Inter-base Electronic Coupling for transport through DNA

H. Mehrez† and M. P. Anantram†

NASA Ames Research Center, Moffett Field, CA 94035-1000, USA

(Dated: December 27, 2021)

We develop a new approach to derive single state tight binding (SSTB) model for electron transport in the vicinity of valence-conduction bands of poly(G)-poly(C) and poly(A)-poly(T) DNA. The SSTB parameters are derived from first principles and are used to model charge transport through finite length DNA. We investigate the rigor of reducing the full DNA Hamiltonian to SSTB model to represent charge transport in the vicinity of valence-conduction band. While the transmission coefficient spectrum is preserved, its position shifts in energy. Thymine is poorly represented and its peak height is substantially reduced. This is attributed to the abstraction of the HOMO-LUMO coupling to other eigen-states in the nearest neighbor DNA bases, and can be corrected within 2^{nd} order time independent perturbation theory. Inter-strand charge transport has also been analyzed and it is found that hopping to the nearest neighbor in the complementary strand is the most important process except in the valence band of poly(G)-poly(C), where hopping to the second nearest neighbor between 3′-ends is the most dominant process. As a result, transport between 3′-ends and 5′-ends in the vicinity of valence band of polyG-polyC is asymmetric.

I. INTRODUCTION

The growing interest in DNA as a molecular device has led to a wide range of experimental as well as theoretical work in the field. However, charge transport results through DNA are still controversial. Experimental it is found that DNA can either be a good conductor, a semi-conductor or even an insulator. Theoretically, while the exact charge transport mechanism is not clear, band transport, polaronic transport, fluctuation facilitated charge migration, and multi-step hopping have been investigated. The narrowing down of the precise charge transport mechanism has been difficult because the base pairs are weakly coupled due to the large inter-base separation of ~ 3.4 Å. This translates into relatively small hopping parameters for electrons between base pairs at the HOMO and LUMO energy levels (less than 125 meV), which results in valence and conduction bandwidths of less than 500 meV. As a result of the narrow bandwidth, charge transport in DNA is easily modified by environment effects, which include counter-ions, impurities, defects as well as hydration. Recently, ab initio calculations have focused on determining inter-base coupling so that DNA can be represented within a single-state tight binding (SSTB) model. However, these derivations suffer from the limitations that they either cannot be extended to describe interaction between different bases, or can only find interactions between HOMO states (valence band) of DNA bases. Moreover, the rigor of reducing the full Hamiltonian to the SSTB model in describing electronic transport through finite length DNA has not been addressed before.

The aim of this work is to develop a general approach to obtain tight binding parameters to describe intrastrand and inter-strand interactions from first principles. This formalism will be applied to poly(G)-poly(C) and poly(A)-poly(T) DNA.

II. METHODOLOGY

We describe our model to determine the tight binding parameters for poly(G)-poly(C) here, and follow an identical approach for poly(A)-poly(T). Initially one GC base pair is constructed using the Nucleic Acid Builder (NAB) software package. The backbone of the structure generated by NAB is replaced by a hydrogen atom, and the position of these H atoms are relaxed at the MP2 level. The optimized structure of the hydrogen terminated DNA bases are shown in Fig.1-a. Using this configuration, double strand B-DNA structures of four, six and eight GC base pairs are constructed. The self consistent Hamiltonian $H_0$ of this structure is calculated using density functional theory, where the B3LYP density functional and 6-31G basis set are used.

To work in an orthogonal basis set, we initially transform the Hamiltonian to $H_1$ such that,

$$H_1 = S_0^{-1/2}H_0S_0^{-1/2},$$

where $S_0$ is the overlap matrix. Every diagonal sub block of $H_1$, which corresponds to a DNA base, is diagonalized and its eigenvectors are used to construct a block diagonal matrix $U$. Following this, $H_1$ is transformed to,

$$H_2 = U^\dagger H_1 U.$$
\[ + \sum_{<n_g,n'_g>} t_{n_g,i,n'_g,j} \left( C^\dagger_{n_g,i} C_{n'_g,j} + c.c. \right) \]
\[ + \sum_{<n_c,n'_c>} t_{n_c,i,n'_c,j} \left( C^\dagger_{n_c,i} C_{n'_c,j} + c.c. \right) \]
\[ + \sum_{i=1 \rightarrow N_g} t_{n_g,i,n_c,j} \left( C^\dagger_{n_g,i} C_{n_c,j} + c.c. \right) \]
\[ + \sum_{i=1 \rightarrow N_g, j=1 \rightarrow N_c} t_{n_g,i,n_c,j} \left( C^\dagger_{n_g,i} C_{n_c,j} + c.c. \right) \]  

