Towards Single Molecule DNA Sequencing

by

Hao Liu

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2013 by the
Graduate Supervisory Committee:

Stuart Lindsay, Chair
Hao Yan
Marcia Levitus

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

Single molecule DNA Sequencing technology has been a hot research topic in the recent decades because it holds the promise to sequence a human genome in a fast and affordable way, which will eventually make personalized medicine possible.

Single molecule differentiation and DNA translocation control are the two main challenges in all single molecule DNA sequencing methods. In this thesis, I will first introduce DNA sequencing technology development and its application, and then explain the performance and limitation of prior art in detail. Following that, I will show a single molecule DNA base differentiation result obtained in recognition tunneling experiments. Furthermore, I will explain the assembly of a nanofluidic platform for single strand DNA translocation, which holds the promised to be integrated into a single molecule DNA sequencing instrument for DNA translocation control.

Taken together, my dissertation research demonstrated the potential of using recognition tunneling techniques to serve as a general readout system for single molecule DNA sequencing application.

DEDICATION

This thesis is dedicated, first of all, to my parents for raising me up and for their tremendous help in my education. Without their support and advice, I would not go this far to pursue my dream, which is to use the scientific knowledge I have learnt to make this world more beautiful.

Secondly, to my dear wife Qianru Gao for her generous support, valuable advice and unconditional love.

ACKNOWLEDGMENTS

I want to express my sincere thanks to all the people who helped me during my PhD residency. The first is Prof. Stuart Lindsay. Because he truly gave me the opportunity to be involved in a real-world cutting-edge scientific research project, and he lets me contribute my efforts to solve problems or try to answer them. He tolerated my crazy ideas, and was kind enough to give me invaluable advice about how to get out of my bad experiment desert and seek my professional oasis. He contributed all of his wisdom and strength to his scientific career, education and business management. His persistent efforts and selfless devotion for the prosperity of his ideas, knowledge and the lab left me a warm and lifetime-lasting impression. I truly and deeply appreciate what he did to educate, help and inspire me, and I sincerely admire his engineering skills, leadership charisma and his dedication. Secondly Prof. Peiming Zhang with his expertise in organic synthesis, surface chemistry and omics research, he helped me tirelessly to refine my experimental procedure and project goal. Thirdly, Prof. Hao Yan and Prof. Marcia Levitus with their outstanding and inspiring coursework, they provided me the technical expertise and extensive research advice. Moreover, Prof. Jin He, Dr. Brett Gyarfas, Dr. Brian Ashcroft, Dr. Pei Pang, Dr. Shuai Chang, Dr. Di Cao, Dr. Feng Liang, Dr. Shuo Huang, Dr. Qiang Fu, Dr. Ashylei Kibel, Dr. Parminder Kaur and my colleagues in the lab Yanan Zhao, Weisi Song, Suman Sen, Padmini Krishnakumar, they all helped me a lot with my experiments and my PhD life. I am so grateful that I have so many talented, helpful and dedicated people around me in these five years of memorable time.

The projects in my PhD work were funded by the DNA sequencing technology program of the National Human Genome Research Institute.

TABLE OF CONTENTS

Page

LIST OF TABLES ix

LIST OF FIGURES x

CHAPTER 1 BRIEF OVERVIEW OF DNA SEQUENCING: TECHNOLOGY AND APPLICATION 1

1.1 Introduction of DNA Sequencing 1

1.1.1 Early Genome Sequencing Technology History 3

1.1.2 The $1,000 Genome Competition 4

1.1.3 High-throughput DNA Sequencing 5

1.1.4 ISFET DNA Sequencing 9

1.1.5 Single Molecule DNA Sequencing 11

1.2 Application of DNA Sequencing 13

1.2.1 Personalized Medicine 13

1.2.2 Better Understanding of Our Neighbors and Ourselves 14

1.2.3 Better and Safer Food 15

1.2.4 Data Storage 15

CHAPTER 2 STM RECOGNITION TUNNELING EXPERIMENTS 16

2.1 Introduction of STM Recognition Tunneling DNA Sequencing Technique 16

2.2 Recognition Molecule Optimization 19

2.3 Explanation of Recognition Tunneling Base Differentiation Mechanism 20

2.4 Self-assembled Monolayer (SAM) Chemistry 22

Page

2.4.1 Monolayer Growth Mechanism 22
2.4.2 Choosing A Better Metal 23
2.4.3 SAM Preparation Experimental Method 26
2.5 Self-Assembled Monolayer Characterization 26
2.5.1 XPS 26
2.5.2 FTIR 28
2.5.3 Ellipsometry 30
2.5.4 Contact Angle 30
2.5.5 STM Surface Scanning Image 31
2.6 Fabrication of STM Tips 32
2.7 STM Recognition Tunneling Experiments 33
2.7.1 Recognition Tunneling Experimental Method 33
2.7.2 Time Traces of DNA Nucleoside Monophosphates Recognition Tunneling Signals 34
2.7.3 SVM Analysis of Tunneling Signals 35
2.7.4 Recognition Tunneling Signals for Monosaccharide Molecules 36
2.8 Conclusions of STM Recognition Tunneling Experiments 40
2.9 Future Work: Building Recognition Tunneling Based Single Molecule DNA Sequencing Nanopore Device 40
CHAPTER 3 TRANSLOCATION OF SINGLE-STRAND DNA IN CARBON NANOTUBES 43

Page

3.1 Why using a carbon nanotube for DNA translocation? 43

3.2 Growth Of Ultra-Long Single-Walled Carbon Nanotubes (SWCNTs) 45

3.2.1 CNT Growth Mechanism and Problem to Be Solved 45

3.2.2 Iron Nanoparticle Method 46

3.2.3 Ferritin-based SWCNT CVD Growth Method 49

3.3 PDMS Nano-Fluidic Device 51

3.3.1 Fabrication of DNA Translocation Chips (My colleagues' work) 51

3.3.2 Fabrication of PDMS Stamp 53

3.4 Design of Single-Strand DNA and PCR primers 54

3.5 Proof of DNA Translocation by Ionic Current Measurement 55

3.5.1 DNA Translocation Experiment Setup 55

3.5.2 Translocation Signal of Single-Strand DNA 56

3.6 Proof of DNA Translocation by Polymerase Chain Reaction(PCR) 57

3.6.1 PCR Experiments and the Results 57

3.6.2 Quantitation of Translocated DNA Molecules Using q-PCR 58

3.7 Conclusions of DNA Translocation in Carbon Nanotubes 61

3.8 Future Work 61

3.8.1 End Modification of CNT and CNT-based DNA Sequencing Device Fabrication 61

3.8.2 Using Graphene for DNA Sequencing 62

Page

CHAPTER 4 BEYOND SEQUENCING: THE MARRIAGE OF BIOCHEMISTRYAND SEMICONDUCTOR INDUSTRY FOR FUTURE BIOMEDICAL DEVICES 65

REFERENCES 66

LIST OF TABLES

Table Page

1. Au/Pt/Pd Physical Properties Comparison..... 23
2. Pt/Pd Grain Size Measured by STM..... 25
3. Time Course Ellipsometry Analysis of SAM Thickness..... 30
4. Time Course Contact Angle Analysis..... 30
5. SVM Analysis Result for dA/dT/dC/dG/dmC Differentiation..... 36
6. SVM Analysis Result for Glucose and Galactose Differentiation..... 39
7. SVM Analysis Result for Ribose and Deoxyribose Differentiation..... 39
8. SVM Analysis Result for Glucose, Ribose and Deoxyribose Differentiation..... 39
9. Quantification Results of Translocated ssDNA Molecules..... 60

LIST OF FIGURES

Figure Page

1. DNA Chemical Structure.[1] 2
2. 454 GS FLX Sequencer Workflow[2] 5
3. Pyrosequencing Chemistry[4] 6
4. Illumina Sequencing Chemistry and Workflow[6] 8
5. ISFET DNA Sequencing Chemistry [3] 10
6. Pacific Biosciences Real-time Single Molecule DNA Sequencing Illustration[4] .. 12
7. STM Recognition Tunneling Experiment Illustration 17
8. Recognition Molecule Form Different Hydrogen Bonding Structures[9] 18
9. Recognition Tunneling for Oligomers[10,11] 18
10. Examples of Recognition Molecule Chemical Structure Optimization 19
11. An Empirical Explanation of Recognition Tunneling Mechanism[3] 21
12. Scheme of the Self-assembled Monolayer Growth Mechanism[5] 23
13. Surface Functionalized Pd Is Resistant to Etchant 24
14. STM Images of Different Metal Surface 25
15. XPS Spectrum of Imidazole Dithiocarbamate on Au Substrate 27
16. XPS Spectrum of Imidazole Dithiocarbamate on Pd Substrate 28
17. An Imidazole Dithiocarbamate Monolayer FT-IR Spectrum 29
18. Comparison of Au Surface without and with SAM Functionalization 31
19. Bare Pd Surface Imaged with an Unfunctionalized Pd Tip in Buffer 32
20. Functionalized Pd Surface Imaged with a Functionalized Pd Tip in Buffer 32
21. STM Tips 33

Figure Page

22. Typical Nucleotides RT Signal (Imidazole dithiocarbamate Reader) 34

23. Dominant Forms of Various Monosaccharide Molecules in Aqueous Solution[7] . 37

24. Typical Ribose and Deoxyribose Recognition Tunneling Signals 37

25. Typical Glucose and Galactose Recognition Tunneling Signals 38

26. Typical Glucuronic Acid Recognition Tunneling Signal 38

27. Illustration of a Parallel Recognition Tunneling Nanopore Device[3] 41

28. CVD CNT Growth Mechanism 46

29. Iron Nanoparticle Synthesis Mechanism 47

30. AFM Image of Iron Nanoparticles on the Substrate 48

31. CNTs Grown by Iron Nanoparticle Method 48

32. Use AFM to Measure CNT Diameter Grown by Iron Nanoparticle Method 49

33. CVD SWCNT Synthesis Using Ferritin-based Method[8] 49

34. AFM Image of The Surface After Treatment of Ferritin Organic Shell Removal 50

35. Long and Clean Growth of SWCNTs Using Ferritin Method 51

36. Ferritin-derived SWCNT on TEM grid. 51

37. CNT Before and After Translocation Experiment 52

38. Chip Design for CNT Mass Transport Experiments 53

39. Assembled Nanofluidic Device for CNT Mass Transport Experiments 54

40. Mfold Prediction of Designed ssDNA Structure 55

41. Typical ssDNA Translocation Ionic Current 57

42. ssDNA translocation experiment PCR results 58

43. 60nt ssDNA Real-time Amplification Result 59

Figure Page

44. 60nt ssDNA qPCR Standard Curve. 59

45. 60nt ssDNA qPCR Standard Melting Peak Plot (only one peak) 60

46. Proposed CNT DNA Sequencing Scheme 62

47. Exfoliated Graphene and Raman Spectrum 63

48. CVD Grown Graphene and Raman Spectrum 63

49. Transfer Graphene to PDMS 64

50. Proposed Graphene DNA Sequencing Scheme 64

CHAPTER 1 BRIEF OVERVIEW OF DNA SEQUENCING: TECHNOLOGY AND APPLICATION

In this chapter, I will first introduce DNA sequencing and the technology evolution history. I will then discuss the broad application of DNA sequencing in short, which is the ultimate motivation of DNA sequencing technology innovation and commercialization.

1.1 Introduction of DNA Sequencing

DNA (deoxyribonucleic acid) is a macromolecule encoding all the genetic instructions used in the development, metabolism and reproduction of all known living organisms and many viruses. Most DNA molecules are double-strand helices. One strand of such helix consists of long spiral polymers of phosphate deoxyribose backbone and nucleobases guanine, adenine, thymine and cytosine represented using the letters G, A, T and C. G and C, or A and T form base pairs through hydrogen bonding. Figure 1 shows such a chemical structure. The human genome, which is the entirety of the genetic information stored as DNA sequences in human egg or sperm cell nuclei, consists of approximately three billion DNA base pairs.[12]

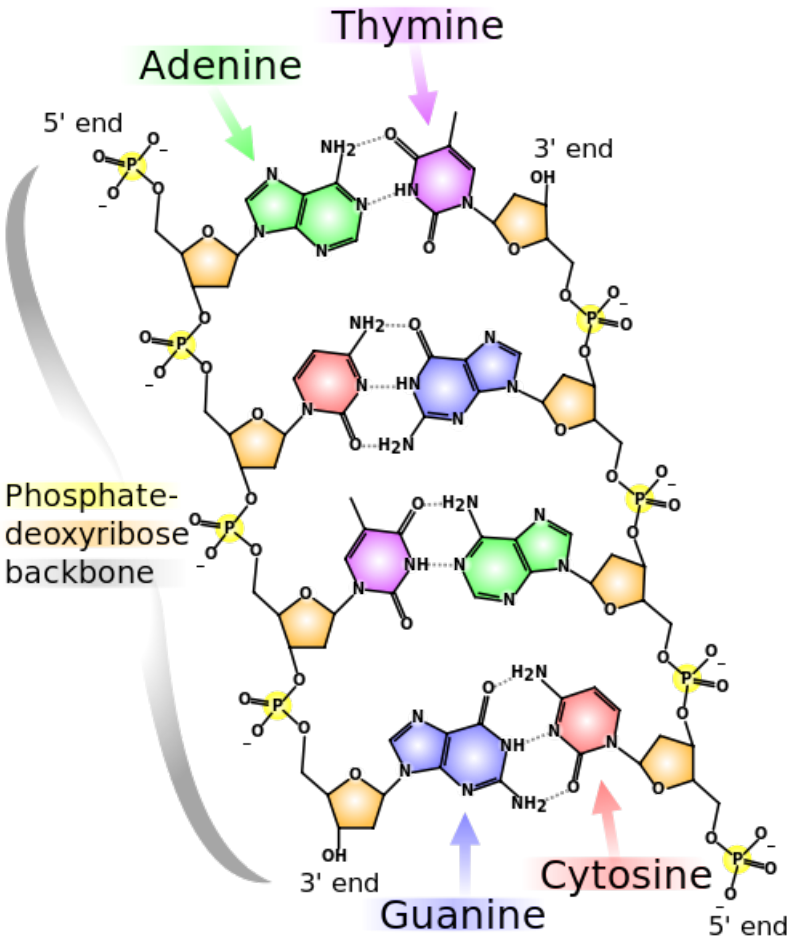


Figure 1. DNA Chemical Structure.[1]

The DNA double helix is stabilized primarily by two forces: hydrogen bonds between nucleobase pairs and hydrophobic base-stacking interactions among aromatic nucleobases.[13] By increasing the temperature of the DNA solution to a point where 50% of the base pairs break, the strength of hydrogen bonding interactions of two helices can be measured. This temperature is the melting temperature, also called T_m value. T_m depends on the length of the DNA molecule and its specific nucleotide sequence composition[14]. Higher melting temperatures are typically associated with higher GC nucleobases percentage.

For primary B-form double-strand DNA, the diameter is around 2nm, and the distance between each nucleotide unit is around 0.3nm.[15] The diameter of a single-strand DNA is narrower, around 1nm. When a single-strand DNA is stretched, the distance between the bases could reach as far as 0.7nm.[16]

Because DNA plays a crucial role in life, methods for manipulating, storing and reading DNA have been the subject of frequent technology innovation ever since the discovery of its double-helix structure. The goal of determining the entire DNA sequence of a human genome was first reached with the essential completion of the Human Genome Project in 2001.[12]

1.1.1 Early Genome Sequencing Technology History

The modern history of DNA sequencing technology began in 1977, when Sanger published the first genome of a bacteriophage with nearly five thousand nucleotides using Sanger chain-terminating DNA sequencing method.[17] The first fully automated DNA sequencer, developed by Applied Biosystems (founded in Foster City, California in 1987, now part of Life Technologies), was proved to be a rapid platform for obtaining short strand DNA sequences with almost 100% accuracy.[18]

For longer DNA, they must be first divided into smaller fragments for sequencing by Sanger's method. Computer programs were then deployed for solving the reassembling puzzle according to the overlapping region to render the entire sequence. 'Shotgun' sequencing method was thus named and was one of the precursor technologies enabling full human genome sequencing. The first human genome draft was generated by an improved version of this simple procedure in 2001.[19]

Today, the shotgun genome sequencing strategy is still under deployment. Other sequencing technologies, so called next-generation DNA sequencing technologies, are invented with the benefit of much higher throughput, longer read length, faster speed and cheaper chemistry.

1.1.2 The $1,000 Genome Competition

The cost of the Human Genome Project was nearly three billion dollars.[12] Among many of the initiatives promoting the development of whole genome sequencing technologies, the series of '$1,000 Genome' grants introduced by National Human Genome Research Institute gain the most attention. The grants aim to promote the development of technologies that will eventually allow a human genome to be sequenced for $1,000 or less. $1,000 genome shows a strong promise for this once-in-a-lifetime expenditure to be affordable by an individual.

To achieve this goal, all makers of sequencing instruments are competing furiously with each other to improve sequencing throughput, read length, read speed and accuracy. In January 2012, Life Technologies unveiled its benchtop Ion Torrent Sequencing platform and chips, designed to sequence a human genome for just $1,000 in a matter of hours. The $1,000 estimated cost per genome includes the materials needed for template preparation, amplification, sequencing itself and the cost of the chips. The next day, Illumina - the market-leading DNA sequencer manufacturer - said the company launched its own updated sequencing machine HiSeq2500 capable of reading a human genome at >30× coverage in 27 hours.[20] Genia Technologies is even proclaiming a $100 genome using its nanopore-based NanoTag sequencing technology.[21]

Foreseeably, the cost of next-generation sequencing will be soon acceptable to the general public, and the application could be widely adopted as a diagnostic routine in clinical laboratories.

1.1.3 High-throughput DNA Sequencing

The developments of high-throughput sequencing (or massively parallel sequencing[6]) are driven by the high demand for low-cost sequencing. Thousands or millions of sequences can be generated at once by parallelized sequencing process. Compared to the traditional Sanger sequencing method, the price has been dramatically reduced, and the throughput has been increased by several orders of magnitude.[6] Since its debut circa 2005, high-throughput DNA sequencing platform has increased its capacity at a speed higher than the Moore's Law rate, which states that the number of transistors per chip will double every two years.[22] The latest methods and programs are developed for sequence assembly, analysis and interpretation because of the large quantities of data

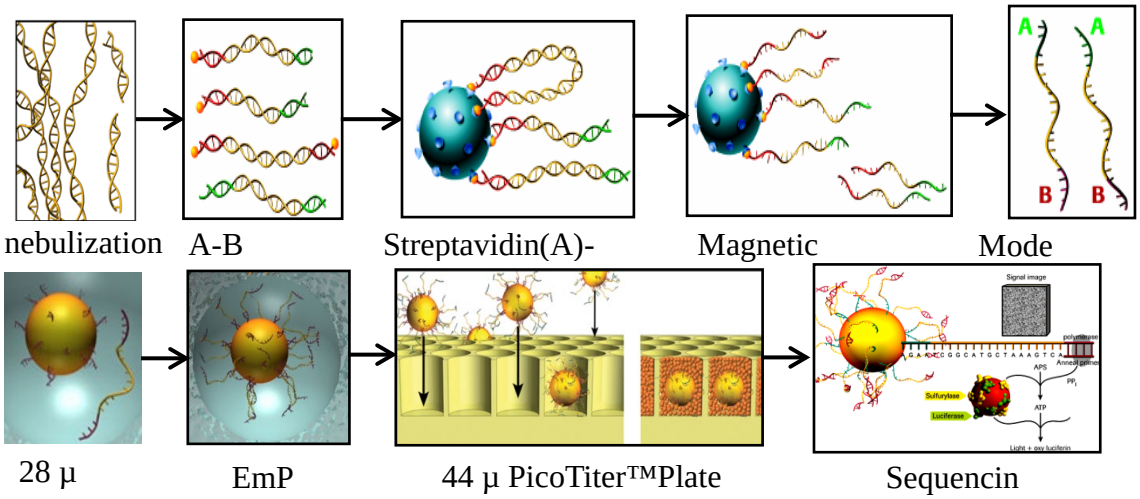


Figure 2. 454 GS FLX Sequencer Workflow[2]

produced by DNA sequencing. In order to manage the extremely large data sets, substantial enhancements in computer infrastructure, data storage and transfer capacity will be needed.

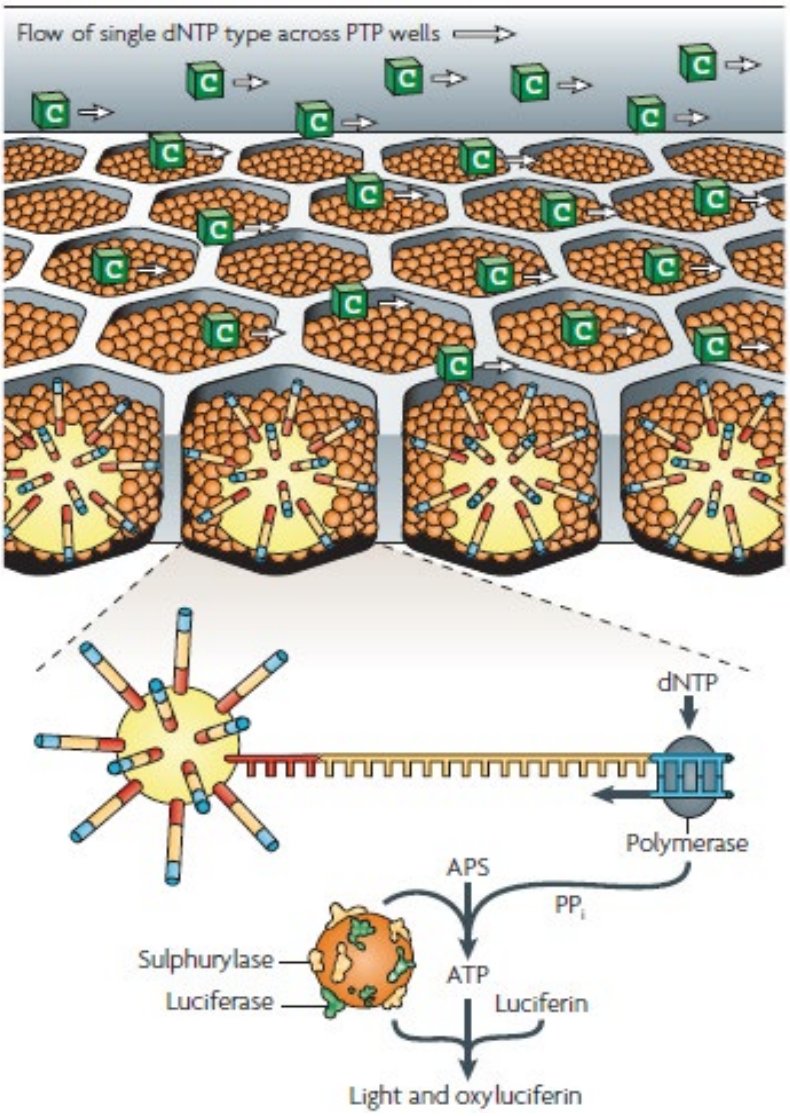


Figure 3. Pyrosequencing Chemistry[4]

The Roche 454 GS FLX series Genome Sequencer is one of the examples of high-throughput DNA sequencing platforms. The latest version GS FLX+ platform produces a mode read length of 700 bp and a typical 700 Mb throughput (= 1M reads per run ×

700bp mode read length) per 23h run. The consensus accuracy at 15× coverage is 99.997%.[2] Figure 2 shows a workflow chart. A typical 500ng genome DNA is first nebulized at 30psi (2.1bar) with vented cap nebulizer. DNA Fragments are polished to generate blunt ends for adaptor ligation. Adaptors containing fluorescent molecule for direct quantitation of the library are ligated. Single-strand DNA templates attached to the magnetic bead are separated for emulsion PCR amplification in water-in-oil microreactors. DNA-positive beads are then enriched and deposited into microwells. Because of the confined well geometry, only one bead is allowed in one well. Layers of packing beads, enzyme beads and PPiase beads are then deposited. The core sequencing chemistry, as shown in Figure 3, is the famous pyrosequencing method. Rather than the chain-termination method in Sanger sequencing, this method relies on the detection of pyrophosphates released in the nucleotide incorporation process.[4] The read length, though shorter than Sanger method and continuously challenged by other high-throughput sequencing method (MiSeq from Illumina, e.g.[23]), is still the longest among all non-single-molecule methods. Long reads, which makes genome assembly easier and more accurate, are essential for *de novo* sequencing of genomes, transcriptomes or amplicons containing repetitive and rear-ranged DNA segments[6]. High reagent cost, long and tedious sample preparation, high error rates in homopolymer repeats and crosstalk between adjacent wells containing single clonally amplified beads are examples of the common drawbacks of this method.[4]

