Topological Line Defects around Graphene Nanopores for DNA Sequencing†

Jariyanee Prasongkit,a Ernane F. Martins,c,d Fábio A. L. de Souza,e Wanderlã L. Scopel,f Rodrigo G. Amorim,g Vittaya Amornkitbamrung,b,h Alexandre R. Rocha,c and Ralph H. Scheicherd

Topological line defects in graphene represent an ideal way to produce highly controlled structures with reduced dimensionality that can be used in electronic devices. In this work we propose using extended line defects in graphene to improve nucleobase selectivity in nanopore-based DNA sequencing devices. We use a combination of QM/MM and non-equilibrium Green’s functions methods to investigate the conductance modulation, fully accounting for solvent effects. By sampling over a large number of different orientations generated from molecular dynamics simulations, we theoretically demonstrate that distinguishing between the four nucleobases using line defects in a graphene-based electronic device appears possible. The changes in conductance are associated with transport across specific molecular states near the Fermi level and their coupling to the pore. Through the application of a specifically tuned gate voltage, such a device would be able to discriminate the four types of nucleobases more reliably than that of graphene sensors without topological line defects.

1 Introduction

The development of DNA sequencing tools has advanced in strides over the past decades.1,2 Nonetheless the cost for sequencing a genome is yet too expensive for widespread use in personalized medicine3–5. Nanopores have been heralded as a possible way of providing low-cost whole-genome sequencing by measuring either ionic currents as DNA translocates through a nanopore on a membrane,6–8 or transverse electronic currents across the membrane itself9,10. A number of materials have been proposed as scaffolds for this type of molecular recognition, including lipid membranes,11 and more recently two dimensional (2D) solid-state materials such as dichalcogenides,12,13 and specially graphene nanopores.14–20 While 2D materials represent the best possible surface to volume ratios - a key ingredient in detection - single molecule recognition remains challenging, specially if one is looking for all-electronic methods.21–24 Graphene combines the feature of being the thinnest possible membrane with very strong carbon-carbon bonds.25,26 However, low signal-to-noise ratios in nanopore experiments have hindered its use in DNA sequencing.27 As an alternative, extended line defects, formed by a combination of octagons and pentagons (OP), have been recently synthesized.28,29 These defects have been predicted to act as conducting quasi-one-dimensional wires embedded in the graphene matrix.28,30 Indeed, reducing the dimensionality is a possible path towards improving sensitivity. An added advantage of this arrangement is that the overall stability of graphene is maintained while allowing for atomically precise edges.

In this work we introduce, extended line defects in graphene to improve the performance of graphene-based DNA sensors. We use a combination of quantum mechanics/molecular mechanics (QM/MM) hybrid methods, and non-equilibrium Green’s functions (NEGF)31–34 to investigate the electronic conductance modulation along extended line defects connecting a nanopore explicitly including the effects of the solvent21,22. By sampling over a
large number of different orientations, we theoretically demonstrate that distinction between four nucleobases can be achieved. We further show that the changes in transport are associated with electron flow across specific molecular states near the Fermi level, and their respective coupling to the pore.

2 Computational Methods

Fig. 1 a) Schematic setup for a DNA sequencing device constructed by embedding line defects around a nanopore in graphene connected to a source and drain (in yellow color), with an additional gate electrode (in blue color) below. The transverse conductance would be measured as a DNA strand passes through the nanopore. The different conductance signals for each nucleobase provide a time-resolved signature for the DNA sequence. b) The device setup used for our first-principles quantum transport simulation consists of a scattering region and two semi-infinite leads. In the scattering region, a nanopore is created between two line defects in graphene, whose edge carbon atoms are passivated by H, and the nucleotides are inserted into the pore. When a solvent is present we also include a buffer layer at the interface between the leads and the scattering region, where the solvent potential drops to zero.

The prototypical sequencing device is conceived as shown in Figure 1a. It consists of a graphene sheet containing two parallel OP line defects extending along the entire system, and a hydrogen-saturated nanopore with a diameter of ~12 Å. The graphene is periodic in the x-direction, while in the z-direction, it is attached to semi-infinite periodic structures (one unit cell of electrodes is shown in blue in Figure 1b). For our calculations, we consider a single nucleotide at a time inside the pore, as we have already shown that the effects of adjacent bases on sensing is small. The four nucleotides; deoxyadenosine monophosphate (dAMP), deoxythymidine monophosphate (dTMP), deoxyguanosine monophosphate (dGMP), deoxycytidine monophosphate (dCMP) are abbreviated simply as A, T, C and G, respectively.