\[ \epsilon_{n_g,i} (\epsilon_{n_c,i}) \text{ is the } i^{th} \text{on-site energy of base } n_g (n_c), \]
\[ t_{n_g,i,n'_g,j} (t_{n_c,i,n'_c,j}) \text{ is the hopping parameter between energy levels } i \text{ and } j \text{ of base pairs } n_g \text{ and } n'_g \text{ (and } n_c \text{ and } n'_c \text{ respectively).} \]
\[ t_{n_g,i,n_c,j} \text{ is the inter-strand hopping parameter between energy levels } i \text{ and } j \text{ of base pairs } n_g \text{ and } n_c. \]
\[ N_g \text{ and } N_c \text{ are the eigen-states in a single guanine and cytosine respectively.} \]
\[ \mathbf{C}^\dagger \text{ and } \mathbf{C} \text{ are the creation and annihilation operators and c.c. is the hermitian conjugate.} \]
\[ < ... > \text{ and } \ll ... \gg \text{ represent 1st and 2nd nearest neighbor interactions, respectively.} \]

In finding the parameters of the Hamiltonian in Eq. 8, we have used the central two base pairs of the simulated system to minimize edge effects. The Hamiltonian in Eq. 8 is referred to as the full DNA model in the remainder of this paper.

III. RESULTS

A. SSTB parameters

We present our results for on-site energy and intra-strand hopping parameters of the HOMO and LUMO states in table I. In poly(G)-poly(C), the HOMO (LUMO) state is localized on guanine (cytosine), and the hopping parameter between consecutive guanines (cytosines) is 115 meV (61 meV). These values are much larger than the ones for poly(A)-poly(T), where the HOMO (LUMO) state is localized on adenine (thymine), and the corresponding hopping parameter between consecutive adenines (thymines) is 21 meV (23 meV). We have also determined the inter-strand hopping parameters which are shown in Table II. Since the helical structure of the DNA breaks reflection symmetry in a plane perpendicular to the axis,\(^\text{25}\) we find that the inter-base hopping parameters depend on the directionality between 3'− and 5'− ends (Fig I(b)).

The most striking difference is for the HOMO state of poly(G)-poly(C), where \( t_H \ll 3' - G - C - 3' \gg = 50 \text{ meV} \) (dotted line in Fig I(b)) and \( t_H \ll 5' - G - C - 5' \gg = 7 \text{ meV} \) (continuous line in Fig I(b)). We find that the directionality dependence of the inter-strand hopping parameters cause significant asymmetry in inter-strand charge transport as discussed in the following sub-section. Finally, we note that both the intra-strand and inter-strand hopping parameters do not depend on the system size as indicated by the results for the four and eight base pair systems shown in Tables I and III.

B. Transport Results

Charge transport experiments typically involve either measuring the current-voltage characteristics of a DNA placed between metal contacts\(^\text{23}\) or measuring the charge transfer between donor and acceptor intercalators placed along DNA strands.\(^\text{22}\) While the parameters we derive here can be used to model both sets of experiments, we will focus on the former experimental configuration in the low bias limit. In presenting our results, we will compare the transmission probability in the vicinity of HOMO-LUMO states obtained using the full DNA (Eq. 8) and SSTB (Tables I and III) models. However, to maintain the same injection rate from the leads to the device in comparing the two models, we account for all eigen-states of the edge base pairs that are coupled to the contacts. The transport calculations are carried out within the Landauer-Büttiker formalism\(^\text{28}\) and the transmission probability is,

\[ T = tr[\Gamma^L G \Gamma^R G^a], \]

where \( G^{r(a)} \) is the retarded (advanced) Green’s function of the isolated DNA attached to the contacts, and \( \Gamma^L(R) \) is the device coupling to the left (right) contact. While the coupling, \( \Gamma^L(R) \), is crucial in determining the transport properties, it depends on the details of DNA-contact coupling, which is difficult to control experimentally. We have considered two limits of the DNA-contact coupling: the weak coupling limit where \( \Gamma^L = \Gamma^R = 10 \text{ meV} \) and the strong coupling limit where \( \Gamma^L = \Gamma^R = 500 \text{ meV}\(^\text{28}\)). However, we present the results for the latter configuration.

Finally, we note that in modeling intra-strand transport both DNA strands are coupled to the contacts at both ends. In the inter-strand representation only one strand is coupled to the contact at each end.

