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: August 13, 2004)

We develop a new approach to derive the DNA tight binding parameters from *first principles*. These parameters are used to model charge transport through finite length *G-C* and *A-T* DNA. We investigate the rigor of reducing the full DNA Hamiltonian to a single HOMO-LUMO state 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 to be strongly asymmetric in PolyG-PolyC in the vicinity of the HOMO state. In the $3'G \rightarrow C3'$ configuration, current is carried through the *second nearest neighbor* DNA basis and it is $\sim 100\%$ larger than current from $5'G \rightarrow C5'$.

The growing interest in DNA as a molecular device[1,2] has led to a wide range of experimental as well as theoretical work in the field. However, transport results through DNA are still controversial[3–16]. The disagreement in the interpretation of DNA results is due to inter-base coupling. It is weak and hence it 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 from *first principles*[2,11,16] so that DNA can be represented within a single-state tight binding model. However, these derivations suffer from the limitations that they either cannot be extended to describe interaction between different basis[16] or can only find interactions between HOMO states (valence bands) of each DNA basis[11]. In addition, the rigor of reducing the full Hamiltonian to single-state tight binding model in describing electronic transport through finite length DNA, has not been addressed before.

It is the aim of this work to develop a general approach to obtain tight binding parameters to describe intra- and inter-strand interactions from *first principles*. This formalism will be applied to PolyG-PolyC (*G-C*) and PolyA-PolyT (*A-T*) DNA structures.

Model: We describe in detail our model to determine tight binding parameters for *G-C* system and we follow the same approach for *A-T*. Initially *one* base-pair *G-C* is constructed using NAB program[17]. The back-bone is replaced by a hydrogen atom. The positions 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, a double strand B-DNA structure[18] of either *four, six or eight G-C* base-pairs (BP) is constructed. The self consistent Hamiltonian $\mathbf{H_0}$ and the overlap matrix $\mathbf{S_0}$ are constructed within B3LYP DFT. These calculation are performed using Gaussian 98 (G98) program[19], and the $6-31G$ basis set is used.

To work in an orthogonal basis set, we initially transform $\mathbf{H_0}$ to $\mathbf{H_1}$ such that, $\mathbf{H_1} = \mathbf{S_0}^{-1/2}\mathbf{H_0}\mathbf{S_0}^{-1/2}$. Every diagonal sub-block of $\mathbf{H_1}$, which corresponds to a DNA basis, is diagonalized and its eigenvectors are used to construct a block-diagonal matrix $\mathbf{U}$. Following this, $\mathbf{H_1}$ is transformed to, $\mathbf{H_2} = \mathbf{U}^\dagger\mathbf{H_1}\mathbf{U}$. In this representation of $\mathbf{H_2}$, diagonal elements represent the localized energy eigen-states of DNA bases and off-diagonal sub-blocks correspond to inter-base interactions. Independent from the simulated DNA system size, being either *four, six or eight* BP, we find the interaction between HOMO/LUMO states and the *second* nearest neighbor basis to be small and have been neglected. Hence the Hamiltonian is written as

$$\begin{aligned}\mathbf{H} = & \sum_{\substack{n_g \\ i=1\rightarrow N_g}} \epsilon_{n_g,i}\mathbf{C}^\dagger_{ng,i}\mathbf{C}_{n_g,i} + \sum_{\substack{n_c \\ i=1\rightarrow N_c}} \epsilon_{n_c,i}\mathbf{C}^\dagger_{n_c,i}\mathbf{C}_{n_c,i} \\ & + \sum_{\substack{<n_g(n_c),n'_g(n'_c)> \\ i,j=1\rightarrow N_g(N_c)}} t_{n_g(n_c),i;n'_g(n'_c),j}(\mathbf{C}^\dagger_{n_g(n_c),i}\mathbf{C}_{n'_g(n'_c),j} + c.c.) \\ & + \sum_{\substack{\ll n_g,n_c \gg \\ i(j)=1\rightarrow N_{g(c)}}} t_{n_g,i;n_c,j}(\mathbf{C}^\dagger_{n_g,i}\mathbf{C}_{n_c,j} + c.c.) \qquad (1)\end{aligned}$$

In Eq. 1, $\epsilon_{n_g(n_c),i}$ is the i^{th} on-site energy on the DNA basis $n_g(n_c)$; $t_{n_g(n_c),i;n'_g(n'_c),j}$ is the hopping parameter between two eigenstates; $\mathbf{C}^\dagger$ and $\mathbf{C}$ are the creation and annihilation operators and *c.c.* is the hermitian conjugate. $< ... >$ and $\ll ... \gg$ represent interactions up to 1^{st} and 2^{nd} nearest neighbor, respectively.

