DNA sequencing by recognition tunnelling

by

Shuai Chang

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

Approved April 2012 by the
Graduate Supervisory Committee:

Stuart Lindsay, Chair
Robert Ros
Peiming Zhang
Nongjian Tao
John Shumway

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

Single molecules in a tunnel junction can now be interrogated reliably using chemically-functionalized electrodes. Monitoring stochastic bonding fluctuations between a ligand bound to one electrode and its target bound to a second electrode ("tethered molecule-pair" configuration) gives insight into the nature of the intermolecular bonding at a single molecule-pair level, and defines the requirements for reproducible tunneling data. Importantly, at large tunnel gaps, there exists a regime for many molecules in which the tunneling is influenced more by the chemical identity of the molecules than by variability in the molecule-metal contact. Functionalizing a pair of electrodes with recognition reagents (the "free analyte" configuration) can generate a distinct tunneling signal when an analyte molecule is trapped in the gap. This opens up a new interface between chemistry and electronics with immediate implications for rapid sequencing of single DNA molecules.

DEDICATION

For my parents, wife, and daughter

ACKNOWLEDGMENTS

TABLE OF CONTENTS

v

LIST OF TABLES

LIST OF FIGURES

Chapter 1

TUNNELING IN MOLECULAR JUNCTIONS

The work described in this thesis depends on the quantum mechanical phenomena

of electron tunneling. In this chapter, I first introduce a simple explanation of

vacuum tunneling, then described tunneling through a tunnel junction containing

a molecule, and finally introduce our invention, which we call "Recognition

Tunneling" a method for chemical recognition in a tunnel junction. Much of the

material in this chapter is reproduced from our research of nanotechnology.

Single molecules in a tunnel junction can now be interrogated reliably using

chemically-functionalized electrodes. Monitoring stochastic bonding fluctuations

between a ligand bound to one electrode and its target bound to a second electrode

("tethered molecule-pair" configuration) gives insight into the nature of the

intermolecular bonding at a single molecule-pair level, and defines the

requirements for reproducible tunneling data. Importantly, at large tunnel gaps,

there exists a regime for many molecules in which the tunneling is influenced

more by the chemical identity of the molecules than by variability in the

molecule-metal contact.   Functionalizing a pair of electrodes with recognition

reagents (the "free analyte" configuration) can generate a distinct tunneling signal

when an analyte molecule is trapped in the gap. This opens up a new interface

between chemistry and electronics with immediate implications for rapid

sequencing of single DNA molecules.

1.1 DNA sequencing

The study of DNA sequencing can be traced back to 1970s when the 1st 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.[1, 2]

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 because it lacks a 3'-hydroxyl group. The ddNTPs (ddATP, ddCTP, ddGTP, ddTTP) can also be radioactively or fluorescently labeled.

The ability of the 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. The quality of the gel results corresponding to the first 15-40 nucleotides is reasonable, however, after 700-900 nuclotides the results are questionable and often blurry. [3] 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. 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.

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

sequencing program 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).

Nanopore sequencing is 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.[4] Several unique approaches have been reported to achieve this goal,[5-9] 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 [7, 10-12] or MspA) and solid state nanopores (Si_3N_4, Graphene [2, 5, 6], Carbon nanotube [7] 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.

3

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 reproducability.

1.2 Tunneling in molecules systems
Let's consider what happens when a ''free" particle hits an "impenetrable barrier".

The spatial part of the wavefunction on the left, where E>V, is the sinusoidal

wave. At the interface, the potential jumps up to a value V>E, so that, according

to classical mechanics, the electron does not have the energy to cross. This is a

classically forbidden region for an electron of energy less than V. However, the

Schrodinger equation does not allow the wavefunction to go to zero abruptly. A

discontinuity in the wave function would lead to an infinite derivative, thereby

leading to an infinite energy [6].

I will begin with a brief overview of the phenomenon of electron tunneling (for

more details, see recent reviews of single molecule junctions[1-4]). Figure 1.2A

shows a pair of metal electrodes, separated by a gap, L. The electrons that

transport current lie near the highest occupied state at the Fermi energy, E_F. The

potential barrier that retains electrons within the metal is V volts so the work

function, ϕ, is given by $\phi = V - E_F$. In the absence of a state in the gap (i.e.,

without the state represented by the thick line ΔE above the Fermi energy) the

tunnel conductance is given approximately by [1]:

$$G \approx G_0 \exp\left(-1.02\sqrt{\phi}L\right) \equiv G_0 \exp(-\beta L) \tag{1.1}$$

where G_0 is the quantum of conductance, $\dfrac{2e^2}{h}$=77.5 µS, ϕ is the work function in units of eV and L is the tunnel gap in Å. Thus, when an atomically-sharp point contact is made, the conductance is approximately 77.5 µS, falling by a factor of 10 for each Ångstrom that the gap is opened for a metal like gold ($\phi \approx 5eV$). This very rapid decay with distance is what makes tunneling such a localized probe and therefore capable of reading single bases in a DNA polymer.
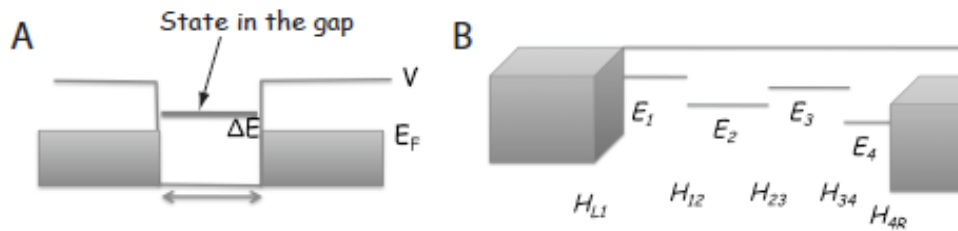
Figure 1-1Factors that control the tunneling signal:

(A) A simple tunnel barrier shown as a 1D structure with a gap where the potential (V) exceeds the Fermi energy (E_F). This potential barrier, V-E_F, can be lowered to a value ΔE by the presence of an atom in the gap with an eigenstate at E_F+ΔE. The extension of this picture to a molecule in the gap is shown in (B) where each atom contributes a level near the gap E_1, E_2 …. and overlap between the atomic states (H_{mn}) leads to a delocalized state that connects the left and right electrodes. This mediates a current proportional to the number of available states on the positive electrode (i.e., proportional to V_{bias}) and the strength of the coupling between the electrodes, as given by the Green's function propagator.

If an electronic eigenstate exists in the gap, then the effective barrier is reduced to approximately ΔE (thick solid line in Figure 1.1A). β, the inverse electronic decay length is then given by $1.02\sqrt{\Delta E}$ if this state connects the left and right electrodes. A real molecule spanning the gap will consist of molecular orbitals often well described in terms of linear combinations of the atomic orbitals of the constituent atoms and the hopping matrix elements between adjacent orbitals. Figure 1.1B is a schematic representation of this situation showing atomic states closest to the Fermi energy (E_1 - E_4) with hopping matrix elements H_{L1}, H_{12}....H_{34}, H_{4R}. Thus, the states of the system, ψ_n, are molecular orbitals, as modified by interactions with the left (H_{L1}) and right (H_{4R}) electrodes. The transmission is calculated with a Green's function:

$$G(E) \sim \sum_n \frac{\langle L|\psi_n\rangle\langle\psi_n|R\rangle}{E - E_{0n}} \qquad (1.2)$$

Here, L and R represent states at the energy E on the left and right electrodes and E_{0n} is the eigenenergy of the nth eigenstate of the system. The overall conductance is calculated from $\dfrac{2e^2}{h}\displaystyle\int_0^v G^2(E)dE$. The key point here is that the states that mediate tunneling are those that overlap *both* the left and right electrodes. Any "break" in the quantum coupling between the electrodes will not allow the two terms in the numerator of equation 2 to be non-zero simultaneously. This has two consequences. The first is that molecules must be coupled to the electrodes, and the most robust way to do this direct chemical bonding between the ends of the molecule and the electrodes.[5] The second consequence is that

6

the coupling strengths throughout the system (the H_{mn}) must be large enough not to be disrupted by thermal fluctuations.

The discussion above appears to imply that $\beta \to 0$ as $\Delta E \to 0$. In solid-state systems, this is one condition for resonant tunneling (the other being tight coupling to the electrodes[6]). In polarizable molecular systems the corresponding process is electrochemical reduction, the charge being trapped on an acceptor level by molecular and solvent reorganization. This is generally an undesirable outcome: Many redox processes are irreversible as a consequence of the generation of reactive species that, in turn, can alter the readout system itself by reacting with recognition molecules. Secondly, redox processes on single molecules can be quite heterogeneous if the environment is heterogeneous.[7]

Given that redox processes are to be avoided, how close to the Fermi level can a state of the combined electrode-molecule-electrode system be? There are two ways to address this, both leading to the same answer. Acceptor states on the molecule will be distributed over a range of energies comparable to thermal energy, $k_B T$. Thus we require $|E - E_{0n}| > k_B T$ to avoid redox processes. However, we also require that the 'weakest link' in the system, $H_{mn}^{weakest} > k_B T$ for quantum coherence to be maintained across the gap. Since the smallest energy splitting ($2\Delta E$ if we assume the Fermi level is in the middle of the molecular gap so that the closest orbital is shifted by ΔE from the Fermi level) is approximately equal to $2H_{mn}^{weakest} = 2\Delta E$, we conclude that we require $\Delta E \gg k_B T$. Thus β cannot be zero.

The foregoing discussion allows us to estimate the largest distance over which single molecule recognition tunneling reads may be made as follows: Let us require $\beta \geq \sqrt{10k_B T} \geq 0.5$ Å^{-1}. Let us assume a minimum read time of 1ms and that our electronics has an input capacitance of 1 pF (it is difficult to do better than this). The current-to-voltage conversion resistor cannot be greater than 1 GΩ (so that RC = 1ms) so the Johnson noise (given by $i_{RMS} = \sqrt{\dfrac{4k_B T \Delta f}{R}}$) in a bandwidth, $\Delta f = 1$ kHz is about 0.1 pA. Assuming a maximum applied bias of 0.5 V (less than this will often be required to avoid redox processes) the smallest useful gap conductance is then 0.2 pS. According to equation 1, this corresponds to a value of $\beta L = 19.7$ which, with $\beta \geq 0.5$ Å^{-1} yields a maximum value for L of about 4 nm. This is the maximum possible value with energy levels in just the right place with respect to the Fermi level. In less favorable situations, this maximum distance will be smaller. Thus Recognition Tunneling is restricted to combinations of analytes and reading reagents that are smaller than 4 nm. Taking the target to be no more than 3 nm diameter would limit the upper read size to proteins of no more than 10^4 Daltons (assuming that the electronic structure was just right).

1.3 Recognition tunneling

The following section of this thesis describes the technique of "Recognition tunneling" developed as a result of the current research. It is adapted from a review published by us in *nanotechnology*. [8]

Overview of recognition tunneling: Reading chemical identity at the single atom or single molecule level has long been a goal of the electron tunneling community. Recognition tunneling achieves that goal, serves as a remarkable probe of chemical bonds at the single molecule level and enables a new type of single molecule electronic sensor.

Atomic identification of sites on clean semiconductor surfaces in ultrahigh vacuum was achieved by atomic resolution tunneling spectroscopy in the mid 1980's[9]. It is difficult to develop an analogous technique for molecules on surfaces where both the nature of the mixing of contact and molecular states, and the degree of localization of "tunneling" electrons on the molecule are unresolved issues,[10] although it does appear to be possible to obtain distinctive tunneling spectra for some types of molecular residue in ultrahigh vacuum studies of molecular adsorbates.[11, 12] Outside of a UHV environment, the distribution of potential between a probe, molecule and underlying surface is completely unknown if the contacts between the electrodes and the molecule are not chemically well-defined, complicating both spectroscopy and even measurement of molecular conductivity.[3, 5] In an ingenious experiment, Tao showed that tunneling transmission could be gated by controlling the potential of a redox active molecule relative to a standard reference electrode.[13] This solves the problem of an unknown potential distribution, and single molecules can be identified, but only if they contain a redox-active species with a distinctive reduction or oxidation feature at an accessible potential. A completely new approach was introduced by Ohshiro and Umezawa in 2006.[14] They

demonstrated that STM images of molecular adsorbates showed chemical contrast if the scanning probe was functionalized with a molecule that bound to specific targets on the surface. In this work, a gold STM probe was functionalized with a thiol-derivatized DNA base, and images were made of monolayers of DNA bases, or PNA (a DNA-like molecule with a peptide backbone) bonded to a gold substrate. Enhanced contrast was found when the base on the probe was the Watson-Crick hydrogen bonding complement of the bases on the substrate. Analysis of the interaction between the functionalized STM probe and the target molecules shows that it is dominated by mechanical adhesion between the probe and target molecules, with essentially no contribution by a "chemically-sensitive" tunnel current.[15] It appears that the STM probe sticks to its complementary target for a longer time as the tip is scanned, resulting in more charge transfer between probe and substrate. Thus the tunnel current serves as a method for monitoring mechanical adhesion, and the method more closely resembles friction-force chemical sensing[16] or single-molecule force spectroscopy[17] as usually implemented with an atomic force microscope. Nonetheless, a variant of the Ohshiro-Umezawa method does permit identification of molecular species on a surface[18] and has been used to "read" the composition of DNA oligomers, albeit with a resolution limited to blocks of about ten bases, owing to the strong mechanical interactions between the probe and surface.[19]
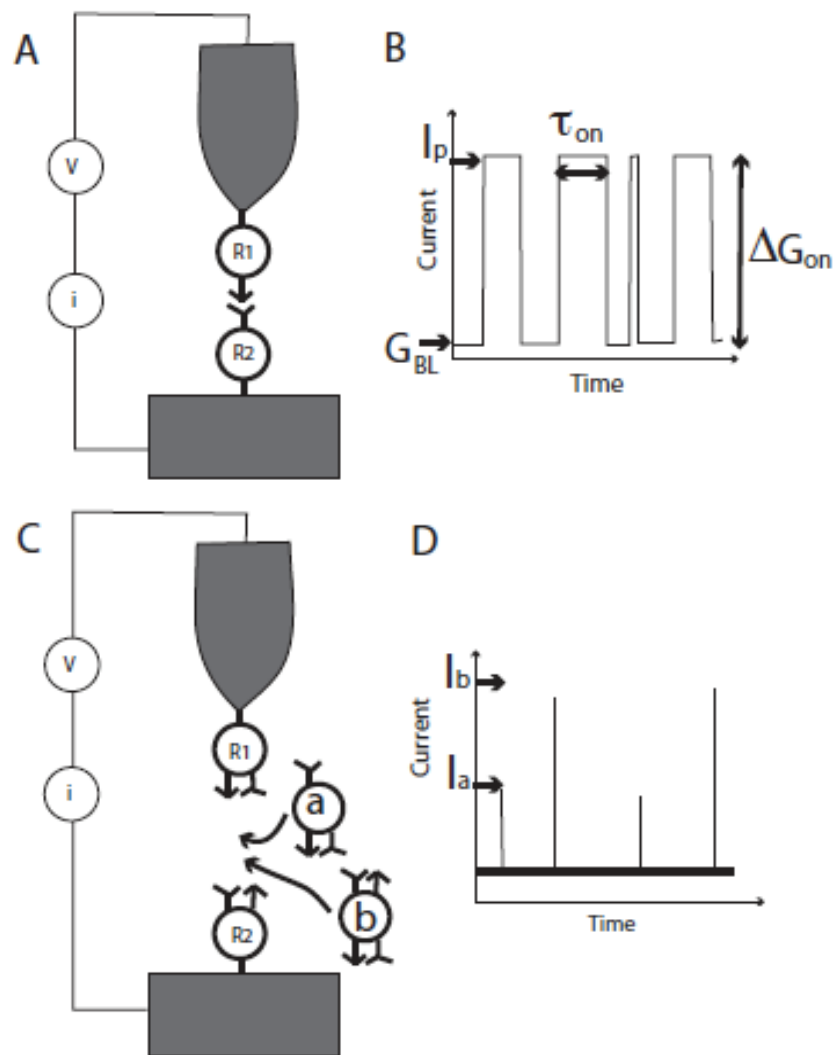
Figure 1-2 Two configurations for Recognition Tunneling.

In (A), the tethered molecule-pair junction, the current is recorded as a probe functionalized with a

first recognition reagent (R1) is held above a surface functionalized with its bonding

partner (R2). (B) Current recorded as a function of time shows switching fluctuations

("telegraph noise") as the junction between R1 and R2 breaks and remakes. The size of

the gap is measured by the "open state current" I_0, which corresponds to a baseline

conductance for the junction given by $G_{BL} = I_0/V$ where V is the junction bias. The

change in conductance when the molecule binds to yield a peak conductance ($G_p = I_P/V$),

$\Delta G_{ON} = G_P - G_{BL}$, is a measure of the conductance of a single molecule pair. The

11

lifetime of the "bound" state can be measured using the width of the individual jumps in current (τ). A second experimental configuration, the free-analyte configuration, is shown in (B). Here each electrode is functionalized with a reagent that presents recognition sites to a target (R1 and R2, but they could be the same reagent, for example one that presents a hydrogen bond donor and a hydrogen bond acceptor). The tunnel gap is set to a large value such that R1 and R2 do not interact with one another directly. Entry of an analyte into the gap ("a" or "b") causes a bonded pathway to be formed across the junction, leading to a "spike" in current through the junction (D). If the electron transmission of "a" and "b" differ significantly, their identity can be read directly from the size of the current spikes that are generated (I_a and I_b).

We have recently refined the method further, and it is now at the point where true single molecule reads are possible with the exquisite degree of recognition that comes from matching sets of hydrogen-bond donors and acceptors on probe and target.[20] Rather than move the probe relative to the substrate, we fix a tunnel gap and monitor the stochastic fluctuations in bonding as the target molecule binds, unbinds and rebinds the probe molecule (Figure 1.2A). Such stochastic switching has been observed before as a consequence of the lability of thiol contacts between a molecule and a gold surface,[21] and was used the basis of an ingenious method for measuring single-molecule conductivity introduced by Nichols and Haiss.[22, 23] This gives rise to a characteristic "telegraph" noise in the tunnel current as illustrated in Figure 1.2B. The gap is set to an initial "baseline" conductance, G_{BL}, the current increasing by an amount corresponding to the "on conductance", ΔG_{ON} when a molecular bridge spans the gap. In the

12

first implementation of "Recognition Tunneling" (which we call the "tethered molecule-pair" configuration, Figure 1.2A, B) we have measured the conductance of Watson-Crick basepairs using a base bonded to a probe and a nucleoside bonded to a surface, and monitoring the telegraph noise generated both by breaking of the hydrogen bonds and by breaking of the gold-thiol contacts used to hold the molecules to electrodes.[20] These measurements permit direct probing of bonds between single molecule pairs, measuring both their lifetimes and their electronic transport properties.

In a further refinement of Recognition Tunneling (which we call the "free analyte" configuration, Figure 1.2C), we have functionalized a pair of tunneling electrodes with molecules that present both hydrogen bond donors and acceptors, setting the gap to be large enough for target molecules to diffuse into it and bridge the sensing molecules on the surface to the sensing molecules on the probe via a network of hydrogen bonds. Differences in the pattern of hydrogen bonding lead to different levels of tunnel current as targets bind transiently in the gap (Figure 1.2D). These signals can be distinctive enough to allow the identity of the four DNA bases to be read with a confidence level that is greater than 0.6 on a single molecule read.[24] This technique could be adapted to read other small molecules, such as amino acid residues and small peptides.

1.4 Bond breaking, bond fluctuations and telegraph noise
It is probably an oversimplification to associate the duration of the "on" state (τ_{on} in Figure 1.2B) with bond lifetimes. The first observation of these bonding fluctuations by scanning tunneling microscopy [21] showed that the conductance

13

fluctuated while the molecule remained in place. This implies that small fluctuations in bonding can disrupt the electronic coupling essential for tunneling while leaving the molecule still "attached" to the surface. Despite this caveat, the distributions of "on" times clearly contain chemical information, as shown in Figure 1.3. Figure 1.3A shows distributions (left sides of the panels) for guanine interacting with deoxycytidine (G:C), 2-aminoadenine interacting with thymine (2AA:T) and adenine interacting with thymine (A:T – structures are shown in Figure 1.3B). The lifetime distributions are clearly bimodal, with one peak near 45 ms (S-Au) and another near 8 ms (HB). Contrasting these distributions to that measured for an octanedithiol (Figure 1.3C) shows that the 45 ms features are a consequence of fluctuations of the (S-Au)-Au bonds, while the 8 ms features must be associated with the hydrogen bonded complexes. The differences in lifetime between the two- and three-hydrogen bonded complexes is small (and the distributions are very wide) but the relative frequency of hydrogen bond breaking is clearly larger relative to (S-Au)-S breaking in the A:T complex than in the G:C and 2AA:T complexes.
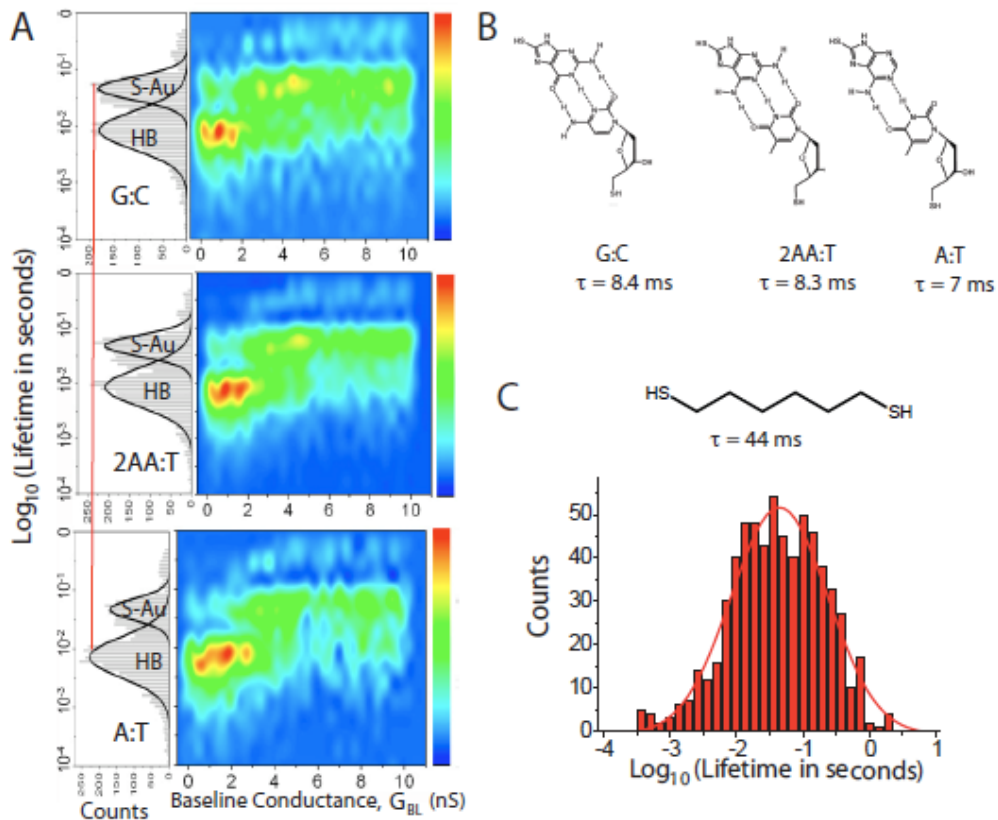
Figure 1-3 Lifetime distributions from telegraph noise.

(A) shows (left) distributions of the "on" state times for three types of DNA base-pairing (B) in a

tunnel gap. Two peaks are observed in the distribution. One is a slower process that

coincides with the peak is observed in junctions with only (S-Au)-Au as the labile bonds

(data for octane dithiol are shown in (C)). The faster process is somewhat dependent on

the nature of the hydrogen bonding, and is relatively more important in the A:T junctions

(two hydrogen bonds) than the 2AA:T and G:C junctions (three hydrogen bonds). The

distributions are broken out as a function of baseline conductance in the color plots to the

right. Hydrogen bond-breaking is more important in the large gap (small G_{BL}) regime

where the molecule is presumably stretched. (S-Au)-Au breaking dominates in smaller

gaps.

The lifetime histograms are resolved as a function of G_{BL} on the right-hand side of Figure 1.3A. Interestingly, there is little evidence of (S-Au)-Au bond breaking at small G_{BL} in the "non-contact" region. Hydrogen bond breaking dominates initially, but then falls off rapidly with increasing G_{BL}. At small gaps (S-Au)-Au bond breaking appears to be the dominant mode, possibly because the fact that (S-Au)-Au bond is an intrinsic property of the Au surface. In contrast, the hydrogen bonded components appear to require to require space to accommodate fluctuations. The intrinsic lifetime of these bonds in a confined space may be very long.

These data clearly demonstrate the ability of this technique to extract bonding information on a single molecule basis. Recent studies of the bonding lifetimes that use force spectroscopy to estimate the off-rate in nanojunctions suggest that the true off rate is very slow. If this is indeed that case, then the system remains "bonded" for long times, with the telegraph noise reporting local fluctuations that interrupt tunnel current, consistent with what was observed by scanning tunneling microscopy.[21] This could explain one of the experimental mysteries of Recognition Tunneling. The stability of the tunnel junctions is quite remarkable even in the absence of servo control, currents remaining stable for periods of up to ten seconds. Perhaps the junctions are held together by bonds that are stable on time scales that are much longer than the timescale of the telegraph noise fluctuations.

1.5 Experimental Methods

1.5.1 Characterization of molecule layer on metal surface

Molecule monolayers were formed by adsorption of thiolated molecules onto flame-annealed Au(111) substrates.[25] STM imaging showed that thiol-derivatized nucleosides normally form disordered monolayers. The thickness of the adlayers was measured by ellipsometry and found to be consistent with a monolayer. Fourier transform infrared spectroscopy (FTIR) was used to demonstrate the presence of the desired compounds and confirm an upright orientation on the gold substrates. The formation of appropriately hydrogen-bonded complexes on these surfaces was confirmed with surface plasmon resonance studies of the binding of hydrogen-bonded targets, and also with studies of the adhesion of functionalized atomic force microscope (AFM) probes.

1.5.2. Tip preparation and functionalization

STM tips were made from 0.25 mm diameter gold wire (purity 99.999% from Alfa Aesar) by either AC (30V, 4.2 KHz) or DC (2V) electrochemical etching in mixed solutions of concentrated HCl and ethanol (50:50, v/v). The gold tips were immersed in Piranha Solution (3:1 H_2SO_4:H_2O_2 - 30% by volume – *use caution as this solution is extremely reactive and will explode on contact with organics*) for 30 s, rinsed with DI water and dried in an N_2 stream prior to use or modification. A typical procedure for functionalization is as follows: tips were immersed in a solution of the target molecules for periods ranging from 2 hours to overnight,

17

then rinsed with clean solvent repeatedly and blown dry in an N_2 stream. Tips for measurements in aqueous solution were coated with wax[26] or high density polyethylene.[27]   High density polyethylene coatings are required when the tip functionalization procedure uses a solvent that dissolves wax insulation. We know of no analytical method sensitive enough to test for successful functionalization of the probes, so tunneling measurements are first carried out with unfunctionalized probes to provide a bench mark for recognizing unsuccessful funtionalization. We currently succeed in functionalizing the majority of the probes we prepare.

1.5.3. STM experiments

Measurements were carried out on a PicoSPM (Agilent, Chandler). The STM tip movement was controlled by custom labview programs and the data were recorded by a Yokogawa digital oscilloscope. 1,2,4-trichlorobenzene (TCB) and aqueous buffered electrolytes solution (0.1x PBS) were both used as working solutions. We performed i-z (STM break-junction) and i-t experiments (STM fixed junction) measurements. For i-z measurements, tunneling current (i) at fixed bias was recorded versus distance (z) while withdrawing tip from the surface in z direction.[15]

For i-t measurement, tunneling current (i) at fixed bias was recorded versus time (t) while holding tip at a fixed distance above the surface. We developed two methods for these "fixed gap" measurements: 1. Open loop without servo control.[20] 2. Closed loop with weak servo control.[24] For the open loop method, the junction is repeatedly set by making ever smaller steps in

18

the PZT voltage to avoid PZT creep.   The junctions are found to be stable (no detectable shift in the background current) over periods of several seconds, perhaps as a result of the bonding process discussed above.

For the "closed loop" method, the STM servo gains were set low so that only spikes of long duration were affected by the action of the current-control servo. We characterized the response time of the servo by recording thermal noise spectra with, and without servo control.[24] The closed-loop method does cause some distortion in signals from bound states of long duration, but it is much easier to implement.

Chapter 2

TUNNELING READOUT OF HYDROGEN-BONDING BASED RECOGNITION

This chapter describes the application of a break-junction version of recognition tunneling to hydrogen bonded interactions and was first published in the journal Nature nanotechnology [15]. The contribution of the current author was conducting of the STM measurements.

