



IFT - UNESP
INSTITUTO DE FÍSICA TEÓRICA



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

Doctoral Thesis

IFT-T.005/18

QM/MM simulations of electronic transport properties for DNA sensing devices based on graphene

Ernane de Freitas Martins

June, 2018



IFT - UNESP
INSTITUTO DE FÍSICA TEÓRICA



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

Doctoral Thesis

IFT-T.005/18

QM/MM simulations of electronic transport properties for DNA sensing devices based on graphene

Ernane de Freitas Martins

Tese apresentada ao Programa de Pós Graduação em Física do Instituto de Física Teórica da Universidade Estadual Paulista como parte dos requisitos para obtenção do título de Doutor em Física.

Alexandre Reily Rocha - Advisor
Gustavo Troiano Feliciano - Co-advisor
Ralph Hendrik Scheicher - Sandwich PhD supervisor

June, 2018

M386q Martins, Ernane de Freitas
QM/MM simulations of electronic transport properties for DNA sensing devices based on graphene/ Ernane de Freitas Martins. – São Paulo, 2018
197 f. : il.

Tese (doutorado) – Universidade Estadual Paulista, Instituto de Física Teórica
Orientador: Alexandre Reily Rocha

1. Grafeno. 2. Dinâmica molecular. 3. DNA. 4. Funcionais de densidade.
I. Título

Ficha catalográfica elaborada pela Biblioteca da UNESP
Instituto de Física Teórica – São Paulo (SP) –
Hellen Sayuri Sato / CRB 8/5888

To my family, parents and wife...

Acknowledgments (in Portuguese)

Em primeiro lugar eu agradeço a Deus por tudo que me proporcionou até aqui. Agradeço à minha família, especialmente meu irmão, minhas irmãs, minha cunhada e meus cunhados pelo apoio de sempre. De modo muito especial agradeço aos meus pais, José e Nely, os quais sempre me incentivaram a continuar com os estudos e não desistir da longa jornada até o doutorado, jornada esta que até poderia ser considerada uma ousadia num país em que filhos de famílias simples sequer iniciavam uma graduação. Foram muitos momentos difíceis desde o início da graduação, em 2009, até chegar neste ponto prestes a finalizar o doutorado. Nós conseguimos. Um muito obrigado a todos vocês!

Agradeço ao meu orientador, Alexandre, por ter me aceitado no processo de seleção para ingresso no doutorado, me dando a oportunidade de executar esses projetos. Agradeço por todos os ensinamentos, discussões, apoio e incentivo. Agradeço também pelo apoio que nos deu para a organização do Congresso Paulo Leal Ferreira de Física (CPLF) no IFT. Agradeço também ao meu coorientador, Gustavo, por todos os ensinamentos das técnicas que utilizamos durante as agradáveis visitas à Araraquara, por sempre ter atendido às minhas dúvidas por mensagens e pelas vídeo chamadas para discutir nossos resultados. Mesmo à distância você sempre foi solícito em me ajudar com o que era preciso.

Thanks to Ralph for receiving me at Uppsala University during the internship. Thank you for all the discussions, corrections in my texts and conferences abstracts, review in this thesis, meetings, the *fkas* and everything else. My internship time under your supervision was really important and fundamental in my PhD. I am very grateful for everything.

Agradeço a todos os professores que contribuíram para minha formação até aqui. Todos vocês foram fundamentais para minha formação profissional e pessoal. Também agradeço a todos os funcionários do IFT que tanto nos ajudam em vários momentos. De modo especial agradeço à Rosane pelo apoio na solução de problemas relacionados às bolsas de estudo que tive. Também agradeço de modo especial à Meire pela solicitude e grande ajuda na organização do CPLF.

Agradeço a todos meus amigos, tanto os de longa data quanto os que conquistei durante o doutorado: Ana, Dérick, Enzo, Gabriel, Henrique, John, Luan, Nathaly, Natália, Segundo, Vivian e tantos outros que em cada conversa, incentivo ou apoio, também contribuíram para que eu chegasse até aqui. Agradeço de modo especial à Nathaly por suas sugestões e correções nesta tese. Considero que tivemos ótimos momentos no IFT e sem dúvidas vocês todos farão muita falta. Espero revê-los.

Agradeço aos amigos que fiz durante o período de doutorado sanduíche na Suécia. Muito obrigado por não me deixar esquecer do calor humano que nós brasileiros temos. Agradeço de modo especial à Débora por ter me recebido de maneira tão solícita e carinhosa. Também

agradeço à Giane por todos os momentos que passamos juntos lá. Como eu gostava das nossas pausas para o café. Agradeço de modo especial ao meu amigo Ageo. Sua amizade foi fundamental para eu concluir o período de sanduíche na Suécia. Agradeço também pelas sugestões e ajuda com o texto nesta tese. Todos vocês me ajudaram muito mesmo a lidar com todas as dificuldades que tive em viver num país totalmente diferente do nosso. Muito obrigado!

Agradeço à Juliana, agora minha esposa. Ela foi um alicerce para mim desde o início do nosso namoro. Sempre estive ao meu lado me aconselhando e me ajudando, o que me tornou uma pessoa melhor. Ela fez parte de todas as difíceis transições que passamos nessa vida de pós-graduandos. Ela foi para mim uma excelente companheira durante o namoro e está sendo uma esposa maravilhosa. Obrigado pela ajuda na correção desta tese e muito obrigado por tudo mais que você fez por mim até aqui e pela vida inteira que ainda vamos viver juntos daqui pra frente.

Agradeço à agência de fomento CNPq pela bolsa de estudos no Brasil e reserva técnica. Também agradeço à agência CAPES pela bolsa de estudos, despesas de viagens, todos os auxílios e suporte necessário para o desenvolvimento do doutorado sanduíche na Universidade de Uppsala, Suécia.

Agradeço ao Laboratório Nacional de Computação Científica (LNCC/MCTI) por fornecer recursos de computação de alto desempenho do supercomputador SDumont, os quais foram fundamentais para o desenvolvimento da maioria das simulações computacionais desenvolvidas no meu doutorado. É importante salientar que a execução desses projetos só foi possível devido aos recursos disponibilizados pelo LNCC. Também agradeço ao *Swedish National Infrastructure for Computing* por disponibilizar recursos computacionais do projeto sob supervisão do professor Ralph Scheicher, os quais foram utilizados para obter parte dos resultados apresentados no Capítulo 4.

O doutorado representou, para mim, um desafio pessoal. Não é fácil ser um pós-graduando e enfrentar todas as adversidades que surgem. O apoio de todos vocês foi fundamental e fica aqui o meu mais sincero agradecimento.

Nanotechnology is an important and very active area of research contributing to many different fields. The development of new devices applied to personalized medicine is one of its applications. When we desire to develop new devices many effort are done, including experimental and theoretical investigations. The theoretical/computational physics can enormously contribute to this area, since the simulations can reveal the working mechanism in these systems being possible to understand and propose new devices with improved performance.

We present an extensive theoretical investigation of the electronic transport properties of graphene-based devices for DNA sensing. We have used a hybrid methodology which combines quantum mechanics and molecular mechanics, the so called QM/MM method, coupled to electronic transport calculations using non-equilibrium Green's functions.

First, we studied graphene in solution in order to understand the effects of polarization on the electronic and transport properties under different salt concentrations. We also studied graphene with Stone-Wales defect in pure water. For these systems we tested a simple polarization model based on rigid rods. Our analysis were also done over different QM/MM partitions including explicit water molecules in the quantum part. Our results showed that the inclusion of the solvent in the electronic transport calculations for graphene decreases the total transmission, showing the important role played by the water. Our results also showed that the electronic transport properties of graphene do not suffer significant changes as we increase the salt concentration in the solution. The inclusion of polarization effects in graphene, despite changing the structuring of water molecules that make up the first solvation shell of graphene, do not significantly affect the electronic transport through graphene.

We then studied DNA sequencing devices. First we focused on sequencing using a nanopore between topological line defects in graphene. Our results showed that sequencing DNA with high selectivity and sensitivity using these devices appears possible. We also address nanogap in graphene. For this we looked at the effects of water on electronic transport by using different setups for the QM/MM partition. We showed that the inclusion of water molecules in the quantum part increases the electronic transmission in several orders of magnitude, also showing the fundamental role played by water in tunneling devices. The electronic transport simulations showed that the proposed device has the potential to be used in DNA sequencing, presenting high selectivity and sensitivity.

We propose an graphene-based biochip for sequence-specific detection of DNA strands. The main idea of this sort of device is to detect hybridization of single-stranded DNA, forming double-stranded DNA. We showed that the vertical DNA adsorption, either through an anchor molecule (pyrene) or using the nucleotide itself as anchor, do not present good re-

sults for detection, since the signals for the single and double strands are quite similar. For the case of horizontal DNA adsorption on graphene our results indicated that the two signals can be distinguishable, showing promising potential for sensitivity and selectivity.

Keywords: Graphene, DNA, nanopore, nanogap, sequencing, DFT, electronic transport, molecular dynamics simulations, QM/MM.

Nanotecnologia é uma importante e muito ativa área de pesquisa contribuindo para muitos campos diferentes. O desenvolvimento de novos dispositivos aplicados à medicina personalizada é uma de suas aplicações. Quando desejamos desenvolver novos dispositivos muitos esforços são feitos, incluindo investigações experimentais e teóricas. A Física teórica/computacional pode contribuir enormemente com esta área, já que simulações podem revelar o mecanismo de funcionamento nesses sistemas tornando possível entender e propor novos dispositivos com desempenho melhorado.

Nós apresentamos uma extensa investigação teórica das propriedades de transporte eletrônico de dispositivos baseados em grafeno para sensoriamento de DNA^I. Utilizamos uma metodologia híbrida que combina mecânica quântica e mecânica molecular, o chamado método QM/MM, acoplado a cálculos de transporte eletrônico utilizando funções de Green fora do equilíbrio.

Primeiramente nós estudamos grafeno em solução de modo a entender os efeitos de polarização nas propriedades eletrônica e de transporte em diferentes concentrações de sal. Também estudamos grafeno com defeito Stone-Wales em água pura. Para esses sistemas, testamos um modelo de polarização simples baseado em bastões rígidos. Nossas análises também foram feitas em diferentes partições QM/MM^{II} incluindo moléculas de água explícitas na parte quântica. Nossos resultados mostraram que a inclusão do solvente nos cálculos de transporte eletrônico para o grafeno diminui a transmissão total, mostrando o papel fundamental desempenhado pelo água. Nossos resultados também mostraram que as propriedades de transporte eletrônico do grafeno não sofrem mudanças significativas na medida em que aumentamos a concentração de sal na solução. A inclusão de efeitos de polarização em grafeno, apesar de mudar a estruturação das moléculas de água que compõem a primeira camada de solvatação do grafeno, não afeta significativamente o transporte eletrônico através do grafeno.

Nós, então, estudamos dispositivos para sequenciamento de DNA. Focamos primeiramente no sequenciamento usando nanoporo entre defeitos de linha topológicos no grafeno. Nossos resultados mostraram que o sequenciamento de DNA com alta seletividade e sensibilidade usando esses dispositivos se mostra possível. Nós também abordamos *nanogap* em grafeno. Para tal, avaliamos os efeitos da água no transporte eletrônico utilizando diferentes configurações para a partição QM/MM. Mostramos que a inclusão de moléculas de água na parte quântica aumenta a transmissão eletrônica em várias ordens de grandeza, também

^IDNA, do inglês *Deoxyribonucleic Acid*, é a sigla utilizada para Ácido Desoxirribonucleico.

^{II}QM/MM, do inglês *Quantum Mechanics/Molecular Mechanics*, é a sigla utilizada para referenciar este método.

mostrando o papel fundamental desempenhado pela água em dispositivos de tunelamento. As simulações de transporte eletrônico mostraram que o dispositivo proposto tem o potencial de ser usado em sequenciamento de DNA, apresentando alta seletividade e sensibilidade.

Propusemos um biochip baseado em grafeno para detecção de sequências específicas de fitas de DNA. A ideia principal desta classe de dispositivos é detectar a hibridização da fita simples de DNA, formando a fita dupla de DNA. Mostramos que a adsorção vertical de DNA, seja utilizando uma molécula âncora (pireno) ou utilizando o próprio nucleotídeo como âncora, não apresenta bons resultados para detecção, já que os sinais para as fitas simples e dupla são bem próximos. Para o caso da adsorção horizontal de DNA em grafeno nossos resultados indicaram que os dois sinais podem ser distinguíveis, mostrando potencial promissor para sensibilidade e seletividade.

Palavras-chave: Grafeno, DNA, nanoporo, *nanogap*, sequenciamento, DFT, transporte eletrônico, simulações de dinâmica molecular, QM/MM.

1.1	General features of a biosensor showing its basic components: biological element, transducer and signal processors. The analyte is detected according to the used bioselective material, then the bio-recognition signal is transformed in the transducer into one that can be measured to obtain the final results. Reprinted from reference [1]. Open access.	35
1.2	Four nucleotides of DNA (not showing Hydrogen atoms). The letters A, T, C and G indicates the nucleobases adenine, thymine, cytosine and guanine, respectively. The color-code chosen to indicate each base will be used throughout the results chapters of the thesis. Carbon in gray, Nitrogen in blue, oxygen in red, and phosphorus in golden. This color-code for the atoms will be used throughout all chapters of this thesis.	36
1.3	Schematic representation of DNA structure. The building (a) blocks of DNA are used to construct the DNA strand (b). An illustration in plane (c) of a double-stranded DNA and the DNA double helix (d). Adapted from reference [4]. Open access.	36
1.4	Scanning tunneling microscopy (STM) images for films of graphene with different thicknesses: (a) 1 mono layer, (b) 3 mono layers and (c) 2 mono layers. The colors shows equivalent and inequivalent atoms in the different layers. Reprinted with permission from reference [6]. Copyright ©(2008) American Physical Society. . . .	37
1.5	Cost per genome of sequencing DNA as a function of the years. Reprinted from reference [23]. Access in may 28, 2018.	38
1.6	DNA sequencing device based on biological pores. The detection is based on ionic current changes and the movement of the DNA inside the pore is controlled by an electric field (a). Structures of α -hemolysin (b) and MspA (c) inserted in lipid bilayer and graphene nanopore (d). Reprinted with permission from reference [25]. Copyright ©(2012) Nature Publishing Group.	39
1.7	Graphene-based devices for DNA sequencing. A nanopore (a) can be used for detecting each nucleotide based on ionic current changes. Nanogaps (b) and nanopores (c) can be used for detection based on electronic changes. Another possibility is the detection based on physisorption process (d). Reprinted with permission from reference [24]. Copyright ©(2016) Nature Publishing Group. . . .	40

1.8	Sketch for the DNA sensing device. The device is connected to a source and drain electrodes. A modified graphene (nanopore or nanogap; only nanogap example shown here) can be used to sequence DNA in such a way that a ssDNA is translocated through the device (upper panel). The DNA detection device can be constructed using pristine graphene with ssDNA attached to the surface. Thus, the ssDNA sample is added to the device and, if this added strand is the complementary DNA strand, the hybridization (and the detection) is reached.	41
2.1	Exemplification of force field terms for Adenine DNA nucleotide. The green atoms/bonds exemplifies each term: (a) bond, (b) angle and (c) dihedral. T large radius spheres in (d) exemplifies the pair defined by the two highlighted atoms. . .	51
2.2	Flowchart for MD calculations.	55
2.3	Flowchart for the DFT self-consistent cycle.	61
2.4	Explanation of QM/MM partitions. The magenta atoms are in the QM part and the cyan ones are in the MM part. This QM/MM partition was defined across a chemical bond. Reprinted with permission from reference [56]. Copyright ©(2013) Springer Nature.	62
2.5	Illustration of the electronic transport problem. In this figure the thermodynamics (a), Hamiltonian (b) and electrostatics (c) pictures of this problem are shown. Reprinted with permission from reference [61]. Copyright ©(2006) American Physical Society.	65
2.6	External potential calculated for the classical part of an example system. The atoms of the classical part (top view) will determine the average potential as a function of the z direction. This figure shows the water molecules and the Na (blue) and Cl (green) ions.	68
2.7	Flowchart summarizing the steps of the methodology from Molecular Dynamics to Electronic Transport calculation.	70
3.1	Different setups used to study graphene in solution. Dry (a) pristine graphene and the same system with one layer of water (e), both for fixed graphene. Initial frame for the polarization model used for graphene (b) and one frame after the begin of simulation in side (c) and top (d) view. Top view for dry graphene with Stone-Wales defect and the same with one layer of water (f). Dry pristine graphene allowing movements in the sheet (g) where we call free graphene instead of fixed graphene.	75
3.2	Electronic transport setup for the study of both graphene in solution and DNA detection. The green areas delimit the leads and the blue ones delimit the buffers. The scattering region is also shown.	79

3.3	Averaged electronic transport calculation (200 frames) for fixed non-polarized pristine graphene in solution (water with 0.2 <i>M</i> of NaCl) and single electronic transport calculation for dry pristine graphene.	80
3.4	Sketch for Density of States (DOS) of pristine graphene to exemplify the case in which there is no scattering.	81
3.5	The angles θ_1 and θ_2 are the angles between the OH vectors showed in (a) and the $\hat{\mathbf{i}}$ unitary vector. Different layers of water (b) with the probability distribution for OH tilt angles (c), for non-polarized fixed graphene with 0.2 <i>M</i> of NaCl. The colors for each layer of water in (b) corresponds to the colors of each curve in (c). Only the half part of the simulation box is shown.	82
3.6	Number density for Oxygen atoms, for non-polarized fixed graphene with 0.2 <i>M</i> of NaCl.	83
3.7	Probability distribution for OH tilt angles of the first layer of water for fixed pristine graphene in different solutions: (a) pure water, (b) water with 0.2 <i>M</i> of NaCl and (c) water with 1 <i>M</i> of NaCl.	84
3.8	Side view of a representative snapshot for non-polarized fixed graphene in water and 0.2 <i>M</i> of NaCl. The graphene and water in <i>ball-and-sticks</i> representation make up the QM part whereas the water in <i>line</i> representation and the NaCl (green and blue <i>ball</i> representation) make up the MM part. This representation pattern will be used throughout this thesis.	85
3.9	Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene in water with 1 <i>M</i> of NaCl.	86
3.10	Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene in (a) pure water and (b) water with 0.2 <i>M</i> of NaCl.	87
3.11	Partial densities for Na and Cl ions with respect to the distance from graphene sheet. Example for fixed non-polarized graphene in water with 0.2 <i>M</i> of NaCl. The graphene sheet is placed at $x = 23.74 \text{ \AA}$	88
3.12	Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene including the first layer of water in the QM part in (a) pure water and (b) water with 0.2 <i>M</i> of NaCl. The cases without QM water are repeated here for comparison.	89
3.13	Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene including the first layer of water in the QM part in water with 1 <i>M</i> of NaCl. The case without QM water is repeated here for comparison.	90

3.14	Convergence analysis example for fixed non-polarized pristine graphene in 0.2 M of NaCl. Electronic transport calculations for all frames and the average over these curves (a), averages over different numbers of frames (b) and evolution of the average for $E - E_F = 0.92 \text{ eV}$ (c).	91
3.15	Top and side views and a zoom at the defect for the water molecules close to the Stone-Wales defect in fixed non-polarized graphene.	92
3.16	Probability for OH tilt angles of the first layer of water for fixed graphene with Stone-Wales defect in pure water comparing the cases with the entire first layer of water and the case including only the water molecules close to the defect.	93
3.17	Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed graphene with Stone-Wales defect.	94
3.18	Comparison between the averages of electronic transport calculations for non-polarized fixed graphene with Stone-Wales defect in pure water and including the first layer of water in the QM part. The case without QM water was repeated here for comparison.	95
3.19	Comparison between the averages of electronic transport calculations for free pristine graphene in liquid (pure water) and in vacuum.	96
3.20	Comparison of the averages for all electronic transport calculations for the systems immersed in pure water (a) and a zoom close to the Fermi energy in (b).	97
4.1	Schematic representation of the nanopore in between line defects (blue area) in graphene with one octagon in blue and one pentagon in red.	102
4.2	(a) Schematic representation for DNA sequencing device using a nanopore between topological line defects in graphene. (b) Setup used in our calculations showing the graphene sheet with a DNA strand inside the nanopore. Reprinted with permission from reference [99]. Copyright ©(2018) American Chemical Society.	103
4.3	Top (zoomed at the nanopore) and side views for representative snapshots of all nucleotides used in the line defect DNA sequencing setup. The <i>ball-and-sticks</i> representation and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the <i>line</i> representations make up the MM part.	105
4.4	Electronic transport setup for the study of DNA sequencing using line defects in graphene nanopore. The green areas delimit the leads and the blue ones delimit the buffers. The scattering region is also shown. The line defects shown in this figure will be further explained.	108
4.5	Electronic transport calculations for the liquid system (nanopore without nucleotide inside - L) and for all frames of each nucleotide defined by the letters used to refer each one. The black curve is the average over the presented curves.	110

4.6	Electronic transport calculations for the system (L) system and for all nucleotides, defined by the letters used to refer each one. The gray shaded region comprises the addition and subtraction of 1 SD.	111
4.7	Electronic transport averages over different number of frames for the liquid system (L) and for all nucleotides, defined by the letters used to refer each one.	112
4.8	Electronic transport calculation for dry empty pore and the same immersed in liquid (average).	113
4.9	Comparison between the average electronic transport calculated for all nucleotides.	114
4.10	Comparison between the single electronic transport calculated for all nucleotides in gas phase using the most favorable structures. Reprinted with permission from reference [99]. Copyright ©(2018) American Chemical Society.	115
4.11	Sensitivity (upper panel) for the device using nanopore in between line defects in graphene for specif energies using the liquid system as reference (Equation 4.1) and selectivity (lower panel) for the same energies (gate voltages) using G as reference (Equation 4.2).	116
5.1	Illustration of the DNA sequencing device using Nitrogen-terminated nanogaps in graphene, showing the DNA strand translocating through the gap. Adapted with permission from reference [17]. Copyright ©(2016) American Chemical Society. . .	119
5.2	Scheme for tunneling process in DNA sequencing using nanogaps. The blue area represents the liquid (water and NaCl) and e^- represents the electron. The Guanine structure was used as example.	120
5.3	Top views for representative snapshots of all nucleotides (A, T, C and G) in the DNA sequencing device using nanogap. In this snapshot the setup 2 for QM/MM partition is shown. The large radius blue spheres and the <i>ball-and-sticks</i> representation make up the QM part while the small radius blue spheres, the green spheres and the <i>line</i> representations make up the MM part.	121
5.4	Top (upper panel) and side (lower panel) view for a representative snapshot for liquid system (nanogap immersed in water and NaCl without nucleotide) showing the unique QM/MM partition setup for this system (the same as setup 3 for the nucleotide systems). The water molecules in <i>ball-and-sticks</i> representation and the whole nanogap sheet make up the QM part, while the remaining atoms make up the MM part.	122
5.5	Side view of the first 2 setups for QM/MM partitions for all nucleotides. The <i>ball-and-sticks</i> representation and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the <i>line</i> representations make up the MM part.	123

5.6	Side view of the third setup for QM/MM partitions for all nucleotides. The <i>ball-and-sticks</i> representation and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the <i>line</i> representations make up the MM part.	124
5.7	Electronic transport setup for the study of DNA sequencing using nanogap. The green areas delimit the leads and the blue ones delimit the buffers. The scattering region is also shown.	127
5.8	Electronic transport averages over different number of frames for all nucleotides in setup 1. The letters defined before identifies each system.	128
5.9	Electronic transport averages over different number of frames for all nucleotides in setup 2. The letters defined before identifies each system. For nucleotide G we used 100 frames because the convergence with 50 frames was not conclusive. . . .	129
5.10	Electronic transport averages over different number of frames for all nucleotides in the setup 3. The letters defined before identifies each system. This is the setup in which we have the liquid (L) system used as reference in our calculations. . . .	130
5.11	Averaged electronic transport calculations in setup 3 for L (filled black line) system and for A, T, C and G in setups 1 (dotted line), 2 (filled line) and 3 (dashed line). The letters and the color-code defined before identifies each system.	132
5.12	Comparison between the average electronic transport calculated in setup 3 for A, T, C, G and L.	133
5.13	Representation of the six bins used to calculate and normalize the local currents. The system for nucleotide A in top and side views without QM water is used as example.	134
5.14	Top and side views for the left-to-right local currents on all nucleotides in QM water. The arrows' thickness are proportional to the value of the current normalized by the largest current in each bin. A cutoff has been introduced for clarity.	135
5.15	Top and side view for the currents of nucleotide T. The arrows' colors identify each subsystem: pink for local currents inside the sheet, blue for the local current from the sheet to water, yellow for the local currents from (to) the water to (from) the nucleotide and green for the local currents from (to) the sheet to (from) the nucleotide. A cutoff has been introduced for clarity.	136
5.16	Top and side view for the currents of nucleotide T. The arrows shows the current flowing from the sheet to the water molecules and through the water molecules. In this case the currents were normalized by the largest current in the subsystem for each bin. A cutoff has been introduced for clarity.	137
5.17	(a) z component of all currents for all nucleotides, (b) the same for subsystem 1: currents from the sheet to the nucleotide and (c) for subsystem 2: from the sheet to the water molecules. The dashed lines show the positions of the gaps.	139

5.18	Top and side view for the snapshots used in the current calculations showed in Figure 5.14 for all nucleotides (A, T, C and G). The water molecules and the Na atom are hidden to better visualize the geometry.	140
5.19	Sensitivity (upper panel) for the device using nanogap in graphene for specific energies using the liquid system as reference (Equation 4.1) and selectivity (lower panel) for the same energies using G as reference (Equation 4.2).	142
6.1	Schematic representation of the DNA hybridization process. When the complementary DNA strand is present the hybridization process occurs. Reprinted from reference [102]. Open access.	146
6.2	Representations of vertical DNA systems using pyrene as anchor molecule. The pyrene molecule is shown in (a), the ssDNA in (b) and the dsDNA in (c). Sequences: 4A + pyrene and 4A + pyrene + 4T.	147
6.3	Snapshot for a ssDNA on graphene highlighting the restraint applied for ssDNA with anchor molecule. Atoms in yellow are restrained in the direction of the yellow axis. There is no restraint in the direction of the blue axes (x direction).	148
6.4	Top and side views for a representative snapshot of vertical ssDNA (a and c) and dsDNA (b and d) without QM water using pyrene as anchor molecule. 4 Na atoms in the QM part for ssDNA and 7 Na atoms in the QM part for dsDNA.	149
6.5	Top and side view for a representative snapshot of vertical ssDNA (a and c) and dsDNA (b and d) with QM water using pyrene as anchor molecule.	150
6.6	Side view for vertical ssDNA (upper panel) and dsDNA (lower panel) without QM water using pyrene as anchor molecule to represent the QM/MM partition idea.	151
6.7	Electronic transport calculations and the averages over the curves for vertical DNA systems using anchor molecule: (a) ssDNA and (b) dsDNA without QM water and (c) ssDNA and (d) dsDNA with QM water.	154
6.8	Electronic transport calculations for vertical DNA systems using anchor molecule: (a) ssDNA and (b) dsDNA without QM water and (c) ssDNA and (d) dsDNA with QM water. The gray shaded region comprises the addition and subtraction of 1 SD.	155
6.9	Evolution of the average and SD for all vertical DNA systems using anchor molecule: (a) ssDNA and (b) dsDNA without QM water and (c) ssDNA and (d) dsDNA with QM water. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \simeq 0.0 \text{ eV}$	157
6.10	Comparison between the averages over 50 electronic transport calculations for vertical ssDNA and dsDNA with (a) and without (b) QM water.	159
6.11	Representations of vertical ssDNA (a) and dsDNA (b) systems not using pyrene as anchor molecule. Sequences: 6G and 6G + 4C.	160

