

Insights into electron tunneling across hydrogen-bonded base-pairs in complete molecular circuits for single-stranded DNA sequencing

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2009 J. Phys.: Condens. Matter 21 035110

(<http://iopscience.iop.org/0953-8984/21/3/035110>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 29/05/2010 at 17:26

Please note that [terms and conditions apply](#).

Insights into electron tunneling across hydrogen-bonded base-pairs in complete molecular circuits for single-stranded DNA sequencing

Myeong H Lee and Otto F Sankey

Department of Physics, Arizona State University, Tempe, AZ 85287-1504, USA

Received 7 August 2008, in final form 13 November 2008

Published 11 December 2008

Online at stacks.iop.org/JPhysCM/21/035110

Abstract

We report a first-principles study of electron ballistic transport through a molecular junction containing deoxycytidine-monophosphate (dCMP) connected to metal electrodes. A guanidinium ion and guanine nucleobase are tethered to gold electrodes on opposite sides to form hydrogen bonds with the dCMP molecule providing an electric circuit. The circuit mimics a component of a potential device for sequencing unmodified single-stranded DNA. The molecular conductance is obtained from DFT Green's function scattering methods and is compared to estimates from the electron tunneling decay constant obtained from the complex band structure. The result is that a complete molecular dCMP circuit of 'linker((CH₂)₂)–guanidinium–phosphate–deoxyribose–cytosine–guanine' has a very low conductance (of the order of fS) while the hydrogen-bonded guanine–cytosine base-pair has a moderate conductance (of the order of tens to hundreds of nS). Thus, while the transverse electron transfer through base-pairing is moderately conductive, electron transfer through a complete molecular dCMP circuit is not. The gold Fermi level is found to be aligned very close to the HOMO for both the guanine–cytosine base-pair and the complete molecular dCMP circuit. Results for two different plausible geometries of the hydrogen-bonded dCMP molecule reveal that the conductance varies from fS for an extended structure to pS for a slightly compressed structure.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The sequencing of the human genome has been completed [1, 2] including the diploid genome [3]. The next step is to sequence the genome of any specific individual. This would provide a major advance toward individualized medical treatment and risk assessment. Sequencing the DNA of an individual patient allows identification of mutations, including those associated with disease. There are an assortment of methods to sequence DNA and most rely on chemical mechanisms. The Sanger process [4] uses chemically altered dideoxynucleotide triphosphates (ddNTPs), which are lacking the 3' –OH group, to terminate the DNA strand elongation at specific bases and separate them by size to read the sequence. Sequencing by synthesis, another chemical means, is actively

being pursued [5]. The theme of much current research is to find an innovative process that is fast and inexpensive. For example, it is proposed that measuring the ionic blockade current when DNA translocates the pore is a potential method to identify each base [6–9]. An alternative technology is to sequence DNA electronically by electron transport through nucleobases to recognize them by their electron transport signature [10].

The electron transport properties of DNA has attracted much attention; they are fundamentally important and they have many potential applications [11–13]. Most experiments on the electron transport in DNA have focused on electron transfer along the axis of the backbone. Initial experiments determined hole transport along the backbone [14] and it soon became clear that the mechanism of transport

involved tunneling over short distances and hopping over long distances [15]. Additionally, there has been much interest in DNA molecules as a functional element in molecular electronics [16]. This led to the design and measurement of electric conduction along the DNA axis through construction of circuits where DNA is attached to electrodes. Porath *et al* [17] measured the current through the double-stranded poly(G)–poly(C) DNA molecules connected to two metal electrodes, and determined the voltage dependence of the differential conductance. A clear peak structure is found which is taken as evidence that electron transport is mediated by the DNA molecular energy bands. Direct conductance measurements have been performed on DNA in solution [18] and have been extended to determine changes in electrical conduction along the DNA axis to identify single nucleotide polymorphisms (SNPs) [19]. Diverse experimental and theoretical results for electronic properties of DNA are summarized by Endres *et al* [20] and the effects of DNA structure and environment such as solvent and counter-ions on the electronic states of DNA using density functional theory are discussed.

Of particular interest to DNA sequencing is transverse electron transfer across the DNA bases; that is transfer perpendicular to the helical axis rather than along it. A nanofabricated prototype device has recently been fabricated [10]. Several theoretical studies have been reported on transverse electron transfer. Yanov *et al* [21] showed that an adenine–thymine DNA base-pair sandwiched between gold electrodes exhibits significant conductance at a small voltage bias. The high conductance is partially a result of the alignment of the Au Fermi level; it is found near the LUMO of the DNA base-pair (although we do not find this to be the case in the present work). Their calculations use a semi-empirical nonequilibrium Green's function method. The effect of protonation on the electron transfer through DNA base-pairs has been investigated by Mallajosyula *et al* [22]. The base-pair is attached to gold clusters by a thiol linker ($-\text{CH}_2\text{S}$). They find that protonation changes the electronic properties of the DNA base-pairs significantly in ways that are unique to the base-pair.

The theoretical work of Di Ventra *et al* [23] concludes that each DNA nucleotide carries a unique signature due to its different electronic and chemical structure when the transverse current is measured through single-stranded DNA sandwiched between gold electrodes. This suggests the possibility of sequencing DNA by measuring the transverse current through single-stranded DNA as it translocates a pore. However, Zikic *et al* [24] argue that the current signature is dominated mainly by the variety of geometrical conformations of the bases relative to the nanoelectrodes. They performed a first-principles calculation of the current–voltage characteristics of the DNA-like nucleotides placed between gold electrodes for different geometries. Large fluctuation of current due to the orientation of the nucleotide between the electrodes makes it impossible to sequence DNA by measuring this way. Their work shows the importance of having control of the contact between the electrodes and the DNA molecule to achieve reproducibility.

Hydrogen bonding is a means to achieve reproducibility. Hydrogen bonds are weak enough to be broken in a sequencing

device yet strong enough to hold the system together, at least momentarily. Ohshiro and Umezawa [25] have shown by scanning tunneling microscopy experiments that electron transfer is enhanced by the hydrogen bond formed between Watson–Crick complementary base-pairs. This breakthrough finding provides for the possibility of using hydrogen-bonded base-pair stabilization as a method to make reliable contacts and circuits for nucleobase recognition. These experiments again emphasize the importance of making chemical contact (not just physical contact) with the molecule. This lesson was demonstrated early on with experiments on alkane chains [26], where chemical linking of both ends of the molecule to metal contacts achieved large conductance enhancements and reproducibility.

Experiments on a precursor of a device based on this bonding principle have been carried out by He *et al* [27, 28]. In the first set of experiments [27], circuits were formed between hydrogen bonding nucleosides and a nucleobase. In later experiments [28] the circuit included a full nucleotide; these experiments are the motivation of this theoretical work. In these latter experiments, monolayers of β -mercaptoethylguanidine are assembled on a gold surface onto which single-stranded DNA with a controlled sequence is deposited. A gold-coated conducting STM tip with a known thiolated nucleobase is brought to the surface, presumably forming hydrogen-bonded base-pairs and completing the circuit. The current is measured as the STM tip is withdrawn from an initial set-point current. The aim of this paper is to provide a theoretical framework to simulate such experiments and to provide estimates of the transverse conductance through hydrogen-bonded pathways. The circuit included the entire nucleotide to simulate a potential circuit of an unmodified single-stranded DNA. We connect to the STM tip by a cytosine–guanine linkage, the strongest hydrogen bonded base-pair. As we will see below, conductances are theoretically predicted to be in the fS range. A slightly compressed structure increases the conductance to the pS range. Geometrical changes of the molecule incurred by the tip may play an important role in experiments.