Illumina's high-end sequencer HiSeq2500 uses 'sequencing by synthesis' (SBS) technology. Novel reversible terminator nucleotides each labeled with different fluorescent dyes[24,25] produce single reads of 150+ base pairs (bp) per end. This machine

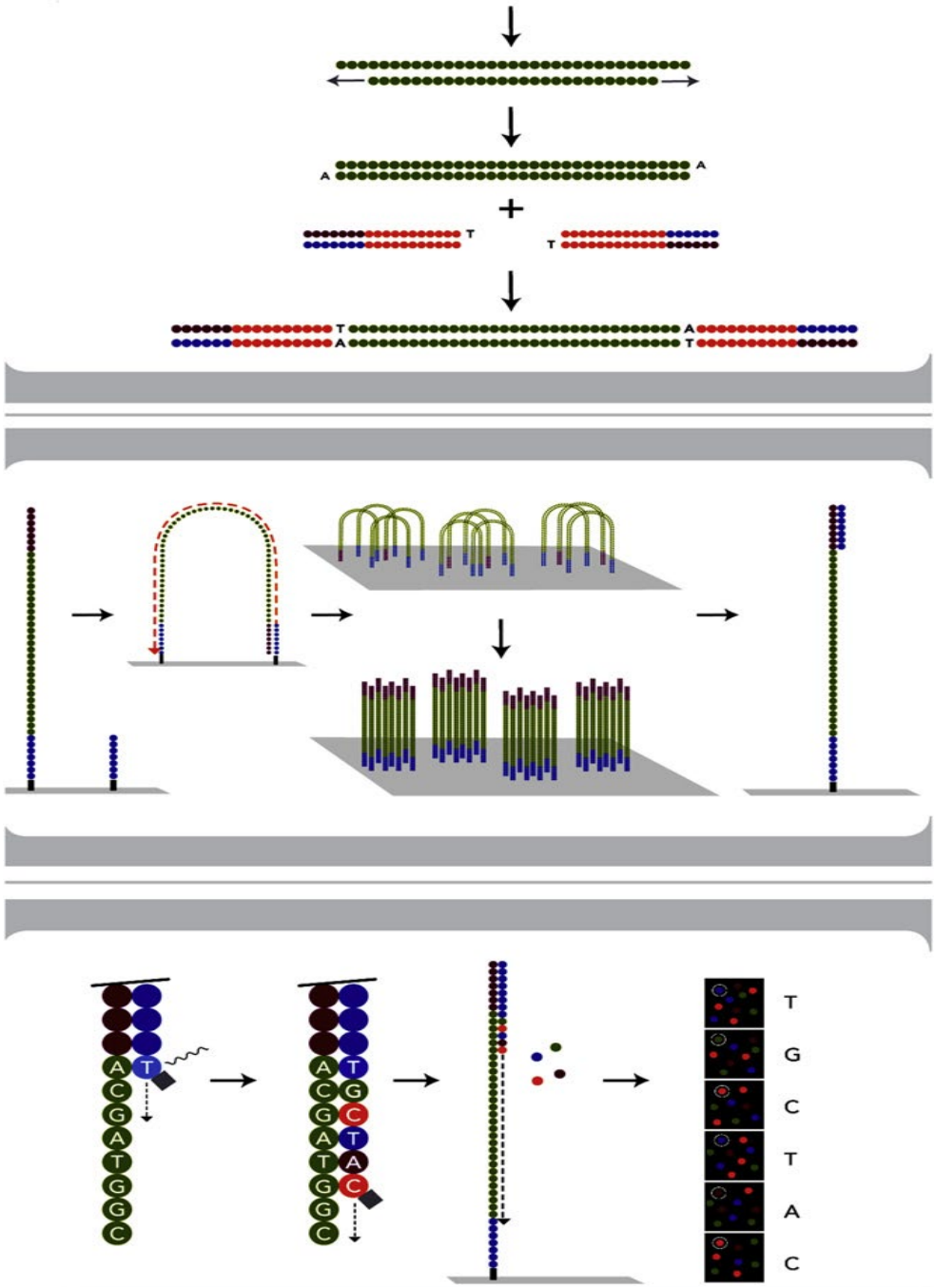


Figure 4. Illumina Sequencing Chemistry and Workflow[6]

is currently capable of generating up to 120 gigabases (Gb) of sequences in 27 hours in a rapid run mode.[20] As shown in Figure 4, fragmented DNA are first ligated with adaptors, and then immobilized to the solid support coated with oligonucleotides complementary to the adaptors in the flow cell chamber. 'Bridge PCR' is performed, and the following denaturation generates dense clusters of single stranded template sequencing library anchored to the surface. The denatured clusters are then added with a universal primer targeting the adaptor sequence of the DNA fragments, fluorescently labeled nucleotides each with their 3'-OH blocked[4], and a special DNA polymerase capable of incorporating the modified nucleotides. The sequencing is done by the cyclic reversible termination technology. After each base is incorporated, the surface is imaged to determine the identity of the incorporated nucleotide. The 3'-OH inactivating residue and the fluorophore are removed, and then the sequencing process is repeated. The resulting 4-color images are used for base calling.

1.1.4 ISFET DNA Sequencing

ISFET, short for ion-sensitive field effect transistor, is used to measure ion concentration in solution. The current passing through the transistor will change according to the change of the ion concentration. An ISFET has a gate electrode connecting only to a passivation layer. The interfacial potential is controlled with respect to the source by means of a reference electrode placed in contact with the electrolyte above the passivation layer. Typical gate materials are SiO_2, Si_3O_4, Al_2O_3, or Ta_2O_5.[26] The mechanism of the oxide surface charge change in response to the local ion concentration can be described by the *site binding model*[27]. The hydroxyl groups on the oxide surface can donate or accept protons.[27] The chemistry of ISFET DNA sequencing

technology is shown in Figure 5. The protons released in the DNA synthesis process in a local chamber can result in falling of pH, which will be further detected by an ISFET. The DNA sequencing library on the bead can be amplified again by emulsion PCR, the same method as shown in 454 sequencing workflow.

Ion Torrent, a division of Life Technologies Inc., recently developed and successfully manufactured ISFET based sequencing chip[22]. Such chips, which contain all the measurement and data collection complexity, are combined with additional automated sample preparation instrumentation, standard sequencing reagents, simple fluidics and adjacent computational hardware to provide a complete, computer-like sequencing

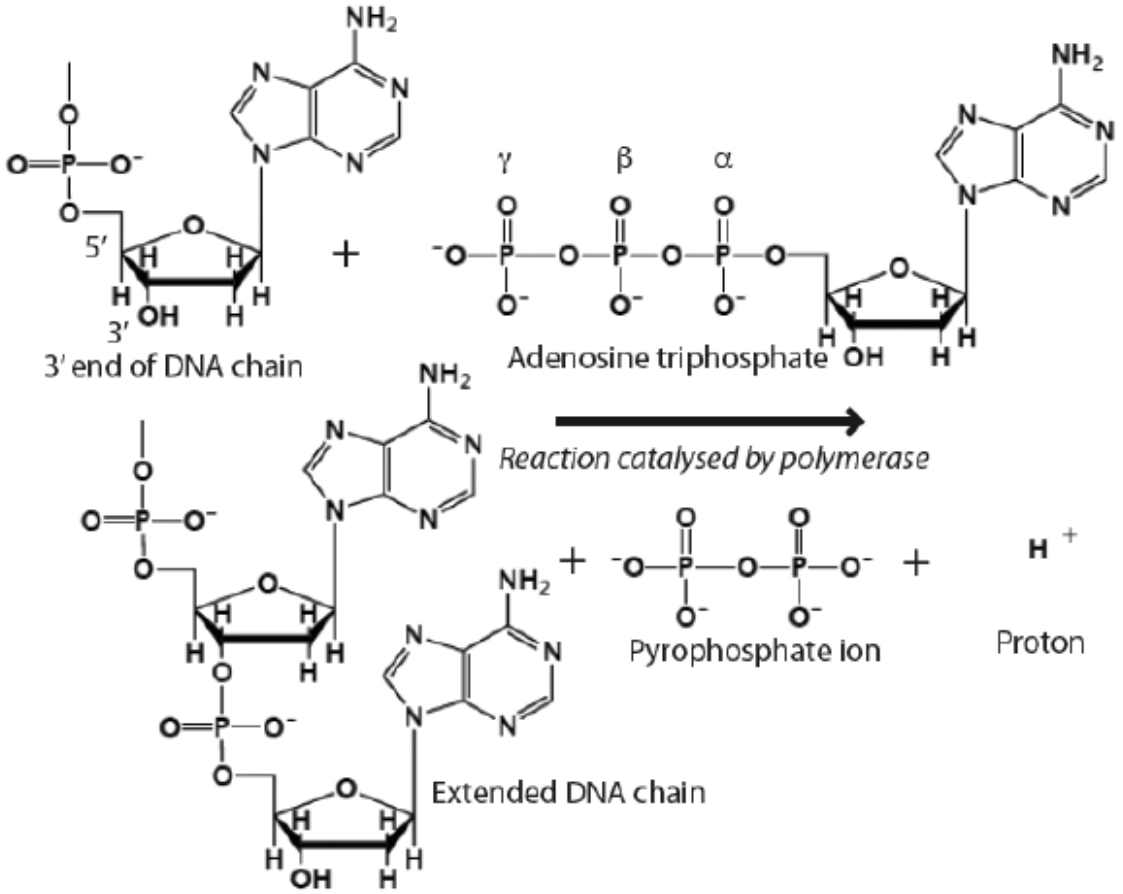


Figure 5. ISFET DNA Sequencing Chemistry [3]

platform. The production of the chip leverages the large scale and low-cost complementary metal-oxide semiconductor (CMOS) chip fabrication facilities currently widely used for computer or cellphone microprocessors manufacturing. CMOS compatible fabrication of the detection circuits makes it super easy for scaling up so the cost of the instrument can be dropped down dramatically.

The company has their own strategy to tackle the "data tsunami" problem and increase genome sequencing pipeline handling efficiency by introducing the stand-alone Ion Proton Torrent Server for data processing, which includes base-calling, alignment, and variant analysis.

Being a topic out of my thesis scope, a detailed comparison between the Ion Torrent Proton and the Illumina HiSeq2500 performance and cost has been done.[28]

1.1.5 Single Molecule DNA Sequencing

In sharp contrast to traditional high-throughput or ISFET DNA sequencing technologies, physical approaches probing DNA molecules at the single-nucleotide level have the potential to deliver faster and low-cost sequencing by cutting out the expensive chemistry needed for library generation and DNA amplification. Nanochannels, nanogaps or nanopores allowing spatial confinement of DNA molecules are central to the single molecule DNA sequencing method.

Nanopore technologies are one of the fast and direct single molecule DNA sequencing methods. Kilobase length single stranded genomic DNA or RNA can be driven through the nanopore by electrophoretic force. The sequence can be reflected by distinct current blockade levels of various bases as DNA traverse through a mutated protein pore's (α-hemolysin[29] or MspA[30]) thin and narrow constriction with high

sensitivity.[30] No DNA amplification or labeling is needed makes affordable and rapid DNA sequencing a possibility. Single-nucleotide resolution and DNA translocation speed control are the two long-standing hurdles to nanopore sequencing.[31] Deconvoluting current traces for underlying sequence extraction is not an easy effort. Identifying extended homopolymer regions with high confidence is still problematic in nanopore sequencing. Robust platform and parallelization needs to be constructed in order to successfully commercialize nanopore sequencing.

A number of other startup companies are vying to commercialize single molecule sequencing technology, including Oxford Nanopore in the UK, NABsys in Providence, RI and Genia Technologies. Oxford Nanopore has not released its commercial sequencing platform yet, but the excitement about its single-molecule nanopore-based sequencing technology prevails.

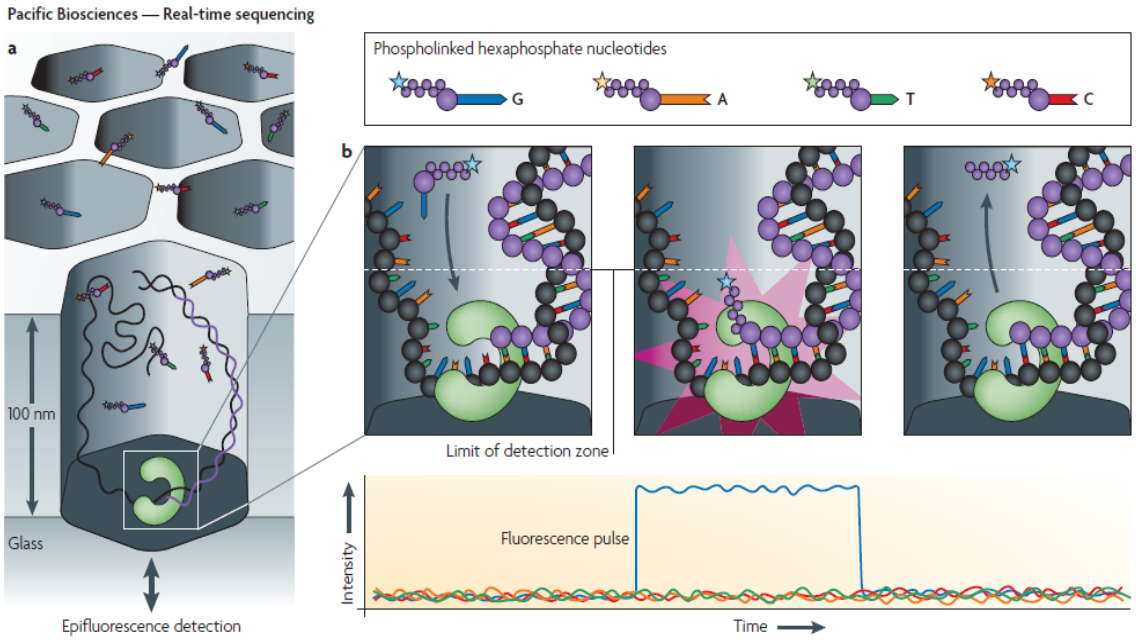


Figure 6. Pacific Biosciences Real-time Single Molecule DNA Sequencing Illustration[4]

Pacific Biosciences, a company leading the effort of single-molecule DNA sequencing uses nucleoside quadraphosphates with fluorescent dyes attached which is further cleaved off during the DNA extension reaction. The process of DNA synthesis will not be halted so the incorporation of bases can be followed in real-time. Single-molecule DNA template without amplification is used by single DNA polymerase molecules attached to the bottom surface of individual zero-mode waveguide detectors (ZMW detectors, Figure 6).[4] Though the technology has the greatest potential for reads exceeding 1kb, the error rates are at the same time the highest compared to other sequencing chemistry.

1.2 Application of DNA Sequencing

The broad application of DNA sequencing is the ultimate drive for faster, cheaper and more accurate DNA sequencing technology innovation.

DNA sequencing applications include de novo sequencing and re-sequencing of genomics, metagenomics[32], RNA analysis, and targeted sequencing of DNA regions of interest.

Below are four examples of popular DNA sequencing applications.

1.2.1 Personalized Medicine

Thirty years ago, the chemistry information in a drop of human blood identified its source as falling into one of four blood type groups. Today, the availability of genetic information in that drop of blood represents one of the most exciting opportunities in the history of biomedicine.[33]

Next generation genome sequencing technologies have allowed a better understanding of the genetic basis of diseases. Recent advances have demonstrated the

clinical potential of sequencing technologies in characterizing the genetic mechanisms of rare inherited diseases, tumor development pathways and response to specific medication.[34] Though challenges in genome analysis remain and more studies are needed to ensure that new technologies will be introduced into clinical practice in a medically and ethically responsible manner, recent genome research discoveries and the resulting clinical genome sequencing applications are showing the promise of personalized medicine and individualization of treatment[35]. Though experts are still debating whether healthy people should have their genome sequenced,[36] consumers will certainly be presented with new information and choices.

Besides pointing the way to new generations of drugs, treatments and diseases prevention methods, next generation sequencing technologies can also be used for a deeper understanding of genotype-phenotype correlations, providing invaluable information about susceptibility to diseases, determining family pedigrees and predicting individuals' vulnerability or adaptability to specific environments and substances. Moreover, rapid non-invasive prenatal and neonatal genome screening tests using next generation sequencing technologies are already on the market[37]. Efforts in applying next generation DNA sequencing technologies to study genes associated with skin aging for development of truly personalized skin care and personal care product are also gaining more and more attention in the beauty industry[38].

1.2.2 Better Understanding of Our Neighbors and Ourselves

We can also use genome sequencing as a tool to understand the species which cannot be cultivated or raised in the lab, archaea in marine sediments as an example, [39] or

to assess the genetic diversity encoded by microbial communities sharing a common habitat[32], bacteria in the human gut, for example.

1.2.3 Better and Safer Food

The whole-genome sequence analysis of food plant or meat animal, bread wheat for example,[40] is crucial to their evolution, domestication and genetic improvement. Fast genomic analysis of foodborne pathogens will help us understand the origin of outbreaks and develop specific diagnostics.[41]

1.2.4 Data Storage

DNA, in nature, is the super stable data storage material that encodes all the information to direct the development and function of living organisms. Therefore, it is possible to store data in the base sequence of DNA. A recent article published in the journal *Nature* reported that over 5 million bits of information were successfully encoded using Agilent Technologies' OLS (oligo library synthesis) process and retrieved with 100% accuracy on Illumina HiSeq 2000.[42] The associated cost is estimated at $\$12{,}400MB^{-1}$ for data storage and $\$220MB^{-1}$ for data decoding. Being a slow process, the method is invented for long-term archival of low access-rate data. However, with the ongoing reduction in DNA synthesis and sequencing costs, DNA based data storage system shows promise as a practical way of digital archiving in the near future.

CHAPTER 2 STM RECOGNITION TUNNELING EXPERIMENTS

Current ISFET based Ion Torrent platform requires time consuming DNA polymerase amplification. Pacific Biosciences' single-molecule real-time DNA sequencing is realized by optical means, which is not easy to scale up compared to CMOS fabrication. So single molecule amplification-free and direct electrical detection sequencing method can be the holy grail because it combines low-cost instruments with straightforward sample preparation. It might be the last generation of DNA sequencing technology innovation, and it is the main focus of my PhD research and work.

This chapter starts with the introduction of recognition tunneling concepts and development history. My experiments and results will be shown after that.

2.1 Introduction of STM Recognition Tunneling DNA Sequencing Technique

In 2008, theoretical calculation showed the possibility of using electron tunneling for DNA base detection with two closely held electrodes.[43] In a scanning tunneling microscopy (STM), because the tunneling current is extremely sensitive to the width between the tip and the substrate (a change of 1Å introduces an order of magnitude difference in tunneling current), it is possible for a macroscopically blunt STM tip to pick out an individual base on a DNA polymer.[44] However, it is almost impossible to use bare STM tips to align individual DNA base with subangstrom precision and the gap distance in the calculation is too small to pass single-strand DNA through easily.

The design of a readout system using tunneling signals to sequence DNA requires knowledge of the conductance across all four DNA bases, and a sensing mechanism for DNA base differentiation. Experiments were first carried out in an organic solvent with a bare STM tip and bare gold substrate separated by a distance of 2nm between each other.

A wide distribution of peak currents were found to be reduced by 10-fold when one of the electrodes was functionalized with a recognition molecule capable of forming specific hydrogen bonding structures with different DNA nucleosides. A density functional calculation predicted that if the second electrodes could be functionalized with the same recognition molecule, the contact resistance to the nucleosides would be reduced allowing electronic signatures of all four DNA nucleosides to be resolved in a tunneling gap (Figure 7). Figure 8 shows the hydrogen bonding energy-minimized structures in computer simulation calculation. Four nucleosides each trapped in a 2.5nm gap with 4-mercaptobenzoic acid as the recognition molecule form four distinct hydrogen bonding complexes. The Further experiments in the lab confirmed this calculation and paved the new way for DNA base differentiation.[9]

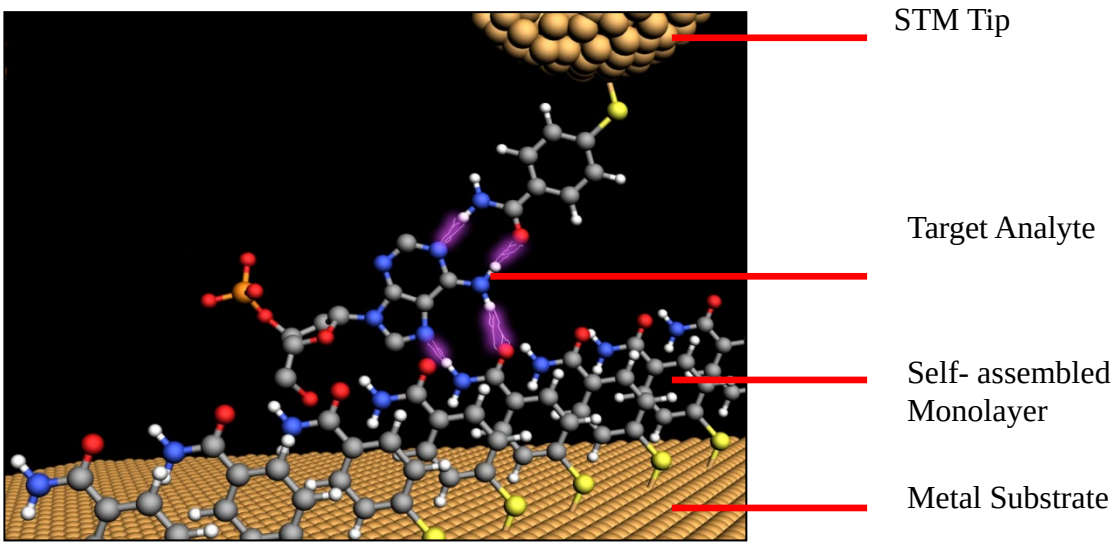


Figure 7. STM Recognition Tunneling Experiment Illustration

For this method to be adopted for DNA sequencing application, experiments must be designed to be done in an aqueous electrolyte solution instead of an organic solvent. Results showing DNA base differentiation in aqueous solution is the minimum proof of

concept requirement. Single base identification in DNA oligomers could be the next milestone. Shown in Figure 9, later experiments successfully addressed these two issues.[10,11,45] In 2010, this technique was named recognition tunneling.[46]

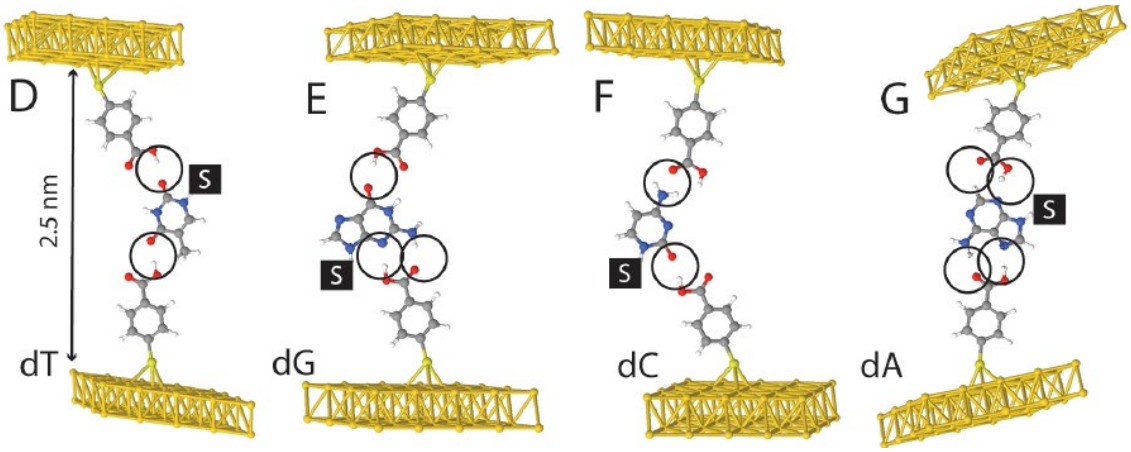


Figure 8. Recognition Molecule Form Different Hydrogen Bonding Structures[9]

