubations in TGB and DI water for chips M0 (blue) and S0 (red). In both cases, the DI response is nearly 10 mV higher than the previous TGB response. This result highlights the need to incorporate blanks into the overall experimental scheme and suggests the use of TGB for both measurement and hybridization.

5.3.2 Dose Response to SKBR3 with Blanking

Our next goal was to produce a dose-response curve for HER2 expression in SKBR3, while including blanks and using the same buffer throughout experiment. As in Section 5.2.1, the result we are looking for is a simple plot that shows an increasing response to increasing concentration of total RNA. In planning this experiment, we needed to decide what range of concentration to use such that a linear or somewhat linear response will be measured. If the concentration range is too low, the plot will likely contain only noise and no discernable signal. If the range is too high, the plot may be a flat-line saturation response.

For the highest concentration, we used the stock solution of 400 ng/μL total RNA in TGB. While confident that we would detect a signal from the stock concentration, and that we probably would detect a signal one decade lower, we were unsure about measuring anything from two decades lower (4 ng/μL). Since the lowest dose was expected to produce the same response as a blank, it provided extra confirmation of the

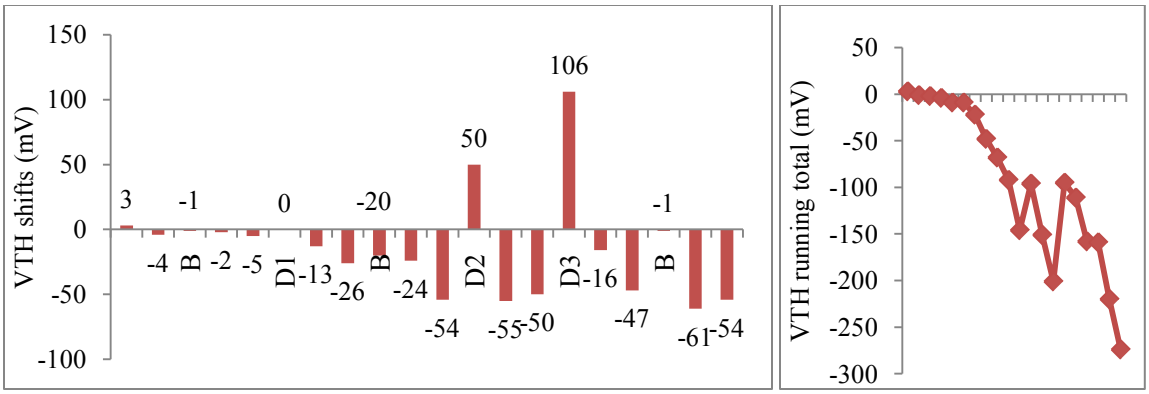


Figure 72. (left) VTH shifts and (right) VTH running total probing for HER2 in three increasing concentrations of SKBR3 total RNA: 4, 40, and 400 ng/µL (D1, D2, D3). It is immediately obvious that the trend of increasing shifts from 0 mV, to 50 mV, and finally 106 mV are an increasing function of dose. Negligible blank shifts (B) were produced throughout the experiment. Dramatic drops in VTH in the latter half of experiment may signal chip aging.

chip stability. Hybridization was performed in a volume of 25 µL of TGB for 1.5 hours at room temperature. We planned to interleave blank hybridizations of only 45 minutes between regular hybridizations. In actuality, one blank was skipped between the two highest doses due to time constraints.

As can be seen in Figure 72, increasing shifts of 0 mV, 50 mV, and 106 mV were produced from the sequentially decade increases in dose. Also, negligible blank responses were achieved throughout the experiment, although the threshold dropped rapidly near the end of experiment over repeated scans. This may signal chip aging, as there were no such drops in the beginning of experiment.

The dose-response due to HER2 is plotted in Figure 73. No blank subtraction or background subtraction was needed since the blanks were negligible. The overall dose-response relationship appears to be a logarithmic function of concentration, with a slope of about 50 mV/decade.

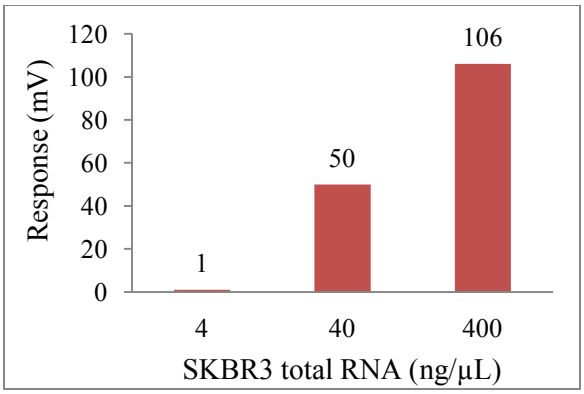


Figure 73. Dose-response curve probing for HER2 in SKBR3 cells. The overall relationship is logarithmic function of concentration, with a slope of about 50 mV/decade.

The mass, concentrations, and signals of HER2 mRNA in the present and previous experiments are listed in Table 10. This time the signal at highest concentration of HER2 was 106 mV. The signal at the same concentration was about 50 mV in the last experiment. Accounting for the significant blank response to DI water of 25 mV, the adjusted signal was 23 mV. In other words, the response this time is 4.6 times as much as last time. This may be due to the poor hybridization conditions provided by DI water, such as extremely low ionic shielding and acidic pH due to carbon dioxide absorption.

Table 10. Mass, concentrations, and signals of HER2 mRNA in previous experiment (in DI) and this experiment (in TGB).

Cell line / buffer	**HER2 (pg)**	**HER2 (fg/µL)**	**Signal (mV)**
SKBR3 / DI	20	780	23
SKBR3 / TGB	19	740	106
	2	74	50
	0.2	7.4	1

This is the first time that we had access to RNA that was eluted in top-gate buffer, which is important in reducing the blank response and providing a proper environment for hybridization. We were able to produce a consistent dose-response curve even without detailed optimization. However, a major improvement will be made. Increasing hybridization temperature to 55 °C will improve the specificity of nucleic acid base-pairing and lead to less non-specific binding.

5.3.3 Dose Response to SKBR3 at Elevated Hybridization Temperature

We next seek to produce another dose-response curve for HER2 in SKBR3, with the main improvement of an elevated hybridization temperature for better specificity. As detailed in Chapter 3, Luminex experiments showed that elevated hybridization temperature was required to achieve discrimination against single-base pair mismatches in synthetic targets. This is reason to suspect that our signals at lower temperatures overestimate the true amount of specific target capture.

We usually take a set of three scans whenever the chip is measured on the multiplexer platform following any treatment. To save time, we reduced the scans to two per set. Since the third of three scans typically did not provide any important data other than chip aging, we expected no significant consequence.

Hybridizations were performed at a temperature of 55 °C for 90 minutes. Blanks were interleaved with hybridizations and shortened to 50 minutes. We applied the same dose concentrations as we used in the room temperature dose response.

As shown in Figure 74, an increasing dose-response is seen among the two highest doses, whereas the preceding blanks are insignificant. However, the first blank (B1) was at the same level as doses D1 and D2. We compensate for the blank shifts by

Figure 74. (left) VTH shifts probing for HER2 in SKBR3 at doses of 4, 40, and 400 ng/µL (D1, D2, D3). It is immediately obvious that highest dose produced the highest response. With the exception of the first blank (B1), the blank shifts (B2, B3) were negative and of similar magnitude with scans immediately following D1 and D2.

computing the net signal as the given dose signal less the average of the "shouldering" blanks. In other words, the compensated dose signal $C_k = D_k - (B_k+B_{k+1})/2$, where D_k is the original signal number k, B_k is the previous blank and B_{k+1} is the following blank. In the case of D3, where there is no following blank, B_4 is taken as equal to B_3.

The adjusted dose-response due to HER2, plotted in Figure 75, and is steadily increasing with increasing concentration but appears to saturate across the two higher concentrations. As a consequence, we will conduct future experiments in a concentration range less than 400 ng/µL.

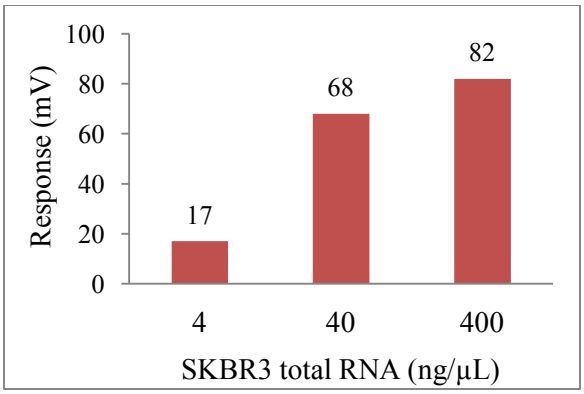


Figure 75. Dose-response curve probing for HER2 in SKBR3 cells, adjusted via shouldering blank average subtraction algorithm. Each signal is the computed as the raw shift minus the average of preceding and following blanks. The dose-response is increasing fast over the two lower concentrations but levels off across the two higher concentrations.

5.4 Determination of HER2 Overexpression in Cancer Cells

Our final goal in this project was to use the CNTFET biosensors to determine the over-expression of HER2 mRNA in real cancer cells. If so, this could pave the way for using the biosensors to determine HER2 status alongside or even competing with FISH, IHC, and RT-PCR.

