

Recognition Tunneling:  
Approaches towards Next Generation DNA Sequencing

by

Shuo Huang

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

Approved April 2011 by the  
Graduate Supervisory Committee:

Stuart Lindsay, Chair  
Otto Sankey  
Robert Ros  
Jeffery Drucker  
Nongjian Tao

ARIZONA STATE UNIVERSITY

May 2011

## ABSTRACT

This thesis describes several approaches to next generation DNA sequencing via tunneling current method based on a Scanning Tunneling Microscope system. In chapters 5 and 6, preliminary results have shown that DNA bases could be identified by their characteristic tunneling signals. Measurements taken in aqueous buffered solution showed that single base resolution could be achieved with economic setups. In chapter 7, it is illustrated that some ongoing measurements are indicating the sequence readout by making linear scan on a piece of short DNA oligomer. However, to overcome the difficulties of controlling DNA especially ssDNA movement, it is much better to have the tunneling measurement incorporated onto a robust nanopore device to realize sequential reading of the DNA sequence while it is being translocated.

## DEDICATION

For my beloved parents and grandparents

## ACKNOWLEDGMENTS

I still recall the moment when I first arrived in Arizona in 2006. The memory is so fresh that makes me believe that it was happening yesterday. And now I am getting my degree and ready to leave.

First of all, I want to thank my research advisor Dr. Stuart Lindsay for his support and trust during all my 5 years of graduate study. At the moment I first talked with Stuart about joining the group, I was already fascinated by his talk and the future development of the sequencing field. I feel so lucky that I made the decision to join such a good research lab which gave me so much fun of doing world's pioneering research here and have learned a lot for my future career.

Besides that I want to thank Maggie Black, Michael Dodson and Steve Woodward who helped me a lot for technical issues. Without them, I could be troubled everyday in the lab without any progress in scientific discoveries.

At the same time I want to thank Jin He, Feng Liang and Brett Gyarfás for the help with my research. As successful and experienced scientific researchers, they are my models and have taught me a lot not only doing science but also pursuing a successful scientific career.

Also I feel grateful to be able to work with these colleagues in the lab. Shuai Chang has always been a good partner working together and it is a great fun to work out things with Shuai who shared a lot of crazy but inspiring ideas in the research. I also feel quite happy to have the experience working with these fellows in Lindsay lab: Lisha lin, Liyun Lin, Di Cao, Pei Pang, Hao Liu, Qiang Fu, Yanan Zhao, Weisi Song, Sen Peng, Padmini Krishnakumar, Parminder Kaur, Shreya Bhattacharyya and Ashley Kibel.

These faculty members also have given me invaluable help for the research. Dr. Peiming Zhang is such an experienced and knowledgeable chemist who can make almost any chemical compound that we are looking for. Dr. Su Lin has helped me with the initial experiment setups for some optics experiments. Dr. Nongjian Tao and his group members all helped me a lot with electronics and hardware. Dr. Hao Yan and his group members helped me and taught me the first knowledge of molecular biology which is so powerful and beneficial to my ongoing research. Dr. Sankey and his group member Myeong Lee and Xiang Chen have finished all the theoretical work along with predictions for future work. I feel proud of working with such a strong theoretical group which saved me a lot of efforts and I learned a lot.

Finally I want to thank my parents, my grandparents and all the relatives living in US who helped me to overcome difficulties both in research and life. They loved me so much which sometimes makes me feel not really deserved. I feel so proud and grateful to have these relatives in the world and they made the transition of my life from China to US very smooth and happy.

I also want to thank everyone who helped me or worked with me but not mentioned here during the 5 years life in Arizona.

I feel proud of all of you. I don't deserve the help from any of you, but I really do appreciate that you did it. Thank you!

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	x
LIST OF FIGURES .....	xi
CHAPTER	
1. INTRODUCTION: DNA AND DNA SEQUENCING .....	1
1.1 Deoxyribonucleic Acid (DNA).....	1
1.1.1 The double helix structure and its chemical structure .....	1
1.1.2 The size of DNA.....	4
1.1.3 DNA mutation .....	5
1.2 DNA sequencing.....	6
1.2.1 Sanger sequencing .....	6
1.2.2 Details of Sanger sequencing .....	7
1.2.3 The drawbacks and limitations of Sanger sequencing .....	10
1.3 Nanopore sequencing.....	11
1.3.1 Protein nanopores.....	12
1.3.2 Solid state nanopores .....	14
2. MOLECULAR ELECTRONICS AND SCANNING TUNNELING	
MICROSCOPY .....	18
2.1 Molecular Electronics.....	18

CHAPTER	Page
2.1.1 Landuaer Formula.....	19
2.2 Scanning Tunneling Microscopy .....	27
2.2.1 The history of STM.....	27
2.2.2 The limitations of STM .....	32
2.3 Why STM for DNA sequencing .....	32
3. THE CONDUCTANCE MEASUREMENT OF DNA BASE ON NUCLEOSIDES .....	34
3.1 Introduction .....	34
3.2 Single molecule Conductance Measurement by STM .....	34
3.2.1 Break Junction: the I(s) method .....	34
3.2.2 Spontaneous On/Off Switching: the I(t) method.....	36
3.3 DNA base on nucleoside measurement.....	39
3.3.1 Materials and Methods.....	39
3.3.2 Results and Discussion.....	41
3.3.3 Conclusion .....	49
4. THE CONDUCTANCE MEASUREMENT OF DNA BASE ON BASE..	51
4.1 Introduction .....	51
4.2 Materials and Methods.....	52
4.3 Results and Discussion.....	63
4.4 Conclusion .....	70



CHAPTER	Page
5. THE RECOGNITION TUNNELING MEASUREMENT IN ORGANIC SOLVENT BY A UNIVERSAL READER.....	72
5.1 Introduction .....	72
5.2 Materials and Methods.....	72
5.3 Results and Discussions.....	76
5.4 Conclusion .....	81
6. THE UNIVERSAL READER IN AQUEOUS BUFFERED SOLUTION .	82
6.1 Introduction .....	82
6.2 Materials and Methods.....	83
6.3 Results and Discussions.....	90
6.4 Conclusion .....	98
7. THE ONGOING AND FUTURE RESEARCH: TO READ THE SEQUENCE OF DNA .....	100
7.1 Introduction .....	100
7.2 Immidazole Reader.....	101
7.3 DNA Sequencing by Linear Scanning.....	102
7.4 Clock Scanning Method .....	105
7.5 Nanopore Tunneling Measurement.....	109
8. CONCLUSION.....	113
REFERENCES.....	115

	Page
BIOGRAPHICAL SKETCH .....	133
APPENDIX	
A AUTOMATED DATA ANALYSIS.....	126
B TIP MAKING AND INSULATING.....	129

LIST OF TABLES

Table	Page
2.1 Single molecule conductance determined in the published paper. ....	29
3.1 Observed frequency of switching.....	42
3.2 Measured conductance compared to calculated values .....	47
4.1 A summary of the measured single molecular conductance compared with theoretical predicted values. ....	69
5.1 Measured and calculated conductance in a functionalized tunnel junction.....	76
6.1 Tunneling signal characteristics for 4 types of single nucleotides.....	92
6.2 Tunneling signal characteristics for DNA oligomers .....	93

## LIST OF FIGURES

Figure	Page
1.1 The chemical structure of DNA.....	2
1.2 The chemical structures of regular and modified DNA bases .....	4
1.3 From DNA to chromosomes. ....	5
1.4 The illustration of Sanger sequencing method.....	8
1.5 A typical gel result of sanger sequencing method. ....	9
1.6 Schematics of $\alpha$ -hemolysine nanopore nucleotide identification based on the exonuclease activity. ....	13
1.7 Sample ionic current signals of DNA nucleotide when translocating an $\alpha$ -hemolysine nanopore.....	14
1.8 Schematic diagrams of DNA translocation through an ultra thin SiN membrane nanopore. ....	16
1.9 SEM image of the translocation reservoir with Single walled carbon nanotube (CNT).....	17
2.1 STM break junction measurement.....	21
2.2 Schematic diagram of electron transport through Metal-Molecule-Metal junction.....	22
2.3 The energy levels calculated for four different DNA bases and DNA base pairing by DFT method. ....	25
2.4 The Schematic diagram of scanning tunneling microscope.....	28
2.5 The first STM image of Si 7x7 lattice structure. ....	30

Figure	Page
2.6 An STM image measured by a modern scanning tunneling microscope. Si 7x7 lattice is clearly shown.....	31
3.1 A typical signal of spontaneous ON-OFF switching during an I(t) measurement. ....	37
3.2 Histograms of 54 switching heights $I_w$ for 1,8-Octanedithiol. ....	38
3.3 Typical comparison of (a) single molecule telegraph switching and (b) multiple molecules trapped in the gap. ....	38
3.4 An intact junction in which the tunnel gap is spanned by a guanine attached to the probe .....	40
3.5 Sample time traces for DNA base on nucleoside (b-d) and control ...	43
3.6 Plot of molecular switching versus baseline conductance for the three base-nucleotide combinations .....	44
3.7 I-V curves for 2AA-thymidine (diamonds), A-thymidine (dots) and G- deoxycytidine (squares).....	45
3.8 Distribution of “on” times for G-deoxycytidine (left histogram plot). ...	46
3.9 Switching time distributions for (a) 2AA on thymidine and (b) A on thymidine. ....	48
3.10 Distribution of switching times for telegraph noise measured in 1,8- octanedithio .....	49

Figure	Page
4.1 The schematic diagram of the base on base experiment embedded in the monolayer of 1'-Octanethiol.....	52
4.2 The thymine monolayer thickness on Au(111) surface as a function of deposition time measured by ellipsometry.....	53
4.3 FTIR absorbance spectra for octanethiol (C8), mercaptothymine (Thymine) and mixed (C8/Thymine) SAMs.....	55
4.4 FTIR absorbance spectra for octanethiol (C8), mercaptocytosine (Cytosine) and mixed (C8/Cytosine) SAMs. ....	56
4.5 High resolution STM images of SAMs .....	57
4.6 Schematics of the AC modulated break junction measurement. ....	57
4.7 DC and AC components of current traces recorded during the pulling away stage: .....	58
4.8 Single molecular base-pairing conductance measurement by I(t) method .....	59
4.9 Sample time traces of single molecule conductance measurement with AC modulation.....	60
4.10 Scatter plots of molecular conductance, $G_m$ , vs the baseline conductance, $G_{baseline}$ .....	64
4.11 Scatter plots of molecular conductance, $G_m$ , vs the baseline tunnel conductance, $G_{baseline}$ .....	65

Figure	Page
4.12 Scatter plots of molecular conductance, $G_m$ , vs the baseline tunnel conductance, $G_{\text{baseline}}$ .....	66
4.13 Gaussian fits to the conductance distributions for the mixed pyrimidine SAMs .....	67
4.14 Conductance histograms for the pure pyrimidine SAMs. The solid lines are Gaussian fits to the conductance distribution.....	68
5.1 Tunneling measurements with functionalized electrodes. ....	74
5.2 Effect of electrode functionalization for pyrimidine reads.....	77
5.3 Summary of the reads. ....	78
5.4 Effect of electrode functionalization on the distribution of current spikes. .....	79
5.5 Comparison between C signal and $^{13}\text{C}$ signal.....	80
6.1 STM images of the self assembled monolayer of benzamide DNA reader on the gold surface.....	83
6.2 FTIR spectrum of 4-mercaptobenzamide monolayer (Purple) and powder (red). ....	84
6.3 Tunnel current decay curves in pure water.....	85
6.4 Histograms of beta in pure water.....	87
6.5 Tunneling signals from nucleotides trapped in a functionalized tunnel gap. ....	89

Figure	Page
6.6 Reading a single base within a heteropolymer.....	90
6.7 The judgement criteria from signal height and burst frequency. ....	92
6.8 Tunneling signal distributions from oligomers resemble those of the constituent nucleotides.....	94
6.9 The dynamic force study of the interaction between dAMP and DNA reader.....	95
6.10 A control curves taken in the absence of dAMP .....	97
7.1 Distributions of current spike heights of the imimidazole reader for 5 different DNA nucleotides.....	101
7.2 Schematic diagram of linear scanning over a piece of ssDNA .....	102
7.3 the typical sequence signal obtained by Imimidazole reader.....	103
7.4 The plot of $T_b$ vs $1/V$ . .....	104
7.5 The schematic diagram of a Clock Scanning.....	106
7.6 Typical signals during the clock scanning.....	107
7.7 DNA self assembly over a benzamide reader on the surface of Au(111) lattice .....	108
7.8 2D intensity plot from measurements over 3 different centre locations on the same chip. ....	109
7.9 A proposed design of a translocation brake in a DNA sensor.....	110
7.10 Hybrid nanopore formed by inserting $\alpha$ -hemolysine into the solid state nanopore. ....	111



Figure	Page
7.11 The schematic diagram of DNA translocation through a grapheme nanopore. ....	112

# **1. Introduction: DNA and DNA Sequencing**

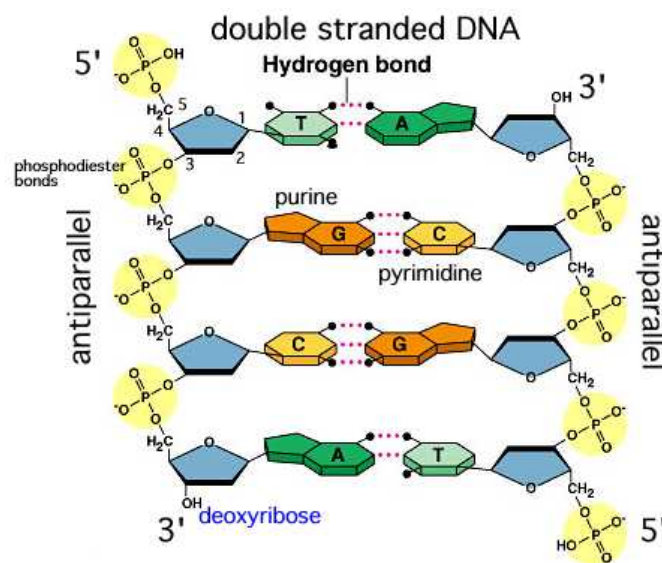
## **1.1 Deoxyribonucleic Acid (DNA)**

The earliest study of DNA can be traced back to 1869 when it was first isolated by Swiss physician Friedrich Miescher.<sup>1,2</sup> DNA is a specific type of nucleic acid that contains the genetic blueprints for the development and function of almost every living organisms on earth with the exception of some viruses which use RNA (ribonucleic acid) as their genome. For this reason, the study of DNA is vitally important for understanding a variety of biological activities.

### ***1.1.1 The double helix structure and its chemical structure***

The double helical model of DNA structure was proposed by Watson and Crick in 1953.<sup>3,4</sup> based on X-ray diffraction data. Today, the double helical structure of DNA is widely accepted and the subject of frequent study. A DNA double helical model is composed of 2 complementary strands of nucleotides with opposite directions (anti-parallel). The backbone of each single stranded DNA is composed of alternating phosphate and sugar rings linked by ester bonds. In addition to one phosphate and one sugar, each nucleotide contains one base. There are 4 types of DNA bases which are two purines (Adenine and Guanine) and two pyrimidines (Thymine and Cytosine). The bases are the coding elements of DNA, arranged in a particular sequence within a single strand. Two single

strands of DNA with complementary sequences and anti-parallel arrangement will bind together (hybridize), with the DNA base-pairs serving as recognition elements and associated by hydrogen bonding. The bases associate in a very predictable and specific way, with A binding to T and G binding to C, thus the total number of purines and pyrimidines are equal in every DNA double helix.

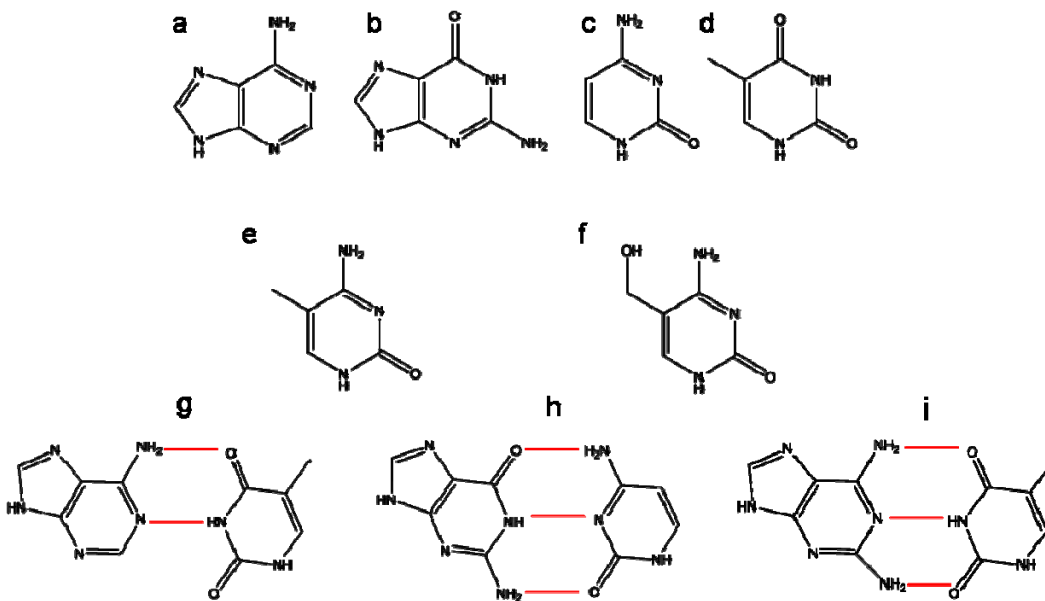


**Figure 1.1** The Chemical Structure of DNA. Each single strand is strictly complimentary in sequence and arranged in opposite direction to the other so that they may form double stranded DNA (dsDNA) via base-pairing. Adenine-Thymine and Guanine-Cytosine form base-pairs by hydrogen bonding.<sup>5</sup>

The chemical structures and the conformation of DNA base pairings are listed in **Figure 1.2**. As demonstrated in the figure, the complimentary bases form basepairs via hydrogen bonding (red bars in the figure). Besides the four regular types of DNA bases, there are several more epigenetic markers like methyl C (<sup>m</sup>C) and hydroxyl methyl C (<sup>hm</sup>C) et al as chemical derivatives of the regular bases. These epigenetic markers are

found in mammalian cells and are believed to regulate gene expression although its exact function is unknown. To understand its exact function is of great importance but challenging by the current generation sequencing method, which is based on polymerase chain reaction (PCR) method. Since the hydrogen bonding groups are exactly the same, the <sup>m</sup>C could form base pairs like its regular version and is not distinguishable from regular C then. Novel techniques that could directly lead to identifications of <sup>m</sup>C on a strand of ssDNA are illustrated in chapter 6.

Another DNA base derivative 2-amino-adenine (2AA) (see **Figure 1.2i**) is used as a control molecule for DNA base-pairing. It is similar in structure as the A but has one more amino group, which helps to form the third hydrogen bond with thymine. This molecule will be discussed in detail in chapter 3 and 4.



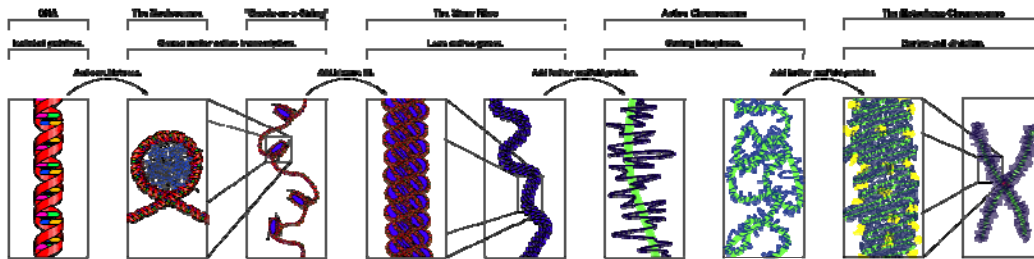
**Figure 1.2** The chemical structures of regular and modified DNA bases (a) adenine (b) guanine (c) cytosine (d) thymine (e) 5'-methyl-cytosine (f) 5-hydroxymethyl-cytosine. (g) A-T basepair. (h) G-C basepair. (i) 2AA-T basepair.

### 1.1.2 The size of DNA

The size of each nucleotide on a piece of ssDNA is about 3 Angstroms in length which is too tiny to be identified by major sensing instruments.<sup>6</sup> However this property of DNA is advantageous for its role as the carrier of the genetic code because billions of bits of information can be contained in a small volume. For example, the relatively simple DNA genome of the  $\lambda$  phage is composed of 48.5 kbp and is only about 15  $\mu\text{m}$  in length.<sup>7, 8</sup> In addition, genomic DNA in the nucleus of a cell forms supercoiled and other bent structures so that even larger amount of DNA can be contained. In most eukaryote cells, DNA is stored in the nuclear envelope as chromosomes, which are combinations of DNA, histone and other proteins.

To form chromosomes, double helix wraps around histones, forming higher order structure, “bead on a string” structure (nucleosome). Nucleosomes arrays are further condensed into 30 nm chromatin fibers to further shrink the structure and eventually packed into the metaphase of chromosomes. The actual size of a whole genome is tiny. However, the volume of coded information is enormous.

For these reasons, unraveling the massive amount of coded information is quite challenging, with consideration of speed and cost paramount to the task.



**Figure 1.3** From DNA to Chromosomes. As the genomic code carrier, DNA must be small enough to fit into the nuclear envelope. For such reason, as such DNA in living cells is compacted into the chromosomes.<sup>9</sup>

### 1.1.3 DNA mutation

As the biochemical molecule that stores genetic information in cells, DNA has to be stable both physically and chemically.<sup>10</sup> However, damage to DNA occurs by several pathways, including exposure to mutagens, oxidizing and alkylating agents and high energy EM radiation like ultra-violet and X-rays.<sup>11</sup> Although it is believed that there are certain mechanisms for DNA to repair itself and minor damage would have quite

minimum effect on the overall function of living organisms. It has also been suggested that the damage or variation of DNA is the ultimate driving force for the evolution of life.<sup>12, 13</sup> It is also true that the damage to DNA or alternation of the genome sequence could lead to severe genomic diseases such as cancer, diabetes et al.<sup>14-16</sup> At this point, it is obvious that understanding the information that is coded in DNA is essential for discovering many of nature's secrets. Thus it is critically important to develop a rapid, effective and a low-cost way to sequence DNA and obtain a better understanding of the process of life.

## **1.2 DNA sequencing**

The study of DNA sequencing can be traced back to 1970s when the 1<sup>st</sup> DNA sequencing was done by 2-dimensional chromatography. The chain termination method developed by Frederick Sanger has become the most popular method for its reliability and simplicity.<sup>17, 18</sup>

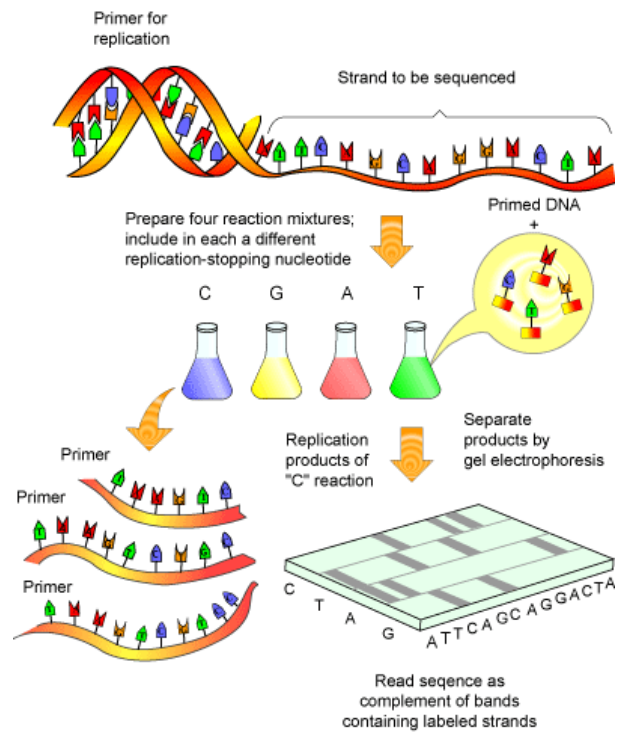
### ***1.2.1 Sanger sequencing***

The Sanger Sequencing method is based on the chain termination ability of a specialized DNA nucleotide: dideoxynucleotide (ddNTP). Unlike regular dNTP, ddNTP will terminate nucleotide elongation during the DNA polymerase reaction. The ddNTPs (ddATP, ddCTP, ddGTP, ddTTP) can also be radioactively or fluorescently labeled.

### ***1.2.2 Details of Sanger sequencing***

In the sanger sequencing method, the DNA polymerase reaction is separated into four separate reactions (see **Figure 1.4**). Each reaction contains a DNA target template, primer, a complete set of dNTPs (dATP, dTTP, dCTP, dGTP) and DNA polymerase. In addition, one of the four fluorescently labeled ddNTPs is also added to the mixture (normally each ddNTP will be labeled with a different color fluorophore). After tens of thermal cycles the DNA target template is amplified by the polymerase reaction. However, the ddNTP arrests the polymerase chain reaction and the newly added ddNTP will occupy the terminal strand position. Thus, the fluorescent label indicates the last nucleotide of the DNA strand and the size of the corresponding DNA indicates the relative position of this labeled base within the strand.

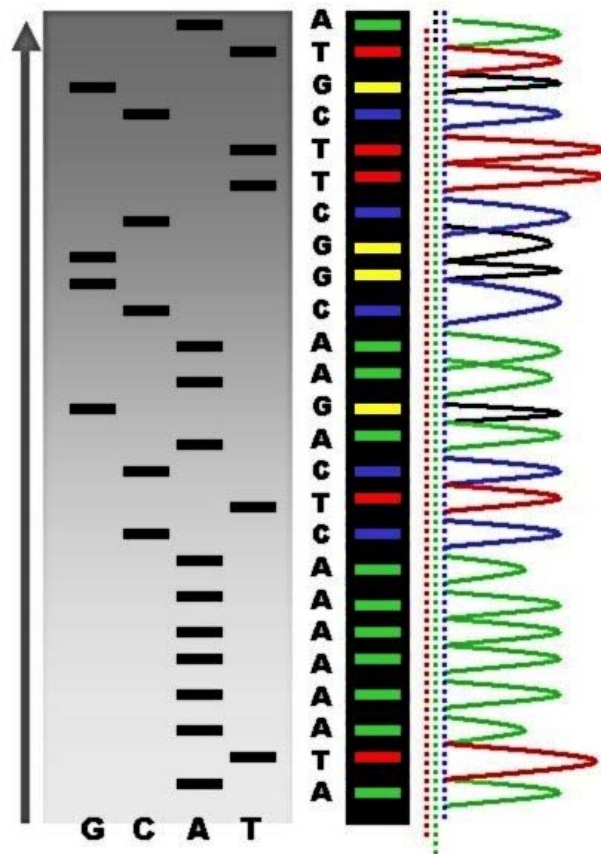




**Figure 1.4** the illustration of Sanger sequencing method.<sup>19</sup>

Subsequent electrophoretic analysis of each reaction mixture (see **Figure 1.5**) will reveal the sequence of the fluorescent labeled dye on the DNA.

The target sequence is the compliment of the labeled strand.



**Figure 1.5** a typical gel result of sanger sequencing method.<sup>5</sup>

The Sanger sequencing method has revolutionized research in several fields, including biology, medical diagnostics, forensics, and many others. It is much more efficient and requires much fewer radioactive or toxic chemical reagents than the Maxim-Gilbert method. The cost and speed of analysis is also significantly lowered due to the simplicity of the experiment design and the necessary instruments. Due to Sanger's contribution for developing the modern DNA sequencing method, he was awarded the Nobel prize in the year of 1980.

### **1.2.3 The drawbacks and limitations of Sanger sequencing**

Despite its utilities, the Sanger sequencing also has several drawbacks and limitations:

1. The ability of Sanger method to distinguish two different nucleotides is entirely based on the relative difference in length of the single strands to which they are attached, as seen in **Figure 1.5**. The quality of the gel results corresponding to the first 15-40 nucleotides is reasonable, however, after 700-900 nucleotides the results are questionable and often blurry.<sup>6</sup>

And each experimental run is limited to 300-1000 bases, which means that sequential readout of long DNA molecules (Genome DNA) is impossible unless they are first cut into small fragments.

2. The Sanger method is based on amplification that is dependent on enzymatic activities. Thus, inaccurate polymerization may also become a problem since enzymatic errors occur and proliferate during the amplification process.

3. Fluorescent labeling, although safe and convenient, is cost prohibitive especially when extremely long DNA must be sequenced.

The next generation sequencing requires a mechanism that provides a fast (24 hour human genome sequencing), accurate, and relatively

inexpensive (under \$1000 genome sequencing) reading device that is enzyme free (no reagent cost and minimum sample preparation).

### **1.3 Nanopore sequencing: the major approaches towards the third generation sequencing**

As illustrated above, the second generation DNA sequencing is efficient to perform but requires an error prone pre-amplification process to enhance the signal. Thus, one critical and challenging requirement for the third generation DNA sequencing is to perform sequencing in single molecule level.<sup>20</sup>

Several unique approaches have been reported to achieve this goal,<sup>20-29</sup> however, there is a consensus that the ideal method would incorporate a nanopore based device that is just large enough for a single strand of DNA to translocate through it. The device would also include a readout mechanism that identifies the sequence of the DNA being translocated.

The two major categories of nanopores are protein based nanopores (alpha hemolysine<sup>17, 30-32</sup> or MspA<sup>33</sup>) and solid state nanopores (SiN<sup>34</sup>, Graphene<sup>22, 35, 36</sup>, Carbon nanotube<sup>37</sup> et al).

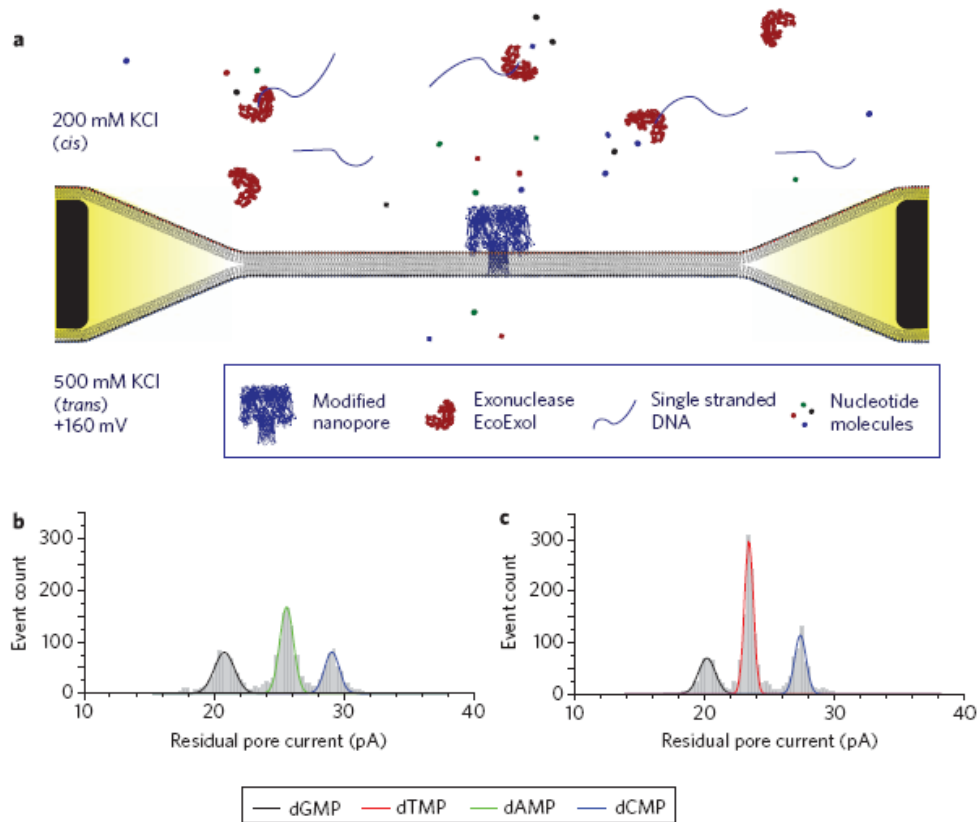
Both types of pores have their advantages and disadvantages. Protein pores are atomically accurate in their structures and size compatible with ssDNA. However, the protein nanopores require a lipid bilayer membrane to support its conformation and maintain its function. In most cases, protein nanopores only remain functional for several hours for measurements. It is also difficult to make a silicon chip based device with protein nanopores due to its inherent instability.