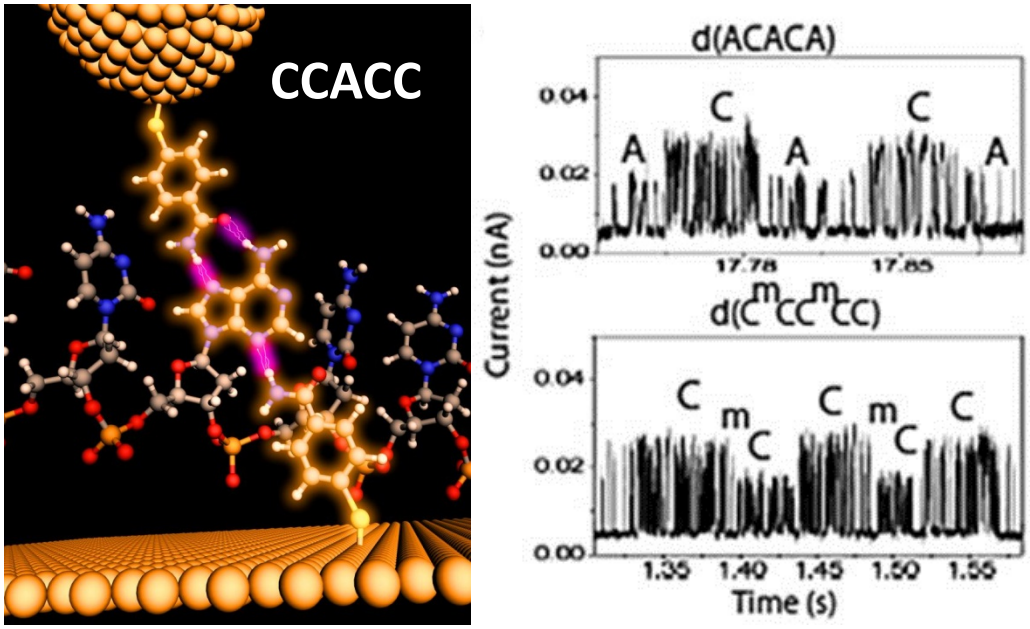


Figure 9. Recognition Tunneling for Oligomers[10,11]

2.2 Recognition Molecule Optimization

Calculations based on density functional theory (DFT) and NMR studies can be performed in designing a better universal base pair reader in a tunneling gap.[47] The desired molecule will form non-covalent complexes with different DNA bases, show distinguishable electronic signatures under an electrical bias, low or no noise in the control experiment (no DNA nucleotides are added) and high solubility in ethanol solution for simple self-assembled monolayer (SAM) preparation.[48]

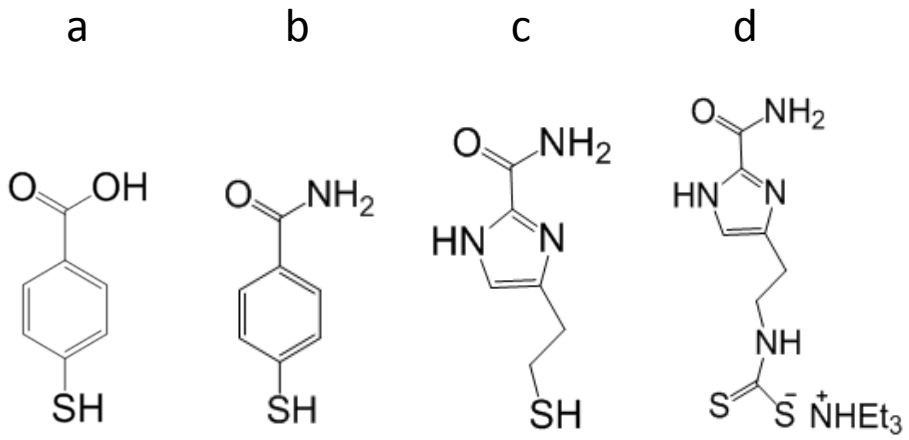


Figure 10. Examples of Recognition Molecule Chemical Structure Optimization

The first generation recognition molecule, shown in Figure 10 (a) was used for DNA base differentiation in organic solvent.[9] To extend this technique to reads in buffered aqueous solution, the reagent 4-mercaptobenzamide (Figure 10b) was synthesized. This molecule, once integrated into electrodes forming a tunneling gap, is capable of identifying individual bases embedded within a DNA oligomer.[10] This is a strong evidence that recognition-tunneling technique can be utilized to resolve single base. However, unfortunately, 4-mercaptobenzamide produced no signals from thymine. A new adaptor molecule, 4(5)-(2-mercaptoethyl)-1*H*-imidazole-2-carboxamide in Figure

10(c) was developed. The detailed synthetic route, physicochemical properties and hydrogen bonding pattern of the thiolated imidazole-carboxamide recognition molecule was reported by my colleagues.[48] With this recognition molecule, all four bases were generating tunneling signals with an excellent base differentiation percentage.[45] An example of another recognition molecule I tried for DNA base differentiation is named thiolated imidazole dithiocarbamate, shown in Figure 10 (d). Because dithiocarbamate-based molecular junctions showed efficient electronic coupling and improved stability[49], this molecule was synthesized by my colleagues for me to try DNA base differentiation and see if this new group of molecules can perform better than the current imidazole carboxamide recognition molecule could. The results will be shown soon in the following sections.

2.3 Explanation of Recognition Tunneling Base Differentiation Mechanism

In recognition tunneling experiments, the STM tips and the substrates were functionalized with a recognition molecule. It is designed to bond strongly with the metal electrodes but to contact the target analytes via only weak, non-covalent interaction. The recognition molecule offers a more specific set of chemical interactions with the target analytes than the bare metal electrode does. The displacement of surface hydrocarbon by the thiol molecules can also reduce the surface energy, and thus reduce the contamination.

The tunneling gap is adjusted to be more than twice the molecular length of the recognition molecule, so the control experiments done in buffer solution without any DNA molecules will essentially produce no spikes. When buffer solution containing target analytes is introduced into the sample well, current spikes will show up as the

analyte molecule diffuses into the tunneling gap and forms a junction with the recognition molecules from both electrogdes. Enhanced tunneling, through the recognition molecules via the target analyte, can happen because the complex that is formed has a smaller HOMO–LUMO gap than through the surrounding water molecules. If the target analytes form different non-covalent interactions with the two recognition molecules as illustrated in Figure 11, each analyte will produce a distinctive 'fingerprint' train of stochastic tunneling current pulses, which can be analyzed by a machine-learning algorithm called a support vector machine (SVM).[45, 11]

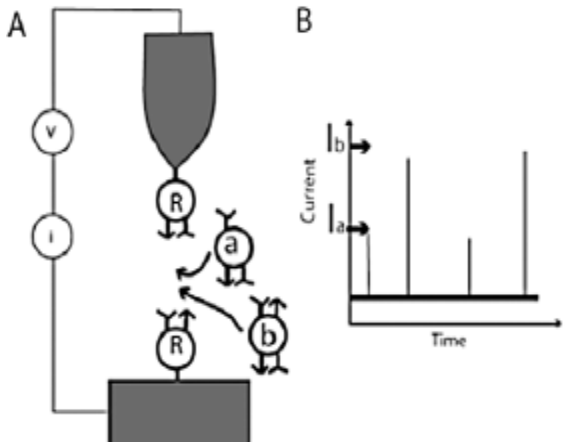


Figure 11. An Empirical Explanation of Recognition Tunneling Mechanism[3]

Simple ideas of hydrogen bonding will be sufficient in explaining the interactions between recognition molecules and target analytes in organic solvents (Figure 8). Predicted order of increasing electronic conductance (T<G<C<A) was followed by experimentally measured distribution of spike heights.[9] Later theoretical prediction and experiments in aqueous buffer solution do not seem to agree with each other very well. NMR titration studies showed that DNA nucleosides interact with imidazole carboxamide recognition molecules through hydrogen bonding in a tendency of dG>dC≫dT>dA.[48]

Peak current analysis from experiments with Au electrodes did not show obvious base differentiation.[45] Water in solution may play a pivotal role in those non-covalent interactions in the solution.

Despite several experimental and theoretical advances, including the utilization of recognition tunneling techniques for DNA nucleotides differentiation, there is still limited correspondence between experimental and theoretical studies of this system.

2.4 Self-assembled Monolayer (SAM) Chemistry

2.4.1 Monolayer Growth Mechanism

Robust organic molecule-based electronic devices require reliable and highly conductive contacts between the metal electrodes and molecules. Thiols and amines are widely used to bind molecules robustly to the metal surface. A paper published in 2005 by George Whitesides's group at Harvard gave a thorough review of self-assembled monolayer chemistry, physical properties and their application in modern nanotechnology.[50] For the recognition tunneling experiments in my thesis, thiolated recognition molecules were functionalized onto various metal surfaces via strong thiol-metal bonds.

The simplest chemical model of this process is illustrated in Figure 12. An initial physisorption step is followed by chemisorption of the desired molecules. Nucleation of the molecules is formed with time, and finally the ordered crystalline domain formation is completed in typically 12~24h.[5]

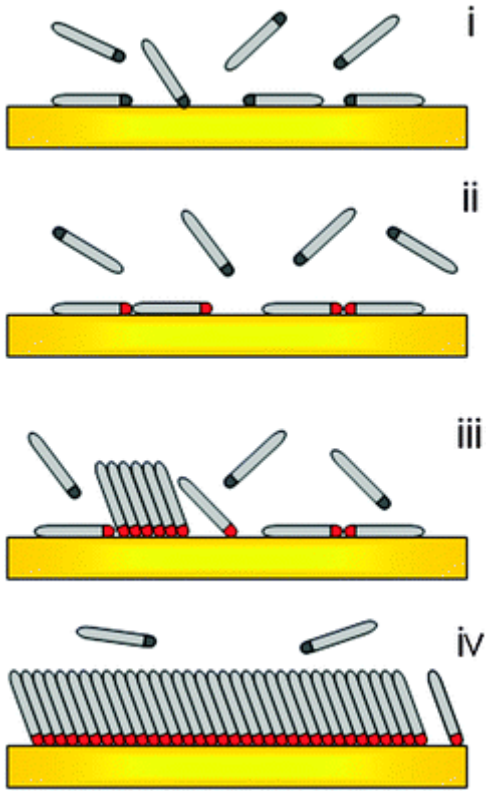


Figure 12. Scheme of the Self-assembled Monolayer Growth Mechanism[5]

2.4.2 Choosing A Better Metal

Table 1. Au/Pt/Pd Physical Properties Comparison

	Au	Pt	Pd
Shear Modulus	27 GPa	61 GPa	44 GPa
Reduction Potential	$Au^{3+} + 3e^{-} \Leftrightarrow Au$: +1.52V	$Pt^{2+} + 2e^{-} \Leftrightarrow Pt$: +1.188V	$Pd^{2+} + 2e^{-} \Leftrightarrow Pd$: +0.915V
Melting point	1064.18 °C (1947.52 °F)	1768.3 °C (3214.9 °F)	1554.9°C (2831°F)
Density at 20°C	19.30g/cm^3	21.45 g/cm^3	12.02g/cm^3
boiling point	2856 °C	3825°C	2963°C
Atomic radius	1.35Å	1.35Å	1.37Å

The physical properties of common metals used in molecular electronics research are listed in Table 1. Most of the surface chemistry and molecular electronics work on metals has focused on gold, a metal not compatible with CMOS fabrication facilities.[51]

Palladium (Pd) has several properties that suggest it would be a better material for SAM-involved molecular electronic devices: (i) It is highly reactive toward thiol-terminated chemicals.[52] (ii) Once formed on the surface, the SAMs will keep the palladium from corrosion by wet etchants such as $FeCl_3$ (Figure 13), because of the PdS interlayer formed between the organic thiol molecules and the underneath palladium. [52](iii) This metal is compatible with the semiconductor industry CMOS fabrication process. (iv) It resists oxidation in air up to 800°C. (v) Theoretical calculation showed a higher tunneling conductance than gold.

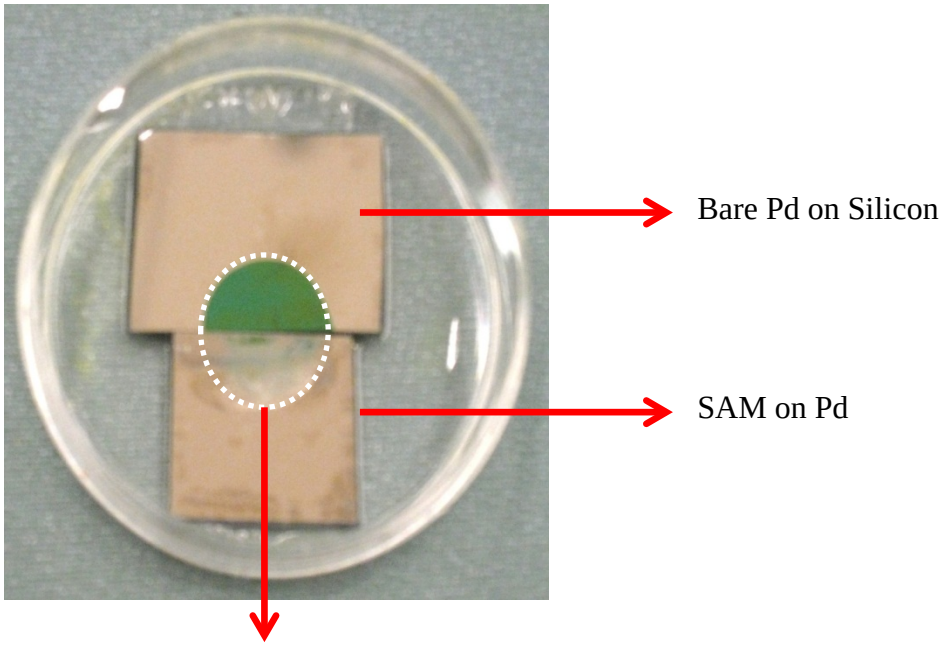
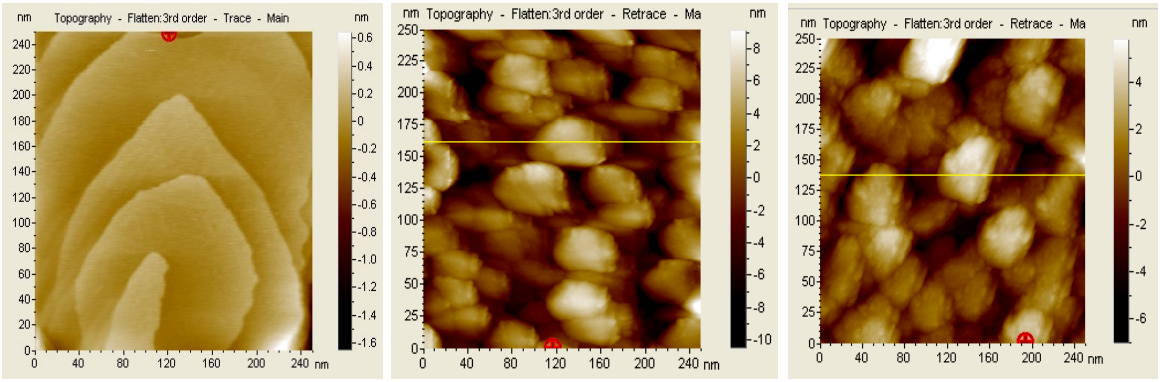


Figure 13. Surface Functionalized Pd Is Resistant to Etchant

In the experiments, Au (111) substrates from Agilent Technologies are prepared by epitaxially grown high purity gold onto green mica in a high vacuum.[53] The resulting gold surface is composed of flat Au (111) terraces up to 280μm. Metals with higher melting point (Pd: 1552°C, Pt: 1772°C) tend to have a smaller grain size than metals with lower melting point (Au: 1064°C)[50]. Polycrystalline Pd/Pt substrates can be made by e-beam evaporation or ion-beam deposition of Pd/Pt onto a Ti adhesive layer on a silicon wafer. The substrates were cut into small pieces with a rough 1×1 cm^2 dimension. STM images for Au/Pd/Pt substrates are shown in Figure 14. The grain size measured by STM is shown on Table 2. The grain size seems to be decreased after the hydrogen flame surface cleaning.



Bare Au(111) Bare Pt Surface Bare Pd Surface

Figure 14. STM Images of Different Metal Surface

Table 2. Pt/Pd Grain Size Measured by STM

	Pt (E-beam)/nm	Pt (Ion-beam)/nm	Pd(E-beam)/nm
Grain width before H flame	~100	~20	~50
Grain depth before H flame	~20	~3	~13
Grain width after H flame	~15	~20	~30

Grain depth after H flame	~12	~5.6	~6
CMOS compatibility	Yes	Yes	Yes

2.4.3 SAM Preparation Experimental Method

Au/Pd/Pt substrates were first annealed with a hydrogen flame for 30-40s. HPLC grade ethanol was sonicated with Argon degasing for 30min before use. SAMs were modified onto Au/Pd/Pt surface by immersing the metal substrates in 2mL 0.1-2.5mM ethanolic solution, DMF:ethanol = 1:10 or DMSO:ethanol = 1:200 solution of various thiol-terminated recognition molecules for 12-24 hours. Longer time was chosen for dithiocarbamate molecules because of the slower SAMs formation kinetics.[54] The substrates were then washed with copious amounts of ethanol and dried under nitrogen prior to following experiments in the phosphate buffer solution.

The resulting monolayers were characterized by XPS, ellipsometry, FTIR, contact angle measurements and STM surface imaging. The next section will show examples for each analytical method done for imidazole dithiocarbamate recognition molecule used in DNA base differentiation work.

2.5 Self-Assembled Monolayer Characterization

2.5.1 XPS

X-ray photoelectron spectroscopy (XPS) can be used to quantitatively analyze the SAM atomic concentration, empirical formula, electronic state of the elements, and the film thickness. The atomic concentration percentage is calculated by dividing the element peak intensity by the relative sensitivity factor (RSF) and then normalized over all of the elements detected. Should the experimental conditions change in any way, for example

the x-ray gun power output changes, then peak intensities would change in an absolute sense, but all else being equal, would remain relatively constant. 10% accuracy is typically quoted for routinely performed XPS atomic concentration percentage (At.%) measurement.[55]

All my X-ray photoelectron spectra were obtained with a VG ESCALAB 200i-XL photoelectron spectrometer. 15keV Al-Kα radiation at 6×10^{-10} mbar base pressure was the source. Wide scan spectra were obtained at a passing energy of 150eV. C1s, Pd3d, N1s, O1s, S2p, Au4f core level high resolution spectra were collected with a passing energy of 20eV. CasaXPS software package was used for the curve fitting of S2p spectra and atomic concentration calculation.

XPS measurements indicated that the sulfur is bound to gold or palladium (Figure 15, 16). The binding energy of the S2p core level electrons is ~162eV, referenced to C1s at ~284.80eV.

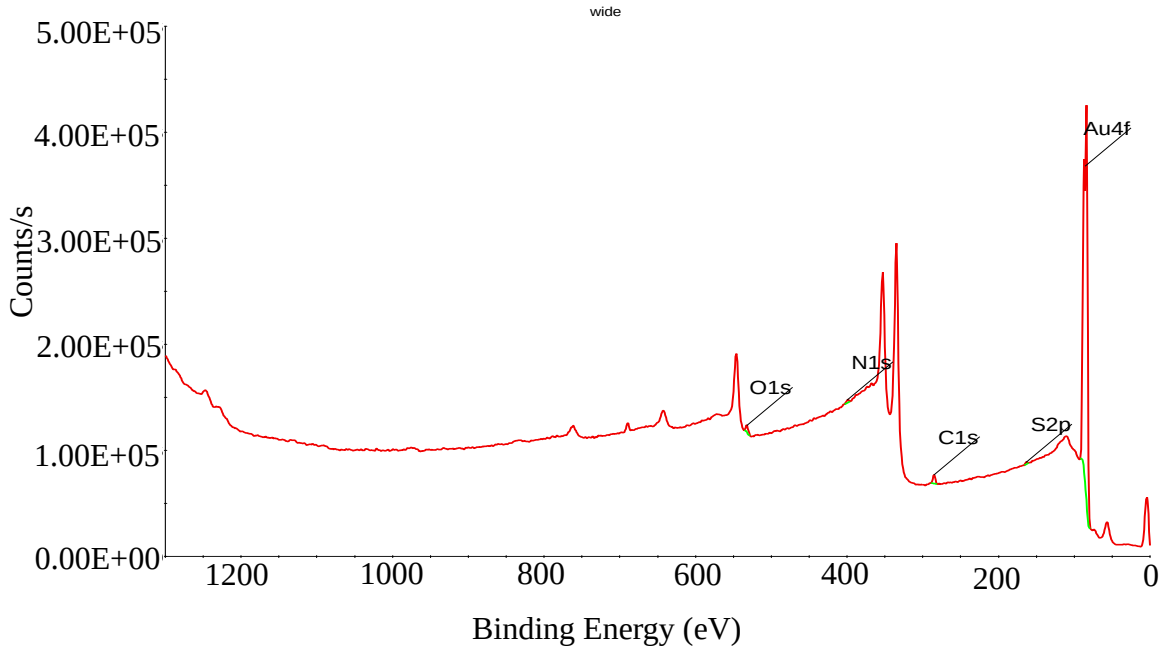


Figure 15. XPS Spectrum of Imidazole Dithiocarbamate on Au Substrate

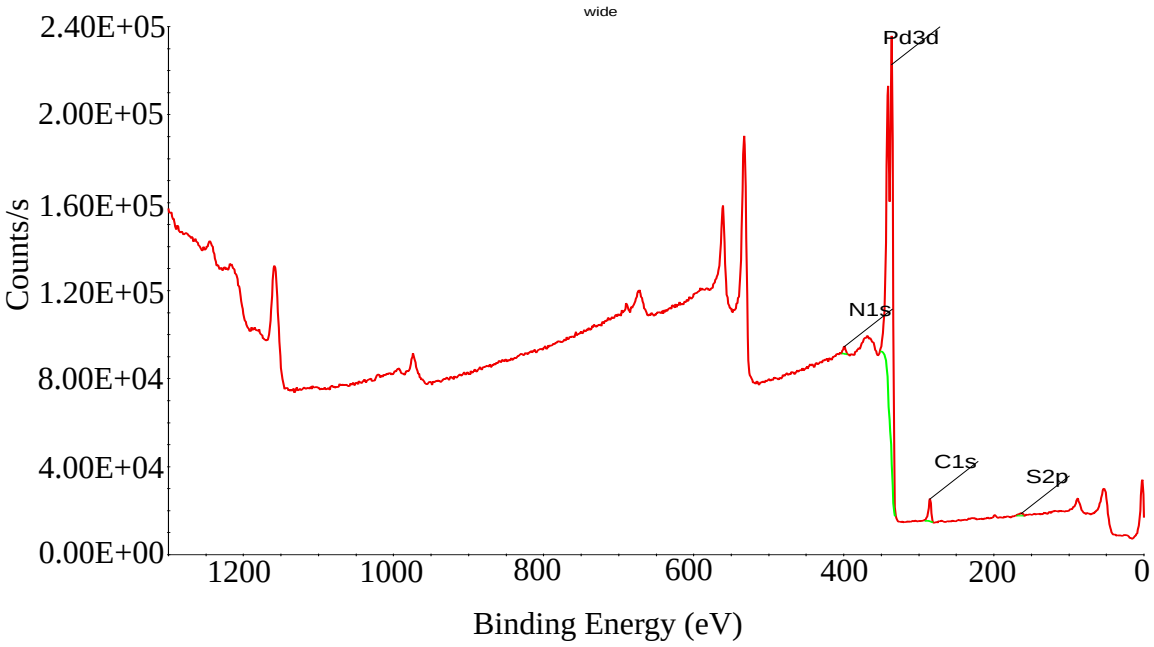


Figure 16. XPS Spectrum of Imidazole Dithiocarbamate on Pd Substrate

2.5.2 FTIR

FTIR measurements can verify the successful SAM modification on the metal surface by checking the particular set of functional groups of the desired recognition molecule. The spectra were recorded using a Nicolet (Thermo-Fisher) 6700 Fourier Transform Infrared Spectrometer[56].

For recognition molecule powder analysis, the single-reflection attenuated total reflection (ATR) accessory (Smart Orbit) will be installed. A durable diamond crystal and a swivel pressure tower on the sample stage will ensure consistent pressure from sample to sample. As for SAM analysis, the instrument is also equipped with an advanced surface gazing angle accessory (Smart SAGA) accessory, designed for the analysis of thin films on reflective substrates. The infrared light incidence for this reflection-absorption accessory is fixed at 80°, allowing sensitive measurements of films as thin as 0.1nm.