First we performed a full dynamical treatment for the four nucleotides (and also the empty pore) in solution (water and 0.2 M of Na and Cl) using classical molecular dynamics. These were carried out in two steps: an initial NPT simulation (Berendsen barostat and Nose-Hoover thermostat) to equilibrate the pressure, followed by an NVT production run simulation (with the same thermostat) to obtain the structures. Both MD simulations were carried out using a time-step of 2 fs to generate 20 ns of trajectories. The whole graphene sheet was kept fixed during the simulation and a harmonic potential was applied to a central atom in the nucleotide to keep it inside the pore. All MD calculations were carried out using GROMACS, with the AMBER99SB force field and single-point-charge (SPC) water model.

From the NVT simulation we extracted 50 snapshots equally separated in time. Typical configurations for each of the nucleobases in the graphene nanopore are shown in Figure 2. We then divided the system into a QM and an MM region to perform the QM/MM simulation. In our case the QM part includes graphene, the nucleotide and one Na atom to equilibrate the charge. All water molecules and the remaining Na and Cl ions were chosen as the MM part.

The QM part of the calculation is performed within DFT as implemented in SIESTA, modified to include the MM contribution as a static external potential. We employed the generalized gradient approximation (GGA) to the exchange-correlation functional, and valence electrons are described using a local basis set with a single-ζ plus polarization orbital (SZP) for all elements, which has already been proven a valid approach for the nucleobase-graphene system. The atomic core electrons are modeled using Troullier-Martins norm-conserving pseudopotentials. For the k-point mesh generation, 5 × 1 × 5 k-points were employed and the mesh cutoff energy was 150 Ry.

The DFT calculation is combined with the NEGF technique to determine the electronic transport. The system is divided into three regions, namely a scattering region and two semi-infinite electrodes as shown in Figure 1. We self-consistently calculate the charge density using the Green’s function for the
scattering region, including the effects of the electrodes via self-energies and assuming the Kohn-Sham Hamiltonian as a single-particle Hamiltonian for our problem, as implemented in the SMEAGOL package\textsuperscript{33,34}. In our particular case the Hamiltonian incorporates the effects of the solvent from our QM/MM calculation. The exchange-correlation functional and basis sets employed in the transport calculation are identical to those described above for the electronic structure calculations. We considered only the $\Gamma$-point for Brillouin zone sampling. Within the NEGF approach, the quantity we are interested in is the transmission coefficient,

$$T(E) = \text{Tr}\left[\Gamma_R(E)G_R(E)\Gamma_L(E)G_A(E)\right],$$

(1)

where $\Gamma_{L,R}(E)$ is the broadening matrix of the left (right) electrode, and $G_{R,A}(E)$ is the retarded (advanced) Green’s function. Then, we simply evaluate the conductance, $G = G_0 T(E_F)$, where $G_0 = 2e^2/h$ is the quantum conductance, $e$ is the charge of the electron and $h$ is Planck’s constant. A more detailed description of the methodology is presented elsewhere\textsuperscript{31,34,41} including the QM/MM transport method\textsuperscript{21,22}.

### 3 Results

Figure 3 shows the transmission averaged over 50 electronic transport calculations for uncorrelated structural configurations of each system. We have also performed calculations for an empty pore in a solvent (see Supplementary Information). For energies close to the Fermi energy, the transmission for the empty pore in liquid is smaller than the transmission for the nucleotides. In fact, we notice that the effect of introducing any of the basis into the pore is an overall shift of the transmission due to the interaction with the nucleobase. As a result the sensitivity of the pore to any nucleobase is expected to be very high. Moreover, we can also see that the four nucleotides present different average transport properties for energies slightly above the Fermi energy (up to 0.1 eV).