6.12	Restraint applied for ssDNA with nucleotide as anchor molecule. Atoms in yellow are restrained in the direction of the yellow axes. There is no restraint in the direction of the blue axis (x direction).	161
6.13	Top and side views for a representative snapshot of vertical ssDNA (a and c) and dsDNA (b and d) without QM water and not using anchor molecule. 5 Na atoms in the QM part for ssDNA and 8 Na atoms in the QM part for dsDNA.	162
6.14	Electronic transport calculations for all frames of vertical ssDNA (a) and dsDNA (b) systems without anchor molecule and the average over these curves.	164
6.15	Electronic transport calculations for vertical ssDNA (a) and dsDNA (b) systems without anchor molecule. The gray shaded region comprises the addition and subtraction of 1 SD.	164
6.16	Evolution of the average and SD for vertical ssDNA (a) and dsDNA (b) systems without anchor molecule. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \simeq 0.0 \text{ eV}$.	165
6.17	Comparison between the average over 50 electronic transport calculations for vertical ssDNA and dsDNA without anchor molecule.	166
6.18	Representations of horizontal ssDNA (a) and dsDNA (b) systems. Sequences 4A and 4A + 4T.	167
6.19	Restraint applied for ssDNA (left) and dsDNA (right) for the horizontal case. All atoms are restrained in both cases in the direction of the yellow axis. There is no restraint in the direction of the blue axes (x and y directions).	168
6.20	Top and side views for a representative snapshot of horizontal ssDNA (a and c) and dsDNA (b and d) without QM water. 3 Na atoms in the QM part for ssDNA and 6 Na atoms in the QM part for dsDNA.	169
6.21	Electronic transport calculations for all frames of horizontal ssDNA (a) and dsDNA (b) systems. The average over these curves is shown in black.	171
6.22	Electronic transport calculations for all horizontal DNA systems: (a) ssDNA and (b) dsDNA. The gray shaded region comprises the addition and subtraction of 1 SD.	171
6.23	Evolution of the average and SD for all horizontal DNA systems: (a) ssDNA and (b) dsDNA. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \simeq 0.0 \text{ eV}$.	172
6.24	Comparison between the average over 50 electronic transport calculations for horizontal ssDNA and dsDNA, average over 200 electronic transport calculations for graphene in liquid and single calculation for dry pristine graphene.	173
6.25	Sensitivity for ssDNA and dsDNA using the results for graphene in solution as reference and selectivity for dsDNA using ssDNA as reference, for the whole energy spectrum.	174

6.26	Sensitivity for ssDNA and dsDNA using the results for graphene in solution as reference and selectivity for dsDNA using ssDNA as reference, for specif energies. . .	174
A.1	Schematic representation of the investigated QM/MM partitions. Graphene is always in the QM part and liquid means water, Cl ions and all Na ions that are not included in the QM part.	191
A.2	Side view of the five different QM/MM partitions for a sequence of 4A connected to pyrene. The <i>ball-and-sticks</i> representations and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the <i>line</i> representations make up the MM part.	192
A.3	Electronic transport calculations for QM/MM partitions (a) 1, (b) 2, (c) 3, (d) 4 and (e) 5. The gray shaded region comprises the addition and subtraction of 1 SD. . . .	194
A.4	Evolution of the average and SD for QM/MM partitions (a) 1, (b) 2, (c) 3, (d) 4 and (e) 5. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \simeq 0.0 \text{ eV}$	195
A.5	Comparison of the averages over 10 and 50 electronic transport calculations for QM/MM partition 5.	196
A.6	Comparison of the averages over 10 electronic transport calculations for 5 different QM/MM partitions.	197

3.1	Simulation details for the study of graphene in solution.	76
3.2	Details of the simulation box for the study of graphene in solution for pure water and water with 0.2 <i>M</i> of NaCl. <i>N</i> is the number of the specified species. Length unit: Å.	77
3.3	Details of the simulation box for the study of graphene in water with 1 <i>M</i> of NaCl. <i>N</i> is the number of the specified species. Length unit: Å.	78
3.4	Details of the simulation box for the study of graphene in solution for graphene with Stone-Wales defect. <i>N</i> is the number of the specified species. Length unit: Å.	78
4.1	Simulation details for the study of DNA sequencing using line defects in graphene nanopore.	106
4.2	Details of the simulation box for the nucleotides in the study of DNA sequencing using nanopore. <i>N</i> is the number of the specified species. Length unit: Å.	107
4.3	Details of the simulation box for the empty pore in the study of DNA sequencing using nanopore. <i>N</i> is the number of the specified species. Length unit: Å.	108
5.1	Simulation details for the study of DNA sequencing using nanogap.	125
5.2	Details of the simulation box for the nucleotides in the study of DNA sequencing using nanogap. <i>N</i> is the number of the specified species. Length unit: Å.	126
5.3	Details of the simulation box for the empty gap in the study of DNA sequencing using nanogap. <i>N</i> is the number of the specified species. Length unit: Å.	127
6.1	Simulation details for the study of DNA detection.	152
6.2	Details of the simulation box for the study of vertical DNA detection using anchor molecule. <i>N</i> is the number of the specified species. Length unit: Å.	153
6.3	Details of the simulation box for the study of vertical DNA detection not using anchor molecule. <i>N</i> is the number of the specified species. Length unit: Å.	163
6.4	Details of the simulation box for the study of horizontal DNA detection. <i>N</i> is the number of the specified species. Length unit: Å.	170

2.1	Time dependent Schrödinger Equation	45
2.2	Time independent Schrödinger Equation	45
2.3	Many body Hamiltonian	46
2.4	Kinetic Energy Operator	46
2.5	Potential Energy Operator	46
2.6	Simplified Hamiltonian	46
2.7	Solution for Schrödinger Equation	46
2.8	Electronic Schrödinger Equation	46
2.9	Electronic Hamiltonian	46
2.10	Electronic and nuclear Hamiltonian	47
2.11	Wave Function applied in SE - 1	47
2.12	Wave Function applied in SE - 2	47
2.13	Wave Function applied in SE - 3	47
2.14	Wave Function applied in SE - 4	47
2.15	Wave Function applied in SE - 5	47
2.16	Wave Function applied in SE - 6	47
2.17	Wave Function applied in SE - 7	48
2.18	Wave Function applied in SE - 8	48
2.19	Wave Function applied in SE - 9	48
2.20	Wave Function applied in SE - 10	48
2.21	Reduced solution for the Schrödinger Equation	48
2.22	Classical nuclei approximation	49
2.23	E'_{el} definition	49
2.24	Electronic Schrödinger Equation	49
2.25	Newton's second law	50
2.26	Potential derivative	50
2.27	Typical Force Field	51
2.28	Ensemble Average	52
2.29	Time Average	52
2.30	Ergodic Hypothesis	52
2.31	Taylor expansion for position	53
2.32	Taylor expansion for velocity	53
2.33	Taylor expansion for acceleration	53
2.34	Leap-frog integrator for position	53
2.35	Leap-frog integrator for velocity	53

2.36	Leap-frog integrator for $\vec{v}(t)$	53
2.37	Verlet integrator for $\vec{r}(t + \delta t)$	53
2.38	Verlet integrator for $\vec{r}(t - \delta t)$	54
2.39	Verlet integrator summation	54
2.40	Verlet/velocity integrator for $\vec{r}(t) + \delta t$	54
2.41	Verlet/velocity integrator for $\vec{v}(t) + \delta t$	54
2.42	Electronic density definition	56
2.43	Hohenberg e Kohn Theorems - 1	56
2.44	Hohenberg e Kohn Theorems - 2	56
2.45	Hohenberg e Kohn Theorems proof - 1	56
2.46	Hohenberg e Kohn Theorems proof - 2	57
2.47	Hohenberg e Kohn Theorems proof - 3	57
2.48	Hohenberg e Kohn Theorems proof - 4	57
2.49	Hohenberg e Kohn Theorems proof - contradiction	57
2.50	Hohenberg e Kohn Theorems proof - universal functional existence	57
2.51	Hohenberg e Kohn Theorems proof - universal functional definition	57
2.52	Hohenberg e Kohn Theorems proof - variational principle	58
2.53	Hohenberg e Kohn Theorems proof - variational principle	58
2.54	Energy as a functional of the density	58
2.55	Minimization of the energy - 1	58
2.56	Minimization of the energy - 2	58
2.57	Minimization of the energy - 3	58
2.58	Minimization of the energy - 4	58
2.59	Exchange and Correlation potential	59
2.60	Hartree potential	59
2.61	Energy as a functional of the density in the KS formalism	59
2.62	Schrödinger-like Kohn-Sham equation	59
2.63	KS Hamiltonian	59
2.64	KS wave function	59
2.65	Spin degree of freedom in the wave function - 1	59
2.66	Spin degree of freedom in the wave function - 2	60
2.67	QM/MM subtractive coupling	62
2.68	QM/MM additive coupling	62
2.69	QM/MM Hamiltonian example	63
2.70	Interaction between solvent atoms - QM/MM	63
2.71	Interaction between electrons and solute atoms - QM/MM	63
2.72	Interaction between solvent atoms and solute nuclei - QM/MM	63
2.73	Electronic Hamiltonian - QM/MM	64
2.74	QM/MM electrostatic coupling	64

2.75	Green's function definition	66
2.76	Right lead Hamiltonian for the scattering problem using Green's functions	66
2.77	Left lead Hamiltonian for the scattering problem using Green's functions	66
2.78	Right coupling matrix for the scattering problem using Green's functions	66
2.79	Left coupling matrix for the scattering problem using Green's functions	66
2.80	Matrix for the scattering problem using Green's functions	67
2.81	Green's function for the scattering region	67
2.82	Self-energy of the right electrode	67
2.83	Self-energy of the left electrode	67
2.84	Current at a given bias V	67
2.85	Total transmission calculated from the Green's functions	67
2.86	Left/Right coupling matrix in the transmission	68
2.87	Conductance at low bias - Fisher-Lee equation	68
2.88	Local current density between two sites	69
4.1	Sensitivity for DNA sequencing devices	115
4.2	Selectivity for DNA sequencing devices	115
6.1	Sensitivity for DNA detection device	173
6.2	Selectivity for DNA detection device	174

List of Abbreviations and Constants

Abbreviations

HF	Hartree-Fock
SE	Schrödinger equation
BO	Born-Oppenheimer
DFT	density functional theory
KS	Kohn-Sham
QM	quantum mechanical
MM	molecular mechanical
SM	statistical mechanics
QM/MM	quantum mechanics/molecular mechanics
MD	molecular dynamics
FF	force field
SPC	simple point-charge
V_{ext}	external potential
NPT	isothermal-isobaric ensemble
NVT	canonical ensemble
LDA	local-density approximation
LSDA	local-spin-density approximation
GGA	generalized-gradient approximation
LCAO	linear combination of atomic orbitals
NEGF	non-equilibrium Greens's functions
SD	standard deviation
DNA	deoxyribonucleic acid
ssDNA	single stranded DNA
dsDNA	double stranded DNA
vdW	van der Waals
SW	Stone-Wales
PBC	periodic boundary conditions

Constants

h	Planck constant, $h = 6.626070040(81) \times 10^{-34} \text{ Js}$
$\hbar$	reduced Planck constant, $\hbar = \frac{h}{2\pi}$
e	elementary charge, $e = 1.60217662 \times 10^{-19} \text{ C}$

1	General introduction	34
1.1	Nanodevices	35
1.2	Deoxyribonucleic acid - DNA	35
1.3	Graphene	37
1.4	DNA detection and sequencing	38
1.5	The role of the solvent	42
1.6	Challenges	42
1.7	Objective	43
2	Methodology and theory	44
2.1	How to simulate a nanodevice	45
2.2	Methods	50
2.2.1	Molecular Dynamics - MD	50
2.2.2	Density Functional Theory - DFT	55
2.2.3	The QM/MM method	61
2.2.4	Electronic Transport - NEGF	64
2.3	General simulation details	69
2.3.1	The step-by-step	69
3	Graphene in solution	72
3.1	Introduction	73
3.2	The systems	74
3.3	The effect of immerse graphene in solution	79
3.4	Polarization effects	81
3.4.1	Structural properties	81
3.4.2	QM/MM partitions	84
3.4.3	Electronic transport results	85
3.4.4	Convergence analysis	90
3.5	Effect of Stone-Wales defect on graphene	92
3.5.1	Structural properties	92
3.5.2	Electronic transport results	93
3.6	Free graphene analysis	95
3.7	Final analysis and conclusions	96
4	Graphene-based DNA sequencing: nanopore	100
4.1	Introduction	101
4.2	The systems	103
4.3	Electronic transport results	108
4.4	Sensitivity and selectivity of the device	115
4.5	Conclusions	116

5	Graphene-based DNA sequencing: nanogap	118
5.1	Introduction	119
5.2	The systems	120
5.3	Electronic transport results	127
5.4	Local currents	134
5.5	Sensitivity and selectivity of the device	141
5.6	Conclusions	143
6	Graphene-based DNA Chips	144
6.1	Introduction	145
6.2	DNA detection with anchor molecule	146
6.2.1	The systems	146
6.2.2	Electronic transport results	153
6.3	DNA detection without anchor molecule - vertical case	160
6.3.1	The systems	160
6.3.2	Electronic transport results	163
6.4	DNA detection without anchor molecule - horizontal case	167
6.4.1	The systems	167
6.4.2	Electronic transport results	170
6.4.3	Sensitivity and selectivity of the device	173
6.5	Conclusions	175
	Final conclusions and perspectives	176
	Bibliography	178
	Appendices	188
A	QM/MM partition test	190
A.1	The QM/MM partitions	191
A.2	Electronic transport results	193

CHAPTER 1	
	General introduction

1.1 Nanodevices

Nanotechnology is an important area of research with applications in health science, computing, engineering, chemistry and physics. One of the possible applications of nanotechnology is the development of novel electronic devices. From the point of view of Physics, electronic transport properties, thermal conductivity and optical phenomena must be understood for the development of new more effective devices.

One particular area where great advances could be achieved is the one of biosensors. Biosensors [1] are devices with a very small active sensing region (order of magnitude of 10^{-9} m) capable of detecting biological materials. Biosensors use a biological recognition element, typically a probe molecule that specifically binds to the target molecule one wishes to detect. There are different techniques for the detection and quantification of biomolecules, as biosensors that use an electrochemical substance, or optical, thermal and mass change mechanisms [1]. As biomolecules usually have nanoscopic dimensions, devices at this scale could yield higher sensitivity, down to the single molecule level [2,3].

The basic components of a biosensor are a sensitive biological element, a transducer and the signal processors. The sensitive biological element defines what kind of biomolecule is going to be detected, the transducer transforms the signal into one that can be measured and somehow turned into a meaningful result. Figure 1.1 shows the basic components of a biosensor.

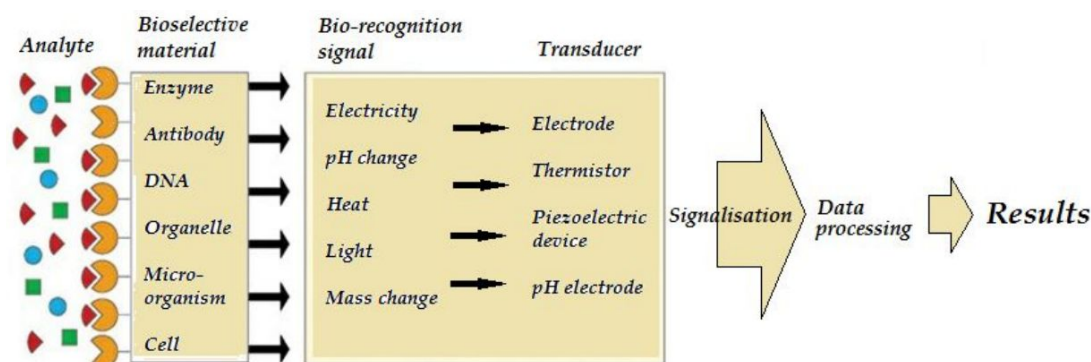


Figure 1.1: General features of a biosensor showing its basic components: biological element, transducer and signal processors. The analyte is detected according to the used bioselective material, then the bio-recognition signal is transformed in the transducer into one that can be measured to obtain the final results. Reprinted from reference [1]. Open access.

1.2 Deoxyribonucleic acid - DNA

deoxyribonucleic acid (DNA) [4] is one of the most studied molecules in biochemistry. The most important role played by DNA is the storage of all the genetic information of an organism. From the DNA, the proteins are synthesized, which are the executors of most biological processes crucial to sustaining life as we know it, such as signal transduction, ob-

taining energy, nerve transmission, etc [4].

DNA is formed by the polymerization of other smaller units, called nucleotides (Figure 1.2), whose composition is a sugar molecule (deoxyribose), a phosphate group and a nitrogenous base, that in the case of DNA may be four: adenine (A), thymine (T), cytosine (C) and guanine (G). The sequence of these bases determines the genetic information, where three nitrogenous bases form a unit called a codon [4]. A representation of each nucleotide can be seen in Figure 1.2. The letters A, T, C and G indicate the nucleobase.

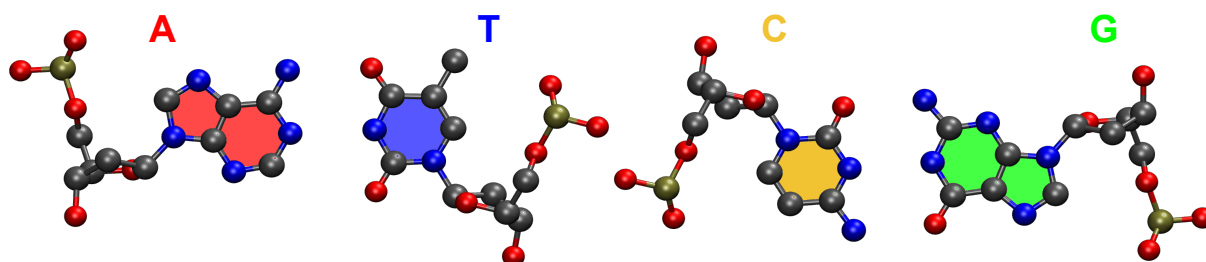


Figure 1.2: Four nucleotides of DNA (not showing Hydrogen atoms). The letters A, T, C and G indicates the nucleobases adenine, thymine, cytosine and guanine, respectively. The color-code chosen to indicate each base will be used throughout the results chapters of the thesis. Carbon in gray, Nitrogen in blue, oxygen in red, and phosphorus in golden. This color-code for the atoms will be used throughout all chapters of this thesis.

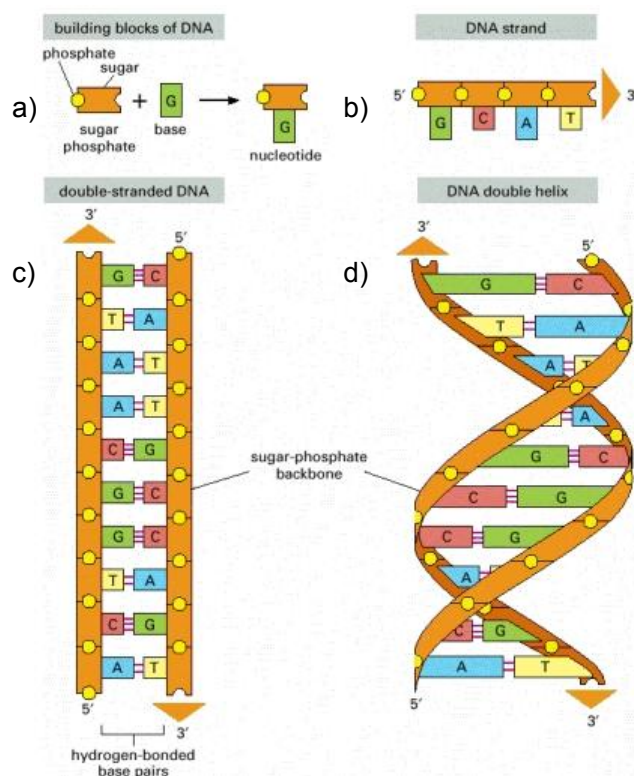


Figure 1.3: Schematic representation of DNA structure. The building (a) blocks of DNA are used to construct the DNA strand (b). An illustration in plane (c) of a double-stranded DNA and the DNA double helix (d). Adapted from reference [4]. Open access.

Figure 1.3 shows the structure of DNA. The nucleobases couple forming a strand via the phosphate group. A DNA double strand is then formed by the interaction between complementary pairs of nucleotides: A binds with T, and C binds with G. These specific interactions are strong due to the hydrogen bonds formed between the pairs. As can be observed, the double stranded DNA (dsDNA) has a double helix structure formed by two strands that coil in opposite directions. The ends of a chain are named either 3' or 5', resulting in different polarities [4].

1.3 Graphene

Graphene consists of a single layer of carbon atoms in a hexagonal arrangement which allows dense packing and high strength. Although proposed in the last century, graphene was isolated for the first time in 2004 by Geim and Novoselov [5]. Since then, several applications have been proposed. An experimental picture for graphene (1, 2 and 3 mono layers) in different scales is presented in Figure 1.4.

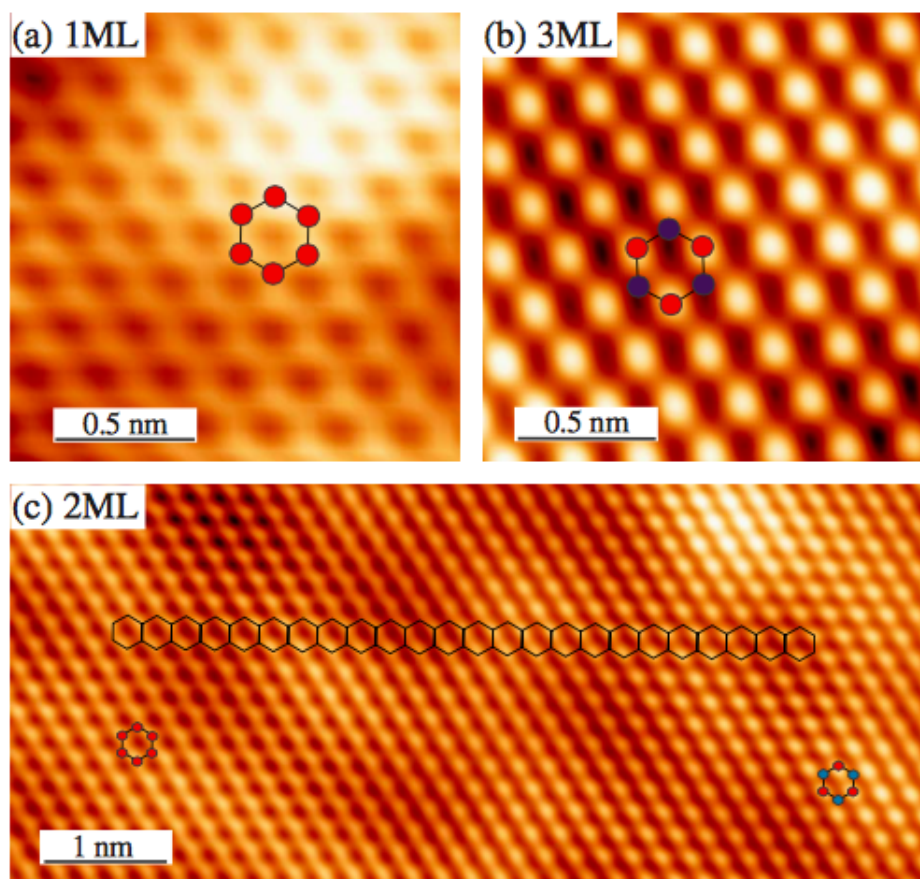


Figure 1.4: Scanning tunneling microscopy (STM) images for films of graphene with different thicknesses: (a) 1 mono layer, (b) 3 mono layers and (c) 2 mono layers. The colors shows equivalent and inequivalent atoms in the different layers. Reprinted with permission from reference [6]. Copyright © (2008) American Physical Society.

Several features [7–9] of graphene (Figure 1.4) - optical, electronic, thermal and mechanical properties, as well as its high surface area - make it a promising material for the active sensing region in biosensors based on different physical phenomena [2, 10–15]. There are numerous experimental works aimed at developing new devices for the detection of biomolecules, but very limited theoretical insight at the atomic level. In our case, the detection is based on electronic transport calculations [16–19], which is equivalent to the transmission probability of the electron passing through the graphene sheet in the Landauer [20] picture. In this way, the detection is related to the presence of the biomolecule which, in principle, locally alters the electronic transport properties of graphene. There are many different ways of interaction between graphene and the biomolecule that can be detected.

1.4 DNA detection and sequencing

DNA sensing is a very active and important area of research. By DNA detection many illnesses can be diagnosed [21], like some cancer types and Hepatitis, just to cite a few examples. DNA detection and sequencing could have a big impact in precision medicine [22] where personalized treatments can be prescribed to different individuals.

Accurate determination of nucleotide sequences is crucial for preventive detection of various diseases with a genetic predisposition. While full genome sequencing would give, in principle, complete information about an individual, it is still somewhat too costly to be part of regular every-day medicine. Since 2001, the cost per genome of sequencing DNA is decreasing substantially but nowadays is still fluctuating around \$ 1000,00, as can be observed in Figure 1.5.

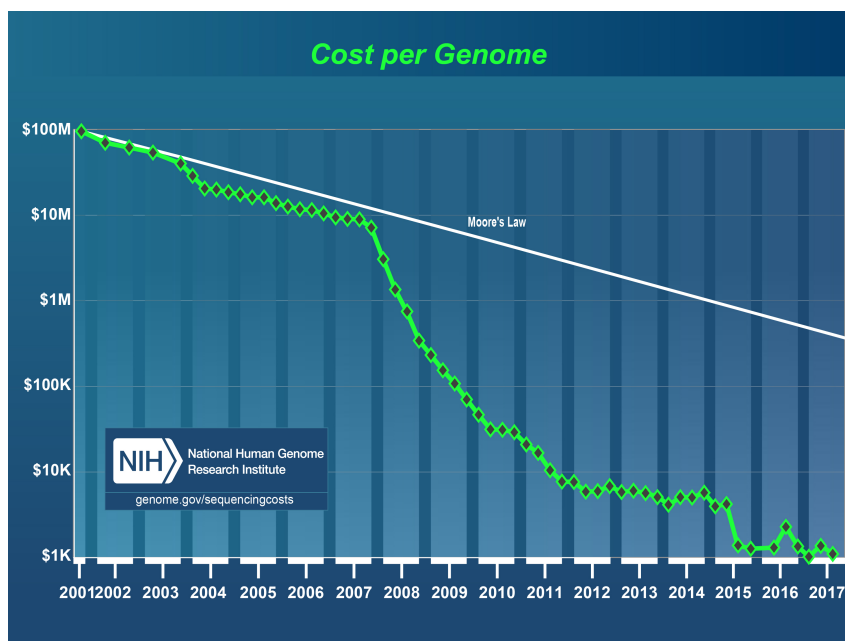


Figure 1.5: Cost per genome of sequencing DNA as a function of the years. Reprinted from reference [23]. Access in may 28, 2018.

An alternative would be to use DNA segments with a specific sequence indicating, for instance, mutation related to a particular illness. These strands can be used as the probe in a biosensor, as we show in Figure 1.1. If the DNA from an individual is dropped on the sensor it could be possible to determine whether that person has a specific corresponding mutation. The question is whether it is possible to design a highly sensitive all-electronic sensor. In this manner, genetic diseases can be detected through fast, cheap and simple DNA analysis.

The DNA sequencing devices aims to detect the sequence of nucleotides in the sample to be analyzed. In the literature [24, 25] there are many theoretical and experimental works proposing new devices for this end. Figure 1.6 shows the device in which a biological pore [25] is used. In this case, the DNA strand is translocated through the pore and its movement is controlled by an electric field. The detection is, thus, based on the ionic current changes. These pores (pink atoms in Figures 1.6b and 1.6c) are placed above biological membranes (lipid bilayer - green atoms in Figures 1.6b and 1.6c). The figure also shows a graphene nanopore.

