Modeling of electron transport in biomolecules: Application to DNA

M. P. Anantram and Jianqing Qi

Department of Electrical Engineering, University of Washington, Seattle, WA, 98195-2500

Tel: 1-206-221-5162, Email: anant@uw.edu and jqqi@uw.edu

ABSTRACT

We review our work aimed at understanding electron transport in DNA molecules and present new results on specific strands. The motivation for our work stems from both: (i) DNA offering a unique framework for nanoscale electronic devices and (ii) recent work to detect diseases by measuring electrical current in DNA. Designer sequences of DNA show the theoretical possibility to behave as superlattices.

Introduction

DNA molecules are quasi one-dimensional structures consisting of two separate strands in a double helix form. Each strand consists of four building blocks, Adenine (*A*), Thymine (*T*), Cytosine (*C*) and Guanine (*G*). The diameter of the DNA molecule is approximately 2 nm (Fig. 1 (a)).

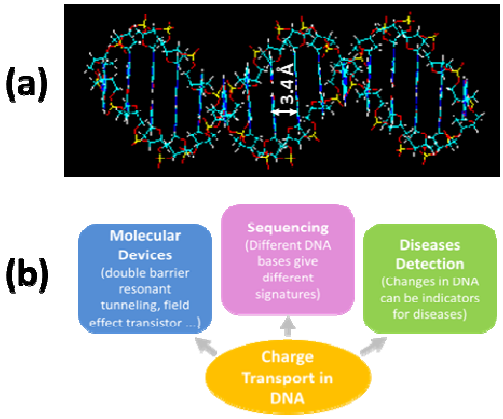


Fig. 1 (a) A double strand DNA consists of two single strands in a double helix form. The distance between the bases is ~ 3.4 Å. (b) Biomolecules (DNA, peptides and RNA) offer a platform for electronic devices. Further, electrical methods for both sequencing and disease detection, where the biomolecule is the channel for electron transport have recently been demonstrated.

These building blocks can be reproducibly engineered in arbitrary sequences of *A*, *T*, *C* and *G*, and chemists have for long studied sequence dependent charge transfer in DNA [1]. The distinct energy levels and ionization potentials of the building blocks provides a framework to construct DNA sequences with interesting device properties. One can think of quantum wells and barriers constructed from DNA having resonances akin to double barrier resonant tunneling diodes and superlattices [2]. Of great interest are also recent experiments which show the potential to detect diseases [3] and sequence DNA [4] by measuring the electrical conductance of a single DNA molecule (Fig. 1 (b)). These studies demonstrate potential for an all-electrical method for disease detection, using very small sample sizes without the amplification step required in PCR.

The modeling of electrical transport in DNA molecules connected to contacts is in early stages compared to solid-state nanotransistors. The main differences between electronic transport in biological molecules and solid state devices arise from: (a) the floppy nature of biomolecules, as a result of which conformational changes have low energy barriers, (b) the relatively large distance between the building blocks of DNA (*A*, *T*, *C* and *G*) and proteins (amino acids) and (c) environment (water molecules and counter ions).

Method

Our model development involves two stages. The first stage involves thousand atom systems that are modeled using a combination of classical molecular dynamics / energy minimization followed by first principles based quantum chemistry calculations [5]. The second stage involves the extraction of a class of tight-binding Hamiltonians that account for a subset of energy levels at each base, and hopping parameters that describe transport both along a single-strand (intra strand) and between complementary strands (inter strand). Fig. 2 shows a typical result from our calculations, which yield a tight binding model for a DNA sequence consisting of eight *G:C* base pairs.

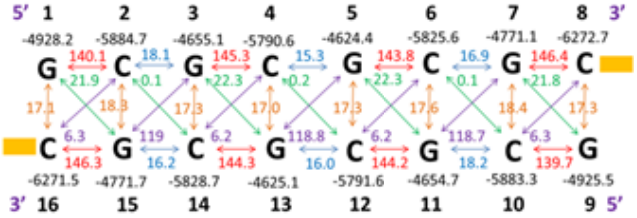


Fig. 2 (**Tight binding model**): An effective tight binding model that is obtained from the full first principles based Hamiltonian of short DNA segments. Tight binding Hamiltonians are useful in studying long strands, which are computationally intractable via first principles based methods. The numbers shown in black are onsite potentials required to mimic the valence (HOMO) band. The numbers shown in different colors are hopping parameters between neighboring bases within and between strands. Units are meV.