5.4.1 Methods

Our colleagues at NCI had already adopted a methodology to determine over-expression of HER2 relative to a "housekeeping gene" called GAPDH, which is expressed at a normal level in both the SKBR3 and MCF7 cell lines. Sagara *et al.*[77] produced HER2 statuses via real-time PCR that were statistically significantly correlated with gene copy amplification as determined by FISH. Their method was to calculate the HER2/GAPDH ratio, which is nearly 3.74 for SKBR3, but only 0.13 for MCF7.

Table 11. Names and sequences of 16-mer anti-sense PNA probes for HER2 and GAPDH developed at NCI. The probe attachment was achieved via an thiol (SH) group at the 5' end.

Name	Sequence
HER2 Probe	5'-/SH/-ACT GGA CCC TAG AGT C-3'
GAPDH Probe	5'-/SH/-GAT GGC AAC AAT ATC C-3'

We adopt the use of the HER2/GAPDH ratio to gauge HER2 over-expression using CNTFET biosensors. This required development of PNA probes for GAPDH, which was also done with our chemist at NCI. Table 11 shows the exact probe sequences that we used to target HER2 and GAPDH.

Since we had no prior experience working with the GAPDH probe, we did not know what concentrations of total RNA to use to achieve the best results in terms of the final HER2/GAPDH ratio. We ended up choosing two concentration ranges: low and high. Our two previous HER2-only dose-response curves were taken at concentrations of 5, 50, and 500 ng/μL. We chose to limit the lowest concentration to 50 ng/μL because 5 ng/μL produced such low HER2 signals from SKBR3 (HER2+) that no signal could be expected for HER2 or GAPDH in MCF7 (HER2−). We also tempered the highest concentration from 500 ng/μL down to 250 ng/μL to avoid saturating the HER2 probe and consuming too much RNA supply.

To increase experimental throughput and reduce chip consumption, we developed the means to section the chip into two separate parts via a small thin-walled polypropylene tube (Figure 76). This allowed the immobilization of two different probes on a single chip. The probe was typically immobilized in the small reservoir at a concentration of 1 mM in 5 μL DI, and in the large reservoir at a concentration of

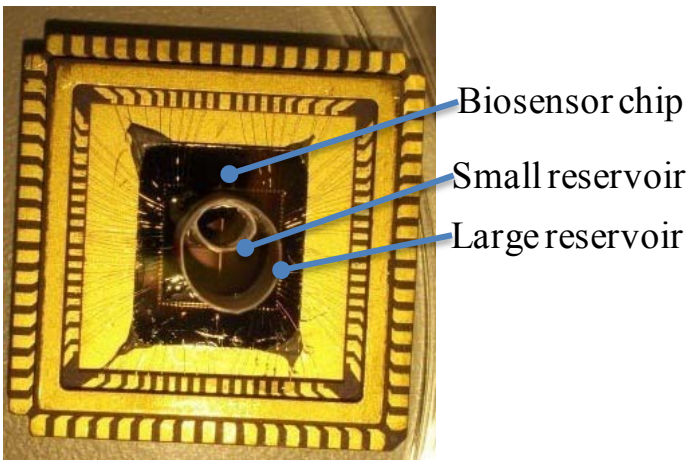


Figure 76. Biosensor with two reservoirs attached: a smaller tube within a larger tube creating two isolated liquid sections, allowing immobilization of probes for HER2 and GAPDH on different areas of the same chip. Larger reservoir is ~6 mm diameter.

0.86 mM in 5.8 µL DI. The transistor measurement scheme was modified to alternate between the top and bottom halves.

Twelve different chips were consumed for a series of experiments spanning four months. All the experiments were performed at the NCI laboratory. The hybridization incubations were performed at 55 °C for 90 minutes. We incorporated the practice of denaturing and snap cooling the RNA to increase hybridization efficiency. The RNA was denatured by heating to 95 °C for one minute, then immediately snap-cooling it in ice-water. This lessens the secondary structure of the RNA, allowing for better hybridization to the probe.

5.4.2 Aggregate Results for MCF7 & SKBR3

Figure 77 plots the dose-response curves based on adjusted data from a series of experiments done over nine different chips. An adjusted signal is generally the raw threshold shift minus the average of the shouldering blanks, as explained before. However, there are cases in which this method is not appropriate. Blanking can cause a

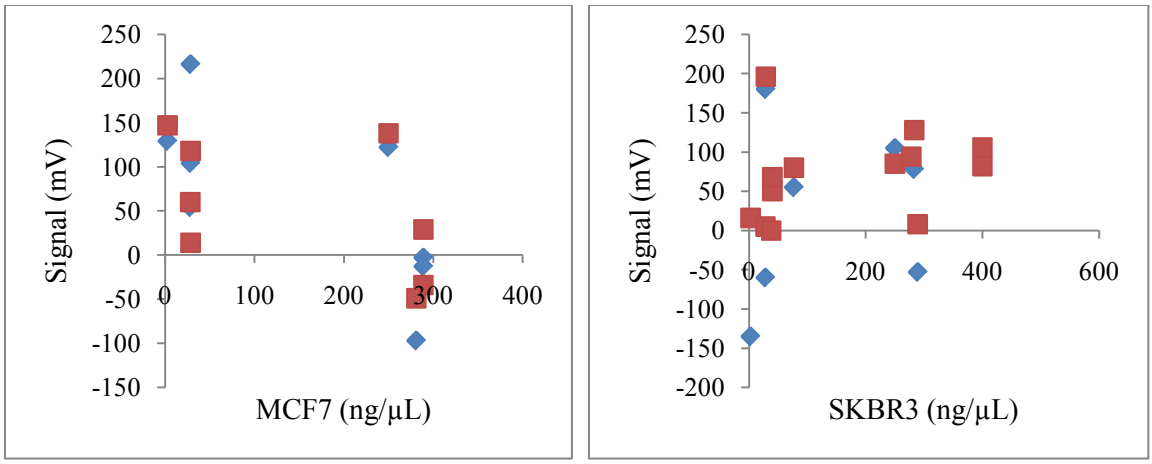


Figure 77. Adjusted signals (without filtering) probing for HER2 (red) and GAPDH (blue) in total RNA from MCF7 (left) and SKBR3 (right). Composite data from four chips on MCF7 and five chips on SKBR3.

negative shift due to the dehybridization of loosely bound RNA, or can cause large shifts due to chip instability. In this case, the data point is either invalid or the shoulder method is reverted.

Figure 78 contains the dose-response curves based on validated, adjusted signals. Careful judgment was used in determining when hybridization signals and blank responses were invalid due to chip or RNA instability. Negative or near-zero adjusted signals were deemed invalid, as were signals considered too high due to chip instability. Also, in SKBR3, the two data points at the highest concentration were removed due to saturation of the probe from experiments at lower concentrations. This data has been condensed into Figure 79, where the data are lumped into low and high concentrations.

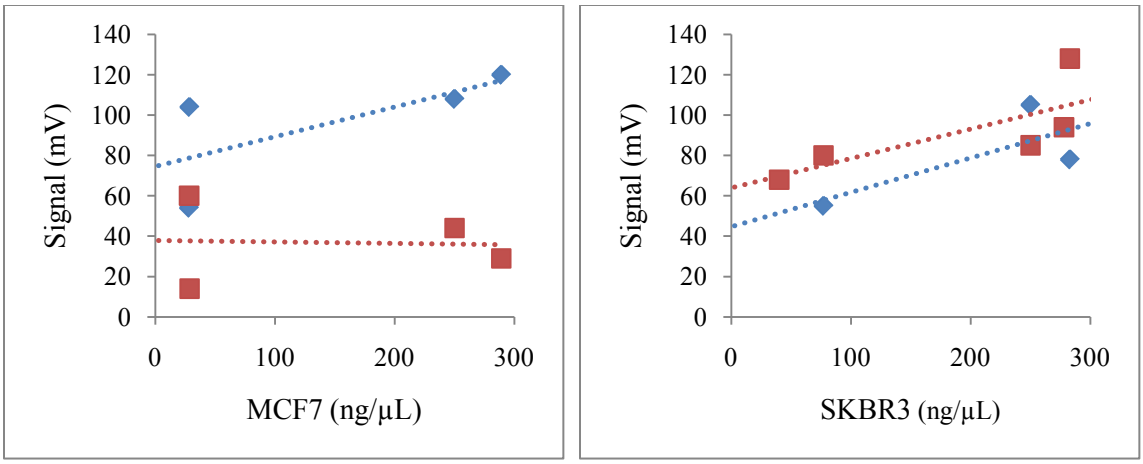


Figure 78. Validated adjusted signals probing for HER2 (red) and GAPDH (blue) in total RNA from MCF7 (left) and SKBR3 (right). The dotted lines are an aid to the eye. (left) Linear-fit dose response to GAPDH is higher than that of HER2. GAPDH content is about 8 times that of HER2 in MCF7. Composite data from four chips. (right) Linear-fit dose response to HER2 is slightly higher than that of GAPDH. HER2 is overexpressed about 4 times as much as GAPDH. Optimization is needed to reproduce the contrast between MCF7 and SKBR3 that RT-PCR was able to achieve. Composite data from five chips.