For SAM modification analysis, a background spectrum is recorded first on a clean bare metal substrate. As the control group, the same substrate will be deposited in pure solvent without adding any recognition molecule. As the experimental group, the background of another substrate is collected before it is treated in the solution with recognition molecules. In all these monolayer analysis, an 8mm diameter circular sampling area on the substrates was exposed to the beam, and the spectra was collected at

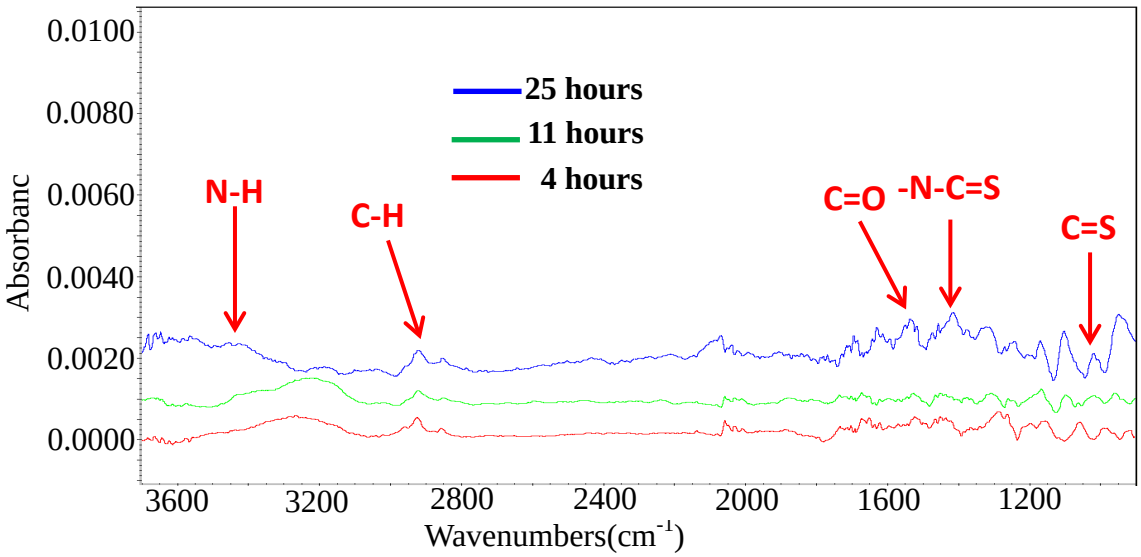


Figure 17. An Imidazole Dithiocarbamate Monolayer FT-IR Spectrum

a 4cm^{-1} resolution. Depending on the signal strength, a typical 128, 256, or 512 data point scan was performed.

Figure 17 shows a FT-IR time course analysis of Pd surface modification using the imidazole dithiocarbamate recognition molecule. The peaks around 2900 cm^{-1} are assigned to anti-symmetric and symmetric stretching of the CH_2 groups on the flexible carbon chain. This is a strong indication that the desired recognition molecule was successfully modified onto the Pd substrates.

FT-IR can only guarantee the modification of desired recognition molecule on the metal surface, but no monolayer thickness data can be obtained from this analysis technique.

2.5.3 Ellipsometry

Optical ellipsometry provides a measure of thickness of the organic layer absorbed onto the metal film.[57] The estimated length of the imidazole dithiocarbamate molecule from a ChemDraw 3D model is about 12Å. These data (Table 3) suggest that there are only single monolayers of molecules on the surface. The thickness data recorded is the average of the thickness values for five points (the center and the four corners) on a surface.

Table 3. Time Course Ellipsometry Analysis of SAM Thickness

12h	**24h**	**48h**
10.3±0.6 Å	12.2±1.4 Å	13.2±1.1 Å

2.5.4 Contact Angle

Table 4 shows an example of a time-course contact angle analysis for the imidazole dithiocarbamate recognition molecule monolayer on Pd surface. The contact angle of the bare palladium substrate is around 6° after hydrogen annealing surface cleaning, as a comparison.

Table 4. Time Course Contact Angle Analysis

12h	**24h**	**48h**
49.7±2.7	49.8±2.4	49.6±2.9

2.5.5 STM Surface Scanning Image

STM imaging can be done to obtain the bare or modified surface topology information. The images were taken using an Agilent STM. Movement of the tips and the tunneling parameter setup were controlled by the PicoSPM software package. Pt-Ir (90:10) tips were typically chosen for imaging done in the ambient air because of the higher image resolution and easy tip fabrication. HDPE coated tips must be used for imaging in the buffer solution. Figure 18 shows the clear difference of an Au surface before and after SAM functionalization (imidazole dithiocarbamate), imaged with a bare Pt-Ir tip in the air. Figure 19 and Figure 20 show the difference of a Pd surface imaged

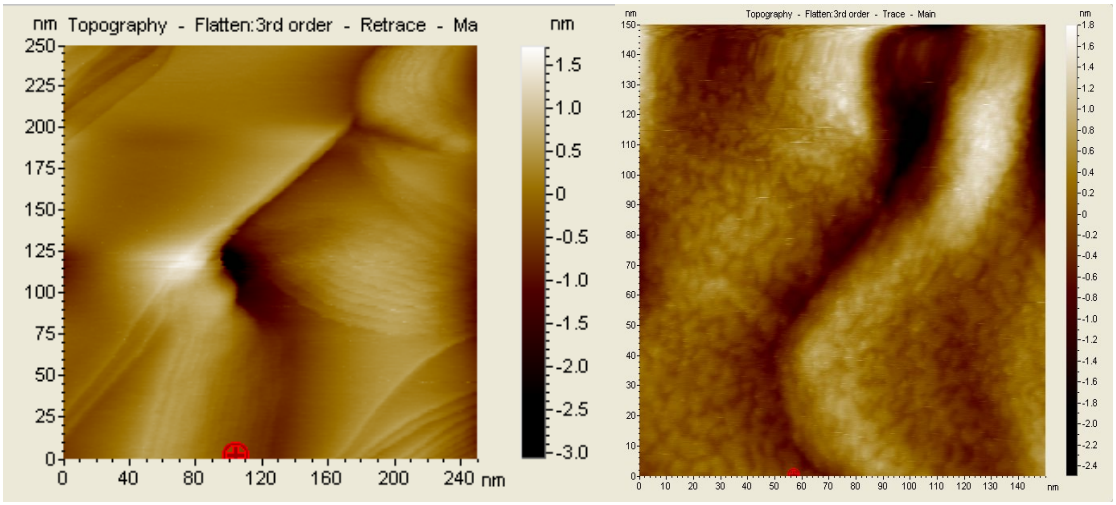


Figure 18. Comparison of Au Surface without and with SAM Functionalization with a HDPE coated Pd tip in 1mM Phosphate buffer solution before and after both were functionalized with a SAM (imidazole carboxamide).

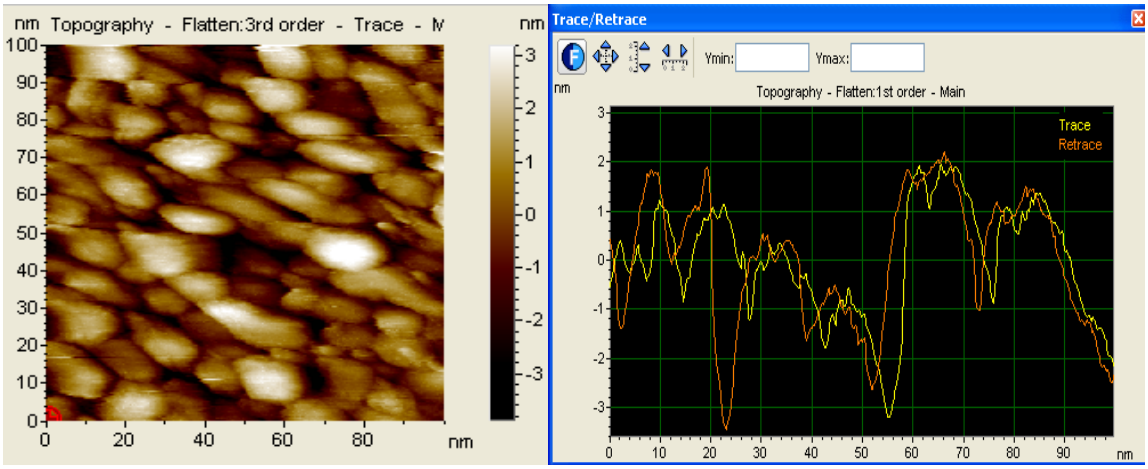


Figure 19. Bare Pd Surface Imaged with an Unfunctionalized Pd Tip in Buffer

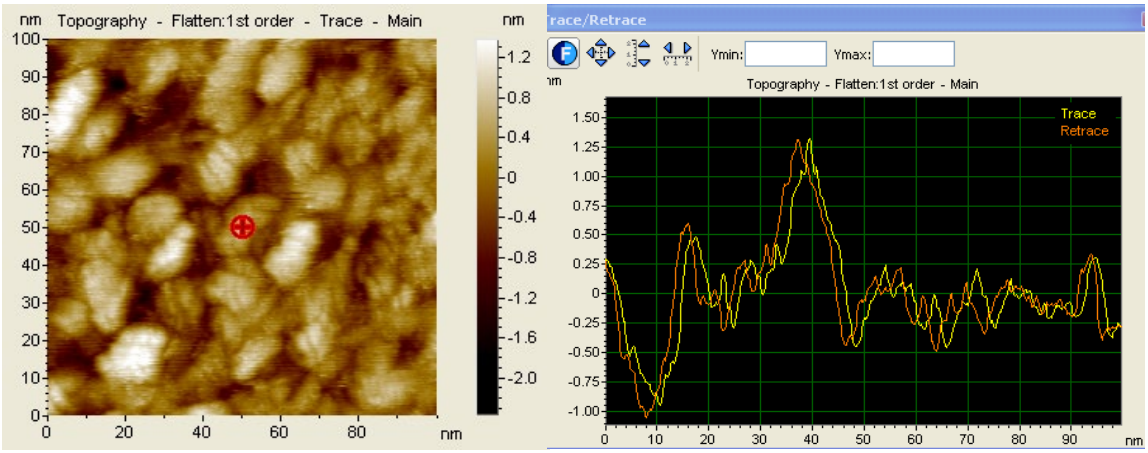


Figure 20. Functionalized Pd Surface Imaged with a Functionalized Pd Tip in Buffer

2.6 Fabrication of STM Tips

The detailed Au STM tip fabrication procedure was published.[58] First of all Au, Pd or Pt tips were etched from 0.25mm Au/Pd wire (California Fine Wires). The tips were then insulated with high density polyethylene (HDPE, from Sigma) to make the very top exposed with a linear dimension a few tens of nm. A quick ethanol and water rinsing was then performed to clean the tips before functionalization. There is no

comparable tool to FT-IR or ellipsometry for the functionalization analysis of a STM tip. Compared to unfunctionalized tips, functionalized tips always show clear background and telegraph noise-like recognition tunneling signals. Control experiments done with bare coated tips were noisier and do not show regular recognition tunneling signals.

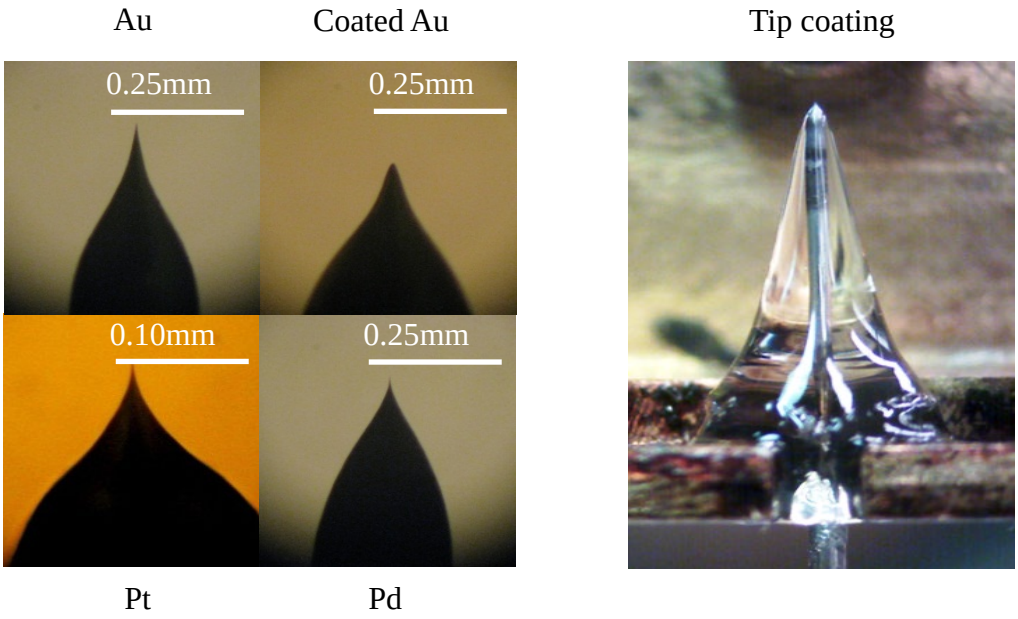


Figure 21. STM Tips

2.7 STM Recognition Tunneling Experiments

2.7.1 Recognition Tunneling Experimental Method

Nucleoside 5'-monophosphates were purchased from Sigma or UBS and used without further purification. They were dissolved in a 1mM phosphate buffer (pH7.4) to a final concentration of 10μM. Control experiments using the buffer solution prepared with Milli-Q Water Purification System (organic carbon contamination TOC<5ppb) were cleaner compared to experiments using Nanopure Water.

Glass vials for making solution and Teflon liquid sample cell on the STM stage were rigorously cleaned between experiments.

Tunneling measurements were carried out in 1mM phosphate buffered aqueous solutions using a picoSPM (Agilent, Chandler, AZ) with a Labview user interface for data acquisition. In -0.5V 6pA tunneling condition with imidazole carboxamide recognition molecules, the tunneling gap distance was measured to be roughly 2.4nm.[59]

2.7.2 Time Traces of DNA Nucleoside Monophosphates Recognition Tunneling Signals

Examples of current versus time traces for the pure phosphate buffer and solution containing one of five DNA nucleoside monophosphates are shown in Figure 22 (dmCMP is 5-methyl-2'-deoxy-cytidine-5'-monophosphate, which is important in

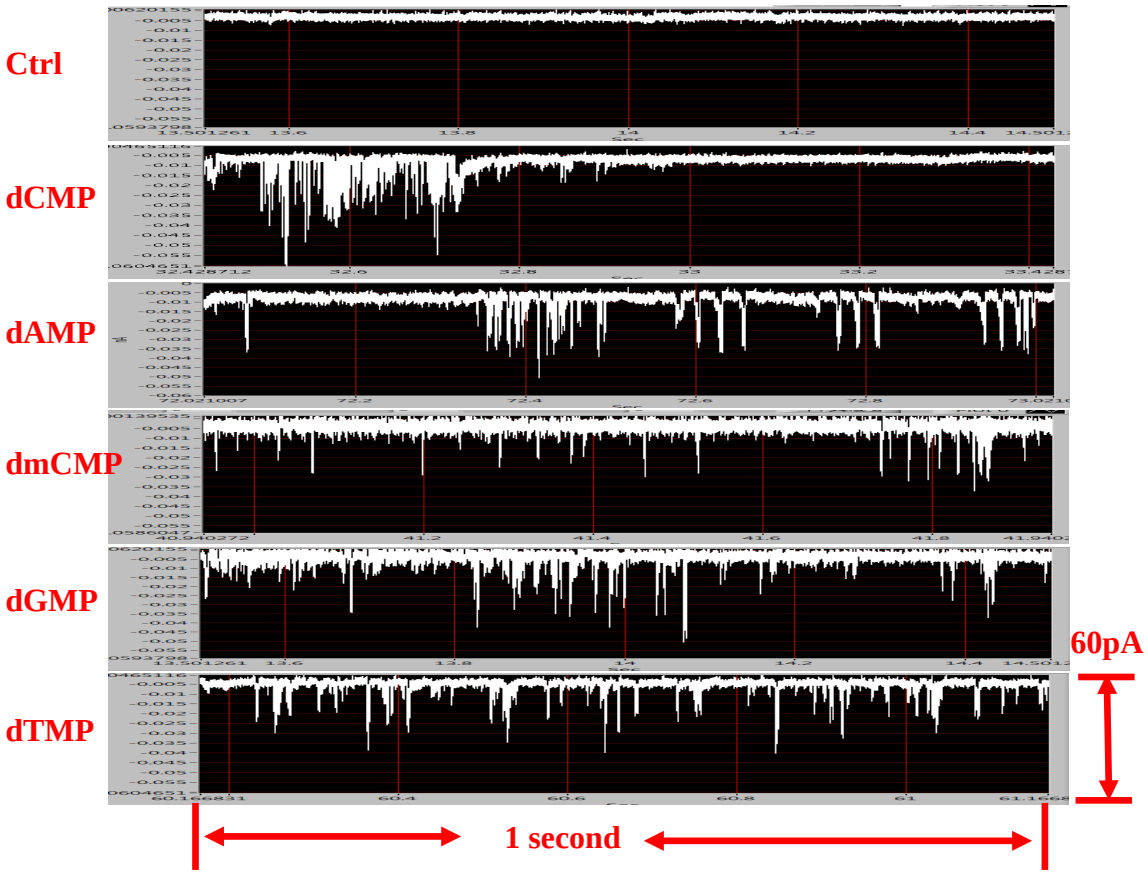


Figure 22. Typical Nucleotides RT Signal (Imidazole dithiocarbamate Reader)

epigenetic study). These signals were collected in a -0.5V, 2pA tunneling condition. The control signal was essentially clean. The five nucleoside monophosphates produced characteristic recognition tunneling signal spikes. The stochastic spikes were composed of both telegraph noise-like spikes and irregular untypical low frequency spikes. All parts of the signal were used for training with SVM machine learning algorithm. Signal sets for five nucleoside monophosphates in other tunneling conditions (like that listed in Table 5) can be easily collected with clean controls. The spike frequency, in different tunneling condition, varies from 1 spike/sec to 20 spikes/sec.

2.7.3 SVM Analysis of Tunneling Signals

The spike physical properties certainly contain valuable analyte-specific chemical information. However, the stochastic nature of the signals and the wide distribution of spike physical properties (like spike amplitude, spike width, spike shape as characterized by Fourier and wavelet analysis[45], spike frequency, on-time, off-time, distribution of spikes in a cluster) makes it extremely hard and time-consuming to manually analyze and differentiate bases according to these properties.

To solve the problem, support vector machine (SVM) a multi-parameter machine learning analysis algorithm[60] was used to analyze data for base calling with high accuracy even on one single molecule read.[11] Firstly, training of a support vector machine is done by taking random selections of the various parameters. Once the parameter set that generates the highest true-positive rate of calling signal spikes is found, the accompanying support vectors are used to assess each spike in the known data stream for base calling accuracy.

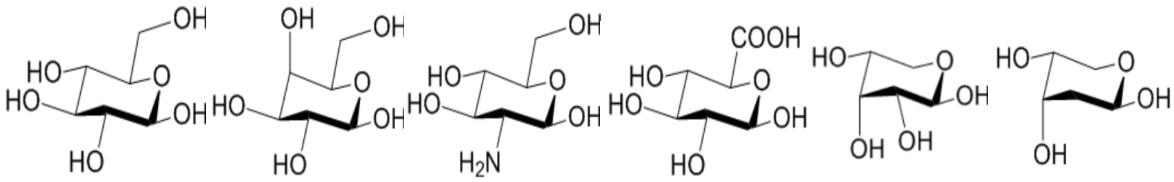
SVM analysis was done to the data obtained from Pd electrodes for imidazole dithiocarbamate recognition molecule in different tunneling conditions listed in Table 5. Results showed that more than 85% of the spikes can be assigned to the correct bases in -0.5V 2pA or -0.5V 6pA tunneling condition. These two tunneling conditions were used for DNA base differentiation with imidazole carboxamide recognition molecule on Pd and Au electrodes.[45] Referring to Figure 10, this can be understood easily because imidazole carboxamide and imidazole dithiocarbamate molecules share the same chemical groups contacting the analytes in the tunneling gap. This may indicate that the functional groups on the top of the monolayer plays a more influential role in DNA base differentiation performance than the underneath metal-recognition molecule interface would.

Table 5. SVM Analysis Result for dA/dT/dC/dG/dmC Differentiation

Tunneling Condition	Conductance	Known Peaks	All Peaks	Useful Peaks	True Positive Rate/Peak
-0.5V 2pA	4pS	25940	30552	84.9%	**87.7%**
-0.4V 2pA	5pS	16300	20295	80.3%	69.9%
-0.3V 2pA	7pS	13230	16184	81.7%	77.2%
-0.5V 4pA	8pS	15966	18730	85.2%	80.1%
-0.2V 2pA	10pS	3709	5266	70.4%	76.7%
-0.4V 4pA	10pS	10208	13813	73.9%	58.7%
-0.5V 6pA	12pS	23862	27344	87.3%	**86.1%**
-0.3V 4pA	13pS	12744	16139	79.0%	67.8%

2.7.4 Recognition Tunneling Signals for Monosaccharide Molecules

Given the simplicity and power of recognition tunneling technique in differentiating DNA bases, I also tried to exploit this technique for monosaccharide molecules differentiation with imidazole carboxamide recognition molecule.



Glucose Galactos Glucosamin Glucuronic acid Ribose Deoxyribose

Figure 23. Dominant Forms of Various Monosaccharide Molecules in Aqueous Solution[7]

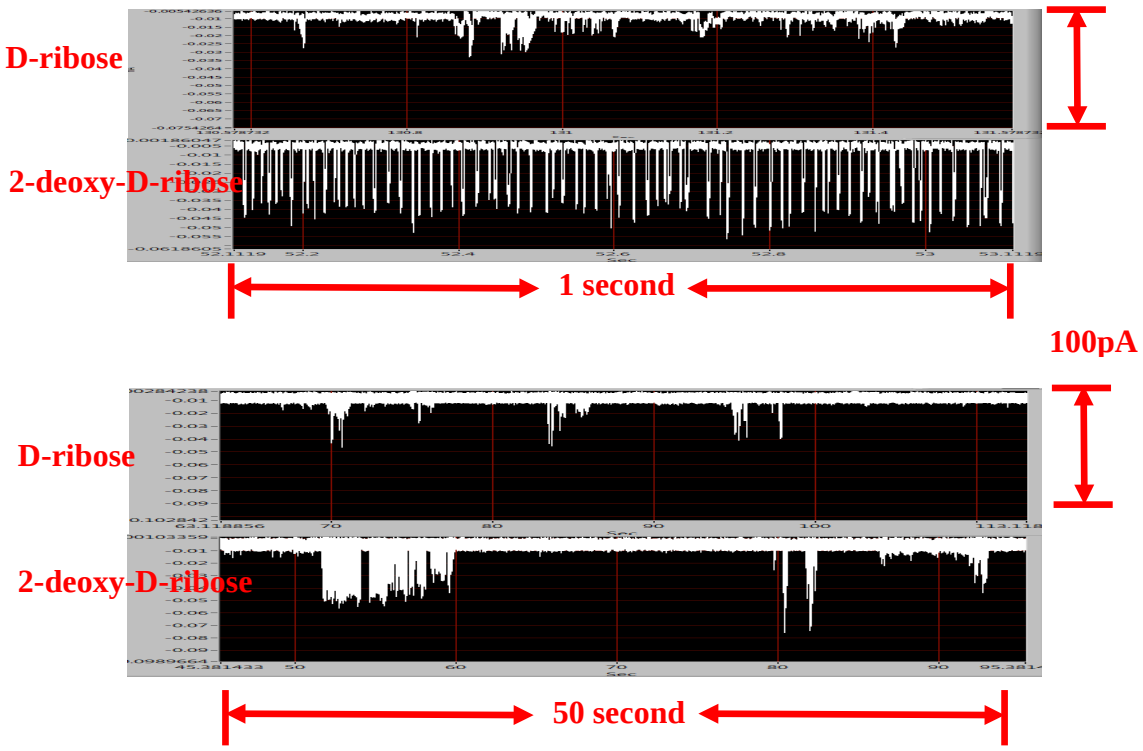


Figure 24. Typical Ribose and Deoxyribose Recognition Tunneling Signals

Figure 23 shows the dominant forms of different monosaccharide molecules in aqueous solution tested in my recognition tunneling experiments. Monosaccharides D-glucose, D-galactose, D-ribose, 2-deoxy-D-ribose, D-glucosamine, D-glucuronic acid

were purchased from Sigma Aldrich (>99% purity). Again, 100uM monosaccharides were dissolved in 1mM phosphate buffer (pH7.4) made using water from a Milli-Q system with minimal organic carbon contamination (TOC<5ppb).