Each base on a single strand is separated from its neighbors by 3.4 Å. This large inter-base distance makes the hopping integral between bases small. To appreciate this, we note from Fig. 2 that the tight-binding parameter to hop between consecutive bases is smaller than 150 meV, which is considerably smaller than nanotubes and nanowires, which have tight-binding parameters that are almost ten times larger (2-3 eV). In comparison to solid state systems, DNA strands are floppy and

melt at low temperatures. As a result, vibrational coupling modifies electronic transport significantly as the length of DNA increases as discussed below.

During the course of modeling the low bias conductance measured in [6], we realized that decoherence plays an important role in obtaining an order of magnitude agreement with experiments. Including decoherence precisely requires knowledge of the Hamiltonians corresponding to various scattering mechanisms and their inclusion in transport calculations, which is an extremely difficult task. Instead, we currently include decoherence using phenomenological Buttiker probes [7, 8], after transforming our Hamiltonian to an orthogonalized form [2]. An independent Buttiker probe is connected to each energy level on the base (*A, T, C, G*) and the current flowing into the Buttiker probe is set equal to zero at every energy point.

Results: Barrier and Superlattice

We now review the construction of tunable barriers and superlattices using engineered DNA sequences.

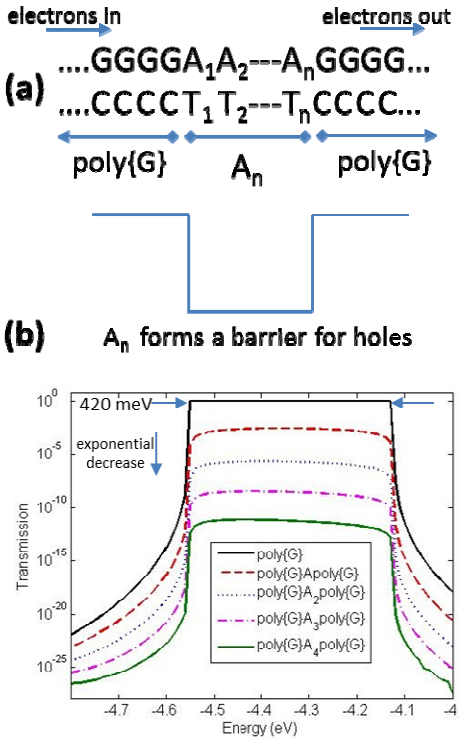


Fig. 3: (a) A suitably designed double strand is a barrier for hole transport. (b) The adenines form a barrier for hole flow between the two poly{G} regions. (c) The black curve is transmission in the poly{G} strand without adenines, which is unity in the valence band. As the number of Adenines (*N*) increases, the transmission in the entire valence band exponentially decreases.

The ionization potential of Guanine is smaller than that of Adenine, Thymine and Cytosine [11]. As a result, a strand that consists of *poly{G}A$_n$ poly{G}* (the complementary strand is *poly{C}T$_n$ poly{C}*) will behave as a barrier for holes. Our calculated transmission for such a strand as a function of *n*, the number of *A:T* forming the barrier is shown in Fig. 3 (a) and (b). The transmission decreases exponentially with increase in the length of the barrier (*n*). When there are no Adenines, the transmission is unity, as expected, within the valence band, which has a bandwidth of only ~ 420 meV. As the number of Adenines increases from 1 to 4, the transmission decreases exponentially (Fig. 3(c)).

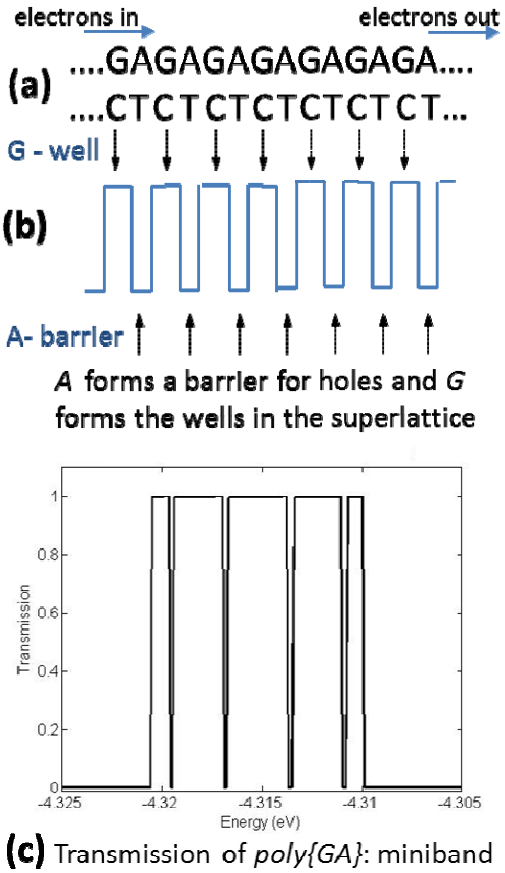


Fig. 4: (a) The *poly{GA}* strand behaves as a quantum well superlattice. (b) Effective potential energy diagram. (c) Transmission versus energy. The width of the miniband is approximately 10 meV.