On the contrary, solid state nanopores have the advantage of incredible stability and the ability to be parallel made on wafer chips. It is also convenient to put nano-electrodes on the devices for electrical measurements. However, the smaller the size of a solid state nanopore, the more challenging it is to obtain high yields and reproducibility.<sup>38</sup>

### ***1.3.1 Protein nanopores***

So far, protein nanopores have proven to be the most reliable and successful nanopore being investigated. It is reported that the  $\alpha$ -hemolysine nanopores can function at extremely high temperatures, even up to the boiling point of water. The size of these pores is quite uniform and well-defined thus only one strand of DNA can pass through at any given time. Several fascinating experimental results and devices have

been reported by the Hagan Bayley group in Oxford, a pioneering team in this field.<sup>17, 39-43</sup> As illustrated in **Figure 1.6**, DNA or nucleotide translocation through the nanopore will block the ion channel temporarily and gives a drop in the ionic current. This level drop is determined by the cross sectional area of the DNA or nucleotide being translocated.

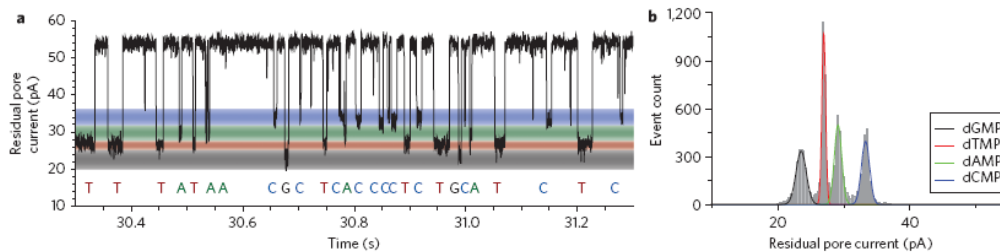


**Figure 1.6** (a) Schematics of  $\alpha$ -hemolysine nanopore nucleotide identification based on the exonuclease activity. (b), (c) statistical histograms of the blockage signals from different nucleotides passing through the protein nanopore.<sup>26</sup>

It has been reported that sequential identification of a DNA sequence can be achieved by reading digested nucleotides<sup>26</sup> by  $\alpha$ -hemolysine as seen in **Figure 1.6**. However, due to its lack of single base resolution, protein

nanopores are not able to achieve single base resolution within a strand of ssDNA.

As seen in **Figure 1.7**, a single DNA nucleotide passing through the  $\alpha$ -hemolysine nanopore generates a signature signal both in height and dwell lifetime and thus gives the information of the composition of DNA. However, the size of the  $\beta$  barrel of  $\alpha$ -hemolysine is too big to resolve single base on a piece of ssDNA. While sequential digestion of nucleotides at the pore end may tell the sequence of the DNA. Therefore, the key point in a  $\alpha$ -hemolysine sequencing is the incorporation of a DNA processing enzyme within the nanopore.



**Figure 1.7** Sample Ionic Current Signals of DNA nucleotide when translocating an  $\alpha$ -hemolysine nanopore.<sup>44</sup>

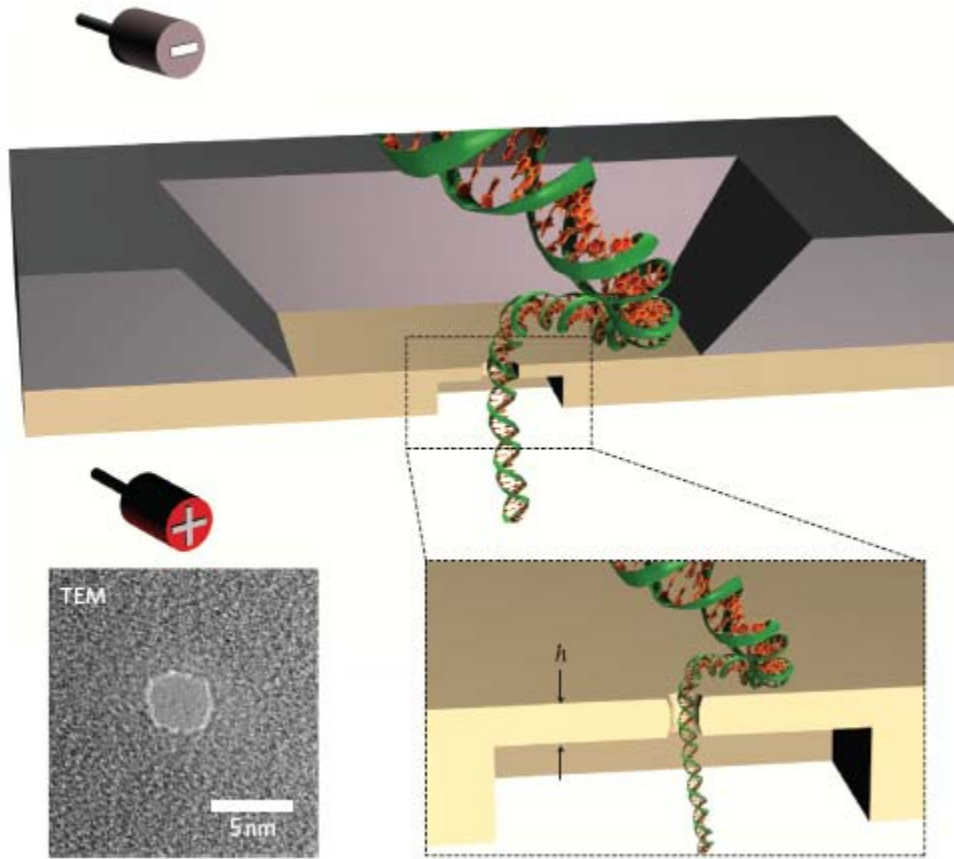
### 1.3.2 Solid state nanopores

Compared to protein nanopores, solid state nanopores have several advantages including the ease of massive production and the overall durability. The most widely used solid state nanopores are based on SiN

membrane formed on a Si based wafer device by chemical etching. Ion beam sculpting and high intensity Transmission Electron Microscope (TEM) illumination can also provide precise control of pore drilling.

Similar to the DNA translocation in protein nanopore, the first DNA translocation through a solid state nanopore was reported at Harvard University a decade ago. However, the major drawback is the size and the thickness of the pores. Also, for most dsDNA, the translocation time for a single molecule could be transient, too fast to provide meaningful information.<sup>6</sup> Synthesizing nanopores with a thickness less than 10 nm is also a considerable challenge for current industrial fabrication techniques, although 10 nm is still 3-4 times larger than the size of a single nucleotide.

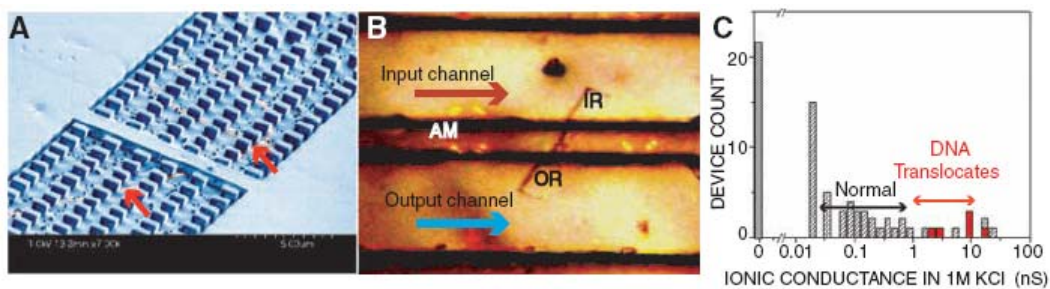




**Figure 1.8** Schematic diagrams of DNA translocation through an ultra thin SiN membrane nanopore. Thickness ( $h$ ) of the pore is below 20nm for ultimate sensitivity of molecular sensing.<sup>34</sup>

It has been reported that ultra-thin SiN membranes (below 10 nm) has successfully discriminated dsDNA, dsRNA and tRNA by its shape difference, which is the extreme achieved by solid state nanopores so far (see **Figure 1.8**).<sup>34</sup> Single base resolution has still not yet been achieved so far. Generating even thinner SiN membranes are beyond the ability of current industrial standard and such thin membranes would be too fragile to handle. Graphene, a material composed of a single layer of carbon, is supposed to replace the traditional SiN for nanopore synthesis. It is the

thinnest material in the world and has the advantage of incredible durability. Three independent groups simultaneously reported the successful DNA translocation through graphene pores.<sup>22, 35, 36</sup> Based on their success, it may be possible to realize single base resolution in the near future. Proposed models are discussed in detail in Chapter 7.



**Figure 1.9** (A) SEM image of the translocation reservoir. Single walled carbon nanotube (CNT) is placed underneath the pillars across the reservoir. (B) Optical image of input and output channel. (C) The yield of DNA translocation event in the prepared device. Red bars indicate the open device with DNA translocation signal (positive spikes). Grey bars represent open device with no DNA signal.<sup>37</sup>

The use of carbon nanotubes is another approach in the nanopore field (see **Figure 1.9**). The size of a single walled carbon nanotube is about 1-2 Å, which is an ideal size for a strand of ssDNA. By CVD growth of a single walled carbon nanotube, a nanopore device which selectively translocates ssDNA can be economically achieved.<sup>37</sup> However, due to the geometrical constraints, to achieve single base resolution on a tube shaped pore (10-20 μm length), requires the attachment of functional reagent on the pore end via covalent linkages.

## **2. Molecular electronics and Scanning Tunneling Microscopy**

### **2.1 Molecular Electronics**

As a newly developed research field, molecular electronics is defined to be the electronics whose behavior is dictated by the chemical, physical and electronics structures of the molecules and thus differ from ensemble properties measured in the macroscopic situation.

With the fast development in the silicon based semiconductor industry, the size limit of moore's law is being approached everyday.<sup>45, 46</sup> People have dreamed for decades to build electronic devices with the size of single molecules which is below 1 nm.<sup>47</sup>

In most reported publications regarding single molecule electronics measurement, setup with two electrodes (source/drain)<sup>21, 48-50</sup> or three electrodes (source/gate/drain)<sup>51-54</sup> are most popularly used. The source/drain electrodes are normally made of gold for the simplicity making gold/thiol linkage thus the sandwich structure (metal-molecule-metal) could be formed.<sup>55</sup>

### **2.1.1 Landauer Formula**

Ohm's law has successfully explained the electrical behavior of macroscopic materials.

$$G = \frac{\sigma A}{L} \quad (2.1)$$

As shown in the expression of this well known formula, the conductance of a rectangular shaped material is proportional to its cross section (A) and is inversely proportional to the length (L) with  $\sigma$  as a characteristic constant for a specific type of the material.

However, the electrical behavior of materials in nano-scale is significantly different from the bulky phase and can only be explained by quantum theory.<sup>56</sup>

There are 2 major reasons for such differences: 1) The mean free path and the phase coherence length of electrons in metals are around nm size. Thus, the electron transport process is ballistic rather than diffusive in nm scale. 2) The contact between the macroscopic electrode and the nanoscale conductor could significantly change the conductance properties of the material and gives uncertainties.

For ballistic electron transport through (metal/molecule/metal) junction, the system could be simply depicted by Landauer theory.<sup>57</sup>

The most simple way to understand Landauer theory is depicted by a 1-D model:

Assuming that there is a 1-D sample of length L, the density of states within the region of k to k+dk is:

$$n(k)dk = 2 \frac{1}{L} \frac{L}{2\pi} dk \quad (2.2)$$

Here the 2 coefficient comes from the electron spin degeneracy. The group velocity of the electron is:

$$v_g = \frac{\partial \omega}{\partial k} = \frac{1}{h} \frac{\partial E}{\partial k} = \frac{k\hbar}{m} \quad (2.3)$$

The current flowing through the sample is then:

$$I = e \int_0^{\infty} v(k)n(k)f_1(k)dk - e \int_0^{\infty} v(k')n(k')f_2(k')dk' \quad (2.4)$$

$$I = e \int_0^{\infty} \frac{\hbar k}{m} \frac{1}{\pi} f_1(E) \frac{dk}{dE} dE - e \int_0^{\infty} \frac{\hbar k'}{m} \frac{1}{\pi} f_2(E') \frac{dk'}{dE'} dE' \quad (2.5)$$

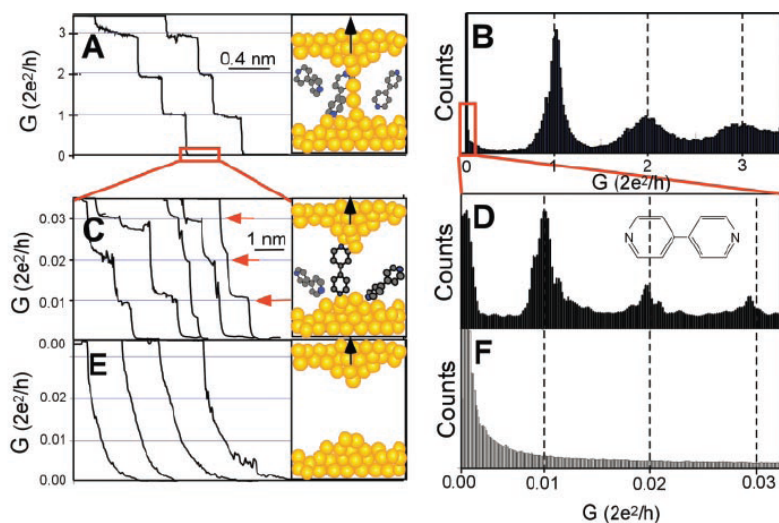
If we assume that the Fermi distribution can be described by a step function at extremely low temperature or small bias conditions, the conductance G of the system at a small bias V is then:

$$G = \frac{2e^2}{h} T(E_F) \quad (2.5)$$

Here e is the electron charge and h the plank constant. For a perfect situation, T=1 and G becomes G<sub>0</sub>.

$$G_0 = \frac{2e^2}{h} = 77 \mu S \approx (12.9 K\Omega)^{-1} \quad (2.6)$$

Here  $G_0$  corresponds to the maximum conductance that could be achieved by a point contact. Landauer first predicted the existence of quantized conductance of point contacts by metal electrode and the phenomenon has been widely observed by several experiments. The most straightforward proof is by a STM break junction measurement pioneered by NJ Tao's group.<sup>57</sup> In their classic paper, the STM tip is forced to advance towards the gold surface and form point contacts. After withdrawing mechanically, a series of plateaus with quantized conductance are detected which corresponds to values multiples of the Landauer conductance.



**Figure 2.1** STM break junction measurement (A) I-S curves showing quantized Landauer conductance as plateaus. (B) The corresponding histograms with peaks at 1, 2 and 3 times of  $G_0$  (C) I-S curves showing plateaus corresponding to the bridging molecules. (D) The corresponding histograms with peaks at multiples of  $0.01 G_0$ . (E) I-S curves showing no plateau when no bridging or gold wire exists. (F) Corresponding histograms showing no peaks.<sup>57</sup>

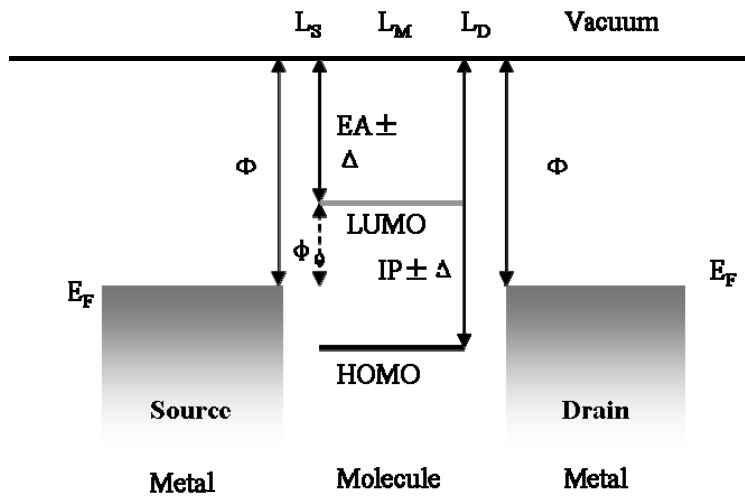
Generally, the Landauer formula at finite bias with a series of quantum modes is written this way.

$$G(E) = \frac{2e^2}{h} \sum_{i,j} T_{ij}(E) \quad (2.7)$$

Here  $T_{ij}$  corresponds to the probability for the electron to transmit from the  $i^{\text{th}}$  mode at the left sample to the  $j^{\text{th}}$  mode at the right sample. For finite bias situation, the current, which represents the intensity of the electron flow, is more generally written as:

$$I = \int_0^\infty \frac{dE}{e} [f(E + eV) - f(E)] G(E) \quad (2.8)$$

Here  $f$  is the Fermi function and  $V$  is the applied bias.



**Figure 2.2** Schematic diagram of electron transport through Metal-Molecule-Metal junction. The Fermi level  $E_f$ , which normally locates between the lowest unoccupied molecular orbital (LUMO) and the highest occupied molecular orbital (HOMO). The vacuum level is above the Fermi level with equal amount energy of the work function ( $\Phi$ ).<sup>58</sup>

From **Figure 2.2**, for a 1-dimensional case, to tunnel from one electrode to another through vacuum, an energy barrier  $\Phi$  is expected. This energy barrier is the energy difference between the metal Fermi level and the vacuum, which is also called the work function. By solving Schrödinger equation, it is known that the current decays exponentially with the gap distance (L) with a decay constant  $\beta$ .

$$\beta = 2\sqrt{\frac{2m\phi}{\hbar^2}} \quad (2.9)$$

Once a molecule is placed within the gap, the vacuum energy level is replaced with a set of molecular orbitals. The Fermi level of metals normally locates between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). In the case that the electron is the charge carrier, the energy barrier is then lowered to the LUMO level of the molecule which spans the gap then. The decay constant  $\beta$  becomes:

$$\beta = 2\sqrt{\frac{2m^*(E_{LUMO} - E)}{\hbar^2}} \quad (2.10)$$

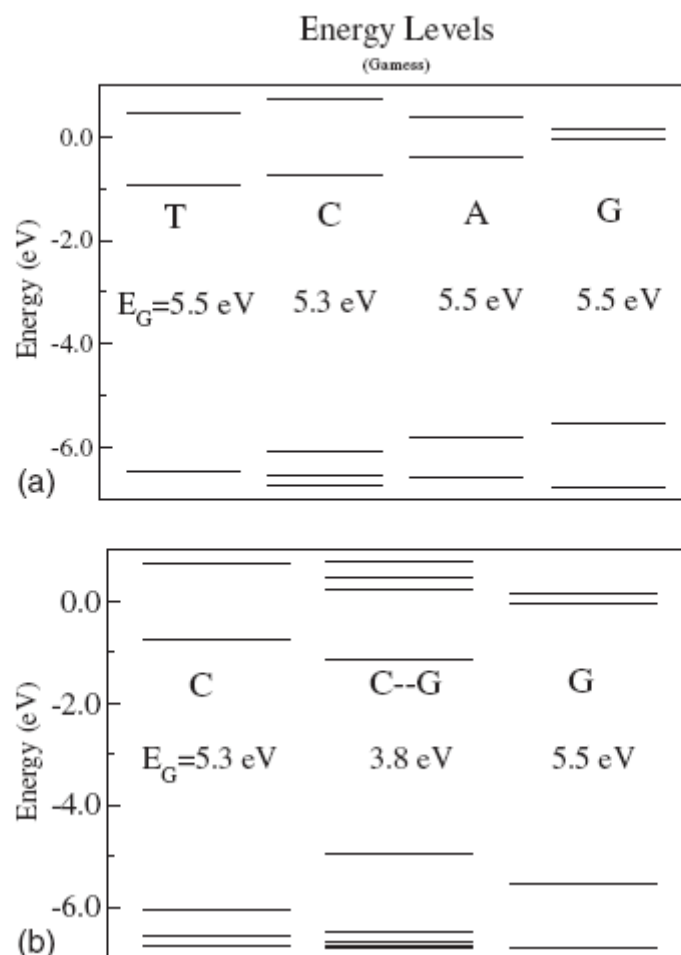
Since the electron is located around the energy extrema, so an effective mass  $m^*$ , same as the effective mass in a semiconductor device, is used in the formula. The effective mass is shown below and it does depend on the energy diagram of the molecule.

$$m^* = \hbar^2 \left( \frac{d^2E}{dk^2} \right)^{-1} \quad (2.11)$$



It is obvious that the decay constant  $\beta$  due to the participation of the bridging molecule could be much less than the value in vacuum tunneling. Thus, the bridging molecule helps to increase the conductivity of the junction.

From this point, we know that the efficiency of this conductivity of the molecule majorly depends on the alignment of the Fermi level of the electrode to the LUMO or HOMO of the bridging molecule and when  $E_{\text{LUMO/HOMO}} - E = 0$ , resonant tunneling occurs.<sup>59</sup> And this simple electrode junction forms a perfect molecular sensor since different molecules, which have different HOMO-LUMO level should behave differently in the aspect of recognition tunneling.



**Figure 2.3** The energy levels calculated for four different DNA bases and DNA base pairing by DFT method.<sup>60, 61</sup>

In **Figure 2.3**, the calculated energy diagram of DNA bases is showing that four different DNA bases are having significantly different LUMO levels. Related experimental measurement of DNA bases conductance is performed in Chapter 5 and 6.

Another important phenomenon is that when DNA base pairs are formed, compared to individual bases, the HOMO-LUMO gap is normally

decreased and thus enhanced its conductance. Based on this phenomenon, measurements are taken and described in Chapter 3 and 4.

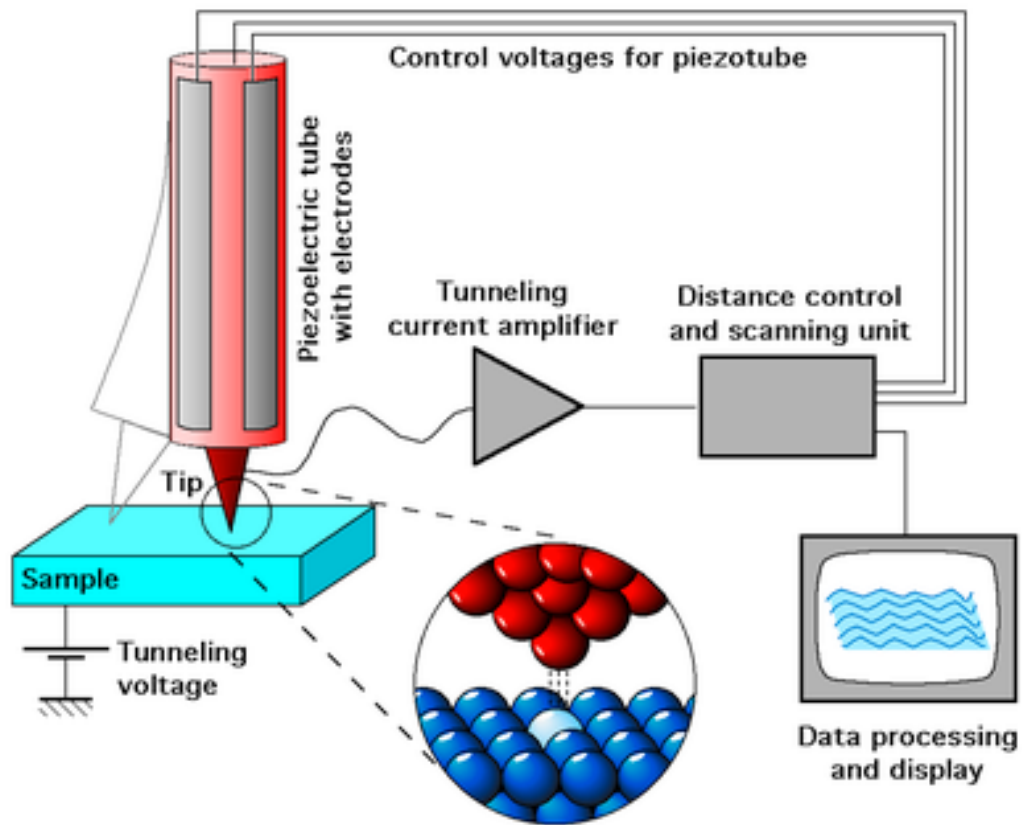
Many published paper have successfully shown that electrical devices based on molecules could be used as electronic components such as diodes<sup>62</sup>, switches<sup>63-65</sup>, rectifiers.<sup>66, 67</sup> The \$1000 human genome research requires a technique that could sequence the entire human genome in 24 hours with less than \$1000 cost. Thus we aim to develop a molecular electronics based DNA sensor which could identify DNA bases and get the information of its sequence by detecting the electrical properties of the DNA thus no reagents are consumed during the detection which makes it a minimum-cost device.<sup>6</sup> The electronic detection itself is in nature fast in speed and due to the nanoscale size of the device, the eventual goal is to make thousands of parallel detection channel on a same tiny chip and thus increase the speed. However, due to the difficulty of making nano-fabricated device at the current situation, we chose STM as a platform to mimic the detection and simplify the system so the yield of data detection could be increased. Below is a brief introduction of Scanning Tunneling Microscope and its application in my measurement.

## **2.2 Scanning Tunneling Microscopy**





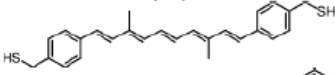
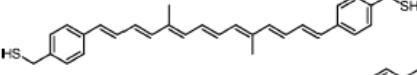
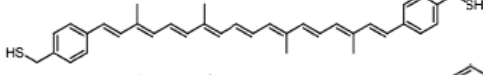
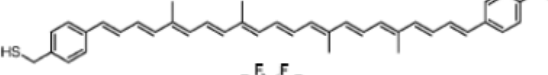

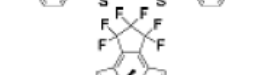
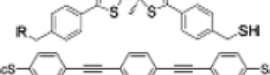
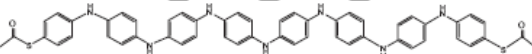
The invention of STM and Atomic Force Microscope (AFM) has for the first time enabled scientists to visualize or even manipulate single molecules and thus measure their properties. In a typical STM setup, a metal-metal junction with only few molecules in the gap is easily achievable with low cost instruments. Several well designed measurement modes have successfully demonstrated the feasibility to do measurement in single molecule level and measure the properties of simple thiolated molecules in close agreement with theoretical predictions as listed in Table 2.1.<sup>68</sup>

### ***2.2.1 The history of STM***

The development in quantum physics is telling us that the physics in the nanoscale or below could be dramatically different from the macroscopic world. Thus studying the world in the scale of nanometers would be of great interests. The curiosity of people studying unfamiliar world has never been satisfied and the tools that facilitated researcher with more ability monitoring the atomic leveled activities has opened up the gate of nanoscience and technology in the recent years. As suggested by the quantum theory, the physical contact of two metals is not the necessary requirement for electron conduction and the tunneling current decays exponentially with distance. The invention of scanning tunneling microscope has demonstrated this theory and the imaging technique developed has increased the resolution to be around atomic level.



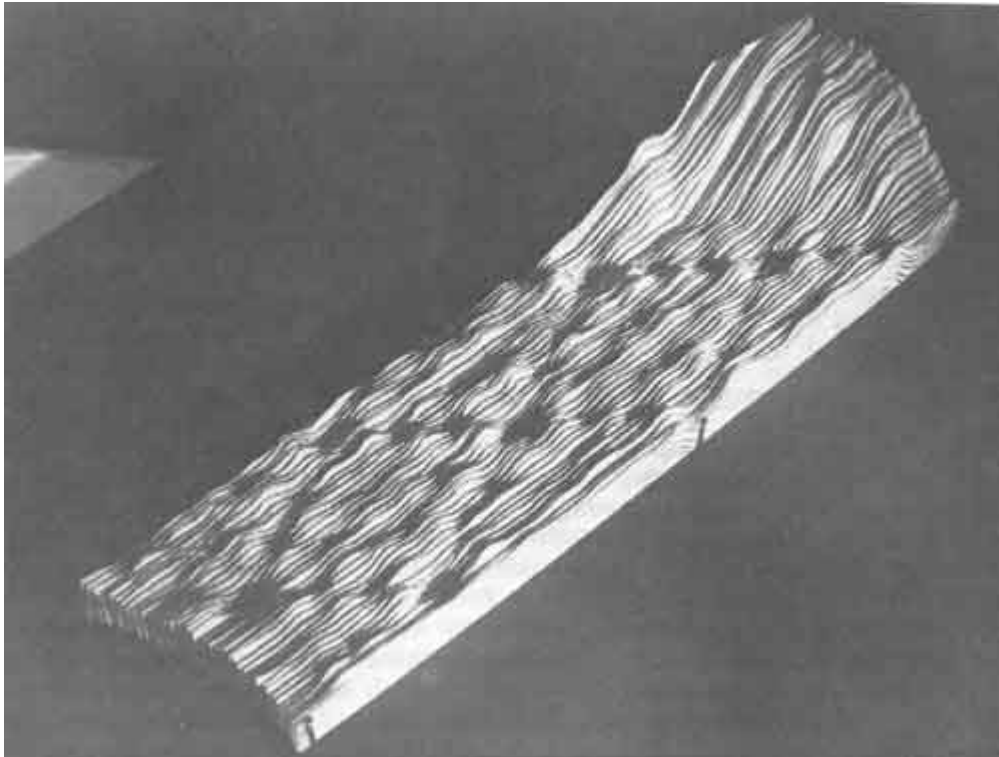
**Figure 2.4** The Schematic Diagram of Scanning Tunneling Microscope. The 2-D sample and a metal probe form the conducting junction. The tip movement is controlled by a piezoelectric tube with a feedback loop. Scanning imaging is performed by forcing the tip to make a raster scan.<sup>58</sup>

	Molecule	G (measured) [nS]	G (theoretical) [nS]	Ratio
1		95 ± 6	185	0.51
2		19.6 ± 2	25	0.78
3		1.6 ± 0.1	3.4	0.47
4		833 ± 90	47 000	0.02
5		2.6 ± 0.05	7.9	0.33
6		0.96 ± 0.07	2.6	0.36
7		0.28 ± 0.02	0.88	0.31
8		0.11 ± 0.07	0.3	0.36
9		1.9 ± 3	0.8	2.4
10		250 ± 50	143	1.74
11		-13	190	0.07
12		0.32 ± 0.03	0.043	7.4

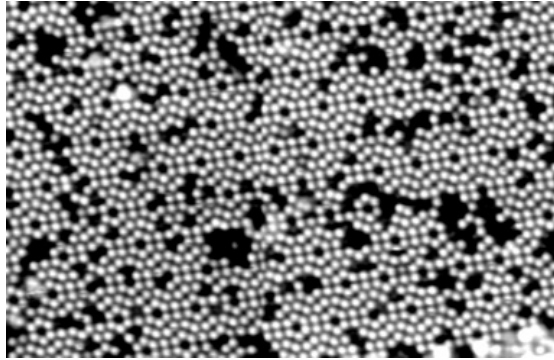
**Table 2.1** Single molecule conductance determined in the published paper. In comparison with predicted theoretical value and the  $G_{\text{measured}}/G_{\text{theory}}$  ratio, a good agreement is obviously seen.<sup>68</sup>

The first STM device was made by Gerd Binnig and Heine Rohrer working for IBM Switzerland in 1979<sup>69-76</sup>. As a graduate student, Binnig has already proposed the idea that scanning by a sharp probe over a clean surface would generate an image from the height change induced tunneling current change. Today, we can still see that the predicted resolution by Binnig was 45 Angstrom in his original proposal. However, when the idea was realized, he noticed that the much higher resolution could easily be achieved and atomic resolution was even detected.<sup>70</sup>

The sharpness of the tip seems not to be a very strict requirement for high resolution imaging. As suggested by the tunneling theory, the current decays exponentially with distance which made the contribution of the closest atoms to the surface be the strongest and thus increased the resolution easily, which means, a not well-made tip monitored under optical microscope is not necessarily a failed one as long as there is a chance for several atoms locating just at the top of the tip and gave satisfied resolution. We will talk about the details of making an STM tips in **Appendix B**.