All interaction parameters of full DNA model, presented in Eq. 1, are derived from the *two* BP at the center of the simulated system so that edge effects are removed. In Table I, we show our results for the on-site energy and intra-strand coupling of the HOMO and LUMO states, derived from *four* and *eight* BP system. Here we note that the on-site energy is scaled so that HOMO eigenstates of G and A match for both calculations. In *G-C*, the HOMO (LUMO) state is localized on Guanine (Cytosine) and coupling between consecutive Gs (Cs) is found to be 115meV (61meV). These values are much larger than the ones corresponding to *A-T* where the HOMO (LUMO) state is localized on A (T) and the correspond-

ing hopping parameter is $21meV$ ($23meV$). We have also determined inter-strand hopping parameters which are also shown in Table I. $<G$-$C>$ and $<A$-$T>$ correspond to nearest DNA basis with hydrogen bonding, and the rest are second nearest neighbor basis interactions. Since DNA basis pitch angle breaks symmetry along the helix[18], inter-base hopping parameter can depend on the direction. For G-C HOMO state, in particular, $t(3' - G - C - 3')$ is $50meV$ while $t(5' - G - C - 5')$ is only $7meV$. This difference can result in asymmetry in the inter-strand charge transport and is investigated later. Finally we note that the difference between parameters obtained from *four* and *eight* BP system is negligible. The HOMO-LUMO gap has changed by less than 40 meV while the hopping parameters undergo a less than 3meV variation.

Transport Results: Transport analysis through DNA can be carried out for two different experimental configuration. Current measurement experiment[3] where metal/DNA contact is fabricated and charge is carried between both contacts. The other experiment corresponds to placing Donor/Acceptor at different DNA sites[20] and charge is transported through DNA. Both of these experiments, can be analyzed within our parameters derived in Eq. 1. However, we will focus on the former experimental configuration and consider *zero-bias* analysis so that the Hamiltonian matrix elements are unperturbed. We will compare transport results in the vicinity of HOMO-LUMO states for full DNA model (Eq. 1) and single state tight binding representation with the parameters shown in Table I. However, to keep the same injection rate from the leads to the device, DNA bases which are in contact with the leads are represented with all eigenstates.

Transport calculations are carried out within Landauer-Büttiker formalism[21]. The transmission coefficient is, $T = tr[\Gamma^L \mathbf{G}^r \Gamma^R \mathbf{G}^a]$, where $\mathbf{G}^{r(a)}$ is the retarded (advanced) Green's function of DNA with the contacts and $\Gamma^{L(R)}$ is the device coupling to the left (right) contact. The coupling of the device to the outside is crucial, but difficult to determine. Experimentally, it is also difficult to control this coupling and hence we parameterize the external lead effect. In our model, we consider the case where only DNA basis at the edges to be coupled to the leads[5,6], and we thoroughly investigate the weakly $\Gamma^L = \Gamma^R = 10$ meV and the strongly $\Gamma^L = \Gamma^R = 500$ meV coupling regimes[22]. However, we present results for the latter configuration.

Intra-strand transport: In this regime both complementary strands are coupled to the contact. Our conductance calculations are shown in Fig. 1. Results for full DNA model are displayed as solid lines, while calculations which incorporate a single state, either HOMO or LUMO, are presented as dashed lines. Clearly, single state tight binding model typifies transport characteristics through G-C. However, the spectrum of the transmission coefficient shifts to higher energy. This is due to the abstraction of the HOMO-LUMO coupling to other eigen-states in the nearest neighbor DNA bases. This energy shift can be approximated within time-independent perturbation theory and the *second order* correction for the HOMO (H) or LUMO (L) state is

$$\Delta E^{(2)}_{n_g(n_c),H(L)} = - \sum_{\substack{\beta=n_g\pm1,n_c\pm1 \\ j\neq H(L)}} \frac{t^2_{n_g(n_c),H(L);\beta,j}}{\epsilon_{\beta,j} - \epsilon_{n_g(n_c),H(L)}} \quad (2)$$