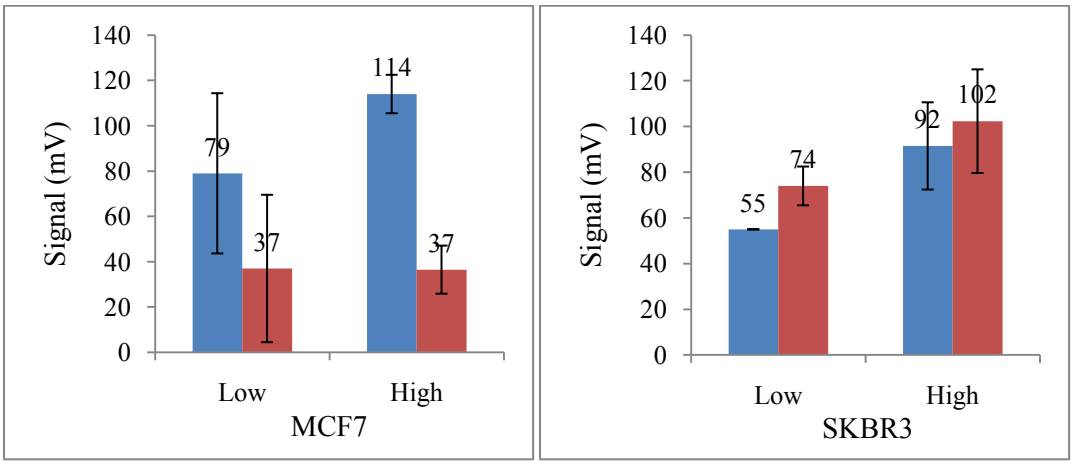


Figure 79. Condensed dose-response signals for MCF7 (left) and SKBR3 (right). GAPDH level is shown in blue, SKBR3 level is in red.

Table 12. HER2 and GAPDH responses and ratios for MCF7 cell line. All data with standard deviation indicated are for 2 samples. Large uncertainty in ratio at the higher concentration was due to the large variation in the HER2 response.

Dose	HER2 (mV)	GAPDH (mV)	HER2 / GAPDH
Low	37±33	79±35	0.47±0.46
High	37±11	114±8	0.32±0.10

Table 13. HER2 and GAPDH responses and ratios for SKBR3 cell line. All data with standard deviation indicated are for 2 samples.

Dose	HER2 (mV)	GAPDH (mV)	HER2 / GAPDH
Low	74±8	55	1.35±0.15
High	102±23	92±19	1.12±0.34

5.4.3 Calculation of HER2/GAPDH Ratios

To determine HER2 status, we calculated the HER2/GAPDH ratio for MCF7 and SKBR3 (Table 12 and Table 13). As in Figure 79, the data was grouped into low and high concentrations. For MCF7, we calculated HER2/GAPDH ratios of 0.47 and 0.32 and the low and high concentrations, respectively. For SKBR3, we obtained HER2/GAPDH ratios of 1.35 and 1.12 at the low and high concentrations, respectively. On the average, we arrive at HER2/GAPDH ratios of 0.4 and 1.24 for MCF7 and SKBR3, respectively. The contrast ratio of SKBR3:MCF7 is 1.24/0.4 = 3.1.

5.4.4 Analysis & Discussion

The results imply that our system, blindly given an unknown sample random chosen from MCF7 or SKBR3, can identify which cell line the sample came from on the basis of whether the HER2/GAPDH ratio is greater than unity. This scheme could be developed further to establish a threshold rule that works over a larger number of cell lines, and would be along the lines of commercializing the system for HER2 detection from real patient cell samples.

We do acknowledge that the experiments need be repeated several more times to provide data for a more meaningful statistical analysis such as would be seen in a clinical trial. We may also need to optimize the probe immobilization and hybridization protocols (buffers, temperatures, etc) to increase sensitivity and improve discrimination between HER2 and GAPDH. We admit that our system does not produce the same contrast between SKBR3 and MCF7 that was obtained by PCR.[77] The SKBR3:MCF7 contrast in HER2/GAPDH ratio by PCR is 3.738/0.131 = 28.5, which is nearly ten times the contrast produced by our biosensor.

The relatively low contrast may be a consequence of the fundamental logarithmic charge-to-voltage response of the biosensor as shown in Figure 64b. Most of our responses (see Figure 78) are in the range of 50 to 125 mV, which is where the logarithmic nature dominates. Thus it is not difficult to imagine that many signals that otherwise would be larger have been diminished. Low contrast may also be due to a common background signal such as non-specific adsorption of RNA onto the gold attachment surfaces. This is supported by the y-axis intercept of the interpolated dose

responses (Figure 78) at a level in the vicinity of 60 mV. Non-specific adsorption may be reduced by blocking residual active gold sites with tRNA or other agent.

The sensitivity of our biosensor may have actually been enhanced by the secondary structure (folding) of RNA, bringing more RNA charge within the Debye length. This enhancement cannot be gained in the fluorescence-based Luminex or microarray methods, since the folding does not enhance fluorescence. We could also increase overall sensitivity to HER2 if we gather more HER2 RNA by immobilizing a mixture of probes targeting different parts of the HER2 gene, rather than just a single probe sequence targeting only one part of the gene.

While our experimental platform has much room for improvement, it does avoid certain pitfalls of PCR, FISH, and other methods of detection. For example, PCR is known to have sequence artifacts and bias.[53] As a result of the greater hydrogen bonding between G-C vs A-T, sequences have a high G-C content may be over-represented in the PCR product. Our colleagues at NCI were not able to successfully use the fluorescence-bead-based Luminex XMAP to detect RNA. This may have been due to failure to attach fluorophores to label the infamously unstable RNA molecule. Though they were able to measure increased HER2 in SKBR3 relative to MCF7 via a PCR-based method (Taqman®), they did not successfully measure the HER2/GAPDH ratio.

5.5 Conclusion

We have demonstrated label-free electronic detection of HER2 mRNA in total RNA extracted from real cancer cells. Based on comparison with GAPDH mRNA levels, our prototype system distinguished which of two breast cancer cell lines is HER2-positive and which is HER2-negative. The range of total RNA concentration was from

50 ng/μL to 250 ng/μL, corresponding to the range of 50,000 to 500,000 cancer cells. This number of cells could easily come from a core biopsy, obviating the need for a larger sample via surgery. We believe these achievements set the stage for development of a rapid, low-cost HER2 biosensor that could serve to augment or even compete with existing expensive and time-consuming HER2 testing methods.

Chapter 6: Future Work & Main Conclusions

6.1 Future Work

6.1.1 Improvement of HER2 Selectivity & Sensitivity

As shown in the previous chapter, we were able to detect a contrast between two different breast cancer cell lines based on the relative expression levels of HER2 with reference to a housekeeping gene. However, our system did not achieve the same level of contrast obtained by Q-RT-PCR. As shown in Chapter 3, the probe was verified to be highly selective against two base-pair mismatches. The remaining weakness in our probing strategy is the lack of any post-immobilization blocking protocol to prevent non-specific adsorption onto the metal (Au) gate.

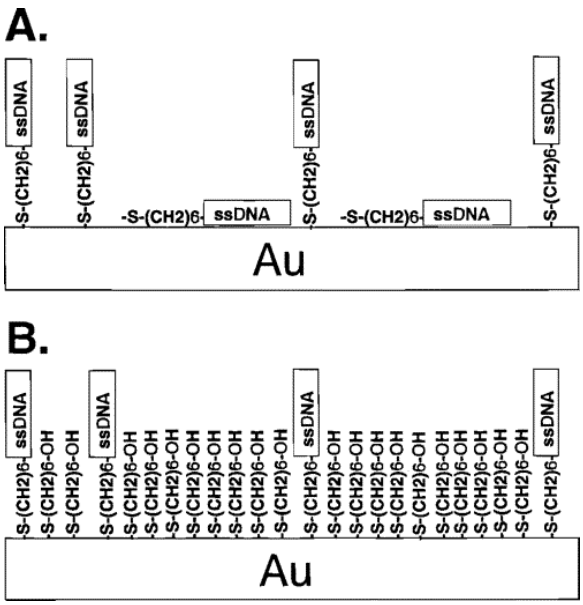


Figure 80. Schematic of (A) HS-ssDNA on Au and (B) both HS-ssDNA and MCH adsorbed on gold. The ssDNA probes are attached to the gold surface through a sulfur-gold linkage. Source: Herne & Tarlov.[74]

As shown in Figure 80, when gold is exposed to thiolated probe DNA, the probe nucleotides will attach through the thiol-gold linkage. As expected, most of the probe molecules will be oriented perpendicular to the surface so as to be accessible for hybridization. However, Steel *et al.*[65] noted that DNA nucleotides can presumably adsorb to gold via multiple amine moieties. Thus some of the probe molecules may lay flat and be unavailable for hybridization, causing reduction in the hybridization signal and increasing the likelihood of false negatives. Furthermore, some fraction of the target molecules may non-specifically adsorb to the CNTFET gate, causing false positives.