**Figure 2.5** The first STM image of Si 7x7 lattice structure. Image comes from manual reconstruction of the collected data.<sup>76</sup>



**Figure 2.6** An STM image measured by a modern Scanning Tunneling Microscope. Si 7x7 lattice is clearly shown.<sup>77</sup>

As seen in **Figure 2.3**, the 1<sup>st</sup> STM is composed of a scanning probe and a tripod consisting of three piezoelectric elements (showing in the figure as Px, Py and Pz). Changing bias applied across the piezoelectric element will generate displacement with nanometer accuracy so the movement of the tip in 3 dimensions is well defined and controlled. Before the imaging, the tip was advanced towards the target surface with a stepper motor and once the threshold current is achieved indicating that tunneling range is reached, the advance is terminated and a feedback circuit functions to maintain the current at the setpoint level by fine tuning the Pz bias which moves the tip vertically. During the scanning, the Px and Py bias is modulated so that a raster scanning is going. The Pz bias is kept changing so that the setpoint current is maintained constant. Topographical images could be reconstructed by the height change along the Z direction.



### **2.2.2 The limitations of STM**

During more than 30 years of nanoscale research, STM has served as one of the few tools with the capability of atomic resolution imaging while the cost making it is surprisingly low. However, it has its limits measuring conducting materials and the target sample has to be flat and clean for optimum results. It also showed disadvantages of stationary imaging which made real time imaging a big challenge and the interpretation of the image can be challenging when complicated materials are studied.

### **2.3 Why STM for DNA sequencing**

DNA sequencing technology has been developed for 30 years and has provided invaluable information for life science development.<sup>19</sup> However, the urgent call for next generation DNA sequencing focuses on the feasibility of single molecule leveled sequencing with long reads.<sup>78</sup>

Being one of the few instruments with nm resolution, achieving single molecule resolution is of no doubt by STM. For years, single molecule conductance measurements on simple organic molecules have been successfully done.<sup>21, 48, 57, 79-93</sup> And in the most recent decades, a bunch of well designed techniques have been developed as potential candidates for single molecule sensing by tunneling.<sup>21, 48, 57, 79-93</sup>

In this work, several STM based single molecule measurement techniques have been developed based on the basic concept of the  $I(t)$  method which we call it recognition tunneling.<sup>94</sup> These modes are specifically designed to fit the measurement of DNA nucleotides in a functionalized metal junction with a fixed size.

In the following chapters, I will illustrate several approaches that I have made to identify all four DNA bases by recognition tunneling.

## 3. THE CONDUCTANCE MEASUREMENT OF DNA BASE ON NUCLEOSIDES

### 3.1 Introduction

DNA, the ultimate genomic code carrier<sup>3, 95</sup>, is known to have strict base-pairing rules between Adenine/Thymine and Guanine/Cytosine base pairs<sup>95</sup>. Here, hydrogen bonds drive the highly specific base-pairing and maintain the double helix structure. Besides being the driving force directing base-pairing, the hydrogen bond is also believed to be a great electron tunneling pathway. In this chapter, I demonstrate approaches measuring tunneling conductivities via DNA base-pairing through thiolated DNA nucleosides and bases using STM. Telegraph switching signals which represent the On/Off state of a single base-pairing is intensively studied and compared. The measured conductance fits exactly with theoretical predicted order of 2AAT>AT>GC (see **Figure 1.2**). This work not only opened a gate for single molecular conductance measurement for DNA base-pairing but also inspired the designing of the universal reader for the next generation DNA sequencing.

### 3.2 Single molecule Conductance Measurement by STM

#### 3.2.1 Break Junction: the $I(s)$ method

$$I = I_0 \exp[-\beta(\Delta x)] \quad (3.1)$$

The tunneling current decays with distance exponentially as depicted in Equation 3.1. Falling by a factor of 10 for opening each 1 angstrom of the tunneling gap, the overwhelming majority of the tunneling current comes from the tiny apex of the scanning probe and could resolve single molecule signals. Such sensitivity has enabled tunneling detection to be an ideal candidate for electronic sensing of DNA sequences<sup>96, 97</sup>. STM has offered a low cost platform for such measurement. However, the bluntness of the scanning tip (see **Figure B.1a**) normally makes the cross section of the junction so big that multiple molecules could stay in it (see **Figure 3.3b**). To gain single molecule resolution, break junction method is commonly used which made the tip almost touching the surface on purpose.<sup>57, 93, 98-101</sup> Step shaped plateau which represents molecular breaking one by one could be detected in the  $I(s)$  curve.

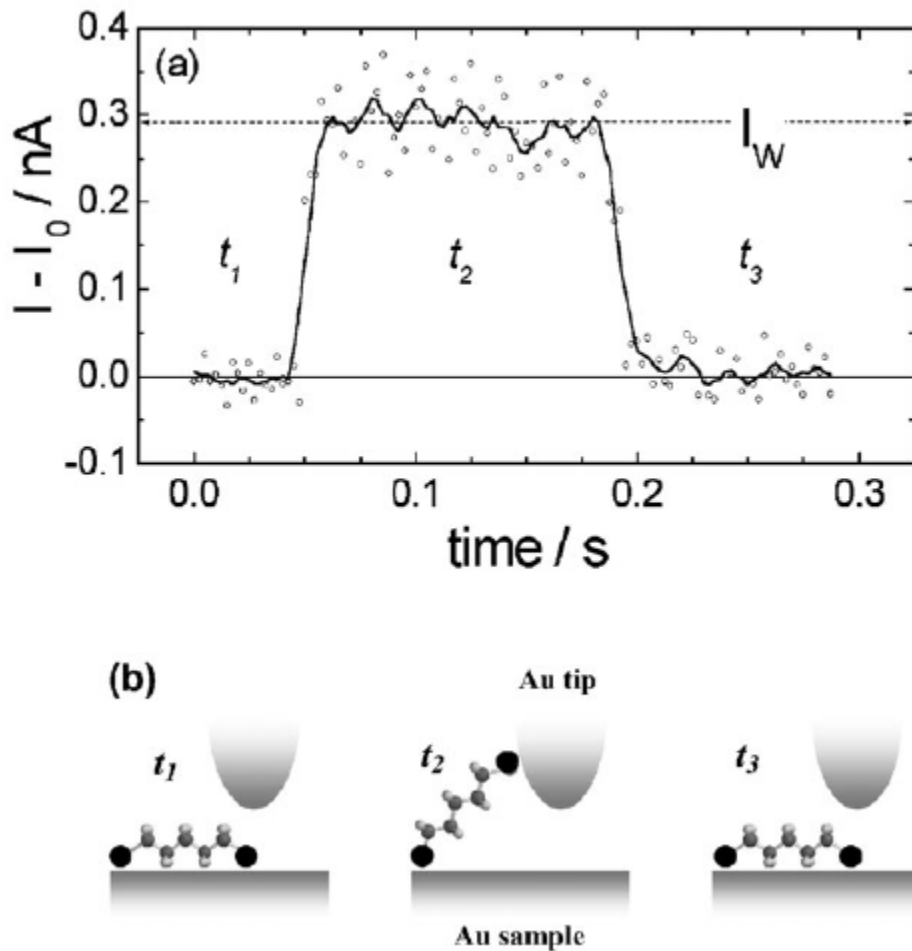
However, the situation for hydrogen bonded tunneling measurement could be slightly different and the weaker bonding force and molecular conformation complexity made the break junction method not appropriate for the measurement and not yet reported to my knowledge.<sup>6</sup>

### 3.2.2 Spontaneous On/Off Switching: the $I(t)$ method

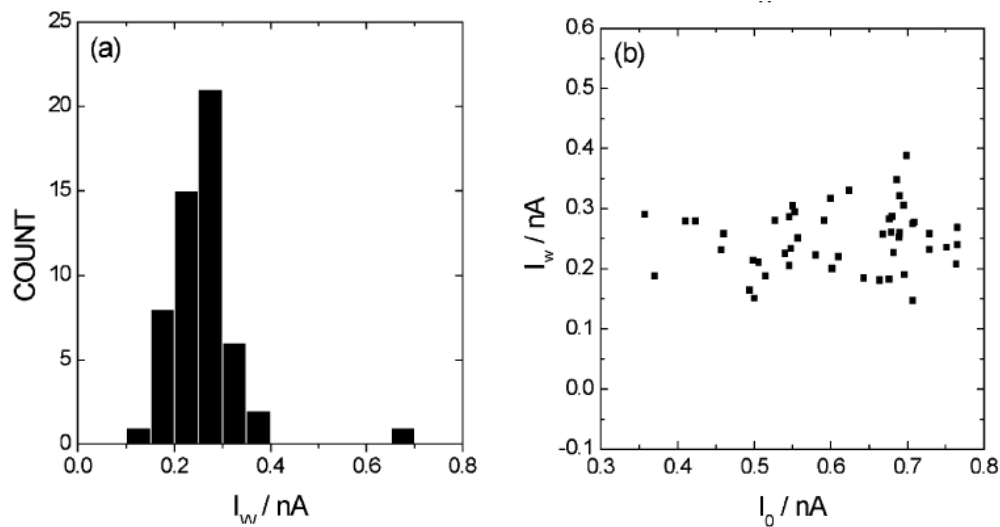
Ramachandran *et al*<sup>102</sup> showed that stochastic fluctuations in the molecule-metal contact give on/off switching in series of STM images. Similar to the results from Ramachandran, Wolfgang Haiss *et al*<sup>84, 86</sup> demonstrated that by placing a gold STM tip above a bis-thiolated alkane chain functionalized gold surface, the time course of the tunneling current (the  $I(t)$  method) through metal/molecule/metal junction shows spontaneous switching between 2 levels. These telegraphic signals show strong evidence of single molecule being trapped in the gap since the chance getting spontaneously synchronized on/off switching from multiple molecules is negligible<sup>21, 48, 87</sup>.

The unambiguous nature of single molecule detection makes  $I(t)$  method to be a better candidate for detecting weak bonding (DNA base-pairing) mediated tunneling. To perform the measurement, the junction size is manually controlled at a fixed size when a characteristic telegraphic signal shows up. 2-level telegraph switching contributed to the majority part of the signals. However, for too small a gap (too high baseline current), multiple molecules signals<sup>88</sup> could occasionally show up (see **Figure 3.3b**), these multiple leveled switching events are complicated to analyze and are normally ignored in the statistics.

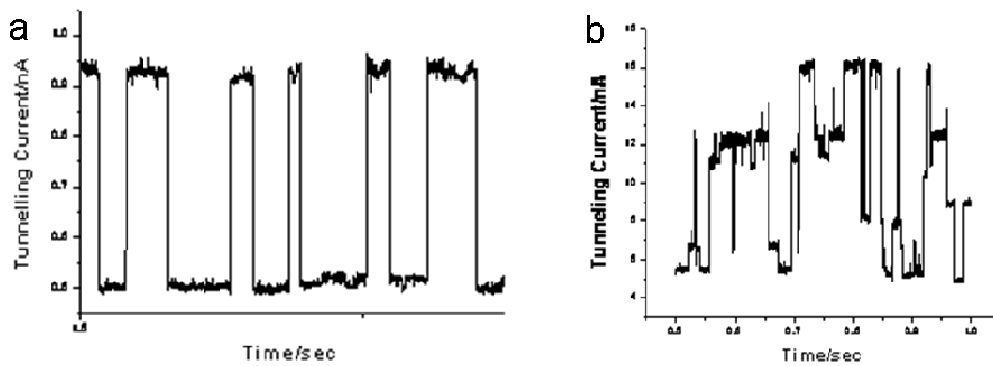
The advantage of this method in our situation is that we can set the gap size manually and when the signal comes up, the size of the gap is fixed so less complicated conformation change could interfere the measurement compared to the regular break junction measurements by which the gap size keeps increasing with time.



**Figure 3.1** (a) a typical signal of spontaneous ON-OFF switching during an  $I(t)$  measurement. The ON-OFF region are marked by the label  $t_1$ (OFF),  $t_2$ (ON) and  $t_3$ (OFF). (b) The corresponding schematic diagram of the bonding situation during the time period of  $t_1$ (unbonded),  $t_2$ (bonded) and  $t_3$ (unbonded)<sup>84</sup>



**Figure 3.2** (a) Histograms of 54 switching heights  $I_w$  for 1,8-Octanedithiol. (b)  $I_0$  vs.  $I_w$  plot for the 1,8-Octanedithiol. The measured switching height is independent of the baseline change.<sup>84</sup>



**Figure 3.3** Typical comparison of (a) single molecule telegraph switching and (b) multiple molecules trapped in the gap.

### 3.3 DNA base on nucleoside measurement

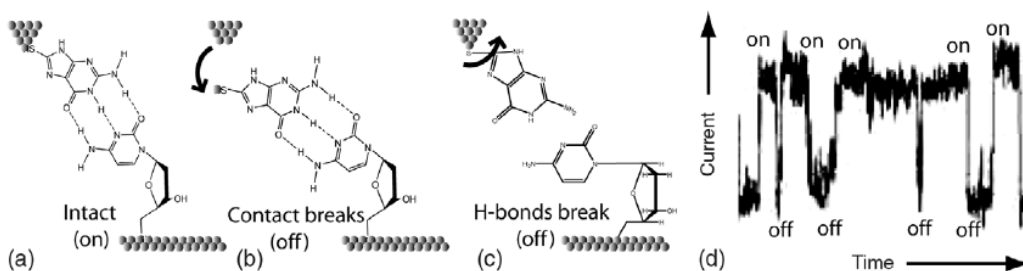
#### 3.3.1 *Materials and Methods*

We customarily synthesized thiolated DNA nucleosides 5'-S-acetyl-5'-thiodeoxycytidine and 5'-S-acetyl-5'-thiothymidine following published procedures<sup>103-105</sup>. A self assembled monolayer (SAM) of thiolated DNA nucleosides is formed on the gold surface by placing freshly annealed gold substrate in the dissolved solution of thiolates for an extended time<sup>49, 91, 106</sup>. 1mM methanol solution of thionucleosides (2 ml) was first treated with pyrrolidine (0.2 ml) for 1 hour to de-protect the thiol group. To keep organic contamination away and form atomically flat gold terraces, the gold substrate (Agilent, Chandler) is hydrogen flame annealed before functionalization in the solution. 2 hours immersion of the chip in the solution plus thoroughly rinsing and drying with methanol makes a monolayer of molecules on the gold substrates. Ellipsometry results give a layer height of  $0.92\pm 0.1$ nm for thiodexoycytidine and  $0.81\pm 0.08$ nm for thiothymidine. Since the calculated heights are 0.96nm (thio-Cytidine) and 0.97nm (thio-Thymidine), we can conclude that a monolayer of molecules are formed on the gold surface. High resolution STM imaging is also carried out by Pt-Ir tips in 1,2,4-trichlorbenzene environment. The image result is blurry implying that a random self assembled monolayer is formed.



The STM probes (see **Appendix B** for details) are functionalized via the same manner<sup>89, 91, 106</sup> but characterization is of challenge to date. Due to its cone shaped geometry and the tunneling principle, it is believed that only the molecules on the apex contribute to the signals so a well ordered SAM is of less importance for the probe then. Bare probes and thiophenol functionalized probes are also prepared for the control experiment.

A picoSTM 5500 (Agilent, Chandler) system is used with the sample and the probe both immersed in 1,2,4-trichlorobenzene. The non-polar and insulating nature of this organic solvent provides a hydrogen bond free environment and the tunneling pathway can only be formed via DNA basepairing.



**Figure 3.4** (a) An intact junction in which the tunnel gap is spanned by a guanine attached to the probe, hydrogen-bonded to a deoxycytidine attached to the substrate. (b) Fluctuations that break the metal-molecule contact will reduce the conductance, as will fluctuations that break the hydrogen bonds (c) An example of the telegraph noise signal produced as bonds break and reform.<sup>48</sup>

The drift of the probe could complicate and interrupt the measurement and could be easily monitored by checking the current fluctuation when turning off the servo feedback. To diminish the thermal drifting of the probe, it is

recommended to stabilize the system for several hours until the drifting becomes negligible.

During the measurement, the servo control is turned off and the tip height is manually controlled by a custom made Labview program which adjusts the Z voltage of the piezoelectric (PZT) element. Bursts of telegraph signals show up spontaneously and the gap size should be fixed during the period of recording.

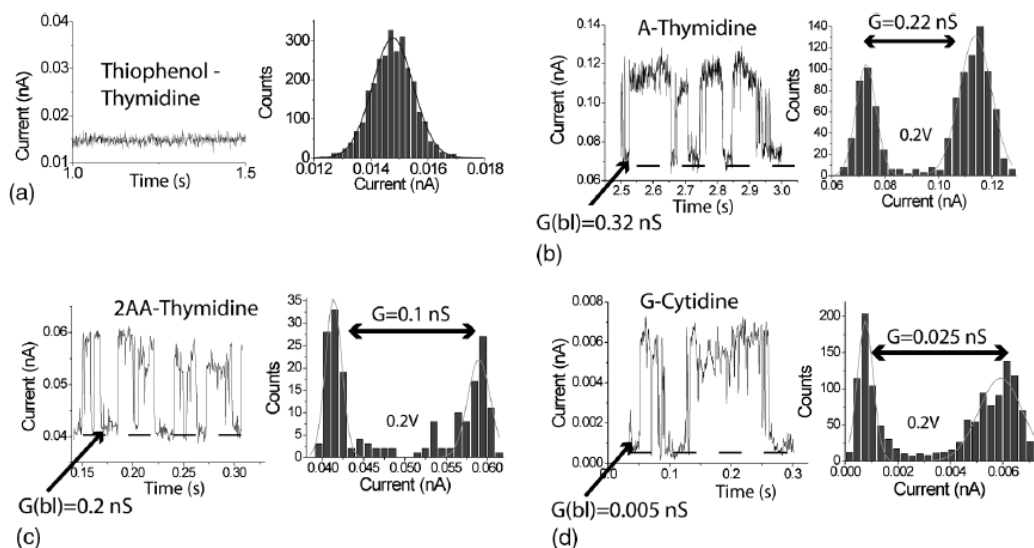
### ***3.3.2 Results and Discussion***

The signals are recorded by a digital oscilloscope (YOKOGAWA DL-750) with 5 sec duration periods. The recorded signals are analyzed by a custom made Labview program (see **Appendix A** for details). Differences between the two levels in a telegraph signal are considered the single molecule contribution of the tunneling current. By making a histogram for the telegraph signals (see **Figure 3.6**), the portion of molecular tunneling is measured by getting the difference between the centers of the two peaks.

	Thiothymidine	Thiodeoxycytidine	Thiophenol	Bare
8-mercapto-adenine	47±8%	-	1±0.3%	0
8-mercapto-2-aminoadenine	51±7%	-	-	0
8-mercapto-guanine	-	52.5±13%	-	0
Thiophenol	2.5±1%	-	0	0
Bare	0	-	0	0
Adenine	0	0	-	0
2-aminoadenine	0	-	-	-

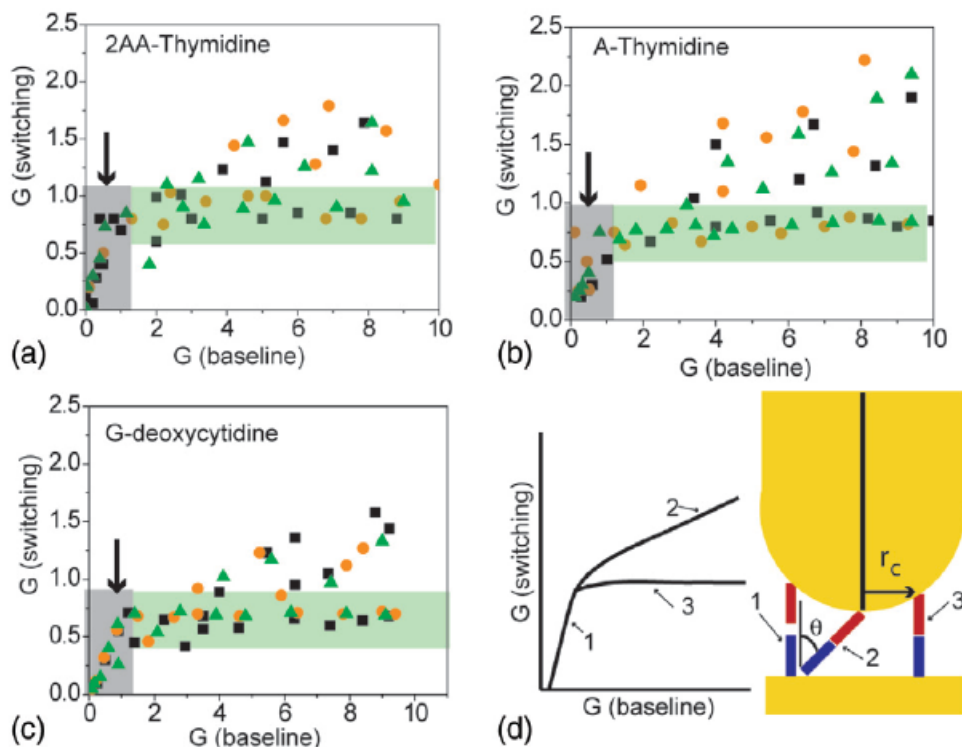
**Table 3.1** Observed frequency of switching for various preparations of the probe and the surface. Errors are 1sd. “-” represents untried combinations. The bottom two rows are for bases physisorbed onto the probes.<sup>48</sup>

Typical control experiment of thiophenol on thiothymidine is shown in **Figure 3.6(a)**, which showed clean traces of tunneling current through Au-Au space tunneling and thus no telegraph signal is detected. The histogram of such signal shows normal Gaussian distribution. In quite rare conditions, telegraph signals with small switching height and high switching frequency could show up when tip bias is set very low around 0.05 volts. Presumably at low bias, it is of more chance getting too close to the surface and the interaction between the aromatic benzene ring and the heterocycle may contribute to tunneling switching. In contrast, the H bonded system show much higher chance getting telegraph signals (see **Figure 3.5**) for all the cases. Summarized in **Table 3.1**, the experiments performed show strong evidence that the metal/base-nucleoside/metal sandwich assembly is the necessary condition to trigger molecular tunneling.



**Figure 3.5** Sample time traces for DNA base on nucleoside (b-d) and control (a) experiment. (a) control junction with thiophenol on the probe and thymidine on the surface, (b) adenine on the probe and thymidine on the surface, (c) 2-aminoadenine on the probe and thymidine on the surface and (d) guanine on the probe and deoxycytidine on the surface. Gaussian fits are shown as solid lines for the two peaks in the signal histogram. The difference between the 2 peaks is used to determine the single molecule conductance in a given run.<sup>48</sup>

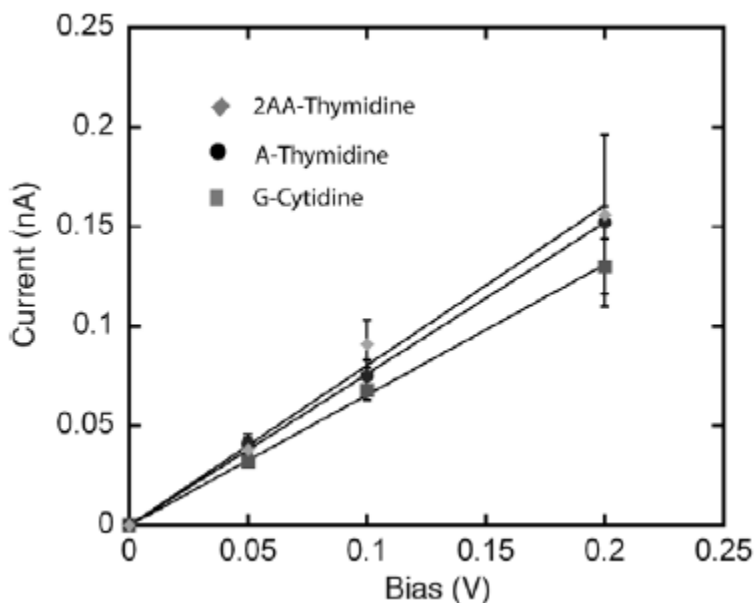
By analyzing all base on nucleoside results, we noticed that the measured switching conductance obviously have a variety of baseline dependence which directly corresponds to the size of the metal gap. Contact geometry and molecular conformation may both change the molecular conductivity in this case and thus more detailed analysis is needed to identify different base-pairing since the conductance is not a unique value.



**Figure 3.6** Plot of molecular switching versus baseline conductance for the three base-nucleotide combinations (a)-(c). (d) Illustrates mechanisms for the various regions. (1) the tunnel gap is larger than the equilibrium molecular pair length. So the molecular pairing is stretched. (2) The gap junction is smaller than the equilibrium molecular pair length, the molecular pairing still forms but may be tilted with an angle relative to the surface. (3) The gap size is just enough for the equilibrium molecular length. A plateau is shown in the plot as labeled by 3.<sup>48</sup>

Molecular conductance  $G_{(\text{switching})}$  and baseline conductance  $G_{(\text{baseline})}$  could be directly deducted by dividing the current  $I_{(\text{switching})}$  and  $I_{(\text{baseline})}$  by the tip bias  $V$ .  $G_{(\text{switching})}$  vs  $G_{(\text{baseline})}$  plot is shown in **Figure 3.6** to demonstrate this dependence. In all the H-bond systems, the  $G_{(\text{switching})}$  all show fast and proportional increasing trend with  $G_{(\text{baseline})}$  in the bottom left regions of the graphs. This part is marked as region 1 in the model graph in **Figure 3.6d**. Passing a critical value of  $G_{(\text{baseline})}$  a long plateau region is obviously seen as marked by region 2 in the model graph. Besides, region

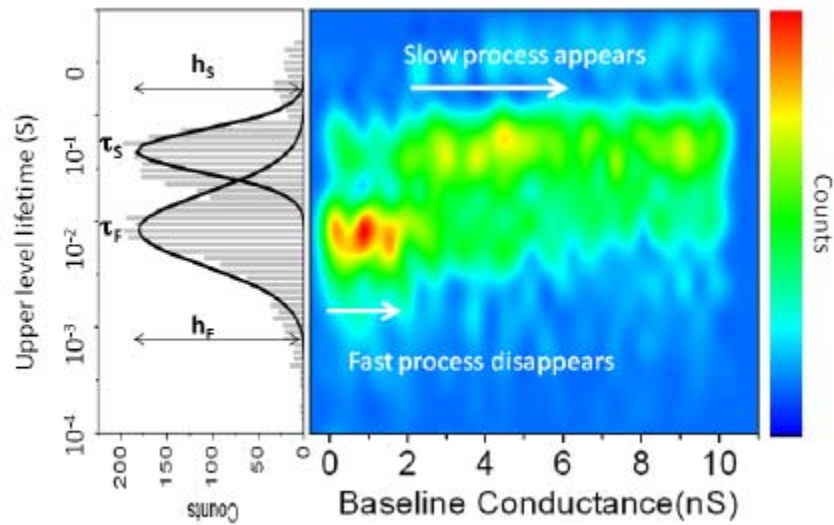
3 corresponds to another increasing trend above the plateau but with non-saturated  $G_{\text{(switching)}}$  value.



**Figure 3.7** I-V curves for 2AA-thymidine (diamonds), A-thymidine (dots) and G-deoxycytidine (squares). Each data point is the mean obtained from data in the plateau regions. The error bars correspond to  $\pm 1\text{sd}$ .<sup>48</sup>

A similar phenomenon has also been reported by Haiss et al<sup>84, 86</sup>, who claimed that the conductance of a rigid molecule increases exponentially as the size of the gap shrinks, similar to region 3 in our case. We interpret that all the 3 regions correspond to different molecular conformation states of the base-pairing as illustrated in **Figure 3.6d**. In region 1, the base-pairing is formed but stretched to fit too large a gap size. The switching conductance increases exponentially and finally saturates after passing a critical point where region 2 begins. After that, the base-pairing is completely formed and no longer depends on the size change of metal

junction. However, there is always a chance for the molecules on the two electrodes to tilt to form base-pairs as depicted in **Figure 3.6d**. In such case, the phenomenon that  $G_{(\text{switching})}$  has an approximately linear dependence of  $G_{(\text{baseline})}$  shows up in region 3. Here the sensitivity of the tunneling detection method complicates the DNA identification because a tiny change in the molecule could change the detected result. Thus, a more uniform and better characterized monolayer may be needed to improve the results.



**Figure 3.8** Distribution of “on” times for G-deoxycytidine (left histogram plot). The distribution is plotted as a function of  $G_{\text{baseline}}$  in the 2D color intensity plot on the right side.<sup>48</sup>

As interpreted above, the plateau in region 2 is considered to be the complete base-pairing so this plateau value is taken as the molecular conductance of the base-pairing. As summarized in **Table 3.1**, the

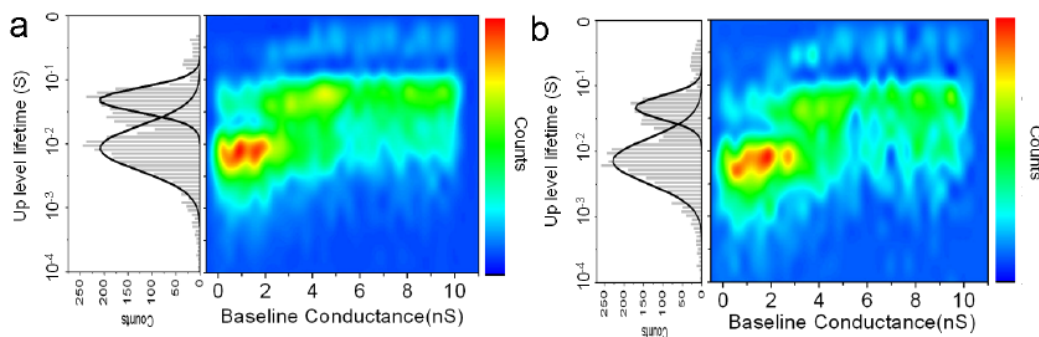
measured conductance fits with the prediction well although an approximate 1:2 ratio is noticeable but consistent with all three base-pairing cases. The trend of AA-T>A-T>G-C is also consistent between the experiment and the theory.<sup>61</sup>

Base-nucleoside pair	$G_{\text{meas}}(\text{nS})$	$G_{\text{pred}}(\text{nS})$	$\tau_f$	$\tau_s$	$h_f/h_s$
A-Thymidine	$0.76 \pm 0.01$ ( $\pm 0.03$ )	1.62	$7 \pm 3.6$	$46 \pm 11$	$1.3 \pm 0.1$
2AA-Thymidine	$0.80 \pm 0.03$ ( $\pm 0.14$ )	1.43	$8.3 \pm 3.9$	$45 \pm 11$	$1 \pm 0.1$
G-deoxycytidine	$0.66 \pm 0.007$ ( $\pm 0.07$ )	0.96	$8.4 \pm 3.4$	$47 \pm 11$	$1 \pm 0.1$

**Table 3.2** measured conductance  $G_{\text{meas}}$  compared to calculated values  $G_{\text{pred}}$ . Values are best fits to the I-V curves in Figure 3.6 with fitting errors shown. Also shown are mean lifetimes ( $\tau_f$  and  $\tau_s$ ) with associated half-widths and ratios of the peak heights,  $h_f/h_s$ , for Gaussian fits to the switching time distributions.<sup>48</sup>

Plot of the duration time for the “on” states are demonstrated in **Figure 3.8** and **3.9** where a logarithmic histograms bin scales are used. The time distribution is also bimodal. The Gaussian fitted peak indicates a fast peak at 8 ms and the slow peak is located at 50 ms. The fitting parameters are listed in **Table 3.2**.



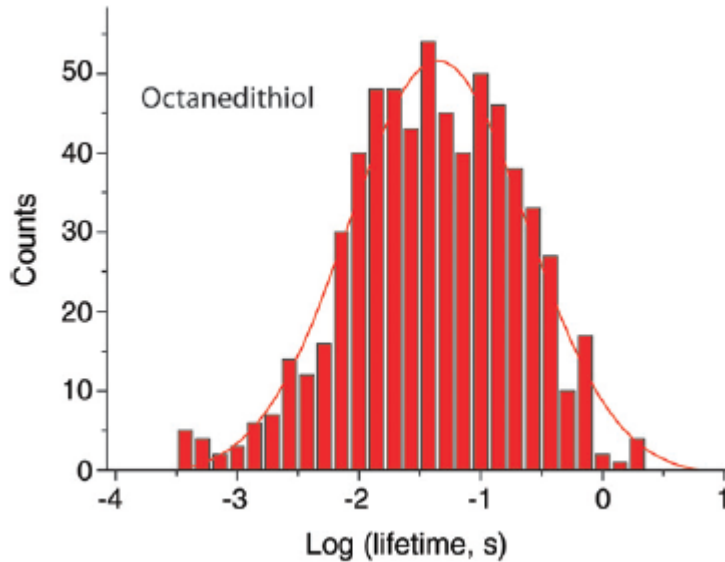


**Figure 3.9** Switching time distributions for (a) 2AA on thymidine and (b) A on thymidine.<sup>48</sup>

Compared with the thiol-/hydrogen bonding system, a control experiment with 1,8-octanedithiol is measured which has thiol- bonds only in the system. The logarithmic time distribution of the “on” time shows only one peak at around 50 ms too (see **Figure 3.10**). This suggests that the similar peak in the lifetime histogram in the thiol-/hydrogen bond- system is coming from the bounding/breaking of the gold-thiol bonding. The other peak comes from the contribution of the hydrogen bond.

