Electronic transport in methylated fragments of DNA

M. L. de Almeida,[1] J. I. N. Oliveira,[1] J. X. Lima Neto,[1] C. E. M. Gomes,[1] U. L. Fulco,[1,a)]
E. L. Albuquerque,[1] V. N. Freire,[2] E. W. S. Caetano,[3] F. A. B. F. de Moura,[4] and M. L. Lyra[4]
[1]*Departamento de Biofísica e Farmacologia, Universidade Federal do Rio Grande do Norte, 59072-970 Natal-RN, Brazil*
[2]*Departamento de Física, Universidade Federal do Ceará, 60455-760 Fortaleza, CE, Brazil*
[3]*Instituto Federal de Educação, Ciência e Tecnologia do Ceará, 60040-531 Fortaleza, CE, Brazil*
[4]*Instituto de Física, Universidade Federal de Alagoas, 57072-900 Maceió-AL, Brazil*

(Received 18 September 2015; accepted 6 November 2015; published online 17 November 2015)

We investigate the electronic transport properties of methylated deoxyribonucleic-acid (DNA) strands, a biological system in which methyl groups are added to DNA (a major epigenetic modification in gene expression), sandwiched between two metallic platinum electrodes. Our theoretical simulations apply an effective Hamiltonian based on a tight-binding model to obtain current-voltage curves related to the non-methylated/methylated DNA strands. The results suggest potential applications in the development of novel biosensors for molecular diagnostics. © *2015 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4936099]

Recently, there has been significant interest in the characterization of electronic processes in biomolecules, with a consequent impact in the development of medical technologies and therapies.[1–4] For instance, investigations of the compatibility of the interface between electronic devices and biological systems are nowadays much encouraged, since the quality of bio-interfaces can play a central role in the prospective success and impact of technologies such as thin film electronics, *in vitro* cell culture models, and medical devices that make use of organic materials in place of conventional semiconductors.[5,6]

Natural semiconductors can serve as templates for the design of synthetic electronic components employing biologically derived nanomaterials, such as quantum dots and nanotubes built up from important classes of biomolecules such as proteins and nucleic acids.[7] Their main advantages are their lower current and power operation, leading to cheaper and simpler devices, while their main disadvantage is their low life-time due to degradation.[8–10]

Among these biological systems, the deoxyribonucleic-acid (DNA) molecule prevails, not only because it carries the genetic instruction used for the functioning and reproduction of all cells in living organisms, but also due to its diversity, versatility, and amenability as part of a nanoelectronic device (for a recent review, see Ref. 11).

A gene is a functional unit of heredity, being formed from DNA segments that represent instructions for protein synthesis in a living organism. After the Human Genome Project, we have a nearly complete list of the genes needed to produce a human being.[12] However, gene expression is a very complex process, relying on a secondary system used by cells to determine when and where a particular gene information will be effectively used during development. This system is overlaid on DNA in the form of epigenetic mechanisms, which refer to heritable and functionally relevant changes in gene activity, but that do not modify DNA.[13] Among them, DNA methylation plays an important role, being a process by which methyl groups (CH_3) are added to the fifth carbon atom of the cytosine base or the sixth nitrogen atom of the adenine base. It is a well-characterized epigenetic signaling tool, typically acting to suppress gene transcription.

Usually, genes are silenced when methylation occurs in their promoter region, where the transcription process is initiated; however, when the methylation occurs in the gene body, they may have a positive correlation to their expression (transcription/translation) process.[14] For example, the 5-methylcytosine-based DNA methylation occurs through the covalent addition of a methyl group at the 5-carbon of the cytosine heterocyclic aromatic ring by an enzyme called DNA methyltransferase resulting in the formation of 5-methylcytosine.[15] This is a typical vertebrate DNA methylation pattern, which occurs predominantly in CpG islands, dinucleotide-rich regions that possess high relative densities of CpG and are positioned at the 5′ ends of many human genes. Most cell types display relatively stable DNA methylation patterns, with 70%–80% of all CpGs being methylated.[15] This event is associated with a number of key processes including embryonic development, chromosome stability, genomic imprinting, X chromosome inactivation, suppression of repetitive elements, and carcinogenesis.[16,17]

It is the aim of this work to suggest a nanoelectronic device able to investigate the electron transport properties of a DNA strand taking into account the contributions of the 5-methylcytosine. Our main purpose is to observe how methylation of cytosine in a DNA strand modifies its charge transport profiles in comparison to a non-methylated DNA strand. Both molecules are set in a linear geometry covalently linked to two platinum electrodes, the source—*S* and drain—*D* contacts (see Fig. 1).

