COMPUTATIONAL PROOF-OF-CONCEPT OF NEXT-GENERATION NUCLEOTIDE–TUNNELING-CURRENT-BASED NANOPORE SEQUENCING DEVICES

Debasis Sengupta, Jerry Jenkins*. Zbignew Sikorski and Shankar Sundaram

CFD Research Corporation, 215 Wynn Drive, Huntsville, AL 35805
Ph: (256)-726-4913, Fax: (256)-726-4806
Email: jwj@cfdrc.com

ABSTRACT

The present work is aimed at computing tunneling current of four constituent bases, namely adenine, thymine, guanine, and cytosine, using quantum chemistry and non-equilibrium Green's function theory. The calculations presented in this work, while under idealized conditions, clearly indicate that the use of tunneling current measurements to sequence DNA as it translocates through a nanopore is indeed possible. However, practical success is shown to depend critically on parameters such as biasing voltage and pore diameter. Future model refinements will include more realistic molecular descriptions that incorporate the effects of solvent and ions present within the system in order to evaluate the I-V characteristics of DNA bases under realistic conditions. Future work will require the development of new computational methods capable of rapidly estimating I-V characteristics of DNA for realistic systems sizes.

Keywords: nanopore, tunneling current, DNA sequencing, non-equilibrium Green's function theory.

1 INTRODUCTION

A complete knowledge of the DNA sequence of an individual can yield fundamental insight into disease diagnosis, treatment, and prevention. The current gold standard for determining a genetic sequence is the Sanger chain termination method; however, a typical mammalian sequence costs millions of dollars and requires many months for completion. Nanopores offer the unique ability to perform a label-free interrogation of individual DNA molecules, in real time. For this reason, nanopores have come under intense investigation (Heng, 2005) due to the promise of inexpensive ultra high-throughput sequencing of DNA (NHGRI goal set at <$1,000 for an entire mammalian sequence http://www.genome.gov).

In practice, a voltage bias applied to a nanopore induces a negatively charged, single-stranded DNA molecule to translocate through the pore (Figure 1a). The ions passing through the pore form an ionic current that is monitored throughout the experiment. As the DNA translocates through the pore, the flow of ions is hindered causing a sharp drop in the ionic current. However, determination of the DNA sequence solely from the ionic current measurements has proven to be problematic, primarily due to the speed of the DNA translocation and the low signal to noise ratio (Fologea, 2005). A recently proposed, novel method to sequence DNA incorporates nanowires at the pore entrance and monitors the change in the local electron tunneling current (Figure 1b). The primary advantage of tunneling current is that the electronics and instrumentation is on the same scale as the DNA (nanometers). However, this method presents many inherent theoretical and fabrication challenges. In order to design such a device, the interplay between molecular conformation, density of electronic states, and electron tunneling must be understood in order to determine the optimal experimental conditions for rapid DNA sequencing.

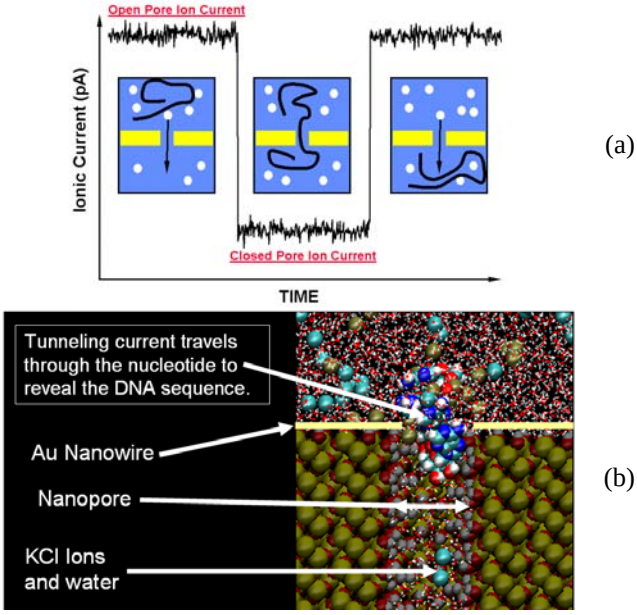

Figure 1: (a) Translocating DNA through a nanopore (b) Molecular structures used in the calculations