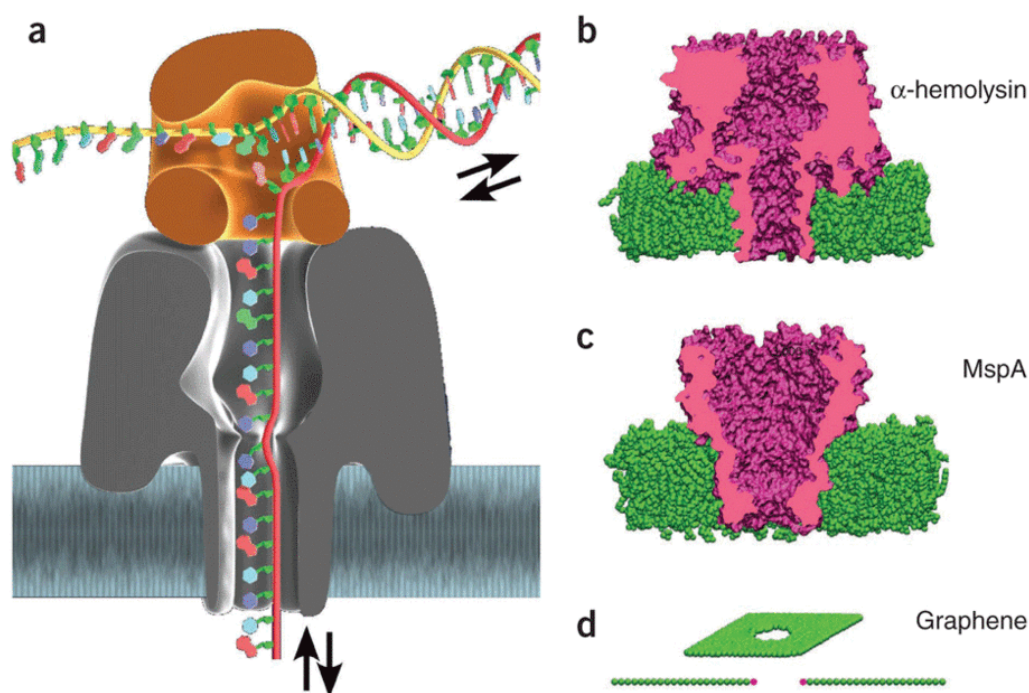


Figure 1.6: DNA sequencing device based on biological pores. The detection is based on ionic current changes and the movement of the DNA inside the pore is controlled by an electric field (a). Structures of α -hemolysin (b) and MspA (c) inserted in lipid bilayer and graphene nanopore (d). Reprinted with permission from reference [25]. Copyright © (2012) Nature Publishing Group.

Another sort of devices for DNA sequencing is based on nanopores and nanogaps in graphene [24]. In these cases, the DNA strand is translocated through the device providing a different electric signal (all-electronic devices) according to the nucleotide which is present in the nanopore/nanogap at a time. Modified graphene is an excellent choice in this sort of device due to its interesting properties and mostly because it could provide single molecule

resolution, since the probe can directly interact with the sheet.

As described in reference [24], there are numerous ways for sequencing DNA using graphene reported in the literature. Some proposed devices are based on ionic current detection using nanopores in graphene (Figure 1.7a), optical detection due to the adsorption of DNA on the sheet or based on electronic changes in devices based on nanogaps (Figure 1.7b), nanopores (Figure 1.7c) or nanoribbons (Figure 1.7d). Thus, graphene can be used for detecting each DNA nucleotide by translocating the strands through the pore (or gap) or by physisorption.

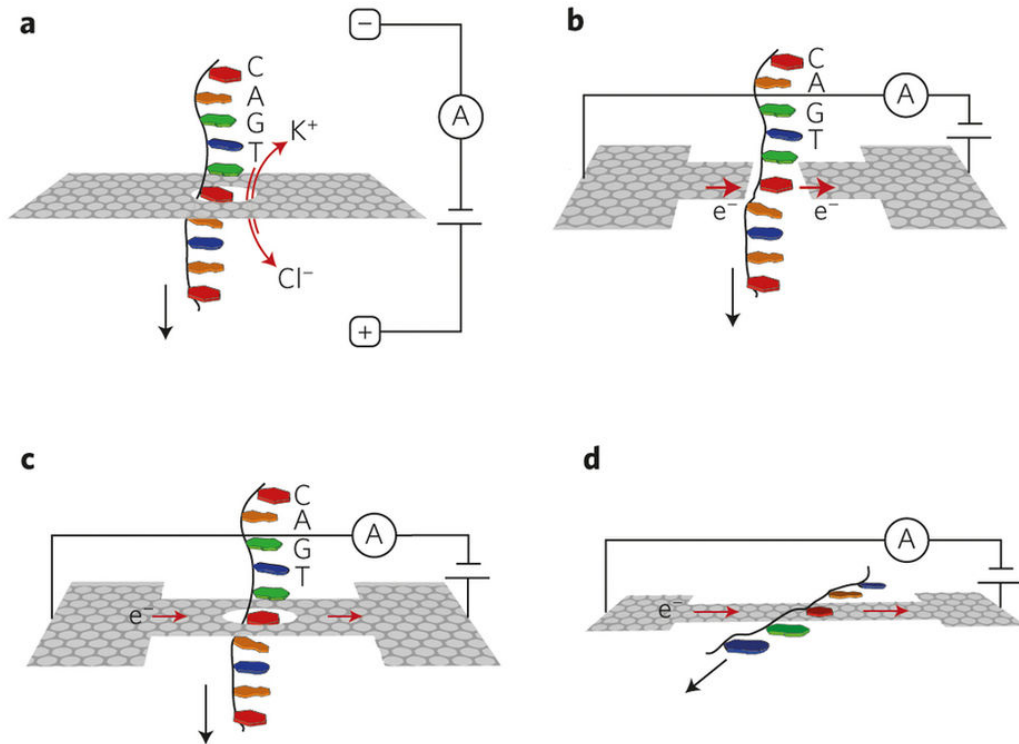


Figure 1.7: Graphene-based devices for DNA sequencing. A nanopore (a) can be used for detecting each nucleotide based on ionic current changes. Nanogaps (b) and nanopores (c) can be used for detection based on electronic changes. Another possibility is the detection based on physisorption process (d). Reprinted with permission from reference [24]. Copyright © (2016) Nature Publishing Group.

The idea for both "detection" and "sequencing" terms, in many ways, is the same, e.g. sensing DNA. We use different terms here because by DNA detection we mean determining specific DNA strands, which are known *a priori*. This is done by using a probe containing a known sequence. Thus, the objective is detecting hybridization (single stranded DNA (ssDNA) forming the complementary dsDNA). By DNA sequencing we mean determining the specific sequences of nucleotides in a segment of DNA for which there is no prior information. This nomenclature will be used throughout this thesis to distinguish the devices for DNA detection and sequencing. The main idea behind devices for DNA sequencing and detection that we studied in this thesis are sketched in Figure 1.8.

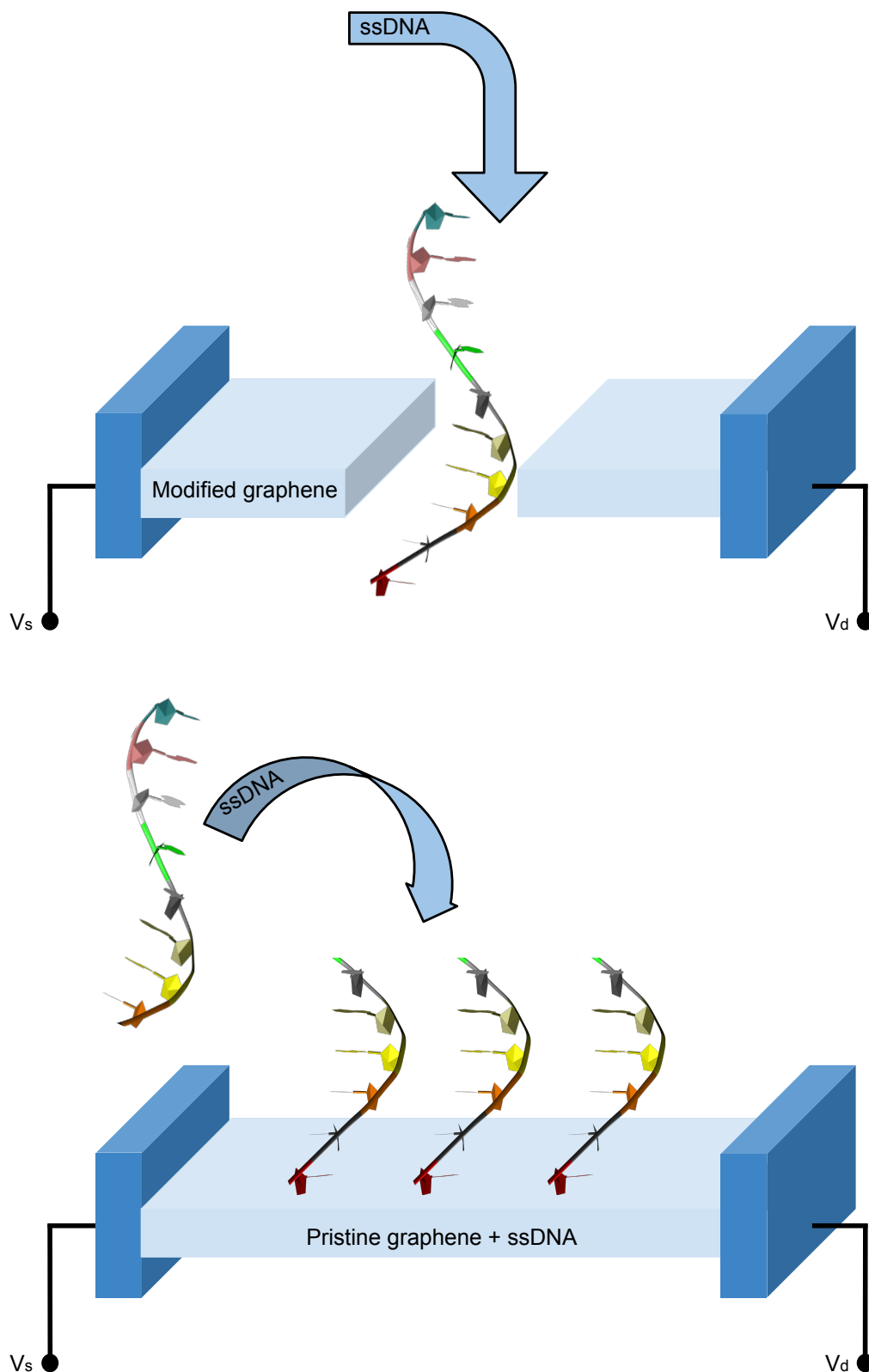


Figure 1.8: Sketch for the DNA sensing device. The device is connected to a source and drain electrodes. A modified graphene (nanopore or nanogap; only nanogap example shown here) can be used to sequence DNA in such a way that a ssDNA is translocated through the device (upper panel). The DNA detection device can be constructed using pristine graphene with ssDNA attached to the surface. Thus, the ssDNA sample is added to the device and, if this added strand is the complementary DNA strand, the hybridization (and the detection) is reached.

1.5 The role of the solvent

The solvent plays a central role on most biological systems. In fact, DNA changes its structure depending on the solution [26]. Furthermore, graphene fabrication is affected by the used solvent as reported by a number of theoretical and experimental works [27–31] also having environmental importance [32]. Therefore, solvent effects must be better understood in these systems.

From the theoretical point of view, solvent effects is frequently neglected. This occurs because of many reasons, like high computational costs, limitations in the used methods, or just to simplify the problem. However, most DNA samples, for example, are immersed in liquid (blood samples for instance). Thus the solvent effect must be taken into account in simulations to investigate biological systems.

There are different ways to simulate the presence of liquid in theoretical investigations of systems in condensed matter. Continuum models [33] are popular in chemistry to include the effect of the solvent in chemical reactions. In those cases a continuum of classical charges are included creating an electrostatic potential that changes depending on the interaction with the system. When we deal with larger systems, coarse-grained models [34] can be used to eliminate degrees of freedom in the liquid that are not essential to describe its influence.

One can also investigate systems in condensed matter using molecular dynamics (MD) [35] calculations. In this case, all particles composing the system are treated via a classical parametrization. However, in a number of cases we are interested in quantum phenomena which are not possible to be analyzed via purely classical calculations. density functional theory (DFT) [36, 37] calculations can be used in these cases, but the size of the systems is a limitation for a purely quantum mechanical description as we aim for an *ab initio* description of a quantum mechanical problem of many particles.

Another way to include the effect of the solvent in theoretical investigations is the one that we have used: the QM/MM [38, 39] method. This method, as we will describe later, combines Quantum Mechanics and Classical Mechanics, in such a way that large systems including thousands of water molecules can be divided into a quantum and a classical part, making possible theoretical investigations including the solvent effect in large systems, also taking into account explicit solvent molecules.

1.6 Challenges

Computational physics represents an important tool in materials research. From the predictive point of view, the computational contribution can provide the development of new devices with optimized performance. From computational physics point of view it is also possible to study the mechanisms of electronic transport in electronic device, which is usu-

ally unknown. Thus, the contribution of theoretical studies in this area may reveal electronic transport mechanisms through nanodevices as well as providing the development of more sensitive and selective devices.

However, the development of theoretical models to describe the electronic transport properties of biomolecules interacting with graphene is extremely challenging. The systems used to simulate devices for detecting biomolecules involve thousands of atoms and, therefore, it is not easily possible to investigate them using purely quantum mechanical calculations. Nevertheless, a quantum description of the system is required to study the electron transport as our ultimate aim is to design electronic devices.

A widely used technique in theoretical biochemistry for simulating systems with thousand of atoms which require a quantum description is the combination of quantum mechanical (QM) methods with molecular mechanical (MM) methods, the so called quantum mechanics/molecular mechanics (QM/MM) procedure, as we mentioned before. In this technique, one must determine the part of the system which will be treated quantum mechanically and all the rest will be treated classically as an external potential acting on the quantum mechanical part. Recent work has shown that this methodology can be coupled to electronic transport method [16], more specifically the non-equilibrium Greens's functions (NEGF) formalism to theoretically investigate electronic transport of solvated systems.

1.7 Objective

Thus, in this work our objective was to theoretically investigate the electronic transport properties of graphene-based devices for detecting DNA using QM/MM coupled to NEGF methods to simulate these systems. Firstly we have studied the effect of the solvent on the electronic transport properties of graphene. Then we have investigated DNA sequencing using nanopore in between line defects in graphene, DNA sequencing using Nitrogen-terminated nanogap in graphene and finally DNA detection based on physisorption using pristine graphene.

In this thesis we firstly made a general introduction, presented in Chapter 1. Then we present the used methodology and the theory foundations in Chapter 2. The results are presented in the Chapters 3, 4, 5 and 6. In chapter 3 we present the results regarding the effect of the solvent on the electronic transport properties of graphene, in chapter 4 we present the results of DNA sequencing using topological line defects around graphene nanopore, in chapter 5 we present the results of DNA sequencing using Nitrogen-terminated nanogap in graphene and in chapter 6 we present the results of DNA detection using pristine graphene. We, thus, present the conclusions and an appendix with a QM/MM partition test.

CHAPTER 2

Methodology and theory

2.1 How to simulate a nanodevice

Realistically simulate a nanodevice designed to detect molecular systems, like DNA, can quickly become a challenging work as system size and complexity increases. In order to obtain detection at the molecular level, we need to simulate these devices using atomistic and molecular methods. In the case of all-electronic devices in which the detection is directly related with the kinetics of the electrons in the system a Quantum Mechanical (QM) treatment is required. However, if on one hand we need a QM treatment, on the other hand the size of these kinds of systems, and the amount of calculations needed to have statistics is a limitation.

All nanodevices simulated by us are based in electronic changes in the material (graphene) used to detect DNA. The DNA systems can be adsorbed on graphene sheet or placed inside a nanogap/nanopore in the sheet. In all these cases, the presence of DNA is supposed to change the electronic properties of graphene-like materials. As we discussed in the Introduction (Chapter 1) the solvent plays a very important role in these systems and we have to take its effect into account. Furthermore one needs to go beyond electronic structure calculations and calculate the electronic transport properties. Below we lay out the theoretical framework for simulating such systems.

The Schrödinger Equation

Since we deal with atomic and molecular systems, from the most fundamental (atomic scale) point of view, simulating these systems means solving the Schrödinger equation (SE) for the system, shown in its time dependent version in Equation 2.1:

$$i\hbar \frac{\partial}{\partial t} |\Psi(\vec{x}, t)\rangle = \hat{H} |\Psi(\vec{x}, t)\rangle, \quad (2.1)$$

where $\hbar$ is the reduced Planck constant, $|\Psi(\vec{x}, t)\rangle$ is the time dependent state and $\hat{H}$ is the Hamiltonian operator¹. If we consider time independent Hamiltonians we can obtain the stationary solutions of the problem using the time independent version of the SE:

$$\hat{H} |\Psi(\vec{x})\rangle = E |\Psi(\vec{x})\rangle, \quad (2.2)$$

where $|\Psi(\vec{x})\rangle$ is the time independent state, and E is the energy eigenvalue of the Hamiltonian operator. The Hamiltonian operator has two components:

$$\hat{H} = \hat{T} + \hat{V}. \quad (2.3)$$

¹The coordinate $\vec{x}$ is a $3(N+P)$ generalized vector that contains the coordinates of N electrons and P nuclei.

The Hamiltonian shown in Equation 2.3 is a many body operator defined by the kinetic energy operator, $\hat{T}$, and potential energy operator, $\hat{V}$, defined as:

$$\hat{T} = - \underbrace{\sum_i^N \frac{\hbar^2}{2m_e} \nabla_i^2}_{\text{Electron's Kinetic energy}} - \underbrace{\sum_I^P \frac{\hbar^2}{2m_I} \nabla_I^2}_{\text{Nuclei Kinetic energy}}, \quad (2.4)$$

$$\hat{V} = \underbrace{- \sum_I^P \sum_i^N \frac{Z_I e^2}{|\vec{r}_i - \vec{R}_I|}}_{\text{Coulomb Interaction between electrons and nuclei}} + \underbrace{\sum_j^N \sum_{i>j}^N \frac{e^2}{|\vec{r}_i - \vec{r}_j|}}_{\text{Coulomb Interaction between electrons}} + \underbrace{\sum_I^P \sum_{J>I}^P \frac{Z_I Z_J e^2}{|\vec{R}_I - \vec{R}_J|}}_{\text{Coulomb Interaction between nuclei}}, \quad (2.5)$$

where $\nabla_{i/I}^2$ is the Laplacian operator for the i -th (I -th) electron (nuclei), e is the elementary charge, m_e is the mass of the electron and m_I , $Z_{I/J}$ is the mass and atomic number of the i -th ion, respectively. Then, the many body Hamiltonian of the system can be rewritten as:

$$\hat{H} = \hat{T}_{Nuc}(\vec{R}) + \hat{T}_e(\vec{r}) + \hat{V}_{eNuc}(\vec{r}, \vec{R}) + \hat{V}_{NucNuc}(\vec{R}) + \hat{V}_{ee}(\vec{r}), \quad (2.6)$$

where Nuc refers to nuclei, e refers to electrons, $\vec{R} = \{R_1, R_2, \dots, R_P\}$ are the generalized coordinates for nuclei and $\vec{r} = \{r_1, r_2, \dots, r_N\}$ are the generalized coordinates for electrons.

An *ansatz* for the wave function can be a combination of electronic and nuclear wave functions:

$$\Psi(\vec{R}, \vec{r}, t) = \sum_n \phi_n(\vec{R}; \vec{r}) \Theta_n(\vec{R}, t), \quad (2.7)$$

where $\Theta_n(\vec{R}, t)$ are the nuclear wave functions and $\phi_n(\vec{R}; \vec{r})$ are the electronic wave functions, which parametrically depend on the nuclei coordinates ($\vec{R}$) and satisfies:

$$\hat{H}_{el} \phi_n(\vec{R}; \vec{r}) = E_{el} \phi_n(\vec{R}; \vec{r}), \quad (2.8)$$

where $\hat{H}_{el}$ is the electronic Hamiltonian:

$$\hat{H}_{el} = \hat{T}_e(\vec{r}) + \hat{V}_{eNuc}(\vec{r}; \vec{R}) + \hat{V}_{NucNuc}(\vec{R}) + \hat{V}_{ee}(\vec{r}). \quad (2.9)$$

The Hamiltonian in Equation 2.6 can be rewritten as:

$$\hat{H} = \hat{T}_{Nuc} + \hat{H}_{el}, \quad (2.10)$$

where $\hat{T}_{Nuc}$ refers to the nuclear Hamiltonian (only kinetic energy of nuclei) and $\hat{H}_{el}$ refers to the electronic Hamiltonian (with the nuclear repulsion included^{II}). The *ansatz* wave function showed in Equation 2.7 should satisfy:

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{R}, \vec{r}, t) = \hat{H} \Psi(\vec{R}, \vec{r}, t) . \quad (2.11)$$

Then, replacing Equation 2.7 for n'' and n' (orthonormal states) in Equation 2.11, we have:

$$i\hbar \frac{\partial}{\partial t} \sum_{n''} |\phi_{n''}(\vec{R}; \vec{r}) \Theta_{n''}(\vec{R}, t)\rangle = \hat{H} \sum_{n'} |\phi_{n'}(\vec{R}; \vec{r}) \Theta_{n'}(\vec{R}, t)\rangle . \quad (2.12)$$

We can project Equation 2.12 on the electronic state $\langle \phi_n(\vec{R}; \vec{r}) |$:

$$\sum_{n''} \langle \phi_n(\vec{R}; \vec{r}) | i\hbar \frac{\partial}{\partial t} |\phi_{n''}(\vec{R}; \vec{r}) \Theta_{n''}(\vec{R}, t)\rangle = \sum_{n'} \langle \phi_n(\vec{R}; \vec{r}) | \hat{H} |\phi_{n'}(\vec{R}; \vec{r}) \Theta_{n'}(\vec{R}, t)\rangle . \quad (2.13)$$

Since n and n'' are also orthonormal states, only the case $n = n''$ in the left side of Equation 2.13 is nonzero. Then, using Equation 2.10 we can rewrite this equation as:

$$i\hbar \frac{\partial}{\partial t} |\Theta_n(\vec{R}, t)\rangle = \sum_{n'} \langle \phi_n(\vec{R}; \vec{r}) | \hat{T}_N + \hat{H}_{el} | \phi_{n'}(\vec{R}; \vec{r}) \Theta_{n'}(\vec{R}, t)\rangle , \quad (2.14)$$

which in turn can be rewritten as:

$$i\hbar \frac{\partial}{\partial t} |\Theta_n(\vec{R}, t)\rangle = \sum_{n'} \{ \langle \phi_n(\vec{R}; \vec{r}) | \hat{T}_N | \phi_{n'}(\vec{R}; \vec{r}) \rangle + \langle \phi_n(\vec{R}; \vec{r}) | \hat{H}_{el} | \phi_{n'}(\vec{R}; \vec{r}) \rangle \} |\Theta_{n'}(\vec{R}, t)\rangle . \quad (2.15)$$

The second term on the right hand side of Equation 2.15 is the electronic energy. Then, moving E_{el} to the left hand side, we can write:

$$\left(i\hbar \frac{\partial}{\partial t} - E_{el} \right) |\Theta_n(\vec{R}, t)\rangle = \sum_{n'} \langle \phi_n(\vec{R}; \vec{r}) | \hat{T}_N | \phi_{n'}(\vec{R}; \vec{r}) \Theta_{n'}(\vec{R}, t)\rangle . \quad (2.16)$$

^{II}Note that the inclusion of nuclei repulsion contribution in the electronic Hamiltonian is an arbitrary choice.

Substituting the nuclear kinetic energy operator, we have:

$$\left(i\hbar\frac{\partial}{\partial t} - E_{el}\right)|\Theta_n(\vec{R}, t)\rangle = \sum_{n'} \langle\phi_n(\vec{R}; \vec{r})| - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 |\phi_{n'}(\vec{R}; \vec{r})\rangle \Theta_{n'}(\vec{R}, t) \rangle . \quad (2.17)$$

We can apply the second derivative in two steps,

$$\begin{aligned} \left(i\hbar\frac{\partial}{\partial t} - E_{el}\right)|\Theta_n(\vec{R}, t)\rangle = \\ - \sum_I \frac{\hbar^2}{2M_I} \sum_{n'} \langle\phi_n(\vec{R}; \vec{r})| \nabla_I \{ \phi_{n'}(\vec{R}; \vec{r}) \nabla_I \Theta_{n'}(\vec{R}, t) + \Theta_{n'}(\vec{R}, t) \nabla_I \phi_{n'}(\vec{R}; \vec{r}) \} \rangle , \end{aligned} \quad (2.18)$$

and

$$\begin{aligned} \left(i\hbar\frac{\partial}{\partial t} - E_{el}\right)|\Theta_n(\vec{R}, t)\rangle = - \sum_I \frac{\hbar^2}{2M_I} \sum_{n'} \{ 2 \langle\phi_n(\vec{R}; \vec{r})| \nabla_I |\phi_{n'}(\vec{R}; \vec{r})\rangle |\nabla_I \Theta_{n'}(\vec{R}, t)\rangle + \\ \langle\phi_n(\vec{R}; \vec{r})| \nabla_I^2 |\phi_{n'}(\vec{R}; \vec{r})\rangle |\Theta_{n'}(\vec{R}, t)\rangle + \nabla_I^2 |\Theta_n(\vec{R}, t)\rangle \} . \end{aligned} \quad (2.19)$$

Equation 2.19 can be rearranged into a Schrödinger-like equation for $|\Theta_n(\vec{R}, t)\rangle$:

$$\begin{aligned} \left(i\hbar\frac{\partial}{\partial t} - E_{el} + \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2\right)|\Theta_n(\vec{R}, t)\rangle = \\ - \sum_I \frac{\hbar^2}{2M_I} \sum_{n'} \{ 2 \langle\phi_n(\vec{R}; \vec{r})| \nabla_I |\phi_{n'}(\vec{R}; \vec{r})\rangle |\nabla_I \Theta_{n'}(\vec{R}, t)\rangle + \\ \langle\phi_n(\vec{R}; \vec{r})| \nabla_I^2 |\phi_{n'}(\vec{R}; \vec{r})\rangle |\Theta_{n'}(\vec{R}, t)\rangle \} . \end{aligned} \quad (2.20)$$

Instead of a linear combination of functions, as proposed in Equation 2.7, we can use a reduced solution for $n' = n$:

$$\Psi(\vec{R}, \vec{r}, t) = \phi_n(\vec{R}; \vec{r}) \Theta_n(\vec{R}, t) . \quad (2.21)$$

This is not an exact solution for this problem because even if the system is initially in this state, the non-diagonal terms in Equation 2.20 would mix the electronic states (inducing electronic excitations) during the time evolution. However, neglecting these non-diagonal

terms means that the time evolution of the nuclei does not induce electronic excitations.

The Born-Oppenheimer approximation

The nuclei can evolve in time in adiabatic and non-adiabatic dynamics. The non-adiabatic case does not neglect the non-diagonal terms in Equation 2.20, and in this case the nuclei time evolution can generate electronic excitations (changes from one potential energy surface to another). In the adiabatic case the nuclei time evolution does not cause electronic excitations and, then, the non-diagonal terms in Equation 2.20 can be neglected. In this case, the Equation 2.20 can be rewritten as:

$$i\hbar \frac{\partial}{\partial t} |\Theta_n(\vec{R}, t)\rangle = \left(E'_{el} - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 \right) |\Theta_n(\vec{R}, t)\rangle, \quad (2.22)$$

where

$$E'_{el} = E_{el} + \sum_I \frac{\hbar^2}{2M_I} \langle \phi_n(\vec{R}; \vec{r}) | \nabla_I^2 | \phi_n(\vec{R}; \vec{r}) \rangle. \quad (2.23)$$

The second term in Equation 2.23 is a diagonal correction to the Kinetic energy. If this term is also neglected we have the so-called Born-Oppenheimer (BO) [40] approximation. In the BO approximation the nuclei and electrons are solved separately. This approximation is corroborated by the fact that nuclei are heavier than the electrons and, because of that, their movement is much slower than the electrons' movements. In this way, one considers that they instantaneously move in an external field generated by the nuclei.

As we mentioned, within the BO approximation the electronic and nuclear degrees of freedom are decoupled, as they can be solved separately. The nuclei kinetics will be discussed using Molecular dynamics (MD) in Section 2.2.1. Thus the point now is solving the Schrödinger equation for the electronic problem,

$$\hat{H}_{el} |\phi_n(\vec{R}; \vec{r})\rangle = E_{el} |\phi_n(\vec{R}; \vec{r})\rangle, \quad (2.24)$$

where E_{el} is the energy eigenvalue.

There are a number of different methodologies to solve the electronic problem. Some of them are based on wave function, like variational methods, perturbation theory and the Hartree-Fock (HF) method. Another class of method is based on a fundamentally different approach where the quantity of interest is the electronic density, instead of wave function. This is the basic principle of density functional theory (DFT) [36, 37, 41, 42].

In this work we have applied classical MD to simulate the time evolution of the nuclei and DFT to solve the electronic structure using structures obtained through MD. Therefore,

these methods will be explained in some details in the next section.