We model the non-methylated/methylated DNA strands adopting a tight-binding Hamiltonian constructed from *ab initio* parameters with a single orbital per site, nearest-neighbor interactions, and neglecting any environment or complex contacts. Besides, the effects of interatomic matrix

a)Author to whom correspondence should be addressed. Electronic mail: umbertofulco@gmail.com. Tel.: +-55-84-3215-3419.

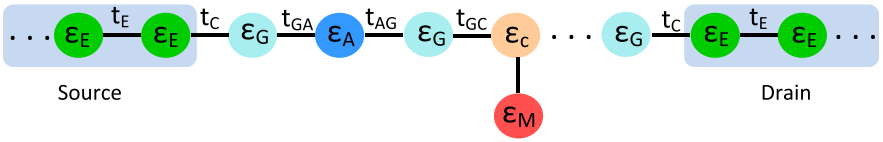

FIG. 1. Schematic representation of the methylated DNA strand sandwiched between two electrodes. We show in red the methylation of the cytosine nitrogen base, the 5-methylcytosine.

elements are replaced by intramolecular hopping parameters obtained from first principles calculations as well, as described in previous works[18,19]

$$H_{total} = H_{DNAm} + H_{electrode} + H_{coupling}. \quad (1)$$

The first term of Eq. (1) describes the intra-strand charge propagation through the methylated DNA, being given by

$$H_{DNAm} = \sum_{n=1}^{N} [\varepsilon_{\alpha}^{n}|n,1\rangle\langle n,1| + t_{n,n\pm1}|n,1\rangle\langle n\pm1,1|], \quad (2)$$

where N is the number of nucleotides in the methylated DNA strand, and ε_{α}^{n} is the on-site ionization potential of the respective base α at the nth site. Here, $\alpha =$ adenine (A), cytosine (C), guanine (G), thymine (T), and 5-methylcytosine (M). Also, $t_{n,n\pm1}$ are the hopping parameters between the adjacent sites n, $n \pm 1$ of the methylated DNA strand.

The second term, related to the two semi-infinite metallic electrodes (S and D contacts), is given by

$$H_{electrode} = \sum_{n=-\infty}^{0} [\varepsilon_{E}^{n}|n,1\rangle\langle n,1| + t_{E}|n,1\rangle\langle n\pm1,1|] + \sum_{n=N+1}^{\infty} [\varepsilon_{E}^{n}|n,1\rangle\langle n,1| + t_{E}|n,1\rangle\langle n\pm1,1|]. \quad (3)$$

Here, ε_{E}^{n} (t_{E}) is the ionization energy (hopping parameter) of the electrode. We consider a platinum electrode whose ionization energy $\varepsilon_{E} = 5.36$ eV is related to the work function of this metal, with hopping integral $t_{E} = 12.0$ eV.[20]

Finally, the third term describes the coupling between the methylated DNA strand and the semi-infinite metallic electrodes, yielding

$$H_{coupling} = t_{c}[|0,1\rangle\langle 1,1| + |N,1\rangle\langle N+1,1|], \quad (4)$$

where $t_{c} = 0.63$ eV is the hopping amplitude between the source (drain) electrodes and the ends of the methylated DNA strand.

