Journal of Physics: Conference Series **1448** (2020) 012017 doi:10.1088/1742-6596/1448/1/012017

The effects of the contacts on the electronic localization properties of DNA

C J Páez[1], J H Quintero [1], and A C Garcia-Castro[1]

[1] Escuela de Física, Universidad Industrial de Santander, Bucaramanga, Colombia

E-mail: cjpaezg@uis.edu.co

Abstract. In this numerically work, we investigate the electronic transport along model deoxyribonucleic acid molecules using an effective tight-binding approach that includes the backbone onsite energies. The transmission probability, localization length and participation number are examined as a function of system size, energy dependence, and the contact coupling between the leads and the deoxyribonucleic acid molecule. On one hand, the transition from an diffusive regime for short systems to a localized regime is identified, suggesting the necessity of a further length scale revealing the system borders sensibility. On the other hand, we show that the localization length and participation number, do not depended of system size and contact coupling in the thermodynamic limit. Finally we discuss possible length dependent origins for the large discrepancies among experimental results for the electronic transport in deoxyribonucleic acid sample.

1. Introduction

The possibility of metallic like conductivity, made deoxyribonucleic acid (DNA) an interesting candidate for roles nature did not intend for this molecule. In particular, DNA could be an useful template in nanoelectronics. Besides an immediate goal of further miniaturizing silicon based electronics by means of molecular electronics, new concepts, like hybrid organic inorganic structures [1] and self assembling of two dimensional DNA memory devices [2] began to appear in the literature. The initial results of the quest for unraveling the electronic transport properties of DNA [3] caused surprise due to the variety of findings, since DNA appeared to be either an insulator [4], a semiconductor [5, 6] or a metal [7]. These initially controversial results suggested eventually a complex scenario not yet completely understood. The transport properties of DNA may be affected by a variety of effects, either environmental like interaction with substrates, sample dryness, counter ions effects; or intrinsic like nucleotide sequencing [8].

In order to cope this variety of experimental results, numerous theoretical works of electrical transport through DNA has been extensively studied, ranging from molecular dynamics investigations, through ab initio approaches, heuristic models, as well as hybrid methods. In the present work we study a model Hamiltonian for describing charge transport through short homogeneous ds-DNA molecules was proposed by Cuniberti, *et al.*, in which they found the existence of a gap in the current-voltage (I-V) characteristics [6].

Several theoretical models,from tight-binding models [9] until *ab-initio* calculus [10], have been focused on understanding the details of the DNA electronic structures, the nature of electronic states, impurities and how the geometrical factors affect the conduction properties of

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

the DNA in the nanoelectronics devices [10]. Here, we focus on the role of Lead/DNA/Lead coupling strength [11], as an important parameter for studying the electronic transport through the DNA as function of size systems.

2. Numerical model

DNA is a macro-molecule formed by double-helix structure consisting by a sugar and phosphate backbone coupled via hydrogen bonds to repeated bases pairs A-T (T-A) or G-C(C-G) made from four different nucleotides: Adenine (A), Thymine (T), Cytosine (C), and Guanine (G). Many mechanisms have been suggested to explain the charge transport along DNA, where generally π-pathway transfer due to the overlap of molecular orbitals of the stacked aromatic bases of DNA can lead to charge propagation even at long distances [3]. Within this framework, an adequate heuristic model to capture essential electronic properties of such structure can be envisaged within a tight binding model ladder, as extensively reported in the literature [12] and schematically shown in Figure 1. Here rather complex molecular units are replaced by effective orbitals representing both the base nucleotides (attached to leads in Figure 1, which will be described latter), as well as the phosphates and sugars in the backbones.