Unreacted surface sites on gold can be blocked with 6-mercapto-1-hexanol (MCH). The benefits are twofold. The MCH rapidly displaces weakly adsorbed probe from the substrate, leaving the probes tethered through the thiol end groups. The alcohol (OH) end groups of MCH tend to repel nucleic acids and do not allow for non-specific adsorption of nucleic acids. Moreover, the use of MCH as a space filler enhances the accessibility of the probes for hybridization, allowing nearly complete hybridization efficiencies.[64,65]

The next improvement in the probing technique is to increase the sensitivity to HER2. So far we have only employed a single 16-mer probe sequence for HER2. However, the HER2 gene is about 41,000 base pairs long, and our collaborators at NCI already selected and fabricated several other probe sequences for HER2. We anticipate that mixing multiple HER2 probes on each biosensor would result in the capture of more parts of the same gene and yield higher hybridization signals.

An alternative to seeking higher signals is to reduce the noise interfering with small signals. Along these lines, we should attempt to improve the uniformity of the

transistor characteristics or to come up with a clever scheme to calibrate the individual signals despite the transistors' nonuniformity.

6.1.2 Single-Walled CNTFET vs. NTNFETs

Currently our sensors are based on spaghetti-like nanotube mats. Each device may have several tubes with different metallic or semiconducting properties in unknown series and parallel combinations.

Star *et al.*[13] saw several advantages to carbon nanotube network field-effect transistors (NTNFETs). The random network avoids the challenges of alignment and averages out inter-device variations due to chirality and precise geometry. Devices can be developed on low-cost flexible polymer substrates by spray deposition, spin-coating, or nanotransfer printing.[9]

To fully utilize the outstanding electronic properties of the nanotubes, well-defined CNTFETs could be made out of aligned single-walled nanotubes grown on quartz.[78] A single-tube CNTFET would be expected to have superior intrinsic sensitivity since it is a true nanoscale sensor. However, it is still unclear if this noble venture would be worthwhile. Nanotubes exhibit a significant level of $1/f$ noise[12] that is inversely proportional to the number of CNTs in the device. A network with a large number of SWNTs reduces the $1/f$ noise by approximately the number of SWNTs raised to the (-1.3) power.[79] In the overall biosensor system, further study is needed to identify the most severe noise sources. If $1/f$ noise is a significant contributor, then large nanotube networks will have a distinct advantage over single-channel devices.

6.1.3 System Integration, Expansion, & Commercialization

Our vision for the future is for a cancer patient to walk into the clinic, provide a biopsy sample, and receives her HER2 results in the same visit. We believe this could only be accomplished with a highly integrated laboratory-on-a-chip system driven by microfluidics. For example, the biopsied tissue sample contains a variety of cell types, such as ductal, lobular, lymph, and adipose (fat). The HER2 status should be taken from the primary cancer cells. A practical application of microfluidics would be to segregate all the different cells in the sample into cancerous and non-cancerous cell types. Then the cancer cells could be lysed, possibly with electric field, to aid in chemical extraction of the RNA. If for some application, PCR amplification is needed, this could be integrated on-chip. Or, a separate stage could separate mRNA from total RNA as a pre-treatment to enhance detection of HER2 mRNA. Other benefits of microfluidics include faster throughput, reduced reagent consumption, and expansion into series/parallel analysis of the same input sample.[80]

Another obvious area of future work is the integration of the detection system with peripheral circuit blocks for multiplexing, signal conditioning/calibration, and final readout. Alternatively portions of the signal processing and final readout/analysis can be done on a notebook computer.

We believe that this work will pave the way for commercialization of a low-cost, rapid cancer detection system that would be licensed as a platform for development of protocols to detect a variety of other diseases, such as dengue, typhoid, and *E. coli*. The method of detection is so versatile that it only needs the proper probe and accompanying protocol to be applied to another new challenge.

6.2 Main Conclusions

We developed a prototype biosensor platform that successfully discriminated between a HER2+ cell line and a HER2− cell line. Each biosensor required no more than ~10 μg of total RNA extracted from no more than ~500,000 cancer cells. This number of cells could be obtained in a core biopsy from a living patient, obviating the need for a larger sample via surgery. To the best of our knowledge, this system is the first CNTFET-based system to directly detect unlabelled HER2 mRNA at clinically relevant levels. This was facilitated in large part due to the use of electrically neutral peptide nucleic acids (PNA), which allow the use of much lower ionic shielding as compared to conventional DNA probes.

Our biosensors are also unique in the use of a gate dielectric and metallic gate surface to isolate the carbon nanotube mats from the hybridization reaction. While sensitivity is traded for selectivity, the aluminum oxide dielectric virtually eliminates the hysteresis caused by water molecules and provides superior resistance against positive mobile ion (sodium, potassium) penetration as compared to silicon oxide. The metallic gate further helps to prevent ionic contamination and also provides a chemically stable surface for probe attachment.

Future incremental improvements may include (i) reducing the device-to-device variation or developing a calibration scheme to overcome this non-uniformity, (ii) employing a set of HER2 probes (as opposed to a single HER2 probe sequence) to increase the hybridization signal, and (iii) optimizing the surface chemistry to increase selectivity. More significant improvements may be oriented at the system level, including integration with microfluidics as well as on-chip measurement circuitry. Such a versatile

platform could easily be expanded to detect other pathogens & biomarkers through an arrangement in which the platform is licensed to companies specializing in assay development. We believe that we have lowered the barrier toward commercialization of a rapid, low-cost, gene expression biosensor.

Appendices

7.1 Custom Instrumentation & Software

A custom measurement and analysis system was designed to allow rapid acquisition of current-voltage (I–V) curves of an array of field-effect transistors. The hardware instrumentation provides the gate and drain voltages and measures the respective currents of one CNTFET at a time. The software component consists of a Windows program to control the measurement routine and a MATLAB program to analyze the collected data.

7.1.1 Hardware Instrumentation

As shown in Figure 81, the hardware instrumentation consists of several devices

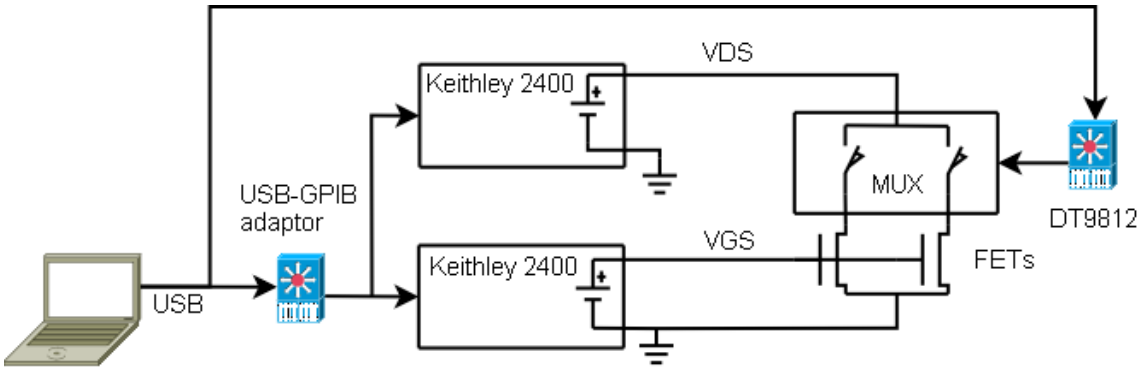


Figure 81. Instrumentation block diagram. The system is controlled by a notebook PC via USB. Two Keithley 2400 SourceMeters provide the drain and gate voltages (V_{DS}, V_{GS}) to the CNTFETs, while an analog multiplexer (MUX) selects one transistor at any given time. The PC interfaces to the SourceMeters via a USB-GPIB adaptor, and to the analog multiplexer via a USB digital I/O module (DT9812). The 64-channel multiplexer is based on two integrated circuit 32-channel analog multiplexers.

controlled by a notebook PC via USB. Two Keithley 2400 SourceMeters provide the drain and gate voltages to the CNTFETs, while an analog multiplexer selects one transistor at any given time. The PC interfaces to the SourceMeters via a USB 2.0 to GPIB (IEEE-488) controller module (488-USB2, ICS Electronics). The PC also interfaces with the analog multiplexer via the digital input/output function of a low cost USB data acquisition module (DT9812, Data Translation). The 64-channel multiplexer is based on two integrated circuit 32-channel analog multiplexers (ADG732BCP, Analog Devices).

Figure 82 shows one of three actual hardware instrumentation setups that we deployed at NCI in Gaithersburg, Maryland. The analog multiplexers and the socket for the biosensor chip were wired together on a hand-wired prototyping breadboard. The biosensor plugs into a socket (3M) in the middle of the measurement platform, and is held in place by a custom-machined "U" clamp. A plastic cap attaches to the socket and houses the top-gate electrodes. The multiplexer ICs in surface-mount Quad Flat Packages (QFP) were hand-soldered to adapter boards (ePBoard) for use in standard "dual in-line package" (DIP) sockets on the breadboard, which also contains the connectors for the SourceMeter and digital I/O cables.