Intra-strand transport: The results for intra-strand transport from the full DNA and SSTB models are shown by the solid and dashed lines respectively, in Fig. 2. The SSTB model reproduces the peak and width of the transmission windows quite well, with the main difference being a shift in both the HOMO and LUMO transmission windows to higher energies in the case of SSTB model. We find that this energy shift occurs due to coupling of the HOMO and LUMO states of a base pair to other eigen-states of neighboring base pairs. Further, because the hopping parameters between the HOMO/LUMO states of a base pair and the other eigen-states of neighboring base pairs is smaller than their energy separation, we find that the energy shift seen in
Can be quite accurately accounted for using second order perturbation theory. The expressions for the second order correction to the HOMO (H) or LUMO (L) eigen-values are

\[ \Delta E_{n_g(n_c),H(L)}^{(2)} = - \sum_{j \neq H(L)} \frac{t_{n_g(n_c),H(L)j}^2 \epsilon_{n_g, H(L)} - \epsilon_{n_g, H(L)} \epsilon_{n_c, H(L)}}{ \epsilon_{n_g, H(L)} - \epsilon_{n_c, H(L)}}. \] (5)

We find that the energy shift for G (C) bases at the center to be \(-49\) \((-69\) meV. The bases at the left and right edges have only one neighbor, and are shifted by \(-29\) \((-29\) and \(-40\) meV, respectively. When this shift is included, the eigenvalues from the SSTB model match the full DNA model. Calculations of the transmission probability which incorporate this second order correction in the SSTB model agree closely with the full DNA model as shown by the open circles in Fig. 2a.

The transmission probability for poly(A)-poly(T) are presented in Fig. 2b. It is clear from the right panel of Fig. 2b that the transmission window of thymine undergoes both a large energy shift and substantial peak height reduction compared to the case of poly(G)-poly(C). The reasons for this are: (i) weak intra-strand hopping parameters for the LUMO level of thymine compared to cytosine (table I) and (ii) large second order energy correction to the LUMO state due to strong interaction with other energy eigen-states. Following the perturbation theory analysis of Eq. (5), we find the second order energy correction to the LUMO eigen-state of thymine is \(-100\) meV while that of the HOMO eigen-state of adenine is only \(-41\) meV, for bases that are away from the edges. The corresponding corrections for bases at the left and right edges are \(-48\) and \(-52\) \((-16\) and \(-25\) meV respectively, for thymine (adenine). Again including these corrections in the Hamiltonian for the calculation of the transmission probability shows that the SSTB model for poly(A)-poly(T) agrees closely to the full DNA model as shown by the open circles in Fig. 2b.

We now focus on the importance of including all eigen-states of the edge base pairs connected to the contact, in modeling charge transport. When the broadening due to the contacts \( \Gamma^L \) and \( \Gamma^R \) is large, there is a non zero density of states in the valence and conduction band energy windows at the edge base pairs due to states other than \( \epsilon_H \) and \( \epsilon_L \). The hopping parameter between these states at the edges and, \( \epsilon_H \) and \( \epsilon_L \) of neighboring base pairs are non zero. As a result, charge injected into the tails of energy eigen-states other than \( \epsilon_H \) and \( \epsilon_L \), contributes to the transmission probability at valence and conduction band energies. This contribution is referred to as the tail effect and can be partitioned out of the total transmission probability. In Fig. 2 we present transmission spectrum between only HOMO (left panels of Fig. 2) and LUMO (right panels of Fig. 2) states at the contact as triangle symbols.

Clearly, this corresponds to \( \sim 90\% \) and \( 25\% \) of the total transmission for \( G-C \) and \( A-T \), respectively. It indicates that contribution of other eigen-states to the total transmission coefficient is important and they must be included by representing the contact with all modes.

Inter-strand transport: Current-voltage measurements of inter-strand transport involve metal contacts connected to complementary strands. That is, contacts are connected to either only 5′-ends or 3′-ends of the DNA as shown in Fig. 1b. We first note that irrespective of which strand an electron is injected into, transport occurs primarily along poly(G) (poly(C)) if electrons are injected energetically into the valence (conduction) band. Inter-strand hopping, which is the transmission limiting step, occurs mainly near the contacts because of the tiny density of states of the valence (conduction) band in poly(C) (poly(G)). We show the transmission probability for inter-strand transport in both the valence and conduction bands of poly(G)-poly(C) in Fig. 3. The inter-strand transmission probability is more than an order of magnitude smaller than the intra-strand case because the density of states is peaked only along one strand in both the conduction and valence bands. Hopping into the strand with a smaller density of states limits the transmission / conductance in inter-strand transport.