2.2 Methods

In the last section we summarized the first approximation to simulate nanodevices - or molecular systems in general - from the most fundamental point of view in atomic scale. Since we have decoupled the nuclear and electronic degrees of freedom now we can use different methods to simulate them. In this section we will briefly summarize these methods in more details.

2.2.1 Molecular Dynamics - MD

Classical Molecular Dynamics (MD) calculations are very useful to simulate large systems ($N_{atoms} > 10^3$) in atomic scale for larger time scales. In this method, classical atoms can interact with each other via a potential that should be defined for each atom, usually called force field (FF) [35, 43]. The evolution of the system is obtained by numerically solving Newton's equations of motion.

The Newton's equations of motion and Force Fields

From Newton's second law, we know that the force $\vec{F}_i$ in a particle (atom) i with mass m_i and coordinate $\vec{r}_i$ is defined, for a system with N particles, by:

$$\vec{F}_i = m_i \frac{\partial^2 \vec{r}_i}{\partial t^2}, \quad i = 1, 2, \dots, N. \quad (2.25)$$

That can be written in terms of the gradient of a scalar field, in this case the potential, V , which defines how the atoms interact between them:

$$\vec{F}_i = - \frac{\partial V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)}{\partial \vec{r}_i}. \quad (2.26)$$

This generates a set of coupled equations that need to be solved. The key point here is to obtain the potential V . In classical MD a typical potential, or FF, includes bonded and non-bonded terms, and can be defined by:

$$V(\vec{r}) = \sum_{bond} K_i^b (b_i - b_0)^2 + \sum_{ang.} K_i^\theta (\theta_i - \theta_0)^2 + \sum_{dihe.} K_i^{dihe.} (1 + \cos[n_i \varphi_i - \varphi_0]) + \sum_{n. bond} \sum_{jk} \frac{A_{jk}}{\vec{r}_{jk}^{12}} - \frac{B_{jk}}{\vec{r}_{jk}^6} + \frac{q_j q_k}{\epsilon \vec{r}_{jk}}, \quad (2.27)$$

where K_i^b is the force constant for the bond, b_i is the bond distance and b_0 is the equilibrium distance for the bond. K_i^θ is the force constant for the angle, θ_i is the angle and θ_0 is the equilibrium angle. $K_i^{dihe.}$ is the dihedral constant, n_i is the multiplicity, φ_i is the dihedral angle and φ_0 is lowest equilibrium angle. A_{jk} and B_{jk} are parameters which depends on the atom pair defining the Lennard-Jones potential, q_j and q_k are the charges of the pairs j, k and $\vec{r}_{jk}$ is the distance between the pair [35].

The first three terms in Equation 2.27 are bonded terms and the last one is a non-bonded term. The first term is a summation over all bonds of the system, the second is a summation over all angles, the third is a summation over all dihedral and the last term is a summation over all pairs (Lennard-Jones and Coulomb potentials). One example for each of these terms is shown for Adenine DNA nucleotide in Figure 2.1.

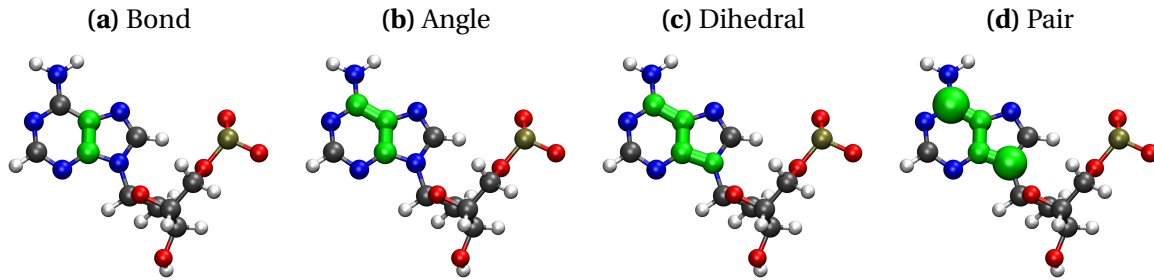


Figure 2.1: Exemplification of force field terms for Adenine DNA nucleotide. The green atoms/bonds exemplifies each term: (a) bond, (b) angle and (c) dihedral. The large radius spheres in (d) exemplifies the pair defined by the two highlighted atoms.

Typically these parameters are obtained from QM methods for small systems or are filled to replicate some experimental property. In the simulations presented in this thesis we have used the AMBER99SB [44] FF and a simple point-charge (SPC) [45, 46] model for water. The SPC model is a three-site model that includes three point charges at each of the three sites. In this model the water molecules move rigidly and interacts via Coulomb and Lennard-Jones potentials, the last one only at the oxygen. This model is widely used in calculations and reported in the literature due to its efficiency on simulations.

Statistical Mechanics

Within MD we can simulate microscopic systems by solving the Newton's equations of motion for all atoms which compose it, as we showed before. The next step is the extraction

of observables from this MD. In this way, statistical mechanics (SM) can connect microscopic simulations and macroscopic properties (observables).

In SM the observables are calculated as ensemble averages, as we can see in Equation 2.28 for a generic observable called A :

$$\langle A \rangle_{ensemble} = \iint d\vec{p}^N d\vec{r}^N A(\vec{p}^N, \vec{r}^N) \rho(\vec{p}^N, \vec{r}^N) , \quad (2.28)$$

where $\vec{p}$ and $\vec{r}$ are respectively the momentum and position of the particles, N is the number of particles and ρ is the probability density function of the observable A . In MD calculations, on the other hand, the observables are calculated as time averages, as we can see in Equation 2.29:

$$\langle A \rangle_{time} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\vec{p}^N(t), \vec{r}^N(t)) dt \approx \frac{1}{M} \sum_{t=1}^M A(\vec{p}^N, \vec{r}^N) . \quad (2.29)$$

The idea behind MD simulations, which calculates averages over (limited) time, is the ergodic hypothesis [43], which states that for a long time simulation all the micro-states have the same probability. This means that in this scenario the ensemble and time averages are the same:

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time} . \quad (2.30)$$

Ensembles

The ensembles can be understood as the probability distribution for a set of micro states that are replicated to describe the system. To simulate dynamical systems we have to define the ensemble that will be used. In this thesis we have used first the isothermal-isobaric ensemble (NPT) and then the canonical ensemble (NVT) [47, 48].

In both NPT and NVT ensembles the number of particles is constant. In NPT ensemble a barostat and a thermostat are used to control the pressure and temperature of the system, respectively, and here the system's volume can change [47, 48]. In the NVT ensemble a thermostat is also used to control the system's temperature but, on the other hand, the system's volume is fixed. Each ensemble should be chosen according to the system/property to be simulated.

Integration Algorithms

The numerical methods to integrate Newton's equations of motion on computers are applied using integration algorithms. In this way, there are many different algorithms that can be used. A Taylor's expansion for position, velocity and acceleration is shown in Equations 2.31, 2.32 and 2.33:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}(t)\delta t + \frac{1}{2}\vec{a}(t)\delta t^2 + \dots, \quad (2.31)$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \vec{a}(t)\delta t + \frac{1}{2}\vec{b}(t)\delta t^2 + \dots, \quad (2.32)$$

$$\vec{a}(t + \delta t) = \vec{a}(t) + \vec{b}(t)\delta t + \dots, \quad (2.33)$$

where $\vec{r}$, $\vec{v}$ and $\vec{a}$ are respectively the position, velocity and acceleration vectors for a time $t + \delta t$. Integrators always consider a truncation in these Taylor's expansions, like the *Leap-Frog* integrator (used in our calculations). In this integration algorithm the position at a time $t + \delta t$ is given by:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}\left(t + \frac{1}{2}\delta t\right)\delta t, \quad (2.34)$$

where the second term is the velocity at the time $t + \frac{1}{2}\delta t$ instead of $t + \delta t$, defined by:

$$\vec{v}\left(t + \frac{1}{2}\delta t\right) = \vec{v}\left(t - \frac{1}{2}\delta t\right) + \vec{a}(t)\delta t. \quad (2.35)$$

In this way, the velocity at time t is defined as an average over the velocities at half t before and half δt after δt :

$$\vec{v}(t) = \frac{1}{2} \left[\vec{v}\left(t - \frac{1}{2}\delta t\right) + \vec{v}\left(t + \frac{1}{2}\delta t\right) \right]. \quad (2.36)$$

The Verlet algorithm considers the position at $t \pm \delta t$:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}(t)\delta t + \frac{1}{2}\vec{a}(t)\delta t^2 + \dots, \quad (2.37)$$

$$\vec{r}(t - \delta t) = \vec{r}(t) - \vec{v}(t)\delta t + \frac{1}{2}\vec{a}(t)\delta t^2 + \dots \quad (2.38)$$

If we sum Equations 2.37 and 2.38 we obtain the Verlet algorithm at time $t + \delta t$:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \vec{a}(t)\delta t^2 + \dots \quad (2.39)$$

The Verlet/Velocity algorithm for position and velocity is defined by:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}(t)\delta t + \frac{1}{2}\vec{a}(t)\delta t^2, \quad (2.40)$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \frac{1}{2}[\vec{a}(t) + \vec{a}(t + \delta t)]\delta t. \quad (2.41)$$

The Molecular Dynamics protocol

The MD protocol is summarized in the flowchart of Figure 2.2. The initial conditions (potential, positions and velocities) for all atoms should be given. In these initial conditions the forces are calculated and then a new configuration is generated. Newton's Equations are solved, and then the required parameters (position, velocity, energies and etc.) are written in the output file. This cycle is repeated until the defined number of steps (simulation time, from ns to μs) is reached.

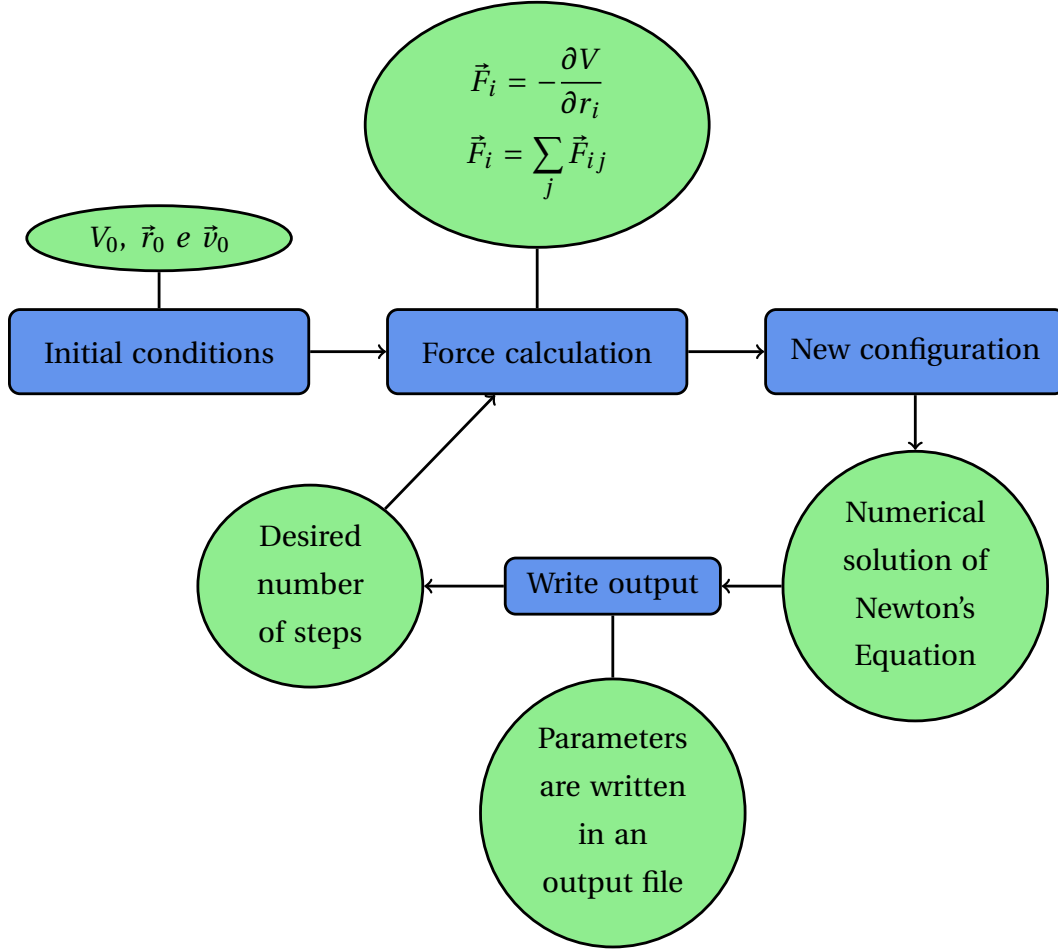


Figure 2.2: Flowchart for MD calculations.

However we need something beyond classical MD, since in this work, we want to simulate electronic transport properties for DNA detecting and sequencing. Since we deal with electrons, a QM treatment is required.

2.2.2 Density Functional Theory - DFT

The electronic density

While the dynamics of the nuclei for our purposes can be well described by a well parametrized FF, electronic transport requires a QM description at same level. There are a number of electronic structure methods that are based on the direct solution - to same level of approximation - for $|\phi_n(\vec{R}; \vec{r})\rangle$ ^{III}. Within such approaches we can cite Hartree Fock, perturbation theory methods, Configuration Interaction method (CI), Coupled Cluster method and etc [49]. All such methods require $3N$ position variables, and the computational requirement typically increases as same power or the number of particles (electrons), N . If this becomes possible, specially for large extended systems.

^{III}Given the anti-symmetry of the wave function the choice of variables for the integration is arbitrary.

At the heart of Density Functional Theory (DFT) is the fact that all calculations are based on the electronic density - $\rho(\vec{r})$ - of the system. In this scenario, DFT provides an enormous simplification of the electronic problem, in such a way that a system of N electrons is reduced into a three-dimensional problem, since the electronic density of the system is defined by:

$$\rho(\vec{r}) = N \int d\vec{r}_2 \cdots \int d\vec{r}_N \Psi^*(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) \Psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) , \quad (2.42)$$

where the integral is made over all $\vec{r}_1, \vec{r}_2 \dots \vec{r}_{N-1}$ vectors.

The Hohenberg and Kohn Theorems

DFT is based on the Hohenberg and Kohn theorems [41]. The first theorem states that the wave function and all the system's observables are uniquely obtained from the electronic density: $\rho(\vec{r})$. The second theorem is based on the variational principle which states that the energy as a functional of the electronic density is a threshold and always equal or higher than the ground state, i.e.

$$E_0 = E[\rho_o(\vec{r})] \leq E[\rho(\vec{r})] , \quad (2.43)$$

given the condition:

$$\int \rho(\vec{r}) d\vec{r} = N . \quad (2.44)$$

The first Hohenberg e Kohn theorem states that the electronic density of the ground state is determined unequivocally. We can prove this statement supposing exactly the opposite. Lets consider a non-degenerate ground state^{IV} and suppose that there are two external potentials $(\nu_{ext}^{(1)}(\vec{r}) \text{ and } \nu_{ext}^{(2)}(\vec{r}))$ that lead to the same density $\rho_0(\vec{r})$ for the ground state. In this manner, we have two Hamiltonians $(\hat{H}^{(1)} \text{ and } \hat{H}^{(2)})$ and two wave functions $(\psi^{(1)} \text{ and } \psi^{(2)})$.

From the variational principle, no wave function can have an energy lower than the one of $\psi^{(1)}$ (assuming that this one is the ground state wave function), that is:

$$E^{(1)} = \langle \psi^{(1)} | \hat{H}^{(1)} | \psi^{(1)} \rangle < \langle \psi^{(2)} | \hat{H}^{(1)} | \psi^{(2)} \rangle , \quad (2.45)$$

and

^{IV}This is also valid for degenerate states.

$$E^{(2)} = \langle \psi^{(2)} | \hat{H}^{(2)} | \psi^{(2)} \rangle < \langle \psi^{(1)} | \hat{H}^{(2)} | \psi^{(1)} \rangle , \quad (2.46)$$

where

$$\langle \psi^{(2)} | \hat{H}^{(1)} | \psi^{(2)} \rangle = \langle \psi^{(2)} | \hat{H}^{(2)} | \psi^{(2)} \rangle + \int d\vec{r} \left[v_{ext}^{(1)}(\vec{r}) - v_{ext}^{(2)}(\vec{r}) \right] \rho_0(\vec{r}) , \quad (2.47)$$

and

$$\langle \psi^{(1)} | \hat{H}^{(2)} | \psi^{(1)} \rangle = \langle \psi^{(1)} | \hat{H}^{(1)} | \psi^{(1)} \rangle + \int d\vec{r} \left[v_{ext}^{(2)}(\vec{r}) - v_{ext}^{(1)}(\vec{r}) \right] \rho_0(\vec{r}) . \quad (2.48)$$

This lead to

$$E^{(1)} + E^{(2)} < E^{(2)} + E^{(1)} . \quad (2.49)$$

As this is contradictory, it proves that there is a unique correspondence between v_{ext} and $\rho_0(\vec{r})$. Thus, the first Hohenberg e Kohn Theorem is proved.

The second Hohenberg e Kohn theorem states that there is a universal functional, $F[\rho(\vec{r})]$, which can be defined in terms of the electronic density $\rho(\vec{r})$ and that the exact ground state is a global minimum of this functional. The energy as a functional of the electronic density can be written as:

$$E[\rho(\vec{r})] = T[\rho(\vec{r})] + E_{int}[\rho(\vec{r})] + \int v_{ext}(\vec{r})\rho(\vec{r}) , \quad (2.50)$$

or

$$E[\rho(\vec{r})] \equiv F[\rho(\vec{r})] + \int v_{ext}(\vec{r})\rho(\vec{r}) , \quad (2.51)$$

and $F[\rho(\vec{r})]$ is an universal functional because kinetic and internal energies, $T[\rho(\vec{r})]$ and $E_{int}[\rho(\vec{r})]$, are the same for all systems. To prove the second Hohenberg e Kohn theorem, considering $\psi^{(1)}$ as the unique ground state, we can write the energy as a functional of the unique ground state electronic density:

$$E^{(1)} = E[\rho_0(\vec{r})] = \langle \psi^{(1)} | \hat{H} | \psi^{(1)} \rangle . \quad (2.52)$$

However, in the variational principle spirit, any different energy will lead to an energy higher than the one of the ground state:

$$\underbrace{\langle \psi^{(1)} | \hat{H} | \psi^{(1)} \rangle}_{E^{(1)}} < \underbrace{\langle \psi^{(2)} | \hat{H} | \psi^{(2)} \rangle}_{E^{(2)}} . \quad (2.53)$$

Energy as a functional of the electronic density

The way for obtaining the energy of the system as a functional of its electronic density was further introduced by Kohn-Sham [42]. One proposal of written the energy as a function of the electronic density in the picture of DFT is given by:

$$E[\rho(\vec{r})] = \underbrace{T_0[\rho(\vec{r})]}_{\text{Kinetic energy}} + \underbrace{V_{xc}[\rho(\vec{r})]}_{\text{Exchange and correlation energy}} + \underbrace{V_H[\rho(\vec{r})]}_{\text{Mean field between electrons}} + \underbrace{v_{ext}[\rho(\vec{r})]}_{\text{External potential due to nuclei}} . \quad (2.54)$$

One can minimize this equation under the condition (Lagrange multiplier) shown in Equation 2.57:

$$\frac{\delta}{\delta \rho(\vec{r})} \left[E[\rho(\vec{r})] - \mu \int \rho(\vec{r}) d\vec{r} \right] = 0 , \quad (2.55)$$

rewriting:

$$\frac{\delta E[\rho(\vec{r})]}{\delta \rho(\vec{r})} = \mu , \quad (2.56)$$

what leads to:

$$\frac{\delta T_0[\rho(\vec{r})]}{\delta \rho(\vec{r})} + v_{eff}(\vec{r}) = \mu , \quad (2.57)$$

where:

$$v_{eff}(\vec{r}) = v_{ext}(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + v_{xc}(\vec{r}) , \quad (2.58)$$

where $v_{ext}(\vec{r})$ is the potential due to the nuclei and the exchange and correlation potential, $v_{xc}(\vec{r})$, is defined by:

$$v_{xc}(\vec{r}) = \frac{\delta E_{xc}[\rho(\vec{r})]}{\delta \rho(\vec{r})}, \quad (2.59)$$

and the Hartree potential, $v_H(\vec{r})$, and , are defined by:

$$v_H(\vec{r}) = \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}'. \quad (2.60)$$

In this manner, the energy as a functional of the electronic density in the Kohn-Sham (KS) formalism is given by:

$$E_{KS}[\rho(\vec{r})] = T_0[\rho(\vec{r})] + \frac{1}{2} \iint \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}d\vec{r}' + \int v_{ext}(\vec{r})\rho(\vec{r})d\vec{r} + E_{xc}[\rho(\vec{r})]. \quad (2.61)$$

The Schrödinger-like Kohn-Sham equation for a system of i electrons, in the DFT fashion, is defined by:

$$h_i^{KS}\psi_i^{KS} = \varepsilon_i\psi_i^{KS}, \quad (2.62)$$

where h_i^{KS} is the KS Hamiltonian for the particle i , given by:

$$h_i^{KS} = -\frac{\hbar^2}{2m}\nabla_i^2 + \underbrace{v_{ext}(\vec{r}) + v_H(\vec{r}) + v_{xc}(\vec{r})}_{v_{eff}(\vec{r})}. \quad (2.63)$$

In Equation 2.63, ε is the energy eigenvalue and, finally, ψ_i^{KS} is the KS wave function, given by the Slater determinant in Equation 2.64:

$$\Psi^{KS} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1^{KS}(\vec{r}_1) & \psi_2^{KS}(\vec{r}_1) & \cdots & \psi_i^{KS}(\vec{r}_1) \\ \psi_1^{KS}(\vec{r}_2) & \psi_2^{KS}(\vec{r}_2) & \cdots & \psi_i^{KS}(\vec{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{KS}(\vec{r}_N) & \psi_2^{KS}(\vec{r}_N) & \cdots & \psi_i^{KS}(\vec{r}_N) \end{vmatrix}. \quad (2.64)$$

The $\psi_i^{KS}(\vec{r}_N)$ also includes spin degree of freedom, given by:

$$\psi_i^{KS}(N) = \phi_i^{KS}(N)\alpha(N), \quad (2.65)$$

and

$$\psi_i^{KS}(N) = \phi_i^{KS}(N)\beta(N) . \quad (2.66)$$

Approximations in DFT

Apart from the BO approximation discussed before, which has been used to separate the electrons and nuclei solutions, we did not use any other approximation^V. However, the exact form of the exchange and correlation functional in KS-DFT, shown in Equation 2.59, is not known exactly. Much effort has been made in order to develop new and more accurate approaches for these functionals in DFT.

The most common approximations [40] are the local-density approximation (LDA), local-spin-density approximation (LSDA) and generalized-gradient approximation (GGA). In LDA one considers that the system's electronic density does not vary much over distance. This is usually true for some particular systems, but it can fail to describe some molecular systems, those which involve long range interactions or open shell systems.

For the latter case, the LSDA approximation improves the description provided by LDA, since it includes the spin degree of freedom. In both LDA and LSDA the basic idea is to use the exchange and correlation energy at a homogeneous electron gas with density $\rho(\vec{r})$ at $\vec{r}$ - thus the term local.

In most real systems, the assumption that the electronic density does not vary with the distance is not true. In these cases, the GGA is used to describe the exchange and correlation functional. Here a gradient term (electronic density's first order derivative) is included to take into account variations in electronic density. In this thesis we have used the GGA PBE [50] functional as well as the VdW DRSL [51], which includes dispersion corrections. Some systems studied by us involve van der Waals (vdW) interactions and the VdW DRSL functional typically improves energies and structures in these cases.

For practical reasons, the KS wave functions (Equation 2.64) are written in terms of a basis set. Thus, the problem becomes one of finding coefficients (from the expansion) that minimizes the KS equations. There are different philosophies to construct and propose basis set, like plane-waves and localized basis functions. In this thesis we have used the latter one.

In localized basis set a combination of atomic orbitals centered in the nuclei is used to describe the molecular orbitals of the systems, the so called linear combination of atomic orbitals (LCAO). In this basis set the interaction range has a cutoff, in such a way that the Hamiltonian elements go to zero as the distance increase. This allow feasible and more efficient calculations as Hamiltonian matrix elements are localized in space in molecular systems [49]. In our particular case it will also allow for the Hamiltonian to be cast in a tight-

^VSome erroneously assign to DFT a character of an effective theory. It is important to point out that the KS Hamiltonian, if solved exactly, will yield the exact many body energy and the exact many body self-energy.

binding-like form which is appropriated for electronic transport calculations.

Finally, for most chemical events, only valence electrons are directly involved. In this way, the use of so-called pseudo potentials [52] to describe the core electrons present a very useful way to simplify the problem, saving computational time. In this manner, the basis set constructed using pseudo potentials treat the ion potential dressed by the core electrons. This basis set is implemented in SIESTA [53–55], the electronic structure code used by us to perform all DFT calculations in this thesis. Here we use Troullier-Martins [52] norm-conserving pseudo potentials to simulate the ion + core electrons.

The size of the basis is arbitrary, and it usually a subset of the full Hilbert space. Typically the orbitals are first chosen as solution to the isolated atoms present in the system. In order to better describe bonds, however, one usually includes more radial functions (multiple ζ) for a particular orbital or more angular functions for a particular radial function (polarization orbitals).

The Density Functional Theory Protocol

The DFT protocol is summarized in the flowchart shown in Figure 2.3. A guess electronic density initializes the process (typically the atomic configuration), then the KS Hamiltonian is calculated, and is solved; a new electronic density is obtained from the calculated wave function and if a convergence criterion is met, the self-consistent procedure ends. If not, the self-consistent cycle is repeated, until that convergence criterion is achieved.

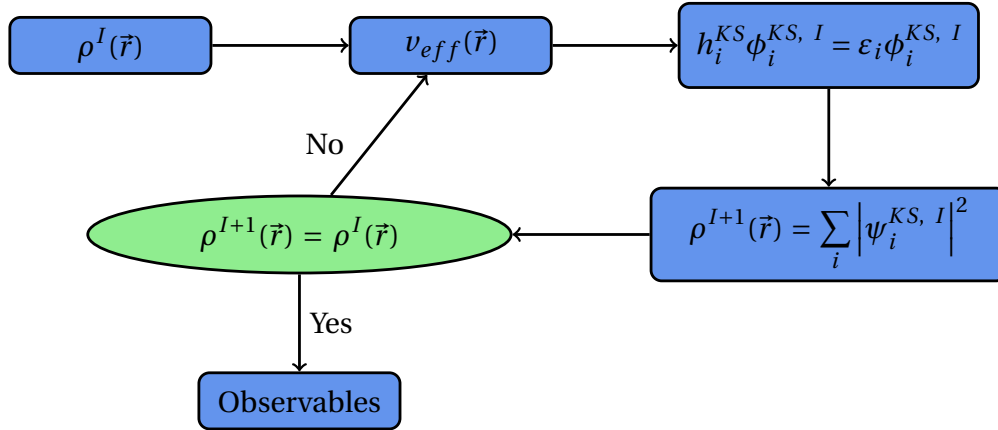


Figure 2.3: Flowchart for the DFT self-consistent cycle.

2.2.3 The QM/MM method

Up to now we showed the approaches used to simulate a nanodevice using different levels of theory. Since we decoupled the electronic and nuclear degrees of freedom we showed how to solve the ion time evolution (MD calculation) and the electronic problem (DFT calculation). Since we want to simulate very large systems, a purely QM simulation - like DFT -

is unfeasible. However, a purely MM simulation - like MD - is not enough because we want to simulate electronic transport. Therefore, we can use the QM/MM method to deal with all these requirements [39].

The main idea in QM/MM is to divide the system into a QM and a MM part, according to the system/property one wants to simulate. The choice of the QM/MM partition is arbitrary but in some systems there is a minimum QM region depending on the chemical event which is going to be investigated. In the ideal limit the entire system should be treated quantum mechanically, so the choice also entails a trade off between accuracy and computational time.