Hydrogen bonding plays a ubiquitous role in molecular recognition, DNA base-pairing being perhaps the best-known example.[28] Scanning-tunneling microscope (STM) images[14] and the decay of tunnel-current as a hydrogen-bonded junction is pulled apart[18] are sensitive to hydrogen-bonded interactions. Here, we show that these tunnel-decay signals can be used to count the number of hydrogen bonds in DNA basepairs. This sensitivity to hydrogen bonding arises predominantly from the molecular-mechanics of the junction. Junctions that are held together by three hydrogen bonds per basepair (e.g., guanine-cytosine interactions) remain intact for longer than junctions held together by two hydrogen bonds per basepair (e.g., adenine-thymine interactions). Similar, but less-pronounced, effects are observed on the approach of the tunneling probe, implying that hydrogen-bond dependent attractive forces also play a role in determining the rise of current. These effects provide new mechanisms for making sensors that transduce a molecular recognition event into an electronic signal.

2.1 Experimental

Electron tunneling *through an analyte molecule* may yield chemical

information via the dependence of current on bias[29] or on surface potential.[13]

Inelastic electron tunneling spectroscopy gives enhanced chemical selectivity[30]

but requires cryogenic temperatures. Given that hydrogen bonds enhance electron

tunneling rates over vacuum tunneling,[31] we have proposed that self-assembled

tunnel junctions based on hydrogen bonds will give good contacts[32] and

chemical selectivity simultaneously.[18]
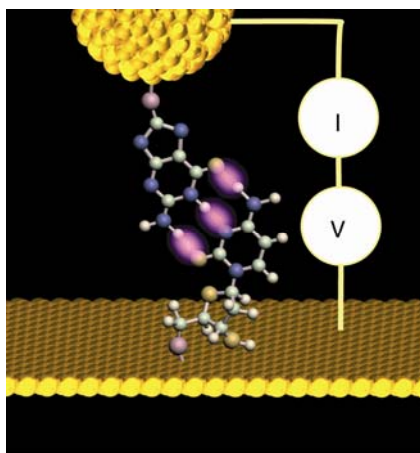
Figure 2-1 Illustrating the STM measurement.

A sharp gold probe, functionalized with a thiolated-base, is approached to a gold (111) surface

functionalized with a monolayer of a thiolated-nucleoside until the desired set-point

current is obtained, and then retracted while the tunnel current is recorded. The current-

distance curves can be used to count the number of hydrogen bonds in the interaction. All

measurements reported here were taken under 1,2,4-trichlorobenzene.

To demonstrate the feasibility of this approach, we functionalized a gold

STM probe with a DNA base that was brought into contact with a monolayer of

nucleosides on a gold surface under 1,2,4-trichlorobenzene (Figure 2.1 and see

21

Methods). We chose 1,2,4-trichlorobenzene as a solvent because of its low

dielectric constant (3.95), high boiling point (214 ˚C) and minimal disruption of

hydrogen bonds. A steady tunnel current set-point (I_{SP}) was established under

servo-control, the servo broken, and the current recorded as the probe was pulled

away from the surface. Current-decay curves for one, two and three hydrogen-

bond molecular junctions are shown in Figure 2.2.   Clearly, the signals from the

adenine (A) and thymidine (T) junctions (two hydrogen bonds - Figure 2.2b)

decay more rapidly than signals from guanine (G) and deoxycytidine (C)

junctions (three hydrogen-bonds - Figure 2.2d). Adding an extra hydrogen bond

to the A-T interaction by using 2-aminoadenine[33] increases the extent of the

signal (Figure 2.2c). Conversely, reducing the number of hydrogen bonds to one

using deoxycytidine as the target and 2-aminoadenine as the probe results in a

more rapid decay (Figure 2.2a). Current decays even more rapidly in solvent

alone. These curves are raw, unselected data with curves from several regions of

the substrate overlaid. Remarkably, the number of hydrogen bonds in an

interaction can often be determined from a single curve.
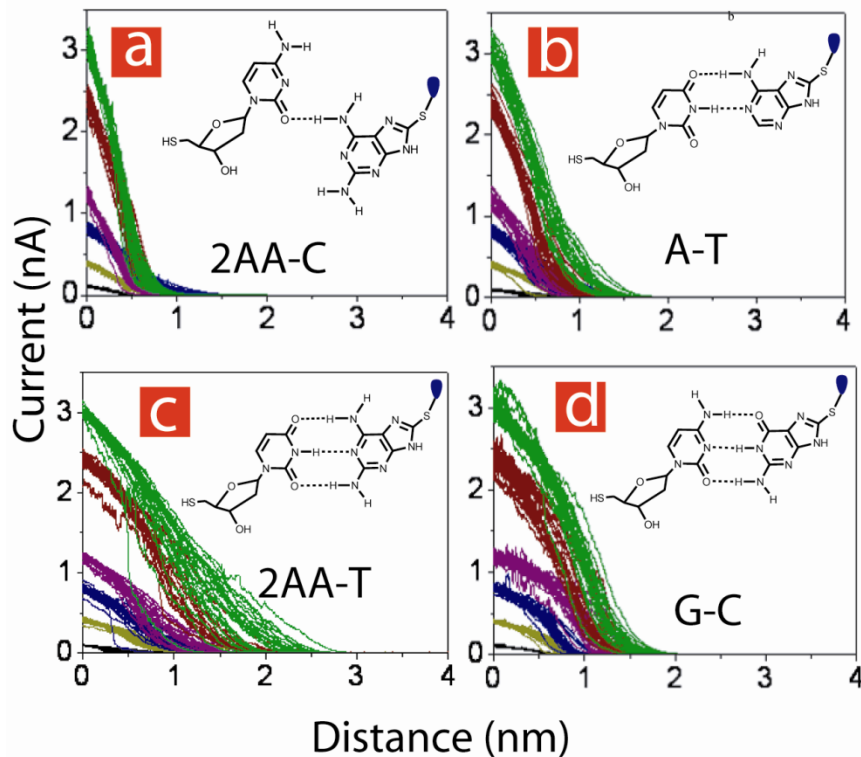
Figure 2-2 Tunnel-current decay curves obtained at a bias (V in Figure 1) of 0.4V as the probe was retracted from a monolayer of nucleosides at a rate of 100 nm/s.

Data are shown for a 2-amino-8-mercaptoadenine functionalized probe interacting with a deoxycytidine monolayer (a), an 8-mercaptoadenine probe interacting with a thymidine monolayer (b), a 2-amino-8-mercaptoadenine functionalized probe interacting with a thymidine monolayer (c) and an 8-mercaptoguanine probe interacting with a deoxycytidine monolayer (all nucleosides were 5'-thiolated). The initial set-point currents were 3 nA (green), 2.4 nA (brown), 1.2 nA (purple), 0.8 nA (blue), 0.4 nA (khaki) and 0.1 nA (black). Data are superimposed from several runs taken over different points on the substrate.

Integration is a simple way to characterize each decay curve. Using time as the horizontal axis, the integral gives the charge transferred in the interaction. This quantity is plotted vs. I_{SP} in Figure 2.3 for zero (triangles), one (dots), two (green dots) and three (red and orange dots) hydrogen bonds. Error bars show the standard deviation, indicating the reliability with which the number of hydrogen bonds can be counted from an individual curve. Linear fits to these data (not

shown) have gradients that are directly proportional to the number of hydrogen

bonds in the interaction, as shown in the inset in Figure 2.3.
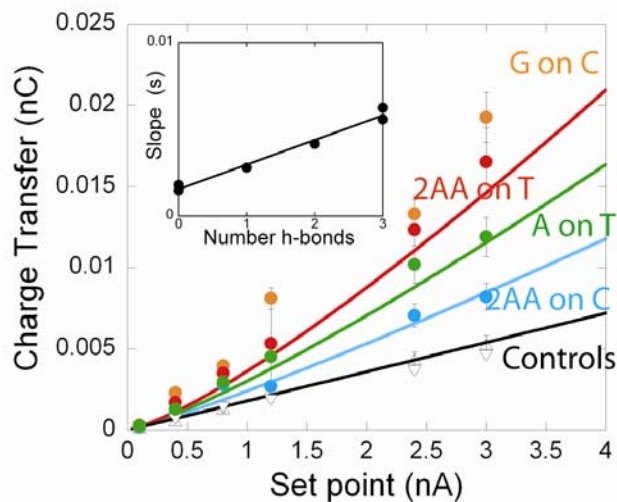
Figure 2-3 Charge transfer obtained from integration of the tunnel current decay curves plotted as a function of set-point current for zero (triangles), one (blue dots), two (green dots) and three (red (2AA-T) and orange (G-C) dots) hydrogen bonds.
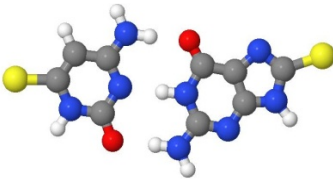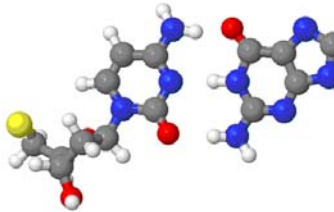
The two sets of controls correspond to a bare probe and a thio-phenol functionalized probe

interacting with a thymidine monolayer. The withdraw speed was 100 nm/s and the error

bars are ± 1sd. The gradient of each of these data sets fitted with a linear regression is an

accurate predictor of the number of hydrogen bonds (inset). The solid lines are calculated

according to equation 2.

2.2 Calculation and discussion

We investigated the conductance of both paired-bases and base-nucleoside

pairs theoretically. The tunneling current was computed using a density

functional theory (DFT) Green's function scattering method [34, 35] based on the

Landauer approach.[36, 37]   Semi-infinite gold electrodes were connected to

sulfur atoms (yellow in the illustration of guanine-cytosine and guanine-deoxycytidine pairs in Table 2.1). Further details are given. Calculated conductances (Table 2.1) are much larger for bases joined by three hydrogen bonds than for those joined by two. However, when one of the bases in the pair is replaced with a nucleoside (right column) the conductance is reduced substantially and is no longer sensitive to the number of hydrogen bonds joining the bases. How can this calculation be reconciled with the exquisite chemical sensitivity of the measurements?

Table 2.1 Conductances calculated with density-functional theory for hydrogen-bonded bases spanning a pair of gold electrodes (middle column) and for nucleosides hydrogen bonded to bases (the sulfur atoms that attach to the gold are shown in yellow for the C:G examples illustrated).

Base-pair (N_{HB}) [Base-nucleoside]	Conductance Base:Base (nS)	Conductance Nucleoside:base (nS)
C:G (3) (Cytidine:G)	83.8	0.96
T:2AA (3) (Deoxythymidine:2AA)	119	1.43
T:A (2) (Deoxythymidine:A)	22.2	1.62
T:G (2) (Deoxythymidine:G)	23.7	0.54

Several factors indicate that the signals are not purely electronic in origin: Firstly, the decay distance is much too long to correspond to an electronic process.[18] Secondly, the change in shape of the decay curves with I_{SP} is not consistent with a simple tunneling process. We illustrate this in Figure 2.4c with a set of averaged curves for each I_{SP}. If the decay followed the I_{SP} = 3nA curve with the data slid over for the lower set-points, it would lie on the dashed lines shown in Figure 2.4c. It clearly does not. Thirdly, there is considerable hysteresis in the data. On approach, the probe must be brought closer to the surface to restore a signal and its growth is much more rapid than the decay was on withdrawal.
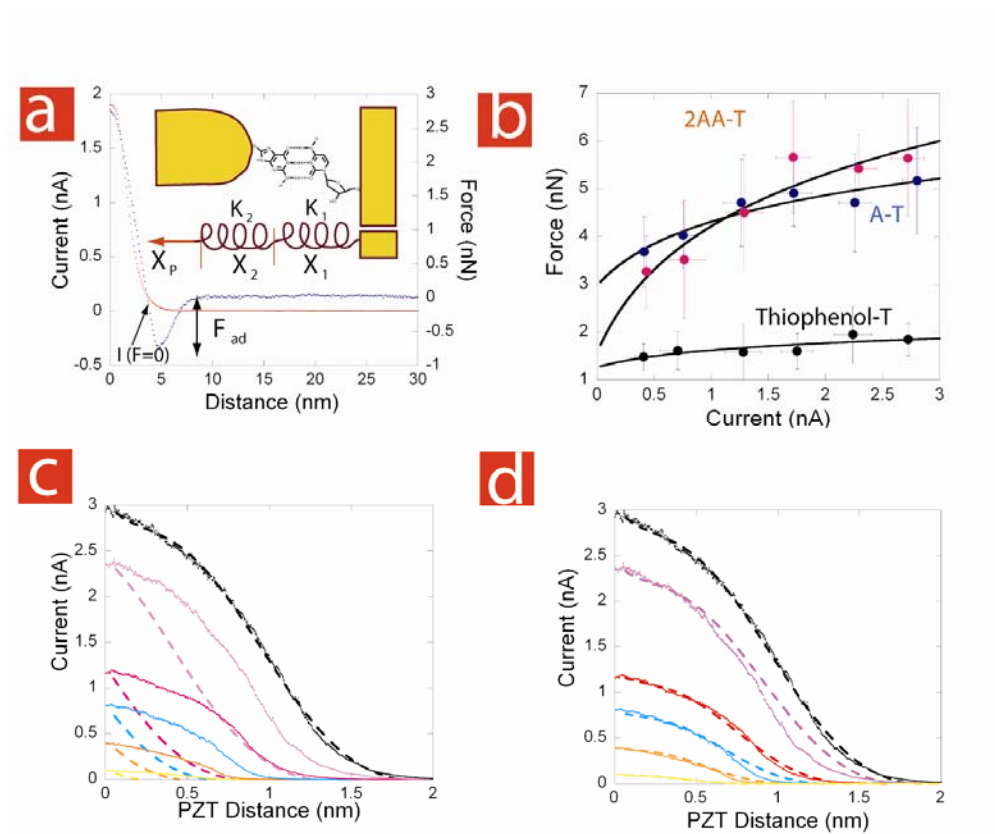
Figure 2-4 The decay of current with withdrawal distance

26

(a) Conducting AFM measurements of interaction force (blue points) and simultaneously acquired tunnel current (red points). The inset illustrates the elastic-distortion model. X_1 represents the stretching of the molecule, X_2 represents the stretching of the probe and X_P is the (measured) displacement of the far end of the probe. (b) Adhesion force plotted vs. current-set point for a 2-amino-8-mercaptoadenine functionalized probe interacting with a thymidine monlayer (red points) and a 8-mercaptoadenine probe interacting with a thymidine monolayer (blue points). The black dots are control data obtained with thio-phenol functionalized probes. The solid lines are fits to $F_{ad} \propto N_{HB} \ln(BI_{SP} + 1) + F_{NS}$ with B set to 2.83 nA^{-1}. The coefficient of the log terms are 1.97±0.3 (AA-T) and 1.0±0.16 (A-T), so their ratio is 1.96 (+0.7, -0.5) a range that spans the expected value of 1.5 ($\frac{N_{HB}(2AA-T)}{N_{HB}(A-T)} = \frac{3}{2}$). (c) Representative STM decay curves for an 8-mercaptoguanine probe interacting with a deoxycytidine monolayer for several set-points (each curve is the average of 26 raw data curves). The dashed-lines show how the current should decay from each set point if it follows the form of the 3nA data (black) as fitted with a 9th order polynomial (black dashed line) and slid over to match the other set-points. (b) Shows the same data with (dashed lines) the set-point current dependence predicted by the elastic distortion model described in the text.

The slow decay of current with withdrawal distance suggests that mechanical interactions play a role, a mechanism for long decay-distances first suggested by Pethica.[38] If the stiffness of the probe (K_2) is smaller than that of the molecular junction (K_1) that spans the gap (inset, Figure 2.4a) then adjustments of the probe position (X_P) result in smaller changes in the tunnel gap (X_1). This results in tunnel signals that appear to decay very slowly when plotted as a function of the external adjustment, X_P. We explored these mechanical

interactions using conducting atomic force microscopy (CAFM). The CAFM

probe was coated with gold and functionalized with either 8-mercaptoadenine (A)

or 2-amino-8-mercaptoadenine (2AA). It was brought into contact with a gold

surface functionalized with a 5'-thiolated thymidine monolayer in 1,2,4-

tricholorbenzene. I_{SP} was set using a custom controller and tip-deflection and

current were recorded as the probe was pulled away from the surface. A typical

current decay curve is shown by the red data points in Figure 4a, with the

simultaneously acquired interaction force shown by the blue points. With this

AFM cantilever, presumably much softer than an STM probe, the "tunneling"

signal now extends out to 4 nm, illustrating the role of the probe stiffness in

determining the extent of the signal. The adhesion force between probe and

surface can be read directly from these plots (F_{ad} in Figure 2.4a). Figure 2.4b

plots values of the adhesion force vs. the set-point current for a 2AA probe (3 H-

bonds, red points) and an A-probe (2 H-bonds, blue points). We used a thiophenol

functionalized probe as a non-hydrogen bonding control (black points).

Adhesion increases rapidly with set-point at low currents, but then more slowly as

the set-point current increases. Since the contact area (and hence the number of

molecular junctions between probe and surface) increases linearly with the

indentation of the probe[39] and the current increases exponentially with the

indentation, we expect that $F_{ad} \propto N_{HB} \ln(BI_{SP} + 1) + F_{NS}$ where N_{HB} is the number

of hydrogen bonds per intermolecular interaction, B is a constant and F_{NS} is a

non-specific adhesion force. This function fits the experimental data quite well

(solid lines). Thus the number of molecular contacts increases with the logarithm

28

of I_{SP}, so that force equilibrium in the elastic-interaction model (inset Figure 2.4a) requires that

$$X_1 = \frac{1}{1 + CN_{HB} \frac{K_1}{K_2} \ln(BI_{SP} + 1)} X_P \qquad (2.1)$$

where K_1 represents the "spring constant" of a single hydrogen bonded interaction, K_2 is the stiffness of the probe and $X_P = X_1 + X_2$ is the total displacement of the probe. Two parameters, B and $\frac{CK_1}{K_2}$, determine changes in the tunnel gap as a function of I_{SP} and the measured probe displacement. These parameters are readily obtained from a subset of the tunneling decay data in Figure 2.4. Using equation 1 to scale the measured displacement to correspond to that in the $I_{SP} = 3\,\mathrm{nA}$ curve leads to a set of predicted decays given by the dashed lines in Figure 2.4d. The agreement with experiment is good, showing that the mechanical interactions revealed by the conducting AFM account for the STM data also. The same model (and parameters) also account for the measured charge transfer values in Figure 3. We find that

$$Q = \frac{I_{SP}}{\beta v}\left(1 + \frac{CK_1}{K_2} N_{HB} \ln(BI_{SP} + 1)\right). \qquad (2.2)$$

Here $v = \dfrac{dX_P}{ds}$ is the retraction speed of the STM probe (100 nm/s) and β is the intrinsic decay constant (taken to be the 6 nm^{-1} measured in solvent). Calculated values of Q are shown by the solid lines in Figure 2.3 for $N_{HB} = 0$ (black), 1 (blue), 2 (green) and 3 (red). The agreement with the measured data shows that

this simple electro-mechanical model works remarkably well. Similar (but smaller) interactions must play a role in the approach curves as well.

How many molecules can be sensed with this method? We can estimate the number of molecules in the junction if we assume that each hydrogen bond contributes about 200 pN to the excess adhesion force over the control sample at the loading rate used here.[40] Thus rupturing 2AA-thymidine requires 600 pN while A-thymidine pairs require 400 pN. Subtracting the background signal (1.7 nN – Figure 2.4b) from the measured adhesion and dividing the remainder by 600 (2AA-T) or 400 pN (A-T) yields a range of 2 to 10 molecules in these junctions over the range of set points (0.1 to 3 nA). Dividing the current at zero force ($I(F=0)$, Figure 2.4a) by the estimated number of molecules in each junction yields an estimate of the conductance of each molecular pair as 100 to 300 pS. This is substantially lower than the calculated values (Table 2.1) but a factor 10 or so disagreement is not unusual in calculated and measured molecular junction tunnel conductances.[3]

In summary, we have demonstrated a method whereby the STM can be used to count the number of hydrogen bonds in few-molecule interactions, and have accounted for the withdrawal signals with an electro-mechanical model. Weaker attractive interactions, presumably owing to transient hydrogen-bonding, result in significant sensitivity in the approach curves also. These mechanisms provides a basis for new types of electronic biosensors and chemosenors based on hydrogen bonding, transducing molecular recognition directly into an electrical signal with a high degree of chemical specificity.

30

2.3 Methods

5'-S-acetyl-5'-thiodeoxycytidine and 5'-S-acetyl-5'-thiothymidine were synthesized as previously described, [18] deacetylated with pyrrolidine before use, and monolayers were prepared and characterized on freshly flame-annealed gold-on-mica substrates following the published procedures.[18] 0.25 mm gold wires were etched, cleaned and functionalized as previously described.[18] They were functionalized with 8-mercaptoguanine (G), 8-mercaptoadenine or 2-amino-8-mercatpto-adenine (2AA) used as received from Aldrich. The tips were rinsed with DMF and ethanol and blown dry in an N_2 stream before use. Tunneling measurements were carried out on a PicoSTM (Agilent, Chandler) with the sample and probe submerged in 1,2,4-trichlorobenzene. Approach and retraction curves were recorded using custom Labview software. Conducting AFM (PicoAFM, Agilent, Chandler) measurements used probes with a nominal spring constant of 0.35 N/m (Mikromasch) sputter coated with alternating layers of chrome and gold (25Å Cr, 50 Å Au, 25 Å Cr, 50 Å Au, 25 Å Cr, 100 Å Au, 25 Å Cr and 300 Å Au). The final probe radii, determined by SEM lay between 100 and 200 nm. Probes were immersed in 1 mM solutions of 8-mercaptoadenine or 2-amino-8-mercaptoadenine in DMF or 1 mM thiophenol in methanol (control experiments) for 2 to 12 h immediately after removal from the sputter-coater. Simultaneous force and conductance measurements were taken at a series of nominal current setpoints between 0.5 and 9 nA with a bias of 0.4V in 1,2,4-trichlorobenzene using thymidine monolayers and a probe retraction speed of

2000 nm/s. After taking measurements, each probe was calibrated using the

thermal-noise method.[41]

Chapter 3

TUNNEL CONDUCTANCE OF WATSON-CRICK NUCLEOSIDE-BASE PAIRS FROM TELEGRAPH NOISE

This chapter describes the application of telegraph noise based recognition tunneling to the measurement of hydrogen bonded systems. It was first published in Nanotechnology [20]. The contribution of the current author was conducting of the STM measurements.

The use of tunneling signals to sequence DNA requires knowledge of the absolute conductance of its components. Hydrogen-bonded DNA base-nucleoside pairs trapped in a gold tunnel junction with thiol bonds display conductance switching between two levels. The absolute values of the conductances determined from this telegraph noise are within a factor two of the predictions of density functional calculations with the relative conductances being in close agreement.

3.1 Introduction

Electron tunneling has been proposed as a sensitive local probe to identify individual bases in DNA, a possible basis for direct electronic sequencing.[29, 42] We have found that current-distance curves collected from junctions containing base-nucleoside pairs[18] and even intact DNA[19] faithfully report the base composition of the target. A careful analysis of these signals shows that they do not arise from single-molecule interactions.[43] Rather, the overall conductance of the tunnel gap is probably set by through-space tunneling across a large-area

junction containing several (two to ten) molecular pairs. Conducting-AFM measurements suggest that the conductance of a base-nucleoside pair is in line with the predictions of density-functional calculations and on the order of nS.[43] Can we measure these conductances directly? The established techniques for measurement of single molecule conductance[32, 44, 45] are difficult to apply to a system as complex as a base-nucleoside hydrogen bonded pair. Further, they do not report the conductance as measured in the type of fixed junction that would be required for DNA sequencing.[46] Stochastic switching of bonds between molecules and the electrodes of a tunnel junction offers another approach to measuring single molecule conductance. Ramachandran et al.[21] demonstrated that, for tunnel-junctions using gold electrodes with thiol-attachment chemistry, fluctuations in the molecule-metal contact result in the stochastic-switching of STM images of molecules embedded in a monolayer. Fluctuations in the C-Si bond were monitored in an STM[47] while the transient binding and unbinding of a carboxylate-EDC complex was monitored via conductance fluctuations in a carbon nanotube.[48] Haiss et al.[22] showed that the time course of bond-fluctuations in a gold-thiol-molecule-thiol-gold tunnel junction could be followed by the simple expedient of placing a gold STM probe above a gold surface funtionalized with bis-thiolated alkane molecules. As the molecules bind and unbind to span the gap, the tunnel current shows a distinctive two-level signal. The single-molecule conductance deduced from the amplitude of tunnel current yielded values for alkane thiols that were in good agreement with the break junction method.[22, 23, 44] In this chapter, we report on Telegraph-Noise

34

measurements made by forming a fixed gap between a probe functionalized with a DNA base and a surface funtionalized with nucleosides.

3.2 Tunneling in hydrogen bonded systems

Hydrogen bond strengths lie between those of covalent and Van der Waal's bonds.[49] With bond strengths of just a few times $k_B T$, hydrogen bonds make an ideal "molecular Velcro" for making transient contacts to analytes for applications like DNA sequencing. Can reasonable tunnel currents be sustained through such "weak" bonds? In fact, this is not the right way to ask the question, for the bond strength is related to the overall lowering of electronic energy in occupied states, whereas the tunneling rate will be determined by the energy difference between the Fermi energy of the metal and the state closest to it that spans the tunnel gap (c.f. equation 2). We need to determine, for a hydrogen bonded system, the β to use in equation 1.

One simple way to determine an upper limit on β is with complex band structure.[50] This technique uses standard (fast) codes for determining the band structure of a periodic system, but uses a complex wavevector, keeping track of the results for both the real and imaginary components. States with an imaginary wavevector are forbidden but the magnitude of the imaginary wavevector is the inverse decay constant, β, for that state (β being a real variable in a decaying wave-function). Assuming that the state that mediates tunneling at the Fermi energy lies halfway between the HOMO and the LUMO, the value of β can be read directly from the imaginary part of the complex band structure. A very simple (and chemically unrealistic) model of an H-bonded system is shown as the

line of planar water molecules on the left of Figure 3.1A. They are arranged in an

infinite lattice of lattice constant a, and the band structure calculated with local

orbitals is shown to the right.[51] The states with imaginary wavevector are

plotted to the right, and the value of β midway between the HOMO and LUMO is

0.99 Å^{-1} (9.9 nm^{-1}). This is very close to the value found in sigma-bonded

systems like alkane chains. Thus, this hydrogen bonded system, despite being

held together by "weak" hydrogen bonds, propagates electrons as well as an

alkane chain. This value of β is calculated for the whole system comprising a unit

cell of the lattice (Figure 3.1A). When the β strictly associated with the H-bond

itself is calculated from the conductance decay as the hydrogen bonds are

stretched, a much larger value (of about 3 Å^{-1}) is calculated.[51] Thus the current

should be very sensitive to straining of these bonds.
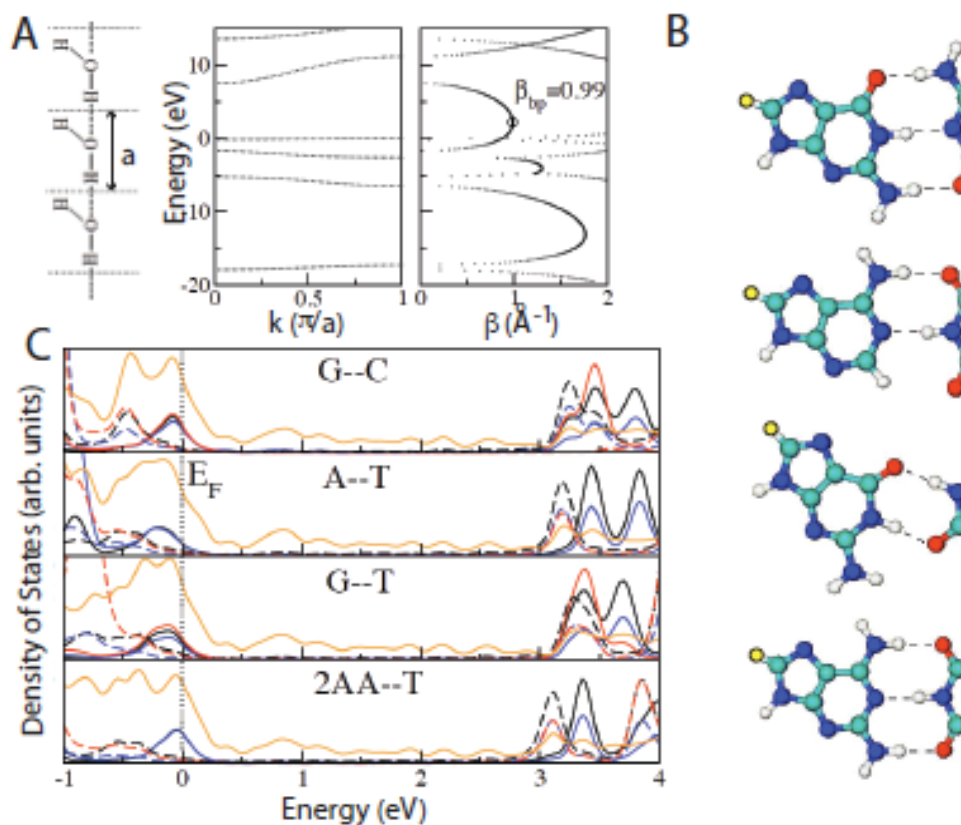
Figure 3-1 Hydrogen-bond mediated tunneling.