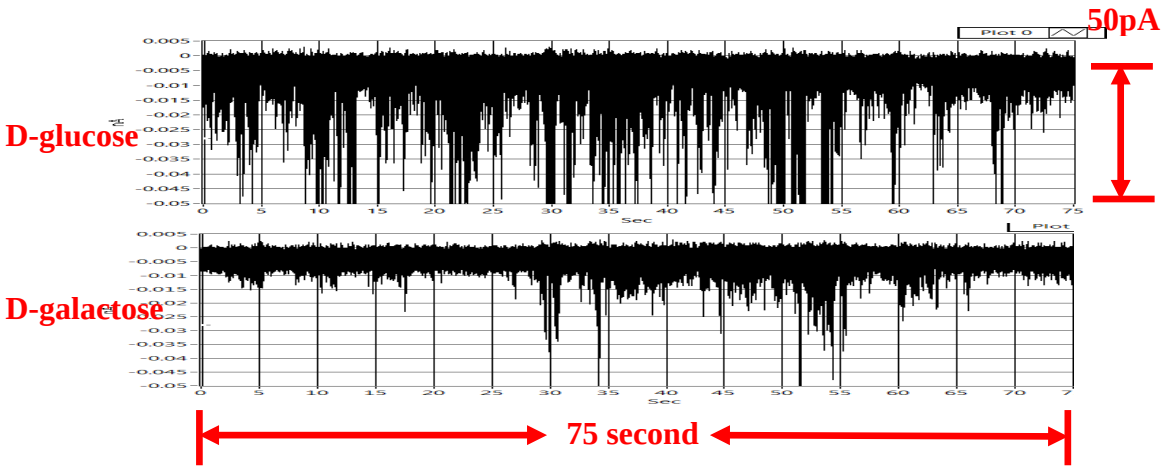


Figure 25. Typical Glucose and Galactose Recognition Tunneling Signals

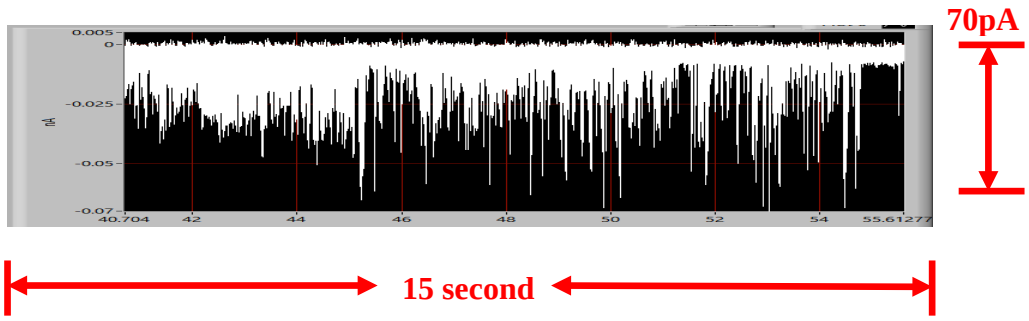


Figure 26. Typical Glucuronic Acid Recognition Tunneling Signal

Figure 24-26 and Table 6-8 show the typical recognition tunneling signals and SVM differentiation results from two pairs of monosaccharide molecules and D-glucuronic acid. Interestingly, for D-glucosamine, a molecule with a positive charge in pH7.4 buffer solution barely showed any telegraph noise like spikes in the experiments, whereas D-glucuronic acid, a molecule with a negative charge in the same buffer solution showed a lot of spikes in the same tunneling condition. This may indicate that for a

negative -0.5V tunneling bias applied to the substrates, negatively charged analyte can be more easily trapped in the tunneling gaps and generate tunneling signals than positively charged analytes could. Probably, negatively charged analytes could collect near the positively biased STM tips, and positively charged analytes would stick to the negatively charged substrates. This is a result to be confirmed by further experiments and theoretical explanation. SVM differentiation for glucuronic acid and glucosamine is trivial.

Table 6. SVM Analysis Result for Glucose and Galactose Differentiation

Tunneling Condition	Conductance	Known Peaks	All Peaks	Useful Peaks	True Positive Rate/Peak
-0.5V 2pA	4pS	890	1272	70.0%	94.9%
-0.5V 4pA	8pS	2294	3122	73.5%	95.2%

Table 7. SVM Analysis Result for Ribose and Deoxyribose Differentiation

Tunneling Condition	Conductance	Known Peaks	All Peaks	Useful Peaks	True Positive Rate/Peak
-0.5V 2pA	4pS	3337	4118	81.0%	92.7%
-0.5V 4pA	8pS	2577	3391	76.0%	92.1%

Table 8. SVM Analysis Result for Glucose, Ribose and Deoxyribose Differentiation

Tunneling Condition	Conductance	Known Peaks	All Peaks	Useful Peaks	True Positive Rate/Peak
-0.5V 2pA	4pS	3606	4648	77.6%	88.8%
-0.5V 4pA	8pS	2880	3791	76.0%	90.3%

In summary, a high recognition tunneling signal differentiation percentage can be easily achieved for various monosaccharide molecules. So it is indeed possible to utilize

this technique for detection of carbohydrates by measuring their distinct electronic tunneling currents in a tunnel gap functionalized with recognition molecules.

2.8 Conclusions of STM Recognition Tunneling Experiments

In conclusion, recognition tunneling junctions can be readily assembled using either gold or palladium electrodes. With functionalized electrodes using imidazole dithiocarbamate recognition molecule, experiments in various tunneling conditions produce tunneling signals from all five bases including 5-methylcytosine. The signals were essentially free of an interfering background signal, and a base separation percentage as high as 87% was obtained.

For summary, monitoring distinct stochastic tunneling signal fluctuations between a target analyte molecule trapped between the recognition molecules gap provides exceptional insights into the chemical nature of the intermolecular interaction at a truly single molecule level. The implication for rapid single molecule DNA sequencing follows immediately.

2.9 Future Work: Building Recognition Tunneling Based Single Molecule DNA Sequencing Nanopore Device

Controlled translocation of ssDNA through nanopores and single base resolution along the DNA molecule are the basic requirements for a recognition tunneling based DNA sequencing device to be assembled and successfully commercialized.

AFM dynamic force spectroscopy measurements implied a long intrinsic bonding lifetime (2-4s) of the hydrogen-bonded DNA base-recognition molecule complex, even without the stabilizing influence of base-stacking.[61] This provides sufficient time for the analyte trapped in a tunneling gap to produce enough signals, and for them to be

analyzed. Besides, experiments with functionalized probe for successful identification of single base in short DNA oligos demonstrated that the recognition complex can form and fall apart rapidly in the tunneling junction. Taken all together, it is indeed possible to use recognition tunneling for single base differentiation along a long ssDNA, an experiment to be done in the future.

Figure 27 shows a proposed DNA sequencing device combining exquisite recognition tunneling techniques, delicate solid-state nanopore fabrication and superior CMOS parallelism. While current DNA sequencing platforms use complicated optics, expensive chemical labeling or fluidics, time-consuming DNA amplification, the versatile recognition tunneling based platform relies on none of those. The proposed platform allows for true single molecule, real-time, direct electrical analysis with massive parallelism possibility.

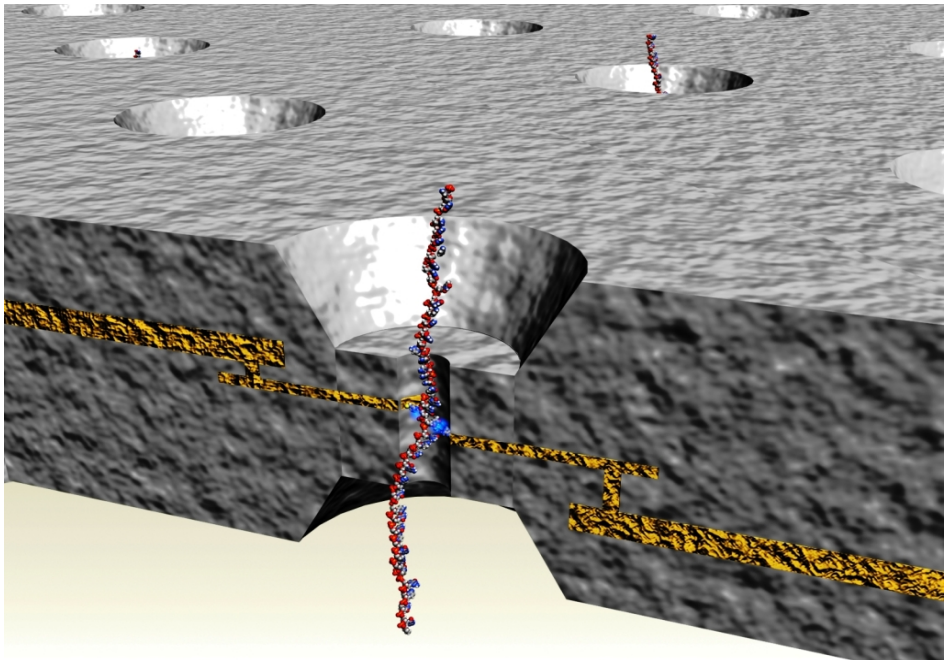


Figure 27. Illustration of a Parallel Recognition Tunneling Nanopore Device[3]

Since no sample amplification is needed before recognition tunneling readout, this technique can also be theoretically applied to sequence any biopolymers like RNAs or proteins. Demonstrated results in my thesis have already shown great promise for this technique to be integrated into rapid carbohydrates detection system.

CHAPTER 3 TRANSLOCATION OF SINGLE-STRAND DNA IN CARBON NANOTUBES

In this chapter, successful integration of single-wall carbon nanotubes (SWCNTs) and microfluidic systems for CNT mass transportation experiments will be shown. My work focuses on CNT growth, PDMS microfluidic device fabrication, ionic current measurement and qPCR quantification of translocated DNA molecules.

First, clean CVD growth methods for long SWCNTs will be explained in detail. Fabrication of the translocation chips was done by my colleagues in the lab. Integration of the devices will be demonstrated. Transient enhancement of ionic current in translocation experiments was observed when DNA translocated through the CNTs driven by electrophoretic force. Single-strand DNA translocation was further verified by PCR experiments. Quantitative PCR (qPCR) was performed to analyze the correlation between DNA translocation quantity and ionic current spike number. More experiments using nano-material for DNA translocation will be discussed at the end of this chapter.

3.1 Why using a carbon nanotube for DNA translocation?

Transport of individual DNA molecules with fixed length through channels having cylindrical symmetry has already been studied using artificial solid-state nanopores[62] or biological nanopores (α-hemolysin[29] or MspA[30]). In all of these studies, by assuming that the presence of the molecule in the channel will disturb the motion of ions that are trying to go through and carry current from one side to the other, the translocation event was signaled by a change in pore electrical conductance. However, for a nanopore-based DNA sequencing device there are four challenges that need to be addressed:

(1) The random motion of DNA in solution due to poor confinement of a DNA strand in a nanopore (e.g., in a 3.4 nm thick pore, only 10 bases or ~0.1% of a 10 kilobase DNA is confined): Majority part of the DNA strand is unrestrained and free to wiggle quickly, form secondary structure, or tangle. This makes it difficult to control the DNA translocation, and creates a lot of noise in the ionic current measurement, which could complicate the signal from a single base for DNA sequencing.

(2) The single molecule sensitivity: It is yet to be seen whether the detection of electrical signals will be sensitive enough to resolve two adjacent unknown DNA bases.

(3) The translocation speed control.

(4) Large-scale accurate and easy fabrication, and assembly of the required nanoscale components.

One way around these problems may be to build nanochannels rather than nanopores. Such nanochannels would restrict the DNA over its entire length rather than at a single point, and if they could be made with a diameter much less than the persistence length, then the energetic cost of a hairpin would be too large and one would expect to see a straight DNA molecule moving by without any kinks.[63] A Double-stranded DNA(dsDNA) has a persistence length (a measure of how far the backbone of a polymer could go in a straight line before thermal fluctuation turns it in another direction) at physiological salt conditions of about 50 nm. A single-stranded molecule (ssDNA) has a significantly smaller persistence length of about 3 nm. Any nanochannels designed to thread a DNA molecule must be no more than 10 nm in diameter for dsDNA and 2 nm in diameter for ssDNA.[63] Fabrication of a 1-D nanogap inside nanofluidic channel for analyzing real-time, label-free dsDNA translocation was demonstrated.[64] dsDNA

translocation in silica nanotubes with accompanying transient changes of ionic current was also shown.[65] 1-D nanoscale structure offers three distinct advantages over traditional nanopore devices. First, these 1-D nanochannels are unique in their high aspect ratio such that they can confine the entire biomolecule, which is likely to result in additional ways to control translocation. Second, these devices extend the time scale of molecule transport events substantially due to friction, which could provide more information on the translocation behavior of DNA molecules. Third, these devices could be more easily fabricated and integrated into nanofluidic circuits.

Carbon nanotube (CNT) provides attractive features ready to be integrated in a nanofluidic device, such as easy fabrication, unique electronic properties, atomically smooth inner surface and low-friction liquid transport due to hydrophobic graphitic surface. The transport of water[66], ions[67] and nucleic acids[68] through CNT has been extensively modeled. The experimental analysis and control of DNA molecule translocation through CNTs will be the foundation stone for such application in the future. Supported by the results shown in detail below, CNT holds the promise to be integrated into next-generation ultra-rapid DNA sequencing device.

3.2 Growth Of Ultra-Long Single-Walled Carbon Nanotubes (SWCNTs)

3.2.1 CNT Growth Mechanism and Problem to Be Solved

All working CNTs used in the experiments were grown by chemical vapor deposition (CVD) process on the silicon oxide or silicon nitride wafers. As shown in Figure 28, the diameter of the nanotube depends largely on the size of the metal nanoparticle catalyst.[69] At the beginning, my colleagues used the block copolymer method[70] for the cobalt catalyst synthesis and ethanol as the carbon source[71]. At high

temperature, precipitated carbon atoms have limited solubility in the metal nanoparticle surface, and tubular carbon in sp^2 bonding structure is thus formed on the catalyst surface leading the SWCNTs growth.

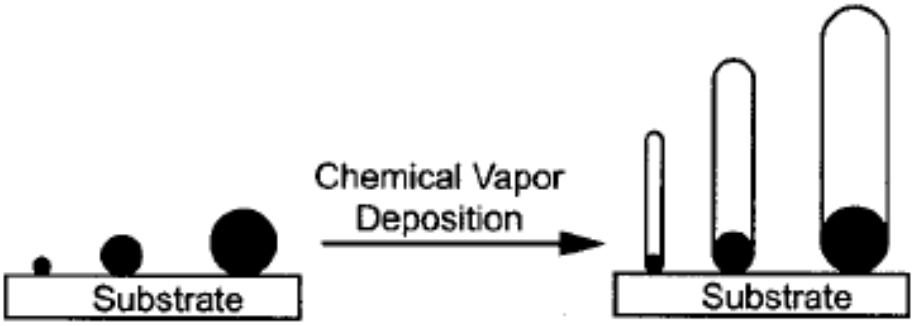


Figure 28. CVD CNT Growth Mechanism

The SWCNTs grown by this block copolymer process tend to be long but small, with a diameter of ~2nm. Still, it is much larger than the width of ssDNA molecule (~1nm), so the theoretical translocation of ssDNA through a carbon nanotube is possible. However, it is hard to fine tune the CNT size by this way. Besides, the cobalt catalyst synthesis involves toxic toluene. The catalyst prepared will not be active any more after one week, so we have to prepare new batches of catalyst every week. This method overall is tedious and time-consuming. An improved and efficient carbon nanotube growth method was desperately needed.

3.2.2 Iron Nanoparticle Method

In order to make a diameter-controlled synthesis of carbon nanotubes, iron nanoparticle method was first tried. Iron nanoparticles with increasingly larger diameters were shown to lead larger diameter carbon nanotube growth.[72]

Figure 29 shows the iron nanoparticle synthesis mechanism, as described in detail in a 2002 article from Charles Lieber's group in Harvard.[72] $Fe(CO)_5$ was mixed with long chain organic acid in dioctyl ether. The mixture was then fluxed at harsh 286°C temperature for 1-3h to yield iron nanoparticle solution with distinct and nearly monodisperse diameters. The length of the organic acid determines the diameter of the iron nanocluster and we can control the diameter of the tubes by adding oleic acid, lauric acid or octanoic acid respectively. It turned out that the iron nanoparticles can be synthesized successfully, as shown in Figure 30. But the carbon nanotubes grown by this method was curly, bended with bamboo nodes, short or multi-walled (Figure 31), though the outside diameter measured by AFM (Agilent, tapping mode) can be large (Figure 32).

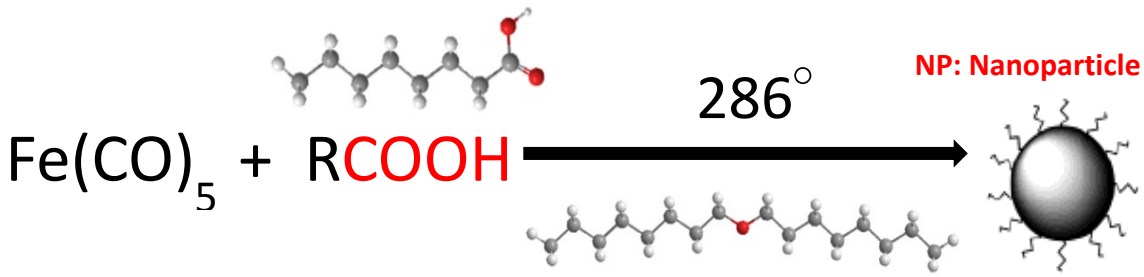


Figure 29. Iron Nanoparticle Synthesis Mechanism

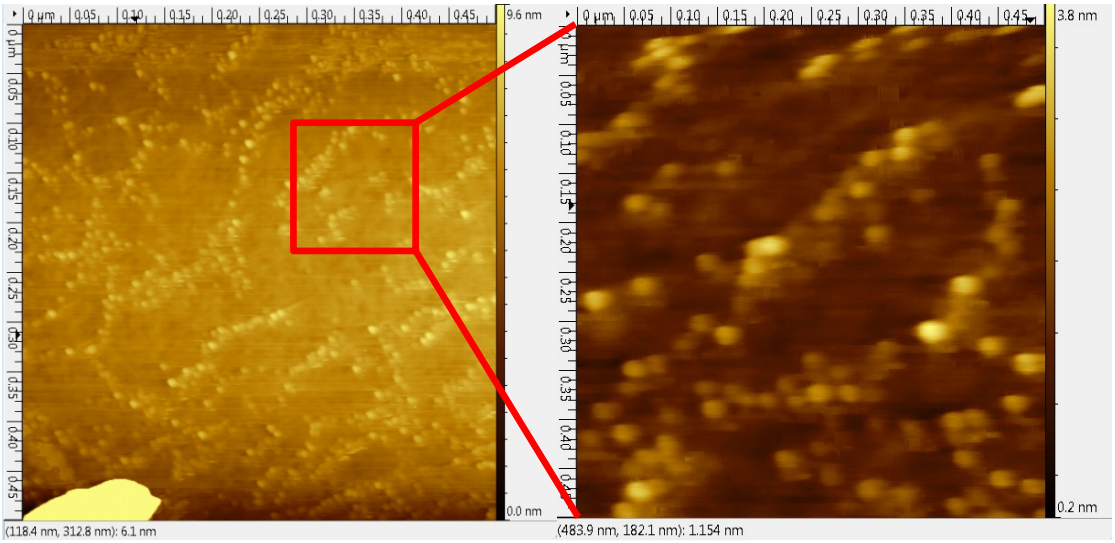


Figure 30. AFM Image of Iron Nanoparticles on the Substrate

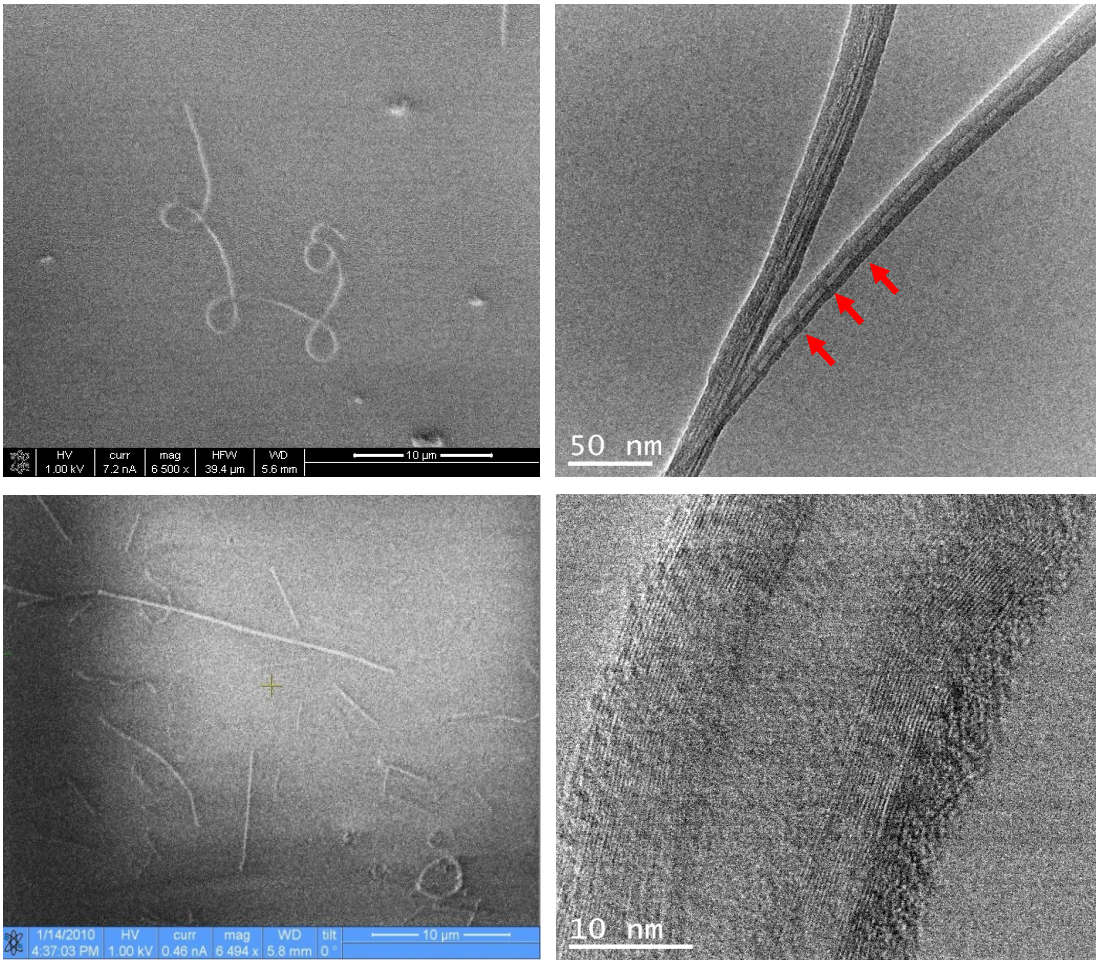


Figure 31. CNTs Grown by Iron Nanoparticle Method

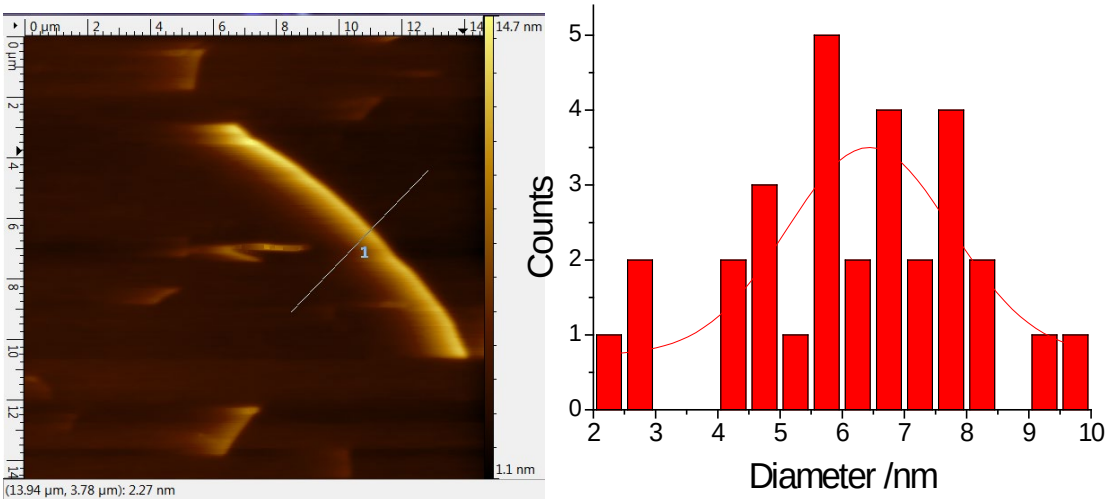


Figure 32. Use AFM to Measure CNT Diameter Grown by Iron Nanoparticle Method

3.2.3 Ferritin-based SWCNT CVD Growth Method

For DNA translocation device fabrication, we need long and clean SWCNTs. Figure 33 illustrates the SWCNT CVD synthesis process using the ferritin-based method. The detailed artificial ferritin synthesis by Fe atoms loading into apoferritin was

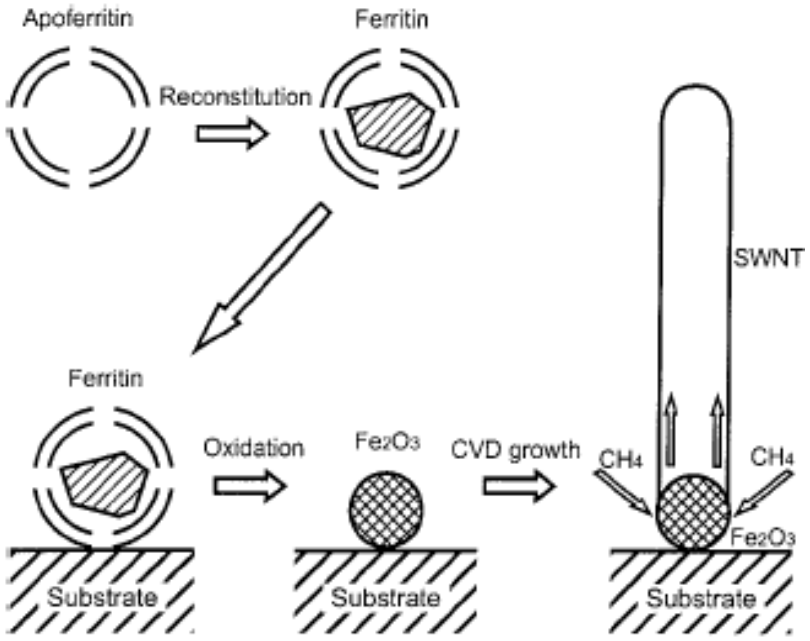


Figure 33. CVD SWCNT Synthesis Using Ferritin-based Method[8]

described elsewhere.[8] Briefly, apoferritin from horse spleen was added into a buffer solution. The most critical step was to control the loading quantity of Fe atoms into apoferritin core by adding different amount of the ammonium iron (II) sulfate into the apoferritin solution. An oxidative reagent was added later to fully oxidize Fe(II) to Fe(III). Prepared ferritins in the solution are chemically active for months, so this method is simple and highly effective. Deposition of ferritin on the silicon chips can be done simply by brushing synthesized artificial ferritin solution onto the edge of the chips. To obtain iron oxide nanoparticles, the chips were heated to 800°C in the air to remove the organic shell of the deposited ferritin protein. Discrete Fe_2O_3 nanoparticles can be thus formed on the chips as the catalyst for SWCNT CVD growth. Figure 34 shows the successful deposition of ferritins onto chip surface after 800°C calcination.