We find that the energy shift for G (C) bases at the center to be -49 (-69) meV, while bases at the left and right edges are shifted by -20 and -29 (-29 and -40) meV, respectively. When this shift is included, the eigenvalues of single eigen-state model match the full DNA representation. Transport calculations, which incorporate this correction, are shown as open circles in Fig. 1-c,d, and result in better transport representation

Transport calculations along A-T are presented in Fig. 1-e,f. It is clear from Fig. 1-f that Thymine spectrum for the transmission coefficient undergoes a large energy shift and substantial peak height reduction. This poor representation is due to weak inter-base coupling and large energy shift due to strong interactions with other energy eigen-states. Following the same perturbation theory analysis of Eq. 2, the 2^{nd} order energy correction to the LUMO eigenstate is found to be -100 meV while that of Adenine is only -41 meV for bases at the center. However T (A) bases at the left and right edges undergo energy correction of -48 and -52 (-16 and -25) meV, respectively. If this is taken into account, as in the case for G-C, single state model for A-T improves the results as shown by open circles in Fig. 1-e,f.

Finally we note that in this regime, $\Gamma^L = \Gamma^R = 500meV$, the states with energy eigenvalues not residing in the valence-conduction band can still have a finite density of states in this energy range. Hence, they contribute to a finite transmission probability. We define transmission through these states as the "tail effect" and has substantial contribution. In Fig. 1 we present transmission spectrum between HOMO (Fig. 1-c,e) and LUMO (Fig. 1-d,f) 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 the "tail effect" is important and must be included by representing the contact with all modes.

Inter-strand charge transport: To investigate the inter-strand transport, donors and acceptors must be placed at complementary strands as performed in recent experiments[20]. Another approach, requires that either $5'-$ends or $3'-$ends[18](Fig. 1-b) have Ohmic contact and inter-strand current measurement is carried out[23]. However, inter-strand charge transport has not been investigated in the latter configuration. The analysis of this inter-strand current transport in which charge is injected from one Ohmic contact of one strand and collected at the opposite end of DNA at the complementary strand, is the focus of the rest of the paper.

In Fig. 2 we show the inter-strand charge transport re-

sults through *G-C*. Calculations that include all interactions are presented as triangle down symbols. The left (right) panel of each graph corresponds to transport through HOMO (LUMO) state. It is clear from the left panels of Fig. 2-a,b that transport through *G-C* is asymmetric in the vicinity of HOMO state. To determine the underlying mechanism for this behavior, we set specific Hamiltonian elements to *zero*, as indicated in the schematic diagram in Fig. 2-c. Since DNA bases in each strand have different energy eigenstates, charge is carried through a single strand and inter-strand hopping occurs only at bases close to contact. Hence, an important inter-strand path would be the hydrogen-bond between bases at the $3' - 5'-$ ends. This bond is set to *zero* for calculations shown as solid lines. Indeed, Fig. 2-a,b shows that eliminating this path would drastically decrease inter-strand *G-C* current flow, except for the left panel of Fig.2-b. However, as can be seen in this figure with dashed lines, eliminating interaction between DNA basis having Ohmic contact with the second nearest neighbor in the complementary strand results in a secondary effect. Therefore, for *G-C*, charge transport occurs through a single strand and inter-strand hopping occurs through hydrogen-bond at the DNA basis at the contact. The exception occurs for *G-C* with $3'-$ ends having Ohmic contact and at an injection energy close to the HOMO eigenstate. This result can also be revealed from inter-strand hopping parameters shown in Table I. It indicates that 1^{st} nearest neighbor ($< ... >$) basis coupling is the strongest except for the *G-C* system at the HOMO state, where 2^{nd} nearest neighbor ($\ll ... \gg$) is more significant. An important consequence of this inter-strand hopping parameter difference is that $3' \rightarrow 3'$ current is more than 100% larger than $5' \rightarrow 5'$ current (see Fig. 1-b) . We have also confirmed that this asymmetric inter-strand current is persistent even for 100 BP (33.8 nm long) *G-C* system. This is a strong motivation to perform such inter-strand transport experiment. However, we note that these results are sensitive to the DNA basis conformation. Hence thermal effects can substantially reduce the asymmetric coupling along the DNA. Therefore, these experiments should be performed at low temperature with ambient conditions that lead to B-DNA conformation.
Inter-strand calculations through *A-T* have also been carried out and showed symmetric behavior. Inter-strand hopping occurs at nearest neighbor basis closest to the contact region.