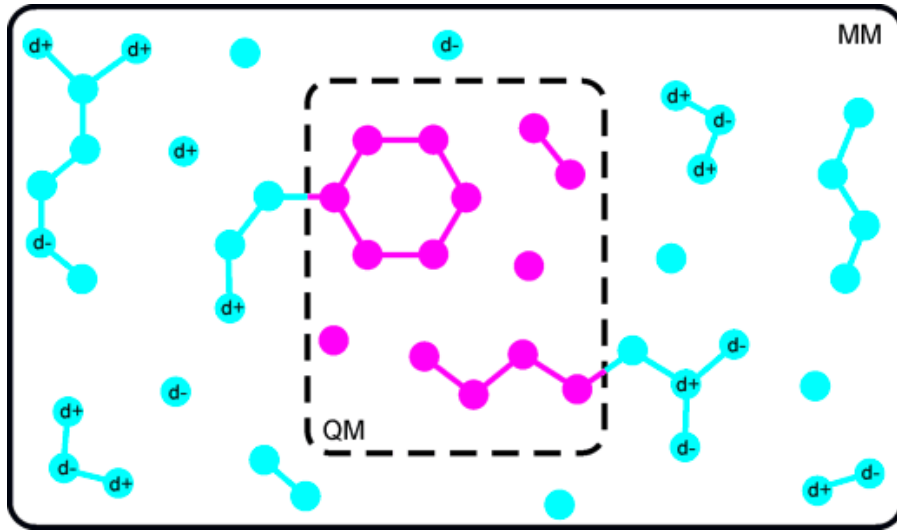


Figure 2.4: Explanation of QM/MM partitions. The magenta atoms are in the QM part and the cyan ones are in the MM part. This QM/MM partition was defined across a chemical bond. Reprinted with permission from reference [56]. Copyright © (2013) Springer Nature.

An example of a QM/MM partition is shown in Figure 2.4. As the figure shows, the atoms represented in cyan make up the MM part of this system whereas the atoms represented in magenta make up the QM part. If a quantum mechanical event is happening in a certain region of a system, such as breaking and formation of chemical bonds, then that region should be included in the QM part. There are two main coupling methods in QM/MM method, the subtractive:

$$V_{QM/MM} = V_{MM}(MM + QM) + V_{QM}(QM) - V_{MM}(QM), \quad (2.67)$$

and the additive:

$$V_{QM/MM} = V_{QM}(QM) + V_{MM}(MM) + V_{QM-MM}(QM + MM), \quad (2.68)$$

where $V_{QM/MM}$ is the hybrid potential of the system, V_{QM} is the potential of the QM part,

V_{MM} is the potential of the MM part and V_{QM-MM} is the interacting potential between the QM and MM parts. This interacting potential, which is required in the additive scheme, is not required in the subtractive one, making it more complex.

To address the philosophy of QM/MM method differentiating the electrons in the system, we can imagine a system composed by a molecule interacting with a solvent. Lets consider that in this system we can put the entire molecule in the QM part and the solvent in the MM part. Furthermore, lets consider that the electrons in the QM part are divided into core and valence electrons. Thus, we can write a Hamiltonian to describe the interactions involved in this problem,

$$\hat{H} = \sum_I^M \frac{\vec{P}_I^2}{2m_I} + \sum_A^N \frac{\vec{P}_A^2}{2m_A} + V_{solv/solv}(\{\vec{R}_I\}) + V_{cor/cor}(\{\vec{R}_A\}) + V_{solv/cor}(\{\vec{R}_I\}, \{\vec{R}_A\}) + \hat{H}_{ele}, \quad (2.69)$$

where $\vec{p}$ is the momentum, I, J are solvent atoms (MM), A, B are the "nuclei" (QM nuclei + core electrons), i, j are the valence electrons (QM), M is the number of classical nuclei and N is the number of quantum nuclei. Thus, m is the mass and $\vec{R}$ is the distance between the particles, according to the label. The interaction between the solvent atoms is given by:

$$V_{solv/solv}(\{\vec{R}_I\}) = \sum_I^M \sum_{J>I}^M \frac{Q_I Q_J}{R_{IJ}} + \sum_I^M \sum_{J>I}^M 4\epsilon_{IJ} \left\{ \frac{\sigma_{IJ}^{12}}{R_{IJ}^{12}} - \frac{\sigma_{IJ}^6}{R_{IJ}^6} \right\}, \quad (2.70)$$

where Q_I and Q_J are the partial charges, σ_{IJ} and ϵ_{IJ} are the Lenard-Jones parameters. The interaction between electrons and solution is given by:

$$V_{cor/cor}(\{\vec{R}_A\}) = \sum_A^N \sum_{B>A}^N \frac{Z_A Z_B}{R_{AB}}, \quad (2.71)$$

where Z_A and Z_B are the effective charges of the nuclei.

The interaction between solvent atoms and solute nuclei is given by:

$$V_{solv/cor}(\{\vec{R}_I\}, \{\vec{R}_A\}) = \underbrace{\sum_I^M \sum_A^N 4\epsilon_{IA} \left\{ \frac{\sigma_{IA}^{12}}{R_{IA}^{12}} - \frac{\sigma_{IA}^6}{R_{IA}^6} \right\}}_{\text{van der Waals}} + \underbrace{\sum_I^M \sum_A^N \frac{Q_I Z_A}{R_{IA}}}_{\text{electrostatic}}. \quad (2.72)$$

And finally, the electronic Hamiltonian is the following:

$$\hat{H}_{ele} = - \sum_i^n \frac{\hbar^2}{2m_e} \nabla_i^2 + \sum_j^n \sum_{i>j}^n \frac{e^2}{\vec{r}_{ij}} - \sum_i^n \sum_I^M \frac{Q_I e^2}{\vec{r}_{iI}} - \sum_i^n \sum_A^N \frac{Z_A e^2}{\vec{r}_{iA}}. \quad (2.73)$$

Thus, the idea of QM/MM simulations for a molecule in a solvent describing the core and valence electrons separately was shown above. Our QM/MM calculations were performed as implemented in SIESTA [38]. Thus the solvent effect (usually the MM region) is taken into account via an electrostatic potential acting in the atoms of the QM region.

The electrostatic potential is calculated via partial charges obtained by the FF used in the MD simulation. It is directly added to the Hartree potential^{VI} shown in Equation 2.60, in each self-consistent step until convergence. Is important to mention that the calculated potential is also periodic because of the boundary conditions. Thus, the QM/MM coupling using the electronic density ($\rho(\vec{r})$), as performed in DFT calculations, is:

$$E_{QM/MM} = \sum_{i=1}^{N_{clas.}} q_i \int \frac{\rho(\vec{r})}{|\vec{r} - \vec{r}_i|} d\vec{r} + \sum_{i=1}^{N_{clas.}} \sum_{j=1}^{N_{quan.}} \frac{q_i Z_j}{|\vec{R}_j - \vec{r}_i|}, \quad (2.74)$$

where $N_{clas.}$ is the number of classical particles (usually refereed by MM part), $N_{quan.}$ is the number of quantum particles (usually refereed by QM part), q_i are the partial charges of the MM particles in r_i positions, R_j are the positions of the QM nuclei with nuclear charge given by Z_j .

Finally, an important aspect when we run QM/MM calculation is to define and correctly calculate the boundaries. In most of our calculations we have chosen the whole molecule as the QM part, thus no chemical bonds are split between the partition. However, breaking bonds in the QM/MM partition some times necessary. In fact, we performed a QM/MM partition test to show the effect of the partition in the electronic transport (shown in Appendix A). In the cases where a covalent bond is broken (like the example illustrated in Figure 2.4) to divide the system into a QM and an MM part we used the H capping [57] method to describe the boundaries. In this method the valence of the atom in the QM region, where the chemical bond was broken, is completed by adding a Hydrogen atom. The partial charge of the MM region calculated through the FF parameters is also corrected.

In summary, the QM/MM method used by us in this thesis is performed over snapshots extracted from a MD calculation. In our cases the MD calculation was performed to generate structures sampling our systems. A QM/MM method to evolve the system is not the best choice in our case, since we need long MD calculations and our systems comprises thousands of atoms even in the QM part.

2.2.4 Electronic Transport - NEGF

As we showed up to now, within the methodology used by us we can include the effect of the solvent on the Hamiltonian generated by means of DFT over geometries sampled by

^{VI}It is important to mention that the v_{ext} in Equation 2.63 is not the same potential calculated for the MM part, referred as V_{ext} , which is summed in the Hartree potential, as we discussed.

MD calculations. As we also discussed, the DNA sensing devices studied by us are based on electronic changes on graphene-based biosensors. Thus in order to treat this problem we require the calculation of the electronic transport properties of the system. A way forward use the NEGF [20, 58–62] formalism which allows for out-of-equilibrium properties of many particle systems, including electronic transport.

An illustration of the electronic transport problem is shown in Figure 2.5. The problem is simulated by defining three regions: the left lead, the extended molecule and the right lead in the spirit first put forward by Caroli [63] et al. Each lead has a chemical potential (μ_L and μ_R in Figure 2.5a). Thus, we have a scattering problem in which an incoming electron (from the left lead) is scattered in the scattering region (EM - extended molecule region in the figure) being, then, transmitted (to the right lead). Figure 2.5c shows this scattering problem from the electrostatic point of view. By applying a voltage V in the system we can have the chemical potentials $\mu_{L/R} = E_F \pm V/2$, where E_F is the system's Fermi energy at equilibrium.

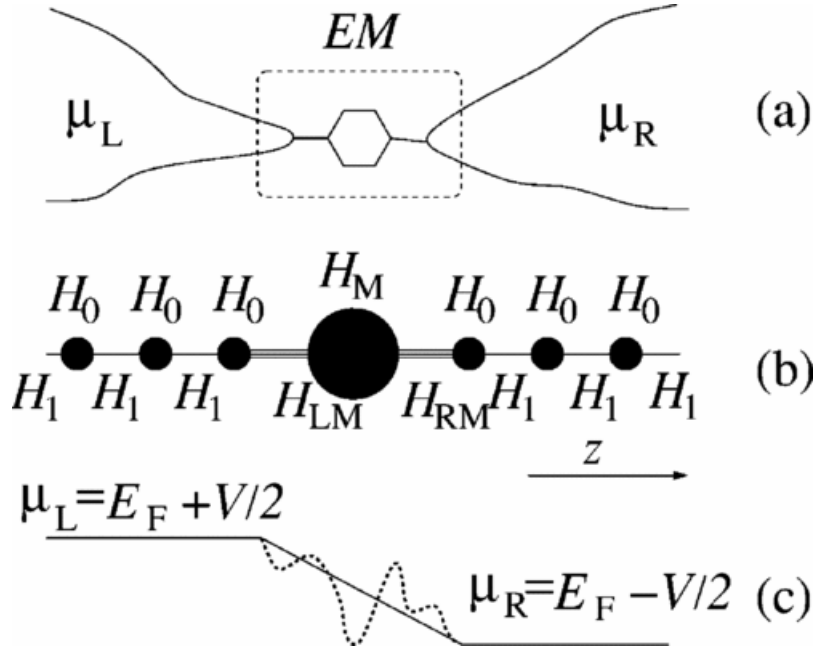


Figure 2.5: Illustration of the electronic transport problem. In this figure the thermodynamics (a), Hamiltonian (b) and electrostatics (c) pictures of this problem are shown. Reprinted with permission from reference [61]. Copyright © (2006) American Physical Society.

This problem can be simulated using two semi-infinite electrodes (not periodic in the transport direction but periodic in the transversal plane) with the scattering region in the center. Figure 2.5b shows a Hamiltonian description to this problem: the semi-infinite Hamiltonians, H_0 , interacting via H_1 , are coupled to the scattering region, H_M (extended molecule Hamiltonian). This coupling is defined by H_{LM} and H_{RM} , which describes the coupling between the left/right electrode and the scattering region.

This description is suitable for SIESTA since it describes the valence electrons using a localized basis set, as we mentioned before, allowing for the partitioning of the system into the

ones cited above. This problem can be solved using Green's functions and, thus, it can be written in terms of the retarded Green's function, defined by:

$$[\epsilon^+ \mathcal{S} - \mathcal{H}] \mathcal{G}^R(E) = \mathcal{I} , \quad (2.75)$$

where $\mathcal{S}$ is the overlap matrix, $\mathcal{I}$ is the identity matrix, E is the energy, $\epsilon^+ = \lim_{\delta \rightarrow 0^+} E + i\delta$ and $\mathcal{H}$ is the Hamiltonian.

In this framework, the Hamiltonian description shown in Figure 2.5b can be written in terms of the left/right semi-infinite Hamiltonians, showing its block-diagonal forms:

$$\mathcal{H}_R = \begin{bmatrix} H_0 & H_1 & 0 & \dots & \dots \\ H_{-1} & H_0 & H_1 & 0 & \dots \\ 0 & H_{-1} & H_0 & H_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} , \quad (2.76)$$

and

$$\mathcal{H}_L = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & H_{-1} & H_0 & H_1 & 0 \\ \dots & 0 & H_{-1} & H_0 & H_1 \\ \dots & \dots & 0 & H_{-1} & H_0 \end{bmatrix} , \quad (2.77)$$

where $H_{\pm 1}$ represents the interaction inside the electrodes with the previous and the next cell. The matrices that defines the coupling between the leads and the scattering region are, thus, given by:

$$\mathcal{H}_{RM} = \begin{bmatrix} H_{RM} \\ 0 \\ 0 \\ \vdots \end{bmatrix} , \quad (2.78)$$

and

$$\mathcal{H}_{LM} = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ H_{LM} \end{bmatrix} . \quad (2.79)$$

These matrices are obtained performing separate DFT calculations on periodic electrodes. Thus, partitioning the Green's functions we can write :

$$\begin{bmatrix} \epsilon^+ - \mathcal{H}_L & \epsilon^+ \mathcal{S}_{LM} - \mathcal{H}_{LM} & 0 \\ \epsilon^+ \mathcal{S}_{ML} - \mathcal{H}_{ML} & \epsilon^+ \mathcal{S}_M - \mathcal{H}_M & \epsilon^+ \mathcal{S}_{MR} - \mathcal{H}_{MR} \\ 0 & \epsilon^+ \mathcal{S}_{RM} - \mathcal{H}_{RM} & \epsilon^+ \mathcal{S}_R - \mathcal{H}_R \end{bmatrix} \begin{bmatrix} \mathcal{G}_L & \mathcal{G}_{LM} & \mathcal{G}_{LR} \\ \mathcal{G}_{ML} & \mathcal{G}_M & \mathcal{G}_{MR} \\ \mathcal{G}_{RL} & \mathcal{G}_{RM} & \mathcal{G}_R \end{bmatrix} = \begin{bmatrix} \mathcal{I} & 0 & 0 \\ 0 & \mathcal{I}_M & 0 \\ 0 & 0 & \mathcal{I} \end{bmatrix}. \quad (2.80)$$

We can solve this problem by eliminating the degrees of freedom of the electrodes keeping only the one of the interface between the electrode and the scattering region. Thus, we can obtain the retarded Green's function for the scattering region, which is the region of interest:

$$G_M^R(E) = [\epsilon^+ S_M - H_M - \Sigma_L^R(E) - \Sigma_R^R(E)]^{-1}, \quad (2.81)$$

where S_M is the overlap matrix and $\Sigma_{L/R}^R$ are the self-energies which describe the coupling to the S_M :

$$\Sigma_R^R(E) = (\epsilon^+ S_{MR} - H_{MR}) g_R^{0R} (\epsilon^+ S_{RM} - H_{RM}), \quad (2.82)$$

and

$$\Sigma_L^R(E) = (\epsilon^+ S_{ML} - H_{ML}) g_L^{0L} (\epsilon^+ S_{LM} - H_{LM}), \quad (2.83)$$

where $g_{L/R}^{0R}$ are the Green's functions for the isolated electrodes [60, 61]. From the Green's function we can calculate the current at a given bias V :

$$I = \frac{e}{h} \int T(E, V) (f_L - f_R) dE, \quad (2.84)$$

where $f_{L/R}$ are the Fermi-Dirac distribution with chemical potential $\mu_{L/R} = E_F \pm V/2$ and we identity:

$$T(E, V) = \Gamma_R(E) G^R(E) \Gamma_L(E) G^A(E), \quad (2.85)$$

where $\Gamma_{L/R}$ are the left and right coupling matrices as a bias dependent Landauer-Büttiker-

like transmission coefficient [20, 64]:

$$\Gamma_{L/R}(E) = \text{Im} \left[\Sigma_{L/R}^R(E) - \Sigma_{L/R}^R(E)^\dagger \right]. \quad (2.86)$$

In the limit of low bias the conductance can be related to the transmission of the Fermi energy by the Fisher-Lee [65] relation:

$$G = \frac{2e^2}{h} T(E_F), \quad (2.87)$$

where h is the Planck constant.

The potentials calculated for the scattering region and for the right/left electrodes should be aligned, since the electronic transport calculations require well defined boundaries. However, the V_{ext} calculated to include the effect of the MM part can present fluctuations in the electrodes' region, what can cause problems in the electronic transport calculations. To address this issue we apply a smoothing function (Fermi-Dirac-like) in this potential taking it to zero in the electrodes region, as can be observed in Figure 2.6. In this manner we guarantee perfectly aligned potentials in the three regions.

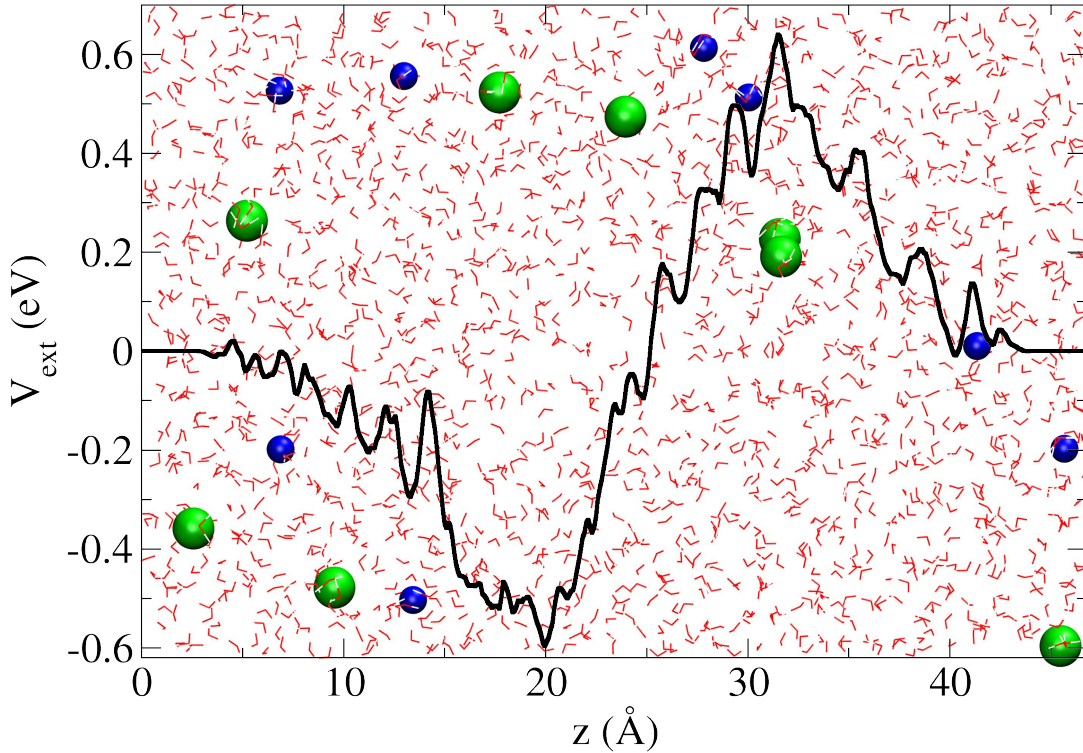


Figure 2.6: External potential calculated for the classical part of an example system. The atoms of the classical part (top view) will determine the average potential as a function of the z direction. This figure shows the water molecules and the Na (blue) and Cl (green) ions.

Finally, for some systems, we also performed local current [59, 66, 67] calculations. This

calculation gives us an insight of the current pathway in the system and is basically the projection of the transmission between two adjacent sites (two atoms in a chemical bond). Defining two sites, I and J , the local current $I(E)$ that flows from I to J is, thus, defined by:

$$I(E)_{I \rightarrow J} = 4 \frac{e}{h} \sum_{\substack{i \in I \\ j \in J}} \text{Im} \left[\{G^R(E) \Gamma_L G^A(E)\}_{ij} H_{ij} \right], \quad (2.88)$$

where H_{ij} is the coupling between the two sites.

2.3 General simulation details

The systems are simulated in a box with periodic boundary conditions (PBC). The box is filled with water and $0.2 \text{ mol } L^{-1}$ of NaCl^{VII} with extra Na ions to equilibrate the charge of the system, accordingly to the system of study. A summary of the methodologies applied by us is presented below.

We initially perform an MD calculation using the NPT ensemble in order to equilibrate the system pressure. From this MD we choose one frame in which the box height (direction to the plane of the graphene sheet) is close to the average height. From this initial structure we perform a second MD simulation using NVT ensemble; The NVT ensemble is chosen in order to have structures with the same amount of vacuum, in such a way that all systems will have equal electrodes. In all MD simulations the graphene sheet is kept fixed^{VIII} during the dynamics and different restraints are applied to the DNA (explained for each case in the chapters where the results are described).

After MD calculations, we divide the system into a QM and an MM part, according to the system, and perform the QM/MM calculation. In the MM part we perform the external potential (V_{ext}) calculation and in the QM part we perform a single point energy DFT calculation reading the V_{ext} calculated before. With the Hamiltonian generated in the DFT calculation (in this case, for scattering region) we can make the electronic transport calculation using the NEGF formalism.

2.3.1 The step-by-step

The step-by-step of this methodology is summarized in the flowchart shown in Figure 2.7 which also shows the programs^{IX} used for each step: Gromacs [35], VMD [69], SIESTA [53–55] and Smeagol [60, 61]. All parameters used for each step will be presented in

^{VII}For graphene in solution we tested different concentrations of NaCl.

^{VIII}For the study of graphene in solution we made one test with free (allowing movements in the sheet) graphene.

^{IX}All figures and animations presented in this thesis and in the presentation slides were made using Tachyon ray tracing library [68] available in VMD program.

the next subsections for each system separately.

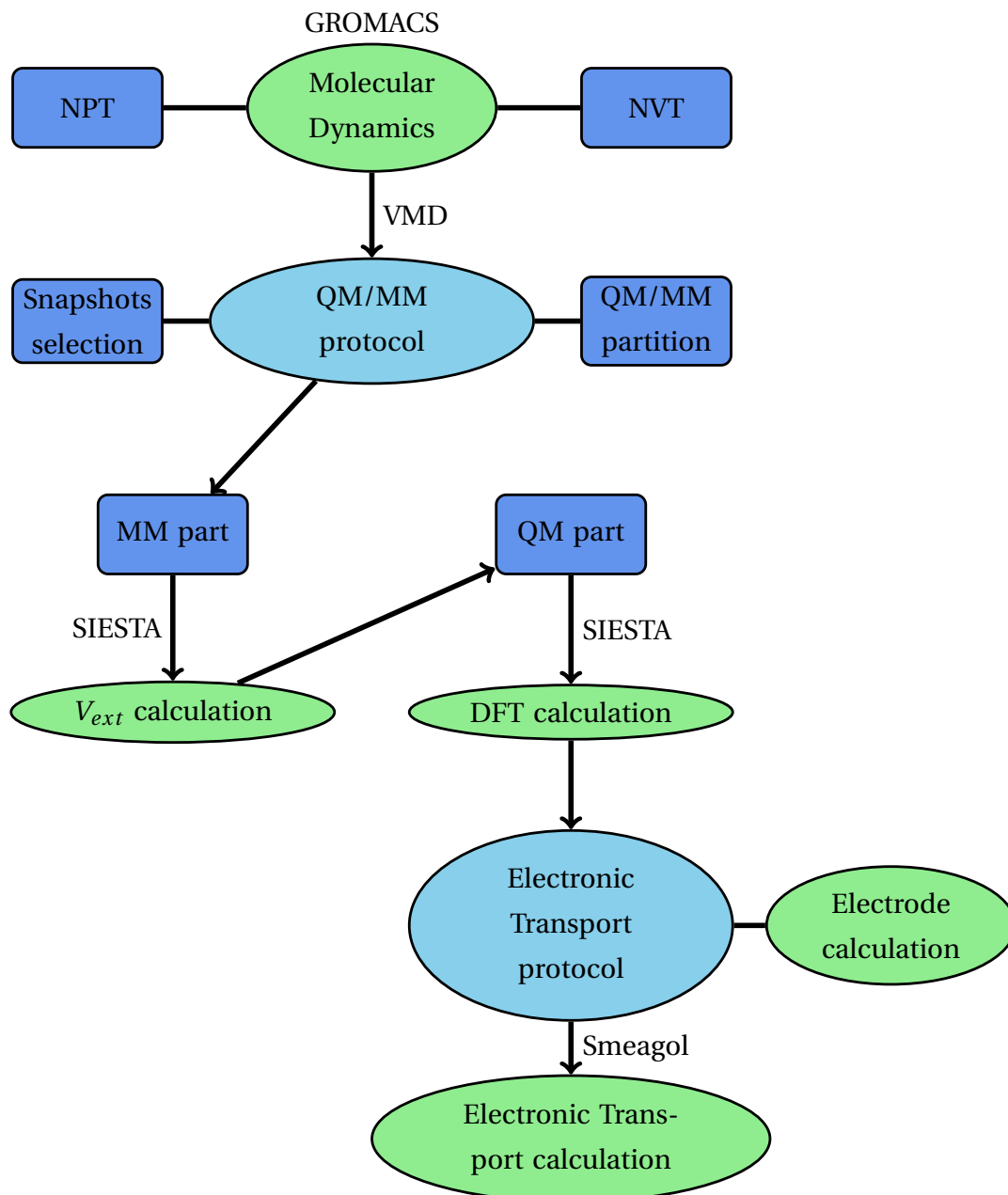


Figure 2.7: Flowchart summarizing the steps of the methodology from Molecular Dynamics to Electronic Transport calculation.

As we mentioned, further details about the employed methodology can be found in the chapters where the results are described according to the systems. All simulations were performed at 298.15 K and 1 atm pressure. Additionally to the references cited in the text, the discussion presented in this chapter is also based on the references [60, 70–76], where detailed descriptions of the applied methods can be found.

CHAPTER 3

Graphene in solution

3.1 Introduction

Before turning to the design of an actual device, it is important to understand some fundamental aspects. The study of solvent effects on graphene represents a fundamental research when it is applied as sensing material because most of these applications are performed in liquid phase [13, 15, 24]. The interface between the liquid and graphene should be clearly understood since it can interfere in the interactions between graphene and the biomolecule one want to detect. Our studies for DNA sensing are performed in liquid phase and the inclusion of the solvent is needed to simulate as realistically as possible these systems.

Solvent effects are neglected in many theoretical works to simulate graphene. If on one hand is hard to introduce solvent effects in this kind of simulation, on the other hand this effect is important and cannot be neglected all the time. Within the methodology we are using here, the solvent effect can be taken into account, since we can simulate very large systems by dividing the problem into a QM and MM part.

From the computational physics perspective, we can simulate the liquid using different models. Here we use MD calculations to sample the liquid dynamics in these systems. The traditional FF's used in MD calculations are based on non-polarizable models. However, some works have shown that the inclusion of polarization on traditional FF's to simulate graphene can alter the interfacial interactions [77, 78] yielding different structuring of water molecules at the interface. These differences in the interfacial water molecules can also interfere in the electronic transport properties of graphene.

An alternative FF to include polarization effects on graphene, called GRAPPA, is discussed in reference [79]. In this FF, two opposite charges (± 0.1) of mass 0.25 amu are added above and below each carbon atom, equally spaced by 0.7 nm . In reference [79] the authors reported that this polarizable FF improves the interfacial interactions presenting better results compared to other models and not increasing significantly the computational cost of the simulations.

We also studied the Stone-Wales (SW) [80] defect on graphene. This defect is constructed by rotating two carbon atoms by 90° , as can be seen in the left panel of Figure 3.1f. The introduction of this new conformation for carbon atoms will alter the electronic structure of graphene and as a consequence can also alter the interaction between graphene and the liquid, since the defect region is more chemically active [81]. We introduced one SW defect in the middle of the sheet to simulate this case.

Here we investigated the effect of the solvent on the electronic transport properties of pristine graphene and graphene with SW defect. We used different aqueous solutions to investigate the interaction between graphene systems and liquid using different FF's for graphene. We vary the salt concentration from zero (pure water) to 0.2 M and to 1 M of NaCl. It is important to mention that in our further investigations for DNA sequencing we include

salt in the solution to equilibrate the charge of the system. Here we address this study with different salt concentrations to understand this effect on the transport through graphene sheet. To investigate these systems we used a number of setups, shown in Figure 3.1.