The poly{*GA*} structure (a periodic repetition of *GA* along with the complementary strand) consists of barriers (*A*) separated by wells (*G*) as shown in Figs. 4 (a) and (b). Based on the knowledge of superlattices in semiconductor heterostructures, one would expect a miniband whose width depends on the strength of the barriers (*A*) and energy level of the wells (*G*). Fig. 4 (c) shows the highest lying valence miniband. The width of the miniband is a little larger than 10 meV, which is indicative of the strength of strong barriers between consecutive *G*s. We remark that we expect a strand that consists of *poly{G}AGGGApoly{G}* to behave as a double barrier resonant tunneling structure, where the number of transmission resonances depends on the width of Guanines [2]. While promising, the experimental realization of device concepts discussed above needs careful attention because weak electronic

coupling between bases (Fig. 2) makes the conductance susceptible to defects and variations in local environment.

Results: Decoherence

Using the full Hamiltonians from [5], we calculate the conductivity of the DNA strands shown in Fig. 5. These strands were used in the novel experiments of [6], and involves a systematic increase in the number of *A:T* base pairs from Sequence 1 to 4 (A:T base pairs are barriers, Fig. 3). The comparison between experiment and modeling in the coherent limit, is shown in Fig. 6 for Sequence 3. We find that the low bias experimental conductance is more than a million times larger than the calculated conductance (independent of position of the Fermi energy).

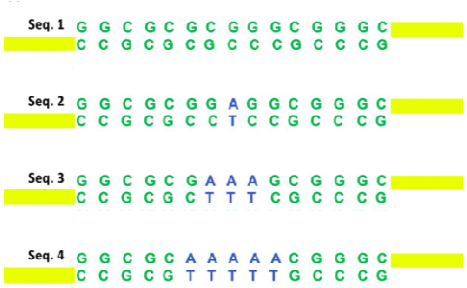


Fig. 5 (**Experimental DNA strands**): Four strands that are fifteen base pairs long in the experiments of [6]. The length of *A:T* region in the middle changes from Sequence 1 to 4. The *A:T* regions are barriers and as a result the conductance from Sequence 1 to 4 monotonically decreases. Our model with decoherence provides insight to understand the values of the experimental conductance. The yellow bars represent metal contacts.

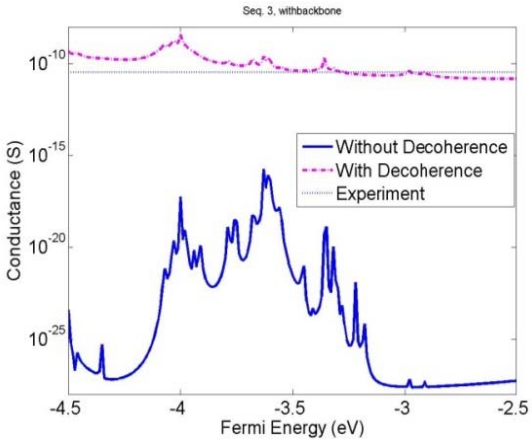


Fig. 6: The value of conductance calculated are significantly smaller (solid blue) than the experimental value of 3.5 E-11 S (dashed blue). Varying the position of Fermi energy and coupling to the contacts does not help in increasing the conductance to values comparable to the experiment. The experimental conductance is a million times larger than the calculated conductance without decoherence (blue). We have found it essential to include decoherence (pink) to obtain values comparable to experiments.

To rationalize this large difference, we have proposed that decoherence should be included [10]. Decoherence helps broaden the density of states at energy levels of neighboring bases, and hence improves electronic coupling. The decoherence rates of 6 meV for *G:C* base pairs and 1.5 meV for *A:T* base pairs are chosen because they give a good match to experiments (Fig. 6). It is important to note that (i) our results hold independent of the strength of coupling between DNA and metal contacts, (ii) the assumed decoherence rates are comparable to those reported in literature for nanotube/wire systems and (iii) these decoherence rates give a relatively good match to all four strands (Fig. 5) in the experiment of [6].