A schematic depiction of a potential sequencing or recognition circuit is shown in figure 1. The single-stranded DNA backbone runs vertically up the center of the diagram. The phosphates along the backbone are labeled 'P' and the DNA bases are generically labeled 'b'. A contact of the molecule to the left metal is achieved by tethering a guanidinium ion (Gd), which forms a hydrogen-bond-like interaction with a phosphate. To complete the circuit, a right contact has a 'reader' base b' which hydrogen bonds with base b . The base-pairing between b and b' need not be complementary in general. Binding to single-stranded DNA is not hampered by the confined space of the double helix, and 0, 1, 2 or 3 hydrogen bonds across base-pairs b and b' are in principle possible. In an actual sequencing device, a nanopore or nanotube will be necessary to straighten out the DNA and to pass it single file past the metallic leads with Gd and b' tethered to them. As the single-stranded DNA passes by the tethered molecules, hydrogen bonds break and reform on the next P–b section of the DNA.

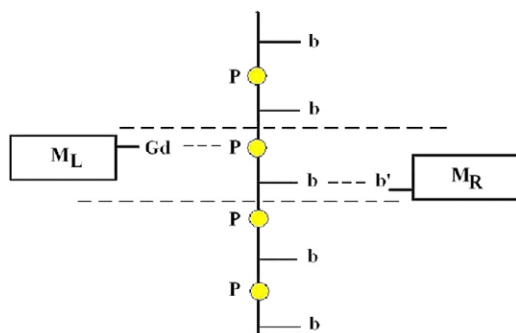


Figure 1. Schematic diagram for a DNA sequencing experiment. The left and right (M_L and M_R) metal electrodes are functionalized with a positively charged guanidinium ion (Gd) on the left side and with a reader nucleobase (b') on the right. The backbone of single-stranded DNA (ssDNA) is described as a straight line with phosphate (P). As the ssDNA translocates a nanopore, the guanidinium ion grabs the phosphate in the ssDNA to form hydrogen bonds. At the same time the reader nucleobase on the right metal (or STM tip) forms hydrogen bonds with the nucleobase on the ssDNA to form a complete molecular electrical circuit.

In this paper we explore the electron transfer through a realistic complete molecular circuit that potentially can be used to sequence unmodified single-stranded DNA polymers. We consider one base-pair, namely the complementary G–C base-pair ($b=C$ and $b'=G$). This represents a first step in simulating the reading and recognition properties of an ssDNA sequencing device which potentially involves 16 base-pair combinations. The G–C base-pair chosen here forms three hydrogen bonds and is exceptionally strong. The effects of the solvent are not included in the electron transfer calculations.

The specific core molecule that we study is shown in figure 2 and contains several components. Starting from the left electrode it contains a linker (L), a guanidinium (Gd), a phosphate (P), a deoxyribose (dR), a cytosine base (C) and a guanine base (G). We refer to it as L–Gd–P–dR–C–G. The terminal hydrogens on the linker and the guanine are replaced by sulfur to bind to gold. The left linker is $(CH_2)_2$. The deoxycytidine-monophosphate (dCMP) portion is P–dR–C and represents a nucleotide on the ssDNA. The L–Gd– is tethered to one metal contact and –G is tethered to the other metal contact.

An important and timely issue is understanding the role that hydrogen bonds play in tunneling. Hydrogen bonding is ubiquitous in biological molecules such as proteins, and is a mechanism for electron transport related to energy transduction in biological systems. In addition to studying the entire circuit, we will also take a ‘reductionist’ approach in that we will study just the base-pair portions of the molecule.

Electron transport through DNA need not be simple [20]. The transfer of charge along the DNA chain is influenced by counter-ions in solution and they play a role in gating electron transport [29]. The effect of solution is also important in electron transfer in proteins [30]. Certainly solvent effects play a role even in the transverse conduction with hydrogen bonding. However, at this initial stage, we focus on the intrinsic transport properties of the molecules themselves.

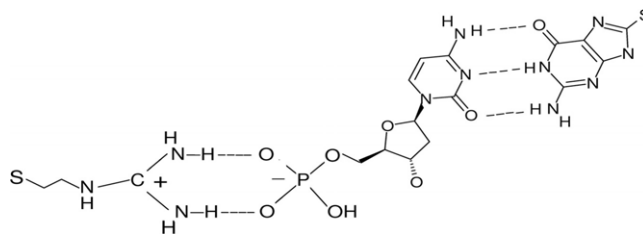


Figure 2. Schematic diagram for the sequencing molecule, L–Gd–P–dR–C–G, that forms a complete molecular electrical circuit. The molecule is composed of a linker $(CH_2)_2$, a positively charged guanidinium ion (Gd), a negatively charged phosphate (P), a deoxyribose (dR), a cytosine base (C) and a guanine base (G). The deoxycytidine-monophosphate (P–dR–C) is part of the unmodified ssDNA. The terminal hydrogen atoms on both ends of the molecule are replaced by a sulfur. The guanidinium and the phosphate form two hydrogen bonds (and are electrostatically attracted) and the cytosine–guanine base-pair forms three hydrogen bonds.

2. Structures

The geometrical structure of the molecules is obtained from energy minimization from quantum chemistry¹ calculations in vacuum using DFT with a 6-31+G(d,p) basis at the restricted Hartree–Fock (RHF) level. Several theoretical investigations related to the hydrogen bond length and bond energy of DNA base-pairs have been performed [32–39] and comparisons made [37, 38] between different DFT exchange–correlation potentials (B3LYP, BLYP, BP86, PBE, PW91, etc). Tsuzuki and Lüthi [40] showed that PW91 performs best for interaction energies of weakly bound systems via van der Waals or hydrogen bond interactions. The hydrogen bond length is the relevant issue in tunneling. We use the hybrid B3LYP GGA exchange–correlation energy functional [41] for geometry optimization. This is because gas phase calculations using B3LYP shows less mean absolute deviation [38] of hydrogen bond lengths of C–G from experiment [42] (0.067 Å for B3LYP and 0.087 Å for PW91), even though B3LYP tends to predict longer hydrogen bond lengths compared to Hartree–Fock-based MP2 results. However, we caution that, for calculations which include nucleobases, water and ions, BP86 and PW91 perform best for both hydrogen bond lengths and bond energies of A–T and G–C base-pairs [38].

To reduce the computational time required for a large molecule like L–Gd–P–dR–C–G, the final geometry was obtained in two steps. First, the relaxed geometry of the G–C base-pair was determined, which determined the three hydrogen bond lengths of the duplex which were $d(H_C \cdots O_G) = 1.74$ Å, $d(N_C \cdots H_G) = 1.89$ Å and $d(O_C \cdots H_G) = 1.91$ Å, and are in excellent agreement with similar calculations [37–39].

The second step is to determine the geometrical structures of the linker–guanidinium–phosphate–sugar–cytosine (L–Gd–P–dR–C) section. This was obtained, keeping in mind how the STM experiments will make contact with the molecule. In an STM experiment, a functionalized tip with a nucleobase

¹ The General Atomic and Molecular Electronic Structure System (GAMESS) is an *ab initio* quantum chemistry package. The version of GAMESS that we used is described in [31].

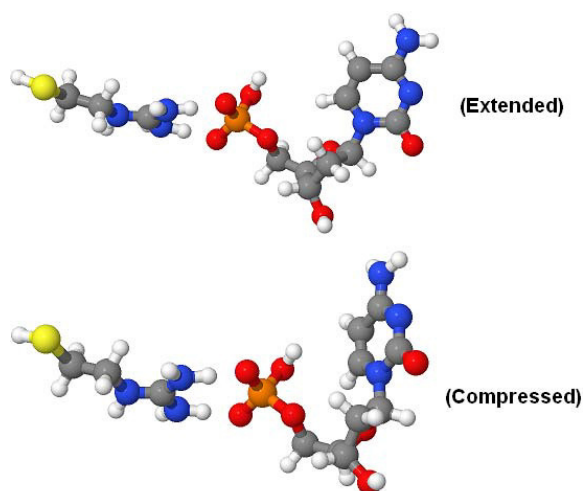


Figure 3. The geometry of L-Gd-P-dR-C for two different configurations. (a) The extended structure and (b) the compressed structures.