Considering the tight-binding Hamiltonian given above, one can evaluate the I-V characteristics by applying the Landauer-Büttiker formulation,[21,22] i.e.,

$$I(V) = \frac{2e}{h}\int_{-\infty}^{+\infty} T_{N}(E)[f_{S}(E) - f_{D}(E)]dE, \quad (5)$$

where $f_{S(D)}$ is the Fermi-Dirac distribution given by

$$f_{S(D)} = [\exp[(E - \mu_{S(D)})/k_{B}T] + 1]^{-1}. \quad (6)$$

Here, $\mu_{S(D)}$ is the electrochemical potential of the two electrodes fixed by the applied bias voltage V as $|\mu_{S} - \mu_{D}| =$ eV. One can note that the current onset may be crucially dependent on the chemical potential of the leads. In view of that, for the sake of simplicity, the chemical potential of the whole system is taken to be zero before the bias voltage V is applied.

The vertical ionization energy and hopping parameters for methylated cytosine and the base pairs, respectively, were obtained from first principles calculations using the Gaussian 09 code within the Density Functional Theory (DFT) framework. For these calculations, we used the Becke's half-and-half hybrid exchange-correlation (BHandHLYP) functional[23] and a Dunning's correlation consistent polarized valence double ζ basis set (cc-pVDZ)[24] to optimize the structures of the base pairs, as well as to calculate their occupied highest occupied molecular orbital (HOMO) energies. The BHandHLYP method is strongly recommended for the study of small peptides and other similar biomolecular systems with hydrogen bonding and charge transfer interactions,[25–27] especially when the objective is to calculate HOMO eigenvalues,[28] as this hybrid functional includes a fraction of the exact orbital exchange.[29]

Most biomolecular interactions take place in an aqueous environment being, therefore, important to consider the effects of the aqueous solvent. In particular, the DNA molecule displays considerable sensitivity to ionic surroundings during its various structural transitions and charge-transfer states, mainly in the base-pairing effect of its double-stranded (native) topology.[30] However, in its single-stranded (denaturated) form, as considered here, its structure and intermolecular interactions are much less affected with a minimal exposure of its hydrophobic groups to the aqueous solvent.[31] In view of that, and to avoid the increase of the computational cost, all calculations were carried out only for molecules in the gaseous phase (no solvation effects were took into account). We used the relation $IE = E(N-1) - E(N)$ to evaluate the ionization potential for 5-methylcytosine to be 7.02 eV, where the ionized [non-ionized] nucleotide with one missing electron was characterized by the total energy $E(N-1)$ [$E(N)$].[32] Taking advantage of the published *ab initio* ionization potential calculation of stacked bases, the values for the other nucleotides are as follows: guanine, 7.75 eV; cytosine 8.87 eV; adenine 8.24 eV; and thymine 9.14 eV.[33]

The hopping parameters are estimated from Ref. 34

$$t_{n,n\pm1} = (1/2)[E_{n,n\pm1}^{HOMO} - E_{n,n\pm1}^{HOMO-1}], \quad (7)$$

where $E_{n,n\pm1}^{HOMO}$ and $E_{n,n\pm1}^{HOMO-1}$ are, respectively, the first and second highest occupied molecular orbital energies for the base pairs formed by n and $n \pm 1$ residues. The hopping parameters between adjacent bases in the methylated DNA strand are listed in Table I.

In order to perform a comparative analysis of the electronic conductance of the standard and methylated DNA

TABLE I. Hopping parameters between adjacent bases in a DNA strand. *X* (column), *Y* (row) represent the location of each base relative to its neighbour in the DNA strand. All energies are expressed in electron volts (eV).

X|*Y*	*A*	*T*	*C*	*G*	*M*
A	0.260	0.179	0.223	0.220	0.227
T	0.173	0.278	0.167	0.245	0.225
C	0.163	0.190	0.198	0.282	0.248
G	0.252	0.221	0.144	0.210	0.201
M	0.369	0.450	0.507	0.145	0.430

forms, current-voltage characteristics of the DNA strands with 16 bases are shown in Fig. 2. When the potential barrier between the metallic contacts and the DNA is set to zero, a steplike feature in the plot I-V is found, roughly depicting its semiconducting characteristics.[35,36] The inset illustrates the trans-conductance dI/dV versus V of the devices, which is highly nonlinear. The approximately linear I-V curves observed for $5 < |V| < 10\,\mathrm{V}$ indicate a semiconductor-metallic transition (SMT), allowing for the possibility to interpret the variation of experimental results found in the standard DNA form. The zero-conductance plateau (current slope) in the resulting I-V characteristics in the region $|V| < 5\,\mathrm{V}\,(5 < |V| < 10\,\mathrm{V})$ is due to the gap (band) width of the electronic band structures of the DNA molecule. From the I-V data, no rectifying behavior was observed indicating that the charge transport features in HOMO states are similar to those found in the lowest unoccupied molecular orbitals (LUMO).

