Theoretical Investigation of Base Pairs-Dependent Electron Transport in DNA System

Alaa Ayad K. Al-mebir**, and Shakir A. A. AL-Saidi*
Department of Physics, College of Education of Pure Sciences, University of Thi-Qar, Nasiriyah, Iraq
**alaa_ash28@yahoo.com, *Shakir.hussain@sci.utq.edu.iq

Abstract: DNA molecules contain high density genetic information that make them beside their self-assembly and self-recognition properties very remarkable topics for many scientific majors, such as medicine, biology and nanotechnology. The mechanism of electron mobility through DNA is important to study DNA-based molecular electronics in nano-bioelectronics. In this study, the influence of number of base pairs on transmission probability, conductance and current for a DNA system has been investigated by using a tight-binding model. Two DNA sequences, (G/C) and (A/T), was studied using the steady state formalism. One active region of DNA molecules assumed to be represented by a fishbone model. The behavior of transmission spectrum \( T \) was discussed for both DNA sequences as a function of the number of base pairs \( N \). In addition, a nonlinear behavior of the current \( I \) was found by studying \( I-N \) curves. Further, the resistance-dependent of number of base pairs is also discussed to describe the resistance behavior and values for both DNA sequences. It demonstrates that the resistance-dependent of number of base pairs represented by nonlinear dependent with small different in its values, and the two DNA sequences gives different \( R-N \) curve characteristics.

Keywords; DNA, Electron transport, Tight-binding model, Fishbone model

1. Introduction

Charge mobility in DNA has taken much attention during the recent decades for its remarkable effect on both biological processes and physical applications [1-3]. The DNA has been used as molecular devices [4], and its properties has been used in self-assembling and self-recognition [5, 6]. Also, DNA is well recognized to behave as molecular wire [7], especially in regular sequences as a poly (GC), or by its doping with metallic ions as in the state of M-DNA [8]. Despite the various experimental conditions and intrinsic complexities within the DNA molecules [9-11], DNA molecules may act as a semiconductor with large band gap and close bandwidth [12]. An energy band shape has been studied [13], and the current-voltage characteristics has been successfully interpreted [14]. The goal of this work is studying the correlation of number of base pairs-dependent with transmission, conductance and current for two DNA sequences, (G/C) and (A/T), using tight-binding model and fishbone model [15].

In the simplest single-particle tight-binding model, it has been assumed that the system can be formalized into a fishbone model in which an effective on-site energy is used for the LUMO energy of each base pair, and for the effect of the coupling between any two nearest neighbor sites. For a homogeneous DNA, this model should be very suitable due to the charge migration occurring along DNA molecule in the process of the electron transport. The schematic model that we considered in this study is presented in Fig (1). The electron is scattered from one bridge that contains of DNA. There is a single conduction channel in which sites represent bases. A coupling interaction occurs in every link between locations. The description of DNA base pairs as single positions represents a simplification of
the fishbone model. This distinction becomes important when hopping between base pairs is considered, especially when the hopping from base pair to base pair is the same.

Fig (1): A schematic illustration of the fishbone model. The left and right ends of the DNA(Bases) are connected to the electrode L and R in addition to D(donor) and A(accepter)

2. Theory and Derivation
We describe the system under consideration that is shown in Fig (1) by using the following time-independent Hamiltonian (using Dirac’s notations). This electronic Hamiltonian considers all the sub-systems interactions. The different indexes D, A, L, R, Br and bq denote the donor, acceptor, left lead, right lead, bridge (DNA base pairs with total number N) and backbone, respectively. The derivation, which you can also check in our previous paper [16], is started with the Hamiltonian as following:

\[
\hat{H} = E_D |D\rangle \langle D| + E_A |A\rangle \langle A| + \sum_{k_{Br}} E_{k_{Br}} |k_{Br}\rangle \langle k_{Br}| + \sum_{k_{L}} E_{k_{L}} |k_{L}\rangle \langle k_{L}| + \sum_{k_{R}} E_{k_{R}} |k_{R}\rangle \langle k_{R}| \\
+ \sum_{q=1,\ldots,N} \sum_{k_{bq}} E_{k_{bq}} |k_{bq}\rangle \langle k_{bq}| + \sum_{q=1,\ldots,N} \sum_{k_{bq}} \left( V_{A_kbq} |A\rangle \langle k_{bq}| + h. c \right) \\
+ \sum_{q=1,\ldots,N} \sum_{k_{bq}} \left( V_{D_kbq} |D\rangle \langle k_{bq}| + h. c \right) \\
+ \sum_{k_{Br}} (V_{A_kR} |A\rangle \langle k_{R}| + h. c) + \sum_{k_{L}} (V_{D_kL} |D\rangle \langle k_{L}| + h. c) \tag{1}
\]