We will now gain some insight into the roles of the nearest and second nearest neighbor hopping parameters in table II in determining inter-strand transport. This is done by calculating the transmission probability by setting specific inter-strand hopping parameters at the edges of the DNA to zero in the Hamiltonian. The solid (dashed) lines in Fig. 3a,b correspond to setting the nearest (second nearest) inter-strand neighbor hopping parameter shown by the solid (dashed) line respectively in Fig. 3c to zero. The solid triangles in Fig. 3a,b is the reference, which corresponds to the full model. Clearly, in the conduction band, the nearest neighbor inter-strand hopping process shown by the solid line in Fig. 3c is the most important in determining the transmission probability between both the 3′-ends and 5′-ends. Further, the conduction band transmission probabilities across the 3′-ends and 5′-ends are comparable. In contrast, in the valence band, the transmission probability across the 3′-ends is twice as large as that across the 5′-ends. The reason for this is that the second nearest neighbor hopping parameter \( t_H \ll 3′ - G - C - 3′ \) determines the transmission probability across the 3′-ends, and is much larger (50 meV) than all other hopping parameters in the valence band of poly(G)-poly(C) as seen in Table II.

We have also investigated this asymmetric inter-strand transport for a hundred base pair (33.8 nm long) poly(G)-poly(C) system. We have found that it is persistent and this represents a strong motivation to perform such inter-strand transport experiment. However, we note that these results are sensitive to the DNA base conformation. Hence thermal effects can substantially reduce the asymmetric coupling along the DNA. Therefore, these ex-
Experiments should be performed at low temperature with ambient conditions that lead to B-DNA conformation. Inter-strand transport through poly(A)-poly(T) have also been carried out and showed symmetric transmission between the 3'-ends and 5'-ends. Inter-strand hopping occurs at nearest neighbor basis closest to the contact region.

IV. CONCLUSION

We have investigated the rigor of reducing the full Hamiltonian of a DNA to a single-state tight binding (SSTB) model. Tight binding parameters for both intra-strand and inter-strand transport have been tabulated. We find that the SSTB model quite accurately reproduces the transmission probability calculated from the full Hamiltonian, when second order corrections to the on-site potential of base pairs is included. As a result, the SSTB model is computationally efficient when compared to the full Hamiltonian. One caveat while applying the SSTB model is that injection of charge from the contacts to the edge base pair should be carefully modeled. This is because charge can be injected into the tails of all eigen-states of the edge base pairs, thereby making conduction due to the tail effect important. We have also investigated inter-strand charge transport and found strong asymmetric inter-strand current in the vicinity of the valence band (HOMO state) of poly(G)-poly(C). Current flowing between Ohmic contacts connecting the 3'-ends is almost twice as large as the current between Ohmic contacts connecting the 5'-ends.

Acknowledgments HM was supported by NASA contract to UARC/ELORET and NSERC of Canada. MPA was supported by NASA contract to UARC.

| Base | $\epsilon_H$ | $\epsilon_L$ | $t_L^{BB}$ | $t_H^{BB}$ |
|------|-------------|-------------|------------|------------|
| G    | -4278 (-4278) | 1137 (1148) | 19 (20)    | -115 (-114) |
| C    | -6519 (-6533) | -1065 (-1072) | -61 (-60) | -24 (-21) |
| A    | -5245 (-5245) | 259 (258)   | 24 (25)   | 21 (21)    |
| T    | -6298 (-6346) | -931 (-972) | -23 (-23) | -98 (-98)  |

| System | $t_L$ | $t_H$ |
|--------|------|------|
| <G-C>  | 63 (63) | 2 (5) |
| <A-T>  | 34 (34) | 26 (26) |
| $\ll 5' - G - C - 5' \gg$ | -12 (-12) | -7 (-8) |
| $\ll 3' - G - C - 3' \gg$ | -16 (-15) | 50 (48) |
| $\ll 5' - A - T - 5' \gg$ | -10 (-10) | -11 (-11) |
| $\ll 3' - A - T - 3' \gg$ | -13 (-13) | 9 (9) |
In the B-DNA structure, base pairs repeat to form a one-dimensional helix with an inter-base separation of 3.38 Å and a pitch angle of 36°. Such a pitch angle breaks symmetry along the helix and defines a direction between the DNA ends. They are called 5′ – and 3′ – ends of each strand. A four-G-C base pair structure with 5′ – and 3′ – ends description is shown in Fig. 14b.