Fig. 2(a) deals with the DNA sequence GAGCTGAC GTTCACGG retrieved from the first sequenced human chromosome 22 (Ch22) entitled NT_{011520}, which contains the basic elements relevant for the analysis of the methylation effect on the electronic transport. Figs. 2(b)–2(d) have some methylated forms, differing in amount of 5-methylcytosine, characterized by one (Fig. 2(b)), two (Fig. 2(c)), and three (Fig. 2(d)) methylated cytosines, respectively. Amazingly, one can see that even a single methylation site in a DNA strand (Fig. 2(b)) reduces the current-voltage curve by one order of magnitude, being a very convenient and cheap way to probe their particular characteristics.

The overall reduction of the saturation current by the presence of methylated cytosine sites is directly associated with the fact that these act as additional impurity centers, thus enhancing the Anderson localization of the electronic states around their closest vicinity. The sensitivity of the saturation current on the position of the methylated cytosine is related to a secondary phenomenon, namely, the impact of cytosine methylation on the hopping amplitudes to the neighboring nucleotides. The data shown in Table I reveal that the hopping amplitude connecting a cytosine to a guanine is reduced upon methylation. In contrast, the hopping amplitudes between cytosine and adenine as well as between cytosine and thymine are enhanced. Hence, electronic transport has an additional suppression when methylation occurs in a cytosine directly connected to a guanine because this leads to a smaller average hopping amplitude. This feature is clearly illustrated in Fig. 2(b) which shows that the saturation current is much smaller when methylation occurs in the cytosines located at position 1, 2, and 4, which are directly connected to a guanine base, as compared to the corresponding saturation current when methylation occurs at the cytosine at position 3, which is connected to adenine and thymine. When methylation takes place at two cytosines, the suppression of the saturation current is smaller when the