3.2 The systems

The Figure 3.1 shows the systems studied by us. We studied the polarization effects for pristine fixed graphene in the three different salt concentrations and also a situation where fixed graphene with SW defect is present in pure water. In all these cases the solution was treated classically. Finally, we also considered one solvation layer of water in the QM part of the system (Figure 3.1e) for both polarized and non-polarized fixed pristine graphene for the three different salt concentrations and for non-polarized fixed graphene with SW defect. The QM/MM partitions will be explained later in the text. We also investigated one case for free graphene in pure water (shown in Figure 3.1).

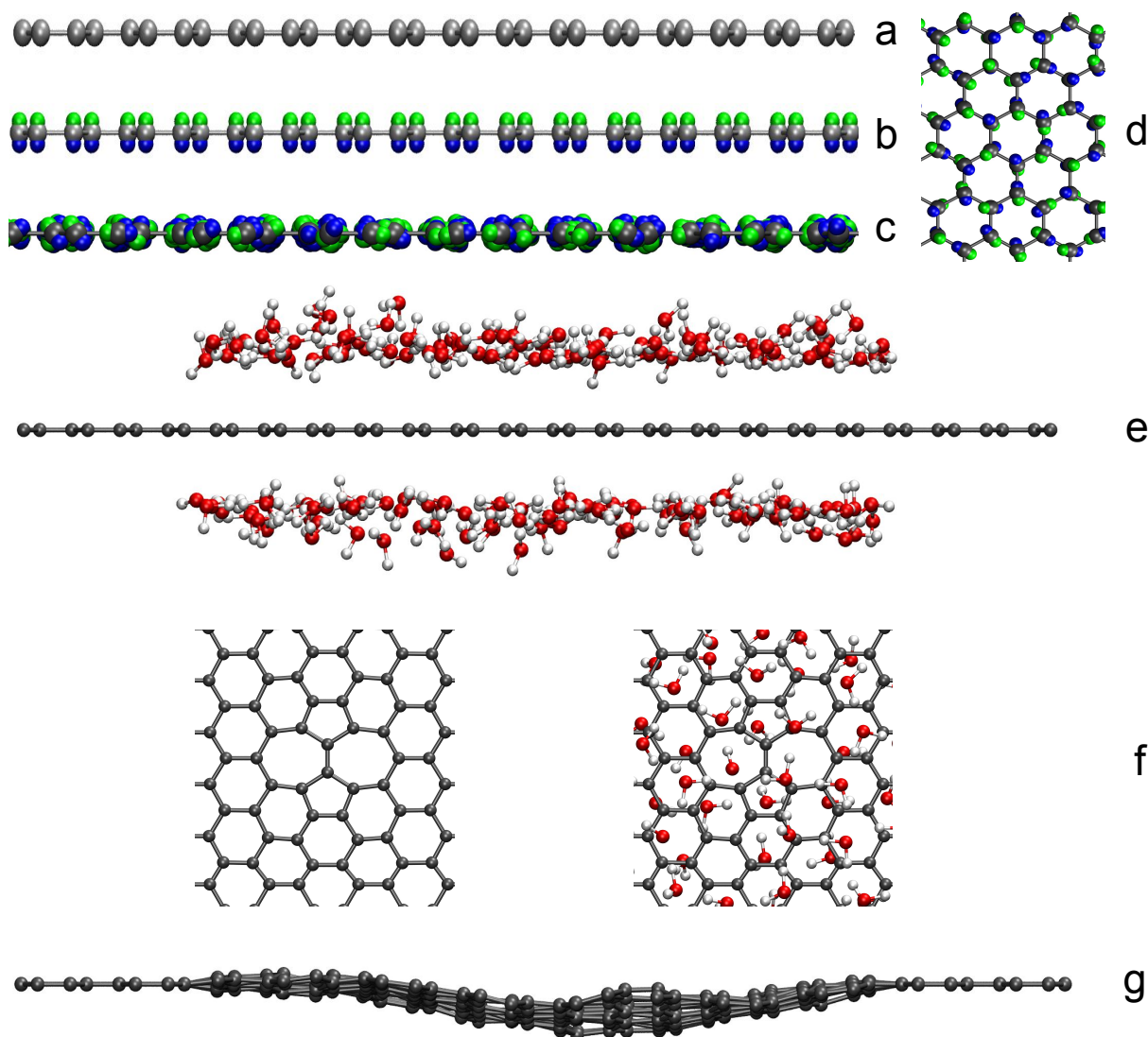


Figure 3.1: Different setups used to study graphene in solution. Dry (a) pristine graphene and the same system with one layer of water (e), both for fixed graphene. Initial frame for the polarization model used for graphene (b) and one frame after the begin of simulation in side (c) and top (d) view. Top view for dry graphene with Stone-Wales defect and the same with one layer of water (f). Dry pristine graphene allowing movements in the sheet (g) where we call free graphene instead of fixed graphene.

The simulation details related with the systems presented in this chapter are shown in Tables [3.1-3.4](#).

MD calculations			
NPT	Parrinello-Rahman barostat	20 <i>ns</i>	
NVT	Nose-Hoover thermostat	20 <i>ns</i>	
Time step	2 <i>fs</i>		
Force Field	AMBER99SB [44] and SPC [45, 46] for water		
Snapshots and QM/MM partition			
Up to 200 snapshots	Equally spaced by 0.1 <i>ns</i>		
QM part I	Graphene sheet		
QM part II	Graphene sheet + first layer of water		
MM part	NaCl ions and the remaining water		
DFT calculations			
Functional	GGA PBE		
Basis set	DZ		
k-points grid (x y z)	1 2 2		
Electronic transport calculations			
Number of energy points	1500		
Number of k-points	16		
Energy range	From -4 <i>eV</i> to 3 <i>eV</i> around Fermi Energy		

Table 3.1: Simulation details for the study of graphene in solution.

pure water - POL	
Box dimensions (x, y, z)	(48.408, 31.980, 46.860)
N_{ions}	0
$N_{water\ molecules}$	2203
$N_{graphene}$	572 C
pure water - NO POL ^I	
Box dimensions (x, y, z)	(47.801, 31.980, 46.860)
N_{ions}	0
$N_{water\ molecules}$	2208
$N_{graphene}$	572 C
water + 0.2 M of NaCl - POL	
Box dimensions (x, y, z)	(48.285, 31.980, 46.860)
N_{ions}	9 Na and 9 Cl
$N_{water\ molecules}$	2185
$N_{graphene}$	572 C
water + 0.2 M of NaCl - NO POL	
Box dimensions (x, y, z)	(47.681, 31.980, 46.860)
N_{ions}	9 Na and 9 Cl
$N_{water\ molecules}$	2190
$N_{graphene}$	572 C

Table 3.2: Details of the simulation box for the study of graphene in solution for pure water and water with 0.2 M of NaCl. N is the number of the specified species. Length unit: Å.

^IThese details are the same for free graphene in pure water.

water + 1 M of NaCl - POL	
Box dimensions (x, y, z)	(48.040, 31.980, 46.860)
N_{ions}	45 Na and 45 Cl
$N_{water\ molecules}$	2113
$N_{graphene}$	572 C
water + 1 M of NaCl - NO POL	
Box dimensions (x, y, z)	(47.396, 31.980, 46.860)
N_{ions}	45 Na and 45 Cl
$N_{water\ molecules}$	2118
$N_{graphene}$	572 C

Table 3.3: Details of the simulation box for the study of graphene in water with 1 M of NaCl. N is the number of the specified species. Length unit: Å.

graphene with SW defect - pure water - POL	
Box dimensions (x, y, z)	(48.265, 31.980, 46.860)
N_{ions}	0
$N_{water\ molecules}$	2196
$N_{graphene}$	572 C
graphene with SW defect - pure water - NO POL	
Box dimensions (x, y, z)	(47.823, 31.980, 46.860)
N_{ions}	0
$N_{water\ molecules}$	2209
$N_{graphene}$	572 C

Table 3.4: Details of the simulation box for the study of graphene in solution for graphene with Stone-Wales defect. N is the number of the specified species. Length unit: Å.

For electronic transport calculation we have to define 3 different regions, shown in Figure

3.2 for the study of graphene in solution^{II}. The green areas comprises both left and right electrodes and the blue areas comprises both left and right buffering regions. The V_{ext} is smoothed out in the buffering region and goes to zero in the electrodes regions in such a way that we have the potentials of electrodes and scattering region aligned. The scattering region showed in Figure 3.2 comprises the QM part of the system. In the figure we can see only the graphene sheet, but in the simulations we have the DNA systems above it. We can also have water molecules, according to the system. The simulation protocol is the same for all the other systems.

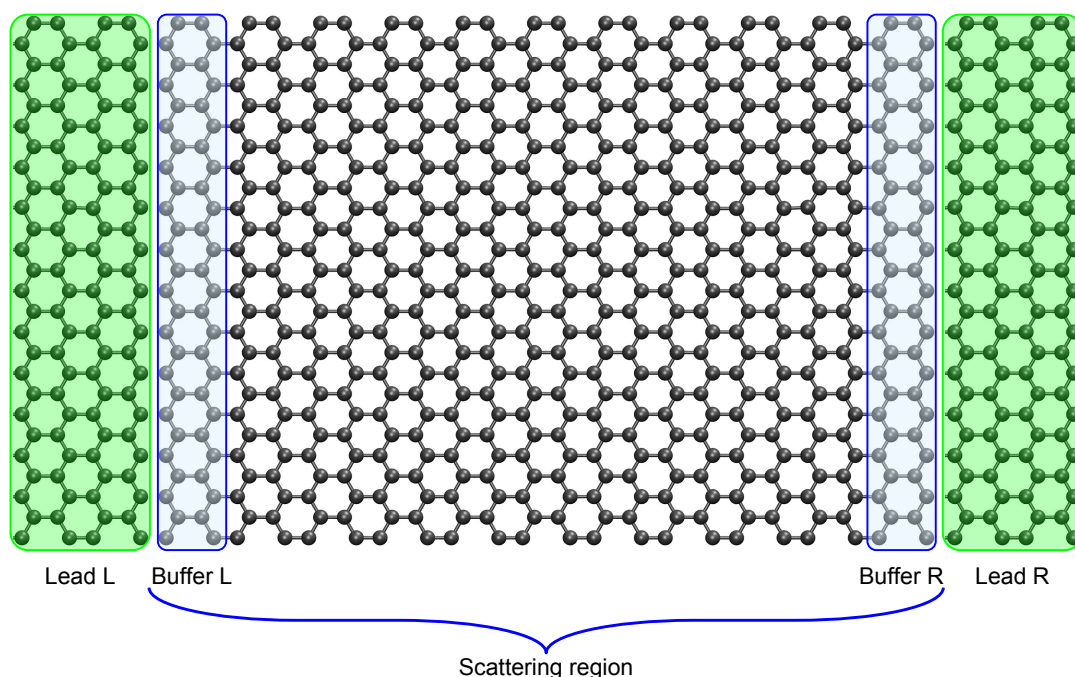


Figure 3.2: Electronic transport setup for the study of both graphene in solution and DNA detection. The green areas delimit the leads and the blue ones delimit the buffers. The scattering region is also shown.

3.3 The effect of immerse graphene in solution

Since we have defined all the systems of study, we are able to apply all the steps of the methodology to study the electronic transport properties of graphene in our three different solutions. The first analysis that we can do is to compare the electronic transport calculation for dry pristine graphene and for the same system in solution (pure water with 0.2 M of NaCl), or, in other words, the effect on the graphene's electronic transport properties of immersing the system in a liquid. A detailed analysis for different solutions will be discussed in the next sections.

^{II}We also investigated Stone-Wales defect in graphene using the same setup for electronic transport calculations.

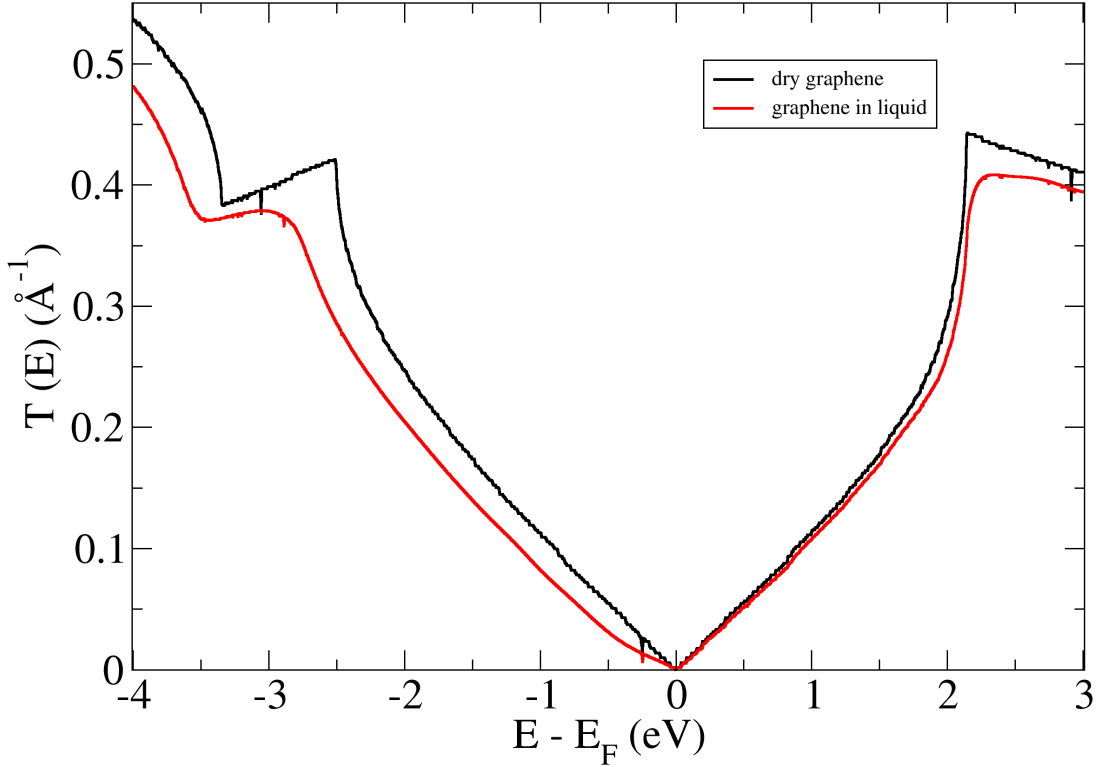


Figure 3.3: Averaged electronic transport calculation (200 frames) for fixed non-polarized pristine graphene in solution (water with 0.2 M of NaCl) and single electronic transport calculation for dry pristine graphene.

Figure 3.3 shows the single calculation for the total transmission of dry graphene (black curve) and the average curve over 200 snapshots for the system immersed in water. $T(E)$ is presented in units of $\text{\AA}^{-1}$, since we divided the total transmission by the length of the box in the y direction (periodic direction in the electronic transport calculation - perpendicular to the transport direction) and multiplied by the number of k -points used to reproduce the periodicity. This length and number of k -points can vary according to the system and this unit will be used in all electronic transport results that we will present throughout this thesis.

As we can observe in Figure 3.3, the overall effect of immersing the system in a solution (red curve) on graphene's electronic transport profile is more scattering, leading to a decrease of the total transmission. The overall trend for the presence of liquid is to induce a chemical gating effect that shifts the local Fermi level of graphene, leading to doping of the sheet. This n -doping in the region containing water shifts the local chemical potential upwards. As all quantities are defined with respect to the Fermi level in the leads, this is equivalent to shifting the band structure downwards. Furthermore, the electronic transport is limited by the maximum number of channels available (which can carry at most one quantum of conductance) either in the electrodes or the scattering region.

The expected resulting transmission without any scattering can be represented by sketch in Figure 3.4. In fact, in Figure 3.3 (black curve) we present the results for a single snapshot showing a very similar curve, especially close to the Fermi energy. Thus, we can conclude

that the overall effect is mostly electrostatic in nature and the presence of the solvent induces small short-range scattering. This result presents the role played by water in the electronic transport properties of graphene, showing the importance of taking into account these effects.

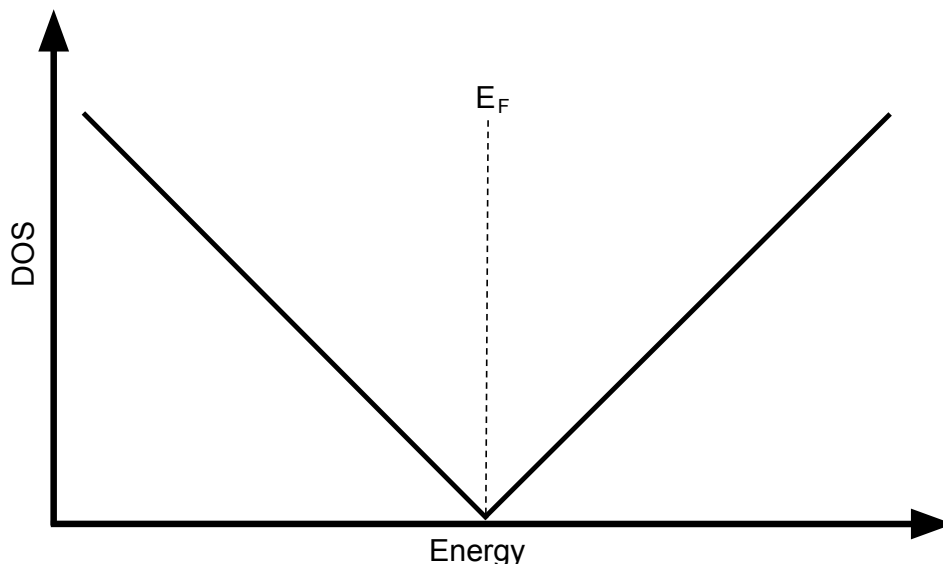


Figure 3.4: Sketch for Density of States (DOS) of pristine graphene to exemplify the case in which there is no scattering.

3.4 Polarization effects

We used the GRAPPA [79] FF to study the polarization effects on graphene, as we discussed. This model is called "rigid rod" model and the initial configuration of the rods is shown in Figure 3.1b. In the first picoseconds of the MD calculation the rods are reoriented in such a way that the opposite charges interact among them, as we can see in Figures 3.1c and 3.1d. It is important to clarify that these charges added in the MD calculation are taken off from the next steps of the methodology: QM/MM and electronic transport.

3.4.1 Structural properties

In order to analyze the water molecules arrangements at the graphene-liquid interface we performed an OH tilt angle analysis. This analysis is also used to define a cutoff radius to include the water molecules in the QM part of the system in the cases where we have different QM/MM partitions. The OH vectors of the water molecules are shown in Figure 3.5a and the OH tilt angles are the angles between these vectors and the unity vector in the x direction (showed in the same figure).

The probability distributions for OH tilt angles showed in Figure 3.5c (for a specific system) are calculated over all frames and all angles for each system (two angles per water

molecule), comprising millions of angles for each system. The data used for these calculations are obtained from the trajectory of the MD calculation. We have developed a code to extract this information and perform the calculation.

As we can see in Figure 3.5c, slice 6 (blue curve) shows a structured layer of water. This layer is shown in Figure 3.5b for the blue layer. We can observe a maximum in the probability around 100° and a small probability around 20° . This kind of structure for water is also observed in other works [78]. Thus, we choose a $5 \text{ \AA}$ cutoff to include the water molecules in the QM part of our calculations with different QM/MM partitions, since this distance is enough to include the water molecules belonging to the blue layer (first solvation shell).

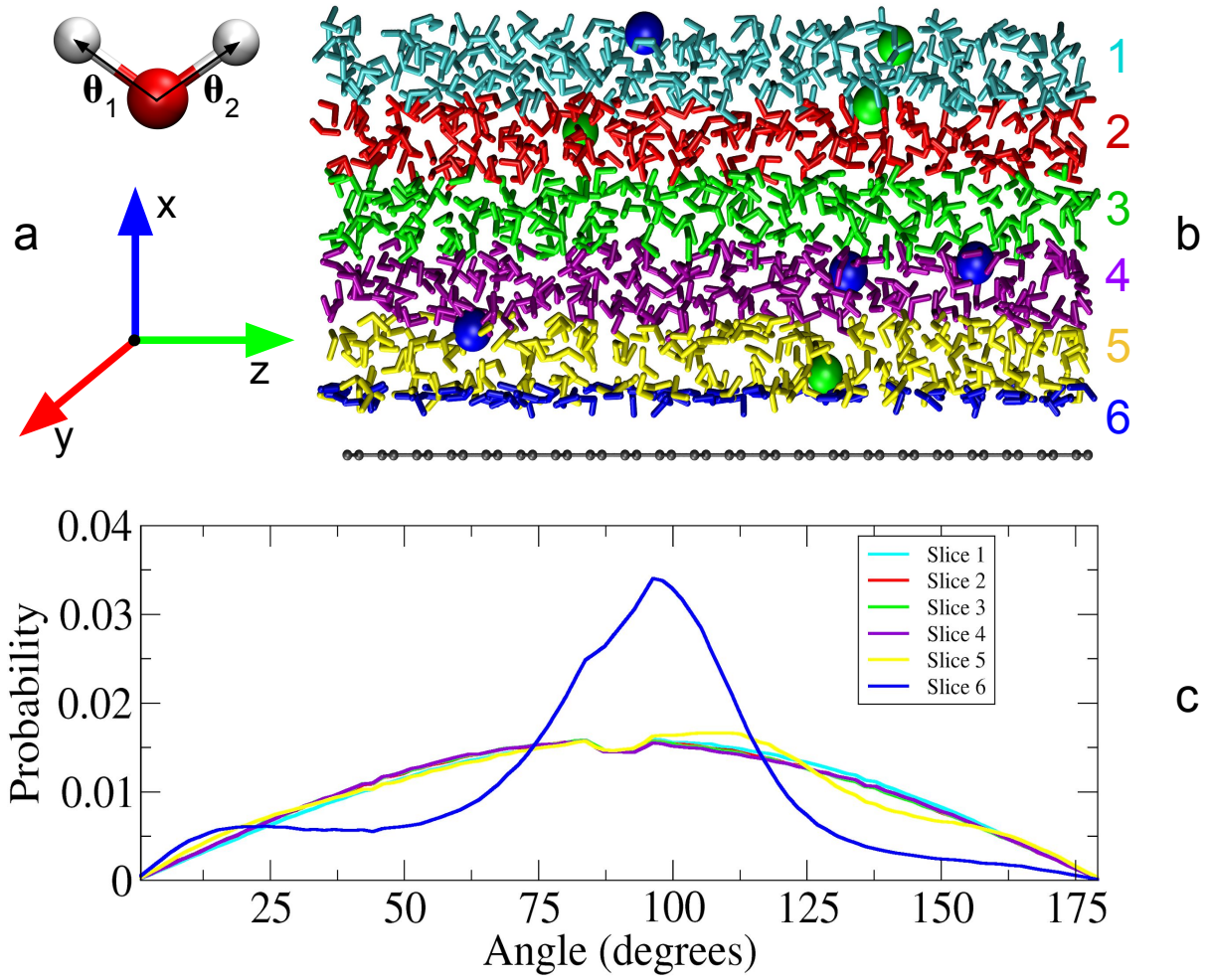


Figure 3.5: The angles θ_1 and θ_2 are the angles between the OH vectors showed in (a) and the $\hat{i}$ unitary vector. Different layers of water (b) with the probability distribution for OH tilt angles (c), for non-polarized fixed graphene with 0.2 M of NaCl. The colors for each layer of water in (b) corresponds to the colors of each curve in (c). Only the half part of the simulation box is shown.

Another analysis that we can do in order to define this cutoff distance is the analysis of number density for the oxygen (or hydrogen) atoms of water in a plane perpendicular to graphene plane. The density analysis can be obtained by an automatic tool provided by GROMACS [35]. In this calculation, the number density for the selected atoms is calculated

using planar surfaces in the system.

Therefore, Figure 3.6 shows the density for oxygen as a function of the box size in x direction (see Figure 3.5). As we can observe, there is a maximum in density around 2.5 Å and a minimum around 5 Å. There are some residual oscillations, but the density converges to bulk quickly in the same way of the angles shown in 3.5.

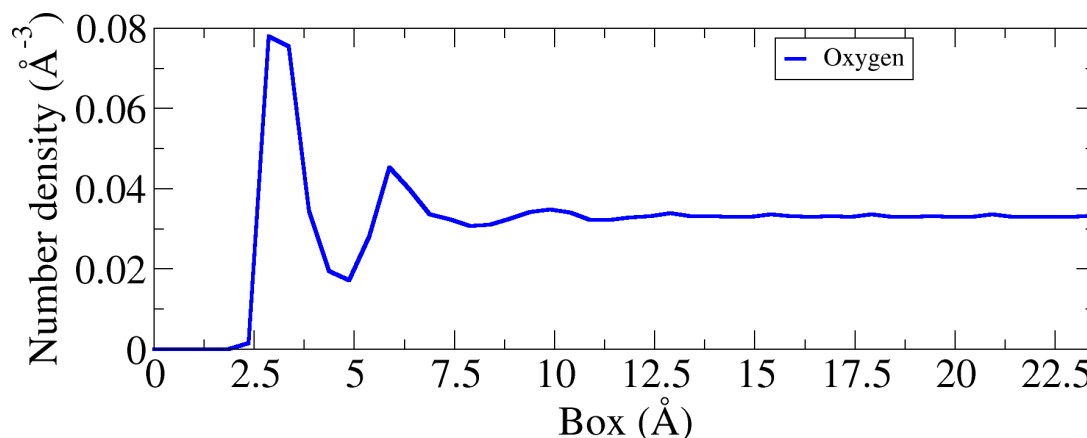


Figure 3.6: Number density for Oxygen atoms, for non-polarized fixed graphene with 0.2 M of NaCl.

This result show that up to 5 Å the structure of the water molecules are affected by the presence of graphene and that this cutoff distance is enough to include all water molecules that make up the first solvation shell in this system. So this is the cutoff distance that we used to choose the water molecules in all cases that we include water in the QM part of the QM/MM partition, as we already mentioned. These water molecules are selected using VMD [69] code, where the program finds the water molecules which are 5 Å away from the nearest atom of the reference group (graphene systems, DNA nucleotides or strands and so on).

For fixed polarized and non-polarized graphene in the three different solutions we analyzed the OH tilt angle for the first solvation shell of water, as shown in Figure 3.7. As can be seen, for fixed graphene in the three different solutions, all systems present a small difference in probability maximum between polarized and non-polarized curves. We notice however, that the remainder of the curves remain essentially unchanged.