The unit cell \( t^{L,R} \) should be compared to \( t_{DNA}^{L,R}/DOS_{DNA}(E_F) \) where \( t_{DNA} \) and \( DOS_{DNA}(E_F) \) are the intra-strand hopping parameter and DNA density of states at Fermi energy. But for DNA semi-infinite one-dimensional system with coupling \( t_{DNA} \), \( DOS_{DNA}(E_F) = 1/\pi/t_{DNA} \). Hence, \( \Gamma \) should be compared to \( t_{DNA} \).

S. Bhattacharya, J. Choi, S. Lodha, D. B. Janes, A. F. Bonilla, K. J. Jeong, and G. U. Lee, IEEE NANO proceedings, 79 (2003).

W. D. Cornell, et al., J. Am. Chem. Soc. 117, 5179 (1995).

In the B-DNA structure, base pairs repeat to form a one-dimensional helix with an inter-base separation of 3.38 Å and a pitch angle of 36°. Such a pitch angle breaks symmetry along the helix and defines a direction between the DNA ends. They are called 5′ – and 3′ – ends of each strand. A four-G-C base pair structure with 5′ – and 3′ – ends description is shown in Fig. 14b.

Gaussian 98 (Revision A.7), M. J. Frisch et al., Gaussian, Inc., Pittsburgh PA, 1998.

S. O. Kelley and J. K. Barton, Science 283, 375 (1999).

R. Landauer, IBM J. Res. Dev. 1, 223 (1957); M. Büttiker, IBM J. Res. Dev. 32, 317 (1988). S. Datta, Electronic Transport in Mesoscopic Systems, Cambridge University Press, New York, (1995).

In the B-DNA structure, base pairs repeat to form a one-dimensional helix with an inter-base separation of 3.38 Å and a pitch angle of 36°. Such a pitch angle breaks symmetry along the helix and defines a direction between the DNA ends. They are called 5′ – and 3′ – ends of each strand. A four-G-C base pair structure with 5′ – and 3′ – ends description is shown in Fig. 14b.

S. Bhattacharya, J. Choi, S. Lodha, D. B. Janes, A. F. Bonilla, K. J. Jeong, and G. U. Lee, IEEE NANO proceedings, 79 (2003).

W. D. Cornell, et al., J. Am. Chem. Soc. 117, 5179 (1995).

In the B-DNA structure, base pairs repeat to form a one-dimensional helix with an inter-base separation of 3.38 Å and a pitch angle of 36°. Such a pitch angle breaks symmetry along the helix and defines a direction between the DNA ends. They are called 5′ – and 3′ – ends of each strand. A four-G-C base pair structure with 5′ – and 3′ – ends description is shown in Fig. 14b.
FIG. 1: (a) Atomic structure of hydrogen terminated DNA bases where arrows indicate H atom which replaces the backbone. (b) 4 base pair G-C arranged in B-DNA configuration. 5', 3'-ends of DNA are shown and the solid (dotted) arrow correspond to $<5' - G - C - 5'> (<3' - G - C - 3'>)$ directional coupling.

FIG. 2: Intra-strand conductance of (a) poly(G)-poly(C) and (b) poly(A)-poly(T). Left (Right) panels correspond to transport through HOMO (LUMO) states. The linear response conductance is calculated at 300K using $T(E) = \int dE' T(E') f(E - E')$ where $f$ is the Fermi function evaluated at 300K. Solid line - full dna model; Dashed lines - SSTB model with all eigen-states at the two edges; Open circles - Same as “Dashed lines” but includes the second order energy correction given in Eq. 4; Solid triangles - full DNA model minus the tail effect. Insets: Same as main figures but conductance is calculated at 50K.

FIG. 3: Inter-strand conductance (calculated at 300K) for poly(G)-poly(C). Left (Right) panels correspond to transport through HOMO (LUMO) states. In (a) only 5'- ends have an Ohmic contact while in (b) only 3'- ends have an Ohmic contact. Solid triangle - full model; Solid (Dashed) lines - Same as the solid triangle, except that the nearest (second nearest) neighbor shown by the solid (dashed) lines in (c) are set to zero. In (c), up-arrow (↑) corresponds to the Ohmic contacts where charge is injected and collected. Each diagram in (c) corresponds to the graph in (a) and (b) directly above it.
Fig. 1
Fig. 2
Fig. 3