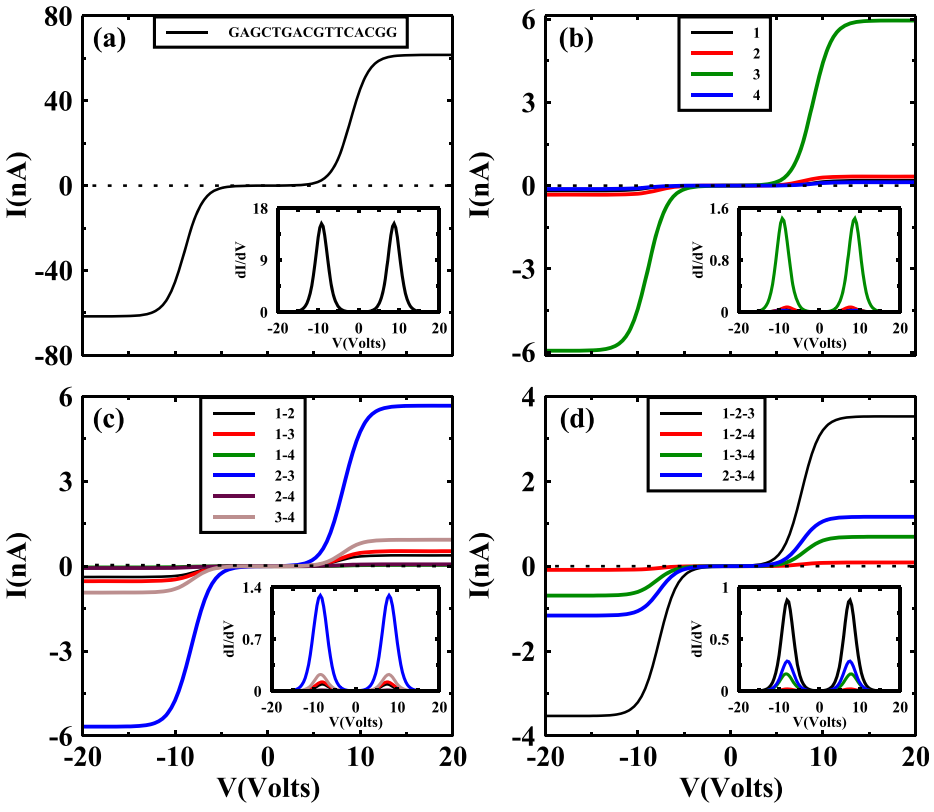

FIG. 2. Current-voltage I (in nA) against V (in volts) curves for a non-methylated/methylated DNA strands. The insets show the transconductance (dI/dV) versus V of the devices. (a) GA GCTGACGTTCACGG chain with four (1, 2, 3, 4, in sequence) non-methylated cytosine bases. (b) Methylation of only one cytosine in the DNA strand (1, 2, 3, or 4). Observe the overlap of the curves 1 (black), 2 (red), and 4 (blue) near the horizontal axis. (c) Methylation of two cytosines in the DNA strand (pairs 1-2, 1-3, 1-4, 2-3, 2-4, and 3-4). (d) Methylation of three cytosines in the DNA strand (labeled by the triplets 1-2-3, 1-2-4, 1-3-4, and 2-3-4).

second methylation site corresponds to a cytosine that, although having one neighboring guanine, is connected to an adenine. This is so because it has the largest increase in the hopping amplitude. Note, however, that cytosines located at positions 2 and 4 have similar neighborhoods (one guanine and one adenine). Fig. 2(c) shows that the saturation current is larger when the second methylated site (besides the one at position 3) is at position 2, corresponding to well separated scattering centers. Finally, when methylation occurs at all cytosine sites except by one, the non-methylated base acts as a defect. As such, the suppression of the saturation current will be larger when the defect (non-methylation) is in cytosine 3, which is the only one that has both hopping amplitudes to the neighboring nucleotides enhanced by methylation, as illustrated in Fig. 2(d). The above physical mechanisms influencing the electronic transport are expected to hold for quite general methylated DNA sequences. Note also the opposite change in the current when one goes from the methylated sites 1–2 to the methylated sites 1-2-3 (increasing of current) as compared to the change when site 4 is methylated from the methylated sites 2-3 (decreasing of current). This feature is related to the distinct influence produced by methylation on the hopping amplitudes of neighboring sites. In fact, the methylation of site 3 enhances the hopping amplitude to both nearest nucleotides, namely, adenine and thymine (see Table I), thus favoring charge transport. On the other hand, the methylation of site 4 enhances (reduces) the hopping to the adenine (guanine) site, leading to a larger barrier for the charge transport.

In conclusion, we have investigated in this paper the I-V characteristics of standard and methylated DNA forms in order to identify and describe the impact of methylation on charge transport. For a single methylation defect, the saturation current is strongly suppressed when cytosine is connected to guanine. If two methylation defects are allowed in a strand, current suppression is smaller, while very specific I-V curves were found for strands with three methylation defects. This strongly suggests the feasibility of using I-V curve measurements to develop biosensors for the diagnosis of human diseases related to aberrant gene expression caused by DNA methylation.

This work was partially financed by the Brazilian Research Agencies CAPES (PNPD) and CNPq (Casadinho/Procad). E.W.S.C. received financial support from CNPq Project 307843/2013-0.