In order to understand the origin of the changes in conductance, and to determine trends in transport, and identify potential descriptors, we performed DFT calculations of the isolated nucleobases in the graphene nanopore. All nucleobases were initially positioned so as to lie in the plane of the graphene. After full relaxation, the interaction between nucleobases and line defects induce a wavy structure on the graphene sheet. A typical final arrangement for a base within the pore is presented in Figure 4a for the case of guanine. Note that the sugar-phosphate backbone, counterions and water molecules were excluded from the quantum transport calculations\textsuperscript{46}. Fig 4b shows favorable binding configurations of four nucleobases and identifies bond lengths between the bases and graphene, which are smaller than 2.5 Å. The binding energy $E_b$ between the nanopore and nucleobase was calculated as

$$E_b = (E_{\text{nanopore}} + E_{\text{base}}) - E_{\text{total}},$$

(2)
Fig. 4 a) The fully optimized structure of the device setup used in the DFT calculations. The wavy structure of graphene is caused by binding between the base and the line defects. b) The most stable configurations for the four nucleobases G, A, C, T. c) Charge density differences (isovalue 0.002 e/Å³ in the y-z graphene plane) of G, A, C, T bound to one side of the line-defect graphene. These binding sites yield the largest binding energies. Blue and red colors represent charge depletion and accumulation, respectively.

where $E_{\text{total}}$ is the total energy of the system (graphene nanopore+base), and $E_{\text{nanopore}}$ and $E_{\text{base}}$ is the energy of graphene nanopore and nucleobase, respectively. The calculated binding energies ($E_b$) of the most favorable configurations shown in Figure 4 are found to be 0.30, 0.33, 0.41 and 0.23 eV for G, A, C and T, respectively.

These binding energies are due to the charge-transfer interactions between the non-bonding electrons (oxygen, nitrogen) and hydrogen. The interaction between oxygen and hydrogen bond is stronger than that of nitrogen and hydrogen since oxygen has a higher electronegativity. However, the binding energy of A is slightly higher than that of G, since A uses its non-bonding electrons (N) to interact with 2 hydrogen atoms. For C, the high binding energy is due to two interacting sites forming H-bonds (O-H and N-H). Finally, compared to G, T has a smaller electronic density around the oxygen atom bound to the graphene due to the presence of a second oxygen, therefore, it has lower binding energy.

We have mapped the total charge density redistribution due to the interaction between device and nucleobase by comparing total charge of the system with its isolated parts: nucleobase and device, using the expression: $\Delta \rho(\vec{r}) = \rho_{\text{device+base}}(\vec{r}) - (\rho_{\text{base}}(\vec{r}) + \rho_{\text{device}}(\vec{r}))$. The charge density difference plotted in Figure 4c reveals that charge fluctuates mostly around the line defect close to the bases. This charge reorganization causes the modulation of charge transport in extended line defects. In Figure 5, we present the zero-bias transmission coefficients as a function of energy responsible for the most favorable configuration shown in Figure 4b. By looking at the transmission coefficients of the empty pore, there are three resonance peaks lying close to the Fermi energy for an energy range $E - E_F \in [0.0, 0.1]$ eV. Peaks 2 and 3 are sharper and taller since both are responsible for the pathway of the wavefunction flowing through the extended line defect (see Supplementary Information of electrodes for details).

We have rotated the transmission coefficients as a function of energy responsible for the most favorable configuration shown in Figure 4c, as compared to that in the absence of nucleobase.

By rotating in the y-z and x-z plane (see Figure 6a), the transport properties can be modulated. Note that each structure is fully relaxed again by using DFT for the final configurations. It is seen that the specific position of oxygen with respect to line defects strongly affects the transmission peak height; the stronger the interaction between the base and graphene, the greater the reduction in peak height. Since the pronounced distinction in $T(E)$ is observed at $E - E_F \sim 0.08$ eV in which the current will flow predominantly along the line defect, we assume a positive gate voltage applied to the device that shifts the Fermi level accordingly.