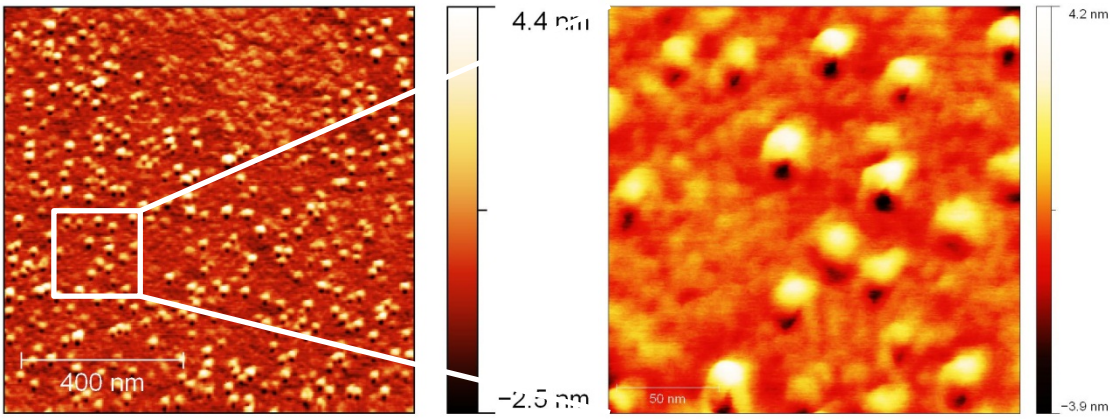


Figure 34. AFM Image of The Surface After Treatment of Ferritin Organic Shell Removal

SWCNTs grown by this method can be extremely long (Figure 35). The length of the tubes was found to be dependent only on the chip dimension and growth time. A centimeter long tube can be easily grown. Several devices can therefore be made on one single tube because of the clean and long growth. TEM can then be used to examine if the

tubes grown this way were single-walled. It turned out that all the tubes grown by this process were indeed single-walled, shown in Figure 36 TEM images.

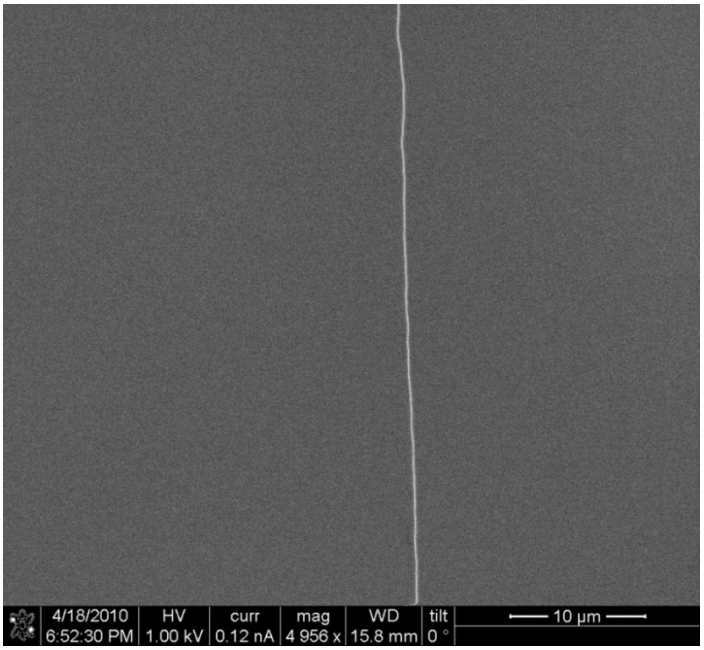


Figure 35. Long and Clean Growth of SWCNTs Using Ferritin Method

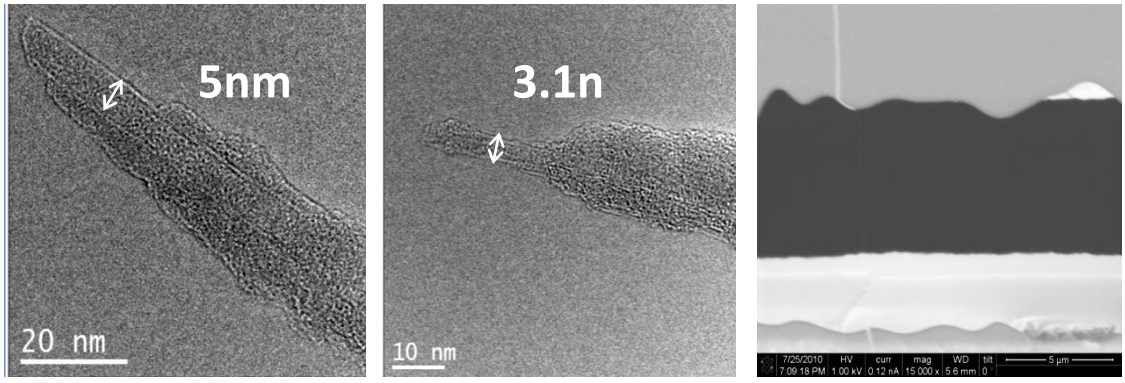


Figure 36. Ferritin-derived SWCNT on TEM grid.

3.3 PDMS Nano-Fluidic Device

3.3.1 Fabrication of DNA Translocation Chips (My colleagues' work)

After CNTs growth, the entire chips were spin-coated with a thin layer of polymethylmethacrylate (PMMA). Gold markers were deposited onto the wafer surface. SEM images were then taken for accurate CNTs location positioning. Ultrahigh-resolution electron beam lithography was used to open two reservoirs on the PMMA according to the gold markers. After development using dichloromethane, the patterned

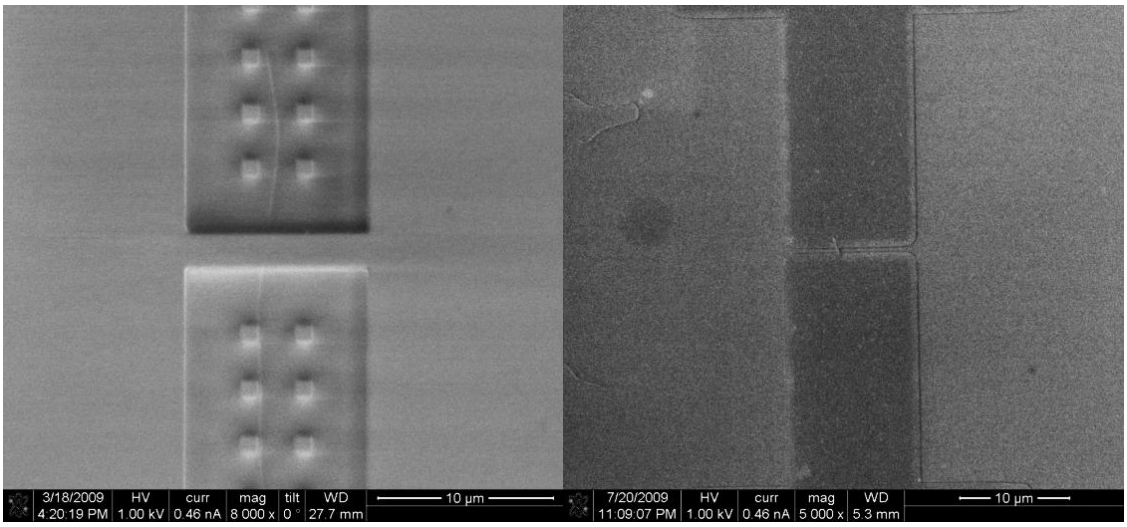


Figure 37. CNT Before and After Translocation Experiment

reservoirs were opened in the air. Shown in Figure 37, Oxygen plasma treatment (7.2W at 550~600 mTorr for 26sec[73]) was used to etch away the regions exposed in the reservoirs and to open the CNTs. Microfluidic PDMS cover introducing buffered DNA solution was aligned onto the reservoirs under the microscope. This nanofluidic device was finally ready for CNT mass transport experiments. Figure 38 is the schematic of a CNT nanofluidic device that features a single SWCNT bridging two microfluidic channels.

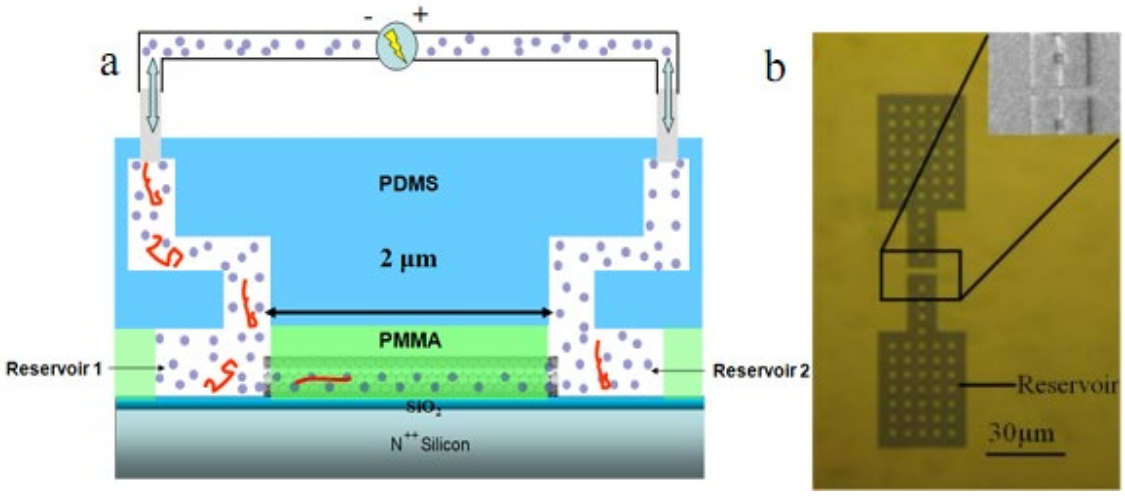


Figure 38. Chip Design for CNT Mass Transport Experiments

3.3.2 Fabrication of PDMS Microfluidic

Polydimethylsiloxane (PDMS) is one of the most common materials used for flow delivery in microfluidics chips.[74] Because it is chemically inert, optically clear, mechanically flexible, and it will not allow aqueous solution to infiltrate and swell the material itself. We use PDMS to create the microfluidic delivery channels, which introduces the DNA molecules into the fluid reservoirs of SWCNT based devices.

The fabrication of the mold of channel pattern by deep reactive ion etching on a silicon wafer was done by my colleagues in the lab. The wafer containing the mold is then salinized overnight with silanizing agent (tridecafluoro-1,1,2,2-tetrahydrooctyl trichlorosilane) to make the surface hydrophobic enough so that the PDMS material will not bind to the mold in the curing process. This wafer will be used as the master for creating PDMS device. The PDMS prepolymer (Sylgard 184 Silicone Elastomer Base, Dow Corning) is mixed with the curing agent at a volume ratio of 10:1. Air bubbles in the mixture can be removed effectively by degassing in a desiccator under vacuum (~40 torr). The mixture will then be poured onto the master wafer and hardened for 2 days at room

temperature before use. Two layers of PDMS microfluidic can be sealed in the oven after the oxygen plasma treatment (1min, 29.6W, at 550~600 mTorr), which can activate the material surface by making it hydrophilic. The fabricated PDMS microfluidic is finally brought into contact with the two PMMA reservoirs on the carbon nanotube translocation device under the microscope.

Figure 39 shows a fully packaged device. Additional features of the apparatus include an integrated flow system with inlet/outlet stainless steel ports that allow for continuous and rapid change of solutions in the reservoirs. PDMS channel surface was used without further surface treatment.

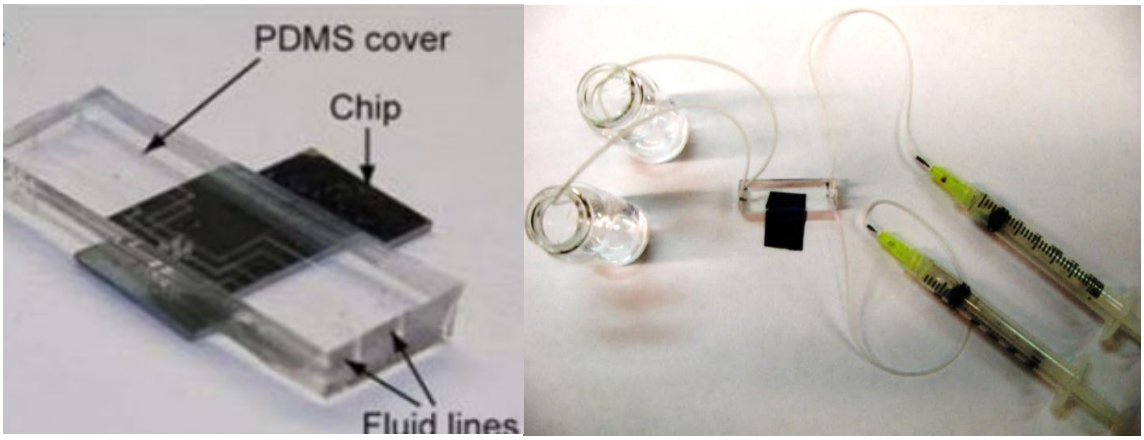


Figure 39. Assembled Nanofluidic Device for CNT Mass Transport Experiments

3.4 Design of Single-Strand DNA and PCR primers

The 60nt and 120nt ssDNA were designed with the help of Mfold software package (http://mfold.rna.albany.edu/?q=mfold). Briefly, a random sequence was inputted into Mfold for secondary structure prediction based on the nearest-neighborhood method.[75] The DNA bases were then changed to reduce the amount of hydrogen bonding pairs according to the secondary structure predicted by Mfold. The changed sequence

was inputted into Mfold software environment again for further optimization. An example of the final secondary structure for one of the 60nt ssDNA used in the experiments is shown in Figure 40. PCR primers were designed with the help of Primer Premier Software package.

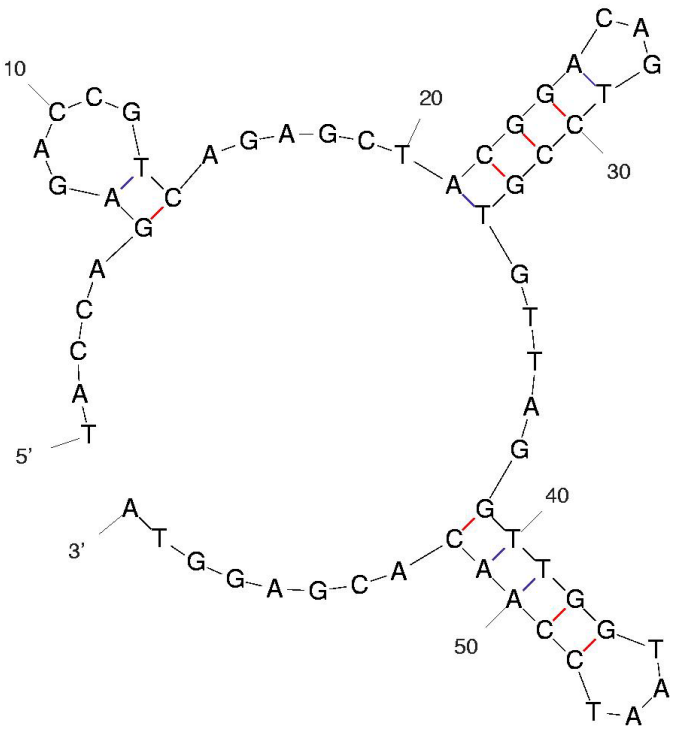


Figure 40. Mfold Prediction of Designed ssDNA Structure

3.5 Proof of DNA Translocation by Ionic Current Measurement

3.5.1 DNA Translocation Experiment Setup

In the experiments, both microfluidic channels were filled with 1M potassium chloride (KCl) buffer solution. All the DNA test solutions in the study were prepared with TE buffer, which consists of 10mM Tris-HCl, and 1mM EDTA in aqueous solution, pH 8.0.

Before adding any DNA solution into the inlet reservoir, we applied a positive bias for more than 10 minutes and collected the outlet solution as the negative control

group for further PCR analysis. The DNA solution was then introduced into the inlet reservoir and the same positive bias was applied. Ionic current was recorded with an applied voltage generated by an Axopatch 200B patch clamp amplifier (Axon Instruments). Translocated DNA molecules in the outlet reservoirs were collected into a pre-UV-sterilized centrifuge tube for further PCR analysis.

All the electric measurements were conducted in a clean condition to avoid dust contamination.

3.5.2 Translocation Signal of Single-Strand DNA

Figure 41 shows the ionic current signals for single-strand 60nt DNA translocation. Before adding any DNA molecules, no transient current change was observed although the baseline had shifted slightly over time. When single-strand 60nt DNA molecules dissolved in 1M KCl buffer were introduced into the negatively biased reservoir while the other one was filled only with 1M KCl buffer, the ionic current exhibited frequent transient enhancement. Moreover, all control experiments without adding DNA or without opening of CNTs by oxygen plasma did not exhibit such phenomena, so we concluded that the current enhancements corresponds to the passage of DNAs through CNT. The typical current change was 80-140pA above the baseline, and the typical spike duration time was 2-4msec. The frequency of current spikes was 4-6 spikes/min. Once the polarity of the applied bias was reversed, no similar spikes could be observed.

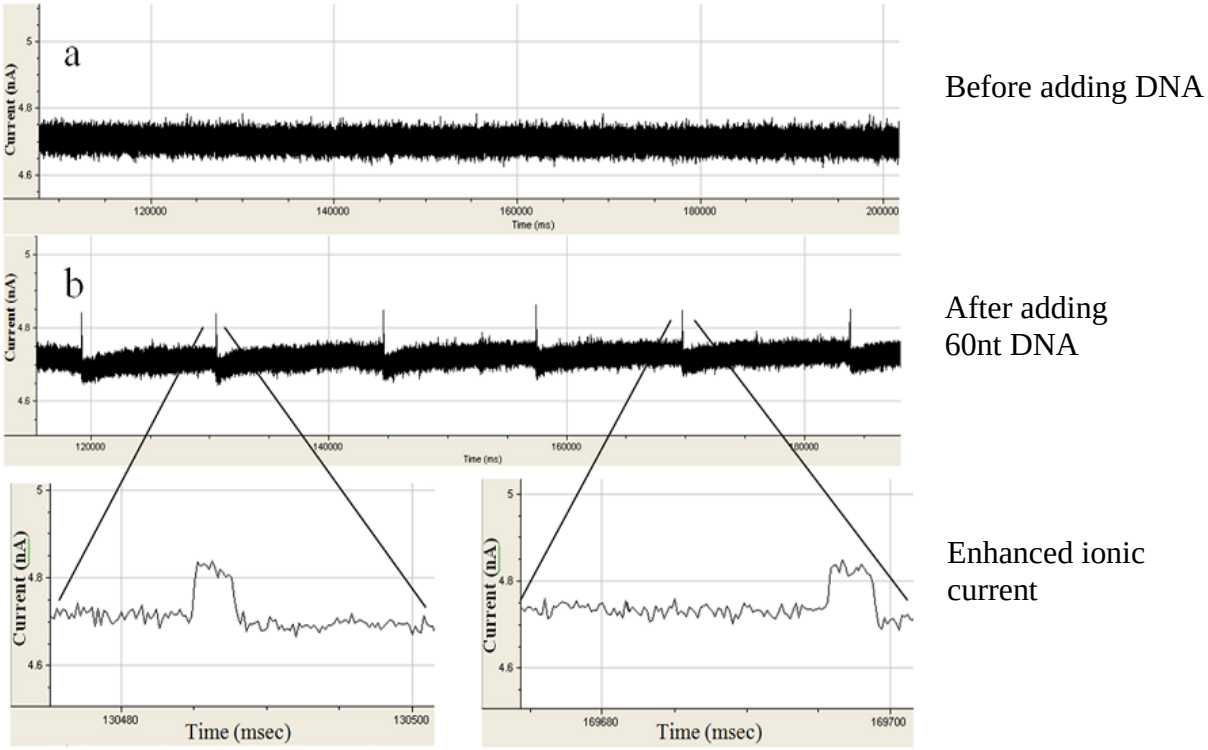


Figure 41. Typical ssDNA Translocation Ionic Current

Although we can verify that ssDNA can translocate through SWCNTs, the explanation of translocation physical process is still quite difficult. Theoretical calculation and computer simulation was done by the lab collaborators, which shows the increased ionic current was induced via ion-selective filtering by a DNA molecule translocated through a carbon nanotube, in contrast to the classic ion current blockade in nanopore DNA translocation experiments.[76] The subsequent work done by my colleagues in the lab explained the origin of the Giant ionic currents in carbon nanotube channels.[77]

3.6 Proof of DNA Translocation by Polymerase Chain Reaction(PCR)

3.6.1 PCR Experiments and the Results

If ssDNA can be translocated through CNTs, the solution collected from the reservoir other than the one injected with DNA containing solution should certainly contain ssDNA, which can be amplified by PCR to show a band on a gel. That was

exactly what happened in the experiments. Shown in Figure 42, clean and uniform 57bp PCR amplification product was verified by native PAGE gel analysis. Experimental group with longer translocation time showed a stronger band on the gel.