From the 2D color coded intensity plots (see **Figure 3.8** and **3.9**) the fast process diminishes rapidly with increasing baseline conductance and the slower process dominates in the higher baseline region. This transition begins at about the same region  $G_{(\text{baseline})}=1\text{ nS}$ , where the molecular conductance change from region 1 to 2 and 3. This result is also indicating another fact that the hydrogen bonded system is more stable in the less stretched or compressed situation and it also approves that the  $I(t)$  method

shows advantage to the break junction method for hydrogen bonded system like DNA base-pairing.



**Figure 3.10** Distribution of switching times for telegraph noise measured in 1,8-octanedithiol. Only one peak in the log scaled time histograms while 2 peaks could be resolved in the thiol- and hydrogen bonded- systems.<sup>48</sup>

### 3.3.3 Conclusion

I(t) method is selected to perform the DNA base on nucleoside measurement. Absolute values of the conductance from different DNA base-pairing (A-T, AA-T, G-C) are determined. The experimentally determined conductance of DNA base-pairs is not single valued but has a strong dependence on the size of the gap. A 3 region model is depicted to interpret the phenomenon. The plateau region is considered to be the representative value for the base-pairing conductance and fits well with the theoretical prediction. The consistency of the measured value and the

theoretical prediction explained and supported the nature of hydrogen bonding mediated tunneling and is potentially a new pathway for next generation DNA sequencing via electrical detection.

For the base on nucleoside case, the molecular conductance is in general small and insensitive to the number of hydrogen bonds among different molecular base-pairing. The deoxyribose sugar ring is believed to be an insulator so by avoiding the tunneling through the ribose sugar may help to solve the problem as illustrated in the coming Chapter.

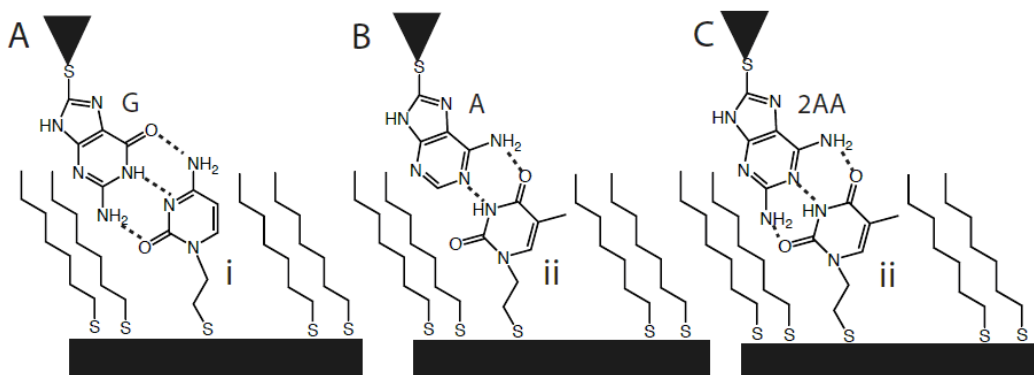
## 4. THE CONDUCTANCE MEASUREMENT OF DNA BASE ON BASE

### 4.1 Introduction

Based on the results in the previous chapter, the  $I(t)$  method is tested to be a powerful tool for tunneling signature detection of different DNA base-pairings. However, the measured conductance difference is diminished due to the insulating deoxyribose sugar ring.<sup>61</sup> It is proposed that the direct tunneling through Watson-Crick base-pairing will increase the molecular conductance dramatically and enhance the hydrogen bonding effect for the purpose of molecule identification and even single molecule DNA sequencing. However, the strong interaction of bases with gold surfaces<sup>107, 108</sup> has complicated the approach in the beginning since even more random results than the base-nucleoside case are encountered and the conductance of base pairing could not be correctly determined. An innovative AC modulation technique is developed to eliminate non-specific bonding signals. A well characterized monolayer is detected by STM with molecular resolution when making a monolayer of two molecules (1-octanethiol and 1-(2-mercaptoethyl) thymine) sequentially. Combining these two techniques, the downstream  $I(t)$  measurement is much improved and plateaus of molecular conductance show up. New features of high molecular conductance have been detected for 2AA-T and G-C base pairing and fit with the theoretically predicted order of conductance: 2AA-T>G-C>>A-T.

## 4.2 Materials and Methods

The schematic picture of the base on base experiment is demonstrated in **Figure 4.1** with the desired geometry of the base-pairing. 8-Mercaptoguanine (G in **Figure 4.1A**), 8-Mercptoadenine (A in **Figure 4.1B**) and 2-amino-8-mercaptoadenine (2AA in **Figure 4.1C**) were used as received from Aldrich. 1-(2-Thioacetate ethyl) cytosine (TE-C) and 1-(2-thioacetate ethyl) thymine (TE-T) were synthesized in the lab.<sup>88</sup>

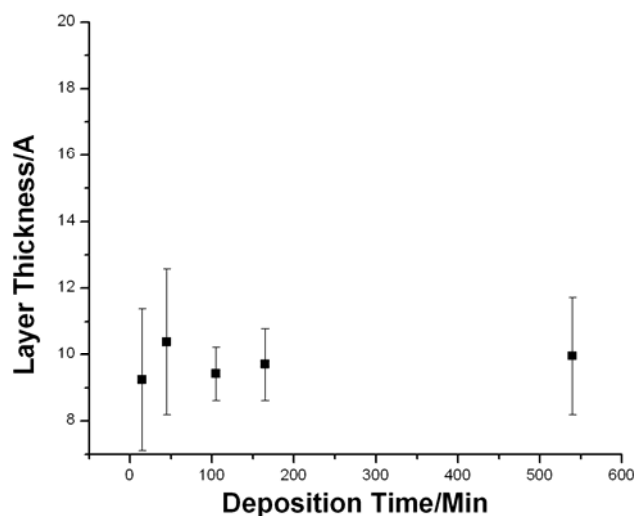


**Figure 4.1** The Schematic Diagram of The Base on Base Experiment embedded in the monolayer of 1'-Octanethiol. (A) guanine-cytosine, (B) adenine-thymine, and (C) 2-aminoadenine-thymine.<sup>88</sup>

Thiolated DNA bases (A, T, C, G and 2AA) are functionalized on the gold substrate or the gold tip simply by immersing the gold material into the ethanol solution of the thiolated bases. The physical properties (Molecular conformation and layer thickness) of the self assembled monolayer (SAM) are controlled by the solvent concentration and deposition time.

Ellipsometry (see **Figure 4.2**), FTIR (see **Figure 4.3, 4.4**) and STM (see **Figure 4.5**) measurements are done to characterize the SAM.

To eliminate any organic contaminations on the gold surface, hydrogen flame annealing of the gold surface is done prior to the functionalization. Monolayers of TE-C and TE-T on Au(111) were prepared by immersing the gold substrate in 1mM dimethyl formamide (DMF) solution of TE-C and TE-T with 0.2 ml of pyrrolidine for 2 hours. The added pyrrolidine eliminates the thio protection group which triggers the functionalization process. The chip is then rinsed with DMF and air dried with compressed nitrogen. Saturating thickness measured by the ellipsometry showed evidence of the formation of a monolayer.



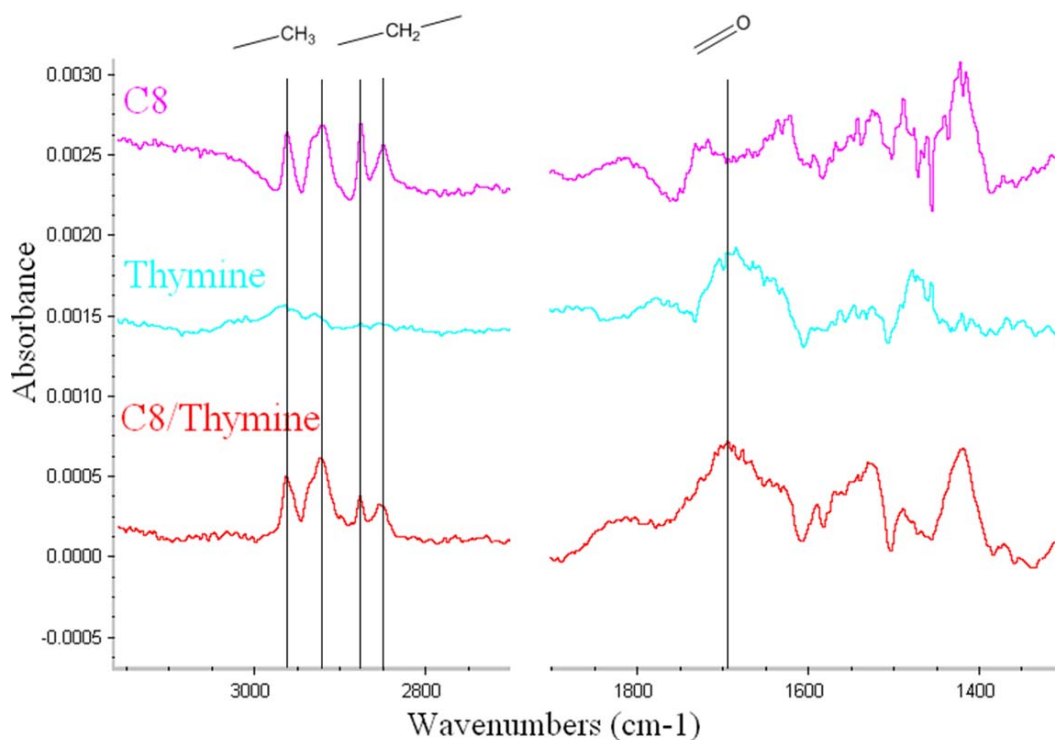
**Figure 4.2** The thymine monolayer thickness on Au(111) surface as a function of deposition time measured by ellipsometry. The layer thickness saturates after 20 minutes of deposition.<sup>88</sup>

The gold probe is customarily made in the lab as described in the **Appendix A**. To functionalize the STM probe, the tip heads are immersed in 1mM DMF solution of A, 2AA and G for at least 2 hours.<sup>89, 91, 106</sup> The

functionalized probe are rinsed sequentially with DMF, DCM and blown-dried with compressed nitrogen gas before use.

FTIR measurement is done with SAGA mode. The corresponding spectrum is illustrated in **Figure 4.3, 4.4**. Vibration modes from carbon chains are easily visible in the spectrum. In the FTIR spectrum, the peak with the highest wavenumber corresponds for the CH<sub>3</sub> group and the other 3 neighboring peaks correspond to various modes from the CH<sub>2</sub> group. C=O group is the only chemical group that gives distinguishable signal in the spectrum when modified as a monolayer.

To prepare mixed SAMs, Au(111) films were immersed in 5mM toluene solution of octanethiol (Sigma Aldrich) for 24 hours, forming a compact and well ordered monolayer.<sup>109</sup> An STM image of pure octanethiol SAM on gold is shown in **Figure 4.5A** where 2 phases of octanethiol are obviously seen in the STM image. The stripe phase in the top left corner corresponds to SAMs with octanethiol lying down on the surface. The compact phase in the top right corner corresponds to the upright geometry of the molecules. Extended immersion time normally results in larger areas of upright phase since higher density of molecules are bonded to the surface.<sup>55</sup>



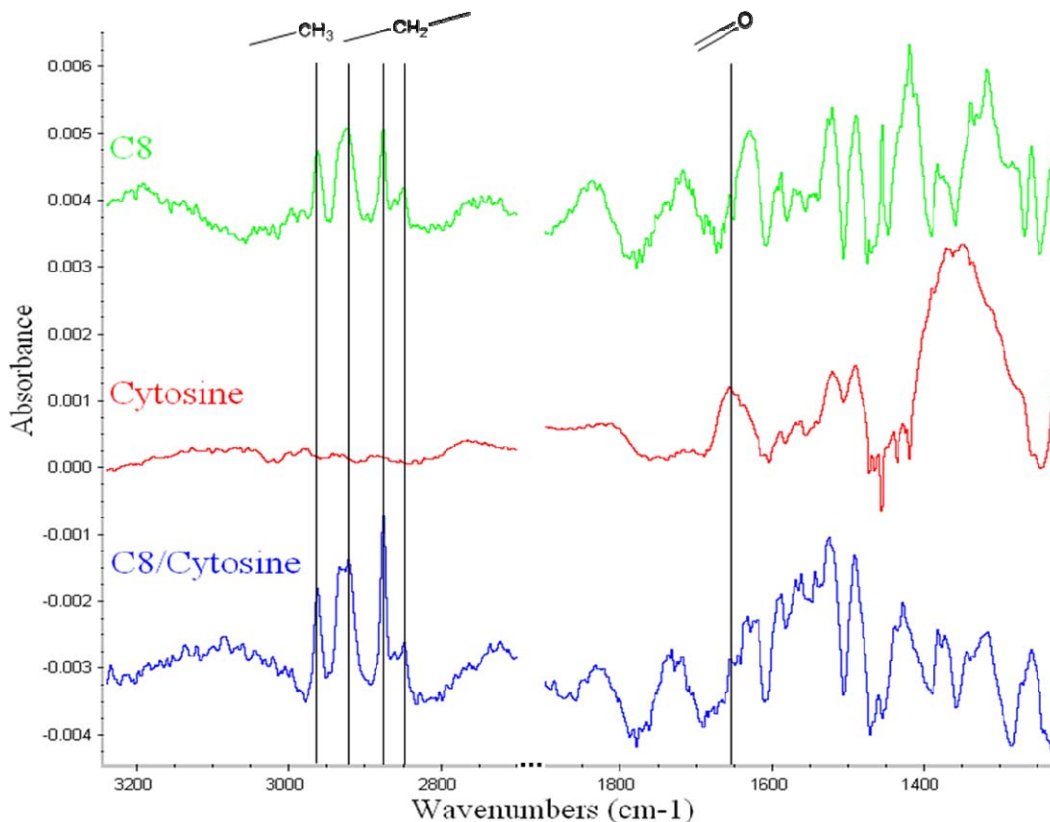
**Figure 4.3** FTIR absorbance spectra for octanethiol (C8), mercaptothymine (Thymine) and mixed (C8/Thymine) SAMs<sup>88</sup>

Such a gold chip with large areas of compact phase is an ideal platform for DNA bases insertion because it restrains the geometry of the DNA bases being embedded. After rinsing this chip with toluene and DMF, another 2 hours immersion of the chip in the 1mM DMF solution of TE-T and TE-C is carried on for base insertion. The film is then rinsed with DMF and toluene and kept in clean toluene for another 30 min for the mixed monolayer to be annealed. This critical process guarantees the formation of mixed SAM since blurry STM images always show up when this step is skipped. My interpretation is that the insertion process is random so the inserted DNA bases spread over the surface. The annealing process



helps the molecules to organize and from the mixed SAM as seen in

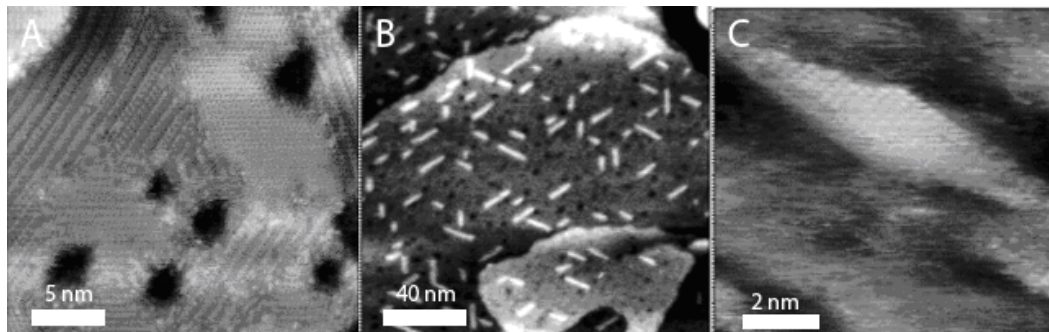
**Figure 4.5B** and **4.5C**.



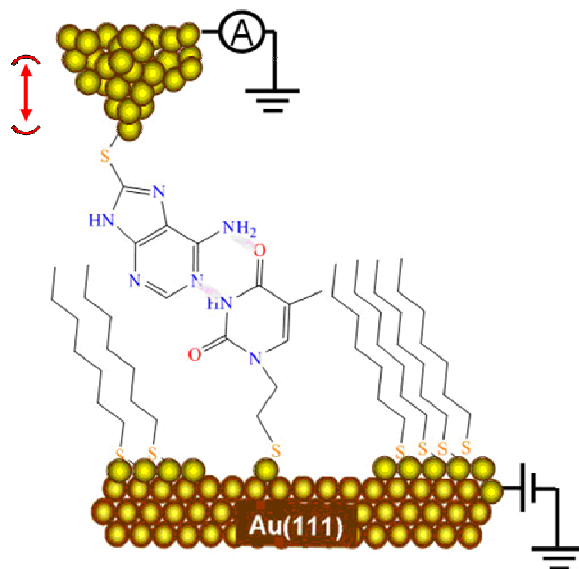
**Figure 4.4** FTIR absorbance spectra for octanethiol (C8), mercaptocytosine (Cytosine) and mixed (C8/Cytosine) SAMs.<sup>88</sup>

In **Figure 4.5B**, the mixed SAM of octanethiol and TE-T shows a remarkable structure in STM images. Bright striped phases were formed (not seen or reported in alkanethiol monolayers alone). Zooming in on them (**Figure 4.5C**), the molecular resolution image suggests rows of stacked thymines. The spacing between adjacent rows was about 0.35nm and the spacing between the disklike features in the rows was about

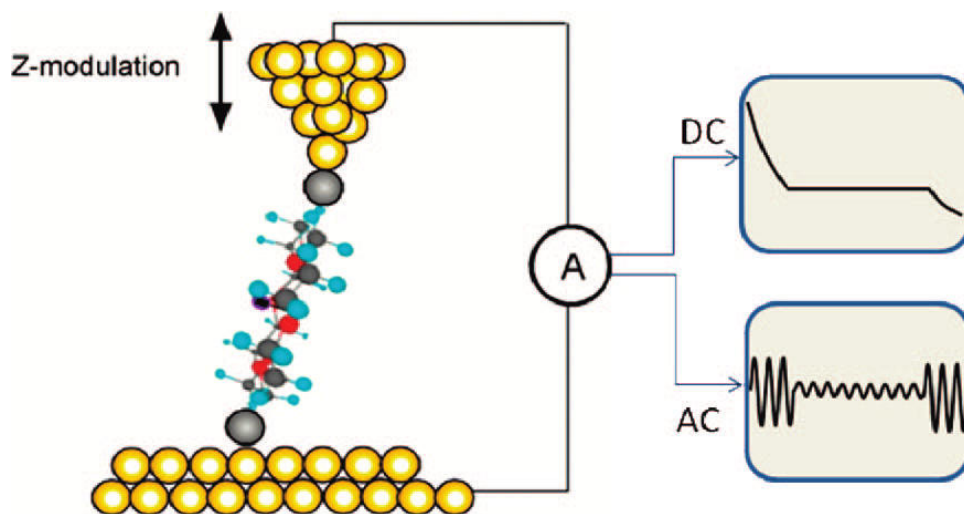
0.25nm. We could not resolve similar features in mixed SAMs of octanethiol and TE-C.



**Figure 4.5** High Resolution STM images of SAMs: (A) octanethiol SAM showing a mixture of domains of upright molecules and molecules lying flat on Au(111); (B) after insertion of 1-(2-thioacetate ethyl) thymine, showing stripes of an ordered thymine phase; (C) zoom-in on one of the thymine stripes showing the molecular-scale structure. Images were obtained in 1,2,4-trichlorobenzene at  $-0.1$  V (tip bias) and 10pA using Pt-IT tips.<sup>88</sup>



**Figure 4.6** Schematics of the AC Modulated Break Junction Measurement. An AC voltage is applied to the z-piezoelectrictransducer of the STM to modulate the tip-substrate distance, and the current is separated into an AC and a DC component. The AC component of the current provides an indicator for one to determine if a molecule is bridged between the tip and the substrate, and the DC component measures the transport current through the molecule.<sup>88</sup>

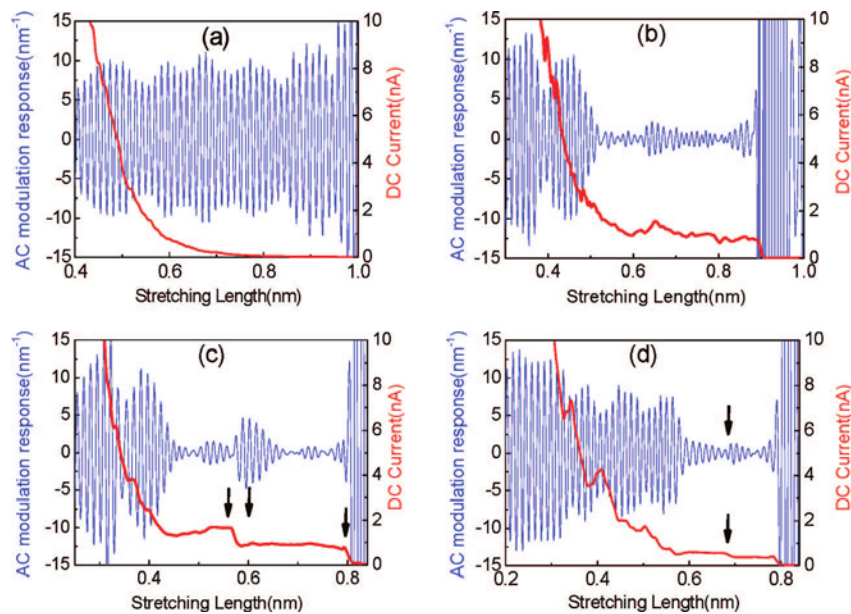


**Figure 4.7** DC and AC components of current traces recorded during the pulling away stage: In (a) pure mesitylene and in (b-d) mesitylene containing 0.2 mM 1,8'-octanedithiol. The AC component is normalized by the DC component and by the distance modulation amplitude, so its amplitude measures  $R$ . The applied tip bias is 0.1 V; the pulling speed is 27 nm/s, and the modulation frequency and amplitude are 2 kHz and 0.05 nm, respectively.<sup>79</sup>

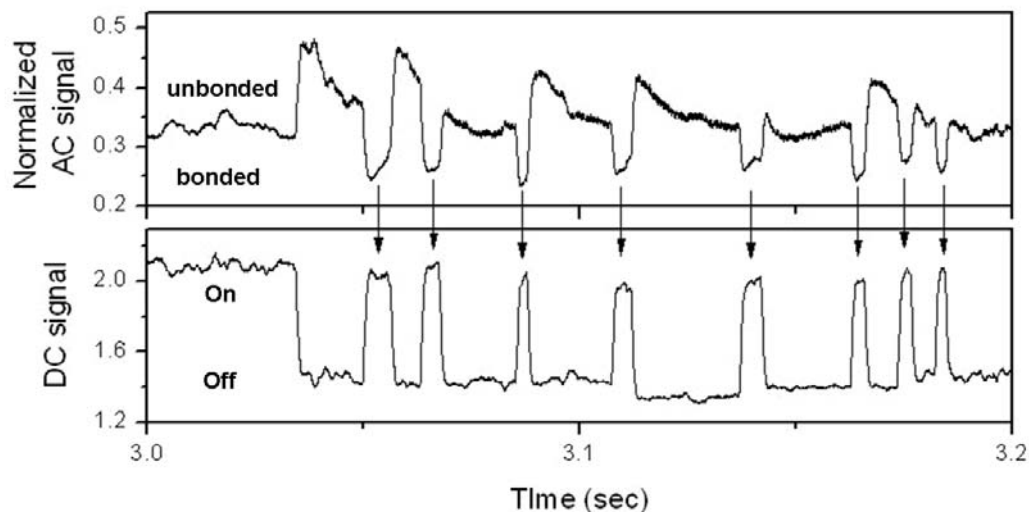
A second innovation in these series of experiments is an AC modulation technique first reported by Tao et al<sup>79</sup> in a break junction measurement (see **Figure 4.6** and **4.7**). Adapted to the  $I(t)$  method in our case, a small modulation is applied to the Z-PZT, resulting in a corresponding ac component in the tunnel current. The magnitude of this response is related both to the stiffness and the change of the electronic decay constant  $\beta$  of the bonded junction. A molecule spanning the entire junction will diminish the modulation relative to the free junction because the molecule will increase the stiffness of the gap significantly. Besides that, the AC signal fall with smaller  $\beta$  so we could visualize the bonding process (refer to **Equation 4.1-4.9**). In the case of break junction measurements, steps in the plot of dc current vs gap extension can be associated with trapped

molecules if the normalized ac modulation is reduced when the molecule is bound.

In the  $I(t)$  method, the expected anti-correlation is also detected between the magnitude of the dc and the normalized ac signal. When the molecule trapping happens, the dc signal jumps up to the upper level while the normalized ac signal drops (see **Figure 4.9**). Most of the detected signal show such anti-correlation which is considered to be the binding signal and goes to the statistical analysis. A dc jump with no change in the normalized ac signal is considered to be a non-specific binding and is normally filtered out of the analysis.



**Figure 4.8** Single Molecular Base-Pairing Conductance Measurement by  $I(t)$  Method with AC Modulation.<sup>79</sup>



**Figure 4.9** Sample Time Traces of Single Molecule Conductance Measurement with AC Modulation. Showing how the AC modulation signals (Top trace) and the DC current (Bottom trace) are anti-correlated, evidence of molecular binding and unbinding in the tunnel gap. The on state in the DC current is always in response to the drop in the normalized AC signal.<sup>88</sup>

During the measurement, A picoSTM (Agilent, Chandler) scanning tunneling microscope is used for both tunneling measurements and STM imaging of the mixed SAMs. Similar to the base-nucleoside experiment, 1,2,4-trichlorobenzene (TCB) is added in the liquid cell as a hydrogen bond free environment for the measurement. To gain atomic resolution in the STM imaging, only Pt-Ir tip could be used for its stiffer texture.<sup>110</sup> The tip preparation is described in the **Appendix A**.

To do the telegraph noise measurement, the functionalized gold probe is first stabilized in close proximity to the surface for at least 30 min or until the tip drift diminishes. After approaching the probe to the surface with a

desired baseline conductance  $G_{(bl)}$ , the servo is turned off. The gap size is then controlled manually via a custom made Labview program which outputs a bias across the Z-PZT. Once the gap was stable for a period of more than 10 sec with telegraph signals showing up, runs of telegraph switching was recorded by a digital oscilloscope (Yokogawa DL750). Data was collected with a variety of  $G_{(bl)}$  ranging from 5nS to 70nS (corresponding to currents between 0.25 and 3.5nA with tip bias of 0.05 volt). No telegraph signal can be detected for  $G_{(bl)}$  less than 5nS.

To apply an AC vibration on the tip, a 2kHz ac signal is applied to the Z-PZT, adjusted to give a peak to peak amplitude of 0.016 nm. The frequency of 2kHz is set much higher than the average telegraph noise frequency of 150 Hz. The amplitude of the in-phase response was recorded using a lock-in (Stanford Research SR 830). Three channels of data are recorded by the oscilloscope. These signals are: 1. 400 Hz low pass filtered DC signal. 2) AC+DC signal with full bandwidth. 3) The amplitude of the AC signal captured by the lock-in amplifier with the time constant set to 3ms.

A simple model to describe the AC modulated  $I(t)$  measurement is as following:

The tunneling current depends on the gap size change exponentially:

$$I = I_0 \text{Exp}(-\beta x) \quad (4.1)$$

Assuming that the applied AC modulation is a Sin function.

$$x = x_0 + A \text{Sin}(\omega t) \quad (4.2)$$

$$I = I_0 \text{Exp}\{-\beta[x_0 + A \text{Sin}(\omega t)]\} \quad (4.3)$$

In the limit that the amplitude of the AC modulation is small enough, the first 2 terms in Taylor series becomes:

$$I = I_0 \text{Exp}(-\beta x_0) - I_0 \text{Exp}(-\beta x_0) \beta A \text{Sin}(\omega t) \quad (4.4)$$

Define:

$$I_{DC} = I_0 \text{Exp}(-\beta x_0) \quad (4.5)$$

$$I = I_{DC} - I_{DC} \beta A \text{Sin}(\omega t) \quad (4.6)$$

$$I_{AC} = -I_{DC} \beta A \text{Sin}(\omega t) \quad (4.7)$$

The magnitude [I] is:

$$[I] = [I_{DC}] - \beta A [I_{DC}] \quad (4.8)$$

Dividing [I<sub>DC</sub>] on both sides gives the expression of the normalized current and its DC and AC component.

$$[I]_{NORM} = 1 - \beta A = [I]_{NORM,DC} + [I]_{NORM,AC} \quad (4.9)$$

The normalized DC component is always 1. The normalized AC component is  $-\beta A$  and thus it is sensitive to the change of the  $\beta$  value. By reading the normalized AC current, we can easily tell whether a bonding is formed or not since the bond junction normally has much smaller  $\beta$  than through space tunneling. (see **Figure 4.9**)

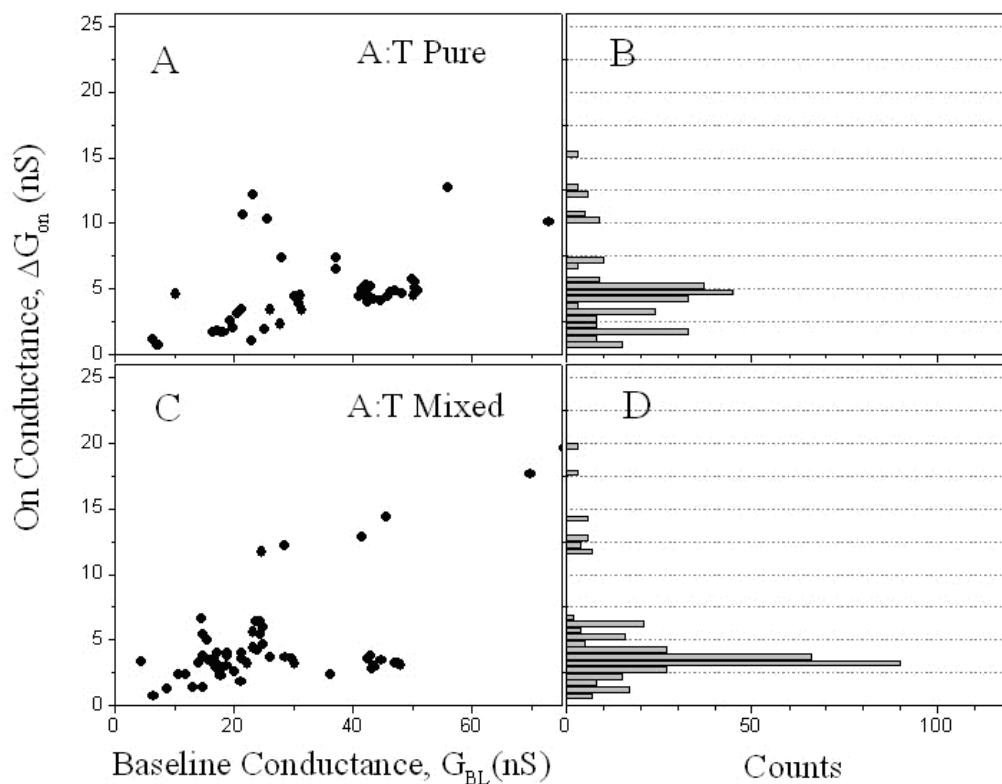
To reduce molecular/surface interaction and force the target molecule to stand up, I tried mixed functionalization with 1-Octanethiol and thiolated thymine. Evidence that the thiolated thymine forms very well ordered line structures with the support of neighboring carbon chains was observed by scanning tunneling microscopy (see **Figure 4.5B**).