(A) A linear water chain model with one water molecule per unit cell. The H-O⋯H angle is 180° and the O⋯H distance is 0.197 nm (left). The panels on the right show the allowed (real k) and gap (imaginary k or real β) states calculated using complex band structure.

(B) Energy minimized configurations for the two Watson-Crick basepairs (G:C, A:T) a GT wobble base pare and a 2-aminoadenine:thymine (2AA:T) pair. Yellow atoms are sulfur atoms that connect to gold slabs. (C) Averaged projected densities of states (DOS) per atom for G--C, A--T, G--T, and 2AA--T base-pairs. The Fermi energy is defined to be zero energy. The projected DOS onto carbon, nitrogen, oxygen, and sulfur atoms are represented in black, blue, red, and orange colors. Solid and broken lines are the projected DOS onto atoms on purines (G, A, 2AA) and pyrimidines (C, T), respectively.

The HOMO is dominated by the orbitals on purines (guanine for G--C and G--T, adenine

for A--T, and 2AA for 2AA--T) and the LUMO is dominated by the orbitals on

pyrimidines (cytosine for G--C, thymine for G--T, A--T and 2AA--T).

In reality, the Fermi level alignment is rarely exactly halfway between the

HOMO and LUMO. Exact Green's function calculations are required for the

entire molecular system as connected to semi-infinite slab electrodes. Structures

for the two Watson-Crick basepairs, a G:T "wobble" base pair and a triply-

bonded 2-aminoadenine:thymine basepair are shown in Figure 3.1B. The yellow

atoms are the sulfur atoms that were place in a three-fold hollow on Au(111)

surfaces (all structures were relaxed prior to calculation of the Green's

function).[52] The calculated density of states for each type of atom (color coded

as explained in the Figure caption) is shown in Figure 3.1C. The solid lines are

projected for atoms in the purines and the dashed lines are for atoms in the

pyrimidines. The Fermi level lies close to the LUMO where states are dominated

by the pyrimidines. In consequence, the effective β for these molecular systems

is quite small and the calculated tunnel conductances quite high.[52]

3.2.1 Probing single molecule junctions with telegraph noise

Haiss and Nichols pioneered the monitoring of stochastic fluctuations in

molecular junctions as a method for determining the single molecule

conductance.[22, 23] In this technique, a molecule with two reactive end groups

(thiols) is trapped between two (gold) electrodes. Stochastic fluctuations of the

metal-molecule interface[21] result in a train of "on"-"off" switching events in the

tunnel current (Figure 3.2B). If a single molecule is trapped, then this current has

just two levels ("on" and "off") and so resembles the telegrapher's Morse code, hence "telegraph noise". Surprisingly, this method has not been widely adopted yet. It has some advantages over the more popular "break junction" methods.[44, 45] Firstly, the signal identifies junctions containing single molecules unambiguously. This information has to be inferred statistically from break junction data. Secondly, the junction remains static, avoiding the constant reforming that occurs with break junctions. And finally, because of the separation of the two electrodes, it is possible to study different molecules on the probe and substrate, essential for the "tethered molecule-pair" version of Recognition Tunneling. These would be rapidly scrambled in a break junction measurement.
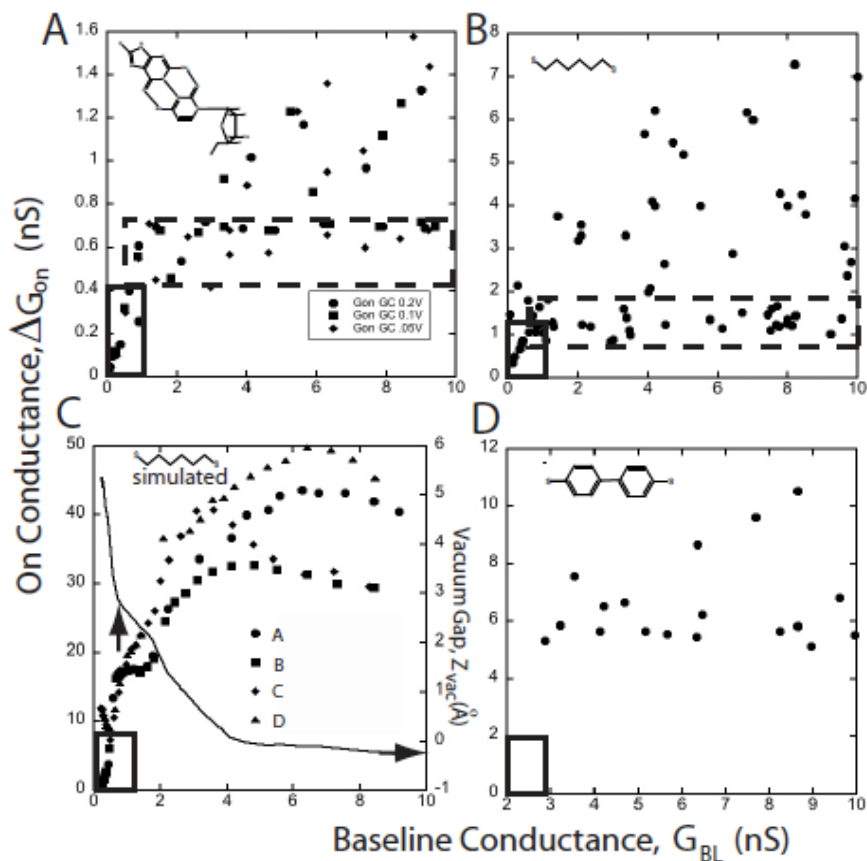
Figure 3-2Interpretation of scatter plots of the molecular conductance ( ) vs. the baseline conductance ( ).

(A) shows typical experimental data for a probe functionalized with a base (guanine) interacting

with a surface-bound nucleoside (deoxycytidine). The dashed box encloses a "plateau"

region where the majority of the data are relatively constant. As G_{BL} is increased, the

number of large conductance points also increases as do the largest values of ΔG_{ON}. At

the smallest values of G_{BL} there is a region (solid box) where $\Delta G_{ON} \propto G_{BL}$ and the

data are single valued. Data for octanedithol show the same general features (B).

Simulations for octanedithiol (C) reproduce these features. This plot was generated for a

series of different contact geometries (A-D, illustrated in figure 5). The scatter increases

as G_{BL} is increased in a way that closely resembles the experimental data in (B) (though

calculated conductances are higher than the measured conductances). The simulations

all converge at small G_{BL}, giving rise to a region where data are single valued and

$\Delta G_{ON} \propto G_{BL}$, as observed in the experiments. The "plateau region" is less densely

occupied in the simulations, probably because of the limited number of tip geometries

explored. In reality, we would expect to find many points around the periphery of the

point forming the smallest gap that can accommodate molecules in their equilibrium

configuration. The solid line in (C) shows the variation in the gap between the gold atom

attached to the sulfur atom and the rest of the gold electrode (Z_{vac}). A nominal contact

point (vertical arrow) is defined by the baseline conductance at which the slope of the

Z_{vac} curve changes abruptly. This is coincident with the transition from contact-

independent conductance to contact dependent conductance. (D) shows data for a

dithiol-diphenol. This "stiff" molecule yields no data in the $\Delta G_{ON} \propto G_{BL}$ regime,

presumably because "stretched" configurations are not energetically possible in this

molecule.

Haiss and Nichols reported that the molecular conductances measured

with telegraph noise varies with the gap, and can become multivalued at small

gaps. [22, 23] For these reasons, we have found it convenient to make a scatter

plot of ΔG_{ON} vs. G_{BL} and data for three combinations of basepairing have been

presented in this form elsewhere.[20] We have now measured several other types

of molecular junction (Chang et al., unpublished) and use some of this data here

to show what appear to be rather universal characteristics. Figure 3.2 shows data

for a guanine-cytidine junction (A) and a junction spanned by just an

octanedithiol (B). The guanine-cytidine junction data are representative of the

two other base-paired systems studied to date. [20] In all these systems, there is a

region where ΔG_{ON} increases approximately linearly with G_{BL} (i.e., exponentially

with decreasing gap) at very small conductances (i.e., large gaps) as shown by the region enclosed by the box on the lower left corners of Figs. 3.2A and 3.2B. As G_{BL} is made larger, measured values of ΔG_{ON} spread out over an increasing range of values between a lower "constant" plateau (dashed boxes) and an upper bound that increase with G_{BL}. Understanding this behavior is critical for the design of Recognition Tunneling readers.

3.2.2 What is going on in these junctions?

We have simulated a junction containing octanedithiol, using the calculated interaction forces in the junctions to predict the types of structure that are likely to arise as the molecule and electrodes influence each other. As we show below, gaps of different sizes are spanned by configurations that involve distortions of both the molecular bonds and the bonds in the metal contact, as well as the molecule-metal bond. We then calculated the electronic conductance for a number of these calculated structures.

Total energy DFT calculations were carried out using the VASP code,[53] within generalized gradient approximation PW91[54]. Plane-wave functions were expanded on a mesh of 12 Monkhorst-Pack k-points on the two-dimensional Brillouin zone with a plane-wave cutoff of 290 eV. The convergence criteria of total energy and force were 10^{-6} eV and 10^{-2} eV/Å, respectively. A Green's function[55] was used for calculating electron transport.

We have studied the system consisting of a single octanedithiol molecule sandwiched between two gold electrodes represented by two asymmetric surfaces having a~(3x3) periodicity in the xy-plane. We also imposed periodic boundary

conditions in the direction perpendicular to the surface, artificially joining the last

two layers of the system. The stretching of the system is simulated by increasing

the distance between the two limiting layers by steps of 0.25Å. After each step,

the system was allowed to relax toward its configuration of minimum energy; in

this relaxation, only the atoms located in the last two layers remain fixed. The

convergence criteria for the relaxation are changes in the total energy per atom

less than 10^{-5} eV and forces on the free atoms less than 0.05 eV/Å. Subsequently,

we used the optimized atomic structures to calculate the current flowing through

molecular junctions. The conductance calculations were carried out using the

Green's function formalism[55] implemented in a local basis set DFT code

Fireball[56, 57] using 64 k-points. The results of these simulations are shown in

Figure 3.3C.
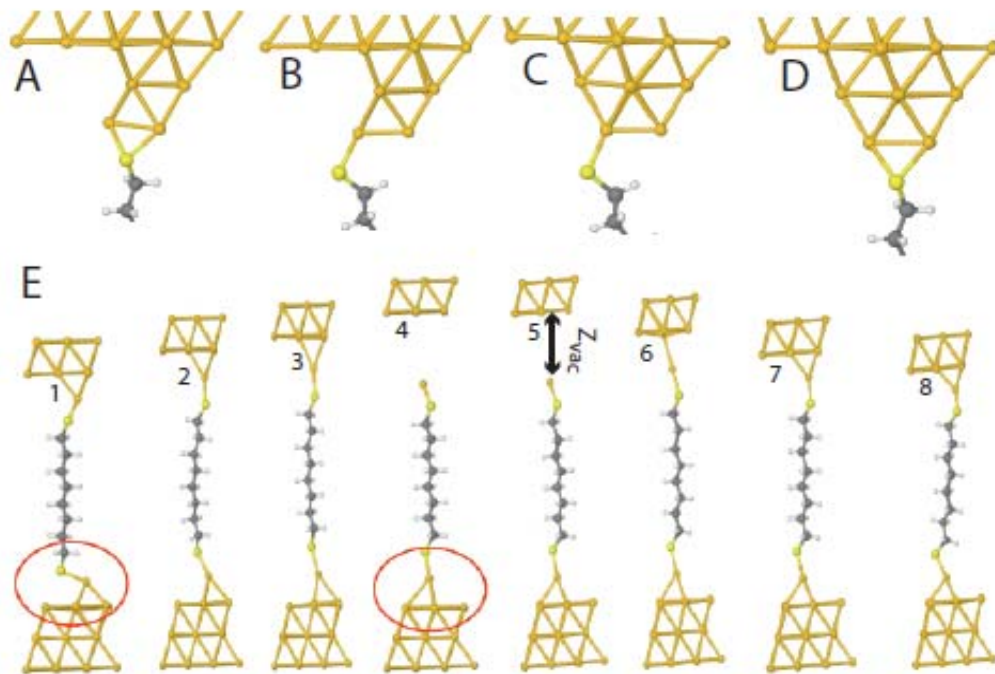
Figure 3-3 Contact geometries.

A-D show contacts of nearly equivalent energy for which the ΔG_{ON} vs. G_{BL} curves in Figure 4

were calculated. E1-E8 demonstrate an intrinsic instability in the surface bonding as a

function of the tunnel gap size (note the differences between the two configurations

circled in red). Note also how the molecule remains "stretched" (4,5 relative to 1,8) even

after the bond to the electrode is "broken", a configuration that is presumably not

available to stiffer molecules.

The calculated conductances are about an order of magnitude higher than

the measured conductances but the general shape of the ΔG_{ON} vs. G_{BL} simulated

plot (Figure 3.3C) is strikingly similar to the measured data (Figure 3.3B). The

simulations were carried out for number of contact geometries, only a

representative selection of which (A-D) are presented here with their structures

illustrated in Figure 3.3.   An important result is that the tunneling current *is*

insensitive to contact geometries at the largest gaps. Figure 3.3C also plots the

"vacuum gap", Z_{VAC} between the gold atom attached to the molecule and the

remainder of the gold cluster (Figure 3.4E). The region of monotonic dependence

of ΔG_{ON} vs. G_{BL} corresponds to the "non-contact" region where Z_{VAC} varies

rapidly with G_{BL}. Figure 3.4E shows an example of how contact geometry can

vary spontaneously over the course of repeated making and breaking of contacts:

the lower contact reorganizes on breaking of the top contact.   Note how this

simulation (and others like it) show clearly that it is an Au-Au bond that breaks

and not an Au-sulfur bond. This result is predicated on the (usually-assumed)

dissociative adsorption of SH on Au, but there is some evidence the hydrogen

remains on the surface.[58] The same picture is recapitulated in our simulations

of non-dissociative adsorption, save that it is the SH-Au bond that is the weak link that stretches (as opposed to the (Au-S)-S bond).

This region of monotonic ΔG_{ON} vs. G_{BL} ("non-contact") sensing may require a flexible molecular junction, for such a region is absent in data obtained with a stiff molecule (dibenzenedithiol -Figure 3.4D).

Finally, this analysis suggests that an "equilibrium" configuration occurs just when a molecule first jumps into contact. In a probe of finite radius, we expect that many such configurations arise, and that these correspond to the "plateaus" observed in the plots of ΔG_{ON} vs. G_{BL} (dashed boxes in Figures 1.4A and 3.4B). For DNA base-pairs, the average values of conductance obtained from these plateaus lie within a factor two of the predicted values, and follow the theoretically-predicted order of conductances.[20]

3.3 Methods
Peiming Zhang and Feng Liang synthesized the nucleosides 5'-thio-deoxycytidine and 5'-thio-thymidine following published protocols,[59-61] deprotecting them and making monolayers on freshly-prepared Au(111) substrates as previously described.[18, 43] Gold STM probes were prepared and functionalized with 8-mercaptoadenine, 8-mercapto-2-aminoadenine and 8-mercaptoguanine (Sigma-Aldrich, St. Louis, Mo.) as previously described.[43] The Watson-Crick pairings are shown below for **i** adenine-thymidine, **ii** 2-aminoadenine-thymidine and **iii** guanine-deoxycytidine.
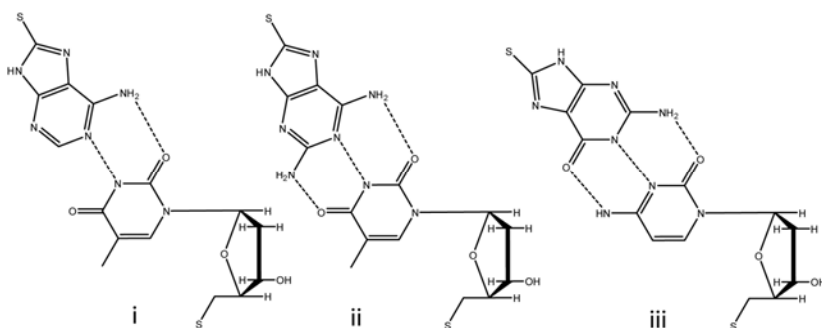
Figure 3-4 The Watson-Crick pairings are shown below for i adenine-thymidine, ii 2-aminoadenine-thymidine and iii guanine-deoxycytidine.

We also prepared probes and gold substrates fuctionalized with a monolayer of thiophenol as controls. Further controls used bare gold probes or substrates. Measurements were carried out on a PicoSTM (Agilent, Chandler) with the sample and probe submerged in 1,2,4-trichlorobenzene. The system was first left to stabilize for three to four hours, and then the probe was advanced towards the surface to achieve a chosen set-point current (I_{SP}) at a tip to substrate bias, V_b. Immediately after the set-point was achieved, the servo-control was broken using custom LabView software, and tunnel current recorded as a function of time using a digital oscilloscope. The tunnel junctions remained stable without servo control for up to about ten seconds. In the hydrogen bonded systems, bursts of "Telegraph Noise" like that shown in Figure 3.4d were recorded in about half the data collection runs. Usually (>95% of the data collected) the noise reflected stochastic switching between two distinct levels, indicative of a single molecular system fluctuating in the junction. Some structural possibilities are illustrated in Figure 3.5.   The molecule-metal contacts may "break" (Figure 3.5b) as in the case of similar recordings from simple bis-thiolated molecules.[22] The motion in the figure is exaggerated for effect, for the frequent re-connection suggests only

a small motion of the contact (which is probably not at the Au-S bond, but rather at the Au-Au bonds that surround the Au atom attached to S[21]). Further, motion in many such thiol-tehtered systems is also affected strongly by interactions with the surrounding molecular matrix.[62] Hydrogen-bond breaking (Figure 3.5c) is yet another possibility for the systems studied here, though we believe this to be unlikely to account for the slow switching we observe (see the discussion below).
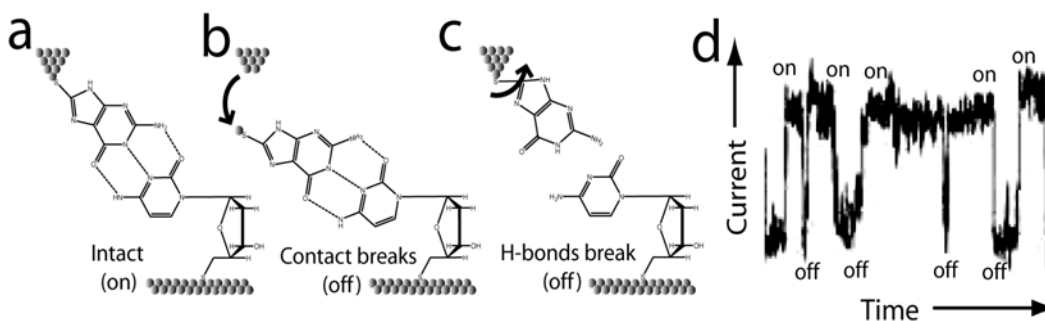
Figure 3-5 An example of the telegraph noise signal produced

(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). (d) shows an example of the telegraph noise signal produced as bonds break and reform.

3.4 Measurement

The result of a typical control experiment (thiophenol probe, thymidine monolayer) is shown in Figure 3.6a with a current trace for 0.5s worth of data on the left, and a histogram of the current distribution on the right. The H-bonded systems (Figures 3.6b, c and d) all show distinct telegraph noise with

corresponding bimodal current histograms. We recorded the fraction of runs that yielded telegraph noise and these are shown as percentages for various combinations of probe and surface in Table 3.1. No switching was seen in all cases where either the probe or the surface was unfunctionalized, demonstrating that both components of the hydrogen bonded system have to be attached to the electrodes in order to generate a switching signal. As further controls, we physisorbed bases onto probes without thiol attachment (bottom rows of Table 3.1) finding no switching signal. Interestingly, thiophenol on the probe or substrate did show rare instances of switching in contact with a base or nucleoside but only at the very lowest bias used. Presumably, interactions between the aromatic benzene ring and the heterocycle can occasionally result in complexation.

	Thio-thymidine	Thio-deoxycytidine	Thiophenol	Bare
Thio-adenine	50%	-	1%*	0
Thio-2-aminoadenine	50%	-	-	0
Thio-Guanine	-	60%	-	0
Thiophenol	3%*	-	0	0
Bare	0	0	0	0
Adenine	0	-	-	-
2-aminoadenine	0	-	-	-

Table 3.1: Observed frequency of switching (percent of measurements) for various preparations of the probe (left column) and the surface (top row). "-" represents untried combinations. * indicates that the observed switching only occurred at low bias (50 mV).
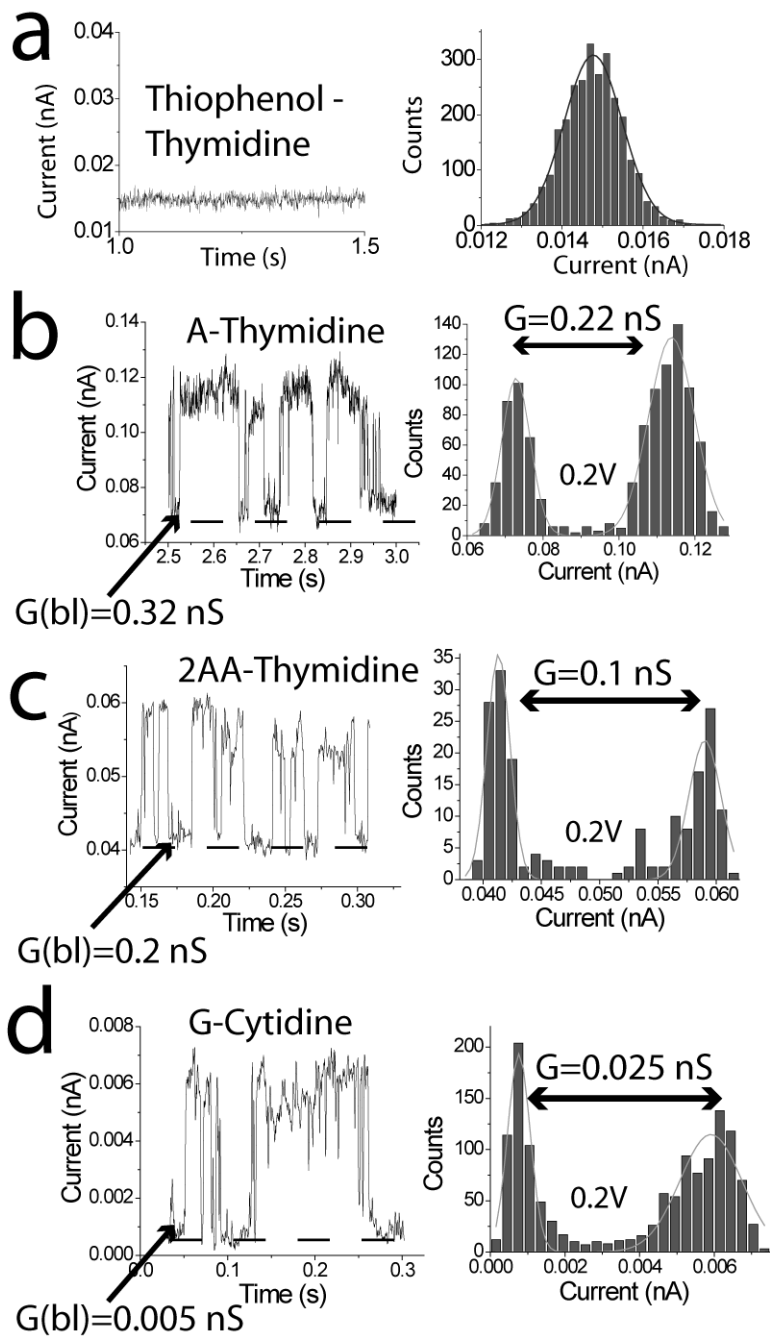
Figure 3-6 Recordings of tunnel current vs. time

:Recordings of tunnel current vs. time (left column) together with the corresponding distributions

of current for (a) 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.

Haiss et al. found that the signals generated by simple alkanedithiols were independent of the initial tunneling conditions.[22] In contrast, we found that the magnitude of the conductance fluctuations in these hydrogen-bonded junctions depended strongly on the initial gap conductance. The signals shown in Figure 3.2 were all collected at low values of the "off" (i.e. baseline) conductance, G(bl), yielding the conductance changes shown on the histograms (obtained by dividing the mean current change by the bias voltage). Data collected at higher values of G(bl) yielded much higher conductances, as shown in Figure 3.7. The distribution of measured conductances is rather wide, shown by the histograms in Figures 3.7a, b and c. However, when the measured conductances are plotted as a function of G(bl) (Figures 3.7d,e and f) it is clear that the conductance rises rapidly with G(bl) initially, but then saturates to a constant value at values of baseline conductance in excess of 0.5 nS. Haiss et al. did not collect data for baseline conductances below this value, and we speculate that our larger molecular system is able to span the larger tunnel junctions (i.e., lower conductances) but only in strained, and hence less-conductive configurations. Presumably the higher conductance junctions, with smaller gaps, offer more possibilities for unstrained molecular pairs to span the gap.
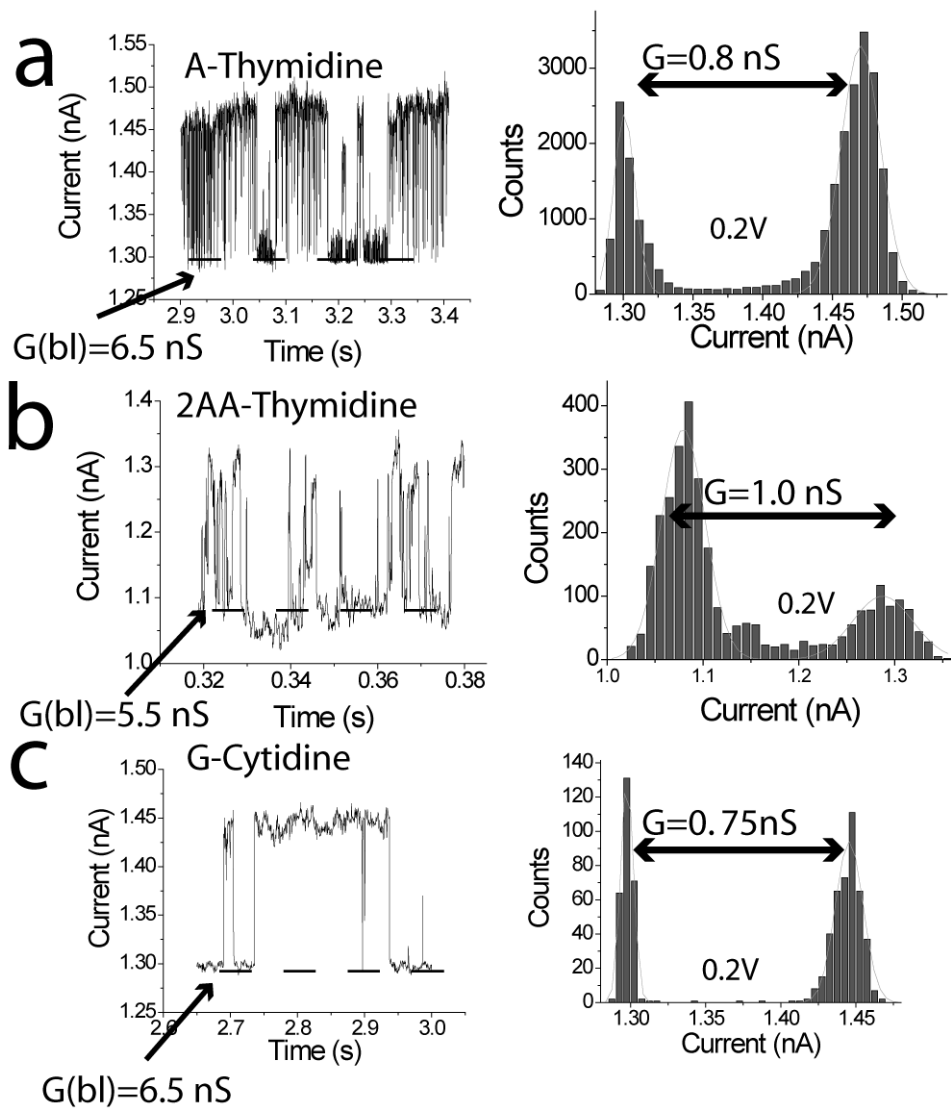
Figure 3-7 Signals collected for the three base-nucleoside pairs (c.f. Figure 3.2 c,d,e) at higher baseline conductance.
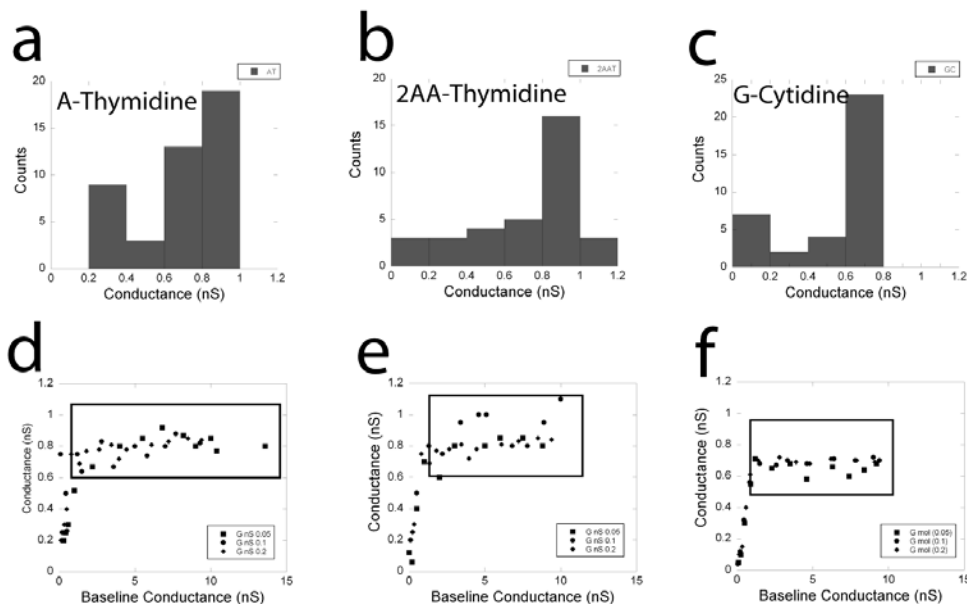
Figure 3-8 Distribution of measured conductances

Distribution of measured conductances for (a) adenine-thymidine, (b) 2-aminoadenine-thymidine and (c) guanine-deoxycytidine. Corresponding plots of these conductances as a function of the baseline conductance. The biases were 0.05V (squares), 0.1V (circles) and 0.2V (diamonds). The data in the boxes are used to derive the currents shown in Figure 3.5.