(G) is pushed onto a gold surface with the L-Gd-P-dR-C molecule attached. The G-C hydrogen bond is formed and the tip is withdrawn from the surface. The process is dynamical and there is no single unique molecular geometry. To gain insight into this complex process, we consider two structures for L-Gd-P-dR-C; an extended one and a compressed one. They were obtained by energy minimization (in vacuum) without gold contacts. Figure 3 shows the structure of these two conformations. The starting point of the energy minimization was a linear-like structure. The local minimum found is the extended structure, while continued relaxation within a relatively flat potential energy landscape leads to the compressed structure. The difference in total energy between the extended and compressed structures was ~ 0.05 – 0.06 eV. This is a small energy difference for a molecule containing 50 atoms and it makes both structures accessible in solution at room temperature were it not for the hindrance and external force provided by the molecular contacts (not included in the free relaxation of the molecule). The major difference between the extended and compressed structures is the distance between phosphate and cytosine, which is reduced in the compressed structure. For example, the closest distance between a hydrogen in phosphate and a carbon atom in cytosine decreases from 4.8 Å in the extended structure to 2.7 Å in the compressed structure. The bonds between guanidinium and the phosphate have negligible changes².

Finally, the geometry of the complete structure, L-Gd-P-dR-C-G, was formed by sliding L-Gd-P-dR-C and G together with C and G in the same plane at the proper duplex hydrogen bond lengths.

There are two hydrogen bonds connecting guanidinium to phosphate and three hydrogen bonds between G and C

² The hydrogen bond donor atoms in guanidinium and the hydrogen bond acceptor atoms in phosphate are aligned more linearly as the relaxation steps go on. The angle between the plane of N, C, N (hydrogen bond donor nitrogen atoms and center carbon atom in guanidinium) and O, P, O (hydrogen bond acceptor oxygen and phosphorus in phosphate) was 25° and 11° for the extended and the compressed structure, respectively.

(figure 2). A comparison is made of the energetics and hydrogen bond lengths of these two cases for gas phase (vacuum) molecules. Of course, these results are not to be interpreted as binding energies in solution (e.g. water) which are much different. Figure 4 shows the binding curve for the Gd-P complex and for the G-C complementary base-pair. The binding curves are obtained by rigidly bringing together the two portions. The G-C base-pair has a hydrogen bond length near 2.0 Å and a binding energy of 26 kcal mol⁻¹, as is found in theoretical gas phase analysis by others [37, 39]. The bond length plotted is that of the central ($H_G \cdots N_C$) hydrogen bond. The hydrogen bond length obtained from the binding curve is slightly higher than the value obtained from the geometry optimization of the G-C base-pair (1.89 Å). The Gd-P cohesive energy in vacuum is 86 kcal mol⁻¹. Thus the guanidinium/phosphate bond is far stronger than the G-C bond. Clearly, such a high binding energy is not due just to the hydrogen bonding, but includes a large charge pairing component. A separation of $+e$ and $-e$ of about 4 Å gives such a binding energy.

It is impossible to know the geometry of the molecules in an STM experiment of a recognition circuit [27, 28]. We believe that the extended geometry is the predominant structure except perhaps when the tip is driven into the surface. We use the extended structure for further calculations and discuss the effect of the compressed geometry on the conduction properties of L-Gd-P-dR-C-G in the final section.

3. Intrinsic electron transport properties

The L-Gd-P-dR-C-G molecule is a long and complex molecule. We will be modeling the electron transport through it using ballistic transport theory via tunneling. In this theory there are no distortions of the molecule or leads before, after or during the transport process and no explicit electron-phonon couplings.

We first evaluate the complex band structure (CBS) of artificially repeated molecular units to understand the overall conductance properties of electron transport through molecular junctions. The complex band structure determines how rapidly the tunneling probability decays (by determining the exponential decay parameter β) along a repeating chain of molecules [43, 44]. Effectively the results of the CBS demonstrate the height of the electron's intrinsic energy barrier to go through the molecular unit. Going even further, the exponential decay parameter β provides a rough estimate for the conductance, g , as $g \approx g_0 e^{-\beta L}$, where g_0 is the quantum of conductance ($g_0 = 77 \mu\text{S}$) and L is the length of the molecule. This is a first approximation and is no substitute for a full scattering theory I - V calculation (given in section 5). The CBS is helpful in identifying how sensitive tunneling is to energy and which molecules are expected to tunnel well and which do not. Since the technique focuses on the decay parameter β , which is in the exponent for the tunneling probability, one is able to simply understand the origins of conductances that vary from 10^{-15} to 10^{-6} S (fS to μS). A plane-wave basis method [44, 45] is used for the electronic structure calculation in the LDA-DFT

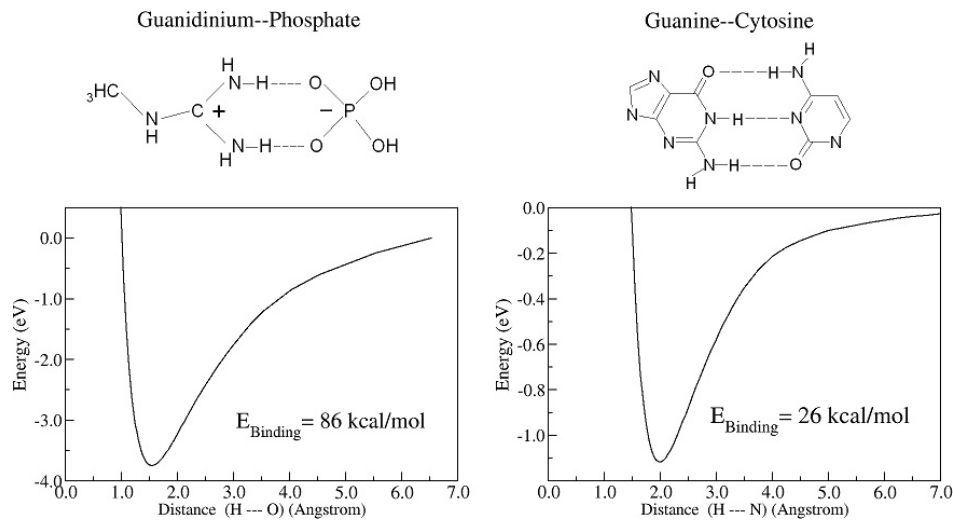


Figure 4. The total energy versus hydrogen bond length (distance between the hydrogen bond donor and acceptor atoms) in guanidinium–phosphate (Gd–P) and guanine–cytosine (G–C) molecules. The Gd–P complex has much higher binding energy (86 kcal mol^{-1}) than the G–C base-pair (26 kcal mol^{-1}) and shorter hydrogen bond length ($1.54\text{--}1.55 \text{ \AA}$) than the G–C base-pair due to the electrostatic interactions.