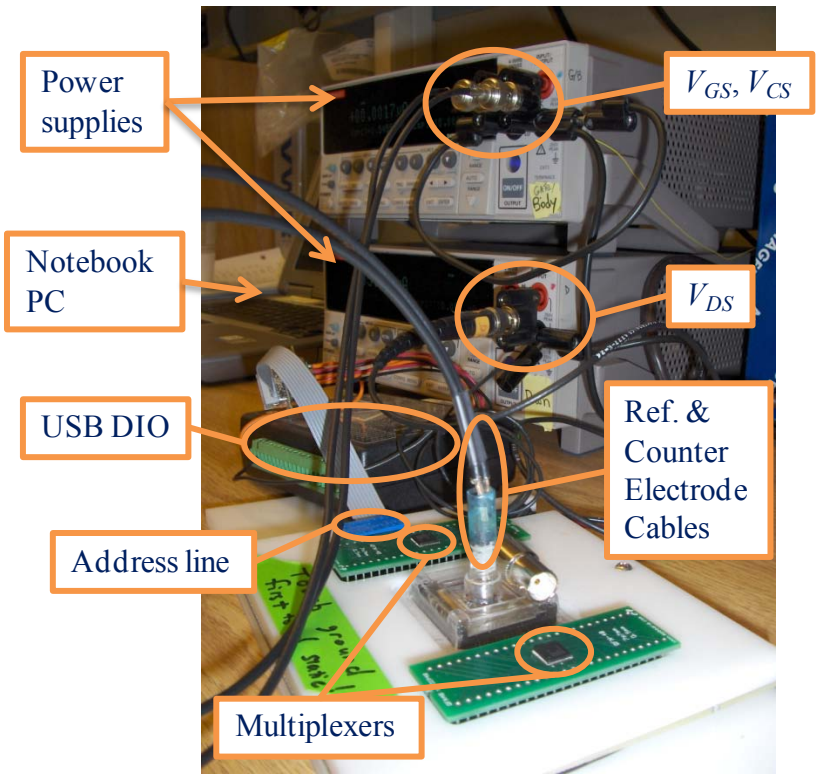


Figure 82. Custom measurement system including (from foreground to background) measurement platform, USB digital I/O box, power supplies, and notebook PC. The chip carrier plugs into the middle of the measurement platform and is covered by a clear protective cap that houses the top gate electrodes. The green printed circuit boards adapt the multiplexers in surface mount packages for use in "dual in-line package" (DIP) sockets. The note on green tape says, "Touch ground first to rid static!"

7.1.2 Measurement Control Software

We wrote custom software to control the batch measurement of transfer characteristics of CNTFETs over the entire chip and store the data to disk for further analysis. The program is called "BMBatch", which stands for "Basic Measure Batch". This program originally had a simple text-mode interface but was upgraded to include a graphical user interface (GUI) as a courtesy by Michael Dreyer at LPS.

The BMBatch user interface, shown in Figure 83, includes real-time plots of the gate and drain I–V curves, as well as controls for automating batch measurement and adjusting critical configuration details. Batch measurement of multiple devices is

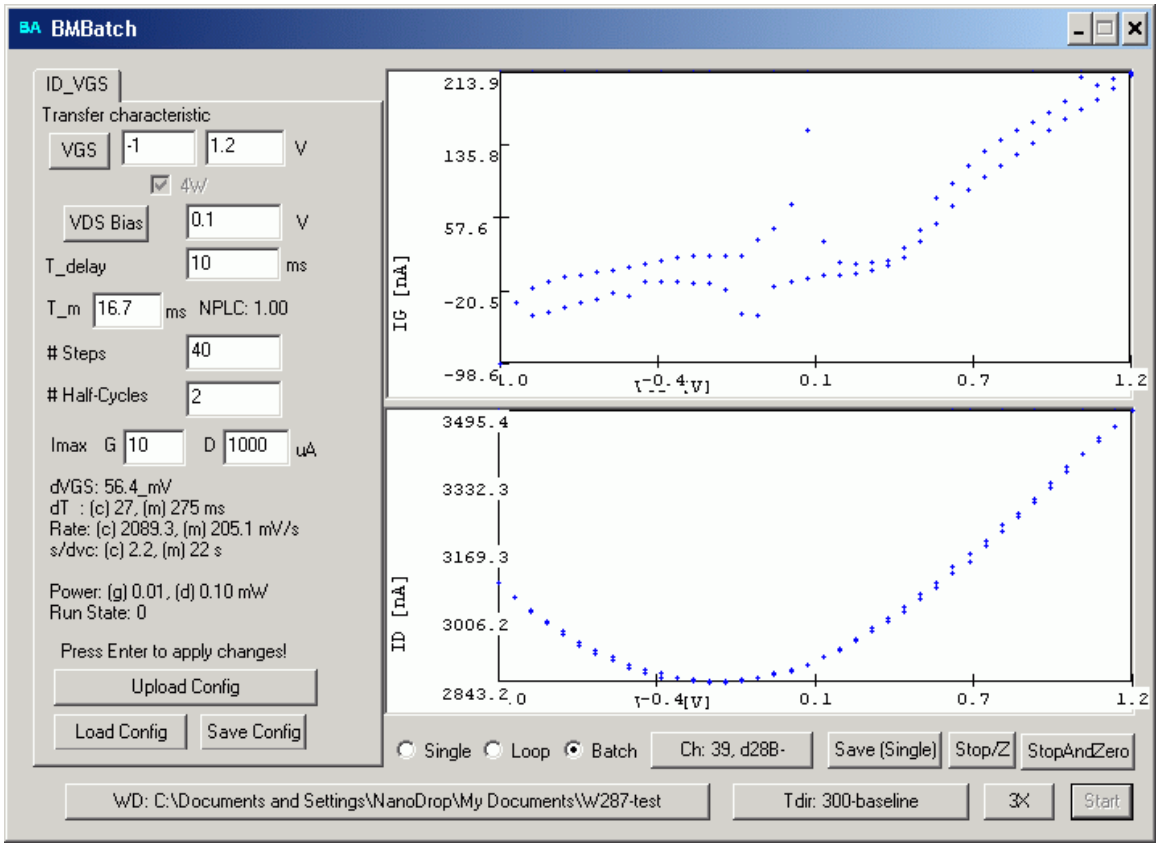


Figure 83. Screenshot of BMBatch custom I–V curve measurement software. On the right hand side are two plots of I–V curves, one for the gate and one for the drain. The top plot shows the gate current (I_G–V_{GS}) for diagnostic purposes. The bottom plot is the transfer characteristic (I_D–V_{GS}) recorded for the currently addressed CNTFET. The pane on the left hand side contains configuration parameters such as voltage range and number of data points. Other parameters such as delay time, measurement time, and current limits are uploaded to the SourceMeters. The bottom section contains buttons for starting and stopping the measurement, as well as options for batch scanning, such as the working directory (WD), current trial directory name (Tdir), and number of repeat scans (3X). Written in C++ using Microsoft Visual Studio 6.0 (1998).

accomplished by loading a batch description text file wherein each line specifies the multiplexer channel number, device number, and molecular probe. For example, “39,d28B-“ indicates channel 39, device 28, probe B. The transfer characteristics will then be saved to disk as comma-separated values (CSV) for analysis in MATLAB in a directory structure as shown in Figure 84.

7.1.3 MATLAB Analysis

After data is acquired via BMBatch, a suite of MATLAB scripts called “ShiftVthBatch” plots the transfer characteristics of all transistors stored in all the trial directories. As shown in Figure 85, a parabola is fitted to the five lowest data points of

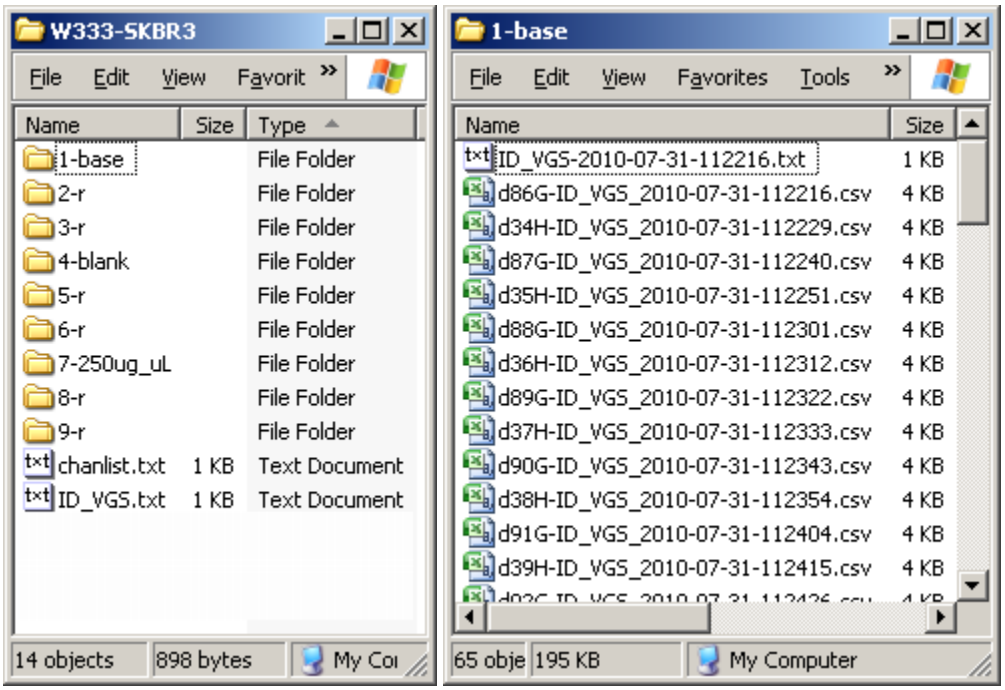


Figure 84. Example directory structure where BMBatch stores acquired data. In analysis, ShiftVTHbatch.m traverses these directories. (left) Chip-level directory, including 9 trial subdirectories, channel-device specification, and measurement configuration. (right) Contents of one trial directory include a configuration file and a CSV file for each device.