We selected current data obtained from junctions for which G(bl) > 0.5 nS and plot the mean currents (error bars are ± 1sd) as a function of V_b in Figure 3.8. The current-voltage relationships for all three Watson-Crick bonded pairs are linear, and the slopes yield the conductances listed in Table 3.2. Density-functional methods[63] have been used to calculate the conductances of these base-nucleoside pairs[43] and the results of these calculations are reproduced in the right hand column of Table 3.2. It is striking that the measured and predicted conductances are within a factor two of one another, remarkably good agreement

in the field of single molecule conductance measurements.[3] Even more

strikingly, the surprising prediction that G-deoxycytidine (3 H-bonds) is less

conductive than that A-thymidine (2 H-bonds) is borne out by experiment. This

unusual result is, in the theory, a consequence of the different relative position of

the Fermi-level in the two cases.[63]

Base-pair	$G_{meas.}$ (nS)	$G_{pred.}$ (nS)
A-deoxythymidine	0.76±0.01 (±0.03)	1.62
2AA-deoxythymidine	0.80±0.03 (±0.14)	1.43
G-deoxycytidine	0.66±0.007 (±0.07)	0.96

Table 3.2: Measured conductances (for G(bl)>0.5 nS) compared to calculated values. Values are best-fits to the I-V curves in Figure 4. Errors are derived from the linear fits to the data in Figure 5. The errors in paranthesis are calculated using the SD/mean for the raw current data (Figure 4) so better represent the spread in single molecule values.
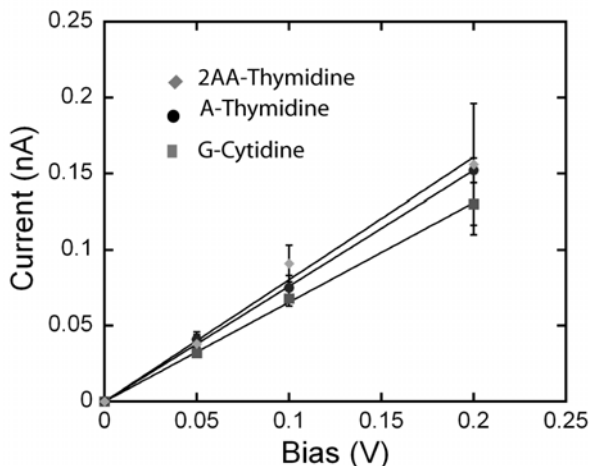
Figure 3-9Current-voltage curves for 2AA-thymidine (diamonds), A-thymidine (circles) and G-deoxycytidine (squares).

Current-voltage curves for 2AA-thymidine (diamonds), A-thymidine (circles) and G-deoxycytidine (squares). Each data point is the mean obtained from data with conductances > 0.5 nS. The error bars correspond to ± 1sd.

53

3.5 Conclusion

We have not addressed the question of which bonds are responsible for the fluctuations. It seems most likely that the observed switching is dominated by the molecule-metal contact for the following reasons: (a) The time-scale of the switching is similar to that reported by Haiss et al. for alkanedithiols.[22] (b) Base-pairs in DNA open on ms timescales, too rapid to be observed here. (c) Analysis of the distribution of opening times does not show significant differences between the two and three H-bond pairs. (d) Fluctuations of H-bonds might be expected to produce intermediate currents corresponding to e.g., one H-bond out of three broken. Switching between more than two levels is rarely observed. On the rare occasions that it is, the levels are all multiples of a common current, much as one might expect if more than one molecular pair spans the junction.

Finally, the close correspondence between the measurements and theory for the present case of base-nucleoside interactions indicates that prediction of very small (fS) conductance across an entire DNA molecule is likely to be valid.[43, 63] We show in a later chapter.

Chapter 4

RECOGNITION TUNNELING MEASUREMENT OF THE CONDUCTANCE OF DNA BASES EMBEDDED IN SELF-ASSEMBLED MONOLAYERS

This chapter describes the application of telegraph noise based recognition tunneling to the measurement of hydrogen bonded systems. It was first published in J Phys Chem C [24]. The contribution of the current author was AC method in STM measurement .

The DNA bases interact strongly with gold electrodes, complicating efforts to measure the tunneling conductance through hydrogen-bonded Watson Crick base pairs. When bases are embedded in a self-assembled alkane-thiol monolayer to minimize these interactions, new features appear in the tunneling data. These new features track the predictions of density-functional calculations quite well, suggesting that they reflect tunnel conductance through hydrogen-bonded base pairs.

4.1 Introduction
Electron tunneling is an extremely localized phenomenon. So much so, that it is believed that tunneling electrodes could detect signals from one (and only one) base in a single-stranded DNA, setting the stage for a new approach to DNA sequencing.[29, 42, 64, 65] However, this same sensitivity renders the detected signals extremely sensitive to the detailed atomic geometry of the contacts,[5] complicating attempts to make reproducible tunneling measurements in the presence of water, ions and the inevitable hydrocarbon contamination of metal

55

electrodes. In order to overcome these problems, we have introduced a new technique we call recognition tunneling.[66] In recognition tunneling, both electrodes are functionalized by recognition reagents attached to the metals by bonds that are strong enough to displace contamination. Tunneling signals are detected via the "telegraph noise" that arises as the reagents stochastically make and break connections with each other,[20] or with a target analyte trapped in the gap.[24] A "two-level" (on-off-on-off…) signal is characteristic of the interaction of a single molecular pair in the tunnel gap, a phenomenon previously exploited by Nichols and Haiss to make measurements of single molecule conductance.[22, 23]

We have determined the tunnel conductance of DNA base-nucleoside pairs using this method, obtaining results that are in reasonable agreement with the predictions of density functional calculations.[20] In this case, theory and experiment agree on an unusual feature: the tunnel current through a guanine-dexoycytidine pair (three hydrogen bonds) is lower than the tunnel current through an adenine-thymidine pair (two hydrogen bonds). Theoretical analysis[67] shows two factors at play. Firstly, transmission through the deoxyribose sugar ring is low, diminishing the importance of the hydrogen bonds in the overall transmission. Specifically, addition of the sugar ring to the molecule alters the molecular levels so as to reduce overall transmission by a large amount. A classical analogy would be to think of the hydrogen bonded system as a low value resistor that depends strongly on the number of hydrogen bonds, in series with a very high value resistor (the sugar) that does not. Secondly, a more

favorable alignment of the Fermi energy in the case of adenine-thymidine swamps any enhancement owing to the extra hydrogen bond in the guanine-deoxycytidine pair.

Theoretical calculations of the transmission through base pairs (directly tethered to electrodes without an intervening deoxyribose ring) show both much larger currents, and a much greater sensitivity to hydrogen bonding, with the predicted currents for triply-hydrogen bonded base pairs being several times the predicted currents through doubly-hydrogen bonded base pairs.[52] Our first attempts to measure currents between base pairs were disappointing. The currents were much smaller than predicted by theory and appeared to be insensitive to the number of hydrogen bonds connecting the base pairs. Knowing that DNA bases interact strongly with gold electrodes,[68, 69] we reasoned that the ribose sugar ring must have played the valuable role of keeping the bases up off the electrode in the case of measurements made with 5'-thiolated nucleosides (as confirmed by FTIR measurements[20]). We therefore modified DNA bases with a thiolated ethylene linker, in the hope that this would keep the bases off the surface, but this also produced small signals that were insensitive to the number of hydrogen bonds holding the pairs together. However, when we first functionalized the gold substrate with an octanethiol monolayer, and subsequently exposed the monolayer to thiolated bases, we discovered new, high current features in the tunnel current data that correlate well with the predictions of density functional calculations, suggesting that this approach mitigates the base-gold interactions, leaving some fraction of the bases in an upright position in the SAM.

The tunnel gaps used in this work had to be made smaller than those used to record signals from base-nucleoside pairs, resulting in an increased risk of artifacts that generate telegraph noise from sources other than a hydrogen-bonding event in the gap. This paper describes a second innovation, which is the adaptation of a combined AC-DC measurement introduced by Tao to verify the interpretation of break junction data.[70] In this method, a small AC modulation is applied to the Z-PZT, resulting in a corresponding AC component in the tunnel current. The magnitude of this modulation is related to the stiffness of the molecular junction. A stiff molecule spanning the electrodes will diminish the modulation relative to the free junction. 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 AC modulation is reduced when the molecule is bound. We have applied the same approach to telegraph noise signals. In this case, we expect to see an anti-correlation between the magnitude of the normalized AC signal and that of the DC signal. As the molecules bind across the tunnel gap, the DC signal jumps up, while the AC signal jumps down. This is precisely what is observed for most of the data, and we have used this signature as our criterion for collecting "good" data.

4.2 Experimental Methods

Reagents, chemicals and characterization: The desired tunneling geometries are illustrated in Figure 4.1. 8-mercaptoguanine (G in Fig. 4.1A) 8-mercaptoadenine (A in Fig. 4.1B) and 2-amino-8-mercatpto-adenine (2AA in Fig. 4.1C) were used as received from Aldrich. 1-(2-Thioacetate ethyl)cytosine and 1-(2-thioacetate

58

ethyl)thymine were synthesized and characterized as described. Monolayers were prepared and characterized on freshly flame-annealed gold-on-mica substrates following the published procedures.[18] Gold STM probes were prepared and functionalized as described elsewhere.[15, 20]
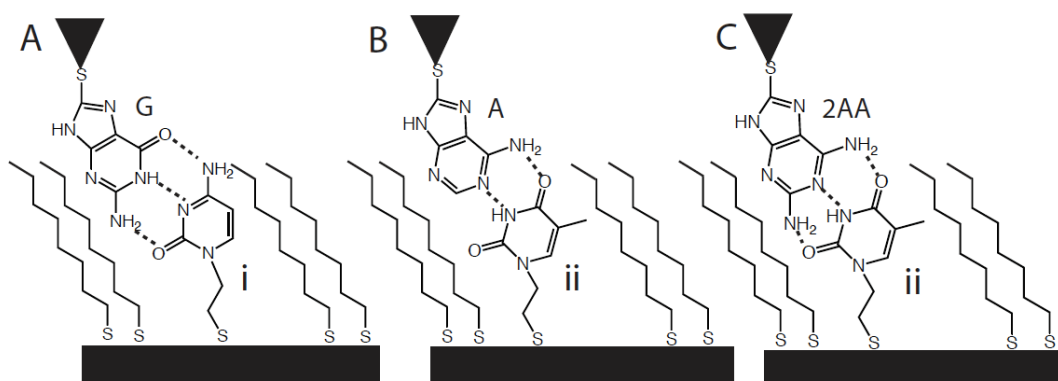
Figure 4-1Interactions between a purine on the STM probe and a pyrimidine embedded in an octanethiol SAM.

Interactions between a purine on the STM probe and a pyrimidine embedded in an octanethiol SAM. (A) Guanine-cytosine, (B) adenine-thymine and (C) 2-aminoadenine-thymine.

Monolayers of 1-(2-mercaptoethyl)cytosine (i in Fig. 4.1A) and 1-(2-mercaptoethyl)thymine (ii in Fig. 4.1B and C) on Au(111) were prepared by immersing the Au substrates in 1mM solutions of 1-(2-thioacetate ethyl)cytosine and 1-(2-thioacetate ethyl)thymine in DMF that were deacetylated with pyrrolidine before use, respectively, for about 2 hours , and then rinsing with DMF and trichlorobenzene sequentially and blown dry with nitrogen. Ellipsometry measurements[18] were indicative of the formation of monolayers

or submonolayers. The quality of FTIR spectra was not adequate to determine orientation of the films.
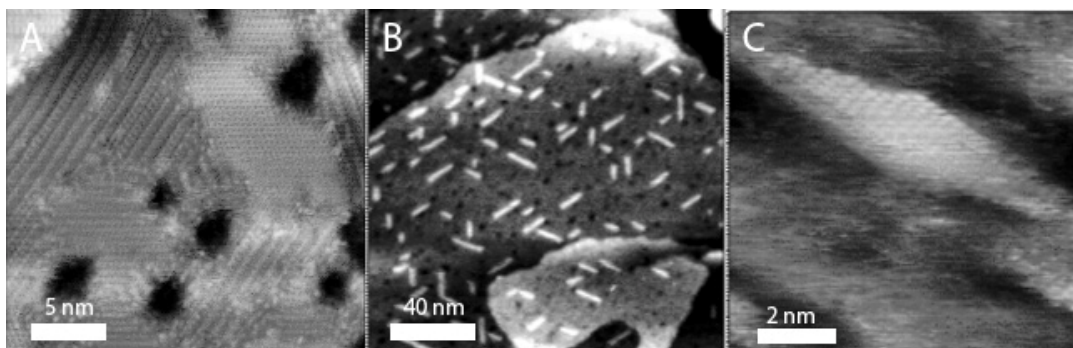
Figure 4-2 High resolution STM images of SAMs

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-mercapto ethyl)thymine, showing stripes of an ordered thymine phase. (C) Zoom in on one of the thymine stripes showing molecular-scale structure. Images were obtained in 1,2,4 trichlorobenzene at -0.1 V (tip bias) and 10 pA using Pt-Ir probes.

We prepared mixed SAMs as follow: freshly annealed Au (111) films were immersed in 5mM octanethiol (Sigma Aldrich) in toluene for about 20 hours, forming highly ordered SAMs (Figure 4.2A). After rinsing with toluene, they were immersed in 1 mM solutions of the thiolated pyrimidines (i and ii) for a further 2 hours. The films were then rinsed with DMF and kept under clean toluene for a another half hour. This final step was critical for the formation of a high quality mixed SAM (Figure 4.2B). The films were characterized by ellipsometry which excluded the possibility of multilayer. The FTIR spectra suggest the coexistence of both molecules in the mixed SAM.

Mixed monolayers containing thymine displayed a remarkable structure in STM images (Figure 4.2B). Bright striped phases were formed (not seen in the alkanethiol monolayers alone) and zooming in on them (Figure 4.2C) showed structures suggestive of rows of stacked thymines. The spacing between adjacent rows was ~ 0.35 nm and the spacing between the disk-like features in the rows was ~0.25 nm. We could not resolve molecular scale features in the mixed SAMs containing guanine.

Tunneling measurements and imaging.

A PicoSTM (Agilent, Chandler) scanning tunneling microscope was used for both tunneling measurements and imaging of the mixed SAMs. During measurements, both tip and sample substrate are submerged in freshly-distilled 1,2,4-trichlorobenzene. For high resolution imaging, we used Pt-Ir probes prepared as described elsewhere.[26]

For measurements of telegraph noise, the functionalized tip was left to stabilize in close proximity to the substrate for at least 30 minutes. After approaching the tip to the surface to obtain a desired baseline conductance, G_{BL}, the current servo was turned off, and the gap was controlled manually via a Labview program which adjusts the z-PZT. Once the gap was stable for periods of 10s or so with no further manual adjustment, runs of telegraph noise were collected. Data were collected over a range of baseline conductances between 5 and 70 nS (corresponding to currents between 0.25 and 3.5 nA with tip biases of 0.05 volts). Telegraph noise was not observed for gap conductances less than 5 nS. Further details are given elsewhere.[66]

For the data reported here, we also applied a 2kHz AC modulation to the z-PZT, adjusted to give a peak to peak amplitude of 0.016 nm was used. This frequency is much higher than the telegraph noise frequency of around 150Hz. The amplitude of the in-phase response was recorded using a lock-in (Stanford Research SR 830). During data acquisition, data from 3 channels was collected simultaneously using a digital oscilloscope (Yokogawa DL750). These signals were (1) DC current (collected through a 400 Hz low pass filter), (2) AC+DC current (full bandwidth) and (3) AC amplitude from the lock-in amplifier with the time constant set to 3 ms.
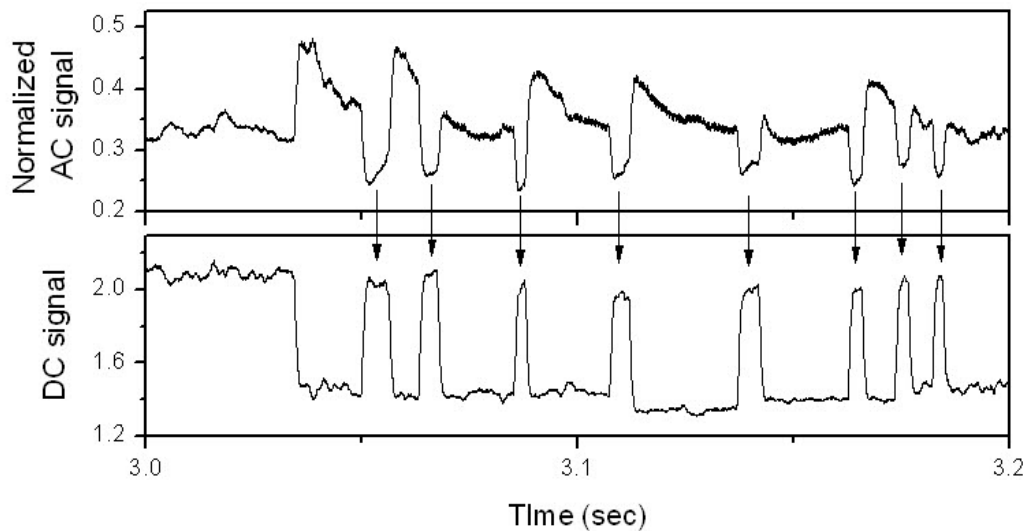
Figure 4-3Showing how the AC modulation signals (top trace) and the DC current signal (bottom trace) are anti-correlated, evidence of molecular binding and unbinding in the tunnel gap.

Showing how the AC modulation signals (top trace) and the DC current signal (bottom trace) are anti-correlated, evidence of molecular binding and unbinding in the tunnel gap.

An example of a typical set of raw data is shown in Figure 4.3. The DC

current (lower panel) shows discrete jumps between the baseline conductance and

a larger value associated with the stochastic bridging of the gap by a bonded

molecular pair. The normalized AC signal (top trace), obtained by dividing the

AC amplitude signal by the DC current signal, remains constant if there is no

molecule trapped in the gap, but falls when a molecule is trapped. [70] This is

because, with $I_{DC} = I_0 \exp(-\beta z)$, the AC

current, for small modulation amplitudes, A, is given by $I_{AC} = \beta A I_{DC}$. Thus, if β

and A do not change, the normalized AC current, I_{AC}/I_{DC}, remains constant

(Chang, manuscript in preparation). A trapped molecule will generally lower

both β and A (the letter because of the increased stiffness of the junction). In the

case of the telegraph noise measurement, this is observed as a fall in the AC

signal when the gap is bridged (i.e., when the DC signal rises) as shown in Figure

4.3.

4.4 Results and Discussion

Scatter plots of the molecular conductance, ΔG_{ON}, vs the baseline conductance,

G_{BL}, are shown for A-T, G-C and 2AA-T in Figures 4.4, 4.5 and 4.6. In each case,

data for the pure SAM are shown in the top panel (4.4A, 4.5A and 4.6A) and data

for the mixed SAMs are shown in the lower panels (4.4C, 4.5C and 4.6C).

Corresponding histograms of the molecular conductances are shown on the right

(4.4B and D, 4.5B and D and 4.6B and D). These data show the broad spread that

comes from variations in the contact geometry, with the trend of increasing ΔG_{ON}

with increasing G_{BL} previously reported for measurements of this kind.[66]

Common (low energy) configurations give rise to "plateau" regions in these plots, represented by a peak in the distribution of conductances[66] (Figure 4.4B and D, 4.5B and D and 4.6B and D).
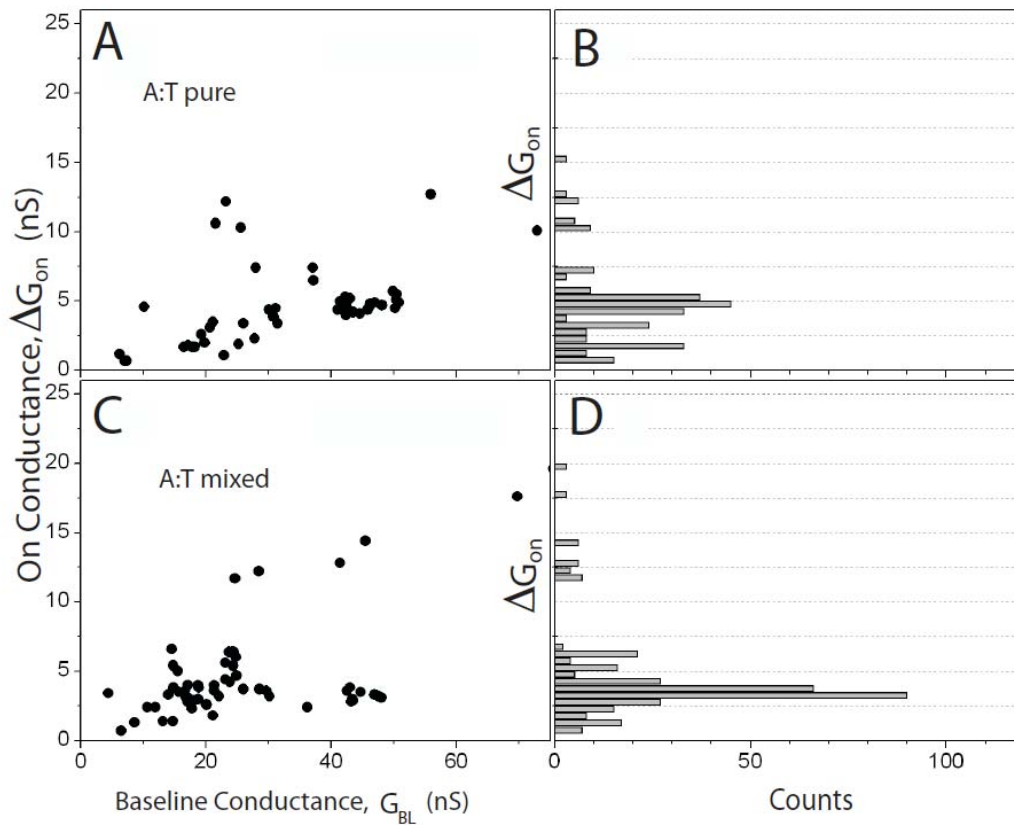
Figure 4-4Scatter plots of molecular conductance,

Scatter plots of molecular conductance, ΔG_{on} vs. the baseline tunnel conductance G_{BL} 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).
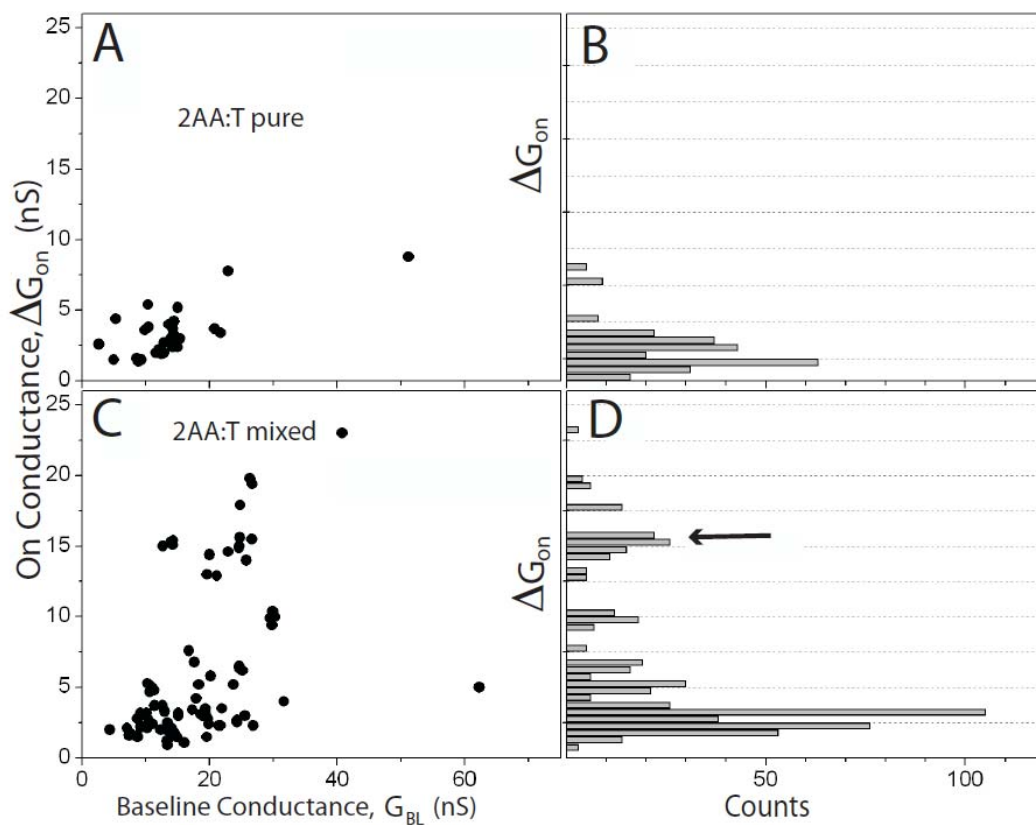
Figure 4-5Scatter plots of molecular conductance,

Scatter plots of molecular conductance, ΔG_{on} vs. the baseline tunnel conductance G_{BL} for (A) 2-aminodenine interacting with thymine monolayers and (C) 2-aminoadenine interacting with thymine 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.

65
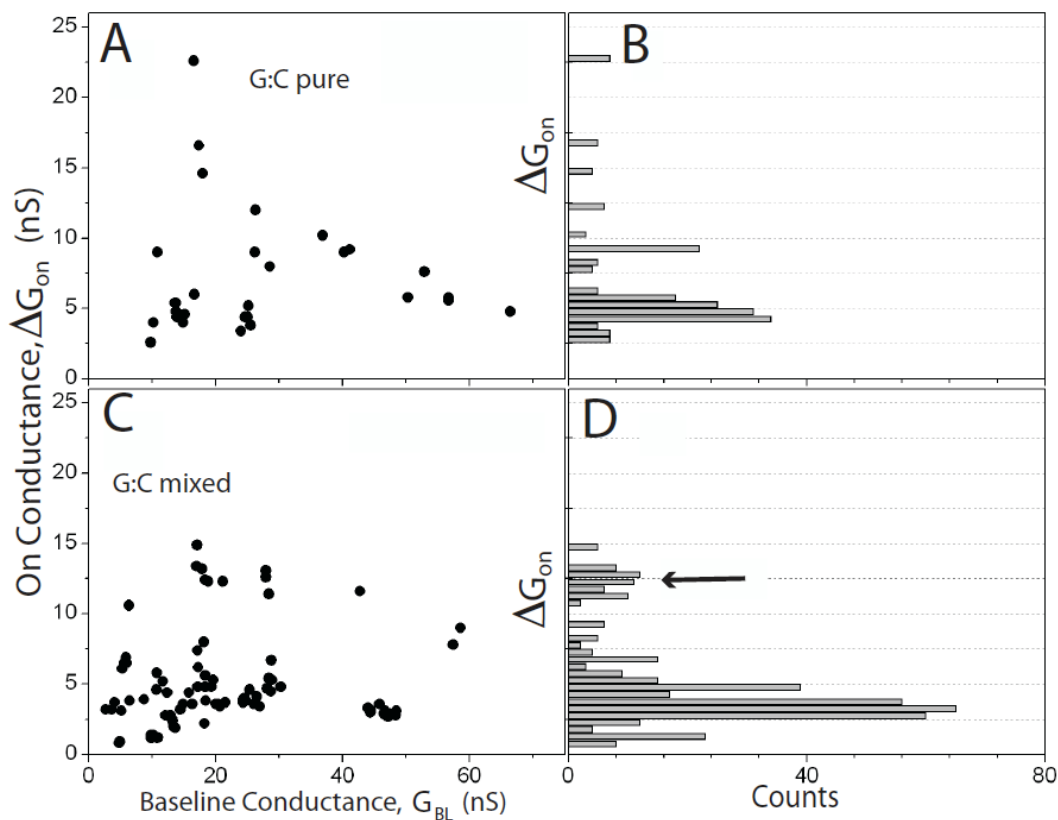
Figure 4-6Scatter plots of molecular conductance

: Scatter plots of molecular conductance, ΔG_{on} vs. the baseline tunnel conductance G_{BL} 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.