Conclusion: We have developed a simple and general approach to determine the DNA tight binding parameters from *first principles*. We used these parameters to model of charge transport in the vicinity of valence-conduction band through a finite length DNA. We have investigated the rigor of reducing the full Hamiltonian to single-state tight binding representation. We find that the transmission coefficient shifts to higher energy and results in a poor description of transport through thymine. The shift is due to the abstraction of HOMO-LUMO coupling to other states and is corrected within *second order* time independent perturbation theory. We have also investigated inter-strand charge transport and found strong asymmetric inter-strand current in the *G-C* system in the vicinity of HOMO state. Current flowing between Ohmic contacts at the $3'$-ends of *G-C* is found to be twice as large compared to current measured between Ohmic contacts at 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.

TABLE I: Upper table, on-site energy and intra-strand DNA electronic coupling (meV) for states near the HOMO-LUMO (H-L) states. Lower table, inter-strand DNA electronic coupling (meV) for L and H states. $< ... >$ and $\ll ... \gg$ correspond to inter-strand hydrogen bond and 2^{nd} nearest neighbor inter-strand bond, respectively. These values are obtained from *four (eight)* BP system.

Base	ϵ_H	ϵ_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)

* also ELORET, Mail Stop 229-1
† also UCSC, Mail Stop 229-1

[1] M. Di Ventra and M. Zwolak, *DNA electronics* in *Encyclopedia of Nanoscience and Nanotechnology* edited by H. Singh-Nalwa, American Scientific Publishers (2004).
[2] R. G. Endres, D. L. Cox, and R. R. P. Singh, *Rev. Mod. Phys.*, **76**, 195 (2004).
[3] H. W. Fink and C. Schönenberger, *Nature* (London) **398**, 407 (1999); K.-H. Yoo, D. H. Ha, J.-O. Lee, J. W. Park, J. Kim, J. J. Kim, H.-Y. Lee, T. Kawai, and H. Y. Choi, *Phys. Rev Lett.* **87**, 198102 (2001); D. Porath, A. Bezryadin, S. de Vries, and C. Dekker, *Nature* (London) **403**, 635 (2000); P. J. de Pablo, F. Moreno-Herrero, J. Colchero, J. Gómez Herrero, P. Herrero, A. M. Baró, Pablo Ordejón, José M. Soler, and Emilio Artacho *Phys. Rev Lett.* **85**, 4992 (2000).
[4] R. Di Felice, A. Calzolari, E. Molinari, and A. Garbesi, *Phys. Rev. B* **65**, 045104 (2001).
[5] Ch. Adessi, S. Walch, and M. P. Anantram, *Phys. Rev. B* **67**, R081405 (2003); Ch. Adessi and M. P. Anantram, *App. Phys. Lett.* **82**, 2353 (2003).
[6] H. Mehrez, S. P. Walch and M. P. Anantram, submitted to *Phys. Rev Lett.* .
[7] E. M. Conwell and S. V. Rakhmanova, *Proc. Natl. Acad. Sci.* **97**, 4556 (2000).
[8] Y. J. Ye, R. S. Chen, A. Martinez, P. Otto, and J. Ladik, *Solid St. commun.* **112**, 129 (1999).
[9] R. Bruinsma, G. Grüner, M. R. D'Orsogna, and J. Rudnick *Phys. Rev Lett.* **85**, 4393 (2000).
[10] Y. A. Berlin, A. L. Burin, M. A. Ratner, *J. Phys. Chem. A* **104**, 443 (2000).
[11] A. A. Voityuk, J. Jortner, M. Bixon, and N. Rösch, *J. Chem. Phys.* **114**, 5614 (2001); A. A. Voityuk, N. Rösch, M. Bixon, and J. Jortner, *J. Phys. Chem. B***104**, 9740 (2000).
[12] S. Roche, *Phys. Rev Lett.* **91**, 108101 (2003); M. Unge and S. Stafstrvm, *Nano Lett.* **3**, 1417 (2003); O. R. Davis and J. E. Inglesfield, *Phys. Rev. B* **69**, 195110 (2004).
[13] R. N. Barnett, C. L. Cleveland, A. Joy, U. Landman, G. B. Schuster, *Science* **294**, 567 (2001).
[14] F. L. Gervasio, P. Carloni, and M. Parrinello, *Phys. Rev Lett.* **89**, 108102 (2002).
[15] P. Maragakis, R. L. Barnett, E. Kaxiras, M. Elstner, and T. Frauenheim, Phys. Rev. B **66**, 241104(R) (2002).
[16] E. Artacho, M. Machado, D. Sánchez-Portal, P. Ordejón, and J. M. Soler, *Molecular Physics* **101**, 1587 (2003).
[17] NAB program (Version 4.5), T. Macke, W.A. Svrcek-Seiler, and D. A. Case.
[18] In B-DNA structure, BP are stuck in a *one dimensional* helix with inter-base separation of 3.38Å and a pitch angle of 36^o. Such a pitch angle breaks symmetry along the helix and defines a direction between the ends. They are called 5' and 3' ends of each strand. A *four G-C* BP structure with $5'$ and $3'$ ends description is shown in Fig. 1-b.
[19] Gaussian 98 (Revision A.7), M. J. Frisch *et.al.*, Gaussian, Inc., Pittsburgh PA, 1998.
[20] S. O. Kelley and J. K. Barton, *Science* **283**, 375 (1999).
[21] R. Landauer, *IBM J. Res. Dev.* **1**, 223 (1957); M. Büttiker, *IBM J. Res. Dev.* **32**, 317 (1988). S. Datta, *Electronic Transport in Mesescopic Systems*, Cambridge University Press, New York, (1995).
[22] Γ should be compared to $t_{DNA}^2/DoS_{DNA}(E_F)$. But for DNA *one dimensional* system with coupling t_{DNA}, $DoS_{DNA}(E_F) = 1/\pi/t_{DNA}$. Hence, Γ should be compared to t_{DNA}.
[23] In D. B. Janes group [S. Bhattacharya et. al. , IEEE NANO proceedings, 79 (2003)], thiol groups are attached to single strand DNA at either $3'-$ or $5'-$ ends to perform electronic transport measurement through double stranded DNA. If thiol groups are attached to either $3'-$ or $5'-$ ends of complimentary strands, inter-strand charge transport can be carried out.