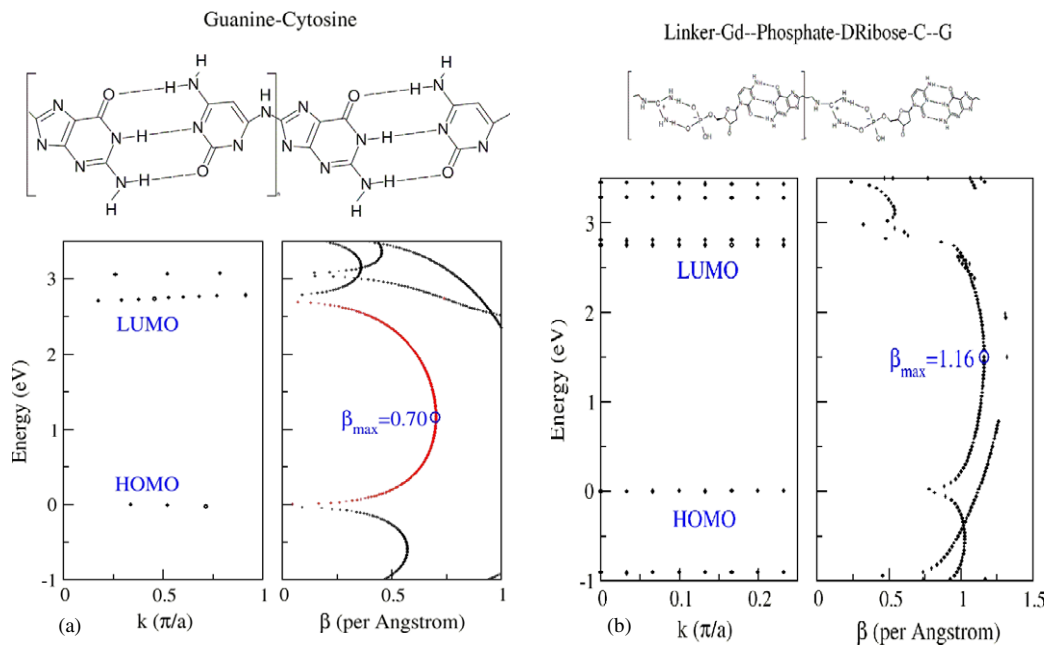


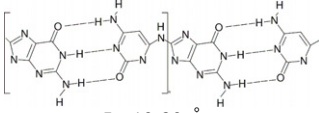
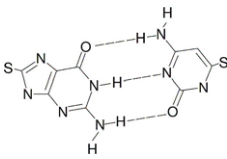
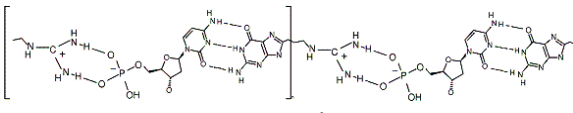
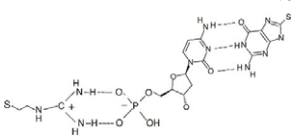
Figure 5. Complex band structure of (a) a guanine–cytosine base-pair and (b) a complete molecular circuit (L–Gd–P–dR–C–G, extended conformation). The molecule is altered using an --NH-- linker for G–C base-pair and $\text{--CH}_2\text{--}$ linker for L–Gd–P–dR–C–G to have periodicity. Two unit cells are shown for each molecule. The left figure shows the conventional band structure for real k and the complex band structure for $\beta = 2 \times \text{Im}[k]$ is shown on the right side. For the guanine–cytosine base-pair (a), the bandgap between the HOMO and the LUMO is 2.70 eV , and the most penetrating state in the bandgap region is depicted as a semi-elliptical curve in red. The maximum value of β for this state is 0.70 \AA^{-1} . The L–Gd–P–dR–C–G complex (b) has a single semi-elliptical curve in the bandgap region and β_{max} is 1.16 \AA^{-1} . The bandgap is 2.75 eV .

pseudopotential approximations. (Other density functionals, e.g. GGA, give similar results for hydrogen bond tunneling and are not reported here [46].) The molecular assemblies that we consider are (a) the hydrogen-bonded guanine–cytosine (G–C) base-pair and (b) the complete L–Gd–P–dR–C–G molecule (extended conformation). The CBS results are shown for these two cases in figures 5(a) and (b), respectively.

We begin by discussing the transport properties through the G–C base-pair (figure 5(a)). A periodic structure is created

by removing a hydrogen atom on each base and connecting the two bases by a linker, to form $\dots(\text{L–G–C})\text{--}(\text{L–G–C})\dots$. Here ‘L’ is the linker molecule, which is NH. The left portion of figure 5(a) shows the ‘real’ band structure (propagating waves, $\psi \sim e^{ikx}$) and the right side shows the imaginary part of the complex band structure (tunneling states, $\psi \sim e^{-\beta x/2}$). The top of the HOMO band is defined to be at 0.0 eV , and the LUMO band starts at about 2.7 eV , which is the HOMO–LUMO gap in this level of theory. Within this gap the k vector

Table 1. A summary of the conductance values and decay parameters for both the G–C base-pair and the L–Gd–P–dR–C–G (extended conformation) molecules. The rightmost column shows the conductance from scattering theory I – V calculations where the molecule is attached to a pair of Au (111) surfaces. Also given is an estimate of the conductance determined from the complex band structure. Two estimates are given—one uses the maximum β decay parameter (giving the lower of the conductance g) and the other uses an estimate of β at the energy of the metallic Fermi level. The length of a periodic unit L is also shown.

	β_{\max} $g_{\min} = g_0$ $\times \exp(-\beta_{\max}L)$	$\beta(E_F)$ $g_{E_F} = g_0$ $\times \exp(-\beta(E_F)L)$	g_{IV} (I – V scattering theory)
G–C	 $L=12.80 \text{ \AA}$ 0.70 \AA^{-1} 9.82 nS	0.35 \AA^{-1} 870 nS	
			84 nS
L–Gd–P–dR–C–G	 $L=24.07 \text{ \AA}$ 1.16 \AA^{-1} 0.056 fS	1.03 \AA^{-1} 1.32 fS	
			8.62 fS

is complex and we define β as $2 \times \text{Im}(k)$. For a molecule of length L , the tunneling probability is proportional to $|\psi(L)|^2$ ($\sim e^{-2\text{Im}(k)L} = e^{-\beta L}$). The decay rate β is zero near the HOMO and LUMO levels and has a maximum decay constant of 0.70 \AA^{-1} in the mid-gap region. Chains with π -bonding have ~ 0.3 – 0.5 \AA^{-1} decay rates [47] while σ -bonded chains have $\sim 0.8 \text{ \AA}^{-1}$ decay rates [48, 49]. Thus the overall decay rate through a hydrogen-bonded G–C base-pair is slightly less than a σ -bonded system. This result represents a quantum mechanical average of a small decay through the π -bonded ring portions of the bases and the hydrogen bonding region connects them.

The CBS result for the G–C base-pair allows a simple estimate to be made of the conductance of this base-pair, which is shown in table 1. We use $g \approx g_0 e^{-\beta L}$, where L is the length of the unit cell. An estimate of the smallest conductance (largest β) for this molecule is 9.82 nS. A more correct value to use is $\beta(E_F)$, the β value at the metallic Fermi level. The alignment of the metal's Fermi level with the molecule's levels will be determined in the next section, but table 1 lists the final result.

The CBS of the complete molecule L–Gd–P–dR–C–G, which is of primary interest in this paper, is shown in figure 5(b). The HOMO–LUMO gap is 2.75 eV. A plot of β versus energy in the HOMO–LUMO gap is largely a single semi-elliptical curve coupling the HOMO and LUMO. The maximum β value is 1.16 \AA^{-1} . The large β value and the long length of L–Gd–P–dR–C–G of 24.07 \AA produces an estimate of its conductance of 0.056 fS (see table 1). This is an extremely small value of conductance. At a bias of 0.1 V,

35 electrons s^{-1} pass through. Stated differently, only one electron passes every 29 ms. Currents this low will create an additional challenge for the development of a device based on the technology described here.