The pure SAMs all show one peak in the conductance histogram, and Gaussian fits yield the peak values and widths shown in Table 4.1. They show no significant changes with the number of bridging hydrogen bonds. We also calculated the average bound state lifetimes directly from the telegraph noise

traces[66] and values are shown in Table 4.1. The bound states persist for longer

for triply hydrogen bonded pairs[20] but no significant differences are observed

here.

Table 4.1: Telegraph noise characteristics for basepairs.

	A on T	AA on T	G on C
Pure SAM Conductance (nS)	4.49±1.28	2.67±1.16	4.77±0.71
Mixed SAM 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

* The values in the parenthesis are calculated for pyrimidines with thiol groups directly
bonded to the heterocycle

Theoretical values for basepairs, including 2AA-T, have been calculated

by Lee and Sankey.[52] These calculations were carried out for pyrimidines

attached to gold electrodes by a thiol group directly bonded to the heterocycle,

and the corresponding values are shown in parenthesis in Table 4.1. Note the

large range in the predicted conductances for the three molecular pairs, and the

large predicted difference between conductances for two- and three- hydrogen

bond connections (a factor between 4 and 5X). This is in sharp contrast to values

calculated for base-nucleoside pairs where the number of hydrogen bonds makes

little difference to the predicted (and measured) molecular conductance.[20]

The theoretical predictions for the bases do not take account of the

ethylene linkages used in the present work. We expect the conductances of the

ethylene-tethered bases (Fig. 4.1) to be lower by a factor e^{-2} because the electronic decay constant β, for alkanes tethered to gold via a thiol linkage is 1 per methylene. [71] Appropriately corrected values of the predicted conductance are shown in Table 4.1.

   Data for the mixed SAMs (Figures 4.4C, 4.5C and 4.6C with corresponding histograms given in 4.4D, 4.5D and 4.6D) show evidence of a second plateau at a higher value of conductance for the two triply hydrogen bonded complexes. Fits to the conductance distributions for these mixed SAMs are shown in Figure 4.7 with the corresponding peak positions and widths listed in Table 4.1.
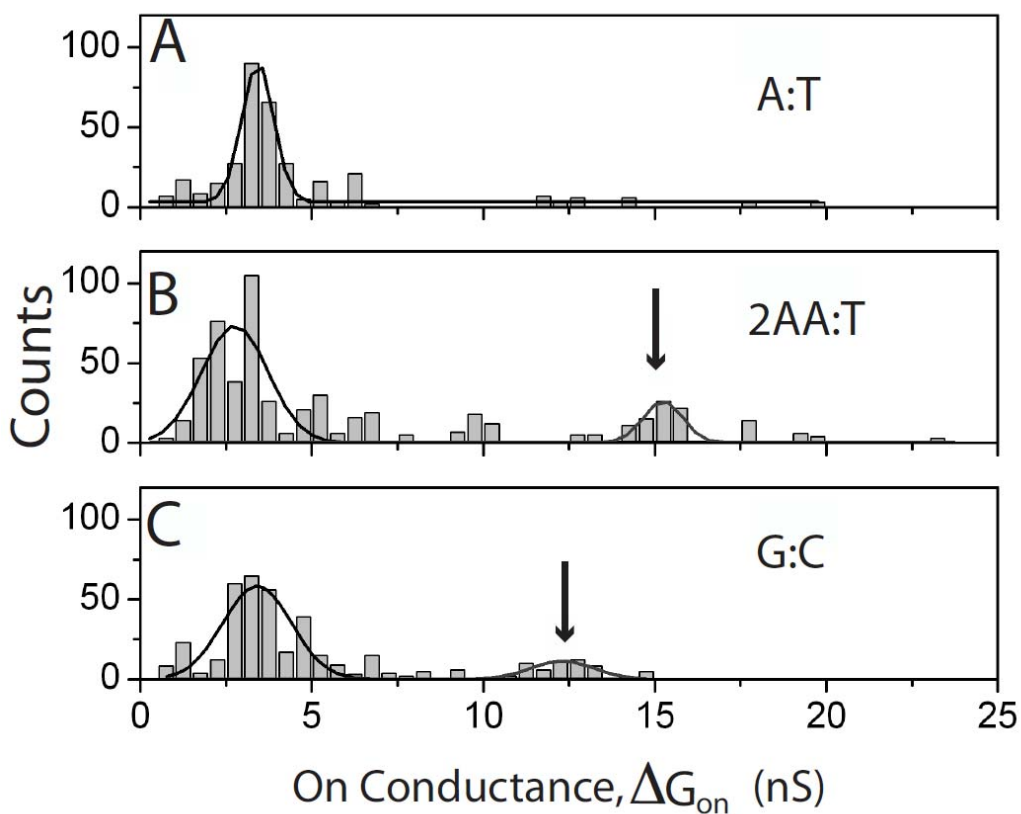
Figure 4-7Gaussian fits to the conductance distributions for the mixed pyrimidine SAMs.

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.

The adenine-thymine data from mixed SAMs are little changed from what was observed in the pure film, both in peak position and peak width (see Table 4.1). Does this mean that the thymine configuration is little different between the pure and mixed SAMs? This is a possibility, in that thymine has the weakest intrinsic interaction with gold[68] so the data obtained from the pure SAM could have come from properly hydrogen bonded pairs. However, we can rule this possibility out, both because of the striking degree of alignment of the thymines in the mixed film (Figure 4.2) and because of the appearance of a new conductance peak in the 2AA-T data (see below). This implies that the thymine orientation must be altered when the molecules are embedded in the octanethiol monolayer. Therefore it appears that the distribution of conductances measured for a vertically oriented array of thymines coincidentally overlaps the distribution measured over the pure film (where, presumably, there are possibilities for even quite exotic interactions like π-stacking[72]).

The appearance of a second, high conductivity peak in the mixed-SAM distributions for 2AA-T and G-C (Figure 4.5 and 4.6) must be a consequence of an increased frequency of vertical orientations that enable proper hydrogen bonding of the purine and pyrimidine. The continued presence of a significant low conductance peak shows that the interactions are still far from uniform. Comparing the theoretical predictions (listed in Table 4.1) to the values of the

69

second peaks for 2AA-T and G-C, and to the only peak for A-T, shows an uncanny degree of agreement. The absolute scale of the agreement cannot be taken too literally because of the many uncertainties in calculations of this sort.[3] The correspondence between predicted ratios of conductance (2AA-T>G-C>>A-T) is significant, capturing the predicted order 2AA-T > G-C for the triply hydrogen bonded complexes.

4.5 Conclusions

The predicted trends in conductance for three DNA basepairings (2AA-T, G-C and A-T) are not observed in single molecule conductance measurements made with monolayer of thiolated bases attached to bare gold electrodes. New features appear in the conductance distributions for 2AA-T and G-C when the bases on the substrate are inserted into an octanethiol SAM. If these features arise from properly hydrogen bonded base pairs, then a remarkable agreement with theory is found. Mixed SAMs may offer a powerful approach for carrying out recognition tunneling with reagents that have a strong intrinsic chemical interaction with the electrodes.

Chapter 5

ELECTRONIC SIGNATURES OF ALL FOUR DNA NUCLEOSIDES IN A TUNNELING GAP

This chapter describes the use of recognition tunneling with two functionalized electrodes, designed to detect DNA bases. It was first published in *Nano Letters*[24]. Nucleosides diffusing through a 2 nm electron-tunneling junction generate current spikes of sub ms duration with a broad distribution of peak currents. This distribution narrows ten-fold when one of the electrodes is functionalized with a reagent that traps nucleosides in a specific orientation with hydrogen bonds. Functionalizing the second electrode reduces contact resistance to the nucleosides, allowing them to be identified via their peak currents according to deoxyadenosine > deoxycytidine > deoxyguanosine > thymidine, in agreement with the order predicted by a density functional calculation.

5.1 Introduction

New approaches to DNA sequencing are required to reduce costs and increase the availability of personalized genomics [42]. In addition, long contiguous reads would help to unravel the long-range structure of the genome [73, 74]. In contrast to Sanger sequencing and next-generation methods, nanopore sequencing [65] is an enzyme-free technique in which DNA molecules are forced through a tiny aperture using electrophoresis, so that a sequence-reading mechanism could maintain its fidelity over the entire length of a molecule. Ion current that passes through the pore is sensitive to the sequence in the nanopore [75-77] but all of the

bases in the nanopore channel contribute to the current blockade [78] as well as those in the region of high field beyond the pore [79, 80]. In consequence, single base resolution has not yet been attained with an ion current readout. Lee and Thundat proposed that electron tunneling across a DNA molecule might be localized enough to sense and identify single nucleotides [64], a conjecture supported by the calculations of Zwolak and Di Ventra [29]. Further calculations show that thermal motion of molecules in the gap broadens the distribution of tunnel currents [81] [82], reducing selectivity substantially. The range of orientations of molecules in a tunnel gap can be greatly reduced by using chemical bonds to tether it to the readout electrodes [5], however, the use of strong bonds is clearly not an option for DNA sequencing where the contact to the electrodes must slide from one nucleotide to the next rapidly. Ohshiro and Umezawa demonstrated that hydrogen bonds can be used to provide chemical contrast in scanning tunneling microscope images [14] suggesting that these weaker bonds can serve as "sliding contacts" to single molecules. We have used chemical reagents that hydrogen-bond to DNA [83] and the nucleosides [15] to measure conductances for single nucleosides bound to an electrode, finding values that are in reasonable agreement with density functional calculations [20, 52].   Here, we report measurements of the current signals generated as free nucleosides diffuse into a tunnel junction in which both electrodes are functionalized with a reagent that presents a hydrogen bond donor and acceptor to the nucleosides (Figure 5.1A). This functionalization serves both to limit the range of molecular orientations in the tunnel gap, and reduces the contact

resistance, increasing the selectivity of the tunneling signal, so that a direct

readout may be possible with a few repeated reads.
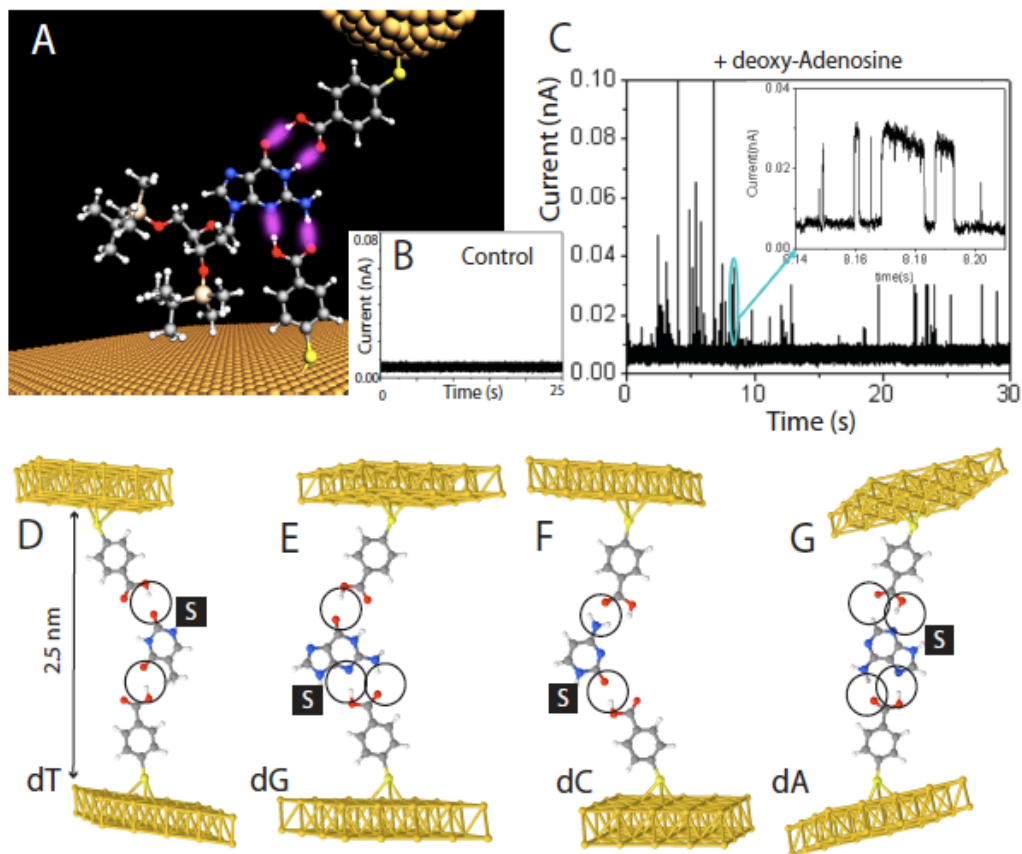
Figure 5-1 Tunneling measurements with functionalized electrodes.

Tunneling measurements with functionalized electrodes. (A) A gold probe and a gold

substrate are functionalized with a monolayer of 4-mercaptobenzoic acid, and 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 (B). When a solution of nucleosides is introduced, current spikes appear,

as shown here for [84] 0.7 μM deoxyadenine in trichlorbenzene with a baseline tunneling

current of 6pA at a bias of 0.5V (C). Hydrogen-bonding schemes for all four nucleosides

are shown in D-F. "S" represents the modified deoxyribose sugar and the hydrogen

bonds are circled.

In order to reduce complications associated with electrochemical leakage,

we operated in an organic solvent (1,2,4-trichlorobenzene, TCB) and the OH

groups of the deoxyribose ring of all four nucleosides were protected with tert-

butyldimethylsilyl (TBDMS) to improve their solubility. Gold electrodes were

functionalized with 4-mercaptobenzoic acid, forming monolayers with the

benzoic acid moiety exposed to solvent [85] [86]. In non-polar solvent, the

benzoic acid is neutral, the O-H group acting as a hydrogen bond donor and the

carbonyl oxygen acting as a hydrogen bond acceptor [49]. Plausible hydrogen

bond configurations for trapping all four nucleosides are shown in Figs 5.1D-G.

Synthesis, preparation and characterization of all reagents, probes and surfaces are

described later.

We carried out tunneling measurements on a PicoSPM scanning probe

microscope (Agilent, Chandler) interfaced to a digital oscilloscope. When both

the probe and a gold (111) substrate were functionalized with 4-mercaptobenzoic

acid, the tunneling background signal in TCB was relatively noise free for set-

points currents, I_{bl} of up to 10 pA at 0.5V bias, a conductance of 20 pS (Fig.

5.1B). A nucleoside solution was placed in the liquid cell, and after the

polarization current had fallen to a small value we re-engaged the probe at a

tunnel current level that had previously given a low-noise background signal.

Current spikes were immediately obvious in the tunneling signal (Fig. 5.1C).

Because neither the surface concentrations of nucleosides nor the efficiency of

74

molecular capture in the gap are known *a priori*, we adjusted the concentrations

of the nucleoside solutions to give approximately equal "spike rates" in the tunnel

gap (Table 5.1). Many of the "spikes" showed the two-level "telegraph noise"

characteristic [20] of binding and unbinding of a single molecule in the gap

(insets, Fig. 5.1C). The STM servo gains were set so that only spikes of the

longest duration were affected by the action of the current-control servo.

	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^{-1})	7.1±1.4	5.5±1.1	5.5±1.1	6.6±1.3

Table 5.1: Measured and calculated conductances in a functionalized tunnel junction at I_{bl} = 6 pA, V = 0.5V. Measured values are the average of three independent runs (errors are ± 1sd). Calculated conductances are for the structures shown in Figure 1 D-G. Read rate is based on counts acquired in a 180 s period for nucleoside concentrations between 0.8 and 4.3 μM. The disparity in the range of values between theory and experiment may reflect neglect of a background contribution via solvent-mediated tunneling into a molecule bound at one electrode [87]. Absolute values will be affected by inaccuracies in the estimate of the gap size.

5.2 Results and discussion

We generated distributions of the peak currents using a custom program to

analyze the height of the spikes. The program captures signals two standard

deviations above the noise on the baseline, and also rejects data of only one or

two points in time (i.e. up to 40 μs duration). The effect of the choice of filtering

parameters on the measured distribution is discussed. Figure 5.2 shows how these

measured distributions are affected by functionalization of the electrodes.

Distributions recorded with bare electrodes are shown in Figs 5.2A and 5.2C. In

order to record signals with bare electrodes, we had to reduce the tunneling gap a

little by operating at a conductance of 20 pS. Even at this smaller gap, reads with

bare electrodes on the pyrimidine nucleosides were much less frequent than reads on purine nucleosides. The measured current distributions were fitted quite well by a Gaussian distribution of the logarithm of the currents (solid lines). The fitted peak currents differ for these two nucleosides (15.9±0.4 pA for dA and 18.7±0.2 pA for dG) but the difference (2.8 pA) is less than the width of the distribution on the high current side (~15pA). When measurements are repeated with a functionalized substrate and a bare gold probe at an increased gap (corresponding to 12 pS) the distribution of measured currents narrows by an order of magnitude (Fig 5.2B - dA, Fig 5.2D - dG) but the peak currents are not significantly different. The distribution of spike lifetimes is quite similar for both bare electrodes and for one functionalized electrode. Thus, it appears likely that the spikes observed with bare electrodes correspond to transiently bound states of the nucleosides also. If this is the case, then the narrowing observed with a functionalized electrode must be a consequence of a reduction in the number of types of bound states in the tunnel gap. When both probe and substrate are functionalized, (Fig 5.2E - dA, Fig 5.2G – dG), the peak current for dA is clearly higher than the peak current for dG. Thus distinctive signals can be generated when both electrodes are functionalized, but do they originate with single nucleosides? The "telegraph noise" signals are characteristic of single-molecule reads and the small size of the peaks assigned to two-molecule reads ("2" in Figs 5.2B,D,E,G) suggests that reads of more than one molecule at a time are infrequent. However, electrochemical leakage currents can introduce current errors that depend on the nucleoside so the measured current may not be generated from single molecule currents alone. A better test of the

fidelity of tunneling reads can be carried out using mixtures of two nucleosides so that any errors owing to an electrochemical background are present in both sets of signals. Fig 5.2F shows the current distribution obtained with a mix of dA and dG. The higher current peak is at essentially the same current as recorded for dA alone, and thus should count the dA molecules in this mixture. This assignment is confirmed by halving the concentration of dA in solution (Fig 5.2H) [84].   Most of the data in this panel were well fitted assuming single molecule reads with only 5% of the reads consistent with both dA and dG in the gap at the same time ("dA+dG", Fig 5.2F).
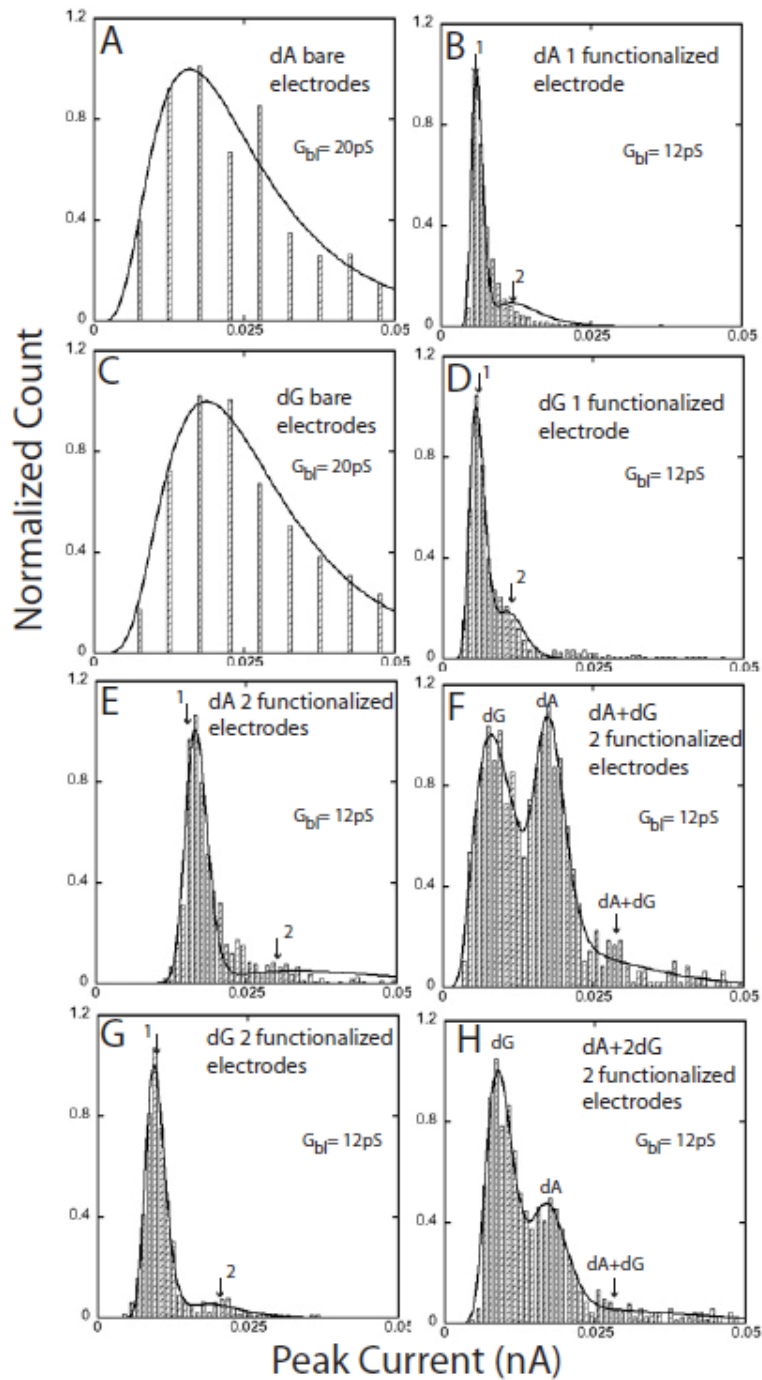
Figure 5-2Effect of electrode functionalization on the distribution of current spikes for purines.

Effect of electrode functionalization on the distribution of current spikes for purines. Bare electrodes (A, dA and C, dG) give broad distributions (gap conductance 20 pS, 0.7 μM dA, 2.9 μM dG in TCB). Fits are Gaussian in the log of the current (figs S6-8).

78

Distributions narrow ten-fold when one electrode is functionalized with 4-mercaptobenzene (B, dA, D, dG) (gap conductance 12 pS, $I_{bl} = 6$ pA, $V = 0.5$V). Fits are to two Gaussians in the log of the current with a peak at i_0 ("1") and a second at $2\ i_0$ ("2") (Eqn. S3). $i_0 = 5.9$ pA for dA and 5.6 pA for dG. When both electrodes are functionalized (E, dA, G, dG) the peak currents are clearly different ($i_0 = 9.4$ pA for dG, $i_0 = 16.5$ pA for dA). F shows the 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. Distributions of the spike widths are given in fig S16.

The same types of features are observed for dC and dT (Fig. 5.3) but the data for a bare gap and bare substrate had to be collected at a yet larger tunnel current (20pA, corresponding to 40 pS) in order to acquire a significant number of reads for the (smaller-sized) pyrimidine nucleosides. dC and dT are also clearly separated in a mixed sample when read with probes that are functionalized (Fig 5.3C).
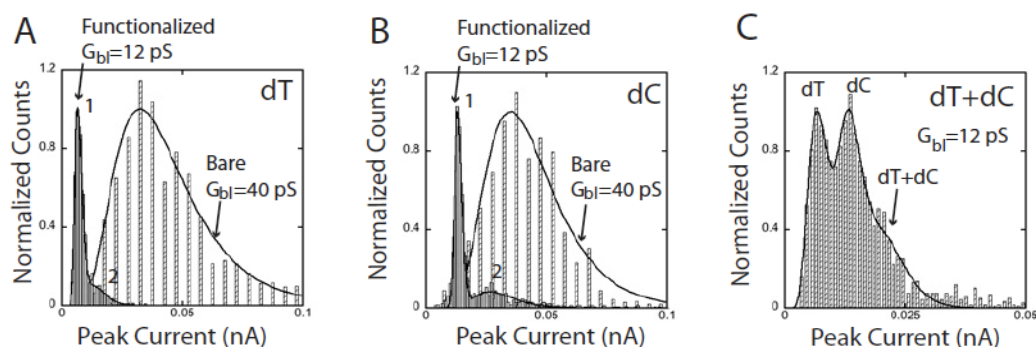
Figure 5-3 Effect of electrode functionalization for pyrimidine reads.

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

At a given bias, the absolute value of peak current is directly proportional to the baseline conductance of the gap (Fig 5.4A), i.e., it increases exponentially as the gap is decreased, similar to what has been reported for other hydrogen-bonded systems in large tunnel gaps [20]. We found evidence of an interesting dependence of the peak currents on bias at a fixed gap size (i.e., gap conductance) indicating the possibility of a non-linear current-voltage dependence for molecules bound to both electrodes. The read frequency also increased as the gap was narrowed. On the other hand, the fraction of multimolecule reads increased rapidly in smaller gaps (Fig 5.4A) so 12 pS appears to be an optimal value for the baseline conductance at a bias of 0.5V.

Values for the peak currents measured at $I_{bl} = 6$ pA, V = 0.5V are summarized by the cross-hatched bars in Fig. 5.4 B. These are the results of three different runs (one carried out, from sample preparation to data analysis, by a different team) on each of the four nucleosides. The peaks for each nucleoside are separated by an amount comparable to the width of the distribution, allowing the fraction that are single-molecule reads with two "good" contacts to identify the base with $p \geq 0.6$.
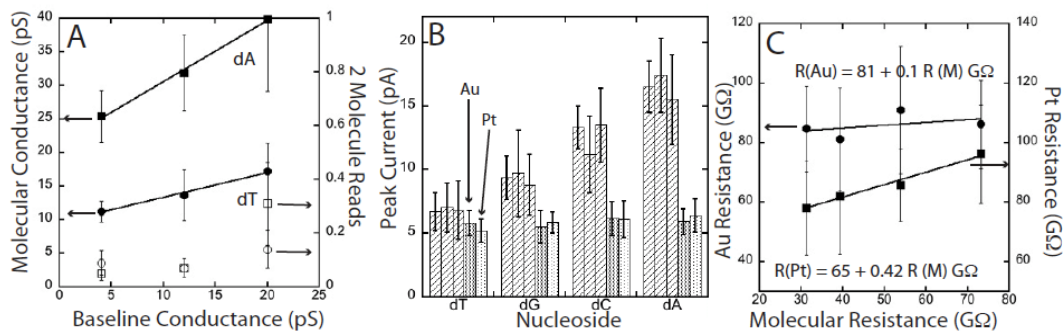
Figure 5-4 Summary of the reads.

Summary of the reads. A. The measured molecular conductance increases linearly with G_{bl} (black circles dT, black squares dA, error bars are ±HWHH). The number of two molecule reads (open circles, dT, open squares, dA) increases at G_{bl} = 20 pS, and the read rate is substantially reduced at G_{bl} = 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 identity of the nucleoside, as shown quantitatively in C where the junction resistance is plotted vs. the molecular resistance determined with two functionalized probes.

We also recorded data with a functionalized substrate and a bare Au (dark shaded bars) or bare Pt (light shaded bars) probe. The peak currents change little from nucleoside to nucleoside, an expected consequence of the resistance, R_c, associated with bare contacts [5] although the lack of selectivity is not accounted for by contact resistance alone. If we assume that reads with two functionalized probes determine a resistance for a single molecule, R_m , then the resistance of a junction with one bare electrode should be given by $R_j = R_c + R_m$. Fig 5.4C shows that the signal with one bare gold electrode is insensitive to the molecular

81

resistance, while a bare Pt electrode is about half as sensitive as the simple

"resistors in series" model predicts, probably reflecting the way in which binding

to the electrodes affects the position of molecular states [88] .

At 12 pS conductance, we estimate the gap to be about 2.5 nm, using

$G = G_0 \exp(-\beta x)$ where G_0 is the quantum of conductance (77 μS) and $\beta = 6.4$

nm^{-1} [18]. Figs 5.1 D-G show what we believe to be the most likely hydrogen

bonded (energy-minimized) structures for the four nucleosides in a gap with both

electrodes functionalized.   We carried out density functional calculations of the

conductance of these four molecular junctions and the predicted conductances are

listed below the measured values in Table 5.1. The predicted order of

conductance agrees with experiment, though the absolute values are significantly

lower, possibly because of an overestimate of the size of the tunnel gap [87].