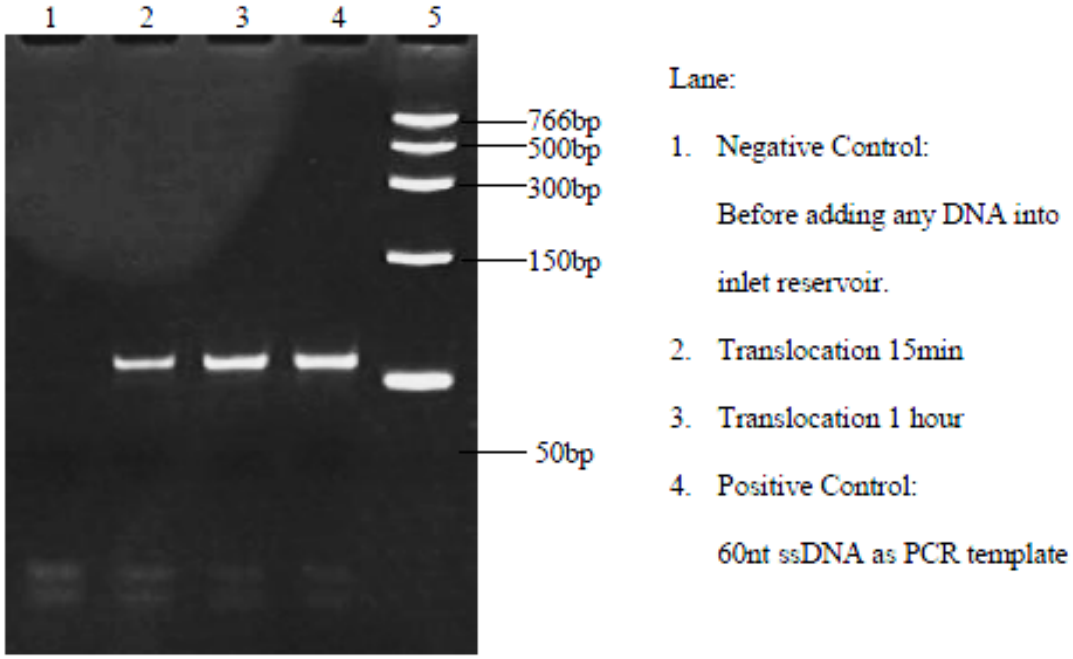


Figure 42. ssDNA translocation experiment PCR results

3.6.2 Quantitation of Translocated DNA Molecules Using q-PCR

Quantification was performed by quantitative real-time polymerase chain reaction (qPCR).[14] Before quantification of real sample, a qPCR standard curve was established using *Eppendorf Mastercycler® realplex* system. To create a standard curve, seven 10-fold dilutions, starting with 10^7 template copies and ending with 10 copies, were prepared. Figure 43 shows the real-time amplification profile using a 60nt ssDNA template. Figure 44 shows the 60nt ssDNA qPCR standard curve by plotting the log of initial template copy number in PCR tubes as the x value and the cycle threshold (Ct) as the y value. The formula $Cycle_{threshold} = -3.787 \log_{10}(\text{Initial copy\# in PCR tube}) +$

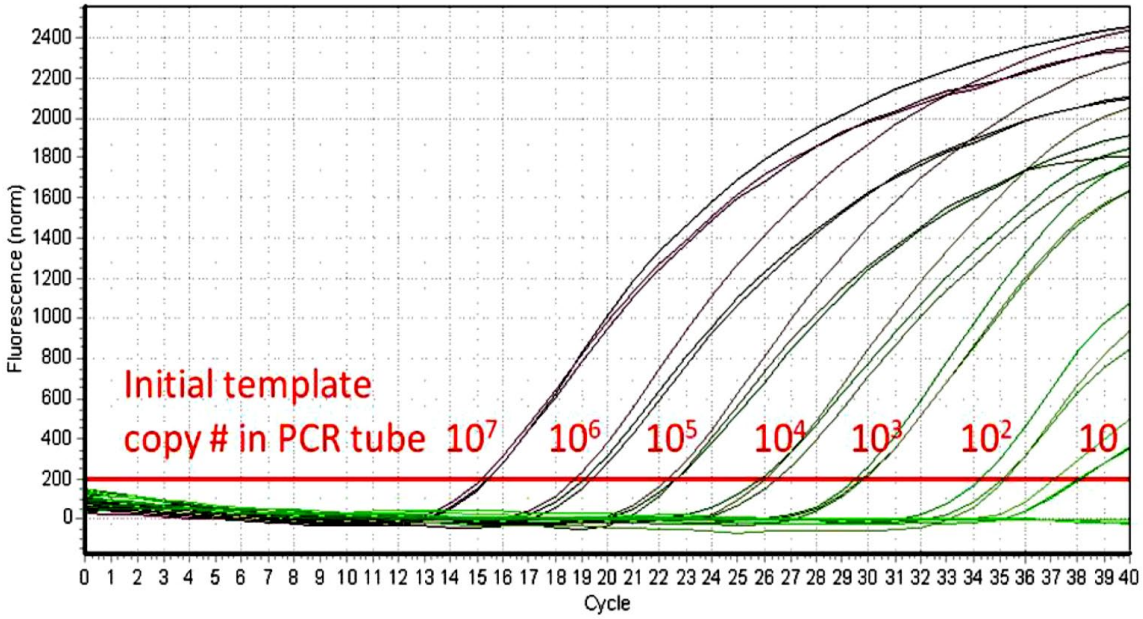


Figure 43. 60nt ssDNA Real-time Amplification Result

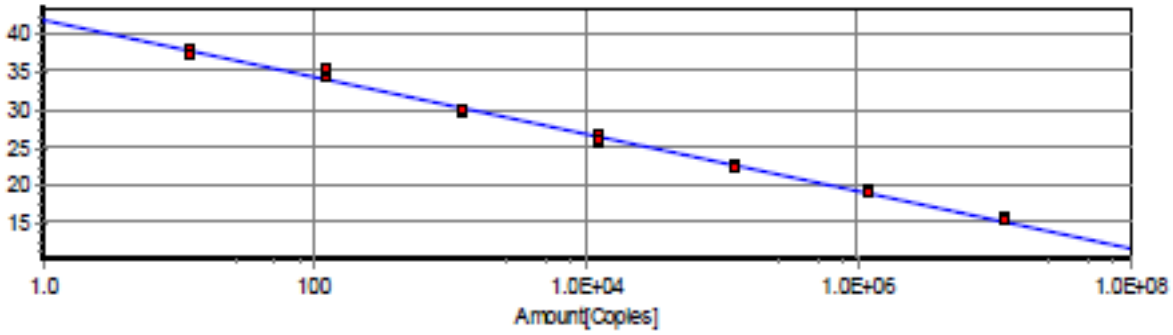


Figure 44. 60nt ssDNA qPCR Standard

$41.99(r^2 = 0.995)$ was used to estimate the initial template number in the solution collected after DNA translocation. The melting curves for the products were also transformed into the first-derivative melting peak as shown in Figure 45. There was only one peak in the melting curve, meaning only one amplicon was generated and there was no false priming or primer dimer.[78]

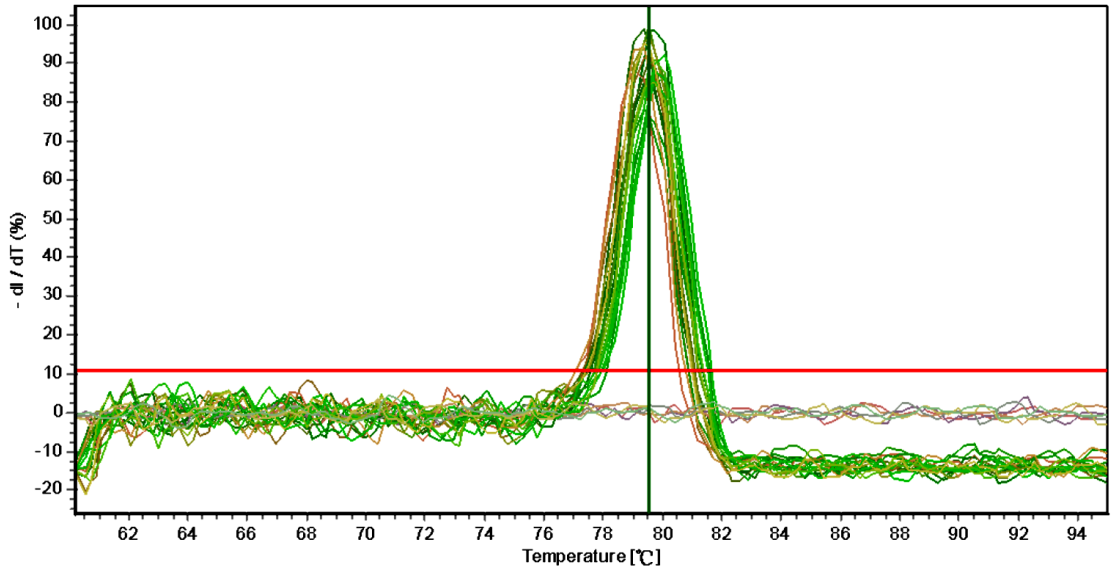


Figure 45. 60nt ssDNA qPCR Standard Melting Peak Plot (only one peak)

Ionic current spikes counting and ssDNA quantification results are shown in Table 9. Interestingly, the number of DNA molecules translocated was at least 10-fold more than the number of spikes for an unknown reason.

Table 9. Quantification Results of Translocated ssDNA Molecules

CNT	CNT Conductance/ nS (1 M KCl)	DNA Sample (nt)	Number of Spikes	Number of DNA Molecules	Molecules per spike
AD1	9.7	60	350 ± 50	8000 ± 2000	23 ± 10
AD2	9.5	60	30 ± 10	400 ± 200	13 (+17, –13)
AA New1	19.6	120	64 ± 10	8500 ± 3100	88 (+126, –88)
AA New2	2.7	120	1500 ± 200	24,400 ± 5700	16 ± 7
HL-4-1-36	9.6	60	36 ± 4	1224 ± 774*	34 ± 21*
A136	1.6	60	46 ± 5	1900 ± 200*	41 ± 10*

3.7 Conclusions of DNA Translocation in Carbon Nanotubes

The observations that the ionic current enhancement appeared (a) only when the inlet reservoir was injected with ssDNA containing solution but not with pure buffer solution, (b) only when an electrophoretic voltage for driving ssDNA through the channel was on but not off or reversed, (c) with correct PCR amplification products shown only in the translocation group but not in the negative control group suggested strongly that the observed signal was caused by ssDNA molecules passing through the CNTs between the inlet and outlet reservoirs.

Translocation of shorter ssDNA will be harder for PCR test, but it is possible to check the translocation of shorter ssDNA with a small diameter fluorophore attached. A combination of optical and electrical tandem detection of fluorescent single-molecule translocation through carbon nanotubes was later demonstrated by my colleagues in the lab.[79] More experimental and theoretical simulation results are still needed for us to understand the detailed physical process of DNA translocation through carbon nanotube device.

Using carbon nanotube nanofluidic devices for translocation geometry and speed control represents a new approach for DNA molecule translocation strudy with the potential for integration into the next-generation ultra-rapid DNA sequencing platform. We still have a long way to go before we can create highly reproducible CNT-based nanofluidic devices with great translocation control for ultra-rapid DNA sequencing platform, but the first steps are being made.

3.8 Future Work

3.8.1 End Modification of CNT and CNT-based DNA Sequencing Device Fabrication

Figure 46 shows one of the proposed CNT-based DNA sequencing readout systems. Once the physical picture of DNA translocation through CNTs is clear, we might use this platform for fine DNA translocation control. Recognition tunneling technique could also be integrated into this system by chemical modification of the CNT

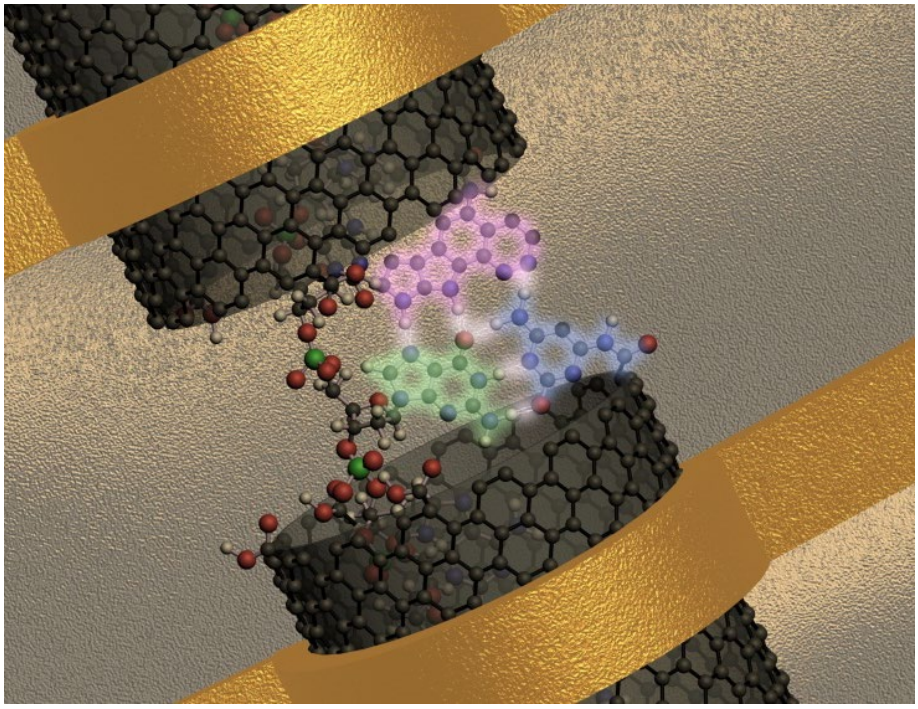


Figure 46. Proposed CNT DNA Sequencing Scheme

ends with recognition molecules via EDC/Sulfo-NHS chemistry[80,81]. Cutting the CNTs with great precision and accuracy might be difficult, and fabrication of the device itself reproducibly is a challenge too. Clearly, more experiments are waiting in front of us.

3.8.2 Using Graphene for DNA Sequencing

Discovery of graphene is announced in 2004 by the journal Science. It is a form of carbon that exists as a sheet, one atom thick. Atoms are arranged into a two-dimensional sp^2 bonding honeycomb structure. Graphene has the most remarkable

properties: it is the most conductive material in the world; it is 200 times stronger than steel, yet it is still bendable, flexible and conducts electricity better than copper; it only absorbs 2.7% of light so it is almost fully transparent; about 1% of graphene mixed into plastics could make them conductive. It is now touted as a possible replacement for silicon in electronics. Flexible touchscreens, lighting within walls, enhanced batteries, highly efficient solar cells[82], water desalination membrane[83], transparent electrodes[84] and supercapacitors[85] are among the likely first graphene based applications. Synthesis of water soluble graphene[86] and CVD process of growing large size, few-layer graphene films on arbitrary substrates[87] are reported.

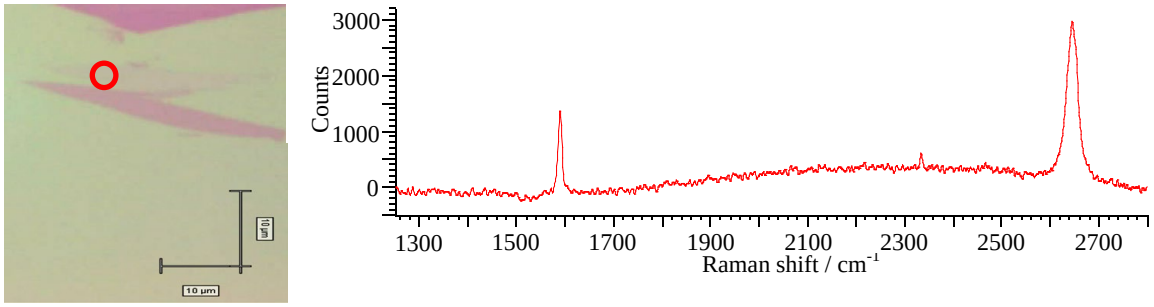


Figure 47. Exfoliated Graphene and Raman Spectrum

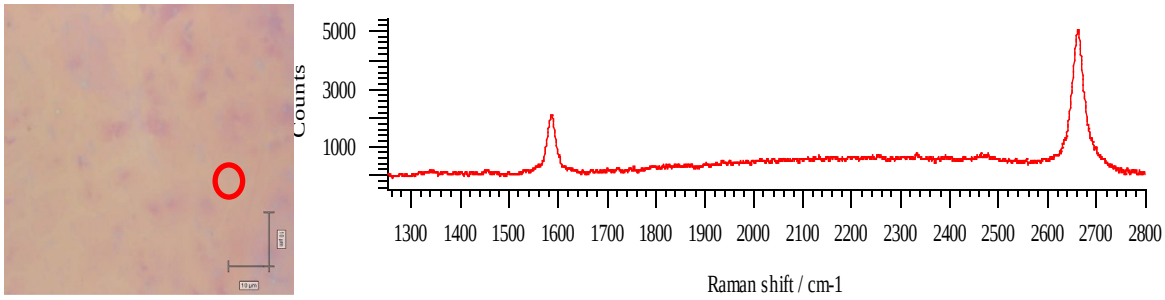


Figure 48. CVD Grown Graphene and Raman Spectrum

I was very lucky to be involved in a graphene synthesis research project during my PhD work. Exfoliated or CVD grown single layer graphene (using Cu foil as the catalyst[88]) were successfully demonstrated and confirmed by Raman spectroscopy[89],

shown in Figure 47 and Figure 48. After growth the copper can be etched away and the remaining graphene film can be easily transferred to an arbitrary surface like PDMS (Figure 49) for device fabrication.[90]

Experiments demonstrating DNA translocation through graphene nanopores were reported.[91] And graphene was shown to be a highly conductive subnanometer trans-electrode membrane.[92] A proposed graphene DNA sequencing platform utilizes these two features of graphene and integrates recognition tunneling base readout technique.

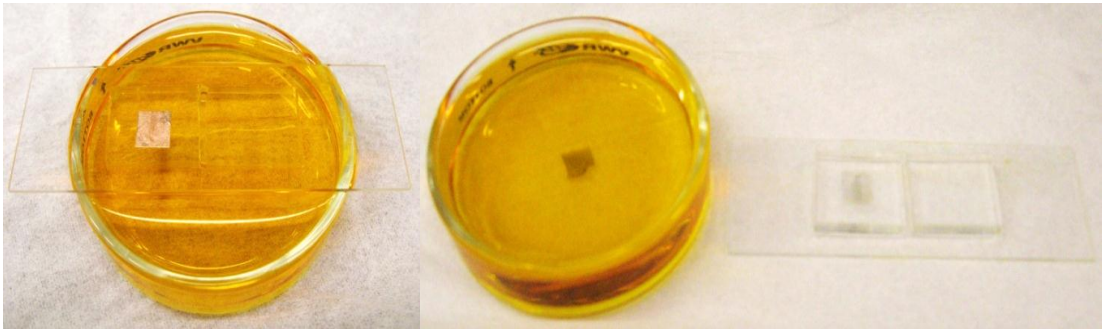


Figure 49. Transfer Graphene to PDMS

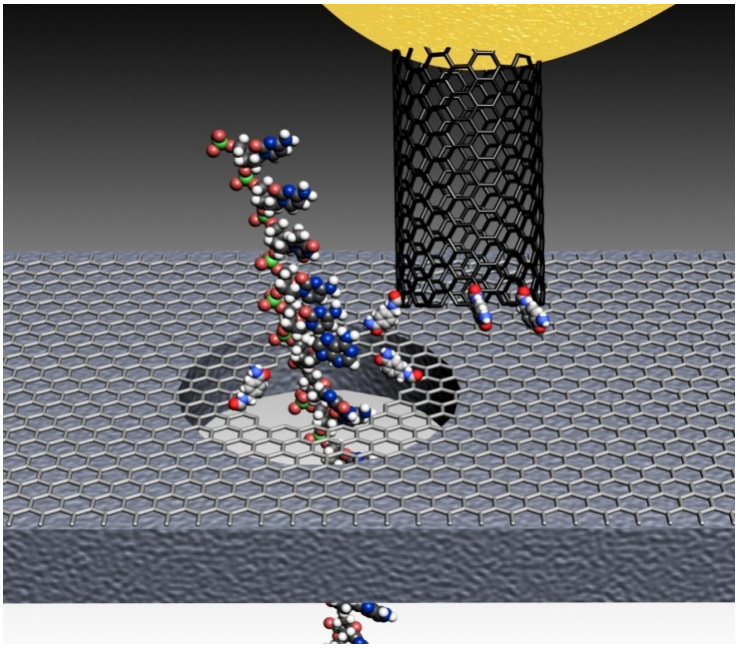


Figure 50. Proposed Graphene DNA Sequencing Scheme

CHAPTER 4 BEYOND SEQUENCING: THE MARRIAGE OF BIOCHEMISTRYAND SEMICONDUCTOR INDUSTRY FOR FUTURE BIOMEDICAL DEVICES

In the late nineteen sixties, K.D. Wise and his colleagues at Stanford University introduced silicon as a substrate for gold microelectrodes designed to measure the neuron electric potential, and opened fascinating new opportunities in medical biomedicine.[93]

Today, the fabrication techniques of microelectronics are leveraged to overcome the limitations on the precision and reproducibility of traditional microelectrodes fabrication methods. The combination of semiconductor fabrication and biochemical sensor usher in the new era of biomedical devices. Once the right coupling between highly specific biochemical sensors and electronic devices is found, complex and massively parallel biochemical signal can be converted into digital electronic data directly. Chemically specific ISFET capable of selectively detecting pyrophosphate ion as a signal of nucleotide addition has been made by immobilizing a synthesized chelator to the surface of a silicon-based field-effect transistor(FET) sensor.[94] ISFET arrays capable of sensing kinase activity in order to find a better kinase inhibitor drug has been proposed.[3] Other silicon based label-free biosensor array platforms with high-sensitivity and low-cost for protein sensing have also been developed.[95] Clearly, its applications will go far beyond DNA sequencing.

In conclusion, future biomedical devices combining highly robust recognition biochemistry and large-scale semiconductor fabrication techniques could revolutionize biomedical diagnosis, and make personalized medicine truly accessible to common consumers.

REFERENCES

1. DNA chemical structure. http://en.wikipedia.org/wiki/File:DNA_chemical_structure.svg (accessed 4/5/2013).

2. Produect - Tehcnology: 454 Life Sciences, a Roche Company. http://454.com/products/technology.asp (accessed 4/5/2013).

3. Lindsay, S., Biochemistry and semiconductor electronics—the next big hit for silicon? *Journal of Physics: Condensed Matter* **2012,** *24* (16), 164201.

4. Metzker, M. L., Sequencing technologies [mdash] the next generation. *Nat Rev Genet* **2010,** *11* (1), 31-46.

5. Vericat, C.; Vela, M. E.; Benitez, G.; Carro, P.; Salvarezza, R. C., Self-assembled monolayers of thiols and dithiols on gold: new challenges for a well-known system. *Chemical Society Reviews* **2010,** *39* (5), 1805-1834.

6. Tucker, T.; Marra, M.; Friedman, J. M., Massively Parallel Sequencing: The Next Big Thing in Genetic Medicine. *The American Journal of Human Genetics* **2009,** *85* (2), 142-154.

7. Angyal, S. J., The Composition and Conformation of Sugars in Solution. *Angewandte Chemie International Edition in English* **1969,** *8* (3), 157-166.

8. Li, Y.; Kim, W.; Zhang, Y.; Rolandi, M.; Wang, D.; Dai, H., Growth of Single-Walled Carbon Nanotubes from Discrete Catalytic Nanoparticles of Various Sizes. *The Journal of Physical Chemistry B* **2001,** *105* (46), 11424-11431.

9. Chang, S.; Huang, S.; He, J.; Liang, F.; Zhang, P.; Li, S.; Chen, X.; Sankey, O.; Lindsay, S., Electronic Signatures of all Four DNA Nucleosides in a Tunneling Gap. *Nano Letters* **2010,** *10* (3), 1070-1075.

10. Huang, S.; He, J.; Chang, S.; Zhang, P.; Liang, F.; Li, S.; Tuchband, M.; Fuhrmann, A.; Ros, R.; Lindsay, S., Identifying single bases in a DNA oligomer with electron tunnelling. *Nat Nano* **2010,** *5* (12), 868-873.

11. Shuai, C.; Shuo, H.; Hao, L.; Peiming, Z.; Feng, L.; Rena, A.; Shengqin, L.; Brett, G.; John, S.; Brian, A.; Jin, H.; Stuart, L., Chemical recognition and binding kinetics in a functionalized tunnel junction. *Nanotechnology* **2012,** *23* (23), 235101.

12. Collins, F. S.; Morgan, M.; Patrinos, A., The Human Genome Project: Lessons from Large-Scale Biology. *Science* **2003,** *300* (5617), 286-290.

13. Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D., Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Research* **2006,** *34* (2), 564-574.

14. Heid, C. A.; Stevens, J.; Livak, K. J.; Williams, P. M., Real time quantitative PCR. *Genome Research* **1996,** *6* (10), 986-994.

15. Richmond, T. J.; Davey, C. A., The structure of DNA in the nucleosome core. *Nature* **2003,** *423* (6936), 145-150.