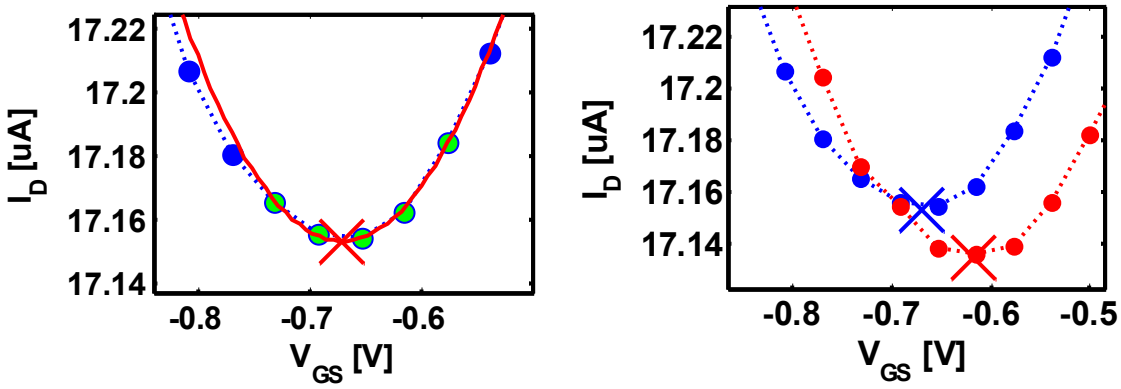


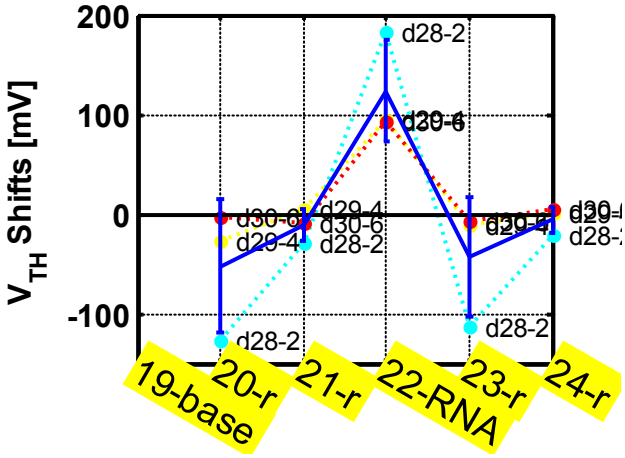
Figure 85. (left) Zoom of region of transfer characteristic used for V_{TH} interpolation. For each transistor, a parabola (red) is fitted to a subset (green) of five points from the original data (blue). The interpolated threshold voltage (red X) is then stored in an array. (right) Transfer characteristics of the same transistor from two consequent trials (1: blue, 2: red). The threshold shift for this FET at trial 2 will be computed as the increase in threshold from trial 1 to trial 2.

each curve, which is warranted because when $V_{GS} \approx V_{TH}$, the transistor is in the saturation regime, and Equation (1) gives:

$$I_D \propto (V_{GS} - V_{TH})^2 \tag{41}$$

The data is fit in a least-squares sense with a second order polynomial using the MATLAB function "polyfit", which returns the coefficients for a parabola of the form: $y = aV_{GS}^2 + bV_{GS} + c$. The interpolated threshold voltage given by $V_{TH} = -\frac{b}{2a}$. Each V_{TH} is stored in an array location corresponding to the trial number and device number.

The threshold shift of one FET at any trial is the increase in threshold from the previous trial. As shown in Figure 86, these shifts are plotted as a function of trial number. The mean V_{TH} shift for the group of FETs at every trial is computed and plotted. The user is given an opportunity to revise the set of transistors that are included in the analysis and the revised data is redisplayed. After successive iterations have been



Num	Trial	mean	stdev
20	r	-52	66
21	r	-11	17
22	RNA	124	51
23	r	-43	60
24	r	-5	13

Figure 86. The V_{TH} shifts of three transistors (yellow, red, cyan) as a function of trial number. The mean shift (blue) and standard deviation (error bars) characterize the group behavior. From this plot, the user can decide which transistors should be excluded from analysis. Once the user is satisfied, the summary data (mean shift and standard deviation) is saved to a spreadsheet file (right).

completed to the satisfaction of the user, the relevant group data (such as mean and standard deviation of the shifts versus trial number) are written to an Excel spreadsheet.

7.2 UHV Iron Evaporation

Chip throughput was originally slow because only two chips could be loaded in parallel on UHV system and pump-down takes about one hour. Throughput was more than doubled after we adopted a multi-chip holder developed at LPS (Figure 87).

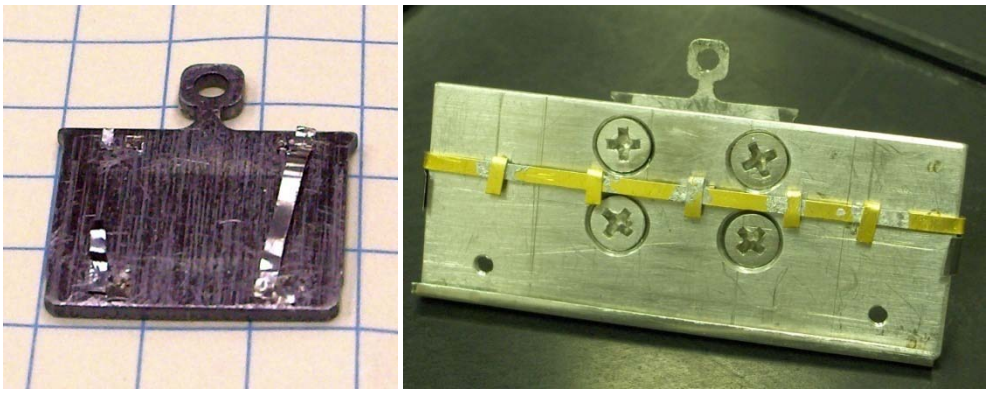


Figure 87. (left) Original UHV chip holder for one chip only (two holders could be loaded for deposition simultaneously). (right) Scaled up UHV chip holder cradling five chips per deposition cycle.

7.3 **Reservoir Attachment Rig**

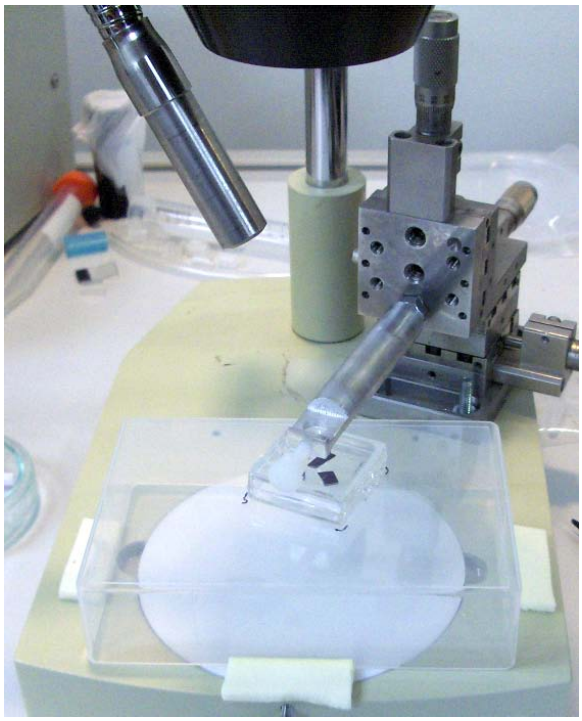


Figure 88. Cup attachment rig (courtesy Prof. Gomez). This setup gives the user a microscopic bird's eye view to precisely position the adhesive-laden reservoir over the active area of the chip without bumping into the fragile bond wires.

List of Publications

1. Konrad H. Aschenbach, Herman Pandana, Jookyung Lee, Javed Khan, Michael Fuhrer, Dan Lenski, and Romel D. Gomez, "Detection of nucleic acid hybridization via oxide-gated carbon nanotube field-effect transistors," *Proceedings of SPIE* (presented at the Micro (MEMS) and Nanotechnologies for Space, Defense, and Security II, Orlando, FL, USA), pp. 69590W-69590W-17, 2008. DOI: 10.1117/12.778531
2. Herman Pandana, Konrad H. Aschenbach, Daniel R. Lenski, Michael S. Fuhrer, Javed Khan, and Romel D. Gomez, "A Versatile Biomolecular Charge-Based Sensor Using Oxide-Gated Carbon Nanotube Transistor Arrays," *IEEE Sensors Journal*, vol. 8, no. 6, pp. 655-660, 2008. DOI: 10.1109/JSEN.2008.922724
3. Herman Pandana, Konrad H. Aschenbach, and Romel D. Gomez, "Systematic Aptamer-Gold Nanoparticle Colorimetry for Protein Detection: Thrombin," *IEEE Sensors Journal*, vol. 8, no. 6, pp. 661-666, 2008. DOI: 10.1109/JSEN.2008.922725
4. Li Gan, Seok Hwan Chung, Konrad H. Aschenbach, Michael Dreyer, and Romel D. Gomez, "Pulsed-current-induced domain wall propagation in Permalloy patterns observed using magnetic force microscope," *Magnetics, IEEE Transactions on*, vol. 36, no. 5, pp. 3047-3049, 2000. DOI: 10.1109/20.908674

Curriculum Vitae

Education

Ph.D. in Electrical Engineering, 2/2011
Development of Carbon Nanotube Field-Effect Transistor Arrays for Detection of HER2 Overexpression in Breast Cancer
Professor Romel D. Gomez
University of Maryland, College Park

M.S. in Electrical Engineering, 12/2003
Low-Cost Distributed Solar-Thermal-Electric Power Generation
Professor Seth R. Sanders
University of California, Berkeley

B.S. in Electrical Engineering, Magna cum Laude, 05/2001
University of Maryland, College Park

Experience

Graduate Research Assistant, 01/2004 - Present
University of Maryland, ECE Department

- Developed prototype biosensor platform based on carbon nanotube field-effect transistor arrays for rapid, cost-effective detection of HER2 breast cancer gene.
- Designed microlithographic masks and optimized fabrication process to create supply of reliable, consistent devices.
- Optimized molecular biology protocols for HER2 probe immobilization and detection of HER2 mRNA from cancer cells.
- Built hardware and wrote software for automated, multiplexed data acquisition and analysis.
- Developed theoretical models for detection mechanism.
- Demonstrated efficacy of technique on real cancer cell lines at clinically relevant sensitivity levels with appropriate controls.