5.3 Conclusions

The present work shows that the two major impediments to sequence

readouts by tunneling – a wide range of molecular orientations and a large contact

resistance – can be overcome using functionalized electrodes [89]. Overlap

between peaks limits the probability of a correct read to ≥ 0.6. The successful

read-rate is further lowered by the presence of a significant fraction of "single

contact" reads (i.e., peaks that appear at the same place as found with just one

functionalized electrodes). These can amount to about half the total reads,

depending on the probe and the point on the substrate that is sampled. The

solution concentrations required to get approximately equal read rates differed by

a maximum factor of 6 so the reading efficiencies for each base are unlikely to

differ by a large amount. These efficiencies will be better measured once a recognition molecule is developed for reads in aqueous electrolytes using oligonucleotide targets. The recent introduction of a naturally-conductive nanopore [90] might facilitate the integration of a tunneling readout into a nanpore device.

Chapter 6

SINGLE BASE RESOLUTION IN TUNNELING READS OF DNA COMPOSITION

This chapter describes the application of recognition tunneling to identification of

DNA bases embedded in DNA polymers in aqueous electrolyte. It was first

published in Nature Nanotechnology [43].

Single-molecule DNA sequencing based on measuring the physical properties of

bases as they pass through a nanopore[42, 65] eliminates the need for the enzymes

and reagents used in other approaches. Theoretical calculations indicate that

electron tunneling could identify bases in single-stranded DNA, yielding long

reads and eliminating enzymatic processing.[29, 81, 91] It was shown recently

that tunneling can sense individual nucleotides[92] and nucleosides.[24] Here, we

show that tunneling electrodes functionalized with recognition reagents can

identify a single base flanked by other bases in a short DNA oligomer. The

residence time of a single base in a recognition junction is on the order of a

second, but pulling the DNA through the junction with a force of tens of

piconewtons would yield reading speeds of tens of bases per second.

6.1 Introduction

Changes in the ion current through a nanopore can be used to identify

translocating nucleotides. This opens the way to DNA sequencing if an

exonuclease can pass each cleaved nucleotide into the pore sequentially.[93] As

an alternative, it has been proposed that the high spatial resolution of electron

tunneling would allow direct reading of bases in an intact DNA polymer. [29, 81, 91] Recent progress in measuring electron tunneling through nucleotides or nucleosides shows that they can be identified by means of characteristic current signals.[24, 92] Recognition tunneling[24, 94] is an approach in which electrodes are functionalized with reagents that bind the target DNA bases. Contact via molecular adsorbates has been used to produce extraordinarily high spatial resolution in atomic force microscopy[95] and, as we show here, single bases can be resolved in a DNA polymer when read by means of a selective chemical contact.
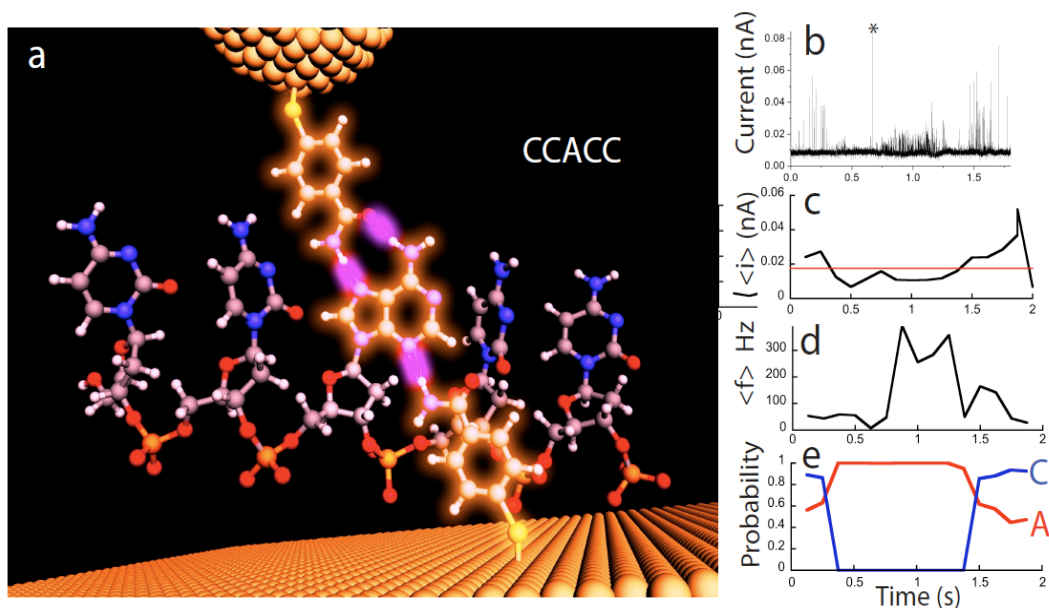
Figure 6-1 Reading a single base within a heteropolymer.

> Reading a single base within a heteropolymer. (a) Benzamide groups on the probe and substrate bind bases in the polymer to give a signal dominated by the shortest tunneling path (highlighted for the connection to the single A in d(CCACC)). (b) Characteristic bursts of tunneling noise with large, infrequent spikes signaling C and smaller, more frequent spikes signaling A. (Background tunnel current is 10 pA, bias + 0.5V). The

spike labeled * (off-scale at 0.11 nA) is non specific, and rejected from the analysis. (c)

Rolling average of the spike height (0.25s window, 0.125 s steps) and spike frequency (d).

C bases generate a negligible number of spikes below 0.015 nA (red line). (e) Probability

that the signal comes from a A (shown by the red line on panel c) or a C (blue line). For

signal amplitudes >0.015 nA, probabilities are calculated as described (values are not

normalized to add to 1 in these regions). This burst of signal was chosen to show a clear

example of transitions between A and C bases. Longer time traces are dominated by

signals from A's which are preferentially trapped in the junction.

6.2 Experiment results and Discussion

To extend recognition tunneling to reads in buffered aqueous electrolyte, we

synthesized the reagent 4-mercaptobenzamide (Fig. 6.1a and Methods) which

presents two hydrogen-bond donor sites (on the nitrogen) and one hydrogen-bond

acceptor site (the carbonyl). Likely binding modes to the four bases are shown in

Fig. 6.2a.[24] A gold (111) substrate and a partially-insulated gold STM probe

were functionalized with this reagent (Method) and characterized in an electron

tunneling junction formed in a scanning tunneling microscope (PicoSPM, Agilent,

Chandler, AZ). Fig. 6.1a shows a d(CCACC) oligomer trapped in a tunnel gap

through hydrogen bonding to one mercaptobenzamide molecule on the probe and

another on the substrate.   In reality, the oligomer is probably held by many

contacts, but only those that complete a short tunneling path (highlighted) will

contribute significantly to the current. In our measurements, the probe is not

deliberately scanned, but moves over the substrate as the microscope drifts.

Alternatively, molecules may diffuse through the gap. Characteristic bursts of

86

current are observed, and an example is shown in Fig. 6.1b. As we show below, the low frequency, large amplitude pulses indicate a C, while the high frequency, small amplitude pulses signal an A. Fig. 6.1c shows a sliding average of the spike amplitudes – values below the red line identify an A base unambiguously. Figure 6.1d shows a sliding average over the pulse frequencies (as defined for each adjacent pair of spikes) – the low frequency regions at each end enhance the confidence with which those regions can be assigned to a C base. The probability of an assignment to A (red line) or C (blue line) is shown in Fig. 6.1e. Calculation of these probabilities is based on our study of nucleotides, homopolymers and heteropolymers as described below. This example clearly shows that a single A base can be identified with high confidence when flanked by C bases in an intact DNA molecule.
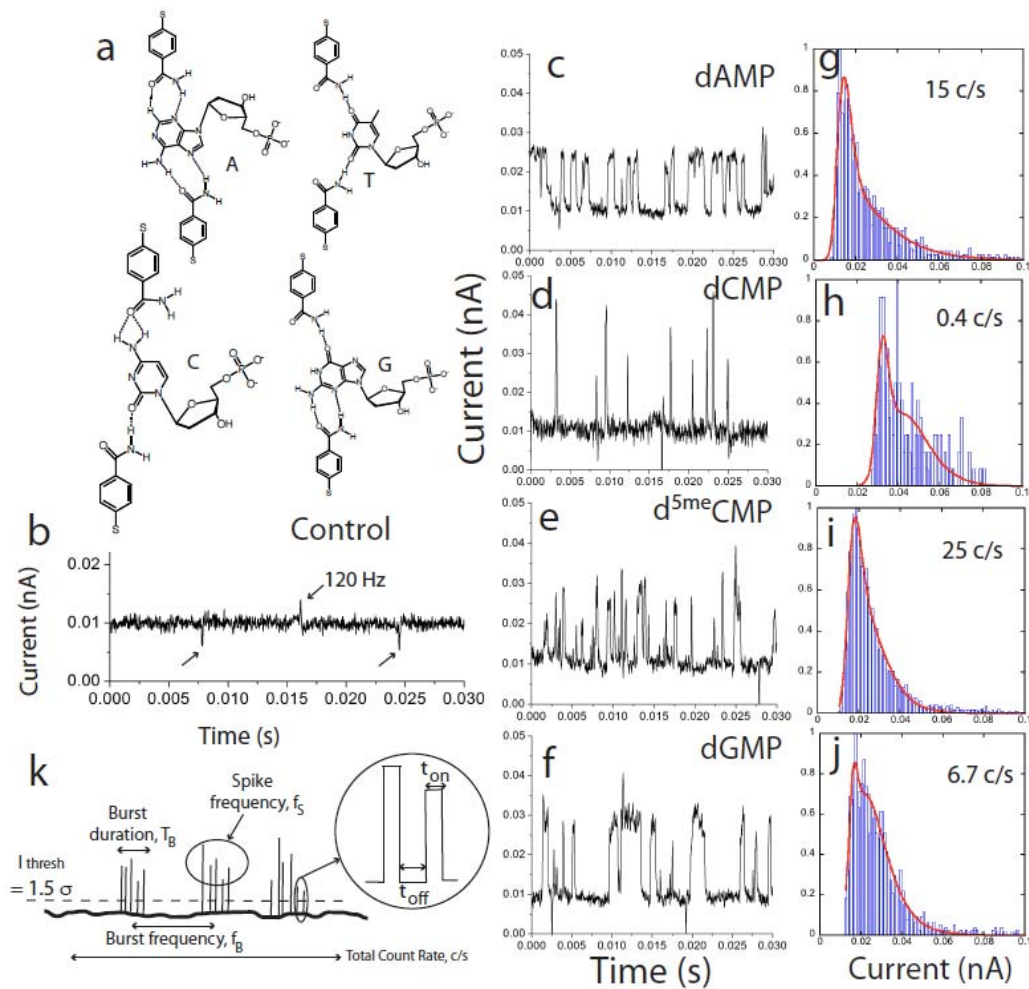
Figure 6-2Tunneling signals from nucleotides trapped in a functionalized tunnel gap.

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 ^{5me}C would not be accounted for by these

structures alone. In phosphate buffered saline, but in the absence of analyte, a 20 pS gap

(i= 10 pA, V = + 0.5V) gave a signal free of features, except for some AC coupled line-

noise pointed by arrows in (b). (c) – (f) Characteristic current spikes produced when

nucleotides dAMP, dCMP, d^{m}CMP and dGMP were introduced (longer signal runs are

given). dTMP gave no signals. (g) – (j) corresponding distribution of pulse heights. Red

lines are fits to two Gaussian distributions in the logarithm of current. (k) Definition of

the parameters used to characterize the tunneling signals. Spikes are counted if they exceed a threshold equal to 1.5 x the standard deviation of the noise on the local background. The signals occur in bursts (duration T_B, frequency f_B) each containing current spikes at a frequency f_S. The spikes stay high for a period t_{on} and low for a period t_{off}. The total count rate (inset in g-j) is the number of spikes in all bursts divided by the measurement time.

We first characterized the tunnel gap using doubly-distilled water and 0.1 mM phosphate buffer (PB – pH=7.4). Small signals were observed from buffer alone with bare electrodes, but they were much rarer when both electrodes were functionalized and the tunnel gap conductance set to 20 pS or less. (Fig. 6.2b). The tunnel decay was much more rapid (decay constant, $\beta = 14.2\pm3.2$ nm^{-1}) with both electrodes functionalized than is the case in water alone ($\beta \sim 6.1\pm0.7$ nm^{-1} – [96]) and we estimate that the tunnel gap at i=10 pA and V = +0.5V is a little over the length of two benzamide molecules (i.e. a little greater than 2 nm).

Introducing DNA nucleotides (10 μM in PB) into the tunnel gap yielded characteristic noise spikes as shown in Figs. 6.2c-f. The signal count rate (defined in Fig. 6.2k) varied considerably from 25 counts/s (5-methyl-2'-deoxycytidine 5'-monophophate, d^{m}CMP) to less than 1 c/s (2'-deoxycytidine 5'-monophophate, dCMP). No signals were recorded at all with thymidine 5'-monophophate (dTMP), the signal looking exactly like the control (Fig. 6.2b). STM images suggest that this nucleotide binds to the surface (and presumably the probe) very strongly, blocking interactions in which a single molecule spans the junction.

The current occurs in bursts of spikes and distributions of the spike heights were quite well fitted with two Gaussians distributions of the logarithm of current[24] as shown in Figs 6.2 g-j. These histograms were generated by counting only pulses that exceeded 1.5x the SD of the local noise background – i.e., typically pulses above 6 pA (a full description of the analysis procedure is given by Chang et el.[24]).

dCMP generates the highest signals and the lowest count rate while 2'-deoxyadenosine 5'-monophophate (dAMP) and d^mCMP produce the smallest signals and the highest count rate (we found little difference between 2'-deoxycytidine and 5-methyl-2'-deoxy cytidine in organic solvent[24]). The three bases with narrower pulse height distributions (dAMP, d^mCMP and GMP) often show bursts of "telegraph-noise" characteristic of sources that fluctuate between two levels[94] (particularly marked for dAMP). Such a two-level distribution is a strong indication that the tunneling signals are generated by a single molecule trapped in the tunnel junction.[94] The characteristics of the tunneling noise from the nucleotides are summarized in Table 6.1.

Table 6.1
Nucleotide tunneling noise characteristics. Parameters are defined in Figure 2k.

Nucleotide	dAMP	dGMP	dCMP	d^mCMP
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
τ_{on} (ms)	0.38±0.01*	0.48±0.02*	0.42±0.02*	0.31±0.09*

τ_{off} (ms)	0.35±0.01*	0.56±0.04*	0.71±0.06*	0.41±0.11*
τ_{on}/τ_{off}	~1	0.9	0.6	0.8
ΔG (kT units)	0	0.1	0.51	0.22

*Error in fit to exponential distribution. §Standard error

dAMP signals are well-separated from dCMP signals, and d^{m}CMP signals are well separated from dCMP signals in spike amplitude and in the time distribution of their signals (Table 6.1). For this reason, we chose to investigate DNA oligomers composed of A, C and ^{m}C bases.

Figs. 6.3a,c and e show representative tunneling noise traces for d(A)$_5$, d(C)$_5$ and d(^{m}C)$_5$ with the corresponding current peak distributions shown in Figs. 6.3b, d and f. Comparing Fig. 6.3b (d(A)$_5$) with Fig. 6.2g (dAMP), Fig. 6.3d (d(C)$_5$) with Fig. 6.2h (dCMP) and Fig. 6.3f (d(^{m}C)$_5$) with Fig. 6.2i (d^{m}CMP) leads to the following startling conclusion: *most of the polymer binding events in the tunnel junction generate signals that resemble those generated by single nucleotides*. That this should be so is not obvious. It requires (1) that single bases are being read and (2) that steric constraints owing to the polymer backbone do not prevent base-binding events from dominating the signals.
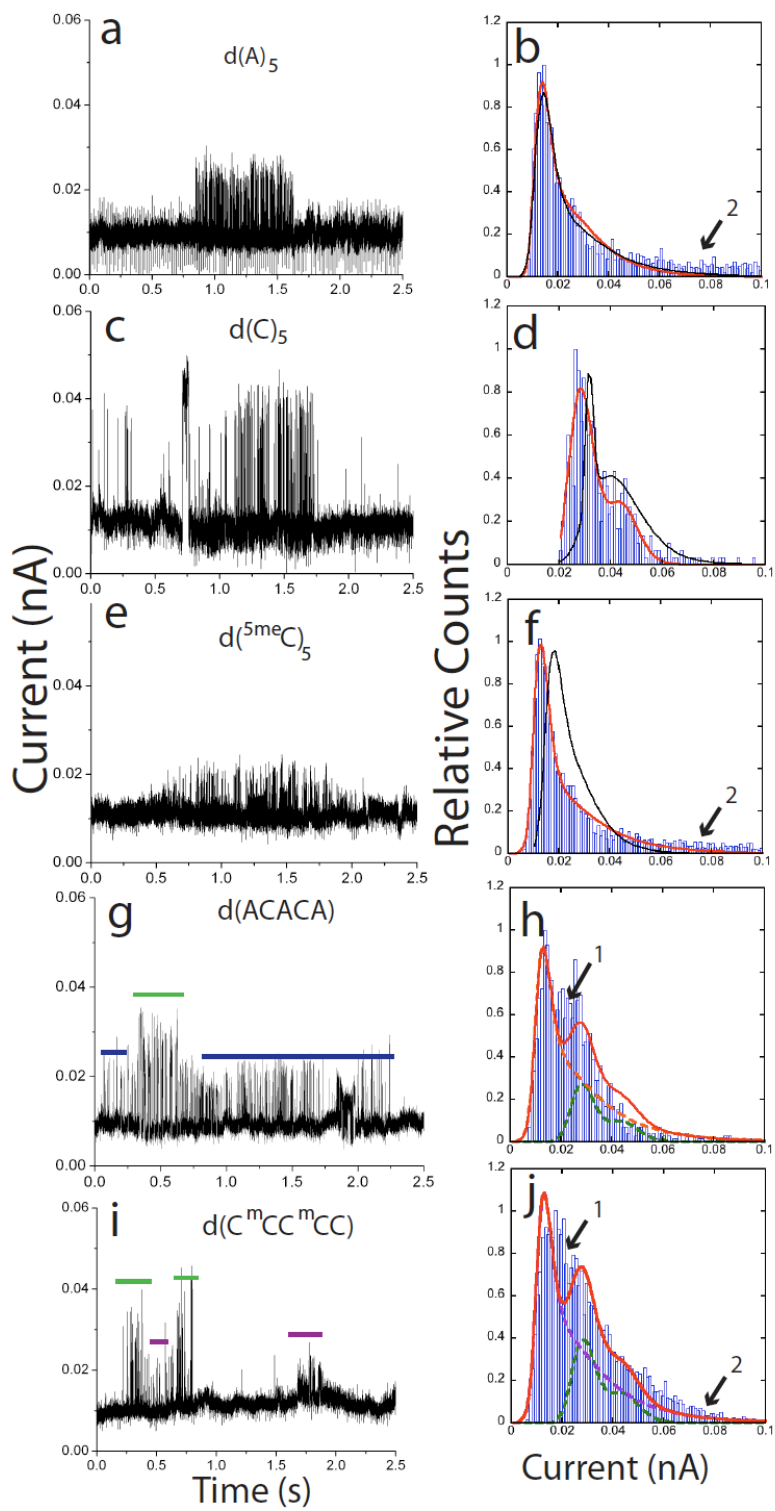
Figure 6-3 Tunneling signal distributions from oligomers resemble those of the constituent nucleotides.

Tunneling signal distributions from oligomers resemble those of the constituent nucleotides. (a, c, e) Representative current traces from $d(A)_5$, $d(C)_5$ and $d(^mC)_5$ with the corresponding distributions shown in b, d and f. Red lines are fits with parameters similar to those used for nucleotides. The black lines are the fits to the corresponding nucleotide distributions shown in Figs. 6.2 g,h and i. "HCT" labels some of the high current features seen in homopolymers but not nucleotides (b and f). (g and i) Current traces from mixed oligomers, $d(ACACA)$ and $d(C^mCC^mCC)$, with corresponding current distributions (h and j). Red lines are scaled homopolymer fits, with the green-dashed line showing the "C" contribution, the orange-dashed line showing the "A" contribution and the purple-dashed line the ^mC contribution. The data are well described by the homopolymer parameters though some intermediate signals ("1") and new high current features (2) show that the sequence context affects the reads a little. The colored bars on the current traces mark bursts of A-like signals (orange), C-like signals (green) and ^mC-like signals (purple).

There are some (small) differences between nucleotide and oligomer signals: (1) Peak positions, widths and relative intensities are altered somewhat (the also the nucleotide distributions which have been replotted on top of the homopolymer distributions as the black lines on Figs 6.3b,d and f.). (2) Almost all of the signals generated by nucleotides are less than 0.1 nA at 0.5V bias (Table 6.1). In contrast, 20% of the total signals generated by $d(A)_5$ and $d(^mC)_5$ are larger than 0.1 nA at this bias (Table 6.2 - this is not obvious in Figure 6.2 where distributions are plotted only up to 0.1nA). These high current (>0.1 nA) features in $d(A)_5$ and $d(C)_5$ are continuously distributed so they do not represent parallel reads of more than one base at a time (where currents would be distributed in

multiples of the single molecule values[3]). Rather, they are new features

associated with the presence of the polymeric structure in the tunnel gap. Such a

non-specific, large amplitude spike is labeled by an asterisk in Fig. 6.1b.

Table 6.2
Oligomer tunneling noise characteristics. Parameters are defined in Figure 2k.

Oligomer	$d(A)_5$	$d(C)_5$	$d(^mC)_5$
Burst Duration (T_B, s)	$0.14\pm0.02^*$	$0.15\pm0.03^*$	$0.41\pm0.03^*$
Burst Frequency (f_B Hz)	$738\pm100^\S$	$320\pm85^\S$	$662\pm116^\S$
Fraction of reads > 0.1 nA	0.20	0.0	0.23
τ_{on} (ms)	$0.33\pm0.01^*$	$0.34\pm0.02^*$	$0.26\pm0.01^*$
τ_{off} (ms)	$0.52\pm0.02^*$	$0.42\pm0.01^*$	$0.47\pm0.01^*$
τ_{on}/τ_{off}	0.6	0.8	0.6
ΔG (kT units)	0.51	0.22	0.51

[*]Error in fit to exponential distribution. [§]Standard error

Features at I > 0.1 nA appear much less frequently in oligomers of mixed

sequence, suggesting that they are associated with base-stacking in the

homopolymers. Fig. 6.3h shows a current distribution for d(ACACA) where 95%

of events are below 0.1 nA. Fig. 6.3j shows a current distribution for d(C^mCC^mCC)

where 99% of events are below 0.1 nA. The solid red lines are the sums of the

distributions measured for the homopolymers corresponding to the constituents

with, scaling aside, only one fitting parameter. This parameter is the ratio, r_{fit} , of

the A/C (r_{fit} =0.48) or ^mC/C (r_{fit} =0.66) contributions. These values differ from

the known composition ratios (0.6 for ACACA and 0.4 for C^mCC^mCC) but are

surprising in as much as the spike rate for dCMP alone is very small, yet C

appears to be quite well represented in the mixed sequence oligomer data. This suggests that Cs surrounded by As are read more frequently, possibly because the C-containing oligomer is better attached to the substrate than the isolated dCMP.

Most importantly, mixed oligomers generate signals that are largely described as the sum of the individual base signals. (Some intermediate current reads, labeled "1" in Figs 6.3h and j, and a small number of additional high current features – labeled "2" - show that sequence context plays a small role.)

In these experiments, the probe drifts randomly over the samples, so the sequence is not "read" deterministically. Nonetheless we can readily find traces in which the signals alternate between "A-like" and "C-like" (Fig. 6.3g) and "^{m}C-like" and "C-like" (Fig 6.3i). The duration of these "bursts" (see Fig. 6.2k) of signals is long (0.14 ± 0.02s in ACACA and 0.15 ± 0.02s in C^{m}CC^{m}CC). Similar bursts are seen in the homopolymers (Table 6.2) and the nucleotides (Table 6.1). This leads us to our second unexpected conclusion: *the lifetime of the bound complex in the tunnel gap is very long* (fraction of a second) compared to either the interval between noise spikes (ms) or the lifetime of the bound-state in solution.
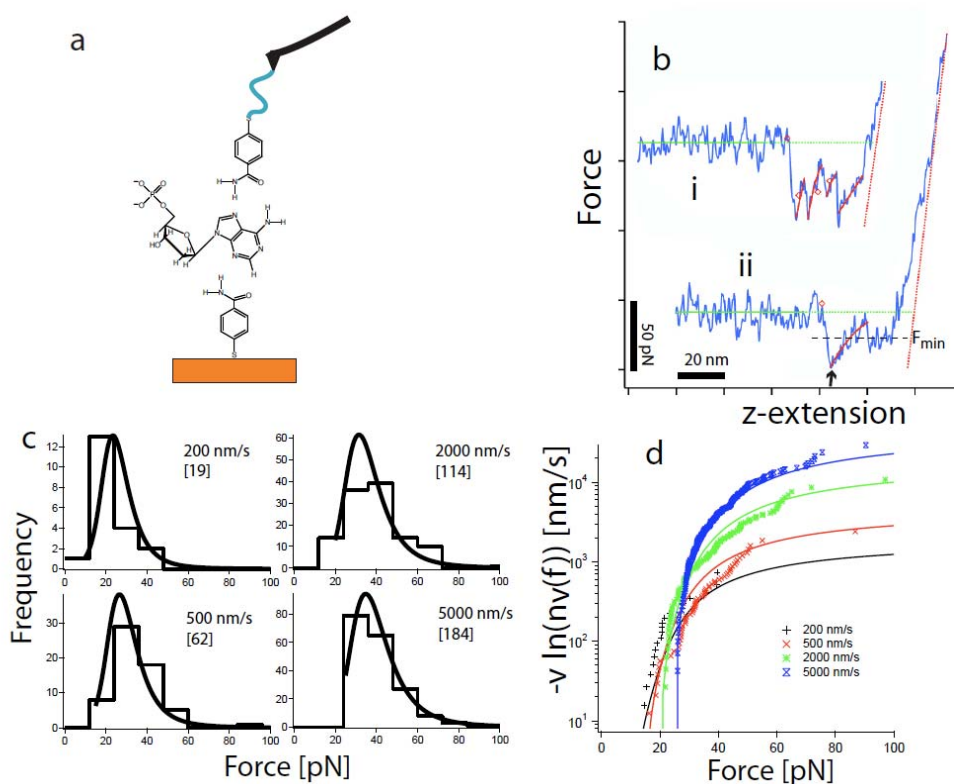
Figure 6-4 The lifetime of the reading complex is on the order of a second at zero force.

The lifetime of the reading complex is on the order of a second at zero force. (a) AFM gap
functionalization where the blue line represents a 34 nm PEG linker. (b) representative
force curves showing: (i) Pulling on more than one molecule at a time – the force
baseline is not restored after each break, and the z-extension (corrected for tip
displacement) is > 34 nm. (ii) A single molecule curve of the type accepted by the
software. The force returns to the baseline after the bond breaks and the corrected
extension is ~ 34 nm. (c) Histograms of bond breaking forces at the pulling speeds
marked. The solid lines are maximum likelihood fits to the heterogenous bond model. (d)
Bond survival probability plotted versus bond breaking force for the four pulling speeds,
fitted by the same heterogeneous bond model parameters (solid lines). These fits yield a
zero-force off rate of 0.28 s^{-1} implying that the assembly lives for times on the order of
seconds in a nanogap, much longer than the lifetime in solution. For details see ref. 14.

We have used dynamic force spectroscopy as an independent test of the unexpectedly long lifetime of the benzamide-base-benzamide complex confined to a nanoscale gap. In these measurements (Fig. 6.4a) one of the recognition molecules was bound to an AFM probe via a 34 nm long polyethyleneglycol (PEG) linker (Methods) while the other formed a monolayer on an Au(111) substrate. dAMP was used as the target analyte to bridge the gap. In the absence of dAMP, adhesion between probe and substrate was extremely small, presumably because the hydrogen bonding sites on the benzamide recognition molecules were stably bound by water molecules. Adhesion features were observed in the presence of a small amount of dAMP, falling as the concentration of dAMP increased (resulting in binding of both probe and substrate by dAMP). Stretching of the PEG tether generated a characteristic signal that permitted multiple binding events (Fig 6.4b (i)) to be separated from single molecule events (Fig. 6.4b (ii)) so that only single-molecule bond-breaking events were analyzed.[97] Single molecule bond-breaking forces as a function of pulling speed are summarized in Fig. 6.4c (solid lines are maximum likelihood fits to a heterogeneous bond model[97],[98]) and the bond survival probability as a function of bond-breaking force is shown in Fig. 6.4d. The solid lines are fits to the same heterogeneous bond model.[98] They yield an off-rate at zero force, K_{off}^{0} = 0.28 s^{-1}. *Thus the intrinsic (zero-force) survival time of this complex is on the order of seconds, not milliseconds.* The analysis also yields the distance to the transition state for dissociation, α = 0.78 nm (as well as its variance, σ = 0.19 nm). We conclude that each base resides in the tunnel junction for a significant fraction

of a second, while generating tunneling signals at kHz rates. Thus the entire

cluster of signals that occur in one burst (burst durations are listed in Tables 6.1

and 6.2) can be used to characterize a base.