A single molecule can conduct electrons when a voltage is applied at the two terminals in contact with the molecule (Figure 2). Evaluation of the current vs. voltage (I-V) characteristics of the individual molecule is the most important aspect of the tunneling current approach. Biological molecules, such as DNA (Fink, 1999), are known to have highly nonlinear electrical conduction characteristics. In order to design an efficient nanopore sequencer based on tunneling current measurements, precise information about the DNA characteristics, such as molecular structure and the density of electronic states, is required. To our knowledge, no such extensive theoretical

analysis of the tunneling-current-based sequencing has been performed to assess feasibility of this detection system.

In response to this challenge, we have undertaken a series of computations to analyze the feasibility of sequencing DNA using a tunneling current approach. We have combined quantum chemistry with a nonequalibrium Green's function approach for electron transport to evaluate I-V characteristics of DNA. The primary outcome is a series of I-V curves for isolated nucleic acids, as well as an assessment of the dependence of the resultant I-V characteristics on pore diameter. Key innovations that mark this research include:

- Use of realistic nucleic acid geometries taken from previous molecular dynamics simulations of DNA translating through nanopores (Jenkins, 2005).
- Assessment of the effect of pore diameter on the resulting I-V characteristics.
- Optimization of the electrode placement and nanopore geometry to maximize the amplitude and variation in the nucleic acid conductances.

The final results of this computation are the I-V curves for the individual nucleotides along with an assessment of the dependence of the I-V curves on pore diameter.

2 COMPUTATIONAL METHODS

In the present work, a non-equilibrium Green's function (NEGF) approach was adopted, using the formalism developed by the Datta group (Zahid, 2003). A molecule is assumed to be located between two gold contacts, and a voltage bias is applied across the contacts (Figure 2). The NEGF technique requires the molecular orbital energies as input. In the present case, the energy levels are obtained using semi-empirical extended Huckel Theory (EHT).

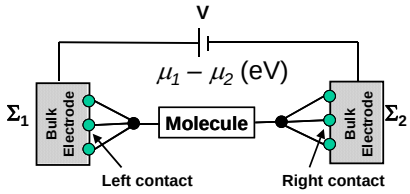

Figure 2: Schematic representation of molecular conduction. See text for definition of the terms.

The total Hamiltonian of the molecule with the contacts is given by Equation 1

$$H' = \begin{bmatrix} H_{11} & H_{1M} & H_{12} \\ H_{M1} & H_{MM} + \Sigma_1 + \Sigma_2 & H_{M2} \\ H_{21} & H_{2M} & H_{22} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{1M} & G_{12} \\ G_{M1} & G & G_{M2} \\ G_{21} & G_{2M} & G_{22} \end{bmatrix} \quad (1)$$

where the subscripts 1 and 2 represent the terms corresponding to left and right contacts; M represents the molecules; Σ represents the self energy term that quantifies the broadening of molecular energy levels, and G is the molecular Green's function. The self energy terms are calculated using the Green's function of the gold contacts. The gold contacts were represented using a 13-atom cluster. A detailed explanation of the contact Green's function can be found in Damle, 2002. The Green's function of the molecule can be expressed as Equation 2

$$G = (ES_M - H_{MM} - \Sigma_1 - \Sigma_2) \quad (2)$$

where, S_M is the molecular orbital overlap matrix, and E is the energy. The Σs are calculated using Equation 3

$$\Sigma_{1,2} = \tau_{1,2}^{\neq} g_{1,2} \tau_{1,2} \quad (3)$$

where, $\tau_{1,2} = ES_{1,2} - H_{1,2}$ and g's are the Green's function of the contacts. The current passing through the molecule can then be calculated using Equation 4 (Datta, 2000)

$$I = \int_{-\infty}^{\infty} \left[Trace(\Gamma_1 G \Gamma_2 G^{\neq}) \{ f(E,\mu_1) - f(E,\mu_2) \} \right] dE \quad (4)$$

where, μ's are the chemical potential of the contacts, *f's* are the fermi functions, and the broadening matrices, Γ is defined in Equation 5.