The index \( k_i \) is the energy wave vector, where \( i \) represents the indexes D, A, Br, bq, L and R. \( E_i \) represents the \( i \)th energy level position, and \( |i\rangle \) and \( \langle i| \) represent the ket and bra states, respectively. \( V_{ij} \) represents the coupling interaction between the subsystems \( i \) and \( j \) where \( V_{ij} = V_{ji} \). (h.c.) represents the harmonic conjugate.

The system wave function can be written as:

\[
\psi(t) = C_D(t) |D\rangle + C_A(t) |A\rangle + \sum_{k_{Br}} C_{k_{Br}}(t) |k_{Br}\rangle + \sum_{k_{L}} C_{k_{L}}(t) |k_{L}\rangle + \sum_{q=1,\ldots,N} \sum_{k_{bq}} C_{k_{bq}}(t) |k_{bq}\rangle + \sum_{k_{R}} C_{k_{R}}(t) |k_{R}\rangle \\
+ \sum_{k_{L}} C_{k_{L}}(t) |k_{L}\rangle + \sum_{k_{R}} C_{k_{R}}(t) |k_{R}\rangle \tag{2}
\]

where \( C_i(t) \) represents the linear expansion coefficients.

The equations of motion for \( C_i(t) \) can be obtained by using time-dependent Schrodinger equation:
\[ \frac{\partial \psi(t)}{\partial t} = -i \hat{H}_0 \psi(t) \]  

(3)

So, we get the following set of related equations:

\[ \dot{\mathcal{C}}_D(t) = -i E_D \mathcal{C}_D(t) - i \sum_{k_L} V_{Dk_L} C_{k_L} C_{k_D}(t) - i \sum_{k_R} \sum_{k_{Rq}} V_{Dk_R} C_{k_R} C_{k_{Rq}}(t) \]  

(4)

\[ \dot{\mathcal{C}}_A(t) = -i E_A \mathcal{C}_A(t) - i \sum_{k_R} V_{Ak_R} C_{k_R} C_{k_A}(t) - i \sum_{k_L} \sum_{k_{Lq}} V_{Ak_L} C_{k_L} C_{k_{Lq}}(t) \]  

(5)

\[ \dot{\mathcal{C}}_{k_{Bq}}(t) = -i E_{k_{Bq}} C_{k_{Bq}}(t) - i \sum_{k_{BR}} V_{k_{BR}k_{Bq}} C_{k_{BR}} C_{k_{Bq}}(t) - i V_{k_{Bq}D} \mathcal{C}_D(t) - i V_{k_{Bq}A} \mathcal{C}_A(t) \]  

(6)

\[ \dot{\mathcal{C}}_{k_{BR}}(t) = -i E_{k_{BR}} C_{k_{BR}}(t) - i \sum_{q=1} V_{k_{BR}k_{q}} C_{k_{q}}(t) \]  

(7)

\[ \dot{\mathcal{C}}_{k_{L}}(t) = -i E_{k_{L}} C_{k_{L}}(t) - i V_{k_{L}D} \mathcal{C}_D(t) \]  

(8)

\[ \dot{\mathcal{C}}_{k_{R}}(t) = -i E_{k_{R}} C_{k_{R}}(t) - i V_{k_{R}A} \mathcal{C}_A(t) \]  

(9)

For steady state, we define \( \mathcal{C}_i(t) \) as \( \mathcal{C}_i(t) = \mathcal{C}_i e^{\int E \, dt} \), where \( E \) representing the system eigen values, so, we put \( \dot{\mathcal{C}}_i = 0 \) and by considering the following separation procedure:

\[ V_{k_{i\alpha}} = v_{ki} V_{i\alpha} \]  

(10)

\[ \overline{c}_{ki} = v_{ki} \mathcal{C}_i \]  

(11)

Where \( \alpha = A, D, Br \) and \( bq \)

By substituting these definitions in eq. (4) and (9), we get

\[ \overline{c}_{A}(E) = \frac{1}{E - E_A} \left( \sum_{k_R} V_{Dk_R}^2 \overline{c}_R + \sum_{k_{Rq}} V_{Dk_{Rq}}^2 \overline{c}_{k_{Rq}} \right) \]  