FIG. 1: (a) Atomic structure of hydrogen terminated DNA bases where arrows indicate H atom which replaces the backbone and its position is fully relaxed. (b) 4 BP G-C arranged in B-DNA configuration[18]. $5', 3'$-ends of DNA are shown and the solid (dotted) arrow correspond to $< 5' - G - C - 5' >$ ($< 3' - G - C - 3' >$) directional coupling. The rest of panels show intra-strand conductance of G-C (c,c',d,d') and A-T (e,e',f,f'). c,c',e,e' (d,d',f,f') panels correspond to transport through HOMO (LUMO) states. Conductance calculated at $300K$ and $50K$ for c,d,e,f and c',d',e',f', respectively; where temperature has only broadening effect in our model. The solid lines are calculations performed using all eigenstates. Dashed lines correspond to calculations with a single mode included inside DNA and all modes at contact. Open circles are with a single mode inside DNA with energy shift of HOMO-LUMO states to take into account interaction with all other modes. Triangle down depicts the transmission of full DNA model from HOMO [LUMO] to HOMO [LUMO] state at the contacts in (c,e) [(d,f)].

FIG. 2: Inter-strand conductance (calculated at $300K$) for G-C (a,b). In (a) only $5'$ ends have Ohmic contact while in (b) $3'$ ends have Ohmic contact. Left (Right) panels correspond to transport through HOMO (LUMO) states. Triangle down symbols correspond to results when all the Hamiltonian is used and continuous as well as dashed lines represent calculations when specific interactions are turned off. In (c), up-arrow ($\uparrow$) corresponds to the Ohmic contact where charge is being injected/collected. Each diagram represents the interactions that are set to zero (solid and dashed lines) to obtain the corresponding graphs in (a) and (b).