$$\Gamma_{1,2} = i\left[\Sigma_{1,2} - \Sigma_{1,2}^{\neq} \right] \quad (5)$$

It should be noted that the charging effect on the orbital energy levels has been taken into account within the NEGF formalism and is assumed to be proportional to the change in the electron density of the molecule due the applied bias. The electron density is calculated using the following expression:

$$\rho = \frac{1}{2\pi} \int_{-\infty}^{\infty} [f(E,\mu_1)(G\Gamma_1 G^{\neq}) + f(E,\mu_2)(G\Gamma_2 G^{\neq})] dE \quad (6)$$

The total number of electrons in the system due to the voltage bias is calculated as the trace of the density matrix. The voltage bias causes a charging of the molecular energy levels. The potential cannot be obtained without a knowledge of the change in the number of electrons. For this reason, the change in number of electrons is solved self consistently for each bias.

3 RESULTS AND DISCUSSION

Following the above methodology, the current-voltage (I-V) characteristics of four bases, namely, adenine, guanine, thymine, and cytosine, have been calculated. Figure 3 shows the structure of each of the bases and their attachment to the gold contacts. The distance between the terminal atoms and the gold contacts is held constant at 3.0 angstroms for all computations. The 3' carbon of the ribose sugar is attached to one of the gold contacts, and the other contact is attached to the closest atom of the base. The adenine base is attached to the 1-N on the purine ring, the 3-N of cytosine pyrimidine ring, the 1-N hydrogen of the guanine purine ring, and the double bonded oxygen on the 4-C of the thymine pyrimidine ring. Figure 4 illustrates the typical density of the states (DOS) of adenine. The states at lower energy (i.e. below –12 eV) are composed of the occupied orbitals. A gap in the DOS between –12 and –8.5 eV represents the gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). The DOS is used as input to calculate I-V profile using the NEGF approach.

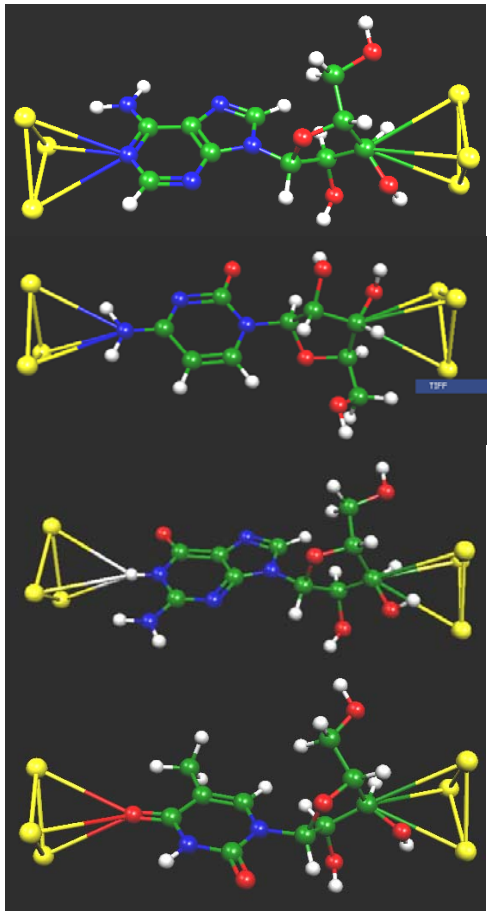

Figure 3: Structure of adenine, cytosine, guanine and thymine base systems. A cluster of three gold atoms are used as contacts.

One of the most important aspects prior to computing current is the evaluation of Fermi energy level (E_f). In the present work, Equation 7 is used to estimate the charge neutralization energy.

$$n(E) = 2\int_0^E D(E)dE \tag{7}$$

The Fermi energy can be estimated from the plot of E vs. n(E) (Damle, 2002). Figure 5 is a plot of energy versus n(E) for adenine, showing that the energy increases sharply when the number of electrons approaches 102. This is the total number of valence electrons in the unperturbed adenine system. The energy corresponding to 102 electrons is the Fermi energy level, and its value is –8.81 eV. It is interesting to note that the Fermi energy level lies closer to the LUMO than the HOMO. Similar results were obtained for other base systems. Figure 6 shows the I-V characteristics of all four bases. All four bases have distinctly different I-V characteristics. Cytosine is the most conductive, and guanine is the least conductive (see inset plot in Figure 6). Adenine and thymine have similar characteristics below 2 volts. However at higher bias (>2volts), all four bases can be distinguished based on the tunneling current. This offers theoretical proof of concept for a notional tunneling-current-based nanopore sequencer.