4. Fermi level alignment

A key parameter that determines electron transport through molecules is alignment of the metal contact's Fermi level with the energy levels of the molecule (e.g. HOMO and LUMO levels). In the previous section, we estimated the tunnel conductance of a molecule using the maximum value of the decay parameter (β_{\max}) in the HOMO–LUMO gap region. Since the electrons that contribute to the conductance are those whose energy levels are near the Fermi energy of the metal, a more appropriate estimate for conductance is made by using $\beta(E_F)$ at the Fermi energy, $g = g_0 e^{-\beta(E_F)L}$. If the Fermi level is aligned near the HOMO or the LUMO, the conductance will increase as the decay parameter $\beta(E_F)$ decreases significantly from β_{\max} .

We determine the Fermi level alignment for the G–C base-pair and for a complete L–Gd–P–dR–C–G (extended conformation) circuit by computing the projected density of states of the molecule connected on each side by thiols to gold surfaces. The electronic structure calculations for a metal–molecule–metal system were performed using a plane wave basis method³ within the DFT–LDA pseudopotential scheme. The ‘metal–molecule–metal’ system is constructed by

³ The Vienna *ab initio* Simulation Program (VASP) was developed at the Institut für Theoretische Physik of the Technische Universität Wien [50].

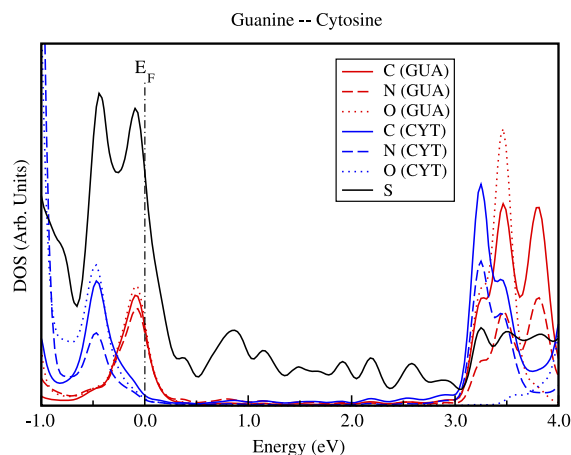


Figure 6. The projected density of states onto carbon, nitrogen, oxygen and sulfur atoms for the di-thiolated G–C base-pair connected across a pair of gold contacts. The figure shows the averaged projection of each atom on the cytosine and guanine. The Fermi energy is defined to be zero energy. The HOMO and the LUMO are localized onto guanine and cytosine, respectively. The Fermi level lies very close to the HOMO (on guanine).

inserting the molecule between (111) gold slabs after terminal hydrogen atoms of the molecule are replaced by sulfur atoms to make chemical bonds with the gold electrodes. The systems are periodic (supercells) both in the molecular axis direction and in the perpendicular directions. The gold layers were 3×3 in size in both cases, and six layers thick for the G–C base-pair and four layers thick for the L–Gd–P–dR–C–G molecule. For the G–C base-pair, the sulfur atom is positioned directly above the gold atom to form a bond to a single gold atom (on-top site contact). The distance between the on-top-Au and sulfur was 2.42 \AA . The on-top site contact was used to prevent atoms (specifically H) on the base-pair from being too close to the gold plane. The (Au–S–C) angles were 120° and 140° for the left and right contacts, respectively. For L–Gd–P–dR–C–G, a sulfur atom is placed 1.95 \AA above a ‘hollow’ position on Au (111) surface equidistant (2.56 \AA) from three gold atoms (hollow site contact). The molecule is inclined rather than being perpendicular to the two Au planes; the (hollow site–S–C) angles were 112° and 135° for the two sides.

Figure 6 shows the projected density of states onto C, N, O atoms on guanine and cytosine, and the sulfur atoms for the G–C base-pair. The HOMO and LUMO for the G–C base-pair are localized onto the orbitals on guanine and cytosine, respectively. The energy levels are shifted so that the Fermi level becomes zero energy. The bandgap is 3.34 eV , which is higher than the bandgap obtained in the complex band structure calculation using the periodic G–C base-pair (2.70 eV). The difference in the bandgap originates from the sulfur and metal–molecule interaction. Figure 6 clearly shows that the Fermi level is aligned very close to the HOMO (about 0.1 eV above the HOMO). This indicates that the conductance of a G–C base-pair will be much higher than what we obtained using the β_{\max} in the previous section. Using the relative position of the Fermi level in the bandgap and the β curve in figure 5(a), we obtain an approximate $\beta(E_F)$ of 0.35 \AA^{-1} . The conductance

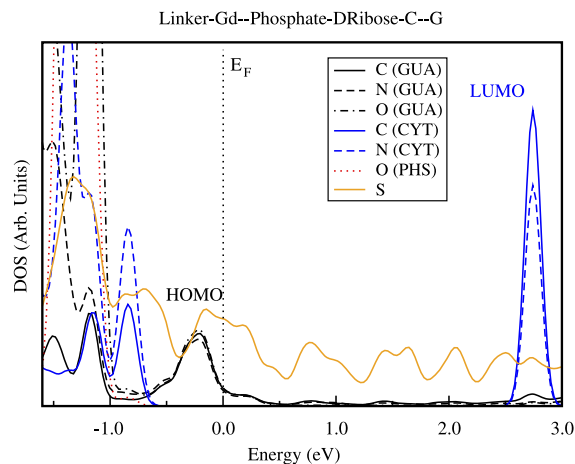


Figure 7. The projected density of states for the di-thiol molecule L–Gd–P–dR–C–G (extended conformation) connected to gold contacts. The HOMO is mostly dominated by orbitals in guanine, and the LUMO is dominated by orbitals in cytosine as in the G–C base-pair. The oxygen orbital states from the phosphate appear much below the HOMO. The orbitals from other components of the molecule contribute to the state outside the bandgap region (not shown in the figure). The Fermi level is aligned near the HOMO (on guanine).

using $\beta(E_F)$ is 870 nS (0.87 \mu S), which is higher than the conductance estimate using $\beta_{\max} = 0.70 \text{ \AA}^{-1}$ by a factor of 10^2 . This clearly shows the effect of the Fermi level alignment on the conductance. We note in passing that the contact geometry also affects the Fermi level alignment. The Fermi level tends to lie a little closer to the HOMO in the on-top site geometry (used here) compared to a hollow site contact.

The projected density of states for L–Gd–P–dR–C–G (extended conformer) is shown in figure 7. The Fermi level is aligned near the HOMO and the bandgap is 2.9 eV . The decay parameter at the Fermi level estimated from figure 5(b) is 1.03 \AA^{-1} and the corresponding estimate of the conductance for the molecule is 1.32 fS . The estimated conductance is higher than the value obtained using β_{\max} (0.056 fS). Figure 7 also shows that the HOMO is dominated by the orbitals in guanine and the LUMO by orbitals in cytosine as in the G–C base-pair. The orbitals from the linker, guanidium and deoxyribose contributed to the states outside the HOMO–LUMO bandgap region (not shown in the figure). The sulfur state is broadly spread due to the interaction with the gold states.

5. Transport properties from I – V calculations

We will assume that the basic transport mechanism from the metal contacts through the molecule is ballistic tunneling. This means that the electron traverses from one electrode to the other through the molecules in a single step, that there is no energy loss (elastic) due to vibrations or other losses, and that the electron does not hop onto the molecule and hop off in a two-(or more) step process. Assuming ballistic transport, methods have been developed by several groups [51–55]. A common features is the Green’s function

that describes quantum mechanical propagation through the molecule; features that differ involve how the infinite contacts are handled, the method used to determine the electronic structure and basis sets, and how the electric field within the molecule is handled. We use a scattering approach based on Fermi's golden rule. Full details are given in [49]. The method assumes a DFT approach where a local orbital basis set with overlap matrix S between orbitals is used. The DFT approximation is generally used in electron transport simulations although this approximation has limitations—self-interaction corrections are neglected [56] and there are errors caused by the use of an equilibrium exchange–correlation functional [57, 58]. The potential drops across the contacts are assumed to be symmetric (gold contacts at each end). The method leads to a transmission function through a barrier as described by the Landauer approach.