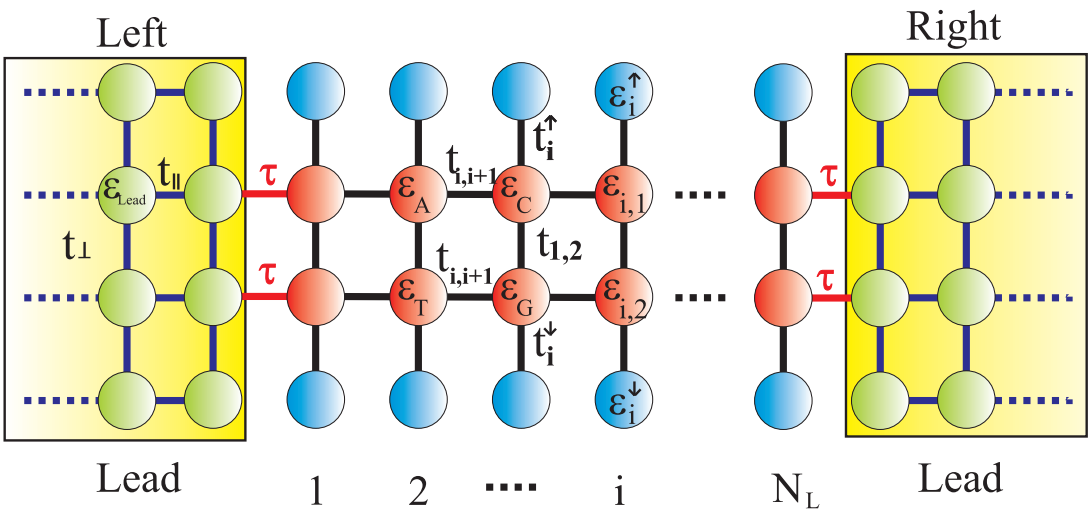

Figure 1. Schematic representation of the ladder model for electronic transport along DNA attached to two semi-infinite electrodes, left (L) and right (R). A nucleobases-pairs sequence are given as red circles representing the four effective nucleotides: A, T, C, and G. Sugar-phosphate (backbone) are given as blue circles and electronic pathways are shown as lines. Throughout this work we will consider the limiting case of a completely random sequencing along the central chains, maintaining the correlation between the base pairs (A-T and C-G).

The electronic properties of the structure depicted in Figure 1 are described by a tight binding Hamiltonian (see Equation (1)).

$$H_{Total} = H_{DNA} + H_{Lead} + H_{Lead-DNA}. \tag{1}$$

Here the total Hamiltonian is separated into three different terms, namely the core of the structure, *i.e.*, the ladder model for a DNA double strand, H_{DNA}, the semi infinite leads, H_{lead}, and the coupling between both sub systems, $H_{lead-DNA}$. The finite ladder structure Hamiltonian is given by Equation (2).

$$
\begin{aligned}
H_{DNA} = \sum_{i=1}^{N_L} \Big[\sum_{j=1}^{2} \varepsilon_{ij} |i,j><i,j| + t_{i,i+1} |i,j><i+1,j| \\
+ \sum_{q=\downarrow,\uparrow} \varepsilon_i^q |i,q><i,q| + t_i^q |i,j><i,q(j)| \\
+ t_{12} |i,1><i,2| \Big] + h.c.
\end{aligned}
\tag{2}
$$

The different terms are better appreciated if referred to Figure 1. $\epsilon_{i,j}$ are the effective orbital energies of the base nucleotides, while ϵ_i^q stand for the backbone molecules. The two DNA strands are connected by the $t_{1,2}$ hopping parameters. Intra strand and backbone to nucleotide couplings are given by $t_{i,i+1}$ and t_i^q, repectively.

Among a great variety of tight-binding based heuristic models, in the present work we will consider heuristic DNA double chain, either with or without backbones. The double chain ladder model used in previous works [13, 14] will be here compared to the more complete description including the backbones [12]. In both cases the base pair sequences considered are completely random.

The choice of the tight binding parameters is far from uniquely determined, being rather a controversial issue, since several parameter sets have being proposed in the literature [15]. The four different values for the effective nucleotide orbital energies, $\varepsilon_A = 8.24$ eV, $\varepsilon_T = 9.14$ eV , $\varepsilon_C = 8.87$ eV and $\varepsilon_G = 7.75$ eV are randomly assigned in one of the strands, with the same probability, while the sites of the second strand are set to follow the DNA pairing (A-T and C-G). In the present phenomenological description, we adopt an early choice [16] which has been recently rather revalidated [8].