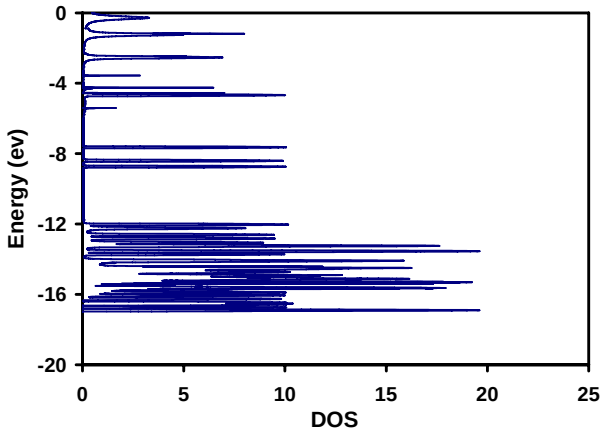

Figure 4: Theoretical density of states for adenine system.

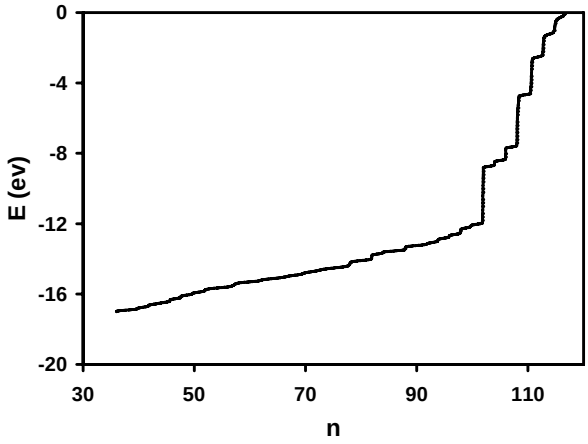

Figure 5: Charge neutralization profile for adenine system,used to estimate Fermi energy level

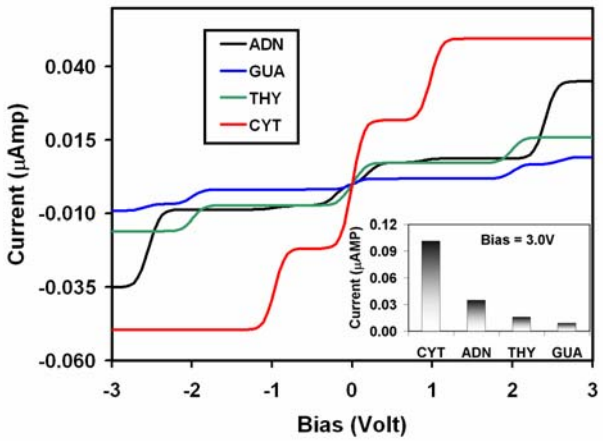

Figure 6: IV characteristics of the four base systems computed using quantum chemistry and NEGF. The inset figure is a plot of the current at a constant bias of 3 V.

Figure 7 shows the conductance of the four base systems. The peak in conductance is due to the proximity of the Fermi level to the LUMO. A change in conductance

is observed as the chemical potential (μ's) of the contacts is close in energy with the orbitals around the LUMO. Since the density of states around the LUMO is sparse, the conductance peaks can be observed individually.