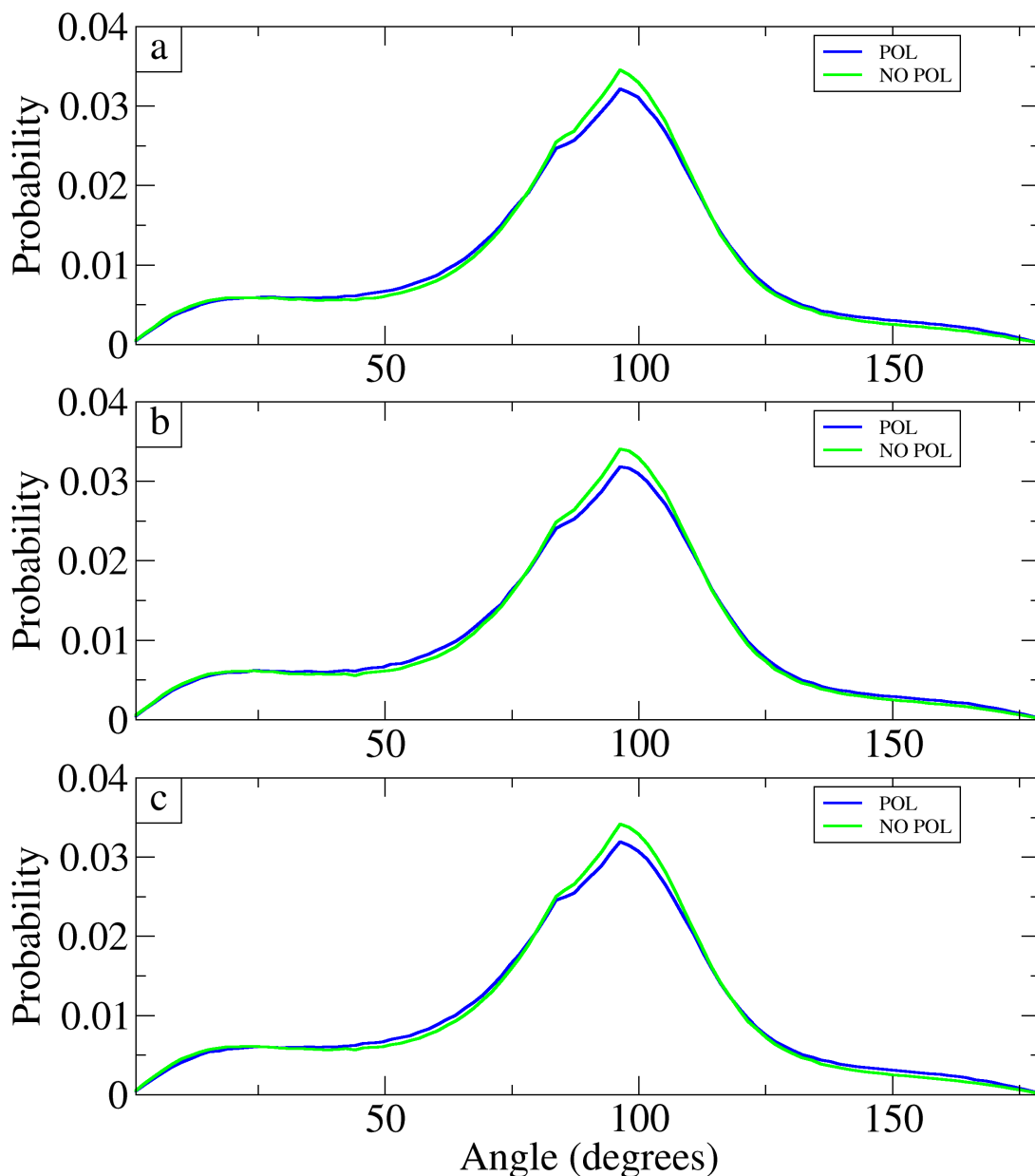


Figure 3.7: Probability distribution for OH tilt angles of the first layer of water for fixed pristine graphene in different solutions: (a) pure water, (b) water with 0.2 *M* of NaCl and (c) water with 1 *M* of NaCl.

Therefore, the polarization model used by us is able to slightly change the organization of the water molecules at the interface. However, increase the salt concentration in the solution does not affect this structuring.

3.4.2 QM/MM partitions

We used two different QM/MM partitions for some systems: the first including only the graphene sheet in the QM part and the second besides the graphene sheet includes also the first layer of water, as we discussed before, in the QM part. To include the first layer of water in the QM part of the system, we should exclude the water molecules above the electrodes,

since we need identical electrodes in all calculations. Therefore, we use the cutoff distance defined by the analysis of angles and number density, as we discussed before, to define the water molecules that will be included in the QM part for each frame.

Since we use a cutoff distance to define the water molecules that will be included in the QM part, we can have a different number of molecules for different snapshots. In Figure 3.8 we can see an example of QM/MM partition including the first solvation shell of water in the QM part, for non-polarized fixed graphene in water with 0.2 M of NaCl.

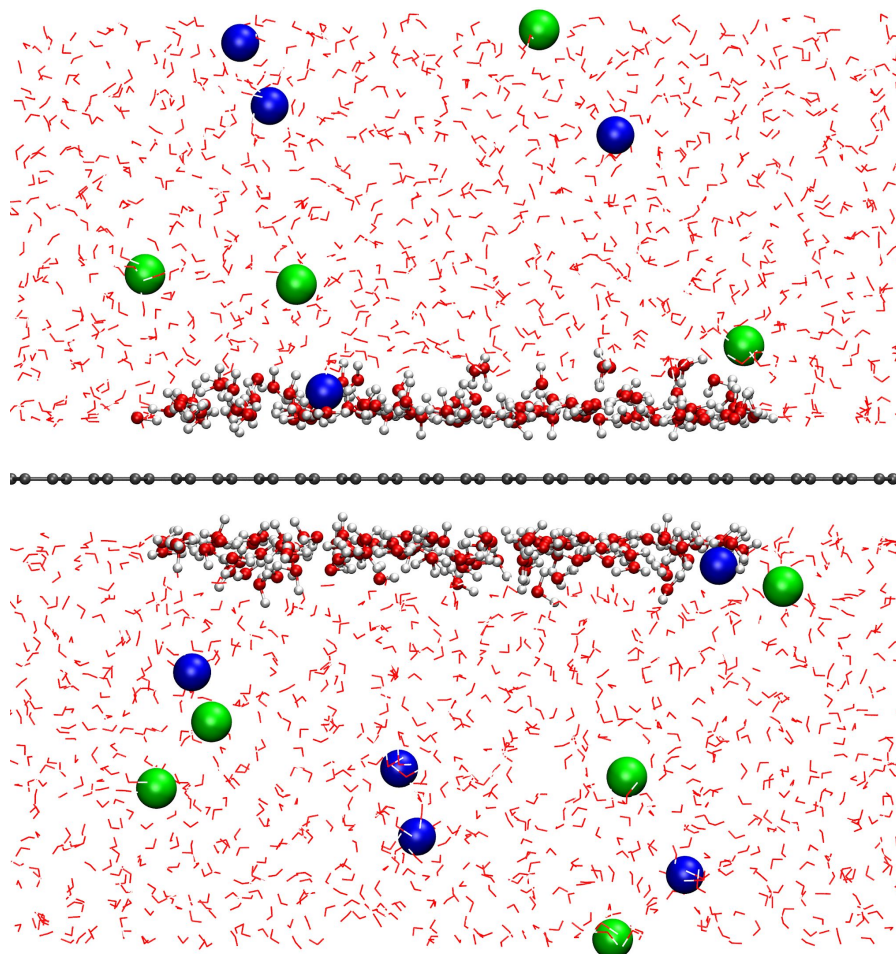


Figure 3.8: Side view of a representative snapshot for non-polarized fixed graphene in water and 0.2 M of NaCl. The graphene and water in *ball-and-sticks* representation make up the QM part whereas the water in *line* representation and the NaCl (green and blue *ball* representation) make up the MM part. This representation pattern will be used throughout this thesis.

3.4.3 Electronic transport results

With the effects of polarization in the water structuring in mind, we turn our attention to the effect of salt concentration and polarization on the electronic transport properties of graphene. Figures 3.10 and 3.9 shows a comparison between the electronic transport calculations for fixed polarized and non-polarized graphene in the three different solutions. These curves are averages over up to 200 snapshots.

As we can observe, the difference between the signals for polarized and non-polarized graphene is rather small. As we showed in Figure 3.7, the inclusion of polarization effects in graphene slightly alters the structure of water molecules in the first solvation shell (closest layer to the sheet), but these changes in the structures do not significantly alter the electronic transport properties of graphene in this QM/MM partition setup (no water included in the QM part).

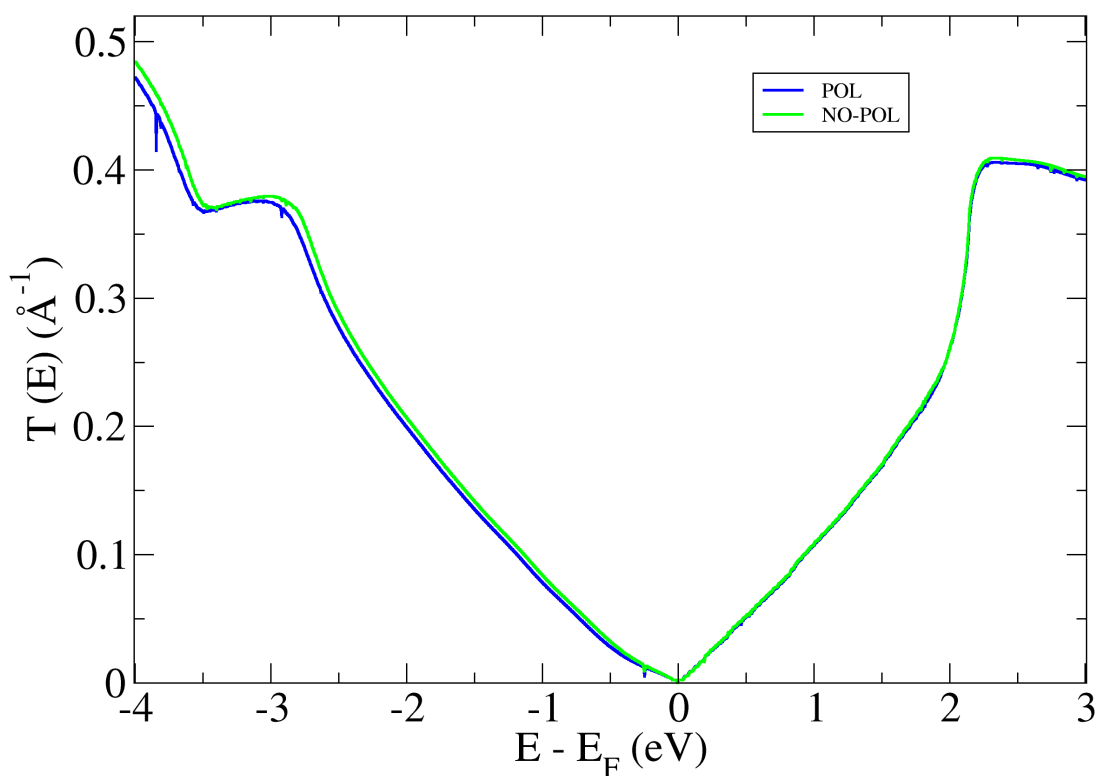


Figure 3.9: Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene in water with 1 *M* of NaCl.

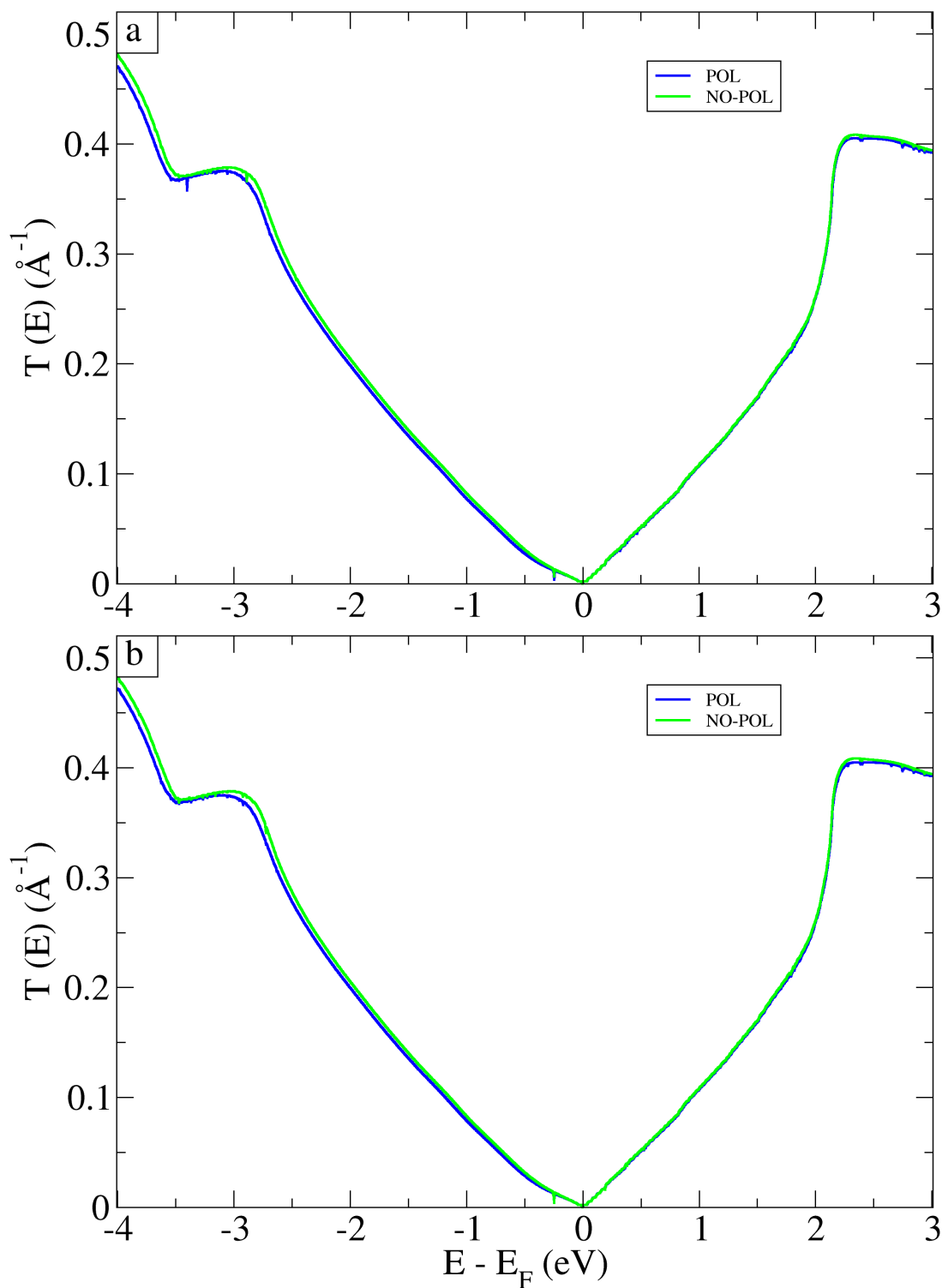


Figure 3.10: Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene in (a) pure water and (b) water with 0.2 M of NaCl.

The presence of polarization does lead to some restructuring of water, but the overall effect is small. On average, those effects are watered down as they are not enough to cause a significant shift in the local chemical potential. The same can be said of salt concentration. The presence of salt does not significantly alter the structure of water close to the surface. Furthermore, Figure 3.11 shows the average concentration of salt as a function of distance.

We notice that both types of counter-ions are, on average, away from the graphene surface, preferring instead to be solvated in water, in a region at the solution, that has little or no influence on the transport properties of graphene.

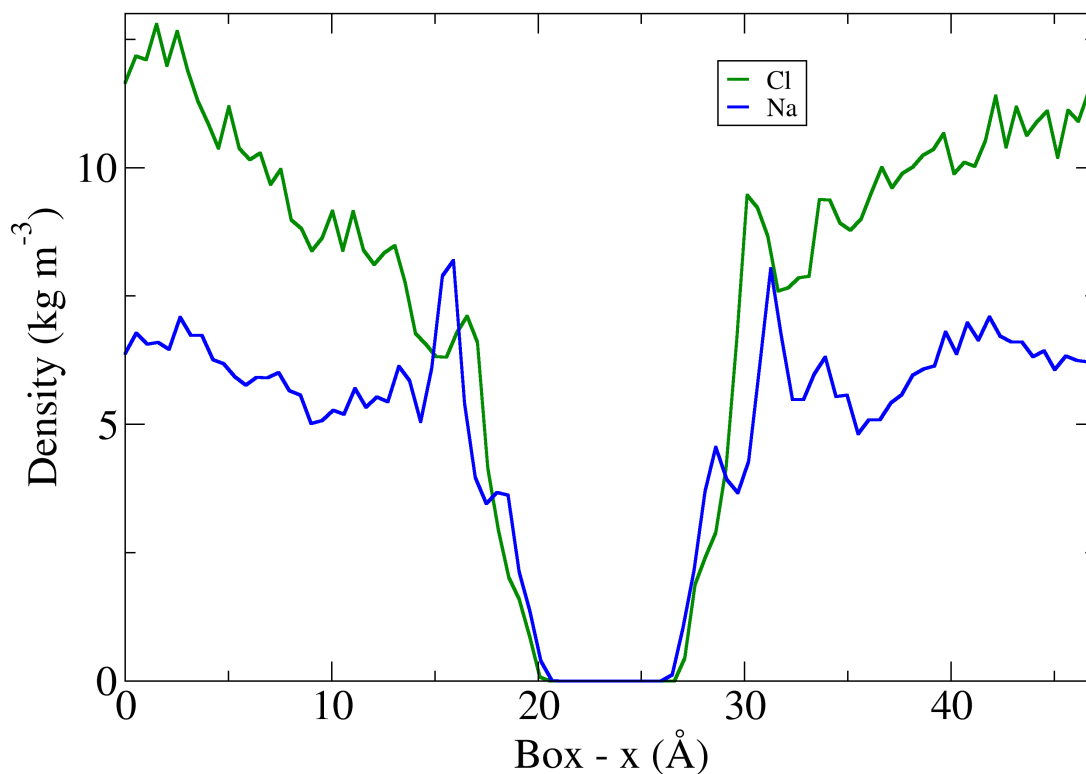


Figure 3.11: Partial densities for Na and Cl ions with respect to the distance from graphene sheet. Example for fixed non-polarized graphene in water with 0.2 M of NaCl. The graphene sheet is placed at $x = 23.74$ Å.

To understand more deeply if the inclusion of polarization effect on the sheet can change the electronic transport properties of graphene, we studied a second QM/MM partition, in which we included this first layer of water (affected by the presence of graphene and polarization effects, Figure 3.7) in the QM part, as we showed in Figure 3.8. These new electronic transport calculations are presented in Figures 3.12 and 3.13, which also present the previous results (without QM water) for comparison.

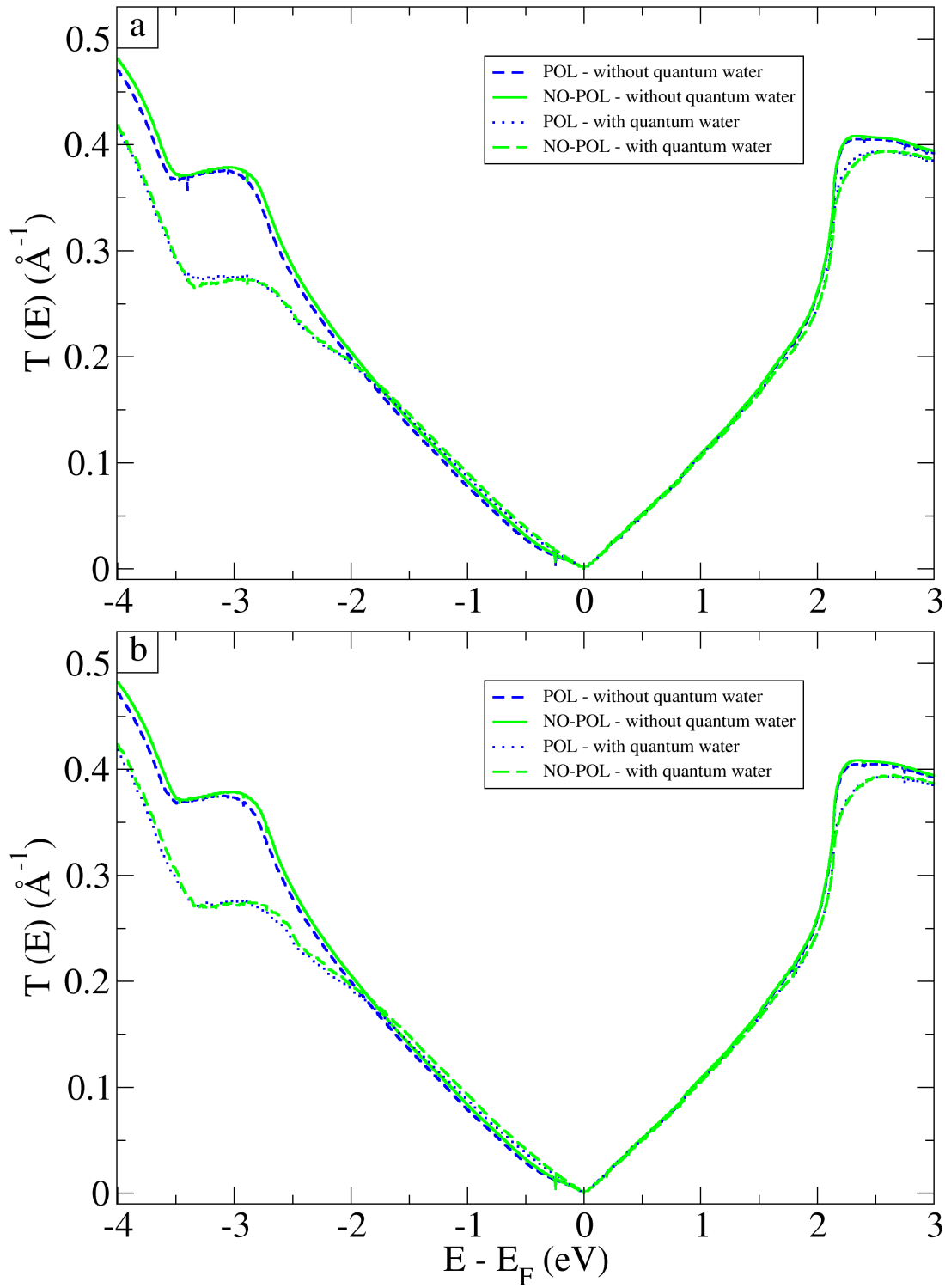


Figure 3.12: Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene including the first layer of water in the QM part in (a) pure water and (b) water with 0.2 *M* of NaCl. The cases without QM water are repeated here for comparison.

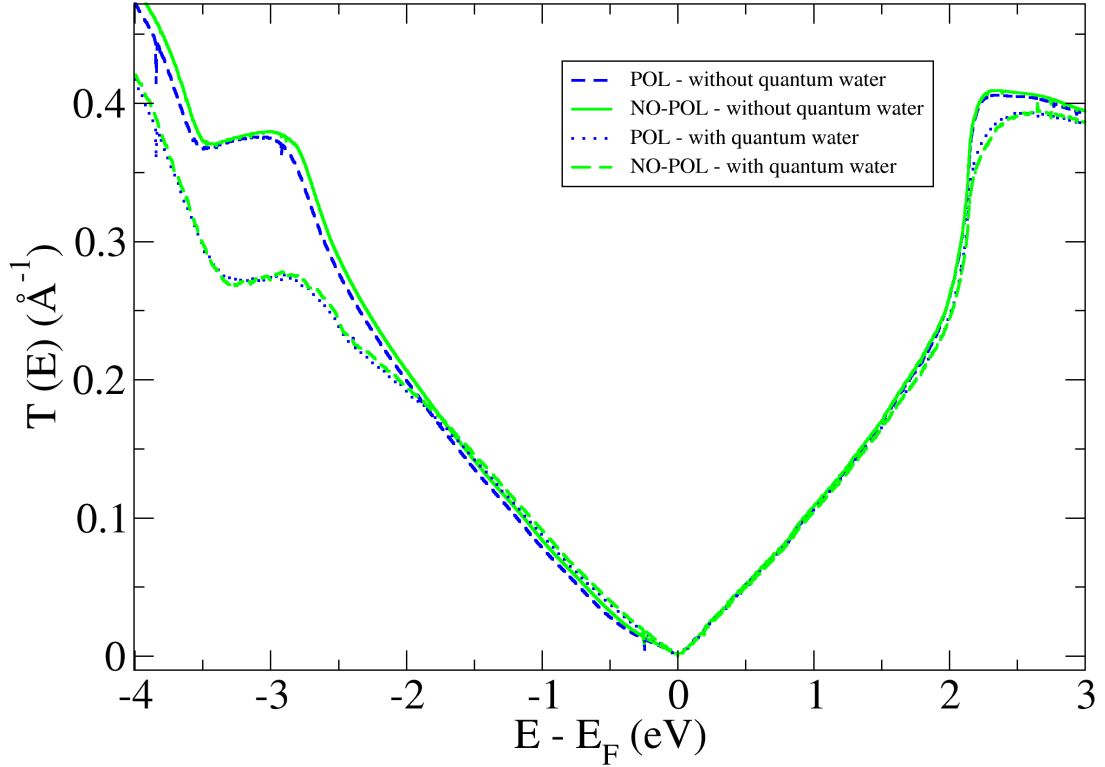


Figure 3.13: Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed pristine graphene including the first layer of water in the QM part in water with 1 M of NaCl. The case without QM water is repeated here for comparison.

As Figure 3.12 and 3.13 shows, the transmission for all cases with QM water are slightly bigger than the transmissions for the cases without QM water, for energies close to the Fermi energy (E_F). However, for energies far from E_F (below -2 eV and above 1.5 eV), this behavior is the opposite. In the previous case, graphene has a strong tendency for n-doping and the electrostatics from water favors that. However, in the presence of QM water, graphene is actually allowed to exchange charge with the environment. In that case, there is a clear competition as water is electronegative and does not favor transferring electrons to graphene, the overall effect is then to return the Dirac cone closes to the Fermi level. At the same time, water levels are more present at approximately 3 eV below the Fermi level thus leading to an increase in scattering.

The differences between polarized and non-polarized transmissions in Figures 3.12 and 3.13 are also small, even in the case in which we include the first layer of water in the QM part. Even including these structures, which are affected by the presence of polarization effects (see Figure 3.7) the electronic transport properties of graphene are not strongly affected.

3.4.4 Convergence analysis

A convergence analysis for the case of fixed non-polarized graphene in 0.2 M of NaCl is presented in Figure 3.14. This analysis was made only for this system as example, but

all the other systems in our study of graphene in solution are similar.^{III} We used up to 200 frames in our calculations and all these curves for this example system are shown in Figure 3.14a. It also shows the average over these curves (green curve). As we can observe, for occupied states (below E_F) the electronic transport presents a large deviation, while for unoccupied states, the curves are close to the average.

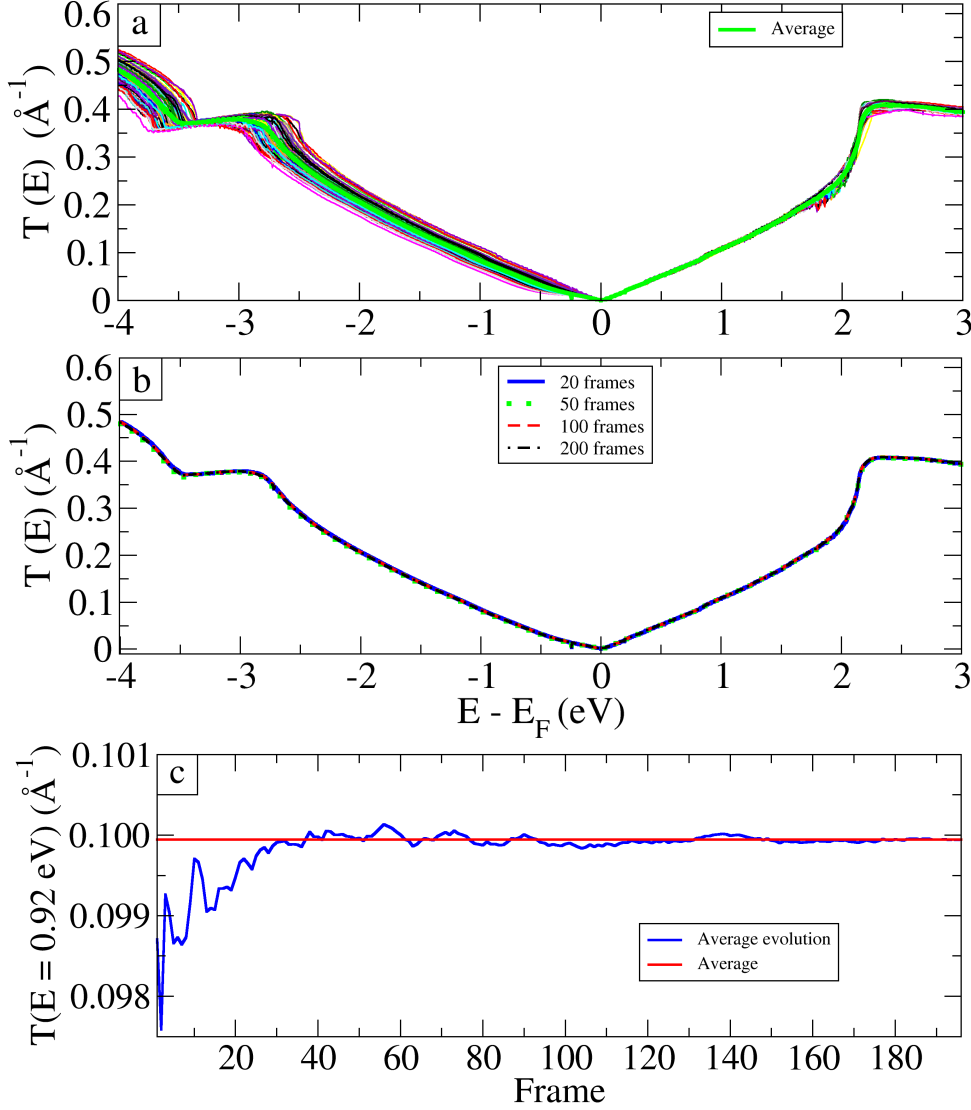


Figure 3.14: Convergence analysis example for fixed non-polarized pristine graphene in 0.2 *M* of NaCl. Electronic transport calculations for all frames and the average over these curves (a), averages over different numbers of frames (b) and evolution of the average for $E - E_F = 0.92$ eV (c).