In order to reduce the number of model parameters and to simplify our computation we have adopted a simple parameterization taking the intra chain hopping parameters and inter chain coupling as $t_{i,i+1} = 1.0$ eV and $t_{12} = 0.5$ eV, respectively. This set of parameters has been used in several ladder models for DNA [17], respecting the expected condition of $t_{12} < t_{i,i+1}$. The additional hopping onto the backbone is $t_i^q = 0.7$ eV [13] and the backbone onsite energy is taken to be $\varepsilon_i^q = 11$ eV.

Having described the central part of the Hamiltonian, we proceed by briefly sketching the leads and the coupling to them. H_{lead} , where $\varepsilon_{lead} = 8.5$ eV represents on site energy in the electrodes (an average of the previously described orbital energies)and $t_\perp = 0.5$ eV, $t_\parallel = 1$ eV is the nearest neighbor coupling constant inside the electrodes. These parameters represent actually an arbitrary choice for an ideal contact lead, *i.e*, with an energy band width which do not hamper the transmission probability for any relevant energy of the central DNA-like band structure. Finally, the molecule-electrode coupling is given by Equation (3).

$$
H_{Lead-DNA} = \sum_{j=2}^{3} \tau |0,j><1,j| + \tau |N_L,j><N_L+1,j|, \tag{3}
$$

where the first (second) term represents the left (right) metal electrode contact to the DNA molecule. Here $N+1(0)$ is the first layer of the right (left) contact adjacent to the DNA molecule. The value of τ is a key parameter in the present work, since its variation mimics changes in the coupling to the leads of a system to be investigated. It is well known that the quality of the contacts strongly affects the transport properties of molecular devices [8].

The transmission probability $T(E)$ through the central DNA-like molecule is indeed the transport property most important, in this case we calculated by the Landauer-Büttiker equation (see Equation (4)).

$$T(E) = \sum_{v=1}^{4} \sum_{v'=1}^{4} \mid t_{v,v'}(E) \mid^2, \tag{4}$$

where $t_{v,v'}$ represents the transmission amplitude for a particle initially in transverse mode v in the left contact to be scattered into transverse mode v' in the right contact. $t_{v,v'}$ is directly related to the total Green's function of the system that was calculated by means of the recursive approach for a lattice of sites treated in a tight-binding approximation [18,19]. Besides that, the local density of states (ρ_j) at a given site j can be straightforwardly extracted from $\rho = -\frac{1}{\pi} G_{j,j}$, where G_{jj} is the total Green's function at site j.

3. Localization length and participation number

The quantities defined in the previous section, $T(E)$ and ρ_j, are relevant ones in defining the localization of a electronic state in the disordered central DNA-like ladder model.

On one hand, the localization length (LL) for the DNA-like ladder model is computed from the exponential decrease in the transmission probability, [20,21], so, it can defined as Equation (5).

$$LL^{-1}(E) = -\lim_{N_L \to \infty} \frac{1}{2N_L} < \ln \sum_{v,v'}^{4} \mid t_{v,v'}(E) \mid^2 >, \tag{5}$$

where $< \cdots >$ means an average over several hundred different chain configurations. This is done to avoid spurious effects due to a particular configuration. Here N_L is the length of the system given by bases pairs number. Hence, for the sake of completeness in the forthcoming discussion, the total number of effective sites in the system is $4N_L$.

Alternatively, another way to define the localization degree of an electronic state is obtained directly from the wave function, namely, the participation number (PN) is defined by Equation (6) [22].

$$PN(E) = \frac{1}{4N_L \sum_{i=1}^{N_L} \sum_{j=1}^{4} |\Psi_{ij}|^4}, \tag{6}$$

where $\Psi_{jj}^2 = \rho_j$ is the wave-function amplitude in the j site. For localized states, $PR \to 0$ in the limit $NL \to \infty$, while truly localized states - extended peridocially modulated states - exhibit $PR = \frac{2}{3}$ in one dimensional systems [22]. A quantity related to the PR is participation number (PN),here, $PN = N_L \cdot PR$, considering a double strand as depicted in Figure 1. While PR is a measure of a fraction, PN would be a measure of the actual number of bases pairs having appreciable wave-function amplitudes at a given energy.