Images from sharp tips could occasionally resolve structures with single molecule resolution (see **Figure 4.5C**).

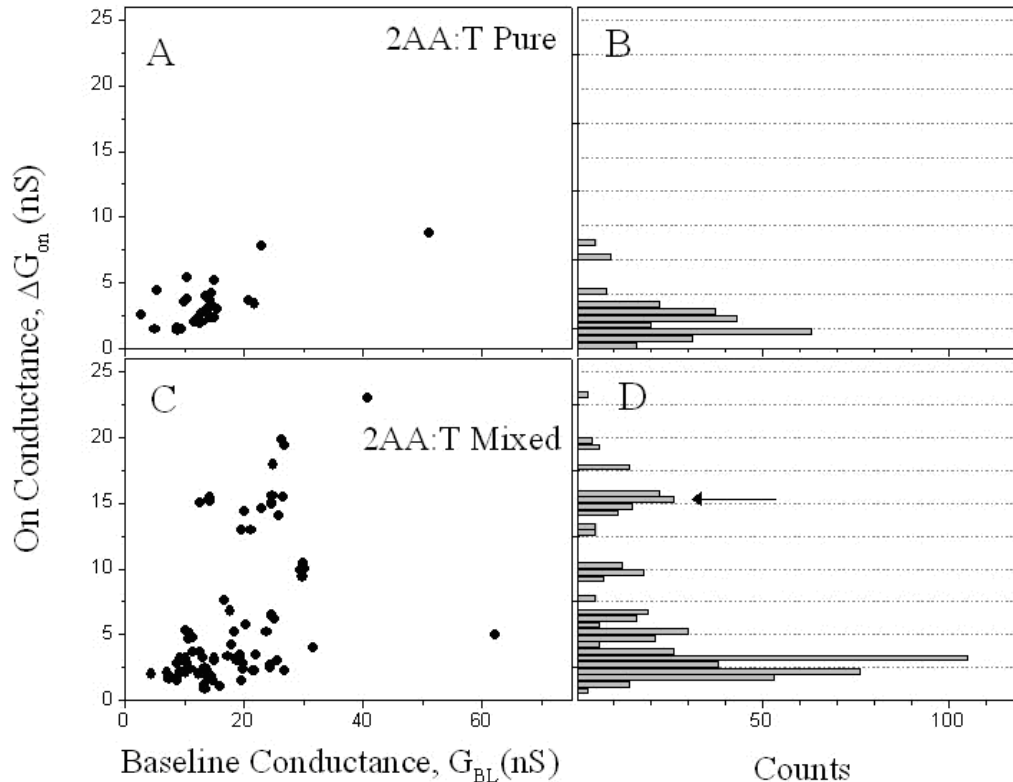
### **4.3 Results and Discussion**

To show the  $G_{(bl)}$  dependence of molecular conductance  $\Delta G_{on}$ , a scatter plot is drawn for  $\Delta G_{on}$  vs.  $G_{(bl)}$  for A-T, AA-T and G-C base-pairing. In each case, data for the pure films are shown in the top (see **Figure 4.10A**) and data for the mixed SAM are shown in the bottom (see **Figure 4.10C**). The corresponding histograms of the molecular conductance are shown on the right (see **Figure 4.10B,D**).





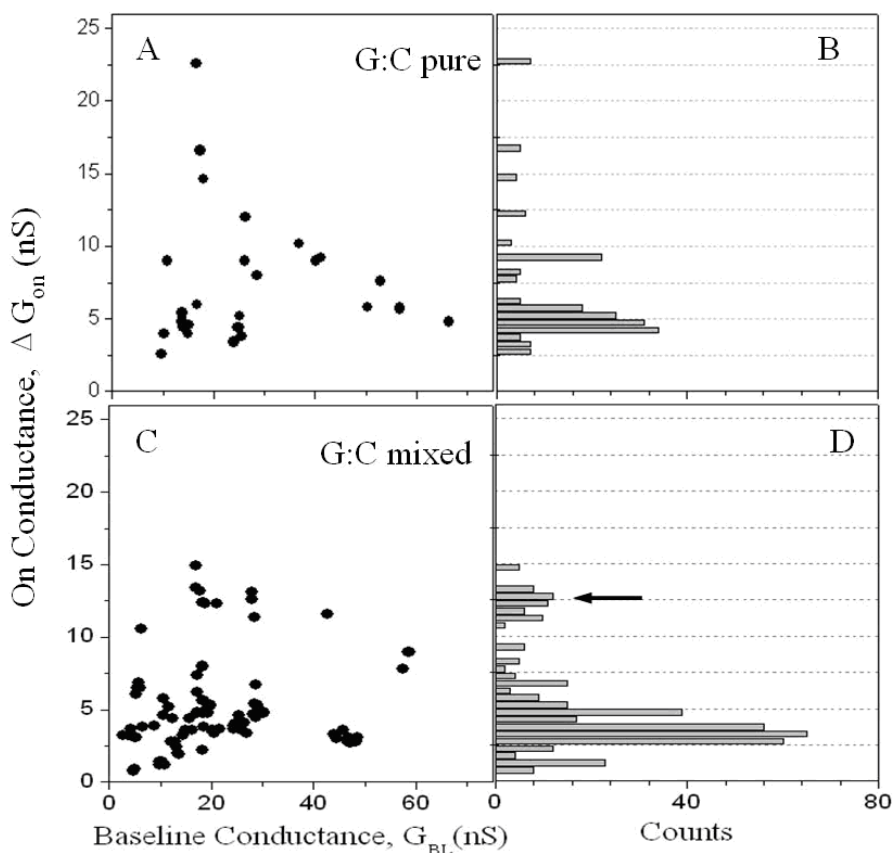
**Figure 4.10** Scatter Plots of Molecular Conductance,  $G_m$ , vs the Baseline Conductance,  $G_{baseline}$ , for (A) adenine interacting with thymine monolayers and (C) adenine interacting with thymine inserted into octanethiol monolayers. Histograms of the molecular conductance are shown to the right (B, pure SAM; D, mixed SAM)<sup>88</sup>



**Figure 4.11** Scatter Plots of molecular conductance,  $G_m$ , vs the baseline tunnel conductance,  $G_{\text{baseline}}$ , for (A) 2-aminoadenine interacting with thymine monolayers and (C) 2-aminoadenine interacting with thymine inserted into octanecethiol monolayers. Histograms of the molecular conductance are shown to the right (B, pure SAM; D, mixed SAM). The arrow points to a new high-conductance feature observed in the conductance distribution in the mixed SAM.<sup>88</sup>

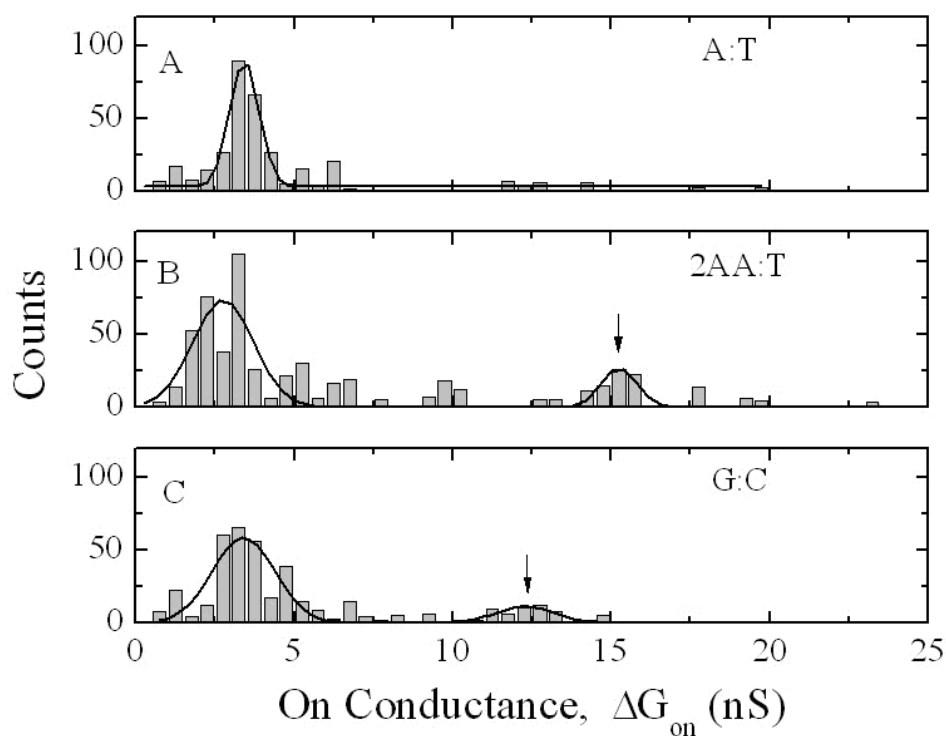
Systematically, scatter plots from the pure films a show broad spread with an increasing trend of  $\Delta G_{\text{on}}$  vs  $G_{(\text{bl})}$ . A variety of contact geometries are very likely to be the cause of that since more uniform data distribution is obtained in each corresponding mixed SAM measurement. A plateau region is more obviously seen in the mixed SAM case too which also indicates that a smaller dependence on baseline is detected and thus more a restrained molecular conformation is formed. A similar

phenomenon has been discussed in the previous chapter for the base on nucleoside cases when the deoxyribose sugar ring played a crucial role supporting the molecular conformation. Lacking the support of the sugar ring, the extra octanethiol worked as a replacement of the deoxyribose sugar ring while not diminishing the total base-pairing conductance too much.

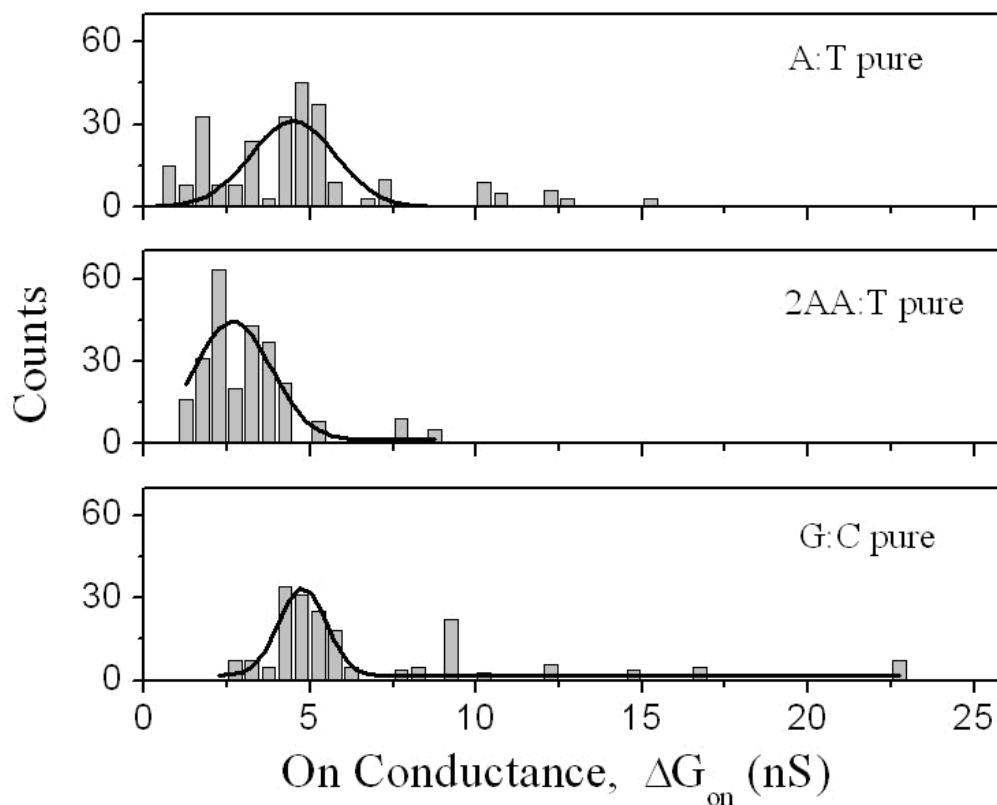


**Figure 4.12** Scatter plots of molecular conductance,  $G_m$ , vs the baseline tunnel conductance,  $G_{baseline}$ , for (A) guanine interacting with cytosine monolayers and (C) 2-aminoadenine interacting with cytosine inserted into octanethiol monolayers. Histograms of the molecular conductance are shown to the right (B, pure SAM; D, mixed SAM). The arrow points to a new high-conductance feature observed in the conductance distribution in the mixed SAM.<sup>88</sup>

By analyzing the histograms (see **Figure 4.10-4.12**), the results from pure SAMs all show one peak in the graph while the histograms of 2AA-T and G-C show an extra high conductance peak each. Gaussian fitting these peaks generates the conductance value and standard deviation (STD) listed in **Table 4.1**.



**Figure 4.13** Gaussian fits to the conductance distributions for the mixed pyrimidine SAMs: (A) A-T, (B) 2AA-T, and (C) G-C. The arrows indicate the high conductance peaks.<sup>88</sup>



**Figure 4.14** Conductance histograms for the pure pyrimidine SAMs. The solid lines are Gaussian fits to the conductance distribution.<sup>88</sup>

Theoretical values for base-pairs, including 2AA-T, have been calculated by Lee and Sankey.<sup>61</sup> These calculations were carried out for pyrimidines attached to gold electrodes directly with no chemical linker connected and the corresponding values are shown in **Table 4.1**. Note the large differences for the three molecular pairs and the large predicted differences between conductance for two and three hydrogen-bonded connections (a factor of 4 and 5). This is in sharp contrast to values calculated for base-nucleoside pairs where the differences among different base-pairs are diminished by the insulating sugar ring.

To compare the theoretical value to the experiment cases, we compensated for the added ethylene linkage. We expect that the conductances of the ethylene-tethered bases to be lowered by a factor of  $e^{-2}$  because the electronic decay constant  $\beta$ , for alkanes tethered to gold via a thiol linkage is 1 per methylene. The approximately corrected values are shown in **Table 4.1** too.

The experimental values for pure SAMs show no significant differences in conductance among 3 base-pairs. However, in the mixed SAMs case, it shows obvious evidence of the difference in molecular conductance with different base-pairs. The 2<sup>nd</sup> plateau at a higher conductance value for the two triply hydrogen-bonded complexes fit exactly to the conductance predicted by the theoretical calculation.

	A on T	AA on T	G on C
<b>Pure SAM</b> Conductance (nS)	4.49±1.28	2.67±1.16	4.77±0.71
<b>Mixed SAM</b> Peak 1 Conductance (nS)	3.42±0.46	2.75±0.97	3.41±1.00
Peak 2 Conductance (nS)		15.24±0.57	12.30±0.94
Theoretical value (nS)*	3.0 (22.2)	16.1 (119)	11.3 (83.8)
Average lifetime (ms)	6.30±2.64	6.65±3.61	7.60±2.73

**Table 4.1** A summary of the measured single molecular conductance compared with theoretical predicted values. Average lifetime is also indicating the bond strength difference among the 3 types of hydrogen bonding.<sup>88</sup> \* The values in the parenthesis are calculated for pyrimidines with thiol groups directly bonded to the heterocycle

Since more hydrogen bonding will in general increase the tunneling conductance because the HOMO-LUMO gap is smaller. That is why I

chose the highest peak for the three hydrogen bond system (2AA-T and G-C) as the conductance value since the lower peak may come from incorrect bonding.

I also summarized the average bond lifetime (the duration time for the on state  $T_{on}$ ) and showed them in **Table 4.1**. The order of bonding strength is G-C>2AA-T>A-T base-pairs which is expected because triply hydrogen bonded state should be stronger and persist longer than the double bonded one.

The position and width of the A-T basepair have no significant differences between pure and mixed SAMs.

#### **4.4 Conclusion**

The theoretical predicted order of base-pair conductance of A-T, AA-T and G-C are not observed in the  $I(t)$  measurement of molecular conductance in the pure film case. The innovative technique of AC modulation is applied to eliminate non-specific signals. New features of molecular conductance arise when doing  $I(t)$  measurement to the mixed SAM instead of pure film. Presumably, these new high conductance values come from correct hydrogen bonding and fit exactly to the theoretical prediction. Several other conclusions are drawn from this experiment: 1, Molecular

conformation could be restrained by the support from neighboring molecules in a mixed SAM. 2, Molecular conformation will induce its conductance change. 3, Pi-Pi stacking might be the driving force for the 1-D structure formation.



## **5. THE RECOGNITION TUNNELING MEASUREMENT IN ORGANIC SOLVENT BY A UNIVERSAL READER**

### **5.1 Introduction**

As demonstrated in the previous 2 chapters, hydrogen bonding mediated tunneling is a great candidate for single molecular recognition. To sequence DNA with this technique, a universal reader which could identify all 4 DNA bases by the conductance difference would be needed. 4-mercaptobenzoic acid immediately came into our mind for its simplicity and the advantage that it has one hydrogen bond donor and acceptor at the same time. Nucleosides diffusing through a 2nm electronic tunneling junction generate current spikes with sub-milliseconds duration time. Statistics of spike height show broad distribution. However, by functionalizing two electrodes with the DNA readers, the distribution is narrowed down by 10-fold allowing different nucleosides to be recognized. The magnitude of the molecular conductance follows the order of deoxyadenosine (dA) > deoxycytidine (dC) > deoxyguanosine (dG) > thymidine (dT), in agreement with theoretical prediction.

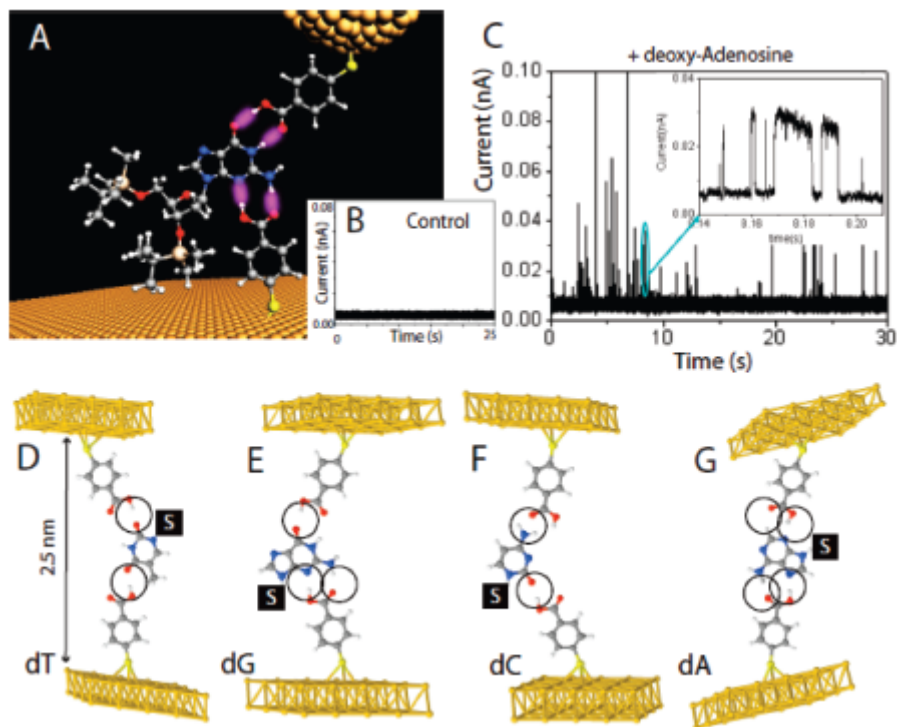
### **5.2 Materials and Methods**

The universal DNA reader: 4-mercaptobenzoic acid (4MBA) is used as received (Sigma Aldrich). To reduce complications from hydrogen bonding

and to lower electrochemical leakage current, experiments with 4MBA are all operated in (1,2,4-trichlorobenzene)TCB environment. Target DNA nucleosides are all dissolved in TCB solution with all O-H groups on the sugar ring protected with tertbutyldimethylsilyl (TBDMS) to improve their solubility. The concentrations of nucleosides used are controlled so that the detection rates are about the same.

The gold probes are prepared by the procedures explained in **Appendix B**. To functionalize the probes and substrates (Agilent, Chandler) with 4MBA, the tip ends and the substrates are immersed in 1mM methanol solution of 4MBA for at least 2 hours. For better SAM formation, 24 hours immersion of the gold substrates is recommended. Both the functionalized tips and substrates are then rinsed with methanol and air-dried by liquid nitrogen.

In TCB, which is a nonpolar organic solvent, the 4MBA is electrical neutral with the O-H group acting as a hydrogen bond donor and the carbonyl group acting as a hydrogen bond acceptor. Assuming the gap size is 2.5nm, which is determined by the experiment below, the most energy favorable hydrogen bond configuration for the trapped nucleosides are shown in **Figure 5.1 D-G**.



**Figure 5.1** Tunneling Measurements with functionalized electrodes. (A) A gold probe and a gold substrate are functionalized with a monolayer of 4-mercaptobenzoic acid, and (B) the size of the gap between two electrodes maintained under servo control at a value such that the two monolayers do not interact with one another, resulting in a tunnel current signal that is free of spikes. (C) When a solution of nucleosides is introduced, current spikes appear, as shown here for 0.7  $\mu\text{M}$  deoxyadenine in trichlorobenzene with a baseline tunneling current of 6 pA at a bias of 0.5 V. Hydrogen-bonding schemes for all four nucleosides are shown in panels (D)-(F). “S” represents the modified deoxyribose sugar and the hydrogen bonds are circled.<sup>21</sup>

The recognition tunneling is carried out on a PicoSTM (Agilent, Chandler) scanning tunneling microscope interfaced with a digital oscilloscope (Yokogawa DL 750). After functionalizing the probes and the substrates with the reader, the recognition tunneling gap as depicted in **Figure 5.1A** could be formed by engaging the tip to the surface until a setpoint current is achieved and maintained. In the control experiment where the measurement environment is plain TCB, the tunneling signal is free of

spikes when the setpoint current  $I_{bl}$  is set as high as 10 pA at 0.5 volts tip bias (see **Figure 5.1B**). I conclude that the gap size in this case is large enough that the reader molecules on the tip and the substrate don't interact and the gold to gold through space tunneling contributes to the entire  $I_{bl}$ .

To do recognition tunneling for DNA, a TCB solution of nucleosides are placed in the Teflon liquid cell just above the gold substrate. After achieving the desired  $I_{bl}$  where we didn't get signals during the control experiment, current spikes immediately show up above the tunneling background (see **Figure 5.1C**).

The spikes show up in groups and a majority of the spikes are in the form of telegraph switching signals which unambiguously show single molecule detection between binding/unbinding states.

To maintain the gap size while not interrupting the spike detection, the STM servo gain is set to be very small so that the majority group of the spikes is not affected due to a general trend of small duration time of the spikes. However, telegraph signals of long duration time may be altered slightly since it mistakenly triggered the servo to expand the gap a little to

maintain the  $I_{bl}$ . One such example is shown in the zoomed in part of

**Figure 5.1C.**

4 kinds of DNA nucleosides solution are tested with the same junction size.

Tunneling signals were collected once they were recorded by a digital

oscilloscope and analyzed by a custom Labview program. For details

about the program, please refer to **Appendix A.**

**5.3 Results and Discussions**

The average spike heights of three different runs by two users are listed in

Table 5.1. Although some overlap exists, the peak value shows an

obvious order of  $dT < dG < dC < dA$  which is also demonstrated in **Figure**

**5.3B.**

	dT	dG	dC	dA
Measured G (pS)	13.6±0.3	18.6±0.9	25.3±2.5	33±1.9
Calculated G (pS)	0.04	0.12	0.51	1.05
Read rate (s <sup>-1</sup> )	7.1±1.4	5.5±1.1	5.5±1.1	6.6±1.3

**Table 5.1** measured and calculated conductance in a functionalized tunnel junction at  $I_{bl}=6pA$ ,  $V=0.5 V$

The distribution is clearly asymmetric. By assuming that the conformation

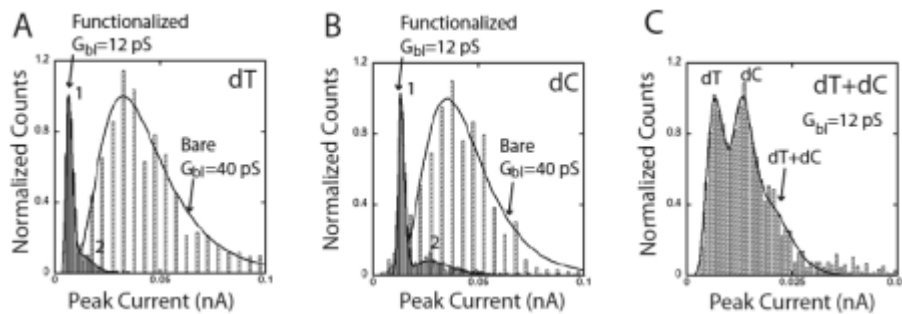
of the molecules being trapped is random and fit Gaussian distribution and

the tunneling current decays exponentially. A new fitting model is

designed and fits well with the signals as shown in **Figure 5.2, 5.4.**

$$N(i) = N_0 \exp\left(-\frac{[\ln(i) - \ln(i_0)]^2}{[\ln(w)]^2}\right) \tag{5.1}$$

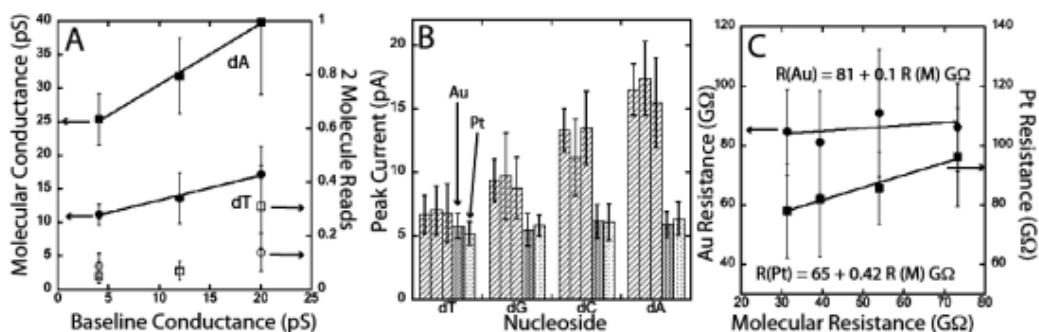
Another control experiment is to run the same procedure with bare electrodes. In order to record signals from bare electrodes, a smaller gap size is set by operating in a higher setpoint  $I_{bl}=10\text{pA}$ . As shown in **Figure 5.2**, results from the bare probes pairs give a broad distribution. By fitting with a Gaussian distribution of the logarithm of the currents, the fitted peaks ( $dA=15.9\pm 0.4\text{ pA}$  and  $dG=18.7\pm 0.2\text{ pA}$ ) differ with  $2.8\text{ pA}$  but less than the width of the distribution ( $\sim 15\text{pA}$ ) of the high current side. However, by running the experiment by a bare probe on the functionalized substrate, the distribution narrows by an order of magnitude for dT and dC as shown in **Figure 5.2A, B** and also for dA and dG as shown in **Figure 5.4 A-D**. The change in the distribution width between one and two sides functionalization is insignificant which means that the reader molecule captures the nucleoside targets and restraints its conformation. In the absence of the DNA reader, nucleosides behave much more randomly.



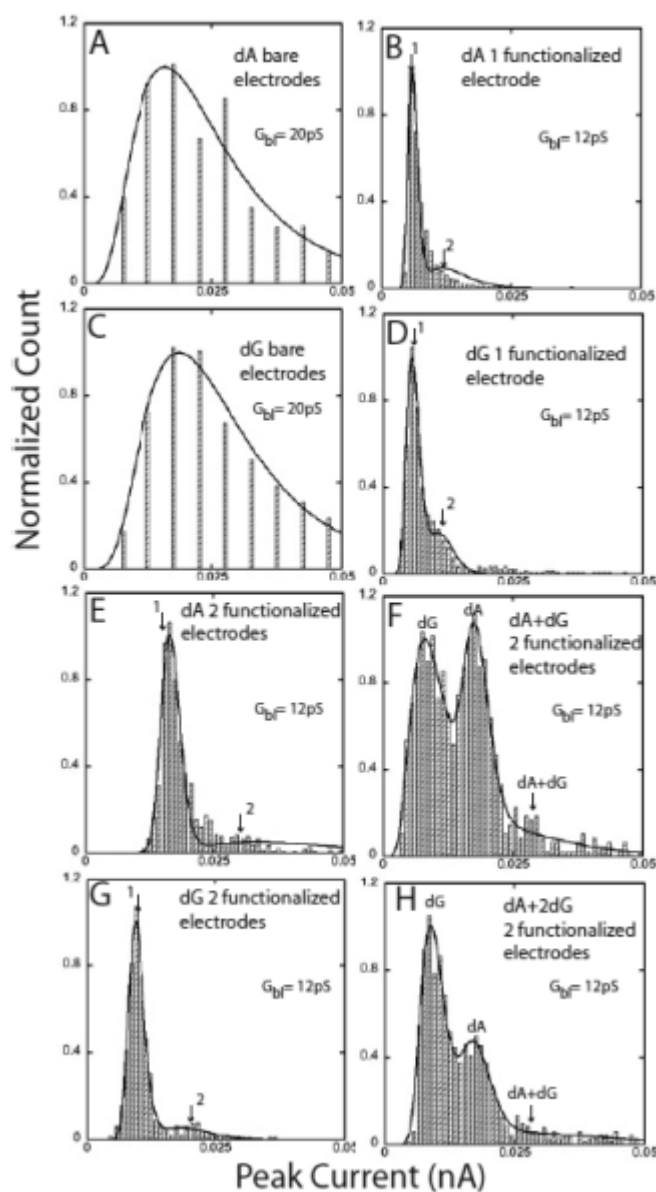
**Figure 5.2** Effect of electrode functionalization for pyrimidine reads. For bare reads (broad distributions in A and B)  $G_{bl}$  was increased to  $40\text{ pS}$  to increase the count rate. The narrow distributions in A and B are taken with both electrodes functionalized and yield  $i_0=6.7\text{ pA}$  for dT and  $13.3\text{ pA}$  for dC ( $G_{bl}=12\text{ pS}$ ,  $I_{bl}=6\text{ pA}$ ,  $V=0.5\text{ V}$ ). In a mixed solution, (C) the dT peak occurs at  $8\text{ pA}$  and the dC peak occurs at  $13.4\text{ pA}$ , an assignment verified by measuring a mixture with half the concentration of dT.<sup>21</sup>

Besides runs with bare gold probes, runs with bare Pt-Ir probes were also carried out. Summarized in **Figure 5.3**, the runs with 2 functionalized electrodes demonstrate the ability to clearly identify all 4 different DNA nucleosides while the differences don't show up in the bare probe (either gold or platinum) runs.

We conclude that the sandwich structure of the metal-reader/DNA nucleoside/reader-metal is able to sense electronic signatures of DNA nucleosides in the single molecule level since most data analyzed are of telegraphic switching shape.