Graduate Research Assistant, 05/2002 – 12/2003
University of California, Berkeley, EECS Department

- Determined cost per watt advantage of non-tracking solar thermal electricity system over conventional photovoltaics.
- Optimized requirements of non-imaging compound parabolic reflector and absorber for fastest payback period.
- Studied feasibility of employing natural convection to drive fluid flow from solar heat collectors to Stirling engine.

Undergraduate Research Assistant, 09/1999 – 05/2001
University of Maryland, ECE Department

- Imaged biological samples using self-built long-range scanning tunneling microscope interfaced with commercial probe microscopy system.
- Built motorized sub-millimeter resolution Hall-effect magnetic imager and programmed data acquisition via Turbo C and visualization via MATLAB.

Publications

Aschenbach, K.H.; Pandana, H.; Lee, J.; Khan, J.; Fuhrer, M.; Lenski, D.; and Gomez, R.D., "Detection of Nucleic Acid Hybridization via Oxide-Gated Carbon Nanotube Field-Effect Transistors," Invited, *Proc. SPIE*, **6959**, (2008).

Pandana, H.; Aschenbach, K.H.; Lenski, D.R.; Fuhrer, M.S.; Khan, J.; Gomez, R.D., "A Versatile Biomolecular Charge-Based Sensor Using Oxide-Gated Carbon Nanotube Transistor Arrays," *Sensors Journal, IEEE*, **8**, 655 (2008).

Pandana, H.; Aschenbach, K.H.; Gomez, R.D., "Systematic Aptamer-Gold Nanoparticle Colorimetry for Protein Detection: Thrombin," *Sensors Journal, IEEE*, **8**, 661 (2008).

Der Minassians, A; Aschenbach, K.H.; Sanders, S.R., "Low-Cost Distributed Solar-Electric Technology," Invited, *Proc. SPIE*, **5185**, 89 (2003).

Gan, L.; Chung, S.H.; Aschenbach, K.H.; Dreyer, M.; Gomez, R.D., "Pulsed-current-induced domain wall propagation in Permalloy patterns observed using magnetic force microscope," *Magnetics, IEEE Transactions on*, **36,** 3047 (2000).

Patent

Gomez, R.D.; Khan, J.; Pandana, H.; Aschenbach, K.; Fuhrer, M.; Wei, J. S., "Apparatus for Microarray Binding Sensors Having Biological Probe materials using Carbon Nanotube Transistors"**,** U.S. Patent Application 20080035494 (2006).

Awards

Electrical & Computer Engineering Distinguished Dissertation Fellowship, 2009–2010.

BIO IT Coalition Scholarship Award, 04/2008.

University of California Regents Fellowship, 09/2001 – 05/2002.

Department of ECE Chair's Award, 2001.

Departmental Honors in Electrical Engineering, 2001.

University Honors Citation, 2001.

Lockheed Martin Academic Excellence Scholarship, 2000/09 – 2001/05.

Consulting Engineers Council of Maryland Scholarship, 2000.

Society of American Military Engineers Scholarship, 1999 & 2000.

Dean's Scholarship Award, 1997-1999

Bibliography

1. Iijima, S. Helical microtubules of graphitic carbon. *Nature* **354**, 56-58 (1991).
2. Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. Surface Studies by Scanning Tunneling Microscopy. *Phys. Rev. Lett.* **49**, 57 (1982).
3. Wildöer, J.W.G., Venema, L.C., Rinzler, A.G., Smalley, R.E. & Dekker, C. Electronic structure of atomically resolved carbon nanotubes. *Nature* **391**, 59-62 (1998).
4. Odom, T.W., Huang, J., Kim, P. & Lieber, C.M. Atomic structure and electronic properties of single-walled carbon nanotubes. *Nature* **391**, 62-64 (1998).
5. Tans, S.J., Verschueren, A.R.M. & Dekker, C. Room-temperature transistor based on a single carbon nanotube. *Nature* **393**, 49-52 (1998).
6. Martel, R., Schmidt, T., Shea, H.R., Hertel, T. & Avouris, P. Single- and multi-wall carbon nanotube field-effect transistors. *Appl. Phys. Lett.* **73**, 2447 (1998).
7. Dürkop, T., Getty, S.A., Cobas, E. & Fuhrer, M.S. Extraordinary Mobility in Semiconducting Carbon Nanotubes. *Nano Letters* **4**, 35-39 (2004).
8. MOSIS file tsmc-035/t48y_lo_epi-params.txt. at <http://www.mosis.com/cgi-bin/cgiwrap/umosis/swp/params/tsmc-035/t48y_lo_epi-params.txt>
9. Hines, D.R. et al. Nanotransfer printing of organic and carbon nanotube thin-film transistors on plastic substrates. *Appl. Phys. Lett.* **86**, 163101 (2005).
10. Kong, J. et al. Nanotube Molecular Wires as Chemical Sensors. *Science* **287**, 622-625 (2000).
11. Collins, P.G., Bradley, K., Ishigami, M. & Zettl, A. Extreme Oxygen Sensitivity of Electronic Properties of Carbon Nanotubes. *Science* **287**, 1801-1804 (2000).
12. Collins, P.G., Fuhrer, M.S. & Zettl, A. 1/f noise in carbon nanotubes. *Appl. Phys. Lett.* **76**, 894 (2000).
13. Star, A. et al. Label-free detection of DNA hybridization using carbon nanotube network field-effect transistors. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 921 -926 (2006).
14. Gui, E.L. et al. DNA Sensing by Field-Effect Transistors Based on Networks of Carbon Nanotubes. *Journal of the American Chemical Society* **129**, 14427-14432 (2007).
15. Maehashi, K., Matsumoto, K., Kerman, K., Takamura, Y. & Tamiya, E. Ultrasensitive Detection of DNA Hybridization Using Carbon Nanotube Field-Effect Transistors. *Jpn. J. Appl. Phys.* **43**, L1558-L1560 (2004).
16. Martel, R. et al. Ambipolar Electrical Transport in Semiconducting Single-Wall Carbon Nanotubes. *Phys. Rev. Lett.* **87**, 256805 (2001).
17. Bergveld, P. Thirty years of ISFETOLOGY: What happened in the past 30 years and what may happen in the next 30 years. *Sensors and Actuators B: Chemical* **88**, 1-20 (2003).
18. Iijima, S. & Ichihashi, T. Single-shell carbon nanotubes of 1-nm diameter. *Nature* **363**, 603-605 (1993).
19. Appenzeller, J. et al. Field-Modulated Carrier Transport in Carbon Nanotube Transistors. *Phys. Rev. Lett.* **89**, 126801 (2002).
20. Cheng, D.K. *Field and Wave Electromagnetics*. (Addison-Wesley: 1989).