We begin by imagining the left and right contacts are uncoupled to the molecule and that the two contacts are far enough away that there is no coupling between them. In this case an electron can reside in either the left contact, the right contact or the molecule, but cannot yet transfer from one region to another. The Green's function propagator for these uncoupled systems is G_0 . The interaction V couples the left contact and the molecule as well as the molecule with the right contact. The Green's function propagator for the coupled system is $G = G_0 + G_0VG_0 + \dots = G_0 + G_0VG$. This propagator has the information as to how an electron from the left contact (say) can tunnel through the molecule and end up on the right contact. The Fermi golden rule transition rate from left to right contact is

$$\Gamma_{L \rightarrow R} = \frac{2\pi}{\hbar} |t_{LR}|^2 \delta(E_L - E_R) \quad (1)$$

where t is the scattering t -matrix [59], $t = V + VGV$. This leads to an expression for the current in terms of the transmission function $T(E)$ and applied voltage bias:

$$I = \frac{2e}{h} \int_{\mu_R}^{\mu_L} T(E) dE. \quad (2)$$

The applied voltage bias is contained in the limits of integration, $\mu_R = E_F - eV/2$ and $\mu_L = E_F + eV/2$ (symmetric voltage drop) where E_F is the equilibrium Fermi level alignment. The transmission function $T(E)$ is given by

$$T(E) = \text{tr}(\Gamma_L G_M \Gamma_R G_M^\dagger), \quad (3)$$

the spectral density of states is

$$\Gamma_L = i(\Sigma_L - \Sigma_L^\dagger) \quad (4)$$

where Σ is the 'self-energy'

$$\Sigma_L(E) = (ES_{LM} - H_{LM})^\dagger G_L^0(E) (ES_{LM} - H_{LM}). \quad (5)$$

Similar expressions exist for the right (R) contact. Here G_M is a Green's function for the molecule, and S_{LM} and H_{LM} are the overlap and Hamiltonian coupling matrix elements between the left contact and the molecule. Note that the propagator of an

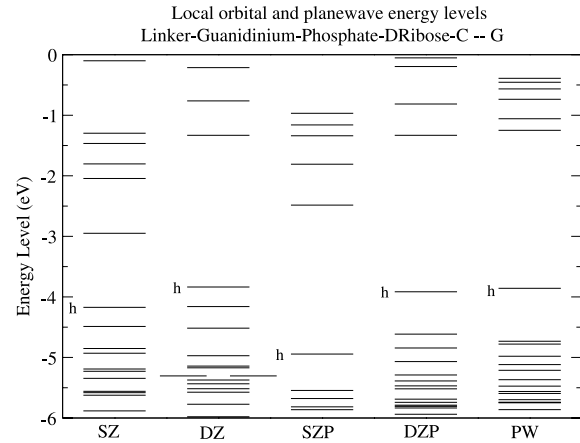


Figure 8. The energy levels for L–Gd–P–dR–C–G (extended conformation) obtained using different local orbital basis sets compared to those using plane waves. The HOMO is indicated as ‘h’. The energy levels obtained using the local orbital DZP basis set show the best agreement with the plane wave results.

uncoupled system is

$$G_L^0(E) = ((E + i\eta)S_L - H_L)^{-1} = (zS_L - H_L)^{-1} \quad (6)$$

where $z \equiv E + i\eta$ and $\eta \rightarrow 0+$. Once the transmission function is determined, the current is computed simply as

$$I = \frac{g_0}{e} \int_{E_F - eV/2}^{E_F + eV/2} T(E) dE \quad (7)$$

where $g_0 = 2e^2/h$ is the quantum of conductance and is numerically $77 \mu\text{S}$. Over the small biases that we apply (0.1 V), we find that the current is linear in all cases. Small biases are used since in the sequencing schemes described here, no complications due to electrochemistry are desired. For small voltages when the system responds linearly, the current is $I = g_0 T(E_F)V$ and the conductance is simply

$$g = g_0 T(E_F). \quad (8)$$

The molecule, with a sulfur at each end, is placed between two gold layers. Periodic boundary conditions are used and the central Au layer is considered ‘bulk’ gold. These ‘bulk’ layers are then mathematically repeated using a recursion method for the Green's function [49]. This exactly reproduces the result of an infinitely thick Au contact layer. A local orbital basis set is used [60], which makes the manipulation of the Green's function natural.

The accuracy of the calculation is only as good as the electronic structure Hamiltonian. To check the adequacy of the local orbital basis and its description of the central L–Gd–P–dR–C–G molecule, we compare the frontier energy levels against those of a plane wave [45] calculation. Figure 8 shows the energy spectrum of the near HOMO–LUMO region of L–Gd–P–dR–C–G using local orbitals with various basis sets [60]. The basis sets used are (for non-hydrogen atoms) sp^3 (single zeta, SZ), $sp^3s^*p^3*$ (double zeta, DZ), sp^3d^5 (single zeta with polarization, SZP) and $sp^3s^*p^3*d^5$ (double zeta with

polarization, DZP). The plane wave (PW) basis is essentially a complete set. The DZP basis set (13 orbitals/non-H atom) gives very good agreement with the plane wave results, and we choose this basis set. For the gold atoms of the metallic slabs we use SZP.

We have performed I - V calculations using Green's function scattering theory of two molecules wired between gold contacts. The two molecular 'wires' are (a) the guanine-cytosine (G-C) base-pair and (b) L-Gd-P-dR-C-G (extended conformer). The hollow site contact was used to connect to Au for L-Gd-P-dR-C-G while the on-top site contact was used for G-C base-pair because of the proximity of the molecule to the gold surface without a linker. Periodic boundary conditions are used for the I - V calculation, with the periodicity being 3×3 lateral Au surface atoms for G-C and 4×4 for L-Gd-P-dR-C-G. In both cases there were six Au layers. The larger lateral size of the Au surface was used for L-Gd-P-dR-C-G to prevent the interaction between the atoms in neighboring lateral supercells.

Figure 9 shows the transmission function ($\ln T(E)$) for the two molecules. (A third molecule, the L-Gd-P-dR-C-G molecule in a compressed geometry, will be described in section 6.) The transmission function for the G-C base-pair shows that the Fermi level is aligned near the HOMO (0.2 eV above), a result very similar to that found using plane waves. The bandgap is about 3.2 eV, which is similar to the bandgap obtained using the plane wave basis set (3.3 eV) in the previous section. The transmission function of the main molecule of interest, L-Gd-P-dR-C-G, shows that the bandgap is about 2.8 eV, a result similar to the bandgap from complex band structure (2.75 eV) and from the projected density of states using the plane wave basis set (2.9 eV). The Fermi level is aligned very close to the HOMO, as found in the previous section.