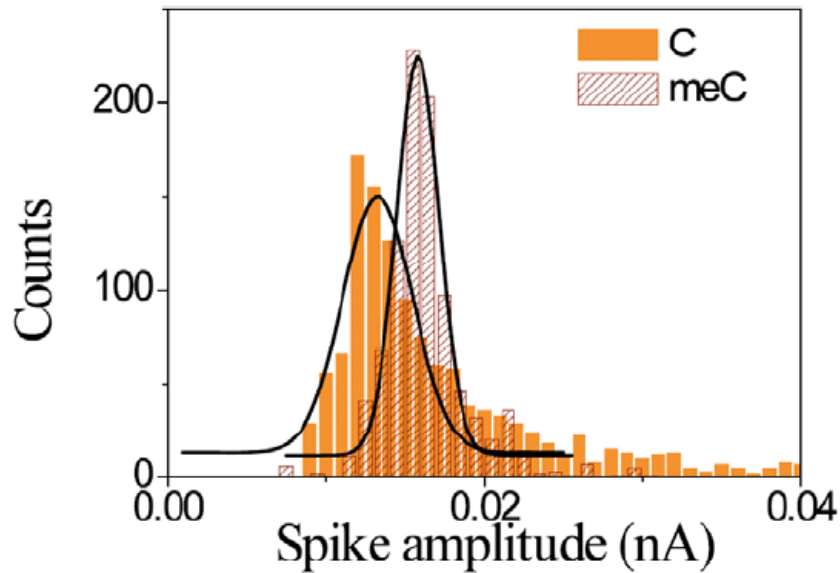
**Figure 5.3** Summary of the reads. (A) the measured molecular conductance increases linearly with Gbl (black circles dT, black squares dA, error bars are HWHH). The number of two molecule reads (open circles, dT, open squares, dA) increase at Gbl = 20 pS and the read rate is substantially reduced at Gbl = 4 pS. (B) peak currents measured in three independent runs for the four nucleosides (cross hatched bars). Error bars represent the HWHH of the current distributions. Reads for a functionalized surface and a bare Pt (light shaded bars) and bare Au (dark shaded bars) probe are relatively insensitive to the identification of the nucleoside, as shown quantitatively in (C) where the junction resistance is plotted vs the molecular resistance determined with two functionalized probes.<sup>21</sup>



**Figure 5.4** Effect of electrode functionalization on the distribution of current spikes. Bare electrodes (A, dA, C, dG) give broad distributions (gap conductance 20 pS, 0.7  $\mu$ M dA, 2.9  $\mu$ M dG in TCB). Fits are Gaussian in the log of the current. Distributions narrow 10-fold when one electrode is functionalized with 4-mercaptobenzoic acid (B, dA, D, dG) (gap conductance 5.9 pA for dA and 5.6 pA for dG). When both electrodes are functionalized (E, dA; G, dG). (F) Distribution for a mixture of dA and dG. The assignment of the higher peak to dA is confirmed by the distribution measured with a reduced concentration of dA (H). The high current tail in (F) and (H) is consistent with a small number of two molecule (dA+dG) reads.<sup>21</sup>



A better test of the fidelity of each peak is to run the experiment with mixtures of known nucleosides. To avoid the complex situation of base-pairing, mixtures of 2 purines (dA+dG) or 2 pyrimidine (dC+dT) are tested and demonstrated in **Figure 5.2C**. Two peaks are clearly resolved in both cases and a lower peak is seen in the statistics when mixing dA and dG with a 1:2 concentration ratio as shown in **Figure 5.4 F, H**. On the contrary, only one single peak shows up in the statistics of the measurement done with a single type of nucleoside. This consistency further confirmed the correct electronic signatures of each nucleoside.



**Figure 5.5** Comparison between C signal and meC signal. The distributions of the spike heights overlaps a lot for C and meC signals.<sup>21</sup>

Methylated C (<sup>m</sup>C), which is an important epigenetic marker, is also measured by the same setup. However, the spike distribution overlaps

with C almost exactly which made this technique not able to identify C and meC in organic solvent TCB (see **Figure 5.5**). Measurement in aqueous buffered solution with another type of DNA reader succeeded in separating them. The results are discussed in detail in chapter 6.

#### **5.4 Conclusion**

The difficulty of the non-specificity in DNA tunneling detection is overcome by functionalizing metal electrodes with thiolated DNA readers. The width of the current distribution narrows by 10 fold when one electrode is modified with the DNA reader. The electronic signatures of all four DNA nucleosides are further identified by functionalizing a pair of readers within a metal junction which forms a sandwich structure of metal-reader/DNA nucleoside/reader-metal. A direct and proportional relationship between nucleoside signals and concentrations confirmed these signatures. We claim that a novel method of tunneling readout of all 4 DNA nucleosides is successfully developed in organic solvent environment. The insolubility of DNA nucleotides and DNA oligomers in this environment has limited this technique and a similar experiment within aqueous buffered solution is needed for the eventual DNA sequencing purposes.

## 6. THE UNIVERSAL READER IN AQUEOUS BUFFERED SOLUTION

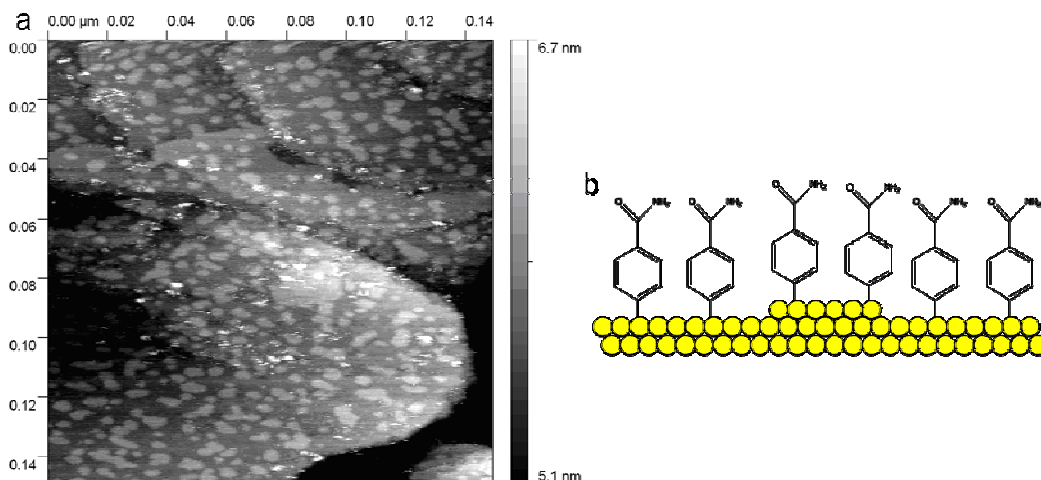
### 6.1 Introduction

The mission of signature identification of DNA nucleosides in organic solvent 1,2,4-trichlorobenzene is successfully achieved and described in the previous chapter along with published results<sup>21, 25, 87, 111</sup>. To do DNA nucleotide recognition and short oligomer sequencing, aqueous buffered solution has to be used for solubility reasons. However, striking challenges show up due to the more complicated situation in the presence of water molecules which compete with the desired hydrogen bonding. A novel DNA reader, 4-mercaptobenzamide, is designed to overcome these difficulties and shows convincing results of single base reads in short DNA oligomers. The tunneling data imply an unusually long lifetime for the recognition complex. A separate force-distance curve measurements by AFM showed that an average residence time of a single base in the junction is on the order of a second which confirmed that bursts of telegraph signals corresponds to tunneling readings of the same single base trapped in the electrical junction. This discovery suggests a solution to the problem of unreasonably fast translocation of DNA through nanopore based DNA sensors. The predicted sequencing speed could achieve tens of bases per second with a force of tens of piconewtons applied on the DNA.

## 6.2 Materials and Methods

The DNA nucleotide reader 4-mercaptobenzamide which presents two hydrogen bond donor sites (the amino group) and one hydrogen bond acceptor site (the carbonyl group) is synthesized following the published protocol.<sup>87</sup> The predicted binding modes and conformations are given in

**Figure 6.5** as optimized by computer simulations in vacuum.

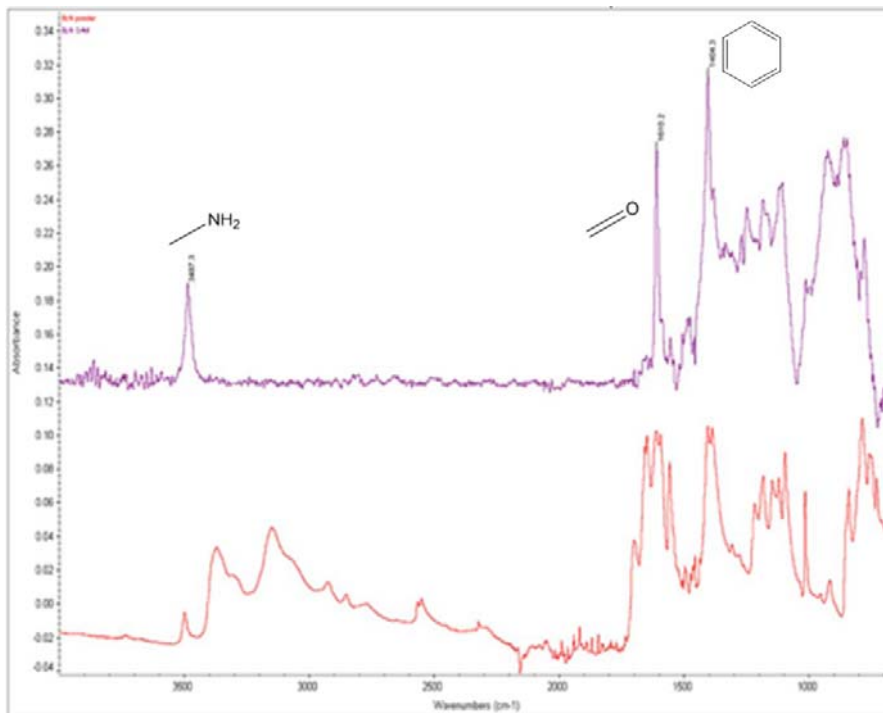


**Figure 6.1** (a) STM images of the self assembled monolayer of Benzamide DNA reader on the gold surface. Patches indicate different domains of the SAM. Image is taken with 1mM PB buffer with 0.5 V tip bias by a HDPE coated gold tip. (b) Schematic picture of 4-mercaptobenzamide molecules on the gold surface<sup>87</sup>

In the case of tunneling detection in aqueous buffered solution, the STM probe needs an extra insulating process so that only a sub micrometer sized area on the apex is exposed. This diminishes the background

electrochemical current. Details about tip insulation are explained in the

## Appendix B.

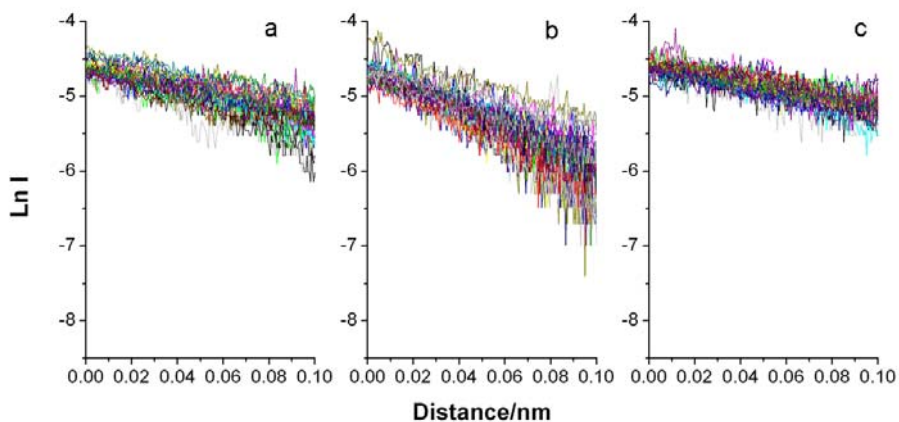


**Figure 6.2** FTIR spectrum of 4-mercaptobenzamide monolayer (Purple) and powder (red). Characteristic peaks are showing up indicating the formation of a well characterized monolayer.<sup>87</sup>

1mM methanol solution of benzamide is used for functionalizing both tips and the substrate. For compact and well ordered SAMs, 24 hours deposition is needed for the substrate. FTIR, ellipsometry and STM imaging are used for surface characterization. Ellipsometry result shows a saturated layer thickness after more than 20 minutes deposition. As shown in **Figure 6.1**, pattern of patches are detected in a typical STM image done in ddH<sub>2</sub>O which are domains of the SAMs on the Au(111) lattice. SAGA mode FTIR spectrum for powder and SAM of benzamide

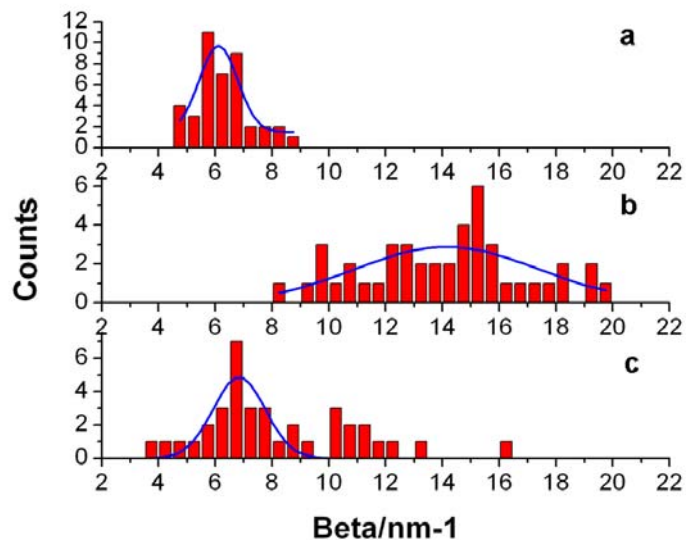
reader is showing that major functional groups (free  $\text{NH}_2^-$  and  $\text{C=O}$  groups) are detectable with sharp and strong distribution as seen as the purple curve in **Figure 6.2**. Normally that indicates the formation of a uniform monolayer otherwise the spectrum would be broadened. For all these reasons, a good monolayer is confirmed and well characterized. Tip immersion time is strictly controlled below 2 hours because an increase of high leakage current occurred at longer time.

DNA nucleotide monomers: 2'-Deoxyadenosine 5'-monophosphate (dAMP), 2'-Deoxycytidine 5'-monophosphate, (dCMP), Thymidine 5'-monophosphate(dTMP), 5-methyl-deoxycytidine 5'-monophosphate ( $\text{d}^{\text{m}}$ -CMP), 2'-Deoxyguanosine 5'-monophosphate (dGMP) are used as received (Sigma Aldrich).



**Figure 6.3** Tunnel current decay curves in pure water. (a) bare gold electrodes. (b) A pair of functionalized electrodes. (c) one bare, one functionalized electrode.<sup>87</sup>

The experimental setup is quite similar to the case described in the previous chapter with a changed measurement environment of 1mM phosphate buffer (PB, pH=7.4). We first characterize the tunneling gap formed by bare electrodes using doubly distilled water (ddH<sub>2</sub>O) and PB. Only rare signals of small heights could be detected. By using a functionalized gold probe and substrate with 4-mercaptobenzamide, clean signals could be obtained with baseline current up to 10 pA with 0.5 volts tip bias. Also a much more rapid decay in the tunneling current is observed with functionalized electrodes on both sides compared to bare electrodes or 1 side functionalized electrode situations (see **Figure 6.3**). This change in decay constant ( $\beta$ ) implies that the physical property of the gap is significantly changed in the water environment by functionalizing electrodes and serves as a test of the functionalization of the probes. The following experiments are all performed at a fixed 10 pA baseline where clean signals are detected during control experiments.

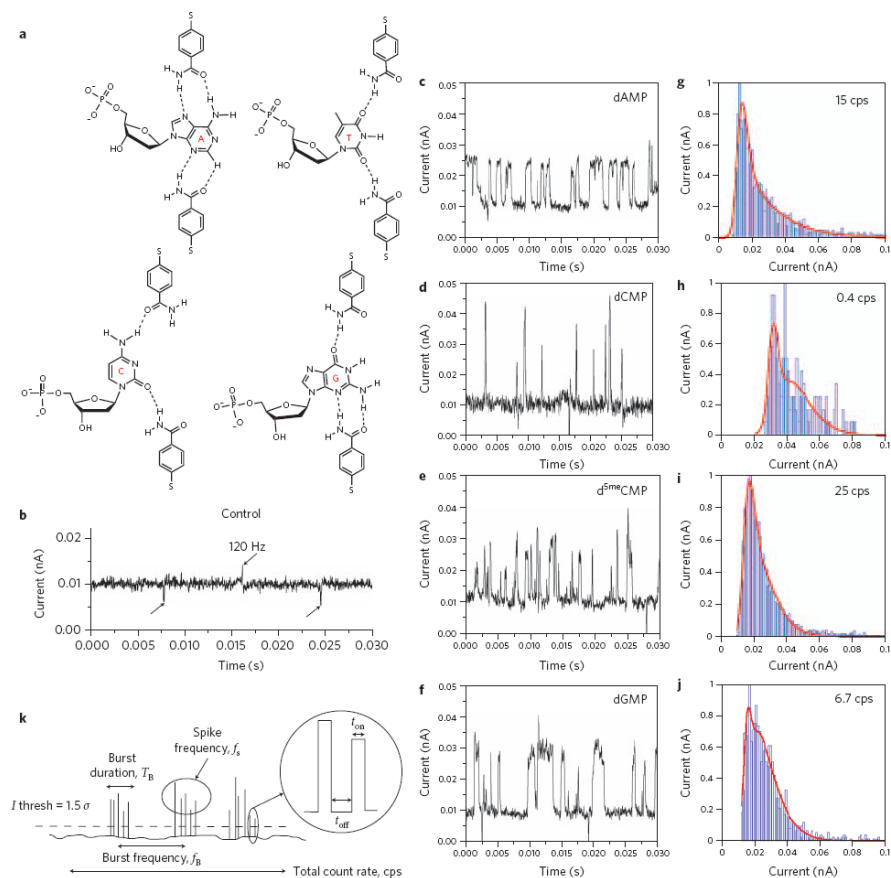


**Figure 6.4** Histograms of beta in pure water for (a) bare gold electrodes. (b) both electrodes functionalized. (c) one electrode functionalized. Gaussian Fitting yields (a)  $6.11 \pm 0.68 \text{ nm}^{-1}$  (b)  $14.16 \pm 3.20 \text{ nm}^{-1}$  (c)  $6.84 \pm 0.92 \text{ nm}^{-1}$ <sup>87</sup>

Introducing DNA nucleotides into the system immediately gives spike signals which represent the diffusion of DNA molecules into the electrode gap. The spike signals appear in groups and characteristic bursts of telegraph signals from different DNA nucleotides make DNA recognition through tunneling feasible. Representative DNA monomer signals are demonstrated in **Figure 6.5c-f** with corresponding statistic results (see **Figure 6.5 g-j**) from series of measurements on the right column. By defining parameters like: total count rates, spike frequency, burst duration,  $t_{\text{on}}$  and  $t_{\text{off}}$  in **Figure 6.5k**, parameter results are listed in **Table 6.1** for all the monomers that are tested. Besides the most probable height of the signals, burst frequency and total count rates are also quite important parameters for judging DNA nucleotides. As seen in **Figure 6.5**, dAMP



signals are small in height with an approximately equal  $t_{on}/t_{off}$  ratio. However, the signals from dCMP are significantly higher in height and much lower in  $t_{on}/t_{off}$  ratio compared to dAMP. Statistical distribution shows that such difference is systematically detectable between the two. Another interesting phenomenon is the significant difference between dCMP and d<sup>m</sup>CMP (see **Figure 1.2**) which is not observed in organic solvent as discussed in chapter 5. d<sup>m</sup>CMP gives the highest burst frequency among all the DNA monomers tested while dCMP gives the slowest switching. It is obvious to identify the two from the spike height since dCMP is the highest and d<sup>m</sup>CMP is the smallest. Thus we conclude that hydration plays an important role separating the two. The signal from dGMP is not significantly different both in height and frequency from dAMP, which made the identifications from these two purines impossible with the present reader molecule.

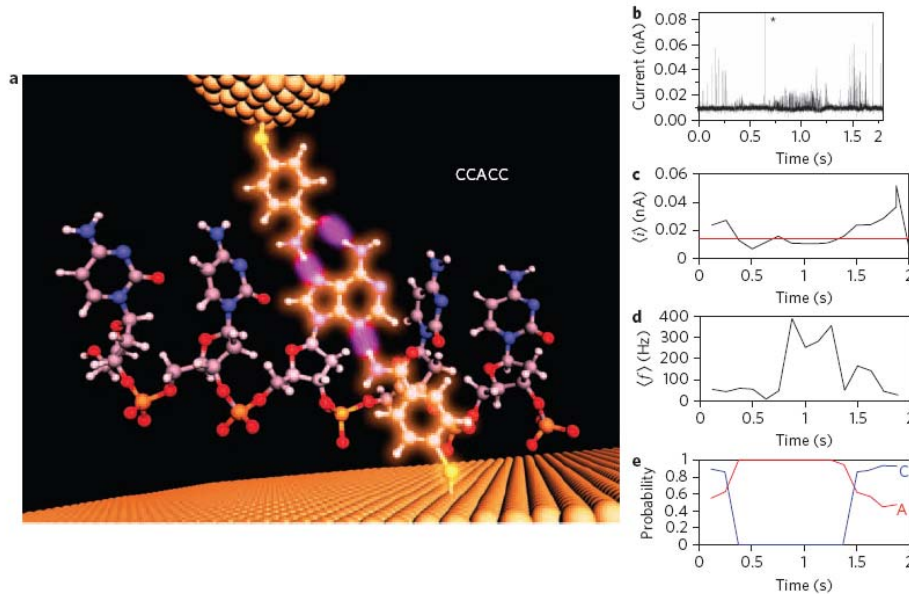


**Figure 6.5** Tunneling signals from nucleotides trapped in a functionalized tunnel gap. a, Proposed hydrogen-bonding modes for all four bases. In practice, water must play a role, because the observed difference between C and MeC would not be accounted for by these structures alone. b, In phosphate buffered saline, but in the absence of analyte, a 20 pS gap ( $i=10$  pA,  $V=+0.5$  V) gave a signal free of features, except for some a.c. coupled line noise. c-f, Characteristic current spikes produced when nucleotides dAMP, dCMP, dGMP and dTMP were introduced. k, Definitions of the parameters used to characterize the tunneling signals. Spikes are counted if they exceed a threshold equal to 1.5 sigma of the noise of the local background.<sup>87</sup>

dTMP is the most strange target among the 5 since no signals could be detected once dTMP is placed in the environment. The tunneling current in the dTMP experiment shows even lower noise levels than the control results indicate that a strong interaction between the tip and the surface. We conclude that dTMP might tend to block the functionalized tip so

strongly that it may inhibit molecule reading that requires a molecule to span the gap.

### 6.3 Results and Discussions



**Figure 6.6** Reading a single base within a heteropolymer. (a) An artistic overview of a piece of DNA oligomer trapped in the junction formed by a pair of electrodes functionalized by 4-mercaptobenzamide reader. (b) Corresponding Tunneling signals for transition between A and C reads on the sequence of 3'-CCACC-5'. (c) Slide window average of spike heights. Red line indicates the judgement criteria for A and C signals (d) Slide window average of switching frequency. (e) Probabilities for reading an A signal or a C signal during the typical time trace.

Customarily synthesized short DNA oligomers were directly ordered from IDT incorporation. 5mers of a series of combination of A, C and <sup>m</sup>C (3'-AAAAA-5', 3'-CCCCC-5', 3'-ACACA-5', 3'-CCACC-5', 3'-C(<sup>m</sup>C)C(<sup>m</sup>C)C-5', 3'-<sup>m</sup>C<sup>m</sup>C<sup>m</sup>C<sup>m</sup>C<sup>m</sup>C-5') are designed for the following experiments.

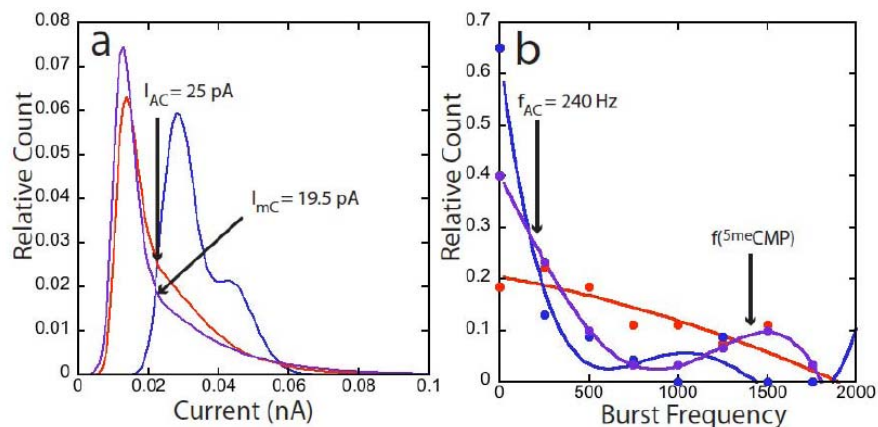
STM measurement is believed to be a highly localized measurement and signals from the neighboring bases should not contribute for average 3Å

spacing between adjacent bases. In reality, thermal drift of both the PZT element and tip materials makes the drifting a problem in highly position-sensitive measurement. Alternatively, this random drift offers an easy way to read signals from one base to another by chance and can therefore demonstrate that single base resolution by tunneling. **Figure 6.6a** shows an artistic picture of a 3'-CCACC-5' oligomer trapped in the benzamide reader junction by hydrogen bonding. **Figure 6.6b** shows a piece of real data in the measurement which corresponds drift reading between A and C signals. As discussed, the low frequency, high spike height signals correspond to C reads while the fast switching, low height signals correspond to A reads. **Figure 6.6c** shows a sliding average of the spike amplitudes. Values below the characteristic red line identify an A base unambiguously. **Figure 6.6d** shows a sliding average over the pulse frequencies. In summary, from the study of nucleotides, homopolymers and heteropolymers, the probability deduced combining these 2 parameters are shown in **Figure 6.6e** which strongly indicates that a single base can be identified with high confidence when flanked by C bases in a whole piece of short DNA oligomers.

Nucleotide	dAMP	dGMP	dCMP	d <sup>m</sup> CMP
Burst Duration $T_b$ (s)	0.19±0.05	0.13±0.02	0.12±0.02	0.06±0.01
Burst Frequency $f_b$ (Hz)	732±82	574±67	306±23	1305±100
Fraction of reads >0.1 nA	0.02	0.001	0.02	0.01
$T_{on}$ (ms)	0.38±0.01	0.48±0.02	0.42±0.02	0.31±0.09
$T_{off}$ (ms)	0.35±0.01	0.56±0.04	0.71±0.06	0.41±0.11
$T_{on}/T_{off}$	1	0.9	0.6	0.8
$\Delta G$ (kT units)	0	0.1	0.51	0.22

**Table 6.1** Tunneling signal characteristics for 4 types of single nucleotides.<sup>87</sup>

In **Figure 6.8**, a complete result is demonstrated for a series of DNA 5mers as combinations of A, C and <sup>m</sup>C elements. Typical time traces are shown on the left column with statistical distribution of pulse heights shown on the right side.



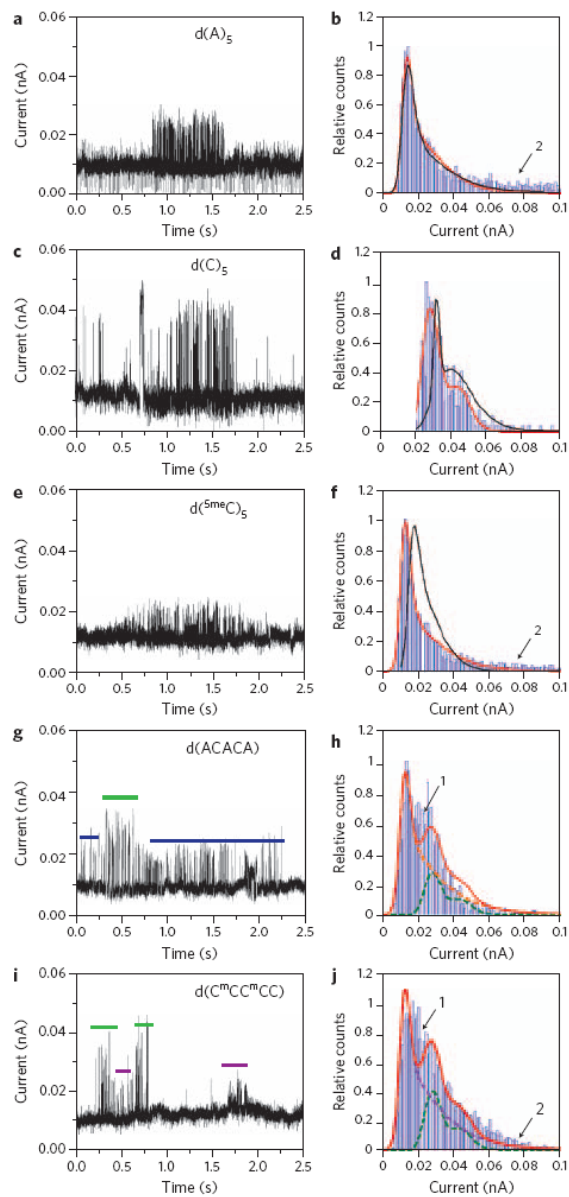
**Figure 6.7** The judgement criteria from signal height and burst frequency. (a) signal height statistics for d(C)<sub>5</sub> (blue), d(A)<sub>5</sub> (red) and d(<sup>m</sup>C)<sub>5</sub> (purple). (b) Burst frequency statistics for d(C)<sub>5</sub> (blue), d(A)<sub>5</sub> (red) and d(<sup>m</sup>C)<sub>5</sub> (purple). The fit is by a gaussian on a log scale as described in chapter 5.<sup>87</sup>

In a 2.5 sec time window, the duration time of bursts of telegraph signals,  $T_b$ , are clearly demonstrated to be on the order of a second. Transition of different types of telegraph signals is commonly detectable although of smaller probabilities since the chance of the random tip drift along the direction of a piece of ssDNA is small.

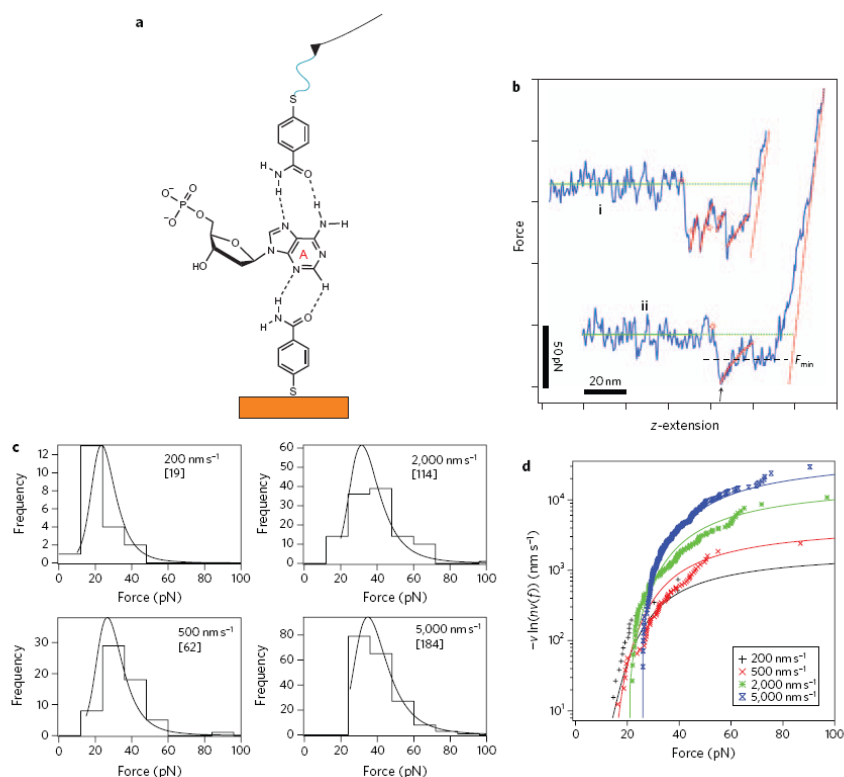
Nucleotide	d(A) <sub>5</sub>	d(C) <sub>5</sub>	d( <sup>m</sup> C) <sub>5</sub>
Burst Duration $T_b$ (s)	0.14±0.02	0.15±0.03	0.41±0.03
Burst Frequency $f_b$ (Hz)	738±100	320±85	662±116
Fraction of reads >0.1 nA	0.20	0.0	0.23
$T_{on}$ (ms)	0.33±0.01	0.34±0.02	0.26±0.01
$T_{off}$ (ms)	0.52±0.02	0.42±0.01	0.47±0.01
$T_{on}/T_{off}$	0.6	0.8	0.6
$\Delta G$ (kT units)	0.51	0.22	0.51

**Table 6.2** Tunneling signal characteristics for DNA oligomers<sup>87</sup>

From the above analysis, bursts of telegraph signals have delivered lots of information for base identification and approved the feasibility of single base resolution by tunneling detection. To understand the nature of telegraph bursts, a separate dynamic force spectroscopy experiment is designed (see **Figure 6.9a**).