21. Kim, W. et al. Synthesis of Ultralong and High Percentage of Semiconducting Single-walled Carbon Nanotubes. *Nano Letters* **2**, 703-708 (2002).
22. Saito, R., Fujita, M., Dresselhaus, G. & Dresselhaus, M. Electronic structure and growth mechanism of carbon tubules. *Materials Science and Engineering: B* **19**, 185-191 (1993).
23. Binnig, G., Quate, C.F. & Gerber, C. Atomic Force Microscope. *Phys. Rev. Lett.* **56**, 930 (1986).
24. Hafner, J.H., Cheung, C., Oosterkamp, T.H. & Lieber, C.M. High-Yield Assembly of Individual Single-Walled Carbon Nanotube Tips for Scanning Probe Microscopies. *The Journal of Physical Chemistry B* **105**, 743-746 (2001).
25. Brintlinger, T. et al. Rapid imaging of nanotubes on insulating substrates. *Appl. Phys. Lett.* **81**, 2454 (2002).
26. Raider, S.I., Gregor, L.V. & Flitsch, R. Transfer of Mobile Ions from Aqueous Solutions to the Silicon Dioxide Surface. *J. Electrochem. Soc.* **120**, 425-431 (1973).
27. Derbenwick, G.F. Mobile ions in SiO2: Potassium. *J. Appl. Phys.* **48**, 1127 (1977).
28. Kim, W. et al. Hysteresis Caused by Water Molecules in Carbon Nanotube Field-Effect Transistors. *Nano Letters* **3**, 193-198 (2003).
29. Gray, P.R., Hurst, P.J., Lewis, S.H. & Meyer, R.G. *Analysis and Design of Analog Integrated Circuits*. (Wiley: 2001).
30. Ishigami, M. et al. Hooge's constant for carbon nanotube field effect transistors. *Applied Physics Letters* **88**, 203116 (2006).
31. Hassibi, A., Navid, R., Dutton, R.W. & Lee, T.H. Comprehensive study of noise processes in electrode electrolyte interfaces. *Journal of Applied Physics* **96**, 1074-1082 (2004).
32. What You Need To Know About™ Breast Cancer - National Cancer Institute. at <http://www.cancer.gov/cancertopics/wyntk/breast/allpages>
33. SEER Cancer Statistics Review, 1975-2007. at <http://seer.cancer.gov/csr/1975_2007/>
34. Slamon, D. et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* **244**, 707-712 (1989).
35. Slamon, D. et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235**, 177-182 (1987).
36. Stark, A. et al. HER-2/neu Amplification in Benign Breast Disease and the Risk of Subsequent Breast Cancer. *Journal of Clinical Oncology* **18**, 267 (2000).
37. Guideline Summary: American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Human Epidermal Growth Factor Receptor HER2 Testing in Breast Cancer. *Journal of Oncology Practice* **3**, 48 -50 (2007).
38. Vinatzer, U. et al. Expression of HER2 and the Coamplified Genes GRB7 and MLN64 in Human Breast Cancer: Quantitative Real-time Reverse Transcription-PCR as a Diagnostic Alternative to Immunohistochemistry and Fluorescence In situ Hybridization. *Clinical Cancer Research* **11**, 8348 -8357 (2005).
39. Alberts, B. *Molecular biology of the cell*. (Garland Science: New York, 2002).
40. Dako HER2 FISH pharmDx™ Kit - Principle of FISH. at <http://lindblad.marketport.dk/learningzone/previewpage.asp?presentationid=6&group=10000&pageid=1000>

41. Vogel, U. Confirmation of a low HER2 positivity rate of breast carcinomas - limitations of immunohistochemistry and in situ hybridization. *Diagnostic Pathology* **5**, 50 (2010).
42. DNA-RNA-Protein. at <http://nobelprize.org/educational/medicine/dna/index.html>
43. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science* **270**, 467-470 (1995).
44. Yager, P. et al. Microfluidic diagnostic technologies for global public health. *Nature* **442**, 412-418 (2006).
45. Bergveld, P. Development of an Ion-Sensitive Solid-State Device for Neurophysiological Measurements. *Biomedical Engineering, IEEE Transactions on* **BME-17**, 70-71 (1970).
46. DNA - Wikipedia, the free encyclopedia. at <http://en.wikipedia.org/wiki/DNA>
47. Poghossian, A., Cherstvy, A., Ingebrandt, S., Offenhäusser, A. & Schöning, M. Possibilities and limitations of label-free detection of DNA hybridization with field-effect-based devices. *Sensors and Actuators B: Chemical* **111-112**, 470-480 (2005).
48. Souteyrand, E. et al. Direct Detection of the Hybridization of Synthetic Homo-Oligomer DNA Sequences by Field Effect. *The Journal of Physical Chemistry B* **101**, 2980-2985 (1997).
49. Fritz, J., Cooper, E.B., Gaudet, S., Sorger, P.K. & Manalis, S.R. Electronic detection of DNA by its intrinsic molecular charge. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14142 -14146 (2002).
50. Perkins, F. et al. An active microelectronic transducer for enabling label-free miniaturized chemical sensors. *Electron Devices Meeting, 2000. IEDM Technical Digest. International* 407-410 (2000).doi:10.1109/IEDM.2000.904342
51. Pouthas, F. et al. Spatially resolved electronic detection of biopolymers. *Phys. Rev. E* **70**, 031906 (2004).
52. Ingebrandt, S. et al. Label-free detection of single nucleotide polymorphisms utilizing the differential transfer function of field-effect transistors. *Biosensors and Bioelectronics* **22**, 2834-2840 (2007).
53. Acinas, S.G., Sarma-Rupavtarm, R., Klepac-Ceraj, V. & Polz, M.F. PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *Appl Environ Microbiol* **71**, 8966-8969 (2005).
54. Nielsen, P., Egholm, M., Berg, R. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **254**, 1497-1500 (1991).
55. Watson, J.D. & Crick, F.H.C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737-738 (1953).
56. Demidov, V. et al. Stability of peptide nucleic acids in human serum and cellular extracts. *Biochemical Pharmacology* **48**, 1310-1313 (1994).
57. Peptide nucleic acid - Wikipedia, the free encyclopedia. at <http://en.wikipedia.org/wiki/Peptide_nucleic_acid>
58. Kaihatsu, K., Janowski, B.A. & Corey, D.R. Recognition of Chromosomal DNA by PNAs. *Chemistry & Biology* **11**, 749-758 (2004).
59. Weiler, J., Gausepohl, H., Hauser, N., Jensen, O.N. & Hoheisel, J.D. Hybridisation

based DNA screening on peptide nucleic acid (PNA) oligomer arrays. *Nucleic Acids Research* **25**, 2792 -2799 (1997).
60. Jacob, A., Brandt, O., Stephan, A. & Hoheisel, J.D. Peptide nucleic acid microarrays. *Methods Mol. Biol* **283**, 283-293 (2004).
61. Jacob, A., Brandt, O. & Hoheisel, J. PNA Chips. *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine* 1422-1425 (2006).at <http://dx.doi.org/10.1007/3-540-29623-9_4960>
62. Hahm, J. & Lieber, C.M. Direct Ultrasensitive Electrical Detection of DNA and DNA Sequence Variations Using Nanowire Nanosensors. *Nano Letters* **4**, 51-54 (2004).
63. Giesen, U. et al. A formula for thermal stability (Tm) prediction of PNA/DNA duplexes. *Nucleic Acids Research* **26**, 5004 -5006 (1998).
64. Steel, A.B., Herne, T.M. & Tarlov, M.J. Electrochemical Quantitation of DNA Immobilized on Gold. *Analytical Chemistry* **70**, 4670-4677 (1998).
65. Steel, A., Levicky, R., Herne, T. & Tarlov, M. Immobilization of Nucleic Acids at Solid Surfaces: Effect of Oligonucleotide Length on Layer Assembly. *Biophysical Journal* **79**, 975-981 (2000).
66. Chan, W.C. & White, P.D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*. (Oxford University Press, USA: 2000).
67. Polk, B.J., Stelzenmuller, A., Mijares, G., MacCrehan, W. & Gaitan, M. Ag/AgCl microelectrodes with improved stability for microfluidics. *Sensors and Actuators B: Chemical* **114**, 239-247 (2006).
68. Gouy, A.J. Constitution of the electric charge at the surface of an electrolyte. *Journal de Physique* **9**, 14 (1910).
69. Chapman, D.L. LI. A contribution to the theory of electrocapillarity. *Phil. Mag. Series 6* **25**, 475-481 (1913).
70. Debye, P. & Hückel, H. Zur theorie der electrolyte. I. Gefrierpunktserniedrigung und verwandte erscheinungen. *Physikalische Zeitschrift* **24**, 185-206 (1923).
71. Stern, O. Theory of the electrolytic double layer. *Zeitschrift fur Elektrochemie* **30**, 9 (1924).
72. Mikkelsen, S.R. Electrochecmical biosensors for DNA sequence detection. *Electroanalysis* **8**, 15-19 (1996).
73. Israelachvili, J.N. *Intermolecular and Surface Forces, Second Edition: With Applications to Colloidal and Biological Systems*. (Academic Press: 1992).
74. Herne, T.M. & Tarlov, M.J. Characterization of DNA Probes Immobilized on Gold Surfaces. *Journal of the American Chemical Society* **119**, 8916-8920 (1997).
75. Engel, L.W. & Young, N.A. Human breast carcinoma cells in continuous culture: a review. *Cancer Res* **38**, 4327-4339 (1978).
76. Soule, H.D., Vazguez, J., Long, A., Albert, S. & Brennan, M. A human cell line from a pleural effusion derived from a breast carcinoma. *J. Natl. Cancer Inst* **51**, 1409-1416 (1973).
77. Sagara, Y. et al. Clinical significance of Caveolin-1, Caveolin-2 and HER2/neu mRNA expression in human breast cancer. *Br J Cancer* **91**, 959-965 (2004).
78. Ding, L., Yuan, D. & Liu, J. Growth of High-Density Parallel Arrays of Long Single-Walled Carbon Nanotubes on Quartz Substrates. *Journal of the American Chemical Society* **130**, 5428-5429 (2008).

79. Snow, E.S., Novak, J.P., Lay, M.D. & Perkins, F.K. *1/f* noise in single-walled carbon nanotube devices. *Appl. Phys. Lett.* **85**, 4172 (2004).
80. Barbulovic-Nad, I., Au, S.H. & Wheeler, A.R. A microfluidic platform for complete mammalian cell culture. *Lab Chip* **10**, 1536 (2010).