(12)

\[ \overline{c}_D(E) = \frac{1}{E - E_D} \left( \sum_{k_{LR}} V_{Dk_{LR}}^2 \overline{c}_L + \sum_{k_{LRq}} V_{Dk_{LRq}}^2 \overline{c}_{k_{LRq}} \right) \]  

(13)

\[ \overline{c}_{Br}(E) = \frac{1}{E - E_{Br}} \left( \sum_{k_{BRq}} V_{Brk_{BR}}^2 \overline{c}_{k_{BRq}} \right) \]  

(14)

\[ \overline{c}_{Bq}(E) = \frac{1}{E - E_{Bq}} \left( V_{BqD} \overline{c}_D + V_{BqA} \overline{c}_A + V_{BqBR} \sum_{k_{BR}} V_{k_{BRq}}^2 \overline{c}_{k_{BRq}} \right) \]  

(15)

\[ \overline{c}_L(E) = \frac{1}{E - E_L} V_{LD} \overline{c}_D \]  

(16)

\[ \overline{c}_R(E) = \frac{1}{E - E_R} V_{RA} \overline{c}_A \]  

(17)

Accordingly, we obtain obvious expression for:
\[ C_A(E) = \frac{X_1(E)}{X_2(E)} C_D(E) \]  

(18)

Where,

\[ X_1(E) = \left\{ \sum_{q=1,4} V^{Abq} \Gamma_{bq}(E) V^{bqD} + \Gamma_{Br}(E) \sum_{q=1,4} V^{Abq} \Gamma_{bq} V^{bqBr} \right\} \]  

(19)

and,

\[ X_2(E) = \left\{ E - E_A - \sum_{q=1,4} \sum_{A_{bq}} \Gamma_{bq}(E) \right\} \left[ 1 - \Gamma_{Br}(E) \sum_{q=1,4} \sum_{Br_{bq}} \Gamma_{bq}(E) \right] \]  

\[- \Gamma_{Br}(E) \left| \sum_{q=1,4} V^{Abq} \Gamma_{bq}(E) V^{bqBr} \right|^2 \]  

(20)

where

\[ \Sigma_{ij}(E) = |\nu_{ij}|^2 \Gamma_j(E) \]  

(21)

is the interaction, self-energy, where \( \Gamma_j(E) \) is given by:

\[ \Gamma_j(E) = \sum_{k} \left\{ \left| \nu_{kj} \right|^2 \right\} \]  

(22)

Where \( k_j = k_L, k_R, k_{Br} \) and \( k_{bq} \).

The transmission amplitude and the transmission probability are respectively defined in the following equations. MATLAB software was designed to run the analysis for our results and discussion.

\[ t(E) = \frac{\tilde{C}_A(E)}{C_D(E)} \]  

(23)

and,

\[ T(E) = |t(E)|^2 \]  

(24)

3. Results and Discussion

3.1. Transmission calculations

Calculations of the transmission spectrum and its properties are the most important step in studying the transport and dynamic properties of the electron mobility mechanism in the field of nanostructures as represented by \[17, 18\]. The eigenvalues of the bases (arranged in homogenous sequence) are calculated by using tight binding model \[15\]:

\[ E_j = E_{\text{basis}} - V_{nm} \cos\left( \frac{n\pi}{N+1} \right) \]  

(25)

Where, \( E_{\text{basis}} \) is the energy of the base, which is given by \( E_G (= -2.63 \text{eV}) \), \( E_C (= -3.75 \text{eV}) \), \( E_A (= -3.25 \text{eV}) \) and \( E_T (= -4.15 \text{eV}) \) \[8\], for the Guanine, Cytosine, Adenine and Thymine respectively. The energy of
the bases was set in the LUMO levels. The coupling interaction between nearest-neighbor bases can be represented by \( V_{nm} \), which can take the values \( V_{nm} = -0.3 \) eV and \(-0.35 \) eV for both (G/C) and (A/T), respectively. The single base is \( j \) and the total number of the bases is \( N \). The energy levels locations in this case set in the range of \( 1.9 \leq E \leq 2.7 \) eV, which set in the LUMO levels. Thus, the density of states is writing as:

\[
\rho(E_j) = \frac{N}{2\pi} \frac{1}{|t| \sin(\frac{E_j}{|t|})},
\]

(26)