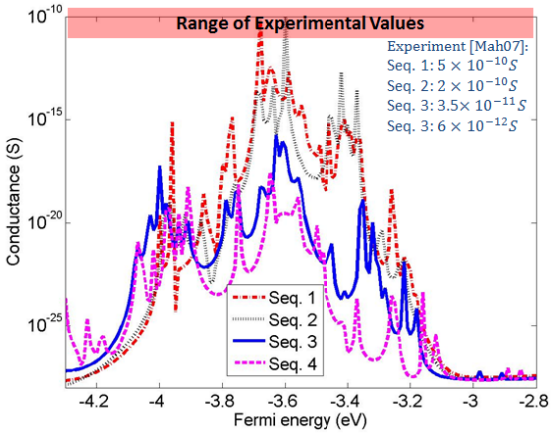


Fig. 7 (**Comparison to experiment, without decoherence):** The red band shows the range of experimental values for the conductance of all four sequences in Fig. 5. Our calculated conductance without decoherence is orders of magnitude smaller than the experimental values, independent of the location of Fermi energy, for all four sequences [10].

Fig. 7 shows that the experimental conductances of the four strands in Fig. 5 are much larger than what modeling predicts in the absence of decoherence. But when decoherence values used in Fig. 6 are used, the magnitude of our calculated conductance is comparable to experiments, and the trend of conductance decrease from Sequences 1 to 4 seen in experiments is captured (Fig. 8).

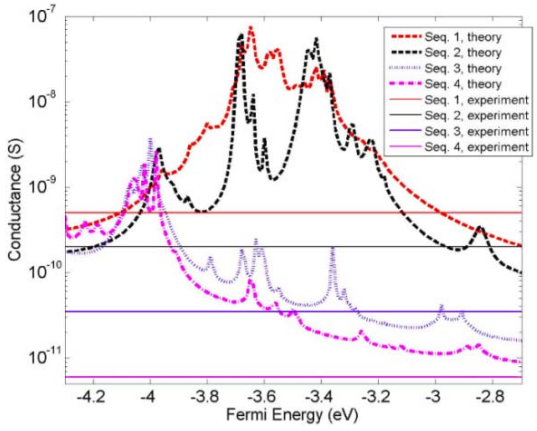


Fig. 8 (**Comparison to experiment, with decoherence**): The conductance as a function of Fermi energy obtained from modeling [10]. The horizontal lines show the experimental values from [6]. The decoherence rates for *G:C* and *A:T* base pairs is taken to be 6 meV and 1.5 meV respectively. The simulations agree in magnitude and trend with experimental conductance.

Results: Detection of epigenomic difference in DNA by electrical transport

Methylation of DNA bases is a wide spread epigenomic mutation in a variety of cells, including stem cells [9]. The methylation of Cytosines in DNA is linked to repression of gene transcription and is regarded to be a step in DNA mutation. Fig. 9 (a) shows methylation, where a Hydrogen in Cystosine has been replaced by a methyl (CH_3) group. The identification of methylation in samples at low concentration is challenging. Conventional techniques such as PCR are not trustworthy at low concentrations and they do not preserve methylation of the original sequence. Two recent experiments have shown that direct electrical measurement of methylation in a single DNA molecule is possible [3, 4]. Reference [3] showed that the conductance of a strand consisting of eight base pairs (Fig. 9 (b)) with methylated Cytosine is smaller than that of an equivalent strand with native Cystosine. We have used our model to calculate the conductance of the two strands from [3] to study factors responsible for the experimental results. Our preliminary calculations show that the conductance of the methylated Cytosine strand is intrinsically smaller than that of the native Cytosine strand (Fig. 9 (c)) when the conformation and coupling to contacts are identical in both cases. This result is in qualitative agreement with experiment but admittedly a clear determination of the Fermi energy at equilibrium has not been possible.