**Figure 6.8** Tunneling signal distributions from oligomers resemble those of the constituent nucleotides. A-f, represent current traces from d(A)<sub>5</sub>, dC<sub>5</sub> and dmC<sub>5</sub>. g-j, from oligomers with mixed sequence.<sup>87</sup>



**Figure 6.9** The dynamic force study of the interaction between dAMP and DNA reader. (a) Schematics of the dynamic force study. (b) Typical force-distance curves for multiple/single molecules pulling. (c) A comparison of bond breaking force with different pulling speed. (d) Master curves reconstructed by the force distance curves.<sup>87</sup>

In the dynamic force measurement, one of the recognition molecules is bound to an AFM probe via a 34-nm long polyethylenecol (PEG) linker, while the other is chemically attached to the Au(111) surface as SAMs<sup>112-115</sup>. To compare with the tunneling measurement, one of the nucleotide monomer dAMP is used as a target analyte to bridge the gap in the 1mM PB buffer environment. In the control buffer absence of dAMP, adhesion between the probe and substrate was extremely small, presumably because the hydrogen binding sites on the benzamide reader is bound by



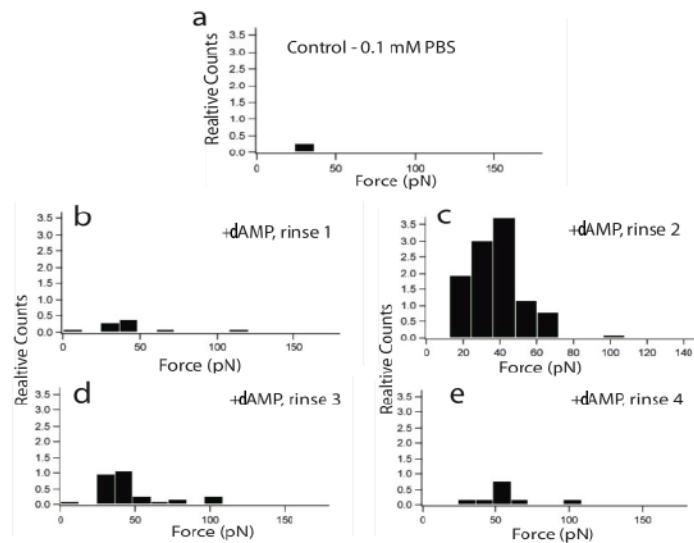
the water molecules which diminished the interactions. To perform dAMP force spectroscopy, the functionalized gold substrate is first dipped in the solution of 10 $\mu$ M dAMP for 10 sec and then rinsed with 1mM PB buffer to get rid of non-specifically bound analyte. Ramp shaped adhesion features (see **Figure 6.9b**) are immediately observed after rinsing the chip for several times.

These adhesion force curves (rupture events) are then analyzed separately. The rupture force, which is the biggest force achieved before the bond breaking is analyzed with different pulling speed. By fitting with the following relationship:

$$F = \frac{k_B T}{x_\beta} \ln\left(\frac{r_f x_\beta}{k_B T k_{off}}\right) \quad (6.1)$$

Here  $r_f$  represents the loading rate in the unit of pN/sec, which equals to the pulling speed times the cantilever spring constant and the linker elasticity.  $x_\beta$  represents the specific length which is 0.78 nm for this case.  $K_{off}$  is the dissociation constant of the bond. Fitted curve generates results of  $K_{off}=0.28 \text{ sec}^{-1}$ , which means the average lifetime for the bonding is on the order of seconds. Inserting numbers into the Bell equation (**Equation 6.2**), by assuming the force applied by the nanopore is 0.24pN/mV, it requires 80mV to achieve 10nucleotide/sec detection speed.

$$k(f) = k_{off} \exp\left(\frac{x_\beta f}{k_B T}\right) \quad (6.2)$$



**Figure 6.10** A control curves taken in the absence of dAMP showed almost no adhesion events between the benzamide molecules, presumably because they are blocked by water. Addition of dAMP led to a number of adhesion events that increased as excess dAMP was rinsed out of the system (b,c) decreasing as the rinsing continued (d,e)<sup>87</sup>

Too high and too low surface concentration both results in small probability of adhesion signals. Low concentration will normally results in small probability of molecule binding however too high concentration may force the molecules to bind the tip and the substrate separately which diminishes the probability of the molecule spanning the gap. Stretching of the PEG tether generated a characteristic signal that permitted multiple binding events (see **Figure 6.9b** top trace) to be separated from single molecule bond breaking events (see **Figure 6.9b** bottom traces) so that multiple binding molecules are excluded from the statistics. Single molecule bond breaking forces as a function of pulling speed are summarized in **Figure 6.9c** and bond survival probability as a function of

bond breaking force is shown in **Figure 6.9c**. Fitting to the same heterogeneous bond model yields an off-rate at zero force to be  $K_{\text{off}}^0=0.28 \text{ s}^{-1}$ . Then I conclude that the ideal duration time trapped analyte should be on the order of seconds at zero-force condition. The analysis also yields the distance to the transition state for dissociation,  $\alpha=0.78\pm0.19 \text{ nm}$ . Although the STM tip is kept fixed, the force applied by random drift is unknown in reality but the dynamic force experiment offered a good estimation of the dissociation constant which makes me certain that the cluster of telegraph burst characterizes a single base read.

#### **6.4 Conclusion**

Telegraph bursts signals with the duration time of a fraction of seconds and hundreds Hz switching rate has provided invaluable information to unpuzzle the mysteries coded in DNA sequences. To further this approach to single molecule DNA sequencing of long reads, a nanopore or nanotube based DNA translocation device has to be incorporated to simplify the sequential read challenge. An active probe scanning approach is being developed in the mean time as a prototype for single molecule DNA sequential reading.

A clear gap signal between adjacent bases may also be needed since the probability detecting the same neighbor bases in the case of a real DNA is

not small and so far this technique could not tell such transition unless it is a transition between dramatically different signals like A and C as depicted in **Figure 6.8g**. This approach is still under study and will be demonstrated in Chapter 7 too.

Although be capable to identify 3 DNA bases by combining several key parameters, the benzamide reader is still not the ultimate DNA reader to be used. A much larger and clearer separation by signals with a complete separation spectrum of 5 DNA bases is required by designing readers based on the information gathered in the preliminary trials of different compounds.

## **7. THE ONGOING AND FUTURE RESEARCH: TO READ THE SEQUENCE OF DNA**

### **7.1 Introduction**

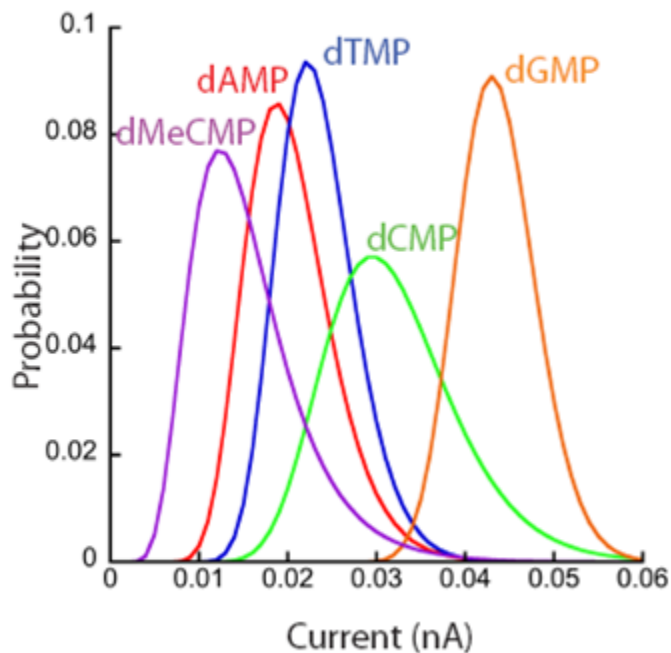
In the previous chapters, I have listed the major work that I have performed to realize next generation DNA sequencing via tunneling readout. The preliminary results have shown convincing evidence that single DNA base could be identified by its characteristic tunneling signals. The composition of DNA nucleotides and DNA oligomers could be successfully obtained. However, the sequence of DNA is not able to be grabbed via tunneling readout so far. The STM based metal junction measurement is convenient due to its simplicity and flexibility however it is difficult to make DNA flowing sequentially through the gap due to the geometrical disadvantage of the gap shape and the difficulty manipulating the DNA in it.

Active scanning of the tip along a strand of ssDNA may deliver the DNA sequence information. Imidazole reader generates the best result so far and results from this specific reader will be discussed intensively in this chapter.

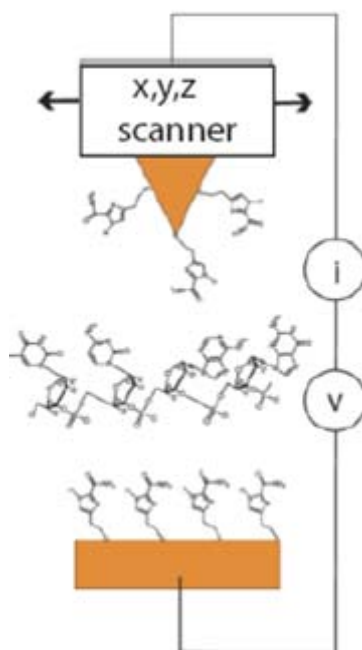
## 7.2 Immidazole Reader

Benzamide reader is the 1<sup>st</sup> reader that is tested to achieve single base resolution on short DNA oligomers but it is not the only reader that is able to do this.

DNA reader with immidazole structure (Immidazole reader) is customarily synthesized in the lab<sup>116</sup> and show advantages of complete spectrum of all 5 bases and slightly better separations between different bases (see **Figure 7.1**). The most important feature is that DNA sequence information could be read out by sliding functionalized tips along the surface. Sequence-like signals are observed occasionally.



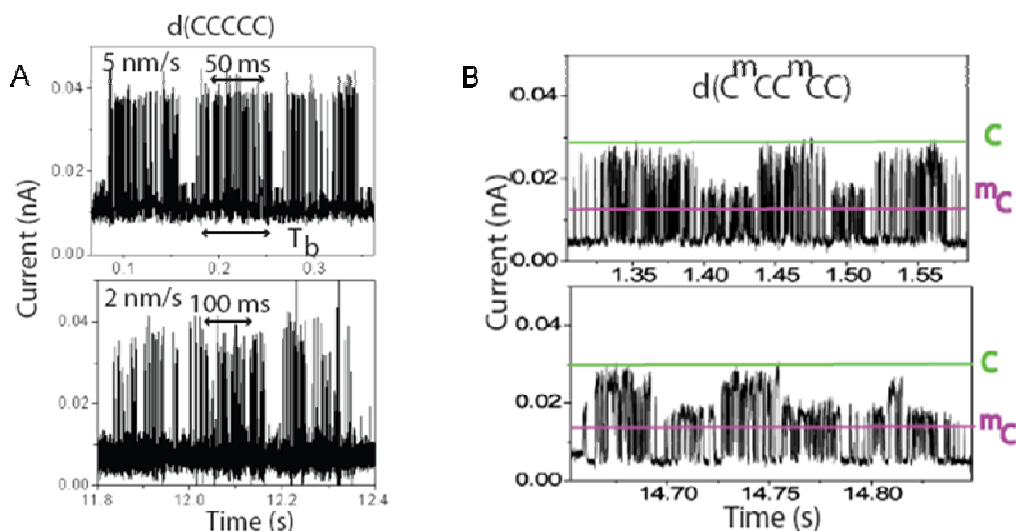
**Figure 7.1** Distributions of current spike heights of the immidazole reader for 5 different DNA nucleotides.



**Figure 7.2** Schematic Diagram of linear scanning over a piece of ssDNA

### 7.3 DNA Sequencing by Linear Scanning

The treatment of tips and Au(111) substrate is quite similar to the procedures depicted in Chapter 6. The tips are piranha treated and cleaned prior to HDPE coating (see **Appendix A** for details). 1mM ethanol (Sigma-Aldrich, HPLC grade) solution of imidazole reader is prepared for tip functionalization. The functionalization normally takes 2 hours to finish by immersing a bare tip into the imidazole solution and forms linkage via gold-thiol reactions. Over immersion of the HDPE coated tip in ethanol solution will result in the damaging of the coating and it is not recommended to functionalize the tip for over 2 hours.



**Figure 7.3** the typical sequence signal obtained by Imidazole reader.  
 (A) d(CCCCC); (B) d(C<sup>m</sup>CC<sup>m</sup>CC)

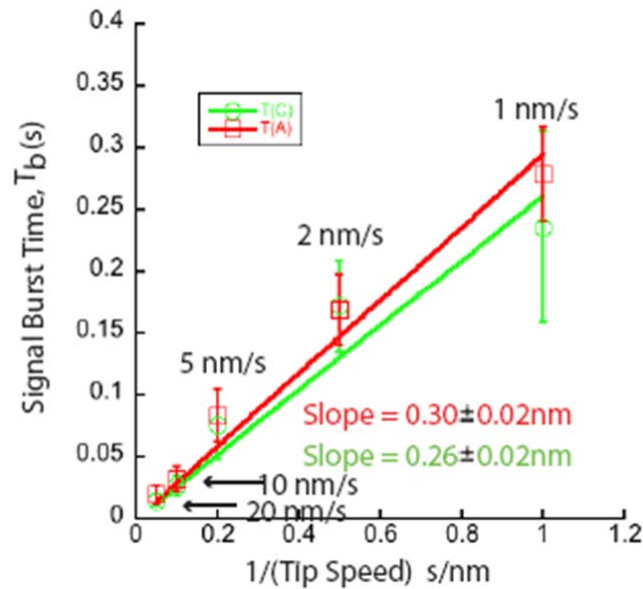
The functionalization of the Au(111) substrate is done by immersing the hydrogen flame annealed chip in the 1mM ethanol solution of imidazole reader for overnight. Extended functionalization time normally results in better assembled and more uniform monolayer.

The measurement is done in 10 mM Phosphate buffered solution (PB buffer) with 50 nM dissolved DNA oligomers in it. In a STM setup, the tip is mechanically approached to the surface and maintained by a piezoelectric servo control at 6 pA baseline. The gap size is tested to be large enough so that the reader on both sides doesn't interact directly and no signals are observed during the control experiment with no DNA added. The



estimated gap size is about 2-3 angstroms so it is about the size for one DNA base to snip in.

During the measurement, the tip is forced to move along an arbitrary linear direction in hope it could coincidentally align with the position and the direction of a piece of ssDNA. Once this arbitrary scanning aligns with the DNA, a typical serial data will show up with information potentially related to the sequence of the DNA (see **Figure 7.3**).



**Figure 7.4** The plot of  $T_b$  vs  $1/V$ . The slope of the fitted line indicates the average distance the tip has traveled corresponding to a single block of telegraph noise. The calculated distance is about the size of a single base (2-3 Angstrom)

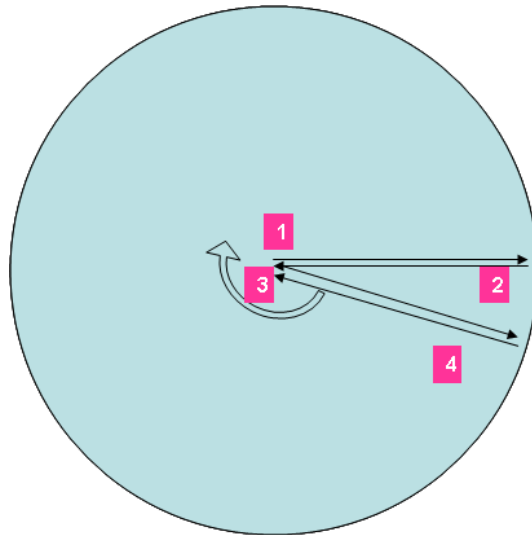
The telegraph signals in this case are always in the form of blocks of burst signals with gaps in between. This is not always clear for homo-polymers but much more obvious when reading the sequence of  $d(C^mCC^mCC)$ . The

periodic changing of telegraph signal height is clearly indicating the transition from one base to another.

By analyzing the duration time of each burst ( $T_b$ ) and put together with the tip speed that we used. A  $T_b$  vs.  $(l/v)$  is drawn as seen in **Figure 7.4**. The slope of the fitted line is the average distance that the tip has traveled during the duration time of a single block. It is very interesting to know that this distance is always a constant which is about the size of a single DNA base. Thus it is indicating the reading of DNA bases on the surface by scanning it.

#### **7.4 Clock Scanning Method**

Due to the difficulties of positioning and aligning the DNA target on the surface, a clock scanning mode is developed by a customarily made labview program.

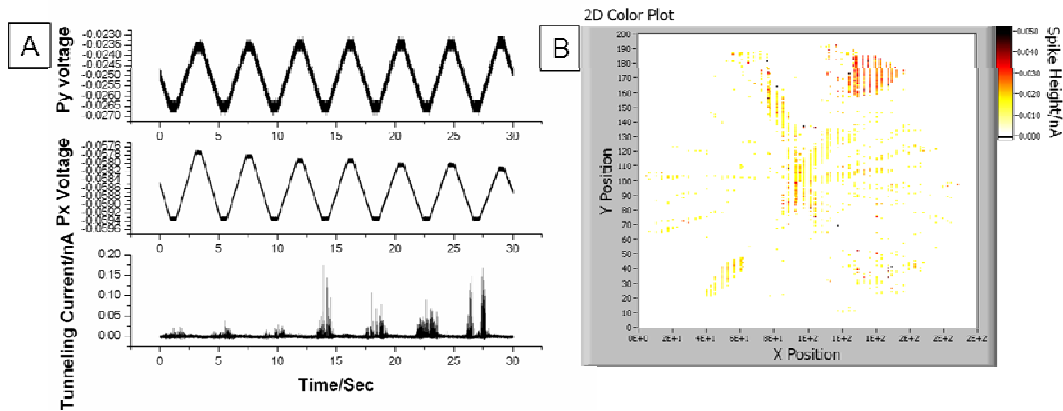


**Figure 7.5** The Schematic Diagram of a Clock Scanning

In a typical cycle of clock scanning (see **Figure 7.5**), the tip is first set in the center position of a circular area and then forced to slide from the center to the boundary and then come back. The following phase is scanning from center to the boundary again but with a small processing angle clockwise relative to the previous phase. This process is repeated one by one until a complete circular area is covered.

During a complete clock scanning cycle, the entire surface within the circular area has been scanned and all the angles have been tested which helps to position and align the DNA target by trying all the combined position and angles within a small area while keeping the gap size constant under the servo control.

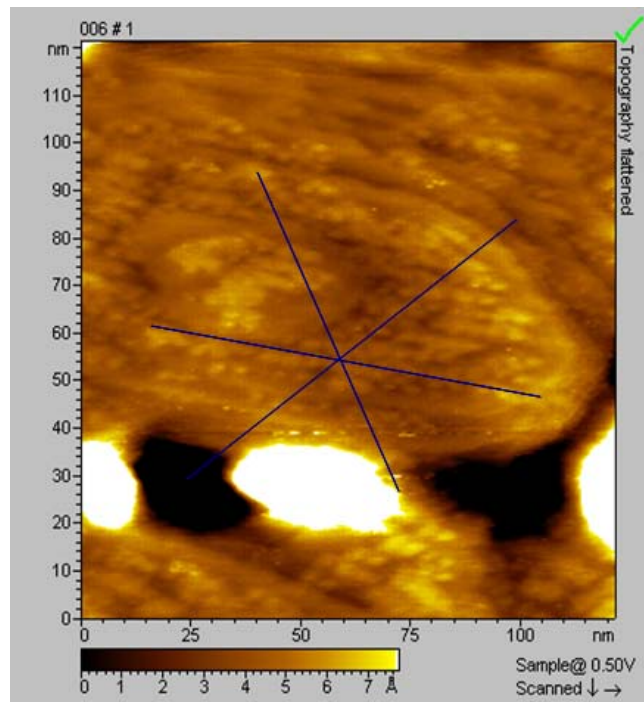
The Px and Py signals, which are the applied bias on the X and Y direction of the piezo-electric material could be used to calculate the position of tip. During an actual data measurement, data of 3 separate channels (Px, Py and the tunneling current, see **Figure 7.6A**) are recorded. By reconstructing a 2D mapping plot using the information from the 3 channels, I could be able to tell the position where a DNA signal is being detected. The intensity plot is color coded and the color represents the signal height strength at that position (see **Figure 7.6B**).



**Figure 7.6** (A) Typical signals during the clock scanning. (B) Reconstructed 2D mapping plot of the scanned area.

Another interesting discovery is that the signal tends to show up in specific directions only, which are 60 degrees relative to each other (see **Figure 7.8**). Similar phenomenon has been observed before during STM imaging (see **Figure 7.7**). After depositing DNA oligomer d(AAAAA) over the surface of benzamide SAM, reconstruction of self assembled monolayer

shows up and forms structures with 6 fold symmetry. This similar phenomenon is quite likely due to the symmetry coming from the Au(111) lattice structure which has the same symmetry. It is straight forward to imagine that the short DNA oligomers tend to align with the Au(111) lattice to achieve energy minimum. Thus the positioning of the DNA on the surface is less random and sequencing over a 2D surface becomes quite achievable.

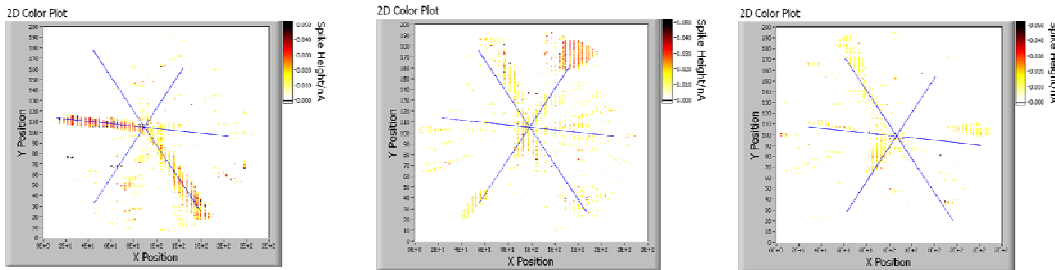


**Figure 7.7** DNA self assembly over a benzamide reader on the surface of Au(111) lattice

Measurements over several different spots on the same chip always show parts of the 6 fold symmetry (see **Figure 7.8**) while quite repeatable for various spots. The overlap of the 3 images in Figure 7.8 forms a complete

set of distribution of the 6 directions. Such uniformity all over the surface on the same chip is also indicating that it is more related to the Au(111) structure.

Currently, our group is working on developing an objective clock scanning program which automates the data analysis.



**Figure 7.8** 2D intensity plot from measurements over 3 different centre locations on the same chip.

## 7.5 Nanopore Tunneling Measurement

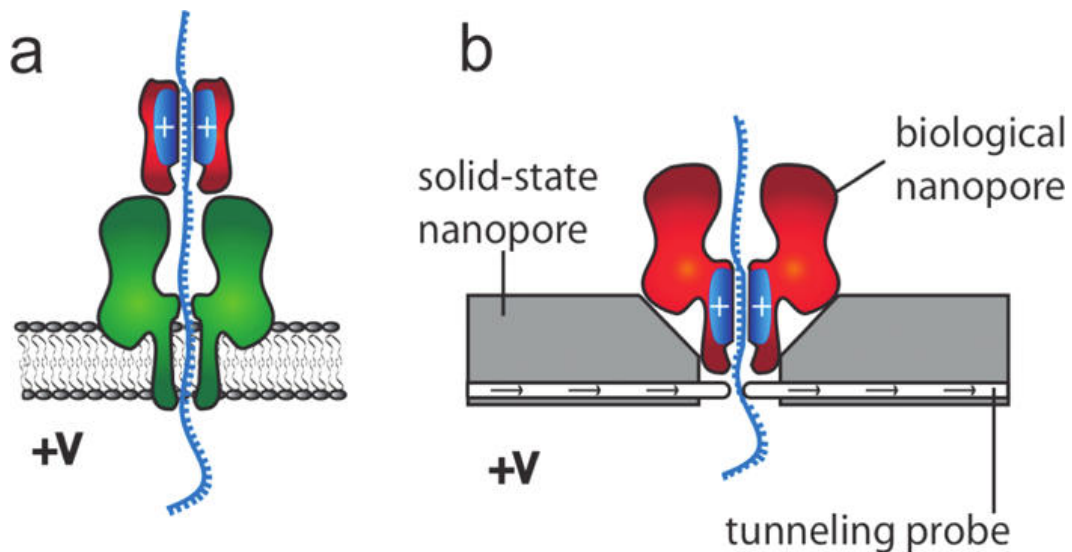
Although the STM based DNA identification system has shown convincing results as a prototype method for single molecule DNA sequencing.

However, the 2D geometry of the STM substrate is limiting its application for sensing the sequence of a 1 dimensional and strand shaped target.

The study in nanopore based sensors have provided an alternative way to combine a tunneling reading device to the pore opening and thus it is able

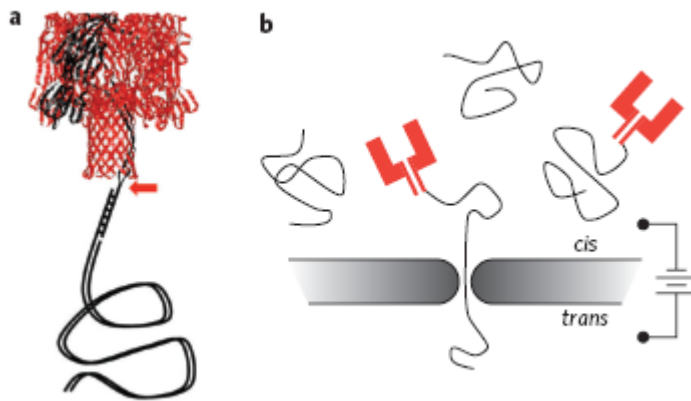
to read out the sequence by letting a piece of ssDNA to floss through it. To achieve such goal, several targets has to be achieved:

1) A nanopore device with the capability to translocate ssDNA only but not dsDNA is required since the recognition tunneling needs a completely open and unbounded DNA but not basepair to read. It is more ideal if the pore could open possible secondary structure of the DNA when it is being translocated. The protein nanopore ( $\alpha$ -hemolysine) which is pioneered by Hagan Bayley's group in University of Oxford fits the need exactly for its atomically accurate structure and its selectivity for ssDNA translocation.



**Figure 7.9** A proposed design of a translocation brake in a DNA sensor. (a) A brake placed outside an  $\alpha$ -hemolysin nanopore (b) A reconstructed  $\alpha$ -hemolysin nanopore placed inside a solid state nanopore with a tunneling readout junction.<sup>117</sup>

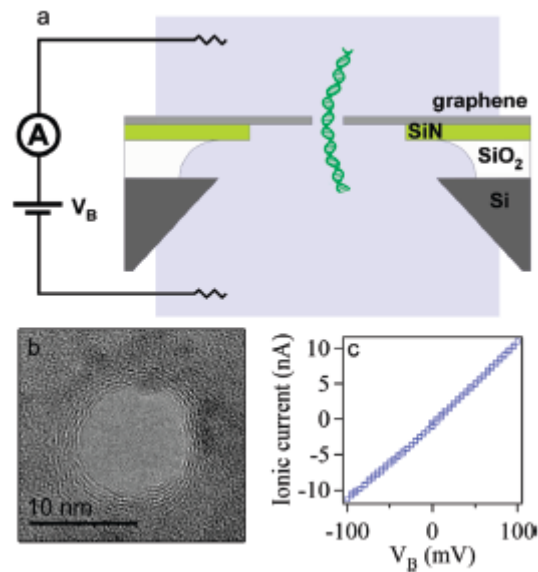
- 2) However, the protein nanopore has its drawback that the pore has to be placed across a lipid bilayer membrane and limit its stability and makes it impossible to put a pair of tunneling electrode around the pore opening. At this point, the solid state nanopore is more able to perform the job.
- 3) A signal gap between neighboring bases has to be well characterized.



**Figure 7.10** Hybrid nanopore formed by inserting  $\alpha$ -hemolysin into the solid state nanopore. (a)  $\alpha$ -hemolysin with a DNA tail (b) The DNA tail directs the insertion of  $\alpha$ -hemolysin into the pore via electrophoresis.<sup>118</sup>

In the most recent published paper, the  $\alpha$ -hemolysin protein nanopore is chemically modified to have positive charge around its  $\beta$ -barrel, which significantly slows down the DNA translocation so that the sensor get enough time resolution to identify each single bases. Almost at the same time, published result showed the possibility that a protein nanopore, which is chemically linked to a dsDNA is able to be inserted into a solid nanopore<sup>118</sup> which potentially solved the durability problem of the  $\alpha$ -hemolysin pore.





**Figure 7.11** (a) The schematic diagram of DNA translocation through a graphene nanopore. (b) TEM image of the graphene nanopore. Layered edges indicate the number of layers of the graphene. Normally more than 5 layers are needed to gain enough robustness. (c) I-V curve for the nanopore. Device in 1M KCl, Ph 9.<sup>22</sup>

The world's thinnest nanopore based on graphene material is also showing great potential as a future candidate due to its stability and size resolution (below 1nm in thickness).<sup>22, 35, 36</sup> It is also exciting that the graphene is an excellent conductor which could be used as one of the electrode with well defined thickness in the near future.

## 8. CONCLUSION

In this thesis, I summarized the major work that I have done for the novel DNA sequencing project during my Phd study. Hydrogen bonded tunneling conductance phenomenon in the single molecule level is observed by detecting the typical telegraph switching signals and verified to be a solid tool for single molecule recognition both in organic and aqueous buffered solution. The difficulty lies in the selection of the optimum DNA readers featuring good self assembly and enhanced conductance difference. Several DNA readers which could mimic the hydrogen bonding between DNA base pairs are intensively studied and compared. Single base resolution on short DNA oligomers could be resolved with DNA reader functionalized tips already.

In the most recent study, tip scanning on short oligomers have shown potential sequence readout signals with correct time constant (time needed for reading one single base). The challenge here is the positioning and alignment of the tip scanning and the DNA target. For such reasons, a clock scanning mode is being developed in the lab to increase the probability although still with low yield of sequence-like signals.

In the near future, assisted by nanopore devices or carbon nanotube devices, the recognition tunneling mechanism could be incorporated into

such a robust nanopore based sensor. Parallel genome reading on such devices would offer a fast, accurate and economic way for decoding genome mysteries. With the feature of identifying methylated and regular cytidine, it is also an invaluable device for epigenetic and cancer research.

## REFERENCES

1. Dahm, R., Friedrich Miescher and the discovery of DNA. *Developmental Biology* **2005**, 274-288.
2. Maderspacher, F., Rags before the riches: Friedrich Miescher and the discovery of DNA. *Current Biology* **2004**, R608-R608.
3. WATSON, J.; CRICK, F., MOLECULAR STRUCTURE OF NUCLEIC ACIDS - A STRUCTURE FOR DEOXYRIBOSE NUCLEIC ACID. *Nature* **1953**, 737-738.
4. WATSON, J.; CRICK, F., GENETICAL IMPLICATIONS OF THE STRUCTURE OF DEOXYRIBONUCLEIC ACID. *Nature* **1953**, 964-967.
5. <http://academic.brooklyn.cuny.edu>
6. Branton, D.; Deamer, D.; Marziali, A.; Bayley, H.; Benner, S.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X.; Jovanovich, S.; Krstic, P.; Lindsay, S.; Ling, X.; Mastrangelo, C.; Meller, A.; Oliver, J.; Pershin, Y.; Ramsey, J.; Riehn, R.; Soni, G.; Tabard-Cossa, V.; Wanunu, M.; Wiggin, M.; Schloss, J., The potential and challenges of nanopore sequencing. *Nature Biotechnology* **2008**, 1146-1153.
7. ESCARMIS, C.; SALAS, M., NUCLEOTIDE-SEQUENCE OF THE EARLY GENE-3 AND GENE-4 OF BACTERIOPHAGE-PHI-29. *Nucleic Acids Research* **1982**, 5785-5798.
8. KING, S.; KROLEWSKI, M.; MARVO, S.; LIPSON, P.; POGUEGEILE, K.; CHUNG, J.; JASKUNAS, S., NUCLEOTIDE-SEQUENCE ANALYSIS OF INVIVO RECOMBINANTS BETWEEN BACTERIOPHAGE-LAMBDA DNA AND PBR322. *Molecular & General Genetics* **1982**, 548-557.
9. <http://en.wikipedia.org/wiki/Chromosome>.
10. Kool, E.; Morales, J.; Guckian, K., Mimicking the structure and function of DNA: Insights into DNA stability and replication. *Angewandte Chemie-International Edition* **2000**, 990-1009.