Long-bound-state life-times accompanied by rapid fluctuations in

electronic signatures have been reported previously in STM images[21] and in the

effect of single-molecule reactions on transport in carbon nanotubes.[48] The

origin of this noise is unclear, save that it appears to be very temperature sensitive,

indicative of small energy barriers to the motion that causes the noise.[21]

Following Goldsmith et al.[48] we have analyzed the distribution of "on" and "off"

times (see Fig. 6.2k). In a limited time range of times, determined by the

amplifier response at one end, and the servo response time at the other,[24] these

distributions are exponential (as expected for a Poisson source) and the 1/e times

(τ_{on} and τ_{off}) are listed in Tables 6.1 and 6.2. They do not differ much, and

calculating an energy difference, ΔG, between the on and off states from

$\Delta G = kT_B \ln(\tau_{off}/\tau_{on})$ yields the values listed in the Tables (in units of thermal

energy, $k_B T$, at 300K). These values are all a fraction of $k_B T$. Thus the

"switching" cannot represent thermal activation over a significant barrier (the

normal source of two-level noise). One possible explanation is Brownian motion

in a bound state sampled by an exponentially-sensitive matrix element.

The "on" and "off" times are so broadly distributed that they are not very

useful for identifying base-signals. However, the frequency within a burst (f_S

Tables 6.1 and 6.2) is a much simpler parameter. These show the current

distributions and frequency distributions for the three homopolymers, normalized

so that the area under each curve is unity. The frequency distribution for $d(^mC)_5$ is bimodal, with many reads in the "C" frequency range and a number at the very fast rate (ca. 1300 Hz) observed for ^mC MP alone (labeled f(^mCMP) on the figure) . This suggests that the binding modes of ^mC are altered significantly in a polymer context (consistent with the larger shift of the polymer signal compared to the nucleotide signal, Fig. 6.3f) so we chose to analyze oligomers containing A and C, in particular the d(CCACC) sequence shown in Fig. 6.1a.

Given an average current in a burst, $\langle i \rangle$ and frequency, $\langle f \rangle$, the distributions shown in, $I_{A,C}(\langle i \rangle)$ and $F_{A,C}(\langle f \rangle)$ determine independent probabilities that a base is an A or a C: $P_{A,C}^i = \dfrac{I_{A,C}(\langle i \rangle)}{I_A(\langle i \rangle) + I_C(\langle i \rangle)}$ and

$P_{A,C}^f = \dfrac{F_{A,C}(\langle f \rangle)}{F_A(\langle f \rangle) + F_C(\langle f \rangle)}$. The current distribution from d(CCACC) is almost completely dominated by A spikes (the component of the C distribution in this fit is 7% or less). This is a surprising result, that more C's in the sequence give a smaller number of C spikes. But it is consistent with our hypothesis that the frequency of C reads is increased when the base is flanked by A's (c.f. the increase in C reads in d(ACACA compared to the dCMP vs. dAMP count rate).

Armed with our analysis of the burst signals, we can now make quantitative assignments of mixed signals (this was done "by eye" in Figs 6.3g and i). d(C)₅ produces no signals below 0.015 nA, so bursts of current below this level (but above the noise) can be unambiguously assigned to A. For larger

amplitude signals we use both the frequency and amplitude data as described. The result is the pair of curves shown in Fig. 6.1e.

Using this approach to sequence DNA requires several further developments. Firstly, the polymer must be pulled through a tunnel junction at a controlled speed, particularly if homopolymer runs are to be read. Since DNA passes through unfunctionalized nanopores too rapidly to be read[65] the long residence time of bases in a functionalized tunnel junction is an asset. At present, movement from one site to another is driven by uncontrolled mechanical drift that generates unknown forces on the reading complex. Our force spectroscopy data can be used to give a crude estimate of the "pulling" force that would be needed to achieve a given read rate (assuming the measured off-rate for dAMP to be representative for all bases). The Bell equation gives the off rate at a force F as

$$K_{off} = K_{off}^0 \exp\left(\frac{F\alpha}{k_B T}\right)$$ so, with $K_{off}^0 = 0.28$ s^{-1} and $\alpha = 0.78$ nm, 19 pN would result

in passage of 10 bases per second. A rate of 10 bases s^{-1} gives about 30 data spikes (on average) for a "C" read, enough to generate an assignment with a reasonable level of confidence. A force of 19 pN can be generated by a bias of just 80 mV across a nanopore[99] so read rates of 10 bases per second per tunnel junction seem feasible.

The second requirement for a practical sequencing system is a better recognition chemistry in which there is a much larger separation of the current distributions from all five bases. New compounds are presently under study in our lab.

6.3 Methods

Nucleoside 5'-monphosphates (from Sigma-Aldrich) were used as supplied.

DNA oligomers were synthesized and characterized by IDT and used without

further purification. Synthesis and characterization of other materials are

described.   Gold probes were etched as described previously[24] and coated with

high-density polyethylene[27] to leave a fraction of a micron of exposed gold

(optical and TEM characterization is described). These probes gave no

measureable DC leakage, important as this can be a source of distortion of the

tunneling signal.[24]   Capacitative coupling of 120 Hz switching signals was a

problem (Fig. 2b) minimized by careful control of the coating profile.   The gaps

were characterized by recording current decay curves as a function of distance

starting at 20 pS, a distance that gave no signals in buffer alone with

functionalized electrodes.   Current signals were recorded using an Agilent

PicoSPM together with a digital oscilloscope controlled by a custom Labview

program. The servo response time was set to about 30 ms as described

previously.[24] This places an upper limit on undistorted measurements of pulse

widths of a few ms. Analysis of current distributions was automated using the

software described elsewhere.[24]   Force spectroscopy was carried out with a

MFP3D AFM (Asylum Research, Santa Barbara). Heterobifunctional PEG linkers

(MAL-SVA 3400 from Lysan Bio) of 34 nm extended length were attached at one

end to silicon nitride AFM probe (Veeco MSNL - spring constant = 0.02 N/m)

and mercapto-benzamide molecules attached to the remaining maleimide as

described elsewhere[100]. Force curves were taken in 1 mM PB buffer with an

initial 10 μM concentration of dAMP in the gap adjusted by rinsing. Force curves

were analyzed using custom software[101].

Chapter 7

GAP DISTANCE AND INTERACTIONS IN A MOLECULAR TUNNEL JUNCTION

This chapter describes the application of AC modulation and break junction to measure the gap distance and interactions in a molecular tunnel. It was first published in Journal of the American Chemical Society [71].

The distance between electrodes in a tunnel junction cannot be determined from the external movement applied to the electrodes because of interfacial forces that distort the electrode geometry at the nanoscale. These distortions become particularly complex when molecules are present in the junction, as demonstrated here by measurements of the AC response of a molecular junction over a range of conductivities from micro Siemens to pico Siemens. Specific chemical interactions within the junction lead to distinct features in breakjunction data and these have been used to determine electrode separation in a junction functionalized with 4(5)-(2-mercaptoethyl)-1H-imidazole-2-carboxamide, a reagent developed for reading DNA sequence.

7.1 Introduction

Adjustable tunnel junctions are widely used to determine the electrical properties of molecules spanning two electrodes.[44, 92, 94, 102] The size of the junction is usually characterized by the measured tunnel conductance of the junction, or by the amount by which a break junction is separated by an externally applied displacement. The actual size of the nanoscale gap is not readily determined from

the external measurement. When the electrodes (usually gold) are metallically

bonded, and then pulled apart to form a break junction, plastic deformation of the

gold leads to the formation of filaments that give rise to a constant conductance

over distances that approach a nm (see Figure 7.3A for an example). When the

filaments break, the metal surface relaxes back to a more stable configuration (i.e.,

they "snap back"). On approach of two metal surfaces, the electrodes can be

drawn together by van der Waals interactions when they are close together,

leading to instabilities in the gap. The electrodes can be repelled as contamination

is trapped in the gap, leading to apparent approach distances of 100 nm or more.

It is important to be able to determine the real value of the nanoscale gap when

the goal of these studies is to make devices that utilize fixed, nanofabricated

tunnel junctions.

We have proposed a read-out system for DNA sequences based on a non-
covalent complex between recognition molecules tethered to fixed electrodes and
the bases of a DNA molecule that is passed through the tunnel junction by
electrophoresis.[103] Here, we report on a study of tunnel junctions based on
gold electrodes functionalized with a new generation of recognition molecule,
with 4(5)-(2-mercaptoethyl)-1H imideazole-2-carboxamide (hereafter
imideazole-2-carbamide - Fig. 7.1B). The synthesis and characterization of this
molecule has been described elsewhere.[104] We have used ac modulation of the
gap[105-107] as a probe of the effective stiffness of the gap, using a logarithmic
current to voltage converter to allow us to probe a range of gap conductances
from close to quantum point contact (conductance = $G_0 = \frac{2e^2}{h} = 77\mu S$) all the way

out to the small conductances used to pass DNA bases between the electrodes (~ 6 pS). These measurements reveal the many interactions taking place in the tunnel junction, interactions that make it impossible to determine the gap from tunnel-current data alone. We determine the gap size using molecules trapped in the gap as a "molecular ruler", molecular tunneling signals ceasing when the gap size exceeds the size of the trapped molecules.[23, 108] The gap size determined in this way is about 24 Å, more than large enough to pass a single-stranded DNA molecule.

7.2 Methods

Monolayers of imideazole-2-carboxamide were formed on freshly flame-annealed Au(111) electrodes and characterized with FTIR, ellipsometry, XPS and STM. Importantly, ellisometry and XPS taken together suggest that the molecules stand upright on the surface with the S bonded to gold, forming a film consistent with the full 8.5 Å length of the recognition molecules. STM probes were etched from gold wire, insulated with high-density polyethylene and functionalized as described previously[103]. The functionalization of the probes was tested by comparing tunneling signals obtained on bare substrates with tunneling signals obtained from bare (i.e., unfunctionalized) probes on functionalized substrates. Tunneling measurements were recorded with an Agilent PicoSPM (Chandler AZ) interfaced to a digital storage oscilloscope and a field programmable gate array controller (PCIe-7842R, National Instruments). The entire junction was submerged in the 1 mM phosphate buffered (pH=7) aqueous electrolyte used for readout of DNA sequence.
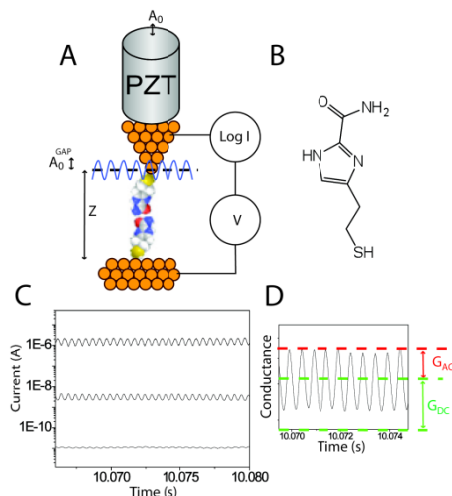
Figure 7-1Set up for determining junction stiffness

Set up for determining junction stiffness: A. An ac modulation applied to the PZT deflects it by an amount A_0. The modulation amplitude of the gap itself, A_0^{GAP} is different because of interaction forces in the tunnel junction. B. Imidazole-2-carboxamide. C. AC modulation of the PZT results in a corresponding modulation of the gap conductance that increases with the DC set-point current according to $G^{AC} = \beta A_0^{GAP} G^{DC}$. D. Showing how the quantities G_{AC} and G_{DC} are extracted.

The modulated-junction method is illustrated in Fig. 7.1. A small distance modulation, A_0, is applied to the gap by adding an ac signal to the voltage used to drive the PZT in the vertical direction. The PZT sensitivity was calibrated by using STM images of single atom steps on Au(111) and an amplitude, A_0, of 0.52 Å was used in the present work. This is small enough that the following expression holds for the AC modulation of the gap conductance: $G_{AC} = \beta A_0^{GAP} G_{DC}$, where G_{DC} is the average gap value and A_0^{GAP} is the modulation of the gap. Thus the AC signal increases in direct proportion to the DC value, as illustrated with traces of the tunneling signal in Fig. 7.1C. The ratio, $\frac{G_{AC}}{G_{DC}}$ (Fig. 7.1D) yields the quantity βA_0^{GAP}. The whole problem lies with the fact that $A_0^{GAP} \neq A_0$, the applied modulation, because of the small compliance of the STM probe.[15] Mechanical

106

interactions in the junction show up as rapid variations in the value of "β" derived

by assuming a constant A_0. In this work, signals were acquired using a

logarithmic current-to-voltage converter[109], calibrated as described.
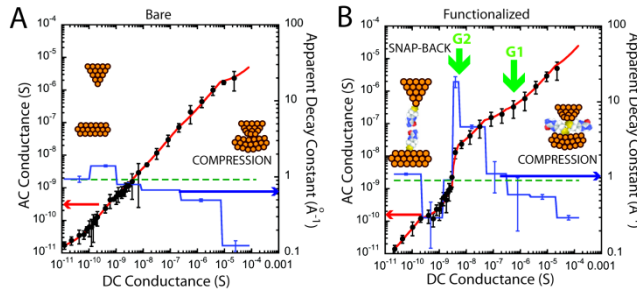
Figure 7-2Showing how GAC varies with GDC for A, a bare gold probe and bare gold surface

Showing how G_{AC} varies with G_{DC} for A, a bare gold probe and bare gold surface in 1 mM phosphate buffer (pH 7.0) and B, both probe and surface functionalized with imidazole-2-carboxamide in the same buffer (black dots). The red lines are linear fits to segments of the plots using the effective decay constants (i.e., the apparent β values) as shown by the blue lines (righthand axes of the plots). Rapid changes at the points labeled G1 and G2 are coincident with peaks in the conductance distributions measured by break junction methods.

Fig. 7.2A shows how G_{AC} varies with G_{DC} for a bare gold probe and a bare

gold substrate in a 1 mM phosphate buffer (pH=7). These data show the mean

(data points) and standard deviation for 39 recordings at each G_{DC}, spanning a

range from about 20 μS to 10 pS (at a probe bias of +0.5V). Data were fitted with

six linear segments, yielding the values of the apparent decay constant, β_{app} (

$\beta_{app} = \frac{1}{A_0}\frac{\partial G_{AC}}{\partial G_{DC}}$). They recapitulate the data first reported for aqueous electrolyte by

Vaught et al.[96] Vaught et al. made measurements using a perchlorate electrolyte,

while we used 1mM phosphate buffer – evidently, these ions do not play a

significant role. At small gap conductances $\beta_{app} \sim 0.9$ E^{-1}, falling as the gap

conductance rises above 10^{-8} S. The fall is particularly marked above 10^{-5} S

where β_{app} falls to almost 0.1 E^{-1}, a consequence of strong mechanical interactions

in the gap. In the absence of such interactions, the gap could be estimated from the sum of the distances, z_{nm}, corresponding to each of the linear segments of constant β_{nm} (i.e., $z = \sum -\frac{1}{\beta_{nm}} \ln\left(\frac{G_m}{G_n}\right)$ where G_n is the conductance at the high side of the linear segment and G_m is the conductance on the low side). Carrying this sum out yields $z = 38$ Å, an unrealistically large gap because it has been exaggerated by the very small value of β_{app} at the highest conductance.

When both electrodes are functionalized with imideazole-2-carboxamide (Fig. 7.2B) the shape of the curve changes dramatically. There is a "bump" between 10^{-7} S> G_{DC} >10^{-9} S where β_{app} reaches the extraordinary value of 18.4 Å^{-1}. Interestingly, β_{app} is not as much reduced at small gaps compared to the unfunctionalized system, possibly reflecting a reduced gold-gold interaction. The very large values of β_{app} are most readily explained by regions in which bonds between the probe and surface break, resulting in a snapping back of the electrode surfaces and a large change in current for a small motion of the probe. Jumps in the slope of this plot at $\sim 10^{-7}$ S and 10^{-9} S are associated with molecular adhesion events as confirmed by break junction data[44]. Figure 7.3A shows some typical current vs. retraction-distance curves for both the probe and surface functionalized. Plateaus are evident at $\sim 10^{-7}$ S (i.e., $\sim 10^{-2}$ G_0, labeled G1) and $\sim 10^{-9}$ S (i.e., $\sim 10^{-4}$ G_0, labeled G2). Curves taken with a bare probe (example in red) show only the plateaus near the quantum of conductance (G0). A histogram (Fig. 7.3B) of the current recorded from one thousand such curves shows distinct peaks at 6.2 ±2.3 x 10^{-7} S (i.e., $\sim 10^{-2}$ G_0 - G1) and 5.8±0.7 x 10^{-9} S (i.e., $\sim 10^{-4}$ G_0

- G2).   The likely origin of the two peaks is immediately clear when similar

curves are collected and histogrammed using a bare tip and a functionalized

surface (Fig. 7.3D).   In this case, only the peak G1 appears in addition to the

metallic contact peak at G0.   Thus, the high current peak (at G1) is assigned to

one molecule spanning the gap (via an amine-gold linkage[110]) while the low

current peak (G2) is assigned to hydrogen-bonded pairs bridging the gap.   Time

traces of the tunnel-current noise taken at conductances above G1 and G2 (Fig.

7.3C) show the characteristic "telegraph" noise fluctuations owing to stochastic

bond-breaking in a molecule spanning the junction.[23, 108] No signals are seen

when the set-point is below G1 or G2 (note that signals owing to hydrogen-

bonded pairs of molecules spanning the gap are too small to be seen in the traces
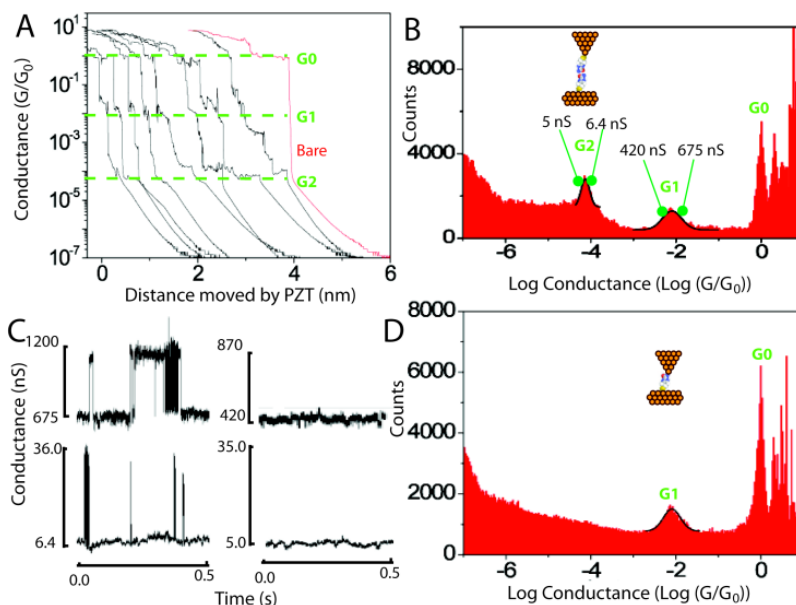
collected near G1).

Figure 7-3Break junction measurements of imideazole-2-carbamide functionalized tunnel junctions.

Break junction measurements of imideazole-2-carbamide functionalized tunnel junctions. A. Typical current-time plots (converted to apparent gap size using the PZT velocity). Distinct plateaus occur near G0, owing to the formation of quantum point contacts.

109

Features near G1 and G2 are associated with molecular structures formed in the gap. The red curve is typical of data collected with unfunctionalized probes and surfaces. No plateaus are seen below G0. B. Distribution of conductances for both probe and surface functionalized. The lower conductance peak (G2) is not present when only the surface is functionalized (D) showing that this peak is associated with pairs of molecules spanning the gap, an interpretation quantitatively consistent with the values of G1 and G2. C. Telegraph noise owing to stochastic bonding of single molecules across the gap (upper left) or pairs of molecules in series (lower left). The noise vanishes when the gap conductance is adjusted to just below G1 or G2 (traces on the right).

Is this assignment of G1 to one molecule and G2 to a pair in series consistent with the observation that $\beta \sim 0.5$ to 0.6 when the molecules are interacting (Fig. 7.2B) $\beta = -\dfrac{\ln\left(\dfrac{G_{1,2}}{G_0}\right)}{z_{1,2}}$ and using G1 = 621 nS (Fig. 7.3B) and $z_1 = 8.5$ E gives $\beta = 0.57$ E^{-1}, while using G2 = 5.8 nS (Fig. 7.3B) and $z_2 = 17$ E gives $\beta = 0.58$ E^{-1}, consistent with one molecular length trapped in the gap at G1 and two molecular lengths trapped in the gap at G2.

The magnitude of the telegraph-noise measures how much the conductance increases when molecules bond the two electrodes together. For the single molecule, the conductance increase on bonding is 432 nS, while for the two molecules in series this value is 28 nS. Based on distance differences alone and assuming that β is still 0.6#E^{-1} leads to an estimate of $\exp(-\beta L + 2\beta L) = \exp(\beta L) = 164 : 1$ for the ratio of the two conductances. This is much larger than the observed ratio (of 15:1) showing that details of the bonding play a large role in determining the size of the telegraph-noise. Note that events that cause "electronic bond fluctuations" of the electrodes are not the same as making and breaking of the chemical bonds between molecules and electrodes, as discussed in detail elsewhere.[103]

We are still left with the question of what the final gap is at the 6 pS conductance (G_{SP}) used for identifying DNA bases. Given that the gap at G2 is 17 Å, and taking the real value of β in water to be 0.92 (Fig. 7.2A - it appears to be less just after the bonds break in the case of the functionalized junction – Fig. 7.2B – because of attraction between the bonding groups) we estimate the additional distance from G1 to G_{SP} to be 7.5 Å for a total gap size of about 24 Å.

7.3 Conclusion

In summary, we have shown how the formation of molecular complexes in a tunnel junction is signaled by changes in the elastic properties of the tunnel junction, evident in the ac response

of the junction, and how the structures themselves can be used as nano-scale "rulers" for determining the size of the nanojunction.

Chapter 8

SUMMARY AND FUTURE DIRECTIONS FOR RECOGNITION TUNNELING

Single molecules in a tunnel junction can now be interrogated reliably using chemically-functionalized electrodes. Monitoring stochastic bonding fluctuations between a ligand bound to one electrode and its target bound to a second electrode ("tethered molecule-pair" configuration) gives insight into the nature of the intermolecular bonding at a single molecule-pair level, and defines the requirements for reproducible tunneling data. Importantly, at large tunnel gaps, there exists a regime for many molecules in which the tunneling is influenced more by the chemical identity of the molecules than by variability in the molecule-metal contact. Functionalizing a pair of electrodes with recognition reagents (the "free analyte" configuration) can generate a distinct tunneling signal when an analyte molecule is trapped in the gap. This opens up a new interface between chemistry and electronics with immediate implications for rapid sequencing of single DNA molecules.

8.1 Recognition Tunneling for Detection of single molecule analytes
The foregoing discussion of the "tethered molecule-pair" configuration (Figure 8.1A) sets the stage for design of a new type of analytical system that traps analytes, the "free analyte" configuration, illustrated in Figure 8.1B. We recently implemented such a scheme for recognition of DNA nucleosides[24] using 4-mercapto benzoic acid as the recognition molecule (here R1=R2 in Figure 8.1C).

In organic solvent (trichlorobenzene) the hydroxyl-group oxygen is protonated, serving as a proton donor, while the carbonyl oxygen serves as a proton acceptor. Pairs of this reagent can form four distinctive hydrogen bonded complexes with all four DNA bases as shown in Figure 8.1 A-D.[24] The choice of gap size is critical. Any larger, and the smaller bases (like T) will not be trapped on both sides. On the other hand, smaller gaps offer the possibility of multiple binding arrangements for larger bases (like A). At this large gap value (2.5 nm, calculated from measured tunnel decay rates[18]) the background tunnel current is just 6 pA (at a bias of 0.5V) so that the small gap conductance (12 pS) ensures that we are in the "non-contact" regime where contact geometry variations in conductance appear to be insignificant. The signal (Figure 1.1E) is also free of the noise spikes that occur in smaller junctions where the benzoic acids molecules can bond with each other across the junction. When nucleosides are injected into the solvent that surrounds the tunnel junction, current spikes are seen almost immediately (Figure 8.1F). These events often show the telegraph noise (inset) characteristic of binding and unbinding events. This particular reading reagent gives distinctive current reads for each of the four bases, and experimentally measured current distributions are shown in Figure 8.1G. Reads are not completely separated, but the overlap is such that a correct assignment can be made at the 0.6 confidence level (or better) on the first read. We have also shown that mixtures of bases give clearly resolved signals.[24]
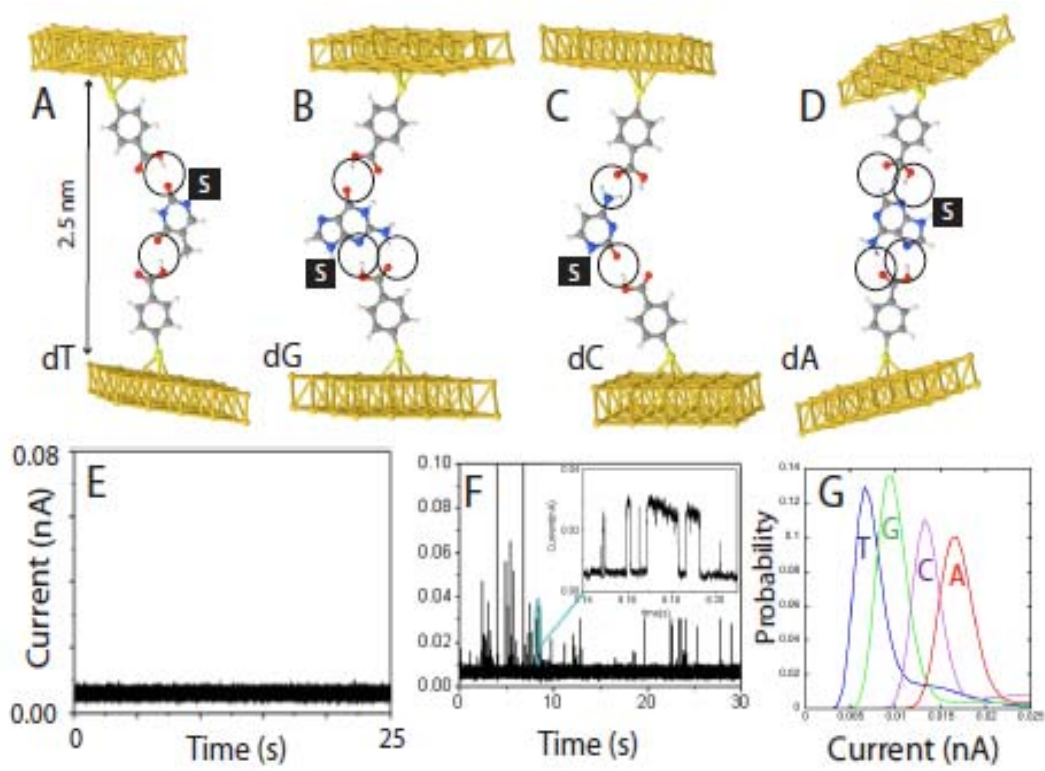
Figure 8-1 "Free-analyte" configuration of Recognition Tunneling for reading DNA bases.

"Free-analyte" configuration of Recognition Tunneling for reading DNA bases. (A-D)

show energy-minimized structures for the four nucleosides bound in a 2.5 nm gap with 4-

mercapto benzoic acid as the reading reagent (R1, R2 in Figure 1.1C). The "S" stands for

the deoxyribose sugar (not shown) and the order (dT, dG, dC, dA) corresponds to the

predicted order of increasing tunnel conductance using density functional theory. (E)

Shows the background tunnel current in organic solvent (trichlorobenzene) with the gap

set to $G_{BL} = 12$ pS (6pA at 0.5V bias). At this gap there is no indication of interactions

between the two benzoic acid readers. (F) Shows an example of the current spikes that

are observed when a solution of dG is injected into the tunnel junction. The inset shows

details of some of the spike on a ms-timescale. Many of them show the telegraph noise

switching characteristic of single molecule binding (the slight slope in the "on" level

reflects the action of the servo used to control the tunnel gap). (G) Measured

distributions of current for the four bases. The order agrees with the density-functional

114

prediction, but the measured currents are larger than predicted. The overlap between

reads limits the probability of a correct assignment on a single read to about 60%.

It is instructive to probe the role of the recognition reagents by comparing

reads with, and without them. Figure 8.2A shows a distribution of currents

measured in a somewhat smaller gap (baseline conductance = 20 pS) with a bare

gold substrate and probe. The target is deoxyguanosine (the gap had to be

decreased somewhat in order to observe current spikes with no functional groups

on the electrodes). When one electrode in the gap is functionalized (and set to a

baseline conductance of 12 pS) the resultant distribution is dramatically narrowed

(Figure 8.2B). A second small peak at twice the current of the first peak appears,

a characteristic feature of junctions like this when two molecule span the gap.[3]

Functionalizing both electrodes does not decrease the width of the distribution

any further. It appears that the target is trapped by the hydrogen bonds,

eliminating motional broadening of the tunnel distribution, but this is not the

correct interpretation of the narrowing. This can be seen from the measured

distributions of on-state lifetimes shown in Figure 8.2C. The distributions are

very similar for bare electrodes (blue line) and one functionalized electrode (red).