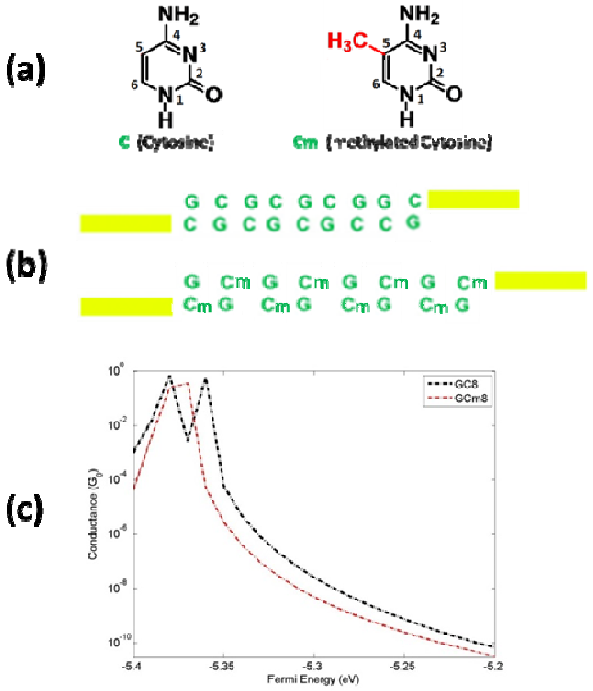


Fig. 9 (**Methylated DNA**): (a) Cytosine and methylated-Cytosine, where the Hydrogen in location 5 has been replaced by a methyl group. Methylation of *G:C* to *G:Cm* is a biomarker that helps detect some diseases. (b) The two DNA strands correspond to *poly{GC}* and *poly{GCm}* sequences that are eight base pairs long. Experiments find that the methylated strand has a lower conductance than the native strand. The yellow bars represent metal contacts. (c) The conductance from our calculations are in qualitative agreement with experimental trends.

Summary

In summary, we use relatively large computational models to derive insight into electron flow in biological molecules using methods used to model nanodevices. We showed examples where engineered DNA heterostructures behave as barriers and superlattices. Using the model developed and experimental data from [6], we discussed the importance of decoherence in obtaining even an order of magnitude agreement with experiments involving long strands of DNA [10]. Finally, we applied the model to recent experiments aimed at detecting epigenomic differences by measuring electrical conductance. Our calculations show that difference in conductance between methylated (*G:Cm*) and native (*G:C*) DNA pair sequences is in principle possible. The determination of Fermi energy through the DNA structures with contacts remains a challenge that requires further attention.

Acknowledgements

We are grateful to Prof. David Janes (Purdue University), and Prof. Josh Hihath (UC Davis) for many useful discussions. We acknowledge support from National Science Foundation under Grant No. 102781.

References

[1] E. Meggers, M. E. Michel-Beyerle, and B. Giese, Sequence dependent long range hole transport in DNA. J. Am. Chem. Soc., vol. 120, p. 12950 (1998)

[2] Ch. Adessi, S. Walch and M. P. Anantram, Environment and structure influence on DNA conduction, Phys. Rev. B, vol. 67, p. 81405(RC) (2003); J. Qi, M. G. Rabbani, S. Edirisinghe, and M. P. Anantram, Transport of charge in DNA heterostructures, 11th IEEE Conference on Nanotechnology (IEEE-NANO), pp. 487-491, 2011; H. Mehrez and M. P. Anantram, Inter-base electronic coupling for transport through DNA, Phys. Rev. B, vol. 71, p. 115405 (2005)

[3] J. Hihath, S. Guo, P. Zhang, and N. Tao, Effect of cytosine methylation on DNA charge transport, J. Phys.: Condens. Matter, vol. 24, 164204 (2012)

[4] M. Tsutsui, K. Matsubara, T. Ohshiro, M. Furuhashi, M. Taniguchi, and T. Kawai, Electrical Detection of Single Methylcytosines in a DNA Oligomer, J. Am. Chem. Soc., vol. 133, p. 9124 (2011)

[5] M. J. Frisch et al. computer code GAUSSIAN 09, Revision A.1, Gaussian, Wallingford, CT, 2009

[6] A.K.Mahapatro, K. J. Jeong, G. U. Lee, and D. B. Janes, Sequence specific electronic conduction through polyion-stabilized double-stranded DNA in nanoscale break junctions, Nanotechnology, vol. 18, 195202 (2007)

[7] M. Buttiker, Four-Terminal Phase-Coherent Conductance, Phys. Rev. Lett., vol. 57, 1761 (1986)

[8] J. L. D'Amato and H. M. Pastawski, Conductance of a disordered linear chain including inelastic scattering events, Phys. Rev. B 41, 7411 (1990)

[9] R. Lister et. al, Human DNA methylomes at base resolution show widespread epigenomic differences, Nature, vol. 462, p. 315 (2009)

[10] J. Qi, N. Edirisinghe, M. G. Rabbani, and M. P. Anantram, Unified model for conductance through DNA with the Landauer-Buttiker formalism, Physical Review B, vol. 87, p. 085404 (2013)

[11] N. Russo, M. Toscano and A. Grand, Theoretical determination of electron affinity and ionization potential of DNA and RNA bases, J. Comput. Chem., vol. 21, p. 1243 (2000)