11. Rothkamm, K.; Lobrich, M., Evidence for a lack of DNA double-strand break repair in human cells exposed to very low x-ray doses. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, 5057-5062.
12. Moore, G.; Maranas, C., Modeling DNA mutation and recombination for directed evolution experiments. *Journal of Theoretical Biology* **2000**, 483-503.
13. BROWN, W.; GEORGE, M.; WILSON, A., RAPID EVOLUTION OF ANIMAL MITOCHONDRIAL-DNA. *Proceedings of the National Academy of Sciences of the United States of America* **1979**, 1967-1971.
14. Aarnio, M.; Sankila, R.; Pukkala, E.; Salovaara, R.; Aaltonen, L.; de la Chapelle, A.; Peltomaki, P.; Mecklin, J.; Jarvinen, H., Cancer risk in mutation carriers of DNA-mismatch-repair genes. *International Journal of Cancer* **1999**, 214-218.
15. NAKAGAWA, Y.; IKEGAMI, H.; YAMATO, E.; TAKEKAWA, K.; FUJISAWA, T.; HAMADA, Y.; UEDA, H.; UCHIGATA, Y.; MIKI, T.; KUMAHARA, Y.; OGIHARA, T., A NEW MITOCHONDRIAL-DNA MUTATION ASSOCIATED WITH NON-INSULIN-DEPENDENT DIABETES-MELLITUS (VOL 209, PG 664, 1995). *Biochemical and Biophysical Research Communications* **1995**, 718-718.
16. KADOWAKI, T.; KADOWAKI, H.; MORI, Y.; TOBE, K.; SAKUTA, R.; SUZUKI, Y.; TANABE, Y.; SAKURA, H.; AWATA, T.; GOTO, Y.; HAYAKAWA, T.; MATSUOKA, K.; KAWAMORI, R.; KAMADA, T.; HORAI, S.; NONAKA, I.; HAGURA, R.; AKANUMA, Y.; YAZAKI, Y., A SUBTYPE OF DIABETES-MELLITUS ASSOCIATED WITH A MUTATION OF MITOCHONDRIAL-DNA. *New England Journal of Medicine* **1994**, 962-968.
17. Bayley, H., Sequencing single molecules of DNA. *Current Opinion in Chemical Biology* **2006**, 628-637.
18. SANGER, F., CITATION CLASSIC - DNA SEQUENCING WITH CHAIN-TERMINATING INHIBITORS. *Current Contents/life Sciences* **1988**, 23-23.
19. SANGER, F., DETERMINATION OF NUCLEOTIDE-SEQUENCES IN DNA. *Science* **1981**, 1205-1210.

20. Shendure, J.; Ji, H., Next-generation DNA sequencing. *Nature Biotechnology* **2008**, 1135-1145.
21. 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**, 1070-1075.
22. Merchant, C.; Healy, K.; Wanunu, M.; Ray, V.; Peterman, N.; Bartel, J.; Fischbein, M.; Venta, K.; Luo, Z.; Johnson, A.; Drndic, M., DNA Translocation through Graphene Nanopores. *Nano Letters* **2010**, 2915-2921.
23. Stoddart, D.; Maglia, G.; Mikhailova, E.; Heron, A.; Bayley, H., Multiple Base-Recognition Sites in a Biological Nanopore: Two Heads are Better than One. *Angewandte Chemie-International Edition* **2010**, 556-559.
24. Stoddart, D.; Heron, A.; Klingelhoefer, J.; Mikhailova, E.; Maglia, G.; Bayley, H., Nucleobase Recognition in ssDNA at the Central Constriction of the alpha-Hemolysin Pore. *Nano Letters* **2010**, 3633-3637.
25. Tsutsui, M.; Taniguchi, M.; Yokota, K.; Kawai, T., Identifying single nucleotides by tunnelling current. *Nature Nanotechnology* **2010**, 286-290.
26. Clarke, J.; Wu, H.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H., Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology* **2009**, 265-270.
27. Gupta, P., Single-molecule DNA sequencing technologies for future genomics research. *Trends in Biotechnology* **2008**, 602-611.
28. McNally, B.; Singer, A.; Yu, Z.; Sun, Y.; Weng, Z.; Meller, A., Optical Recognition of Converted DNA Nucleotides for Single-Molecule DNA Sequencing Using Nanopore Arrays. *Nano Letters* **2010**, 2237-2244.
29. Singer, A.; Meller, A., Nanopore-based Sensing of Individual Nucleic Acid Complexes. *Israel Journal of Chemistry* **2009**, 323-331.
30. Astier, Y.; Bayley, H.; Howorka, S., Protein components for nanodevices. *Current Opinion in Chemical Biology* **2005**, 576-584.
31. Bayley, H.; Jayasinghe, L.; Wallace, M., Prepore for a breakthrough. *Nature Structural & Molecular Biology* **2005**, 385-386.

32. Bayley, H.; Jayasinghe, L., Functional engineered channels and pores - (Review). *Molecular Membrane Biology* **2004**, 209-220.
33. Derrington, I.; Butler, T.; Collins, M.; Manrao, E.; Pavlenok, M.; Niederweis, M.; Gundlach, J., Nanopore DNA sequencing with MspA. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, 16060-16065.
34. Wanunu, M.; Dadosh, T.; Ray, V.; Jin, J.; McReynolds, L.; Drndic, M., Rapid electronic detection of probe-specific microRNAs using thin nanopore sensors. *Nature Nanotechnology* **2010**, 807-814.
35. Garaj, S.; Hubbard, W.; Reina, A.; Kong, J.; Branton, D.; Golovchenko, J., Graphene as a subnanometre trans-electrode membrane. *Nature* **2010**, 190-U73.
36. Schneider, G.; Kowalczyk, S.; Calado, V.; Pandraud, G.; Zandbergen, H.; Vandersypen, L.; Dekker, C., DNA Translocation through Graphene Nanopores. *Nano Letters* **2010**, 3163-3167.
37. Liu, H.; He, J.; Tang, J.; Liu, H.; Pang, P.; Cao, D.; Krstic, P.; Joseph, S.; Lindsay, S.; Nuckolls, C., Translocation of Single-Stranded DNA Through Single-Walled Carbon Nanotubes. *Science* **2010**, 64-67.
38. van den Hout, M.; Hall, A.; Wu, M.; Zandbergen, H.; Dekker, C.; Dekker, N., Controlling nanopore size, shape and stability. *Nanotechnology* **2010**, -.
39. Bayley, H., NANOTECHNOLOGY Holes with an edge. *Nature* **2010**, 164-165.
40. Bayley, H., MEMBRANE-PROTEIN STRUCTURE Piercing insights. *Nature* **2009**, 651-652.
41. Bayley, H., Engineered protein pores for nanotechnology. *Biophysical Journal* **2007**, 163A-164A.
42. Bayley, H.; Romesberg, F., From therapeutic nucleic acids to redox hydrogels: the diverse world of biopolymers. *Current Opinion in Chemical Biology* **2006**, 598-600.
43. Bayley, H., Ion channels get flashy. *Nature Chemical Biology* **2006**, 11-13.

44. Astier, Y.; Braha, O.; Bayley, H., Toward single molecule DNA sequencing: Direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *Journal of the American Chemical Society* **2006**, 1705-1710.
45. Lundstrom, M., Moore's law forever? *Science* **2003**, 210-211.
46. Tour, J., Molecular electronics. Synthesis and testing of components. *Accounts of Chemical Research* **2000**, 791-804.
47. Thompson, S.; Parthasarathy, S., Moore's law: the future of Si microelectronics. *Materials Today* **2006**, 20-25.
48. Chang, S.; He, J.; Lin, L.; Zhang, P.; Liang, F.; Young, M.; Huang, S.; Lindsay, S., Tunnel conductance of Watson-Crick nucleoside-base pairs from telegraph noise. *Nanotechnology* **2009**, -.
49. He, J.; Lin, L.; Liu, H.; Zhang, P.; Lee, M.; Sankey, O.; Lindsay, S., A hydrogen-bonded electron-tunneling circuit reads the base composition of unmodified DNA. *Nanotechnology* **2009**, -.
50. Cui, X.; Zarate, X.; Tomfohr, J.; Sankey, O.; Primak, A.; Moore, A.; Moore, T.; Gust, D.; Harris, G.; Lindsay, S., Making electrical contacts to molecular monolayers. *Nanotechnology* **2002**, 5-14.
51. Xia, J.; Chen, F.; Wiktor, P.; Ferry, D.; Tao, N., Effect of Top Dielectric Medium on Gate Capacitance of Graphene Field Effect Transistors: Implications in Mobility Measurements and Sensor Applications. *Nano Letters* **2010**, 5060-5064.
52. Li, X.; Xu, B.; Xiao, X.; Yang, X.; Zang, L.; Tao, N., Controlling charge transport in single molecules using electrochemical gate. *Faraday Discussions* **2006**, 111-120.
53. Xiao, X.; Brune, D.; He, J.; Lindsay, S.; Gorman, C.; Tao, N., Redox-gated electron transport in electrically wired ferrocene molecules. *Chemical Physics* **2006**, 138-143.
54. Xu, B.; Xiao, X.; Yang, X.; Zang, L.; Tao, N., Large gate modulation in the current of a room temperature single molecule transistor. *Journal of the American Chemical Society* **2005**, 2386-2387.



55. STRONG, L.; WHITESIDES, G., STRUCTURES OF SELF-ASSEMBLED MONOLAYER FILMS OF ORGANOSULFUR COMPOUNDS ADSORBED ON GOLD SINGLE-CRYSTALS - ELECTRON-DIFFRACTION STUDIES. *Langmuir* **1988**, 546-558.
56. Xu, H.; Shorubalko, I.; Maximov, I.; Seifert, W.; Omling, P.; Samuelson, L., A novel device principle for nanoelectronics. *Materials Science & Engineering C-Biomimetic and Supramolecular Systems* **2002**, 417-420.
57. Xu, B.; Tao, N., Measurement of single-molecule resistance by repeated formation of molecular junctions. *Science* **2003**, 1221-1223.
58. He, J. Ph.D thesis. Arizona State University, Tempe, 2005.
59. CHANG, L.; ESAKI, L.; TSU, R., RESONANT TUNNELING IN SEMICONDUCTOR DOUBLE BARRIERS. *Applied Physics Letters* **1974**, 593-595.
60. Lee, M.; Sankey, O., Insights into electron tunneling across hydrogen-bonded base-pairs in complete molecular circuits for single-stranded DNA sequencing. *Journal of Physics-Condensed Matter* **2009**, -.
61. Lee, M.; Sankey, O., Theory of tunneling across hydrogen-bonded base pairs for DNA recognition and sequencing. *Physical Review E* **2009**, -.
62. Derosa, P.; Guda, S.; Seminario, J., A programmable molecular diode driven by charge-induced conformational changes. *Journal of the American Chemical Society* **2003**, 14240-14241.
63. Mizuta, H.; Oda, S., Bottom-up approach to silicon nanoelectronics. *Microelectronics Journal* **2008**, 171-176.
64. Lu, W.; Lieber, C., Nanoelectronics from the bottom up. *Nature Materials* **2007**, 841-850.
65. Kang, S., Silicon nanoelectronics: Precise fabrication via a bottom-up approach. *Jom* **2004**, 19-19.
66. Sentein, C.; Fiorini, C.; Lorin, A.; Nunzi, J., Molecular rectification in oriented polymer structures. *Synthetic Metals* **1997**, 81-82.

67. Sentein, C.; Fiorini, C.; Lorin, A.; Nunzi, J., Molecular rectification in oriented polymer structures. *Advanced Materials* **1997**, 809-&.
68. Lindsay, S.; Ratner, M., Molecular transport junctions: Clearing mists. *Advanced Materials* **2007**, 23-31.
69. BINNIG, G.; ROHRER, H., SCANNING TUNNELING MICROSCOPY. *Helvetica Physica Acta* **1982**, 726-735.
70. BINNIG, G.; GERBER, C.; STOLL, E.; ALBRECHT, T.; QUATE, C., ATOMIC RESOLUTION WITH ATOMIC FORCE MICROSCOPE. *Surface Science* **1987**, 1-6.
71. BINNIG, G.; ROHRER, H., SCANNING TUNNELING MICROSCOPY - FROM BIRTH TO ADOLESCENCE. *Reviews of Modern Physics* **1987**, 615-625.
72. MARTI, O.; BINNIG, G.; ROHRER, H.; SALEMINK, H., LOW-TEMPERATURE SCANNING TUNNELING MICROSCOPE. *Surface Science* **1987**, 230-234.
73. BINNIG, G.; QUATE, C.; GERBER, C., ATOMIC FORCE MICROSCOPE. *Physical Review Letters* **1986**, 930-933.
74. BINNIG, G.; ROHRER, H., SCANNING TUNNELING MICROSCOPY. *Surface Science* **1983**, 236-244.
75. BINNIG, G.; ROHRER, H., SCANNING TUNNELING MICROSCOPY, AN ATOMIC PROBE. *Scanning Electron Microscopy* **1983**, 1079-1082.
76. BINNIG, G.; ROHRER, H.; GERBER, C.; WEIBEL, E., 7X7 RECONSTRUCTION ON SI(111) RESOLVED IN REAL SPACE. *Physical Review Letters* **1983**, 120-123.
77. Emundts, A.; Coenen, P.; Pirug, G.; Voigtlander, B.; Bonzel, H.; Wynblatt, P., Combination of a Besocke-type scanning tunneling microscope with a scanning electron microscope. *Review of Scientific Instruments* **2001**, 3546-3551.
78. Metzker, M., APPLICATIONS OF NEXT-GENERATION SEQUENCING Sequencing technologies - the next generation. *Nature Reviews Genetics* **2010**, 31-46.

79. Xia, J.; Diez-Perez, I.; Tao, N., Electron transport in single molecules measured by a distance-modulation assisted break junction method. *Nano Letters* **2008**, 1960-1964.
80. Tao, N., Electron transport in molecular junctions. *Nature Nanotechnology* **2006**, 173-181.
81. Tao, N.; Xu, B.; Xiao, X., Measurement of electron transport in single redox molecules. *Abstracts of Papers of the American Chemical Society* **2005**, U726-U726.
82. Xiao, X.; Xu, B.; Tao, N., Changes in the conductance of single peptide molecules upon metal-ion binding. *Angewandte Chemie-International Edition* **2004**, 6148-6152.
83. Xu, B.; Zhang, P.; Li, X.; Tao, N., Direct conductance measurement of single DNA molecules in aqueous solution. *Nano Letters* **2004**, 1105-1108.
84. Nichols, R.; Haiss, W.; Higgins, S.; Leary, E.; Martin, S.; Bethell, D., The experimental determination of the conductance of single molecules. *Physical Chemistry Chemical Physics* **2010**, 2801-2815.
85. Haiss, W.; Martin, S.; Leary, E.; van Zalinge, H.; Higgins, S.; Bouffier, L.; Nichols, R., Impact of Junction Formation Method and Surface Roughness on Single Molecule Conductance. *Journal of Physical Chemistry C* **2009**, 5823-5833.
86. Haiss, W.; Wang, C.; Grace, I.; Batsanov, A.; Schiffrin, D.; Higgins, S.; Bryce, M.; Lambert, C.; Nichols, R., Precision control of single-molecule electrical junctions. *Nature Materials* **2006**, 995-1002.
87. 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. *Nature Nanotechnology* **2010**, 868-873.
88. Huang, S., Recognition Tunneling Measurement of the Conductance of DNA Bases Embedded in Self-Assembled Monolayers. *The Journal of Physical Chemistry C* **2010**, 114 (48), 20443–20448.

89. Chang, S.; He, J.; Kibel, A.; Lee, M.; Sankey, O.; Zhang, P.; Lindsay, S., Tunnelling readout of hydrogen-bonding-based recognition. *Nature Nanotechnology* **2009**, 297-301.
90. He, J.; Forzani, E.; Nagahara, L.; Tao, N.; Lindsay, S., Charge transport in mesoscopic conducting polymer wires. *Journal of Physics-Condensed Matter* **2008**, -.
91. He, J.; Lin, L.; Zhang, P.; Lindsay, S., Identification of DNA basepairing via tunnel-current decay. *Nano Letters* **2007**, 3854-3858.
92. Li, X.; He, J.; Hihath, J.; Xu, B.; Lindsay, S.; Tao, N., Conductance of single alkanedithiols: Conduction mechanism and effect of molecule-electrode contacts. *Journal of the American Chemical Society* **2006**, 2135-2141.
93. Cui, X.; Primak, A.; Zarate, X.; Tomfohr, J.; Sankey, O.; Moore, A.; Moore, T.; Gust, D.; Harris, G.; Lindsay, S., Reproducible measurement of single-molecule conductivity. *Science* **2001**, 571-574.
94. Lindsay, S.; He, J.; Sankey, O.; Hapala, P.; Jelinek, P.; Zhang, P.; Chang, S.; Huang, S., Recognition tunneling. *Nanotechnology* **2010**, -.
95. CRICK, F.; WATSON, J., THE COMPLEMENTARY STRUCTURE OF DEOXYRIBONUCLEIC ACID. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences* **1954**, 80-&.
96. Zwolak, M.; Di Ventra, M., Electronic signature of DNA nucleotides via transverse transport. *Nano Letters* **2005**, 421-424.
97. Zwolak, M.; Di Ventra, M., Colloquium: Physical approaches to DNA sequencing and detection. *Reviews of Modern Physics* **2008**, 141-165.
98. Venkataraman, L.; Klare, J.; Nuckolls, C.; Hybertsen, M.; Steigerwald, M., Dependence of single-molecule junction conductance on molecular conformation. *Nature* **2006**, 904-907.
99. Venkataraman, L.; Klare, J.; Tam, I.; Nuckolls, C.; Hybertsen, M.; Steigerwald, M., Single-molecule circuits with well-defined molecular conductance. *Nano Letters* **2006**, 458-462.

100. Venkataraman, L.; Park, Y.; Whalley, A.; Nuckolls, C.; Hybertsen, M.; Steigerwald, M., Electronics and chemistry: Varying single-molecule junction conductance using chemical substituents. *Nano Letters* **2007**, 502-506.
101. Xiao, X.; Xu, B.; Tao, N., Conductance titration of single-peptide molecules. *Journal of the American Chemical Society* **2004**, 5370-5371.
102. Ramachandran, G.; Hopson, T.; Rawlett, A.; Nagahara, L.; Primak, A.; Lindsay, S., A bond-fluctuation mechanism for stochastic switching in wired molecules. *Science* **2003**, 1413-1416.
103. Zheng, T.; Burkart, M.; Richardson, D., A general and mild synthesis of thioesters and thiols from halides. *Tetrahedron Letters* **1999**, 603-606.
104. Henningfeld, K.; Arslan, T.; Hecht, S., Alteration of DNA primary structure by DNA topoisomerase I. Isolation of the covalent topoisomerase I-DNA binary complex in enzymatically competent form. *Journal of the American Chemical Society* **1996**, 11701-11714.
105. Yelm, K., A simple method for in situ generation of thiols from thioacetates. *Tetrahedron Letters* **1999**, 1101-1102.
106. He, J.; Lin, L.; Zhang, P.; Spadola, Q.; Xi, Z.; Fu, Q.; Lindsay, S., Transverse tunneling through DNA hydrogen bonded to an electrode. *Nano Letters* **2008**, 2530-2534.
107. Kimura-Suda, H.; Petrovykh, D.; Tarlov, M.; Whitman, L., Base-dependent competitive adsorption of single-stranded DNA on gold. *Journal of the American Chemical Society* **2003**, 9014-9015.
108. Ostblom, M.; Liedberg, B.; Demers, L.; Mirkin, C., On the structure and desorption dynamics of DNA bases adsorbed on gold: A temperature-programmed study. *Journal of Physical Chemistry B* **2005**, 15150-15160.
109. Love, J.; Estroff, L.; Kriebel, J.; Nuzzo, R.; Whitesides, G., Self-assembled monolayers of thiolates on metals as a form of nanotechnology. *Chemical Reviews* **2005**, 1103-1169.

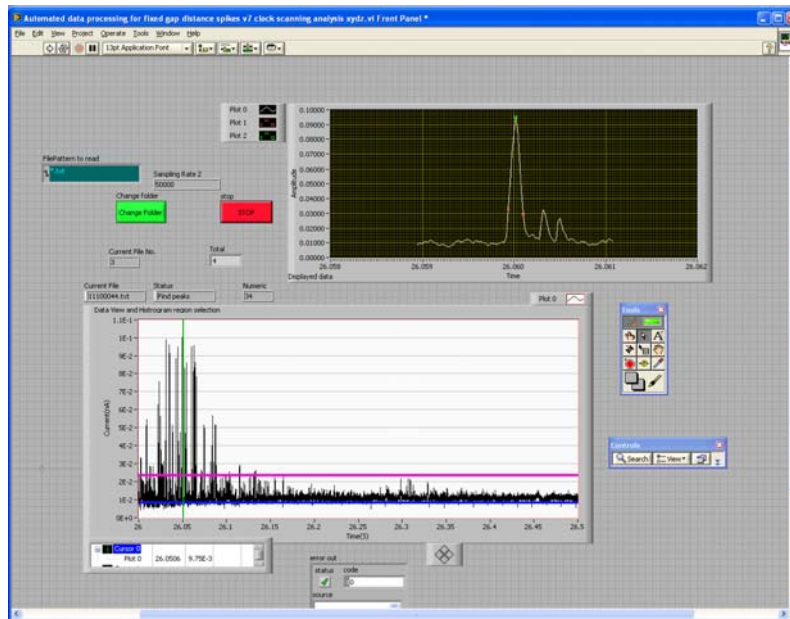
110. NAGAHARA, L.; LINDSAY, S.; THUNDAT, T.; KNIPPING, U., TIP-BIAS INDUCED SURFACE MODIFICATION ON GOLD SURFACES. *Journal of Microscopy-Oxford* **1988**, 145-147.
111. Tanaka, H.; Kawai, T., Partial sequencing of a single DNA molecule with a scanning tunnelling microscope. *Nature Nanotechnology* **2009**, 518-522.
112. Fuhrmann, A.; Ros, R., Single-molecule force spectroscopy: a method for quantitative analysis of ligand-receptor interactions. *Nanomedicine* **2010**, 657-666.
113. Fuhrmann, A.; Schoening, J.; Getfert, S.; Eckel, R.; Reimann, P.; Anselmetti, D.; Staiger, D.; Ros, R., Single molecule Protein-RNA interactions. *Biophysical Journal* **2007**, 168A-168A.
114. Fuhrmann, A.; Anselmetti, D.; Ros, R.; Getfert, S.; Reimann, P., Refined procedure of evaluating experimental single-molecule force spectroscopy data. *Physical Review E* **2008**, -.
115. Fuhrmann, A.; Schoening, J.; Anselmetti, D.; Staiger, D.; Ros, R., Quantitative Analysis of Single-Molecule RNA-Protein Interaction. *Biophysical Journal* **2009**, 5030-5039.
116. Liang, F.; Zhang, P.; Lindsay, S., Unpublished data.
117. Rincon-Restrepo, M.; Milthallova, E.; Bayley, H.; Maglia, G., Controlled Translocation of Individual DNA Molecules through Protein Nanopores with Engineered Molecular Brakes. *Nano Letters* **2011**, 746-750.
118. Hall, A.; Scott, A.; Rotem, D.; Mehta, K.; Bayley, H.; Dekker, C., Hybrid pore formation by directed insertion of alpha-haemolysin into solid-state nanopores. *Nature Nanotechnology* **2010**, 874-877.

APPENDIX A  
AUTOMATED DATA ANALYSIS

The measured tunneling signals in this thesis are either in the form of telegraphic switching or random spikes. Both of these signals could be analyzed by an automated and custom coded Labview program.

The typical data analysis environment is shown in Figure A.1. The X-Y graph in the middle of the window is a general view of the imported data with X axis as the time scale and Y axis as the current scale. The peak detection process is completely automated. During the detection, the baseline current  $I_{bl}$  is acquired by getting the fitted Gaussian peak of the lower part of the data. The blue cursor indicates the level of the baseline and tracks the baseline change during the detection. The purple cursor above the baseline is a peak detection threshold which is set at the level of  $I_{bl} + 2 \text{ sigma}$  of the  $I_{bl}$  so that the threshold is always higher than the noise level thus no white noise can counted. The green vertical cursor indicates the peak current been detected at a specific moment and it keeps tracking the detected current from left to the right.





**Figure A.1** The Labview Based Peak Detector Program. The middle part is a general view of the imported I-t time traces normally in the length of 0.5 msec. The top-right graph indicates a zoomed in region around a peak detected by the green cursor in the general view box.

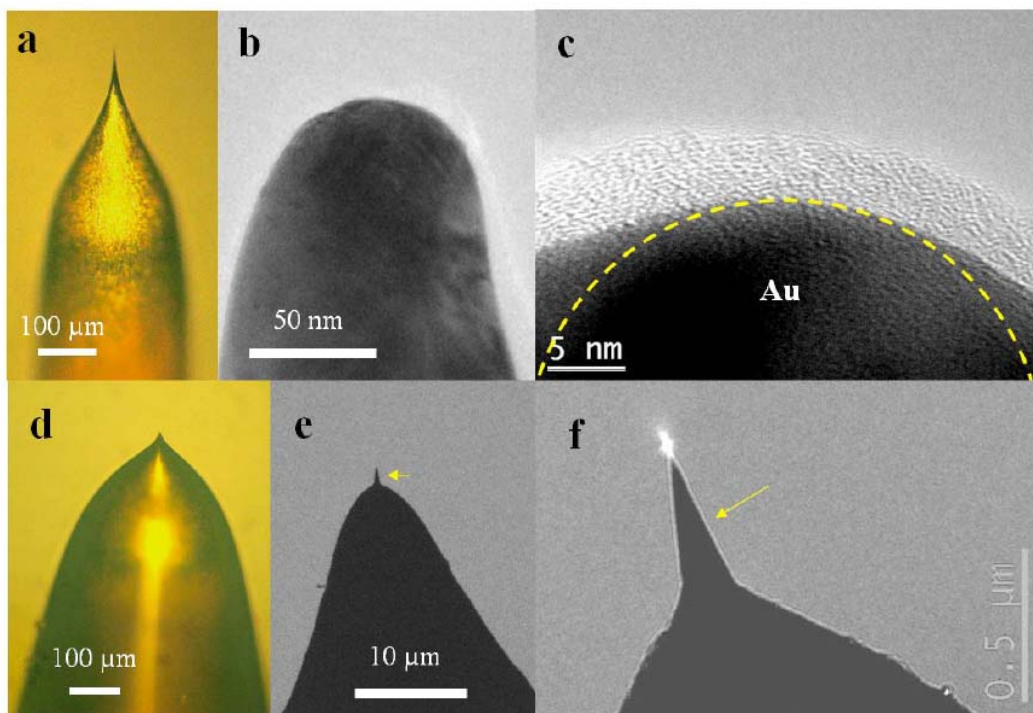
In the top-right part of the window, there is another X-Y graph showing a zoomed in image of a single peak indicated by the green vertical cursor in the general view window. In this graph, both the peak height which is defined to be the highest green point above the threshold and the peak threshold which is indicated as the time scale difference between the 2 red dots are shown simultaneously with the captured signal.

The peak information  $T_{on}$   $T_{off}$ ,  $\Delta I$  are finally listed in a table for downstream data analysis.

APPENDIX B  
TIP MAKING AND INSULATING

Due to the physical nature of the electron tunneling, the current drops 10 fold with every 1 angstrom opening of the gap size. It is thus believed and tested that atomic resolution imaging quality could be achieved even with a “blunt looking” tip. However, a “sharp looking” tip (see Figure B.1a) shows significantly higher yield for single molecule switching than blunt ones. Thus the tip quality is a critical factor of a successful experiment.

The STM tip could be easily made by a homemade electrochemical etching station. Gold wires (Aeser 99.999%) are first cut into small pieces (3-4 cm in length) and electrochemically etched using a mixture of ethanol and HCl (volume ratio=1:1). The sharpness of the tip is judged by an optical microscope with 300X magnification. A typical sharp-looking tip is shown in Figure B.1a. However, a “sharp looking” tip apex is not perfectly sharp under the Transmission Electron Microscope (TEM) as shown in Figure B.1b and B.1c. The radius curvatures for the best tips range from 5 to 20 nm. It is surprising that a tip like this could deliver measurement resolution in single base resolution (2-3 Å).



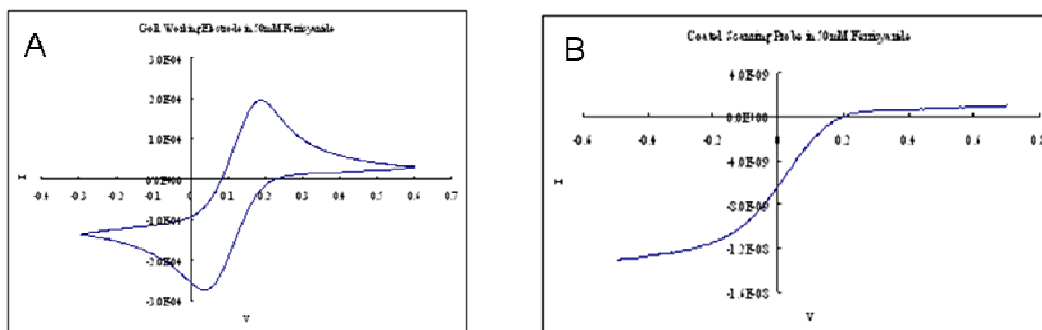
**Figure B.1** The optical microscope and TEM image of bare (a-c) and HDPE coated tips (d-f). (c) the carbon layer (the white layer covering the gold tip) was deposited during TEM imaging. The yellow dashed arc has a radius of 16 nm. (e-f) the yellow arrows indicate the location of exposed gold tip apex.

One critical factor for successful measurement in aqueous buffered solution is a well made High Density Polyethylene (HDPE) coated tip (see Figure B.1d). Without a well coated tip, the electrochemical noise will dominate the detection and stop the approaching of the tip.

Prior to the insulation, the gold tip needs to be cleaned with piranha (mixture of hydrogen peroxide and sulfuric acid with 1:3 volume ratios. Caution: In reaction with organic materials will cause severe explosion.) to get rid of organic contaminations and rinsed with DI water, ethanol and air-dried under compressed nitrogen gas. To perform the insulation, HDPE is

first melted under 250 °C on a homemade HDPE coating holder. The tip is mechanically approached and pierced into the melted HDPE back and forth for several times to fully coat the tip. The tip apex is always exposed (see Figure B.1e and B.1f) due to the surface extension force difference on the apex.

The success of the insulation is normally judged by optical microscopy. The estimation of the exposed area is done by cyclic voltammetry (CV) in potassium ferricyanide (see **Figure B.2**). The area covered by the CV curve significantly shrinks due to the decrease in the exposed area.



**Figure B.2** The Cyclic Voltammetry of (A) bare and (B) HDPE coated tip. Assuming a hemispherical exposed tip shape and using the formula  $I_{\max} = 2\pi RnFCD$ , the typical exposed area of the coated scanning probe is on the order of  $10^{-2} \mu\text{m}^2$

## BIOGRAPHICAL SKETCH

Shuo Huang was born on June, 22<sup>nd</sup>, 1984 in Nanjing, China and is the only child in the family. His father, a professor and also a senior engineer in Nanjing University has influenced him to have interest in natural science. Being a little boy, Shuo has showed strong interest in the fascinating discoveries of natural world and showed talented thoughts in discovering the unknown. In the year 1996, Shuo entered “the attached middle school of Nanjing University” and finished the middle school study in 2002. In 2002, Shuo entered Nanjing University majored in condensed matter physics as a bachelor student. In 2006, Shuo applied the Ph.D program of physics in Arizona State University and got admitted to pursue his Ph.D degree in Biophysics. In the first year of study, Shuo has maintained the highest GPA in the department and showed well evaluated talent in the teaching assignment for PHY 101 lab section. He was awarded the Molecular Imaging best graduate student award for that year. He started working on the \$1000 genome project in 2007 and has participated in the development of the recognition tunneling technique for DNA identification proposed by Dr. Stuart Lindsay. The major research results have been published in the Journal of Physical Chemistry C, Nano Letters and Nature Nanotechnology. The work which showed single base resolution of DNA detection in aqueous buffered solution has been selected as the cover story of the December issue of 2010 in the journal of Nature Nanotechnology. He is graduating in May, 2011 and is going to be a postdoctoral researcher in Hagan Bayley’s group in University of Oxford, United Kingdom to keep doing the research on the next generation DNA sequencing study by protein nanopore.