Current-voltage (I - V) curves resulting from Green's function scattering theory produced linear I - V curves over the small⁴ (± 0.10 V) applied voltage bias. The conductances of these molecular wires are shown in table 1, along with the estimates based on the complex band structure for comparison. The most significant result is the extremely small conductance of the extended L-Gd-P-dR-C-G; it is just 8.62 fS. The complex band structure estimates support the conclusion of a very small conductance—the complex band structure, using β_{E_f} , is estimated to give a conductance of 1.32 fS. A conductance in the fS range means that it is difficult to measure in a typical laboratory set-up. Although one expects a small conductance for a molecule of this great length, the extremely small conductance (in the fS range) is somewhat surprising. The base-pair conductance, which involves hydrogen bonding, has a β value in the range of 0.35 – 0.70 \AA^{-1} (table 1). Using these values instead of the β for the L-Gd-P-dR-C-G molecule (extended conformer) gives a conductance of 4 pS–17 nS; thus a minimum of one thousand times larger. Thus we conclude that the conduction through the base-pair region itself is relatively efficient—rather, the bottleneck is in the L-Gd-P-dR portion of the molecule. This is at least qualitatively similar

⁴ Experiments of DNA in solution for sequencing will generally apply low voltages to avoid reactions produced by electrochemistry.

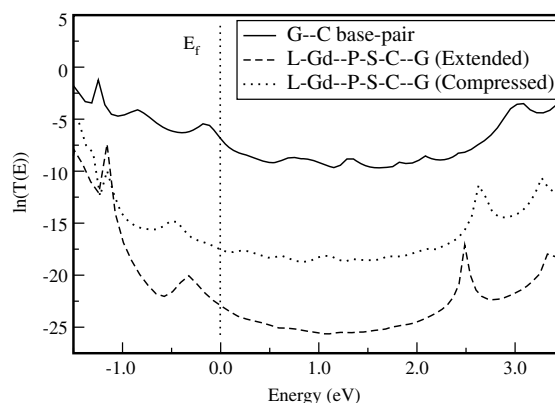


Figure 9. The natural log of the transmission function for G-C and L-Gd-P-dR-C-G in the extended and the compressed geometries. The gold Fermi level is shown as a dotted line and for both G-C and L-Gd-P-dR-C-G it is aligned very close to the HOMO. The transmission probability increases significantly for the compressed geometry of L-Gd-P-dR-C-G.

to the results of Zikic *et al* [24]. These authors computed the conductance of a single neutral nucleotide between electrodes separated by 15 \AA and obtained 0.1 – 1.0 pS. Their geometry did not have covalent (e.g. sulfur) contacts.

In experiments [28] to measure the conductance of L-Gd-P-dR-C-G, a variety of geometries for the molecule are to be expected. The STM tip is first pushed into the surface, and then withdrawn. Geometries such as the compressed conformer may play a role in at least portions of the experiment. The L-Gd-P-dR-C-G conformer may be compressed upon approach and extended upon withdrawal. The issue of structure/geometry will now be explored.

6. Geometry dependence of conductance

Scanning tunneling microscopy (STM) experiments [27, 28] on circuits like L-Gd-P-dR-C-G are performed by driving the STM tip (functionalized with G) onto a Au surface containing L-Gd-P-dR-C, with the sulfur attached to gold. The tip is brought to the surface until a set-point current is achieved. The location of the tip relative to the surface is not known, nor is the geometrical configuration of the L-Gd-P-dR-C-G molecule. Possible effects of molecular geometry on electron transfer properties will be probed in this section; specifically the conductance of the compressed conformer will be studied.

We evaluate the tunneling characteristics of L-Gd-P-dR-C-G in its compressed structure to determine conductance-structure relationships. Figure 9 shows the natural log of the transmission functions for the compressed structure and compares it to the extended structure. The energy levels in the figure were adjusted so that the Fermi level is zero energy for each system individually. The change in transmission probability upon compression is quite large, approximately e^{+5} at the Fermi level, which increases the conductance substantially. The computed I - V curve (not shown) was linear within the range of ± 0.1 V. The computed conductance of the compressed structure is 1.89 pS, which is more than 200 times

greater than for the extended structure (8.62 fS). The reason for such a huge increase in conductance for the compressed structure is because of the reduced distance between the phosphate and cytosine. This makes it possible for the electrons to transport directly from the phosphate to cytosine without transferring through the deoxyribose. This shows that deoxyribose greatly reduces the conductance. These results have important consequences for experiments and devices that measure transverse conductance by hydrogen bonding. Fairly minor changes in the structure make a large difference in the current and conductance. In fact, if the structure changes from a slightly compressed form to an extended form, the measured signal is likely to change from being observable to being non-observable.

7. Conclusions

In this paper we investigated the electron tunneling properties of deoxycytidine-monophosphate chemically bonded to metal electrodes through hydrogen bonds. The molecule studied here consists of a linker ((CH₂)₂), guanidinium, phosphate, deoxyribose, cytosine and guanine (L–Gd–P–dR–C–G), where electrode-tethered guanidinium has both an electrostatic and hydrogen bond interaction with the phosphate and a guanine nucleobase tethered to the metal electrode forms hydrogen bonds with cytosine. A complete molecular circuit for electron tunneling is formed when the molecule is sandwiched between gold electrodes. This work is motivated by ongoing experiments for this circuit using an STM probe to provide a new DNA sequencing scheme by recognition [27, 28]. Two different methods, complex band structure (using periodicity) and I – V calculation using the Green's function scattering methods, were used for analysis. The decay constant of the tunnel current obtained from the complex band structure for the base-pair alone (G–C) as well as for a complete molecular circuit (L–Gd–P–dR–C–G, extended conformer) shows that a G–C base-pair is as conductive as a σ -bonded molecule ($\beta_{\max} = 0.70 \text{ \AA}^{-1}$) while the complete molecular circuit is far less conductive with a high decay constant of ($\beta_{\max} = 1.16 \text{ \AA}^{-1}$). Tunneling is energy-dependent and, for the case of Au contacts, the Fermi level alignment is near the HOMO. I – V calculation were performed for a G–C base-pair and for a complete L–Gd–P–dR–C–G circuit, and the result is consistent with conductance estimates obtained from the complex band structure. The results are summarized in table 1.

We conclude that the calculated conductance of the molecule L–Gd–P–dR–C–G (in an extended structure) is very low (on the order of fS) while the hydrogen bonded base-pair gives a moderate conductance (on the order of tens of nS). The addition of the guanidinium–phosphate unit and deoxyribose to the base-pair decreases the conductance significantly. Conductances in the fS range present a challenge for measurement in an experiment. The conductance from the I – V calculation obtained for two geometries (an extended geometry and a compressed geometry) shows a significant increase (from fS to pS) as the structure of the molecule becomes compressed rather than extended. Some form of compression or folding is likely to occur in STM experiments

where an STM tip is brought to the surface to achieve a given set-point current. The present work studied the tunneling characteristics of a complete molecular circuit involving the G–C base-pair, and forms the first step in understanding the 16 base-pair combinations necessary for base-pair recognition of ssDNA.

Acknowledgments

We are pleased to acknowledge a close and fruitful collaboration with the experimental sequencing team including J He, L Lin, Q Spadola, Z Xi, Q Fu, S Chang, S Huang, A Kibel, D Cao, F Liang, P Pang, H Liu, P Pelletier, K Reinhart, P Zhang and S Lindsay. We acknowledge support from the NIH (NIH R21 HG004378-01).