The transmission probability \( T(E) \) as a function of number of base pair for the active region, which consists of the DNA bases only, was obtained. Fig (2) and (3) show the transmission results obtained when considering the fishbone model for the two sequences (G/C) and (A/T), respectively. The factors \( V_{Abq} = -0.9 \) eV, \( V_{bqD} = -0.9 \) eV and \( V_{bqBr} = V_{Brbq} = 0.7 \) eV were used to calculate the results. It can be noted that many of the states close to the active region edges has a very low transmission probability. While high transmission probability is shown at locations near energies of base pairs since the coupling between them is turn on. Also, the shapes of the curves are changed with the kind of sequence. This is reasonable because the coupling interactions of the active region with the donor and accepter are equal, but the energy location depends on the type of the sequence. The number of peak-dip transmissions seem to be less than \( N \) due to the degeneracy in some energy levels. Further, it can be noted that there is a remarkable emerge of these resonances, which can be as a result of the interference of coupling interactions. Moreover, the transmission spectrum magnitude and location are different depending on the sequence type.

![Fig (2): Transmission as a function of number of base pairs for sequence (G/C)](image)
3.2. Tunneling current calculations

The transmission spectrum calculations found in the previous section will be used to find the electric current through the DNA molecule. The steady state electric current for the active region can be measured by using the Landauer formula [19]:

$$I = \frac{2e}{h} \int_{-\infty}^{\infty} T(E) [f_L(E) - f_R(E)] dE$$  \hspace{1cm} (27)

$f_\alpha (E)$ is a Fermi distribution function of electrons in the lead $\alpha$, with $\alpha=L, R$, given by:

$$f_\alpha (E) = \left[ 1 + \exp \left( \frac{E - \mu_\alpha}{k_B T_\alpha} \right) \right]^{-1}$$  \hspace{1cm} (28)

Where $\mu_\alpha$ is the chemical potential of the lead $\alpha$, with $\mu_L = \frac{V}{2}$ and $\mu_R = -\frac{V}{2}$, where $V(3eV)$ is the bias voltage. $T_\alpha$ is the temperature of the lead $\alpha$, where $T_L=T_R=T$, fixed at 300K, this means that both leads are in thermal equilibrium state.

The tunneling current properties as a function of number of base pairs for the two sequences (G/C) and (A/T) are shown in the Fig (4) and (5), respectively. The investigation was carried out by changing the number of base pairs from 1 to 50. It can be clearly noticed that the $I$-$N$ curves show a nonlinear dependence. This is due to the overlap of electrons in single molecule is an innovator effect and has some like overlap on dual slits. Also, the two pathways about the molecule display a finite phase variation relying on the aligning of the contacts. The $I$-$N$ dependent is not monotonic relation. Further, we can evidentially distinguish two orders in both figures because of the active region has steady order. The current level and conductivity can be explained by the difference of resistance values. These results show a good agreement with the electronic structure (i.e. the transmission spectrum) of the system.
3.3. Resistance calculations

The resistance can be obtained after calculating the transmission probability using our model derived in the previous sections. It can be obtained as a function of number of base pairs, i.e. the lead temperature, in the thermal equilibrium state by using the following formula [20]:

\[
G = \frac{2e^2}{h} \int_{-\infty}^{\infty} dE \frac{T(E) \delta f(E)}{2E}
\]  \hspace{1cm} (29)

Where \( f(E) \) is defined in eq (28). The thermal equilibrium is considered in our calculation \( T_R = T_L = T \) where \( 0 < T < 300 \) K. The resistance results are presented using fishbone model for the sequence (G/C).

---

Fig (4): Current as a function of number of base pairs for sequence (G/C)

Fig (5): Current as a function of number of base pairs for sequence (A/T)
and (A/T) in Fig (6) and (7), respectively. As shown in Fig (6) for sequence of (G/C), the resistance is dependent on the number of base pairs at 1 to 3 base pairs. After increasing of number of base pairs, the resistance is decreased dramatically. The resistance become nearly constant with increasing number of base pairs at 4 to 50 base pairs approximately for (G/C) sequence. Further, the behavior of the $R-N$ curves shown in Fig (7) for sequence of (A/T) is inverse to the behavior shown in Fig (6) for the sequence (G/C). When increasing of number of base pairs from 1 to 5 base pairs, the resistance is extremely increased. While at 6 to 50 base pairs numbers, the resistance become constant. The cementing of the resistance can result from the hopping transport mechanism between the chemical potentials and active region energy levels. Also, it is well known that at low temperature, the electrons in the molecules have not enough energy to participate of conductance.