Thus the trapping time of molecules in the gap does not depend strongly on the

chemical functionalization of the gap. Similar results were obtained for the other

three bases.[24] It is known that the bases interact quite strongly with gold via the

lone-pair electrons on the imines and amines, so the reads taken with bare

electrodes almost certainly reflect transient binding of the targets to the gold

115

electrodes. Thus the observed narrowing of the current distribution on

functionalizing one of the electrodes reflects a reduction in the range of *static*

bound conformations. Tunneling reads require an interaction between the target

molecules and the electrodes. Recognition Tunneling makes this a controlled

interaction.
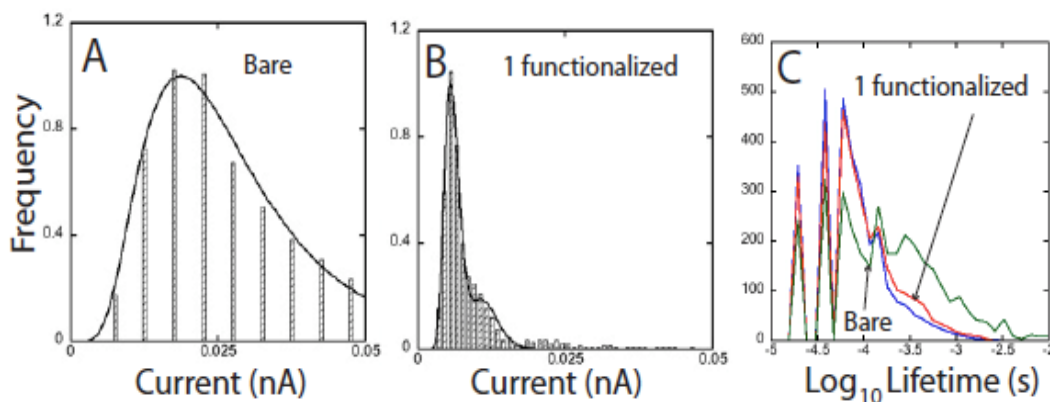
Figure 8-2Current distributions and binding modes

Current distributions and binding modes: (A) Shows the current distribution measured for

a pair of bare gold electrodes for dG. G_{BL} was increased to 20 pS to obtain these reads.

(B) Shows the current distribution for dG with just one electrode functionalized. (C)

Shows the distribution of on-state lifetimes for bare electrodes (blue), one electrode

functionalized (red) and both electrodes functionalized (green). The fine structure

reflects data binning. The three distributions are similar, implying that the narrowing that

occurs on functionalizing one electrode comes from a smaller range of bound

configurations in the gap and not a slowing of DNA motion.

8.2 The future

Clearly, Recognition Tunneling is a powerful new tool for investigating bonding

in a small junction. It is clearly capable of single molecule detection and might

enable a new class of single molecule sensors. As far as DNA sequencing goes, three questions remain: (1) Can the bases be distinguished with reads made in aqueous buffered electrolyte? (2) Can single base resolution be obtained in the context of a DNA polymer where each base is separated by about 0.3 nm? (3) Can chemical modifications of the bases (important as epigenetic coding) be read? Preliminary work in our laboratory suggests that the answer to all three questions is positive. Electrochemical leakage in aqueous buffer is not a problem with good insulation[26] of tunneling probes. Signal levels appear to be even larger that those found in organic solvent, despite the expectation that hydrogen bonding between the recognition elements is expected to be weaker in water because of competition from water molecules.

Evidence for single base reads is seen in long runs of telegraph noise as small DNA oligomers with an alternating base sequence are trapped in the gap. If more than one type of base contributed to the signal, this two level switching would be hard to account for. Finally, experiments with cytosine monophosphate and 5-methylcytosine monophosphate show that there are clear differences in the size of the tunneling signals. It appears that hydration is playing a role here, because no difference was observed between deoxycytidine and 5-methyl-deoxycytidine measured in organic solvent.[24] Methylation has long been known to affect the melting temperature of double helical DNA, reflecting its effects on the bonding between helices in water.[111]

Recognition tunneling is not limited to DNA targets. Half of the amino acid residues have hydrogen-bonding sites in addition to those on the peptide

117

backbone itself. Five (asn, glu, gln, his, arg) have at least two or more sites while another five (lys, ser, thr, try, trp) have one other site available. This raises the possibility of partial reads of peptide sequences using Recognition Tunneling. Many published paper have successfully shown that electrical devices based on molecules could be used as electronic components such as diodes[112], switches[113-115], rectifiers.[116, 117] 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.[118] 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.
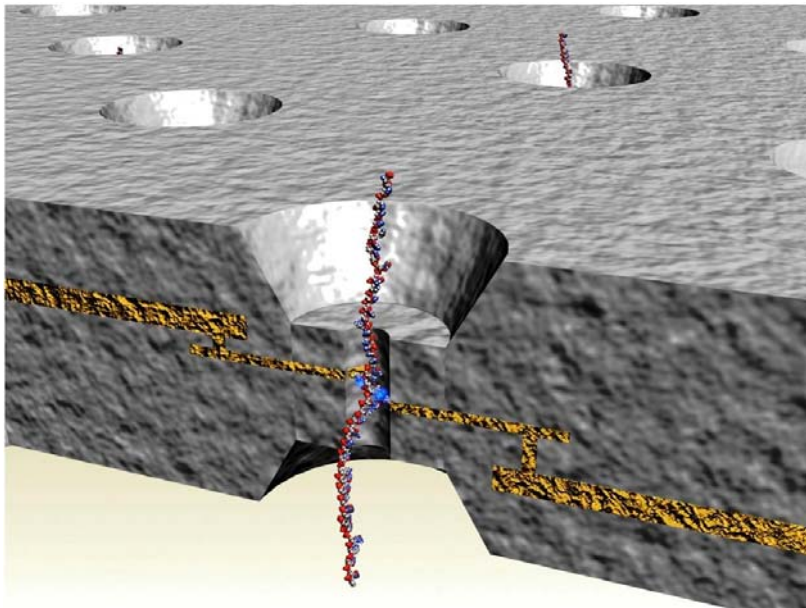
Figure 8-3The schematic diagram of DNA translocation through a nanopore.

In this thesis, I described the work depends on the quantum mechanical phenomena of electron tunneling. In this chapter, I first introduce a simple explanation of vacuum tunneling, then described tunneling through a tunnel junction containing a molecule, and finally introduce our invention, which we call "Recognition Tunneling" a method for chemical recognition in a tunnel junction. Much of the material in this chapter is reproduced from our research of nanotechnology. I mainly show that 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.

119

Single base resolution on short DNA oligomers could be resolved with DNA reader functionalized tips already.

In recent study, tip scanning on short oligomers have shown potential sequence readout signals with correct time contant (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, see Figure 8.3, 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]   Datta S 1990 *Electronic transport in mesoscopic systems* (Cambridge: Cambridge University Press)

[2]   Datta S 2005 *Quantum Transport: Atom to Transistor*: Cambridge Univeristy Press)

[3]   Lindsay S M and Ratner M A 2007 Molecular Transport Junctions: Clearing Mists *Advanced Materials* **19** 23-31

[4]   Moth-Poulsen K and Bjørnholm T 2009 Molecular electronics with single molecules in solid-state devices *Nature Nanotechnology* **4** 551-6

[5]   Cui X D, Primak A, Zarate X, Tomfohr J, Sankey O F, Moore A L, Moore T A, Gust D, Harris G and Lindsay S M 2001 Reproducible measurement of single-molecule conductivity *Science* **294** 571-4

[6]   Lindsay S M 2009 *Introduction to Nanoscience* (Oxford: Oxford University Press)

[7]   Chen F, He J, Nuckolls C, Roberts T, Klare J and Lindsay S M 2005 A molecular switch based on potential-induced changes of oxidation state *Nano Letters* **5** 503-6

[8]   Stuart L and et al. 2010 Recognition tunneling *Nanotechnology* **21** 262001

[9]   Hamers R J, Tromp R M and Demuth J E 1986 Surface Electronic Structure of Si (111)-(7×7) Resolved in Real Space *Phys. Rev. Lett.* **56** 1972-5

[10]   Kuznetsov A M and Ulstrup J 1999 *Electron transfer in chemistry and biology* (New York: Wiley)

[11]   Shapir E, Cohen H, Calzolari A, Cavazzoni C, Ryndyk D A, Cuniberti G, Kotlyar A, Felice R D and Porath D 2008 Electronic structure of single DNA molecules resolved by transverse scanning tunnelling spectroscopy *Nature Materials* **7** 68-74

[12]   Tanaka H and Kawai T 2009 Partial sequencing of a single DNA molecule with a scanning tunnelling microscope *Nature Nanotechnology* **Published online July 5**

[13]   Tao N 1996 Probing potential-tuned resonant tunneling through redox molecules with scanning tunneling microscopy. *Phys. Rev. Letts.* **76** 4066-9

[14]   Ohshiro T and Umezawa Y 2006 Complementary base-pair-facilitated electron tunneling for electrically pinpointing complementary nucleobases *Proc. Nat. Acad. Sci.* **103** 10-4

[15]   Chang S, He J, Kibel A, Lee M, Sankey O F, Zhang P and Lindsay S M 2009 Tunneling readout of hydrogen-bonding based recognition *Nature Nanotechnology* **4** 297-301

[16]   Frisbie C D, Rozsnyai F, Noy A, Wrighton M S and Lieber C M 1994 Functional group imaging by chemical force microscopy *Science* **265** 2071-4

[17]   Hinterdorfer P, Baumgartner W, Gruber H J, Schilcher K and Schindler H 1996 Detection and localization of individual antibody-antigen recognition events by atomic force microscopy *Proc. Natl. Acad. Sci. (USA)* **93** 3477-81

[18]   He J, Lin L, Zhang P and Lindsay S M 2007 Identification of DNA base-pairing via tunnel-current decay *Nano Letters* **7** 3854-8

[19]   He J, Lin L, Liu H, Zhang P, Fu Q, Lee M, Sankey O and Lindsay S 2009 A hydrogen-bonded electron-tunneling circuit reads base composition of unmodified DNA *Nanotechnology* **20** 075102 1-8

[20]   Chang S, He J, Lin L, Zhang P, Liang F, Young M, Huang S and Lindsay S 2009 Tunnel conductance of Watson-Crick nucleoside-basepairs from telegraph noise *Nanotechnology* **20** 075102-10

[21]   Ramachandran G K, Hopson T J, Rawlett A M, Nagahara L A, Primak A and Lindsay S M 2003 A Bond-Fluctuation Mechanism for Stochastic Switching in Wired Molecules *Science* **300** 1413-5

[22]   Haiss W, Nichols R J, Zalinge H v, Higgins S J, Bethell D and Schiffrin D J 2004 Measurement of single molecule conductivity using the spontaneous formation of molecular wires *Phys. Chem. Chem. Phys.* **6** 4330 – 7

[23]   Haiss W, Wang C, Grace I, Batsanov A S, Schiffrin D J, Higgins S J, Bryce M R, Lambert C J and Nichols R J 2006 Precision control of single-molecule electrical junctions *Nature Materials* **5** 995 - 1002

122

[24] Chang S, Huang S, He J, Liang F, Zhang P, Li S, Chen X, Sankey O F and Lindsay S M 2010 Electronic Signature of all four DNA Nucleosides in a Tunneling Gap *Nano Letters* **10** 1070-5

[25] DeRose J A, Thundat T, Nagahara L A and Lindsay S M 1991 Gold Grown Epitaxially on Mica: Conditions for Large Area Flat Faces *Surf. Sci.* **256** 102-8

[26] Nagahara L A, Thundat T and Lindsay S M 1989 Preparation and Characterization of STM Tips for Electrochemical Studies *Rev. Sci. Instrum.* **60** 3128-30

[27] Visoly-Fisher I, Daie K, Terazono Y, Herrero C, Fungo F, Otero L, Durantini E, Silber J J, Sereno L, Gust D, Moore T A, Moore A L and Lindsay S M 2006 Conductance of a biomolecular wire *Proc. Nat. Acad. Sci.* **103** 8686–90

[28] Jeffrey G A and Saenger W 1991 *Hydrogen bonding in biological structures* (Berlin: Springer)

[29] Zwolak M and Di Ventra M 2005 Electronic Signature of DNA Nucleotides via Transverse Transport *Nano Lett.* **5** 421-4

[30] Troisi A and Ratner M A 2005 Modeling the inelastic electron tunneling spectra of molecular wire junctions *Phys. Rev. B* **72** 033408 1-4

[31] Wuttke D S, Bjerrum M J, Winkler J R and Gray H B 1992 Electron-Tunneling Pathways in Cytochrome c *Science* **256** 1007 - 9

[32] Cui X D, Primak A, Zarate X, Tomfohr J, Sankey O F, Moore A L, Moore T A, Gust D, G. H and Lindsay S M 2001 Reproducible measurement of single-molecule conductivity *Science* **294** 571-4

[33] Cantor C R and Schimmel P R 1980 *Biophysical Chemistry* (San Francisco: W.H. Freeman)

[34] Tomfohr J K and Sankey O F 2004 Theoretical analysis of electron transport through organic molecules *J. Chem. Phys.* **120** 1542- 15554

[35] Mujica V, Kemp M and Ratner M A 1994 Electron conduction in molecular wires. I. A scattering formalism *J. Chem. Phys.* **101** 6849-55

[36] Landauer R 1989 Conductance determined by transmission: Probes and quantized constriction resistance *J. Phys. Condens. Matter* **1** 8099-110

[37]   Imry Y and Landauer R 1999 Conductance viewed as transmission *Revs. Mod. Phys.* **71** S306-S12

[38]   Pethica J B and Oliver W C 1987 Tip Surface Interactions in STM and AFM *Physica Scripta* **T19** 61-6

[39]   Israelachvilli J N 1991 *Intermolecular and Surface Forces* (New York: Academic press)

[40]   Williams J M, Han T and Beebe T P 1996 Determination of Single-Bond Forces from Contact Force Variances in Atomic Force Microscopy *Langmuir* **12** 1291-5

[41]   Hutter J L and Bechhoefer J 1993 Calibration of atomic force microscope tips *Rev. Sci. Instrum.* **64** 1868-73

[42]   Zwolak M and Di Ventra M 2008 Physical approaches to DNA sequencing and detection *Reviews of Modern Physics* **80** 141-65

[43]   Chang S, He J, Kibel A, Lee M, Sankey O F, Zhang P and Lindsay S M 2009 Tunneling readout of hydrogen-bonding based recognition *Nature Nanotechnology* **accepted for publication**

[44]   Xu B and Tao N J 2003 Measurement of Single-Molecule Resistance by Repeated Formation of Molecular Junctions *Science* **301** 1221-3

[45]   Venkataraman L, Klare, J.E., Tam, I.W., Nuckolls, C., Hybertsen, M.S., Steigerwald, M.L. 2006 Single-Molecule Circuits with Well-Defined Molecular Conductance *Nano Lett.* **6** 458-62

[46]   Branton B, Deamer D, Marziali A, Bayley H, Benner S A, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich S B, Krstic P S, Lindsay S, Ling X S, Mastrangelo C H, Meller A, Oliver J S, Pershin Y V, Ramsey J M, Riehn R, Soni G V, Tabard-Cossa V, Wanunu M, Wiggin M and Schloss J 2008 Nanopore Sequencing *Nature Biotechnology* **26** 1146-53

[47]   Basu R, Tovar J D and Hersam M C 2005 Scanning tunneling microscopy study of single molecule motion on the Si(100)-2 X 1 surface *J. Vac. Sci. Technol. B* **23** 1785-9

[48]   Goldsmith B R, Coroneus J G, Kane A A, Weiss G A and Collins P G 2008 Monitoring single-molecule reactivity on a carbon nanotube *Nano Letters* **8** 189-94

124

[49]   Jeffrey G A 1997 *An Introduction to Hydrogen Bonding* (Oxford: Oxford University Press)

[50]   Tomfohr J and Sankey O F 2002 Complex bandstructure, decay lengths and Fermi level alignment in simple molecular electronic systems *Phys. Rev. B* **65** 245105--12

[51]   Lee M 2009 Electronic Structure Simulations od DNA base recognition and Vibrational Property Analysis of Polyanionic Hydrides. In: *Physics,* (Tempe: Arizona State University) p 178

[52]   Lee M H and Sankey O F 2009 Theory of tunneling across hydrogen-bonded base pairs for DNA recognition and sequencing *Phys. Rev .E* **79** 051911 1-10

[53]   Kresse G and Furthmuller J 1996 Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set *Phys. Rev. B* **55** 11169-86

[54]   Perdew J P and Wang Y 1992 Accurate and simple analytic representation of the electron-gas correlation energy *Phys. Rev. B,* **45** 13244

[55]   Blanco J M, Flores F and Perez R 2006 STM-theory: Image potential, chemistry and surface relaxation *Prog. Surf. Sci.* **81** 403

[56]   Jelinek P, Wang H, Lewis J P, Sankey O F and Ortega J 2005 Multicenter approach to the exchange-correlation interactions in ab initio tight-binding methods *Phys. Rev .B* **71**:235101

[57]   Jelinek P, Perez R, Ortega J and Flores F 2004 Mechanical properties and electrical conductance of different Al nanowires submitted to an homogeneous deformation: a first-principles simulation *Surf. Sci,* **566-568** 1-13

[58]   Rzeznicka I I, Lee J, Maksymovych P and Yates J T 2005 Nondissociative Chemisorption of Short Chain Alkanethiols on Au(111) *J. Phys. Chem. B* **109** 15992-6

[59]   Henningfeld K A, Arsian T and Hecht S M 1996 Alteration of DNA primary structure by DNA topoisomerase I. Isolation of the covalent topoisomeraseI - DNA binary complex in enzymatically competent form *J. Am. Chem. Soc.* **118** 11701 - 13

[60]   Yelm K E 1999 A Simple Method for in situ Generation of Thiols from Thioacetates *Tetrahedron Letters* **40** 1101-2

[61] Zheng T-C, Burkart M and Richardson D E 1999 A General and Mild Synthesis of Thioesters and Thiols from Halides *Tetrahedron Letters* **40** 603-6

[62] Lewis P A, Inman C E, Yao Y, Tour J M, Hutchison J E and Weiss P S 2004 Mediating Stochastic Switching of Single Molecules Using Chemical Functionality *J. Am Chem Soc* **126** 12214-5

[63] Lee M H and Sankey O F 2008 Insights into electron tunneling across hydrogen-bonded base-pairs in complete molecular circuits for single-stranded DNA sequencing *J. Phys. C: Condensed Matter* **21** 035110 1-11

[64] Lee J W and Thundat T 2005 DNA and RNA sequencing by nanoscale reading through programmable electrophoresis and nanoelectrode-gated tunneling and dielectric detection (US

[65] Branton D, Deamer D, Marziali A, Bayley H, Benner S A, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich S B, Krstic P S, Lindsay S, Ling X S, Mastrangelo C H, Meller A, Oliver J S, Pershin Y V, Ramsey J M, Riehn R, Soni G V, Tabard-Cossa V, Wanunu M, Wiggin M and Schloss J 2008 Nanopore Sequencing *Nature Biotechnology* **26** 1146-53

[66] Lindsay S, He J, Sankey O, Hapala P, Jelinek P, Zhang P, Chang S and Huang S 2010 Recognition Tunneling *Nanotechnology* **accepted**

[67] Lee M 2009 Electronic Structure Simulations of DNA base recognition and Vibrational Property Analysis of Polyanionic Hydrides. In: *PhD in Physics,* (Tempe: Arizona State University) p 178

[68] Kimura-Suda H, Petrovykh D Y, Tarlov M J and Whitman L J 2003 Base-Dependent Competitive Adsorption of Single-Stranded DNA on Gold *J. Am Chem Soc* **125** 9014-5

[69] Ostblom M, Liedberg B, Demers L M and Mirkin C A 2005 On the Structure and Desorption Dynamics of DNA Bases Adsorbed on Gold: A Temperature-Programmed Study *J. Phys. Chem. B.* **109** 15150-60

[70] Xia J L, Diez-Perez I and Tao N J 2008 Electron Transport in Single Molecules Measured by a Distance-Modulation Assisted Break Junction Method *Nano Lett.* **8** 1960-4

[71] Li X, He J, Hihath J, Xu B, Lindsay S M and Tao N J 2006 Conductance of Single Alkanedithiols: Conduction Mechanism and Effect of Molecule-Electrode Contacts *J. Am. Chem. Soc.* **128** 2135-41

126

[72]   Wu S M, Gonzalez M T, Huber R, Grunder S, Mayor M, Schoenenberger C and Calame M 2008 *Nature Nanotechnology* **3** 569-74

[73]   Pennish E 2007 Human Genetic Variation *Science* **318** 1842-3

[74]   Sharp A J, Cheng Z C and Eichler E E 2006 Structural variation of the human genome *Annu. Rev. Genomic Hum. Genet.* **ARI** 407-42

[75]   Akeson M, Branton D, Kasianowicz J J, Brandin E and . D W D 1999 Microsecond timescale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules *Biophys J.* **77** 3227-33

[76]   Meller A, Nivon L, Brandin E, Golovchenko J and Branton D 2000 Rapid nanopore discrimination between single polynucleotide molecules *Proc. Natl. Acad. Sci. (USA)* **97** 1079-84

[77]   Ashkenasy N, Sanchez-Quesada J, Bayley H and Ghadiri M R 2005 Recognizing a single base in an individual DNA strand: A step toward DNA sequencing in nanopores. *Angew. Chem. Int. Ed.* **44** 1401-4

[78]   Meller A, Nivon L and Branton D 2001 Voltage-driven DNA translocations through a nanopore *Phys. Rev. Lett.* **86** 3435-8

[79]   Aksimentiev A, Heng J B, Timp G and Schulten K 2004 Microscopic kinetics of DNA translocation through synthetic nanopores *Biophysical Journal* **87** 2086-97

[80]   Muthukumar M and Kong C Y 2006 Simulation of of polymer translocation through protein channels *Proc. Natl. Acad. Sci. (USA)* **103** 5273-8

[81]   Lagerqvist J, Zwolak M and Di Ventra M 2007 Influence of the Environment and Probes on Rapid DNA Sequencing via Transverse Electronic Transport *Biophys J.* **93** 2384–90

[82]   Zikic R, Krstić P S, Zhang X-G, Fuentes-Cabrera M, Wells J and Zhao X 2006 Characterization of the tunneling conductance across DNA bases *Phys. Rev. E* **74** 011919 1-9

[83]   He J, Lin L, Liu H, Zhang P, Lee M, Sankey O F and Lindsay S M 2009 A hydrogen-bonded electron-tunneling circuit reads the base composition of unmodified DNA *Nanotechnology* **20** 075102-10

[84]   Concetration Conctrations are nominal - see supporting information

[85]   A. H. Schäfer, Seidel C, Chi L and Fuchs H 1998 STM Investigations of Thiol Self-Assembled Monolayers *Adv. Mat.* **10** 839-42

[86]   Creager S E and Steiger C M 1995 Conformational Rigidity in a Self-Assembled Monolayer of 4-Mercaptobenzoic Acid on Gold *Langmuir* **11** 1852-4

[87]   Current_Range Disparity in range less if we note T baseline is for non bined and add

[88]   Meunier V and Krstić P S 2008 Enhancement of the transverse conductance in DNA nucleotides *J. Chem. Phys.* **128** 041103-1-4

[89]   Bare_reads Bare-electrode reads show some sensitivity to the identity of the nucleosides (with purine generating more current than pyrimidines – Table S1)

[90]   Liu H, He J, Tang J, Liu H, Pang P, Cao D, Krstic P S, Joseph S, Lindsay S and Nuckolls C 2010 Translocation of single-stranded DNA through single-walled carbon nanotubes *Science* **in press**

[91]   Krstic P S, Wells J C, Fuentes-Cabrera M, Xu D and Lee J W 2007 Toward Electronic Conductance Characterization of DNA Nucleotide Bases *Solid State Phenomena* **121-123** 1387-90

[92]   M.Tsutsui, Taniguchi M, Yokota K and Kawai T 2010 Identification of Single Nucleotide via Tunnelling Current *Nature Nanotechnology* **5** 286-90

[93]   Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S and Bayley. H 2009 Continuous base identification for single-molecule nanopore DNA sequencing *Nature Nanotechnology* **4** 265 - 70

[94]   Lindsay S, He J, Sankey O, Hapala P, Jelinek P, Zhang P, Chang S and Huang S 2010 Recognition Tunneling *Nanotechnology* **21** 262001-13

[95]   Gross L, Mohn F, Moll N, Liljeroth P and Meyer G 2009 The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy *Science* **325** 1110-4

[96]   Vaught A, Jing T W and Lindsay S M 1995 Non-exponential tunneling in water near an electrode *Chemical Physics Letters* **236** 306-10

[97]   Fuhrmann A, Anselmetti D, Ros R, Getfert S and Reimann P 2008
         Refined procedure of evaluating experimantal single-molecule force
         spectroscopy data *Phys Rev. E* **77** 03912-22

[98]   Getfert S and Reimann P 2007 Optimal evaluation of single-molecule
         force spectroscopy experiments *Phys Rev E* **76** 052901-5

[99]   Keyser U F, Koelman B N, van Dorp S, Krapf D, Smeets R M M, Lemay
         S G, Dekker N H and Dekker C 2006 Direct force measurements on DNA
         in a solid-state nanopore *Nature Physics* **2** 473-7

[100]   Ashcroft B, Spadola Q, Qamar S, Zhang P, Kada G, Bension R and
         Lindsay S M 2008 An AFM/Rotaxane Molecular Reading Head for
         Sequence-Dependent DNA Structure *Small* **4** 1468-75

[101]   Fuhrmann A 2010 Force Spectroscopy, from Single Molecules to Whole
         Cells: Refined Procedures of Data Analysis. In: *Physics,* (Tempe: Arizona
         State University)

[102]   Venkataraman L, Klare J E, Nuckolls C, Hybertsen M S and Steigerwald
         M L 2006 Dependence of single-molecule junction conductance on
         molecular conformation *Nature* **442** 905-7

[103]   Huang S, He J, Chang S, Zhang P, Liang F, Li S, Tuchband M, Fuhrman
         A, Ros R and Lindsay S M 2010 Identifying single bases in a DNA
         oligomer with electron tunneling *Nature Nanotechnology* **5** 868-73

[104]   Liang F, Li S, Lindsay S and Zhang P 2011 Chemical and Hydrogen
         Bonding Properties of Imidazole-2-carboxamide, a reagent for DNA
         Sequencing by Recognition Tunnelling *Chemistry* submitted

[105]   Xu B 2007 Modulating the Conductance of a Au–octanedithiol–Au
         Molecular Junction *Small* **3** 2061-5

[106]   Xia J L, Diez-Perez I and Tao N J 2008 Electron Transport in Single
         Molecules Measured by a Distance-Modulation Assisted Break Junction
         Method *Nano Lett.* **8** 1960-4

[107]   Zhou J, Chen G and Xu B 2010 Probing the molecule-electrode interface
         of single molecule junctions by controllable mechanical modulations *J.
         Phys. Chem. C* **114** 8587-92

[108]   Haiss W, ChangshengWang, Jitchati R, Grace I, Martın S, Batsanov A S,
         Higgins S J, Bryce M R, Lambert C J, Jensen P S and Nichols R J 2008
         Variable contact gap single-molecule conductance determination for a

series of conjugated molecular bridges *J. Phys: Condens. Matter* **20** 374119-28

[109] He J, Sankey O F, Lee M, Tao N J, Li X and Lindsay S M 2006 Measuring Single Molecule Conductance with Break Junctions *Faraday Discussions* **131** 145 - 54

[110] Quek S Y, Neaton J B, Hybertsen M S, Venkataraman L, Choi C H and Louie S G 2007 Amine-Gold Linked Single-Molecule Circuits: Experiment and Theory *Nano Letters* **7** 3477-82

[111] Ehrlich M, Ehrlich K and Mayo J A 1975 Unusual properties of the DNA from Xanthomonus phage p12 in which 5-methylcytosine copletely replaces cytosine. *Biochimica et Biophysica Acta* **395** 109-19

[112] Derosa P, Guda S and Seminario J 2003 A programmable molecular diode driven by charge-induced conformational changes *Journal of the American Chemical Society* 14240-1

[113] Mizuta H and Oda S 2008 Bottom-up approach to silicon nanoelectronics *Microelectronics Journal* 171-6

[114] Lu W and Lieber C 2007 Nanoelectronics from the bottom up *Nature Materials* 841-50

[115] Kang S 2004 Silicon nanoelectronics: Precise fabrication via a bottom-up approach *Jom* 19-

[116] Sentein C, Fiorini C, Lorin A and Nunzi J 1997 Molecular rectification in oriented polymer structures *Synthetic Metals* 81-2

[117] Sentein C, Fiorini C, Lorin A and Nunzi J 1997 Molecular rectification in oriented polymer structures *Advanced Materials* 809-&

[118] 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 and Schloss J 2008 The potential and challenges of nanopore sequencing *Nature Biotechnology* 1146-53