16. Dalibard., D. C. S. C. R. M. D. T. J., Multiple Aspects of DNA and RNA: from Biophysics to Bioinformatics. **2005**.

17. Sanger, F.; Air, G. M.; Barrell, B. G.; Brown, N. L.; Coulson, A. R.; Fiddes, J. C.; Hutchison, C. A.; Slocombe, P. M.; Smith, M., Nucleotide sequence of bacteriophage [phi]X174 DNA. *Nature* **1977,** *265* (5596), 687-695.

18. Gocayne, J.; Robinson, D. A.; FitzGerald, M. G.; Chung, F. Z.; Kerlavage, A. R.; Lentes, K. U.; Lai, J.; Wang, C. D.; Fraser, C. M.; Venter, J. C., Primary structure of rat cardiac beta-adrenergic and muscarinic cholinergic receptors obtained by automated DNA sequence analysis: further evidence for a multigene family. *Proceedings of the National Academy of Sciences* **1987,** *84* (23), 8296-8300.

19. Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Francesco, V. D.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R.-R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z. Y.; Wang, A.; Wang, X.; Wang, J.; Wei, M.-H.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S. C.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.;

Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y.-H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N. N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigo, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y.-H.; Coyne, M.; Dahlke, C.; Mays, A. D.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X., The Sequence of the Human Genome. *Science* **2001,** *291* (5507), 1304-1351.

20. Performance and Specifications for HiSeq 2500/1500. http://www.illumina.com/systems/hiseq_2500_1500/performance_specifications.ilmn.

21. Press Releases | Genia Technologies. http://www.geniachip.com/about/press-releases/ (accessed 4/5/2013).

22. Merriman, B.; D Team, I. T.; Rothberg, J. M., Progress in Ion Torrent semiconductor chip based sequencing. *ELECTROPHORESIS* **2012,** *33* (23), 3397-3417.

23. MiSeq Support - Questions & Answers. http://support.illumina.com/sequencing/sequencing_instruments/miseq/questions.ilmn (accessed 4/5/2013).

24. Fedurco, M.; Romieu, A.; Williams, S.; Lawrence, I.; Turcatti, G., BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Research* **2006,** *34* (3), e22.

25. Bentley, D. R.; Balasubramanian, S.; Swerdlow, H. P.; Smith, G. P.; Milton, J.; Brown, C. G.; Hall, K. P.; Evers, D. J.; Barnes, C. L.; Bignell, H. R.; Boutell, J. M.; Bryant, J.; Carter, R. J.; Keira Cheetham, R.; Cox, A. J.; Ellis, D. J.; Flatbush, M. R.; Gormley, N. A.; Humphray, S. J.; Irving, L. J.; Karbelashvili, M. S.; Kirk, S. M.; Li, H.; Liu, X.; Maisinger, K. S.; Murray, L. J.; Obradovic, B.; Ost, T.; Parkinson, M. L.; Pratt, M. R.; Rasolonjatovo, I. M. J.; Reed, M. T.; Rigatti, R.; Rodighiero, C.; Ross, M. T.; Sabot, A.; Sankar, S. V.; Scally, A.; Schroth, G. P.; Smith, M. E.; Smith, V. P.; Spiridou, A.; Torrance, P. E.; Tzonev, S. S.; Vermaas, E. H.; Walter, K.; Wu, X.; Zhang, L.; Alam,

M. D.; Anastasi, C.; Aniebo, I. C.; Bailey, D. M. D.; Bancarz, I. R.; Banerjee, S.; Barbour, S. G.; Baybayan, P. A.; Benoit, V. A.; Benson, K. F.; Bevis, C.; Black, P. J.; Boodhun, A.; Brennan, J. S.; Bridgham, J. A.; Brown, R. C.; Brown, A. A.; Buermann, D. H.; Bundu, A. A.; Burrows, J. C.; Carter, N. P.; Castillo, N.; Chiara E. Catenazzi, M.; Chang, S.; Neil Cooley, R.; Crake, N. R.; Dada, O. O.; Diakoumakos, K. D.; Dominguez-Fernandez, B.; Earnshaw, D. J.; Egbujor, U. C.; Elmore, D. W.; Etchin, S. S.; Ewan, M. R.; Fedurco, M.; Fraser, L. J.; Fuentes Fajardo, K. V.; Scott Furey, W.; George, D.; Gietzen, K. J.; Goddard, C. P.; Golda, G. S.; Granieri, P. A.; Green, D. E.; Gustafson, D. L.; Hansen, N. F.; Harnish, K.; Haudenschild, C. D.; Heyer, N. I.; Hims, M. M.; Ho, J. T.; Horgan, A. M.; Hoschler, K.; Hurwitz, S.; Ivanov, D. V.; Johnson, M. Q.; James, T.; Huw Jones, T. A.; Kang, G.-D.; Kerelska, T. H.; Kersey, A. D.; Khrebtukova, I.; Kindwall, A. P.; Kingsbury, Z.; Kokko-Gonzales, P. I.; Kumar, A.; Laurent, M. A.; Lawley, C. T.; Lee, S. E.; Lee, X.; Liao, A. K.; Loch, J. A.; Lok, M.; Luo, S.; Mammen, R. M.; Martin, J. W.; McCauley, P. G.; McNitt, P.; Mehta, P.; Moon, K. W.; Mullens, J. W.; Newington, T.; Ning, Z.; Ling Ng, B.; Novo, S. M.; O/'Neill, M. J.; Osborne, M. A.; Osnowski, A.; Ostadan, O.; Paraschos, L. L.; Pickering, L.; Pike, A. C.; Pike, A. C.; Chris Pinkard, D.; Pliskin, D. P.; Podhasky, J.; Quijano, V. J.; Raczy, C.; Rae, V. H.; Rawlings, S. R.; Chiva Rodriguez, A.; Roe, P. M.; Rogers, J.; Rogert Bacigalupo, M. C.; Romanov, N.; Romieu, A.; Roth, R. K.; Rourke, N. J.; Ruediger, S. T.; Rusman, E.; Sanches-Kuiper, R. M.; Schenker, M. R.; Seoane, J. M.; Shaw, R. J.; Shiver, M. K.; Short, S. W.; Sizto, N. L.; Sluis, J. P.; Smith, M. A.; Ernest Sohna Sohna, J.; Spence, E. J.; Stevens, K.; Sutton, N.; Szajkowski, L.; Tregidgo, C. L.; Turcatti, G.; vandeVondele, S.; Verhovsky, Y.; Virk, S. M.; Wakelin, S.; Walcott, G. C.; Wang, J.; Worsley, G. J.; Yan, J.; Yau, L.; Zuerlein, M.; Rogers, J.; Mullikin, J. C.; Hurles, M. E.; McCooke, N. J.; West, J. S.; Oaks, F. L.; Lundberg, P. L.; Klenerman, D.; Durbin, R.; Smith, A. J., Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **2008,** *456* (7218), 53-59.

26. ISFET - Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/ISFET (accessed 4/5/2013).

27. Yates, D. E.; Levine, S.; Healy, T. W., Site-binding model of the electrical double layer at the oxide/water interface. *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* **1974,** *70* (0), 1807-1818.

28. The Ion Torrent Proton compared to the HiSeq 2500. http://www.yuzuki.org/the-ion-torrent-proton-compared-to-the-illumina-hiseq-2500/ (accessed 4/5/2013).

29. Stoddart, D.; Heron, A. J.; Mikhailova, E.; Maglia, G.; Bayley, H., Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proceedings of the National Academy of Sciences* **2009,** *106* (19), 7702-7707.

30. Manrao, E. A.; Derrington, I. M.; Laszlo, A. H.; Langford, K. W.; Hopper, M. K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H., Reading DNA at single-

nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat Biotech* **2012,** *30* (4), 349-353.

31. Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X.; Jovanovich, S. B.; Krstic, P. S.; Lindsay, S.; Ling, X. S.; Mastrangelo, C. H.; Meller, A.; Oliver, J. S.; Pershin, Y. V.; Ramsey, J. M.; Riehn, R.; Soni, G. V.; Tabard-Cossa, V.; Wanunu, M.; Wiggin, M.; Schloss, J. A., The potential and challenges of nanopore sequencing. *Nat Biotech* **2008,** *26* (10), 1146-1153.

32. Hugenholtz, P.; Tyson, G. W., Microbiology: Metagenomics. *Nature* **2008,** *455* (7212), 481-483.

33. Schuster, S. C., Next-generation sequencing transforms today's biology. *Nat Meth* **2008,** *5* (1), 16-18.

34. Jones, S.; Hruban, R. H.; Kamiyama, M.; Borges, M.; Zhang, X.; Parsons, D. W.; Lin, J. C.-H.; Palmisano, E.; Brune, K.; Jaffee, E. M.; Iacobuzio-Donahue, C. A.; Maitra, A.; Parmigiani, G.; Kern, S. E.; Velculescu, V. E.; Kinzler, K. W.; Vogelstein, B.; Eshleman, J. R.; Goggins, M.; Klein, A. P., Exomic Sequencing Identifies PALB2 as a Pancreatic Cancer Susceptibility Gene. *Science* **2009,** *324* (5924), 217.

35. Robinson, J. T.; Thorvaldsdottir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov, J. P., Integrative genomics viewer. *Nat Biotech* **2011,** *29* (1), 24-26.

36. Should Healthy People Have Their Genomes Sequenced? - WSJ.com. http://online.wsj.com/article/SB10000872396390443884104577645783975993656.html?goback=%2Egde_1902623_member_214610947 (accessed 3/5/2013).

37. (a) IVF Clinic Deploys Ion Torrent Sequencing in Embryo Screening - Bio-IT World. http://www.bio-itworld.com/2013/2/22/ivf-clinic-deploys-ion-torrent-sequencing-embryo-screening.html (accessed 4/1/2013); (b) Saunders, C. J.; Miller, N. A.; Soden, S. E.; Dinwiddie, D. L.; Noll, A.; Alnadi, N. A.; Andraws, N.; Patterson, M. L.; Krivohlavek, L. A.; Fellis, J.; Humphray, S.; Saffrey, P.; Kingsbury, Z.; Weir, J. C.; Betley, J.; Grocock, R. J.; Margulies, E. H.; Farrow, E. G.; Artman, M.; Safina, N. P.; Petrikin, J. E.; Hall, K. P.; Kingsmore, S. F., Rapid Whole-Genome Sequencing for Genetic Disease Diagnosis in Neonatal Intensive Care Units. *Science Translational Medicine* **2012,** *4* (154), 154ra135.

38. P&G Beauty & Grooming | Genomics Research at P&G Beauty. http://www.pgbeautygroomingscience.com/genomics-research-at-pg-beauty.php (accessed 3/5/2013).

39. Lloyd, K. G.; Schreiber, L.; Petersen, D. G.; Kjeldsen, K. U.; Lever, M. A.; Steen, A. D.; Stepanauskas, R.; Richter, M.; Kleindienst, S.; Lenk, S.; Schramm, A.; Jorgensen, B. B., Predominant archaea in marine sediments degrade detrital proteins. *Nature* **2013,** *advance online publication*.

40. Ling, H.-Q.; Zhao, S.; Liu, D.; Wang, J.; Sun, H.; Zhang, C.; Fan, H.; Li, D.; Dong, L.; Tao, Y.; Gao, C.; Wu, H.; Li, Y.; Cui, Y.; Guo, X.; Zheng, S.; Wang, B.; Yu, K.; Liang, Q.; Yang, W.; Lou, X.; Chen, J.; Feng, M.; Jian, J.; Zhang, X.; Luo, G.; Jiang, Y.; Liu, J.; Wang, Z.; Sha, Y.; Zhang, B.; Wu, H.; Tang, D.; Shen, Q.; Xue, P.; Zou, S.; Wang, X.; Liu, X.; Wang, F.; Yang, Y.; An, X.; Dong, Z.; Zhang, K.; Zhang, X.; Luo, M.-C.; Dvorak, J.; Tong, Y.; Wang, J.; Yang, H.; Li, Z.; Wang, D.; Zhang, A.; Wang, J., Draft genome of the wheat A-genome progenitor Triticum urartu. *Nature* **2013,** *advance online publication*.

41. Chin, C.-S.; Sorenson, J.; Harris, J. B.; Robins, W. P.; Charles, R. C.; Jean-Charles, R. R.; Bullard, J.; Webster, D. R.; Kasarskis, A.; Peluso, P.; Paxinos, E. E.; Yamaichi, Y.; Calderwood, S. B.; Mekalanos, J. J.; Schadt, E. E.; Waldor, M. K., The Origin of the Haitian Cholera Outbreak Strain. *New England Journal of Medicine* **2011,** *364* (1), 33-42.

42. Goldman, N.; Bertone, P.; Chen, S.; Dessimoz, C.; LeProust, E. M.; Sipos, B.; Birney, E., Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **2013,** *advance online publication*.

43. Zwolak, M.; Di Ventra, M., Colloquium: Physical approaches to DNA sequencing and detection. *Reviews of Modern Physics* **2008,** *80* (1), 141-165.

44. Lindsay, S., *Introduction to Nanoscience*. OUP Oxford: 2010.

45. Shuai, C.; Suman, S.; Peiming, Z.; Brett, G.; Brian, A.; Steven, L.; Hongbo, P.; Stuart, L., Palladium electrodes for molecular tunnel junctions. *Nanotechnology* **2012,** *23* (42), 425202.

46. Lindsay, S.; Jin, H.; Otto, S.; Prokop, H.; Pavel, J.; Peiming, Z.; Shuai, C.; Shuo, H., Recognition tunneling. *Nanotechnology* **2010,** *21* (26), 262001.

47. Liang, F.; Lindsay, S.; Zhang, P., 1,8-Naphthyridine-2,7-diamine: a potential universal reader of Watson-Crick base pairs for DNA sequencing by electron tunneling. *Organic & Biomolecular Chemistry* **2012,** *10* (43), 8654-8659.

48. Liang, F.; Li, S.; Lindsay, S.; Zhang, P., Synthesis, Physicochemical Properties, and Hydrogen Bonding of 4(5)-Substituted 1-H-Imidazole-2-carboxamide, a Potential Universal Reader for DNA Sequencing by Recognition Tunneling. *Chemistry – A European Journal* **2012,** *18* (19), 5998-6007.

49. von Wrochem, F.; Gao, D.; Scholz, F.; Nothofer, H.-G.; Nelles, G.; Wessels, J. M., Efficient electronic coupling and improved stability with dithiocarbamate-based molecular junctions. *Nat Nano* **2010,** *5* (8), 618-624.

50. Love, J. C.; Estroff, L. A.; Kriebel, J. K.; Nuzzo, R. G.; Whitesides, G. M., Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. *Chemical Reviews* **2005,** *105* (4), 1103-1170.

51. Tavendale, A. J.; Pearton, S. J., Deep level, quenched-in defects in silicon doped with gold, silver, iron, copper or nickel. *Journal of Physics C: Solid State Physics* **1983,** *16* (9), 1665.

52. Love, J. C.; Wolfe, D. B.; Chabinyc, M. L.; Paul, K. E.; Whitesides, G. M., Self-Assembled Monolayers of Alkanethiolates on Palladium Are Good Etch Resists. *Journal of the American Chemical Society* **2002,** *124* (8), 1576-1577.

53. Hydrogen Flame Annealing | Agilent http://www.home.agilent.com/agilent/editorial.jspx?cc=US&lc=eng&ckey=914989&nid=-33980.0&id=914989 (accessed 3/5/2013).

54. Morf, P.; Ballav, N.; Putero, M.; von Wrochem, F.; Wessels, J. M.; Jung, T. A., Supramolecular Structures and Chirality in Dithiocarbamate Self-Assembled Monolayers on Au(111). *The Journal of Physical Chemistry Letters* **2010,** *1* (5), 813-816.

55. XPS Spectra. http://www.casaxps.com/help_manual/manual_updates/xps_spectra.pdf (accessed 3/6/2013).

56. FTIR User Guide. http://mmrc.caltech.edu/FTIR/Nicolet/Nicolet%20Software/Nicolet%202/4700_6700_User.pdf (accessed 2/7/2013).

57. Tompkins, H. G.; Irene, E. A., *Handbook of ellipsometry*. William Andrew Pub.: 2005.

58. Tuchband, M.; He, J.; Huang, S.; Lindsay, S., Insulated gold scanning tunneling microscopy probes for recognition tunneling in an aqueous environment. *Review of Scientific Instruments* **2012,** *83* (1), 015102-4.

59. Chang, S.; He, J.; Zhang, P.; Gyarfas, B.; Lindsay, S., Gap Distance and Interactions in a Molecular Tunnel Junction. *Journal of the American Chemical Society* **2011,** *133* (36), 14267-14269.

60. Chang, C.-C.; Lin, C.-J., LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011,** *2* (3), 1-27.

61. Fuhrmann, A.; Getfert, S.; Fu, Q.; Reimann, P.; Lindsay, S.; Ros, R., Long Lifetime of Hydrogen-Bonded DNA Basepairs by Force Spectroscopy. *Biophys. J.* **2012,** *102* (10), 2381-2390.

62. Dekker, C., Solid-state nanopores. *Nat Nano* **2007,** *2* (4), 209-215.

63. Austin, R., Nanopores: The art of sucking spaghetti. *Nat Mater* **2003,** *2* (9), 567-568.

64. Liang, X.; Chou, S. Y., Nanogap Detector Inside Nanofluidic Channel for Fast Real-Time Label-Free DNA Analysis. *Nano Letters* **2008,** *8* (5), 1472-1476.

65. Fan, R.; Karnik, R.; Yue, M.; Li, D. Y.; Majumdar, A.; Yang, P. D., DNA translocation in inorganic nanotubes. *Nano Letters* **2005,** *5* (9), 1633-1637.

66. Zhu, F. Q.; Schulten, K., Water and proton conduction through carbon nanotubes as models for biological channels. *Biophys. J.* **2003,** *85* (1), 236-244.

67. Joseph, S.; Mashl, R. J.; Jakobsson, E.; Aluru, N. R., Electrolytic Transport in Modified Carbon Nanotubes. *Nano Lett.* **2003,** *3* (10), 1399-1403.

68. Yeh, I.-C.; Hummer, G., Nucleic acid transport through carbon nanotube membranes. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (33), 12177-12182.

69. Kong, J.; Soh, H. T.; Cassell, A. M.; Quate, C. F.; Dai, H., Synthesis of individual single-walled carbon nanotubes on patterned silicon wafers. *Nature* **1998,** *395* (6705), 878-881.

70. Sreekar, B.; Alfonso, R.; Jifa, Q.; Jing, K.; Angela, M. B., Block-copolymer assisted synthesis of arrays of metal nanoparticles and their catalytic activities for the growth of SWNTs. *Nanotechnology* **2006,** *17* (20), 5080.

71. Huang, L.; Cui, X.; White, B.; O'Brien, S. P., Long and Oriented Single-Walled Carbon Nanotubes Grown by Ethanol Chemical Vapor Deposition. *The Journal of Physical Chemistry B* **2004,** *108* (42), 16451-16456.

72. Cheung, C. L.; Kurtz, A.; Park, H.; Lieber, C. M., Diameter-Controlled Synthesis of Carbon Nanotubes. *The Journal of Physical Chemistry B* **2002,** *106* (10), 2429-2433.

73. Guo, X.; Gorodetsky, A. A.; Hone, J.; Barton, J. K.; Nuckolls, C., 1 - Conductivity of a single DNA duplex bridging a carbon nanotube gap. *Nat Nano* **2008,** *3* (3), 163-167.

74. Whitesides, G. M., The origins and the future of microfluidics. *Nature* **2006,** *442* (7101), 368-373.

75. Zuker, M., Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **2003,** *31* (13), 3406-3415.

76. Jae Hyun, P.; Jin, H.; Brett, G.; Stuart, L.; Predrag, S. K., DNA translocating through a carbon nanotube can increase ionic current. *Nanotechnology* **2012,** *23* (45), 455107.

77. Pang, P.; He, J.; Park, J. H.; Krstić, P. S.; Lindsay, S., Origin of Giant Ionic Currents in Carbon Nanotube Channels. *ACS Nano* **2011,** *5* (9), 7277-7283.

78. Ririe, K. M.; Rasmussen, R. P.; Wittwer, C. T., Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. *Analytical Biochemistry* **1997,** *245* (2), 154-160.

79. Song, W.; Pang, P.; He, J.; Lindsay, S., Optical and Electrical Detection of Single-Molecule Translocation through Carbon Nanotubes. *ACS Nano* **2012,** *7* (1), 689-694.

80. Banerjee, S.; Hemraj-Benny, T.; Wong, S. S., Covalent Surface Chemistry of Single-Walled Carbon Nanotubes. *Advanced Materials* **2005,** *17* (1), 17-29.

81. Hirsch, A.; Vostrowsky, O., Functionalization of Carbon Nanotubes. In *Functional Molecular Nanostructures*, Schlüter, A. D., Ed. Springer Berlin Heidelberg: 2005; Vol. 245, pp 193-237.

82. Miao, X.; Tongay, S.; Petterson, M. K.; Berke, K.; Rinzler, A. G.; Appleton, B. R.; Hebard, A. F., High Efficiency Graphene Solar Cells by Chemical Doping. *Nano Letters* **2012,** *12* (6), 2745-2750.

83. Cohen-Tanugi, D.; Grossman, J. C., Water Desalination across Nanoporous Graphene. *Nano Letters* **2012,** *12* (7), 3602-3608.

84. Wang, X.; Zhi, L.; Mullen, K., Transparent, Conductive Graphene Electrodes for Dye-Sensitized Solar Cells. *Nano Letters* **2007,** *8* (1), 323-327.

85. Liu, C.; Yu, Z.; Neff, D.; Zhamu, A.; Jang, B. Z., Graphene-Based Supercapacitor with an Ultrahigh Energy Density. *Nano Letters* **2010,** *10* (12), 4863-4868.

86. Si, Y.; Samulski, E. T., Synthesis of Water Soluble Graphene. *Nano Letters* **2008,** *8* (6), 1679-1682.

87. Reina, A.; Jia, X.; Ho, J.; Nezich, D.; Son, H.; Bulovic, V.; Dresselhaus, M. S.; Kong, J., Large Area, Few-Layer Graphene Films on Arbitrary Substrates by Chemical Vapor Deposition. *Nano Letters* **2008,** *9* (1), 30-35.

88. Li, X.; Cai, W.; An, J.; Kim, S.; Nah, J.; Yang, D.; Piner, R.; Velamakanni, A.; Jung, I.; Tutuc, E.; Banerjee, S. K.; Colombo, L.; Ruoff, R. S., Large-Area Synthesis of High-Quality and Uniform Graphene Films on Copper Foils. *Science* **2009**, 1171245.

89. Ferrari, A. C., Raman spectroscopy of graphene and graphite: Disorder, electron-phonon coupling, doping and nonadiabatic effects. *Solid State Communications* **2007,** *143* (1-2), 47-57.

90. Lee, Y.; Bae, S.; Jang, H.; Jang, S.; Zhu, S.-E.; Sim, S. H.; Song, Y. I.; Hong, B. H.; Ahn, J.-H., Wafer-Scale Synthesis and Transfer of Graphene Films. *Nano Letters* **2010,** *10* (2), 490-493.

91. Merchant, C. A.; Healy, K.; Wanunu, M.; Ray, V.; Peterman, N.; Bartel, J.; Fischbein, M. D.; Venta, K.; Luo, Z.; Johnson, A. T. C.; Drndić, M., DNA Translocation through Graphene Nanopores. *Nano Letters* **2010,** *10* (8), 2915-2921.

92. Garaj, S.; Hubbard, W.; Reina, A.; Kong, J.; Branton, D.; Golovchenko, J. A., Graphene as a subnanometre trans-electrode membrane. *Nature* **2010,** *467* (7312), 190-193.

93. Wise, K. D.; Angell, J. B.; Starr, A., An Integrated-Circuit Approach to Extracellular Microelectrodes. *Biomedical Engineering, IEEE Transactions on* **1970,** *BME-17* (3), 238-247.

94. Liu, D. J.; Credo, G. M.; Su, X.; Wu, K.; Lim, H. C.; Elibol, O. H.; Bashir, R.; Varma, M., Surface immobilizable chelator for label-free electrical detection of pyrophosphate. *Chemical Communications* **2011,** *47* (29), 8310-8312.

95. Iqbal, M.; Gleeson, M. A.; Spaugh, B.; Tybor, F.; Gunn, W. G.; Hochberg, M.; Baehr-Jones, T.; Bailey, R. C.; Gunn, L. C., Label-Free Biosensor Arrays Based on Silicon Ring Resonators and High-Speed Optical Scanning Instrumentation. *Selected Topics in Quantum Electronics, IEEE Journal of* **2010,** *16* (3), 654-661.