**Fig (6): Resistance as a function of number of base pairs for sequence (G/C)**

**Fig (7): Resistance as a function of number of base pairs for sequence (A/T)**
4. Conclusion
An investigation of the influence of number of base pairs on the transmission probability, conductance and current for a DNA system has been carried out in this study. The DNA system with two DNA sequences, (G/C) and (A/T), was derived using a tight-binding model and the steady state formalism. It has been assumed that one active region of DNA molecules can be represented by a fishbone model. It was found that the number of peak-dip transmissions was noticed to be less than N, because of the degeneracy in some energy levels. The transmission spectrum (magnitude and onsite) changes with the type of sequence. Furthermore, the current-dependent curves clearly show a nonlinear dependence. The I-N curves shapes are different in behavior for the two sequences, but the values of current are nearly converged. Moreover, the resistance-dependent of number of base pairs is also give nonlinear dependent with small different in its values. While the curve of R-N does not behavior similarly for both sequences (G/C) and (A/T).

Acknowledgments
We thank Prof. Jenan Majeed AL-Mukh and Prof. Shakier Ibraheem Easa from Basrah University for useful discussions. We also thank Iraqi Ministry of Higher Education and Scientific Research for its support of scientific researches through the Iraqi Virtual Science Library.

References
1. Lakhno, V., DNA nanobioelectronics. International Journal of Quantum Chemistry, 2008. 108(11): p. 1970-1981.
2. Chakraborty, T., Charge migration in DNA: perspectives from physics, chemistry, and biology. 2007: Springer Science & Business Media.
3. Baranovski, S., Charge transport in disordered solids with applications in electronics. Vol. 17. 2006: John Wiley & Sons.
4. Linko, V., A. Ora, and M.A. Kostiainen, DNA nanostructures as smart drug-delivery vehicles and molecular devices. Trends in biotechnology, 2015. 33(10): p. 586-594.
5. Li, Z., et al., Self-assembled DNA nanomaterials with highly programmed structures and functions. Materials Chemistry Frontiers, 2018. 2(3): p. 423-436.
6. Joshi, A., N. Singh, and G. Verma, Preparation and applications of self-assembled natural and synthetic nanostructures, in Fabrication and Self-Assembly of Nanobiomaterials. 2016, Elsevier. p. 29-55.
7. Berlin, Y.A., A.L. Burin, and M.A. Ratner, DNA as a molecular wire. Superlattices and Microstructures, 2000. 28(4): p. 241-252.
8. Chen, Z., et al., DNA metallization: principles, methods, structures, and applications. Chemical Society Reviews, 2018. 47(11): p. 4017-4072.
9. Porath, D., et al., Direct measurement of electrical transport through DNA molecules. Nature, 2000. 403(6770): p. 635-638.
10. Nogues, C., et al., Sequence dependence of charge transport properties of DNA. The Journal of Physical Chemistry B, 2006. 110(18): p. 8910-8913.
11. Xu, B., et al., Direct conductance measurement of single DNA molecules in aqueous solution. Nano letters, 2004. 4(6): p. 1105-1108.
12. Maia Jr, F., et al., Anhydrous crystals of DNA bases are wide gap semiconductors. The Journal of chemical physics, 2011. 134(17): p. 058601.
13. Hjort, M. and S. Stafström, Band resonant tunneling in DNA molecules. Physical review letters, 2001. 87(22): p. 228101.
14. Jo, Y.-S., Y. Lee, and Y. Roh, *Current–voltage characteristics of λ-and poly-DNA*. Materials Science and Engineering: C, 2003. 23(6-8): p. 841-846.

15. Cuniberti, G., et al., *Tight-binding modeling of charge migration in DNA devices*, in *Charge Migration in DNA*. 2007, Springer. p. 1-20.

16. AL-Saidi, S.A., A.A.K. Al-mebir, and M. Hallool, *Characteristics of Thymine Molecule System Behave as Molecular Electronic Device*. Misan Journal of Academic Studies, 2019. 18(35): p. 130-143.

17. Joe, Y.S., S.H. Lee, and E.R. Hedin, *Electron transport through asymmetric DNA molecules*. Physics Letters A, 2010. 374(23): p. 2367-2373.

18. Lee, S.-H., *Electron transport through one and four-channel DNA models*. 2010.

19. Brisker-Klaiman, D. and U. Peskin, *Coherent elastic transport contribution to currents through ordered DNA molecular junctions*. The Journal of Physical Chemistry C, 2010. 114(44): p. 19077-19082.

20. Gutiérrez, R., S. Mandal, and G. Cuniberti, *Dissipative effects in the electronic transport through DNA molecular wires*. Physical Review B, 2005. 71(23): p. 235116.