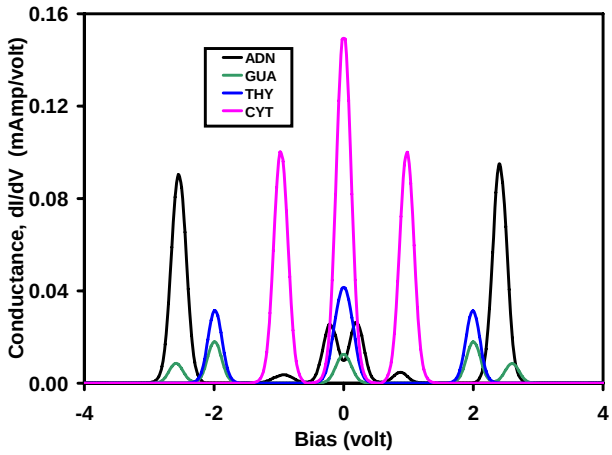

Figure 7: Conductance vs. applied bias profile for four base systems.

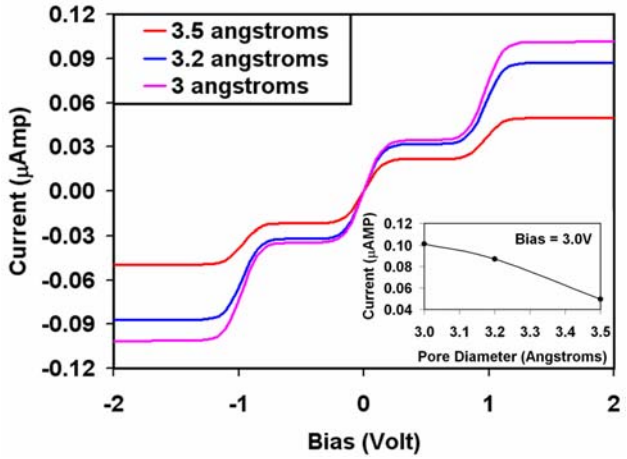

Figure 8: Variation of IV profile for cytosine system when distance between the left contact and the terminal atom is varied.

Finally, the effect of pore diameter on the tunneling current is evaluated in Figure 8. This figure shows that the tunneling current decreases with an increase in pore diameter. In the computations, the contact on the left side is moved away from the terminal atom of cytosine base. The inset plot in Figure 7 illustrates that the tunneling current is quite sensitive to the diameter of the nanopore. The current drops by a factor of 2 with a 0.50 angstrom increase in the pore diameter, indicating a scaling of current with the -4.75 power of distance. This result strongly indicates that a smaller pore diameter with closely spaced electrodes will optimize molecular contact and maximize the resultant tunneling current.

4 CONCLUSIONS / FUTURE WORK

The present work is aimed at computing tunneling current of four constituent bases, namely adenine, thymine, guanine and cytosine, using quantum chemistry and non-equilibrium Green's function theory. The calculations presented in this work, while under idealized conditions, clearly indicate that the use of tunneling current measurements to sequence DNA as it translocates through a nanopore is indeed possible. However, practical success is shown to depend critically upon parameters such as biasing voltage and pore diameter. Future model refinements will include more realistic molecular descriptions that incorporate the effects of solvent and ions present within the system in order to evaluate the I-V characteristics of DNA bases under realistic conditions in order to:

- Calculate the magnitude of open pore tunneling current as background noise.
- Estimate the effect of thermal and geometric fluctuations on the resultant tunneling current.
- Evaluate the I-V characteristics of DNA bases under realistic conditions.

One major theoretical hurdle is that accurate *ab initio* methods are too costly for realistic system sizes. Future work will require the development of new computational methods capable of rapidly estimating I-V characteristics of DNA for the purposes of in-silico design and testing of a realistic device.

5 ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Air Force Research Laboratory (FA8750-05-C-0055, Technical Monitor: Dr. Tom Renz, P.I. Zbigniew Sikorski) for funding.

REFERENCES

[1] Damle, P.; Ghosh, A.; and Datta, S. Chem. Phys. 281:171 (2002).
[2] Fink, H. and Schonenberger, C. Nature 398:407 (1999).
[3] J. Jenkins, D. Sengupta, S. Sundaram, LNCS 3516:309-316 (2005).
[4] S. Datta; Superlattices and Microstructures; 28 (2000) 253
[5] J. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. Grinkova, S. Sligar, K. Schulten, G. Timp. Nanoletters **5**:1883-1888 (2005).
[6] D. Fologea, J. Uplinger, B. Thomas, D. McNabb, and J. Li. Nanoletters **5**:1734-1737 (2005).