4. Results and discussion

Since several transport measurements in DNA systems have been undertaken for relatively short strands, one has to raise the question of how the coupling to the contacts - within a manifold of other environmental variables - could affect the transport results. Here we will focus on an heuristic approach, by simply investigating the transmission probability as a function of energy, having the coupling between the DNA-like finite chain to the contacts, τ, as defined in Equation (3). In Figure 2(a) we see that switching the coupling lead to dramatic effects on the transmission probability through a short $50bp$ long chain for the entire energy window. As could be expected, weakening the coupling hampers the transmission probability. It should be noticed that the ordered backbones, considered in Figure 2(a), open a semiconducting gap, absent in the simple

double chain model. Including disorder in the backbones -representing environmental effects, like dryness - will smooth out these threshold states [6].

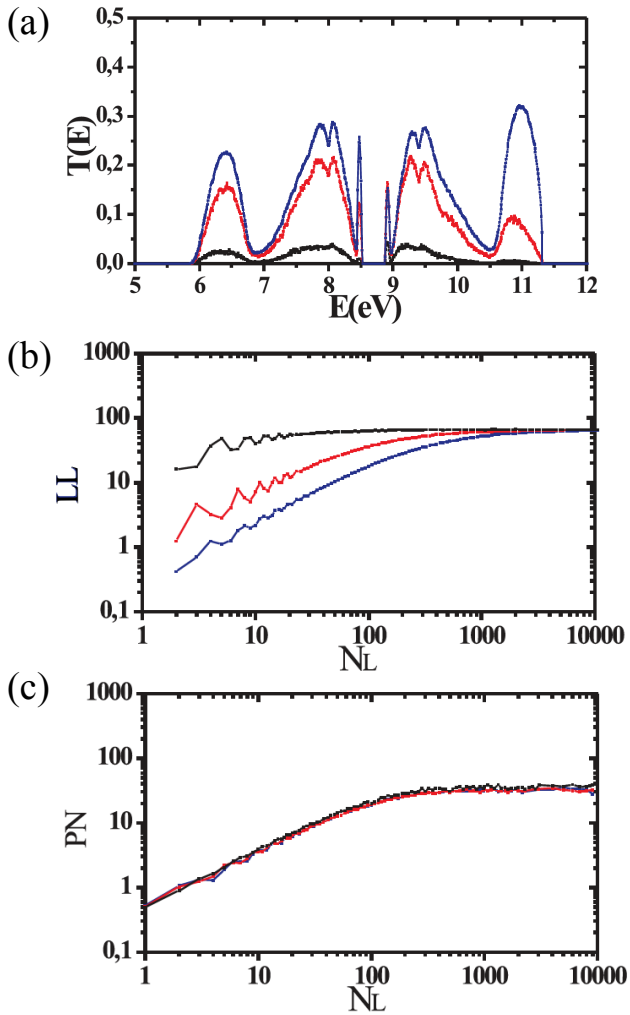

Figure 2. (a) Average transmission as a function of energy for a double-strand DNA of 50*bp* length for different coupling constant between contact and DNA: $\tau = 0.1$ (blue), $\tau = 0.2$ (black) and $\tau = 0.6$ (red). (b) localization length, LL, and (c), participation number, PN, at E = 10.95 eV, as a function of length of DNA-like double strands, with backbones (see Figure 1) length, N_L.

In Figure 2(b) and Figure 2(c) we observe the behavior of both definitions for localization degree considered throughout this work. PN and LL as a function of length of DNA-like double strands, when the coupling τ is varied. We look at weaker, $\tau = 0.2$ eV, and $\tau = 0.6$ eV, couplings that the values for PN are independent of τ, since the participation ratio is not related to the transmission probability. Somehow surprising is the fact that changing the coupling is not relevant, even for very short systems, for which the modification of the probability density at the neighborhood of the contacts could affect the Participation ratio results. On the other hand, the localization length reflects directly the transmission probability, hence the contacts play an important role here. One aspect should be kept in mind, regarding Figure 2(b) and Figure 2(c), namely the coupling strength becomes irrelevant for long systems.