References

- [1] Lander E S *et al* 2001 International human genome sequencing consortium *Nature* **409** 860–921
- [2] Venter J C *et al* 2001 *Science* **291** 1304–51
- [3] Levy S *et al* 2007 *PLoS Biol.* **5** 2113–44 e254
- [4] Sanger F and Coulson A R 1975 *J. Mol. Biol.* **25** 441–8
Sanger F, Nicklen S and Coulson A R 1977 *Proc. Natl Acad. Sci. USA* **74** 5463–7
- [5] Aksyonov S A, Bittner M, Bloom L B, Reha-Krantz L J, Gould I R, Hayes M A, Kiernan U A, Niederkofler E E, Pizziconi V, Rivera R S, Williams D J B and Williams P 2006 *Anal. Biochem.* **348** 127–38
- [6] Akeson M, Branton D, Kasianowicz J K, Brandin E and Deamer D W 1999 *Biophys. J.* **77** 3227–33
- [7] Deamer D W and Akeson M 2000 *Trends Biotechnol.* **18** 147–51
- [8] Deamer D W and Branton D 2002 *Acc. Chem. Res.* **35** 817–25
- [9] Kasianowicz J K, Brandin E, Branton D and Deamer D W 1996 *Proc. Natl Acad. Sci. USA* **93** 13770–3
- [10] Gierhart B C, Howitt D G, Chen S J, Zhu Z, Kotecki D E, Smith R L and Collins S D 2008 *Sensors Actuators B* **132** 593–600
- [11] Fink H-W and Schönenberger C 1999 *Nature* **398** 407–10
- [12] Drummond T G, Hill M G and Barton J K 2003 *Nat. Biotechnol.* **21** 1192–9
- [13] Zvolak M and Di Ventra M 2008 *Rev. Mod. Phys.* **80** 141–65
- [14] Lewis F D, Liu X, Liu J, Miller S E, Hayes R T and Wasielewski M R 2000 *Nature* **406** 51–3
- [15] Berlin Y A, Burin A L and Ratner M A 2001 *J. Am. Chem. Soc.* **123** 260–8
- [16] Bhalla V, Bajpai R P and Bharadwaj L M 2003 *EMBO Rep.* **4** 442–5
- [17] Porath D, Bezryadin A, de Vries S and Dekker C 2000 *Nature* **403** 635–7
- [18] Xu B, Zhang P, Li X and Tao N 2004 *Nano Lett.* **4** 1105–8
- [19] Hihath J, Xu B, Zhang P and Tao N 2005 *Proc. Natl Acad. Sci. USA* **102** 16979–83
- [20] Endres R G, Cox D L and Singh R R P 2004 *Rev. Mod. Phys.* **76** 195–214
- [21] Yanov I and Leszczynski J 2004 *Int. J. Quantum Chem.* **96** 436–42
- [22] Mallajosyula S S and Pati S K 2007 *J. Phys. Chem. B* **111** 11614–8
- [23] Zvolak M and Di Ventra M 2005 *Nano Lett.* **5** 421–4
- [24] Zikic R, Krstić P S, Zhang X-G, Fuentes-Cabrera M, Wells J and Zhao X 2006 *Phys. Rev. E* **74** 011919
Zhang X-G, Krstić P S, Zikic R, Wells J C and Fuentes-Cabrera M 2006 *Biophys. J.* **91** L04–6

- [25] Ohshiro T and Umezawa Y 2006 *Proc. Natl Acad. Sci.* **103** 10–4
- [26] Cui X D, Primak A, Zarate X, Tomfohr J K, Sankey O F, Moore A L, Moore T A, Gust D, Harris G and Lindsay S M 2001 *Science* **294** 571–4
- [27] He J, Lin L, Zhang P and Lindsay S M 2007 *Nano Lett.* **7** 3854–8
- [28] He J, Lin L, Zhang P, Spadola Q, Xi Z, Fu Q and Lindsay S 2008 *Nano Lett.* **8** 2530
He J, Lin L, Liu H, Zhang P, Lee M, Sankey O F and Lindsay S M 2008 *Nano Lett.* submitted
- [29] Barnett R N, Cleveland C L, Joy A, Landman U and Schuster G B 2001 *Science* **19** 567–71
- [30] Lin J, Balabin I A and Beratan D N 2005 *Science* **310** 1311–3
- [31] Schmidt M W, Baldrige K K, Boatz J A, Elbert S T, Gordon M S, Jensen J H, Koseki S, Matsunaga N, Nguyen K A, Su S J, Windus T L, Dupuis M and Montgomery J A 1993 *J. Comput. Chem.* **14** 1347–63
- [32] Brameld K, Dasgupta S and Goddard W A III 1997 *J. Phys. Chem. B* **101** 4851–9
- [33] Guerra C F, Bickelhaupt F M, Snijders J G and Baerends E J 1999 *Chem. Eur. J.* **5** 3581–94
- [34] Bertran J, Oliva A, Rodríguez-Santiago L and Sodupe M 1998 *J. Am. Chem. Soc.* **120** 8159–67
- [35] Šponer J, Jurečka P and Hobza P 2004 *J. Am. Chem. Soc.* **126** 10142–51
- [36] Šponer J, Leszczynski J and Hobza P 1996 *J. Phys. Chem.* **100** 1965–74
- [37] Mo Y 2006 *J. Mol. Model.* **12** 665–72
- [38] van der Wijst T, Guerra C F, Swart M and Bickelhaupt F M 2006 *Chem. Phys. Lett.* **426** 415–21
- [39] Fellers R S, Barsky D, Gygi F and Colvin M 1999 *Chem. Phys. Lett.* **312** 548–55
- [40] Tsuzuki S and Lüthi H P 2001 *J. Chem. Phys.* **114** 3949–57
- [41] Becke A D 1993 *J. Chem. Phys.* **98** 5648–52
- [42] Saenger W *Principles of Nucleic Acid Structure* (Berlin: Springer) pp 123–4 and references therein
- [43] Tomfohr J K and Sankey O F 2002 *Phys. Rev. B* **65** 245105
- [44] Picaud F, Smogunov A, Dal Corso A and Tosatti E 2003 *J. Phys.: Condens. Matter* **15** 3731–40
- [45] Baroni S, Corso A D, de Gironcoli S and Giannozzi P <http://www.pwscf.org>
- [46] Lee M H and Sankey O F 2008 in preparation
- [47] Tomfohr J K and Sankey O F 2002 *Phys. Status Solidi b* **233** 59–69
- [48] Xu B and Tao N J 2003 *Science* **301** 1221–3
- [49] Tomfohr J and Sankey O F 2004 *J. Chem. Phys.* **120** 1542–54
- [50] Kresse G and Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15
Kresse G and Hafner J 1993 *Phys. Rev. B* **47** 558
Kresse G and Furthmüller J J 1996 *Phys. Rev. B* **54** 11169
- [51] Brandbyge M, Mozos J L, Ordejon P, Taylor J and Stokbro K 2002 *Phys. Rev. B* **65** 165401
- [52] Damle P S, Ghosh A W and Datta S 2001 *Phys. Rev. B* **64** 201403
- [53] Taylor J, Guo H and Wang J 2001 *Phys. Rev. B* **63** 245407
- [54] Mujica V, Kemp M and Ratner M A 1994 *J. Chem. Phys.* **101** 6849–55
- [55] Di Ventura M, Lang N D and Pantelides S T 2002 *Chem. Phys.* **281** 189–98
- [56] Toher C, Filippetti A, Sanvito S and Burke Kieron 2005 *Phys. Rev. Lett.* **95** 146402
- [57] Evers F, Weigend F and Koentopp M 2004 *Phys. Rev. B* **69** 235411
- [58] Koentopp M, Burke K and Evers F 2006 *Phys. Rev. B* **73** 121403
- [59] Lippmann B A and Schwinger J 1950 *Phys. Rev.* **79** 469–80
- [60] Ordejón P, Sánchez-Portal D, Artacho E, Soler J M and García A 2006 The SIESTA code <http://www.uam.es/departamentos/ciencias/fismateria/c/siesta/>
Ordejón P, Artacho E and Soler J M 1996 *Phys. Rev. B* **53** R10441–4
Sánchez-Portal D, Ordejón P, Artacho E and Soler J M 1997 *Int. J. Quantum Chem.* **65** 453–61