[1]E. Werner, M. Reiter-Schad, T. Ambjörnsson, and B. Mehlig, Phys. Rev. E **91**, 060702 (2015).
[2]C. Liu, K. Kim, and D. L. Fan, Appl. Phys. Lett. **105**, 083123 (2014).
[3]S. Kim, D. Baek, J. Y. Kim, S. J. Choi, M. L. Seol, and Y. K. Choi, Appl. Phys. Lett. **101**, 073703 (2012).
[4]F. Shao and J. K. Barton, J. Am. Chem. Soc. **129**, 14733 (2007).
[5]A. Reinhardt and D. Frenkel, Phys. Rev. Lett. **112**, 238103 (2014).
[6]S. Angioletti-Uberti, P. Varilly, B. M. Mognetti, and D. Frenkel, Phys. Rev. Lett. **113**, 128303 (2014).
[7]H. Lee, J. Park, J. Kim, H. Jung, and T. Kawai, Appl. Phys. Lett. **89**, 113901 (2006).
[8]M. W. Shinwari, M. J. Deen, E. B. Starikov, and G. Cuniberti, Adv. Funct. Mater. **20**, 1865 (2010).
[9]R. G. Endres, D. L. Cox, and R. R. P. Singh, Rev. Mod. Phys. **76**, 195 (2004).
[10]D. Porath, G. Cuniberti, and R. Di Felice, Top. Curr. Chem. **237**, 183 (2004).
[11]E. L. Albuquerque, U. L. Fulco, V. N. Freire, E. W. S. Caetano, M. L. Lyra, and F. A. B. F. de Moura, Phys. Rep. **535**, 139 (2014).
[12]M. V. Olson, Proc. Natl. Acad. Sci. **90**, 4338 (1993).
[13]S. B. Baylin, Nat. Rev. Clin. Oncol. **2**, S4 (2005).
[14]A. Sadakierska-Chudy, R. M. Kostrzewa, and M. Filip, Neurotoxic. Res. **27**, 84 (2015).
[15]A. Bird, Genes Dev. **16**, 6 (2002).
[16]K. D. Robertson, Nat. Rev. Genet. **6**, 597 (2005).
[17]Z. D. Smith and A. Meissner, Nat. Rev. Genet. **14**, 204 (2013).
[18]L. M. Bezerril, U. L. Fulco, J. I. N. Oliveira, G. Corso, E. L. Albuquerque, V. N. Freire, and E. W. S. Caetano, Appl. Phys. Lett. **98**, 053702 (2011).
[19]J. I. N. Oliveira, E. L. Albuquerque, U. L. Fulco, P. W. Mauriz, and R. G. Sarmento, Chem. Phys. Lett. **612**, 14 (2014).
[20]Y. A. Berlin, A. L. Burin, and M. A. Ratner, Chem. Phys. **275**, 61 (2002).
[21]R. Landauer, IBM J. Res. Dev. **1**, 223 (1957).
[22]M. Buttiker, Phys. Rev. B **35**, 4123 (1987).
[23]A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
[24]T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[25]W. Yu, L. Liang, Z. Lin, S. Ling, M. Haranczyk, and M. Gutowski, J. Comput. Chem. **30**, 589 (2009).
[26]J. Csontos, N. Y. Palermo, R. F. Murphy, and S. Lovas, J. Comput. Chem. **29**, 1344 (2008).
[27]M. S. Liao, Y. Lu, and S. Scheiner, J. Comput. Chem. **24**, 623 (2003).
[28]G. Zhang and C. B. Musgrave, J. Phys. Chem. A **111**, 1554 (2007).
[29]D. J. Tozer and F. de Proft, J. Phys. Chem. A **109**, 8923 (2005).
[30]H. Yin, Y. Ma, J. Mu, C. Liu, and M. Rohlfing, Phys. Rev. Lett. **112**, 228301 (2014).
[31]D. X. Macedo, I. Guedes, and E. L. Albuquerque, Physica A **404**, 234 (2014).
[32]R. Maul, M. Preuss, F. Ortmann, K. Hannewald, and F. Bechstedt, J. Phys. Chem. A **111**, 4370 (2007).
[33]H. Sugiyama and I. Saito, J. Am. Chem. Soc. **118**, 7063 (1996).
[34]K. Dedachi, T. Natsume, T. Nakatsu, S. Tanaka, Y. Ishikawa, and N. Kurita, Chem. Phys. Lett. **436**, 244 (2007).
[35]A. V. Malyshev, Phys. Rev. Lett. **98**, 096801 (2007).
[36]M. S. Xu, S. Tsukamoto, S. Ishida, M. Kitamura, Y. Arakawa, R. G. Endres, and M. Shimoda, Appl. Phys. Lett. **87**, 083902 (2005).