Fig. 5 Zero-bias transmission coefficients as a function of energy responsible for the most favorable configuration shown in Figure 4c, as compared to that in the absence of nucleobase.
The sensitivity is assessed by the change in conductance; $S = \frac{(G - G_0)}{G_0} \times 100\%$, where $G$ and $G_0$ are the conductance of the nanopore+base and the empty nanopore, respectively. The sensitivity variations for four nucleobases are presented in Figure 6b. The maximum conductance follows the hierarchy: $C > A > G > T$. The change in tone, from dark to light, indicates high to low sensitivity for each nucleobase. In Figure 6c, we observe a linear relationship between the sensitivity and the binding energy, confirming that the interaction strength significantly affects the sensitivity$^{46}$. In the experiment, manipulating the translocation speed is critically important for nanopore sequencing$^{47,48}$. The translocation time strongly depends on the orientation of bases$^{49-51}$, relating to the binding energy.

Based on our results, we thus note that the existence of line defects in graphene brings about pronounced improvement. By using first-principle investigations, we find that there is an opening of a new conducting channel due to the existence of a metallic wire embedded in a graphene sheet, resulting in higher conductance compared to using only a graphene nanopore. Although the influence of the solvent on the dynamics reduces the conductance by about 50%, there is a significant improvement in selectivity.

4 Conclusions

In conclusion, we have demonstrated that using extended line defects in graphene may present advantages over the most conventional choice based purely on graphene nanopores. The conductance modulation fully accounting for solvent effects has been studied by us using a combination of QM/MM and NEGF methods. Based on DFT formalism, we have investigated the interaction strength between line defects in graphene and nucleobases related to the selectivity and sensitivity of the devices. Furthermore, we have demonstrated the possibility to improve sensing performance by applying a gate voltage to the device. We are confident that our results will encourage future studies to fabricate the proposed device, as well as follow-up theoretical investigations in a search for higher sensing performance depending on pore edge termination by different functional groups.

Acknowledgements

The authors acknowledge support from the Thailand Research Fund (MRG5980185), FAPESP (Proj. # 2016/01343-7), FAPES, CAPES, CNPq-Brazil and the National Laboratory for Scientific Computing (LNCC/MCTI, Brazil) for providing HPC resources of the SDumont supercomputer, which have contributed to the research results reported within this paper: http://sdumont.lncc.br. ARR acknowledges support from the ICTP-Simmons Foundation associate scheme. RHS thanks the Swedish Research Council (VR) for financial support. The Swedish National Infrastructure for Computing (SNIC) provided computing time for this project.

References

1. S. J. Heerema and C. Dekker, Nat. Nanotechnol., 2016, 11, 127–136.
2. E. M. Bahassi and P. J. Stambrook, Mutagenesis, 2014, 29, 303–310.
3. D. A. Wheeler, M. Srinivasan, M. Egholm, Y. Shen, L. Chen, A. McGuire, W. He, Y.-J. Chen, V. Makhijani, G. T. Roth, X. Gomes, K. Tartaro, F. Niazi, C. L. Turcotte, G. P. Iryzk, J. R. Lupski, C. Ch Gaines, X.-z. Song, Y. Liu, Y. Yuan, L. Nazareth, X. Qin, D. M. Muzny, M. Margulies, G. M. Weinstock, R. A. Gibbs and J. M. Rothberg, Nature, 2008, 452, 872–876.
4. A. Ziegler, A. Koch, K. Kronenberger and A. Großhennig, Hum. Genet., 2012, 131, 1627–1638.
5. J. Shendure, S. Balasubramanian, G. M. Church, W. Gilbert, J. Rogers, J. A. Schloss and R. H. Waterston, Nature, 2017, 550, 345–353.
6. C. Dekker, Nat. Nanotechnol., 2007, 2, 209–215.
7. D. Fologea, J. Uplinger, B. Thomas, D. S. McNabb and J. Li, Nano Lett., 2005, 5, 1734–1737.
8. D. Fologea, M. Gershov, B. Lieden, D. S. McNabb, J. A. Golovchenko and J. Li, Nano Lett., 2005, 5, 1905–1909.
9. S. Howorka and Z. Szwy, Chem. Soc. Rev., 2009, 38, 2360–2384.
10. S. Howorka, S. Cheley and H. Bayley, Nat. Biotechnol., 2001, 19, 636–639.
11. J. Bello, Y.-R. Kim, S. M. Kim, T.-J. Jeon and J. Shim, Microchim. Acta, 2017, 184, 1883–1897.
12. A. H. Loo, A. Bonanni, A. Ambrosi and M. Pumera, Nanoscale, 2014, 6, 11971–11975.