The central idea, concerning this and previous observations, is that a new figure of merit, defining the transition from localized to diffusive regimes, may be given by the length at which PN and LL converge. However, we clearly see that this convergence happens only for energies at which the sates are spread out across both kinds of effective sites: backbone and nucleotide-like, Figure 2(b) and Figure 2(c).

5. Conclusions

The electronic states are more selectively distributed on either backbone or nucleotide-like sites, leading to aforementioned modulations in PR and consequently a not so clear interpretation for the PN as defined here. As previously suggested can be, within the new results, further verified. We observe that, although the collapsing of both quantities to a common value for sufficiently long systems is not always fulfilled, PN and LL indeed saturates at approximately the same lengths for all energies investigated, as illustrated in Figure 2. This common saturation occurs for the present case at $N_L \approx 300$ and should be considered length of localisation for DNA.

Finally, having in mind that experiments suggest that random DNA is not a good conductor over average distances of 300 nucleotides, our results indicate that, indeed, a few hundred base pairs is the transition from diffusive to localized transport regimes. However, DNA samples, a few tents of base pairs long, may also show a diffusive transport regime, but at such small lengths, the strength of the contacts to the leads may severely influence the results. On the other hand, discrepancies between experimental results for longer systems should have an origin in other environmental aspects than the contact to the leads.

References

[1] Johnson R R, Johnson A T C, Klein M L 2008 *Nano Lett.* **8** 69
[2] Yan H, Park S H, Finkelstein G, Reif J H, LaBean T H 2003 *Science* **301(5641)** 1882
[3] Endres R G, Cox D L, Singh R R P 2004 *Rev. Mod. Phys.* **76** 195
[4] Bockrath M, *et al.* 2002 *Nano Lett.* **2(3)** 187
[5] Porath D, Bezryadin A, Vries S, Dekker C 2000 *Nature* **403** 635
[6] Cuniberti G, Craco L, Porath D, Dekker C 2002 *Phys. Rev. B* **65(24)** 241314
[7] Fink H W, Schoenenberger C 1999 *Nature* **398** 407
[8] Shinwari M W, Deen M J, Starikov E B, Cuniberti G 2010 *Adv. Funct. Mater.* **20** 1865
[9] Mehrez H, Anantram M P 2005 *Phys. Rev. B* **71(11)** 115405
[10] Pemmaraju C D, Rungger I, Chen X, Rocha A R, Sanvito S 2010 *Phys. Rev. B* **82** 125426
[11] Zhang W, *et al.* 2010 *Phys. Rev. B* **81(21)** 214202
[12] Klotsa D, Romer R A, Turner M S 2005 *Biophys. J.* **89** 2187
[13] Páez C J, Schulz P A 2013 *Eur. Phys. J. B* **86** 103
[14] Páez C J, Rey-González R, Schulz P A 2010 *Phys. Rev. B* **81(2)** 024203
[15] Hawke L G D, Kalosakas G, Simserides C 2010 *Eur. Phys. J. E* **32** 291305
[16] Roche S, Bicout D, Maci E, Kats E 2010 *Phys. Rev. Lett.* **91** 228101
[17] Diaz E, Sedrakyan A, Sedrakyan E, Dominguez-Adame F 2007 *Phys. Rev. B* **75(1)** 014201
[18] Sols F, Macucci M, Ravaioli U, Hess K 1989 *J. Appl. Phys.* **66** 3892
[19] Lewenkopf C H, Mucciolo E R 2013 *J. Comput. Electron* **12** 203
[20] Kissel G J 1991 *Phys. Rev. A* **44** 1008
[21] Johnston R, Kunz H 1983 *J. Phys. C* **16** 3895
[22] Edwards J T, Thouless D J 1972 *J. Phys. C* **5** 807