Since we have a broad spectrum of electronic transport curves for these systems, we need to use a number of structures (liquid fluctuation) in orders to guarantee convergence. An analysis to verify that is done by plotting the averages over different numbers of equally spaced curves (uncorrelated), as Figure 3.14b shows. In this figure we can see averages over

^{III}For the next systems of study that we will present in this thesis we analyze the convergence for all cases because of the presence of DNA molecules. In these cases we have to guarantee that the number of different conformations for DNA are enough to sample the phase space of the systems.

20, 50, 100 and 200 curves. We notice that from this graph the curves look virtually converged for a small numbers of curves.

A further analysis can be done by choosing one specific energy and following the evolution of the average transmission. As we can observe in Figure 3.14c, for a representative energy value does not present significant changes in the transmission after approximately 50 curves and is very similar to the case with 200 curves.

3.5 Effect of Stone-Wales defect on graphene

Up to this point we have presented our results for solvation effects using pristine graphene. As we discussed, the presence of defects in graphene can create chemical active areas, what is supposed to change the interfacial interactions at that region. In this manner, we also investigated solvation effects on graphene with SW [80] defect.

3.5.1 Structural properties

For the case of graphene with SW defect we analyzed the structural properties using all the water molecules of the first solvation shell and also using only the water molecules close to the defect, as we can see for a representative snapshot in Figure 3.15, in order to understand the effect of the defect in graphene on the water molecules organization.

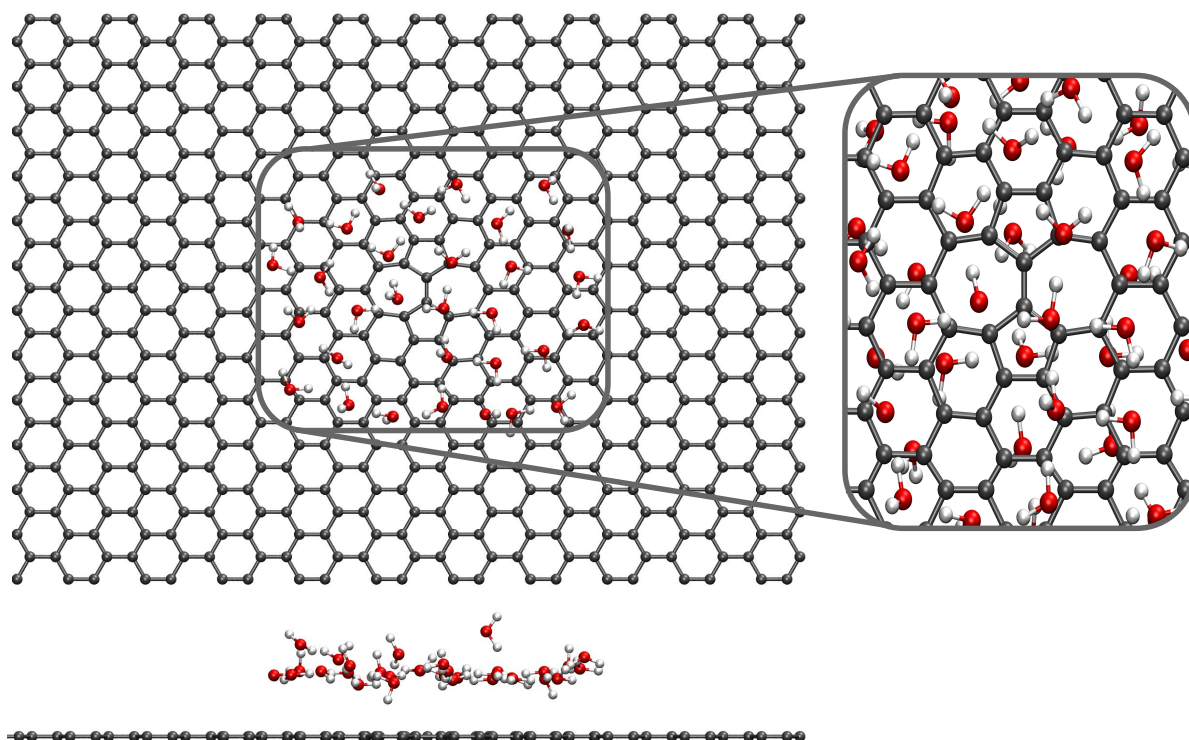


Figure 3.15: Top and side views and a zoom at the defect for the water molecules close to the Stone-Wales defect in fixed non-polarized graphene.

The distribution probability for OH tilt angles for the case of graphene with SW defect are shown in Figure 3.7. In this case, we can also observe the difference between polarized and non-polarized curves for the probability distribution, as we observed in the cases of pristine graphene (see Figure 3.7). The same can be seen for the case where we take into account only the OH tilt angles for the water molecules close to the defect.

Another aspect in Figure 3.16 is the fact that in the calculation using only the OH tilt angles for the water molecules close to the defect, we can see a decrease in the maximum probabilities, showing that the water molecule structure close to the defect are altered by its presence.

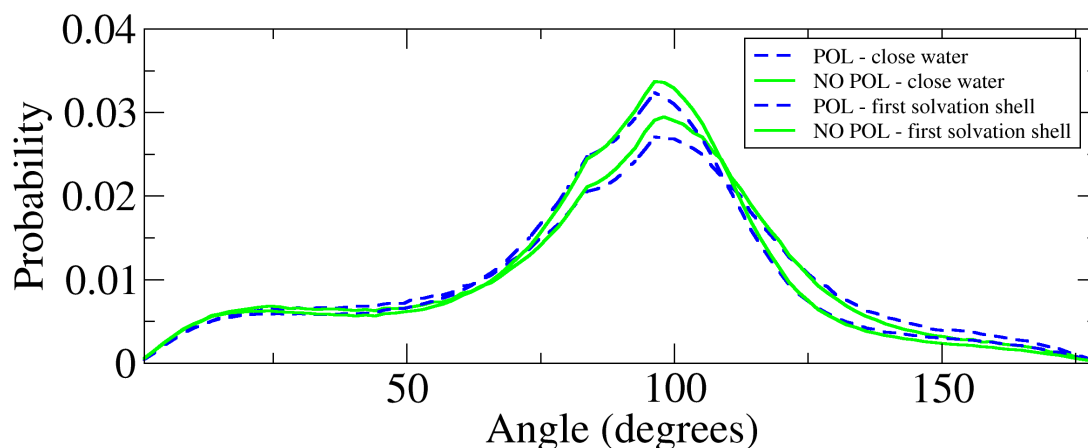


Figure 3.16: Probability for OH tilt angles of the first layer of water for fixed graphene with Stone-Wales defect in pure water comparing the cases with the entire first layer of water and the case including only the water molecules close to the defect.

3.5.2 Electronic transport results

The electronic transport calculations for polarized and non-polarized graphene with SW defect is shown in Figure 3.17. The difference between polarized and non-polarized curves is also small for this case with defect, which is supposed to alter the structure of water molecules close to the defect (see Figure 3.16).

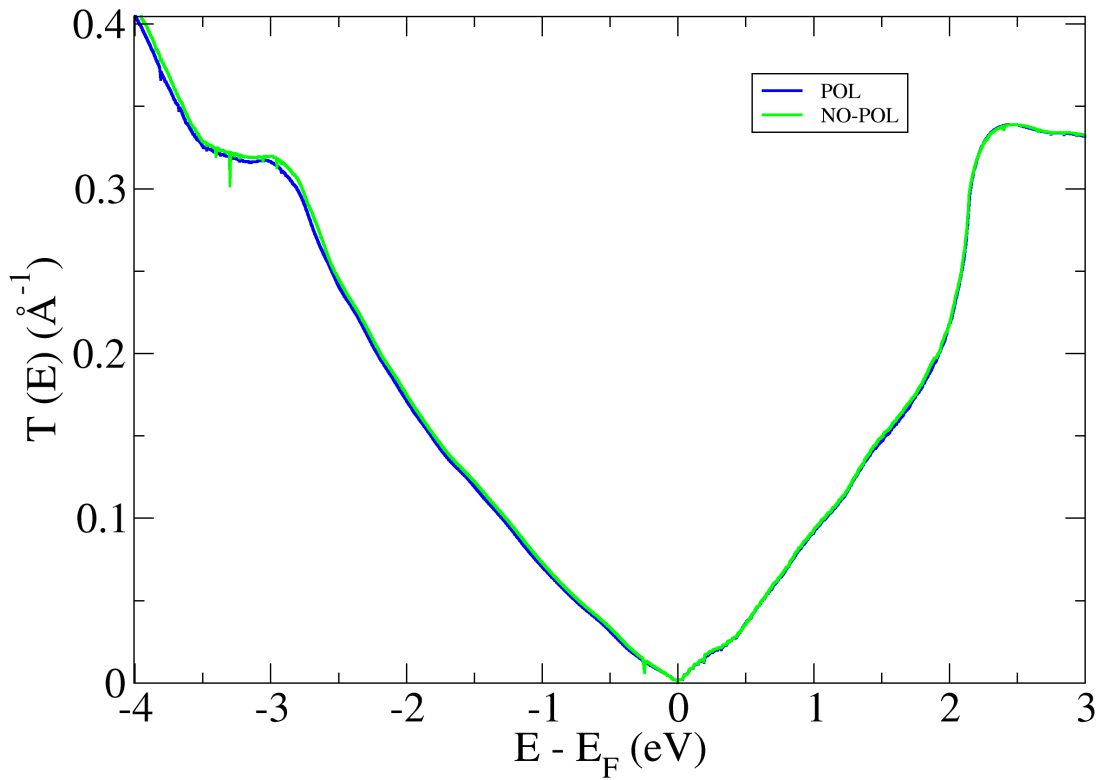


Figure 3.17: Comparison between the averages of electronic transport calculations for polarized and non-polarized fixed graphene with Stone-Wales defect.

In the same way we did before, here we also included one layer of water in the QM part, in order to explicitly count the effect of the water molecules with different structuring due to the presence of defect. We did this calculation only for non-polarized graphene. Thus, the electronic transport calculations are shown in Figure 3.18:

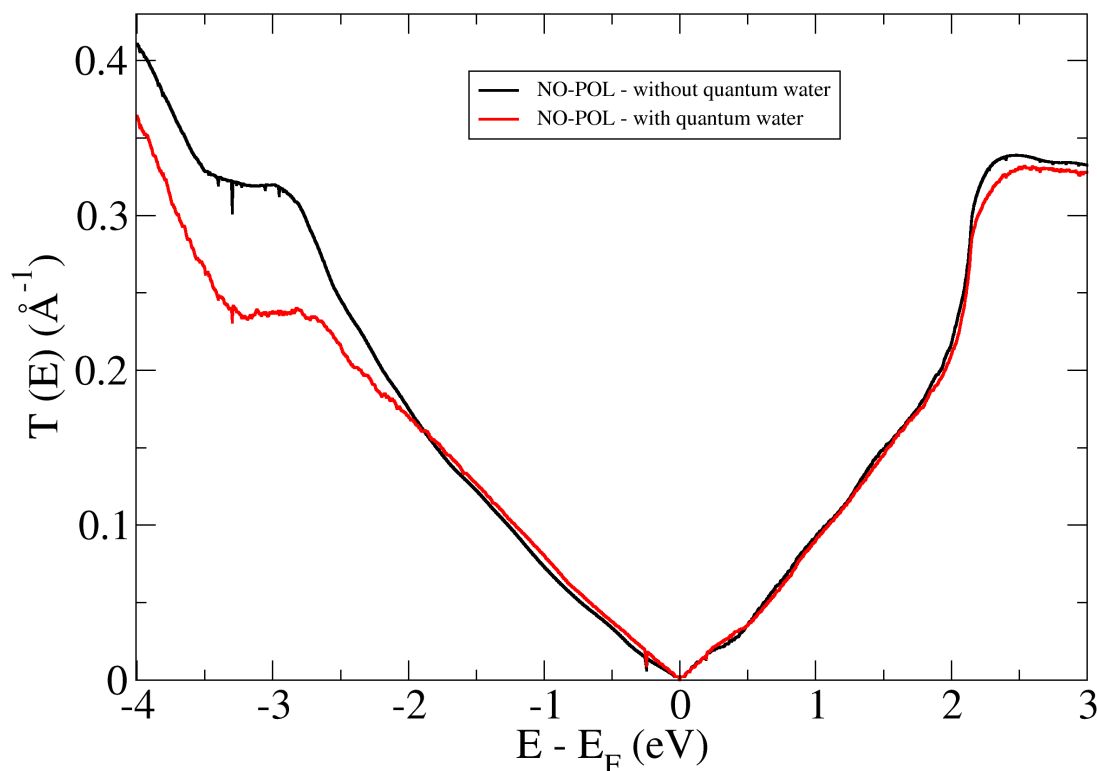


Figure 3.18: Comparison between the averages of electronic transport calculations for non-polarized fixed graphene with Stone-Wales defect in pure water and including the first layer of water in the QM part. The case without QM water was repeated here for comparison.

For this case, the effect of include water molecules in the QM part is the same observed before. As we can observe in Figure 3.18, for occupied states the inclusion of water molecules slightly increases the transmission, whereas the non-occupied states have virtually the same transmissions for curves with and without water molecules in the QM part.

3.6 Free graphene analysis

We also tested non-polarized free pristine graphene (see Figure 3.1g), as shown in Figure 3.19. In this figure the electronic transport calculations were made in two different conditions: in pure water and in vacuum. This is performed in the DFT step, either including or not the external potential (V_{ext}) calculated for the MM part. Thus, the structures for the QM part are the same. As in the case of the calculations including the first layer of water in the QM part, here only the carbon atoms from the central part of graphene are allowed to move, in such a way that the electrodes are fixed during the MD simulation. This guarantees equal electrodes for all calculations (or all snapshots).

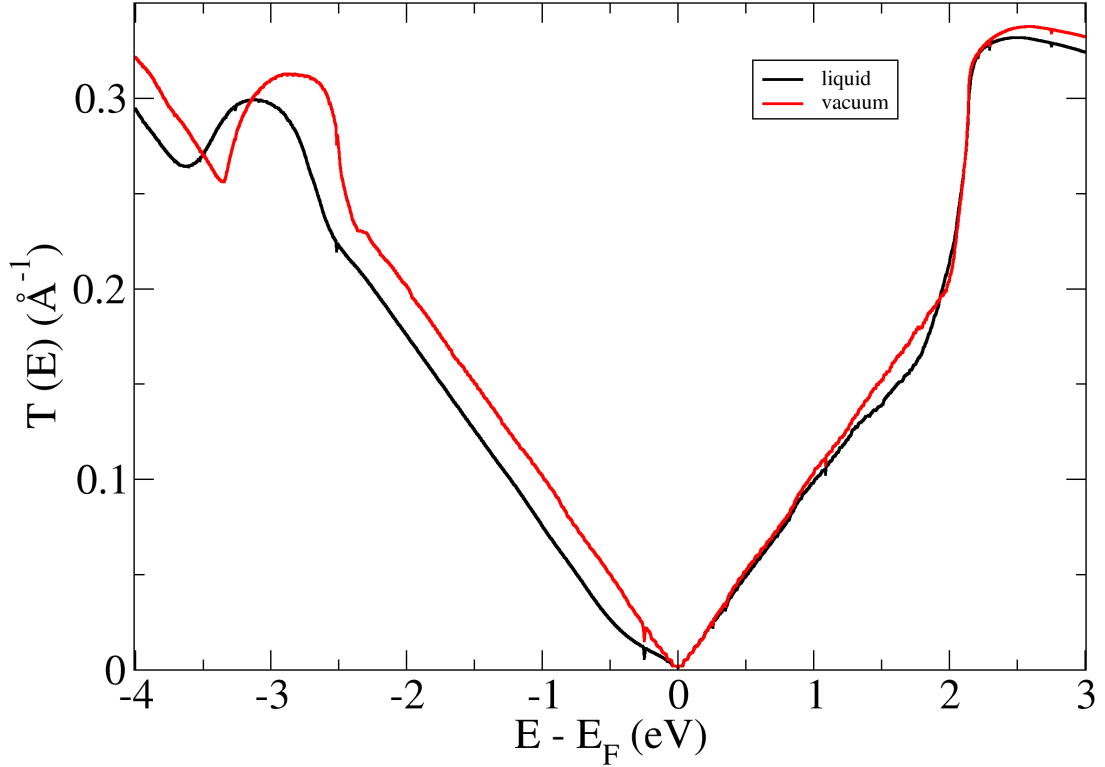


Figure 3.19: Comparison between the averages of electronic transport calculations for free pristine graphene in liquid (pure water) and in vacuum.

Figure 3.19 shows the average curves over 200 snapshots. There we can see that the average electronic transport calculations for free pristine graphene immersed in liquid and in vacuum are rigid shifts compared to each other. Moreover, this graphic also show that the scattering for the system in liquid is larger than the scattering for vacuum. This result is expected and it is in agreement with the results presented before for the same systems but for fixed graphene.

3.7 Final analysis and conclusions

In Figure 3.20 we summarize all results for electronic transport calculations in pure water: fixed polarized and non-polarized graphene with and without QM water (1, 2, 3 and 4), fixed polarized graphene with SW defect without QM water (7), fixed non-polarized graphene with SW defect with (9) and without (8) QM water, free non-polarized graphene in pure water (5) and in vacuum (6) and, finally, dry fixed graphene with SW defect and dry fixed pristine graphene as references. In Figure 3.20 all electronic transports are averaged over up to 200 frames, except for dry graphene with SW defect (11) and dry pristine graphene (10), in which we do not have statistics because these structures are fixed and there is no fluctuation due to the presence of a solvent.

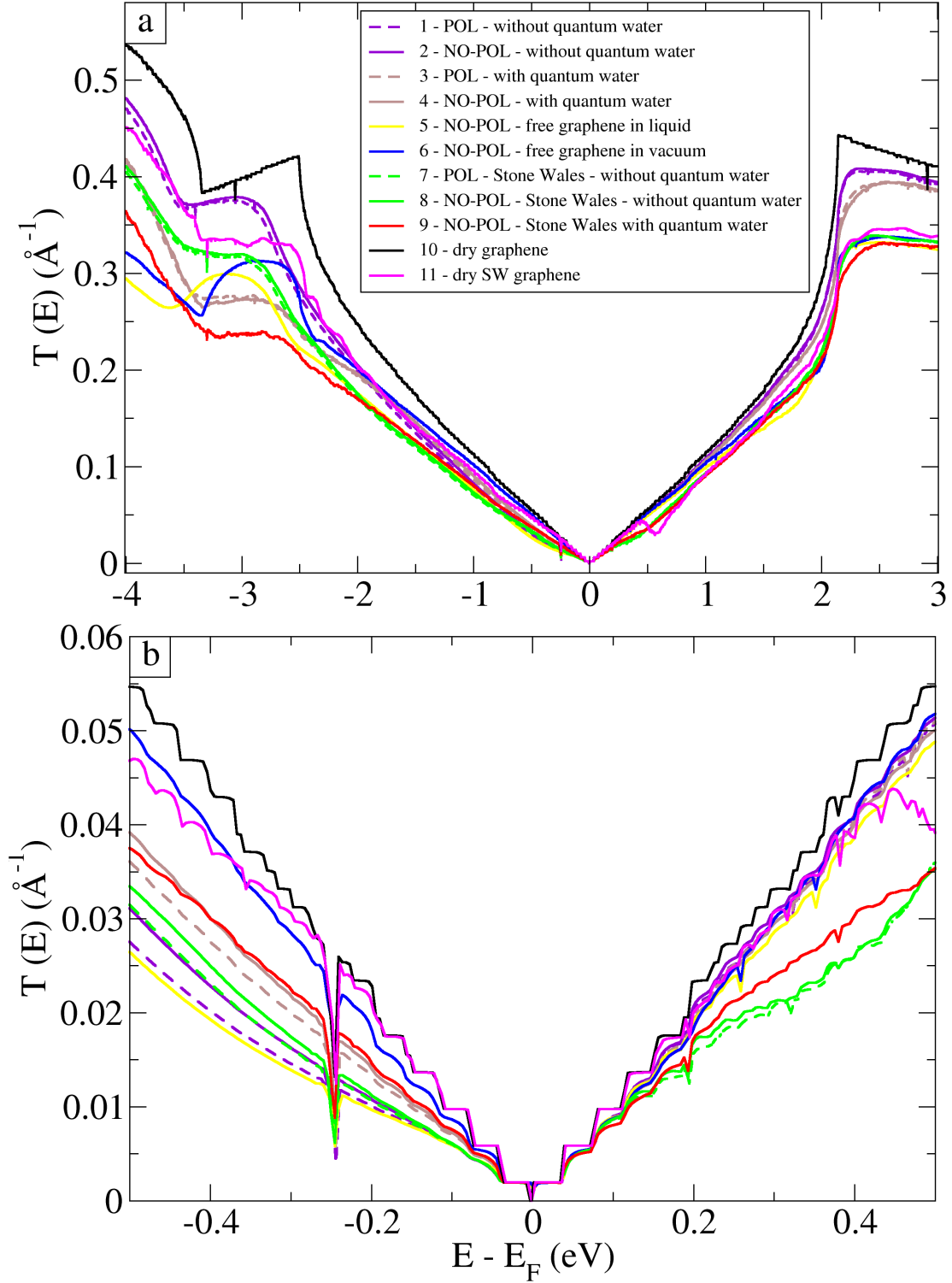


Figure 3.20: Comparison of the averages for all electronic transport calculations for the systems immersed in pure water (a) and a zoom close to the Fermi energy in (b).

The first aspect that we can see in Figure 3.20 is that the fixed dry pristine graphene has the largest transmittance (black curve), which is expected as the limiting factor for transmission in the number at available channels. The second aspect is that the free dry pristine graphene (blue curve) has a transmittance smaller than the case of fixed dry pristine graphene (black curve) for energies far from the E_f (Figure 3.20a) and a similar transmit-

tance for energies close to the E_F (Figure 3.20b).

The third aspect, still focusing on dry systems, is that the electronic transport for fixed dry graphene with SW defect (pink curve) is also larger than the cases of electronic transport for systems with SW defect in liquids (green dotted and solid lines and red line), which is also expected for the same reasons as in the cases without SW defect.

Thus, we showed that the presence of solvent has a strong influence on the electronic transport properties of pristine graphene and graphene with SW defect, but the salt concentration does not significantly alter that effect. We also showed that the inclusion of polarization effects on graphene does not significantly change the electronic transport properties of these systems, although we see differences in the interfacial water structuring by the OH tilt angle analysis. Even though the salt concentration has low effect on the electronic transport for graphene, its inclusion is important when we simulate biological systems, like DNA, since we have to simulate physiological conditions (≈ 0.2 M of NaCl) in which DNA is found in real systems. Then, in all simulations for DNA that we will present in the next chapters, we include the effect of the solvent simulating DNA in water with 0.2 M of NaCl.

CHAPTER 4

Graphene-based DNA sequencing: nanopore

4.1 Introduction

A big effort has been made in developing new DNA sensing devices in recent years [22, 24]. One possible way for sequencing DNA is using nanopores [3, 82–86], which can provide low-cost whole-genome sequencing. The proposals vary from measuring ionic currents as DNA trans-locates through a nanopore [87–89] to measuring transverse electronic currents across the device itself [90–92]. In this kind of device we use the term DNA sequencing, since here the idea is identifying the sequence of nucleotides.

As already mentioned, graphene is the ideal candidate for nanopore sequencing devices as it is an extremely strong membrane with excellent conductivity [6, 7]. In these devices, a nanopore saturated with Hydrogen (or other chemical element) is made on a graphene sheet in such a way that the DNA can translocate through the nanopore [88, 93]. Thus, for an all-electronic device using this topology to sequence DNA, the basic principle is detecting one nucleotide at a time, while the translocation process is occurring. This is possible because the electronic transport through the sheet with the nanopore will be altered by the presence of the nucleotides, since there will be interactions between the nucleotide and the nanopore.

The traditional DNA sequencing devices using nanopores have presented good results. However, from the electronic transport point of view, the non-localized electrons in graphene can be a problem in this kind of device, since the effect of introducing a nanopore can be low. This problem can be addressed decreasing the dimensionality. In this manner, the electric current can be controlled in the device. One way to do this is using graphene nanoribbons, but there are some limitations to fabricate it [94].

Alternatively to the traditional proposals for DNA sequencing devices using nanopores in graphene we can introduce line defects (Figure 4.1), formed by a combination of octagons and pentagons (OP) [95, 96] shown to be relatively stable. [97] The presence of line defects in graphene can lead to an enhancement in the selectivity of the device since has been shown that it acts as a conducting quasi-one dimensional (1D) wire, which decreases the number of scattering channels, in such a way that the electrons are confined in the defects [95, 98]. Thus, another way for detecting DNA is using nanopores in between line defects on graphene, as shown in Figure 4.1.

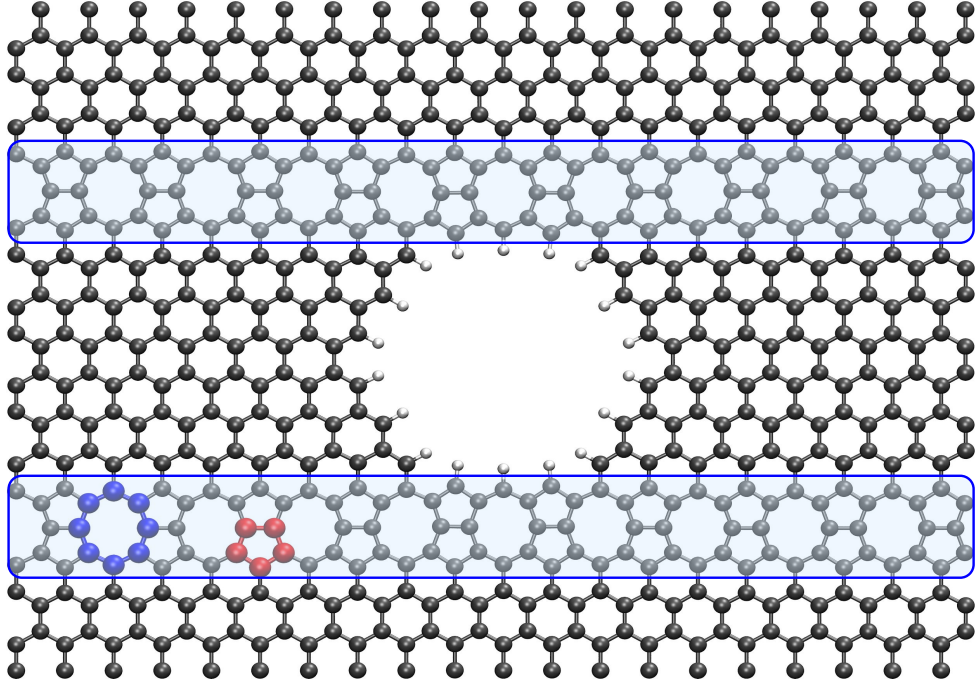


Figure 4.1: Schematic representation of the nanopore in between line defects (blue area) in graphene with one octagon in blue and one pentagon in red.

In this manner, we studied two different cases for DNA sequencing: the first one using topological line defects in graphene nanopore (Chapter 4) and the second one using graphene nanogap which will be presented in Chapter 5. In both cases our simulation model is based on the detection of isolated nucleotides. To simulate these systems using our protocol we always add water with 0.2 M of NaCl, as we explained, with one extra Na atom to equilibrate the phosphate group charge of DNA.

In all MD simulations, we used restraints applied to one central atom in the nucleotides to keep them inside the nanopore (or inside the nanogap in the next chapter). The force constant applied was $10 \text{ kJmol}^{-1} \text{ nm}^{-2}$. Therefore, we still allow conformational changes for the nucleotides. The whole sheet containing the line defects and the nanogap is kept fixed (the same for nanogap sheet in the next chapter) during the simulation.

4.2 The systems

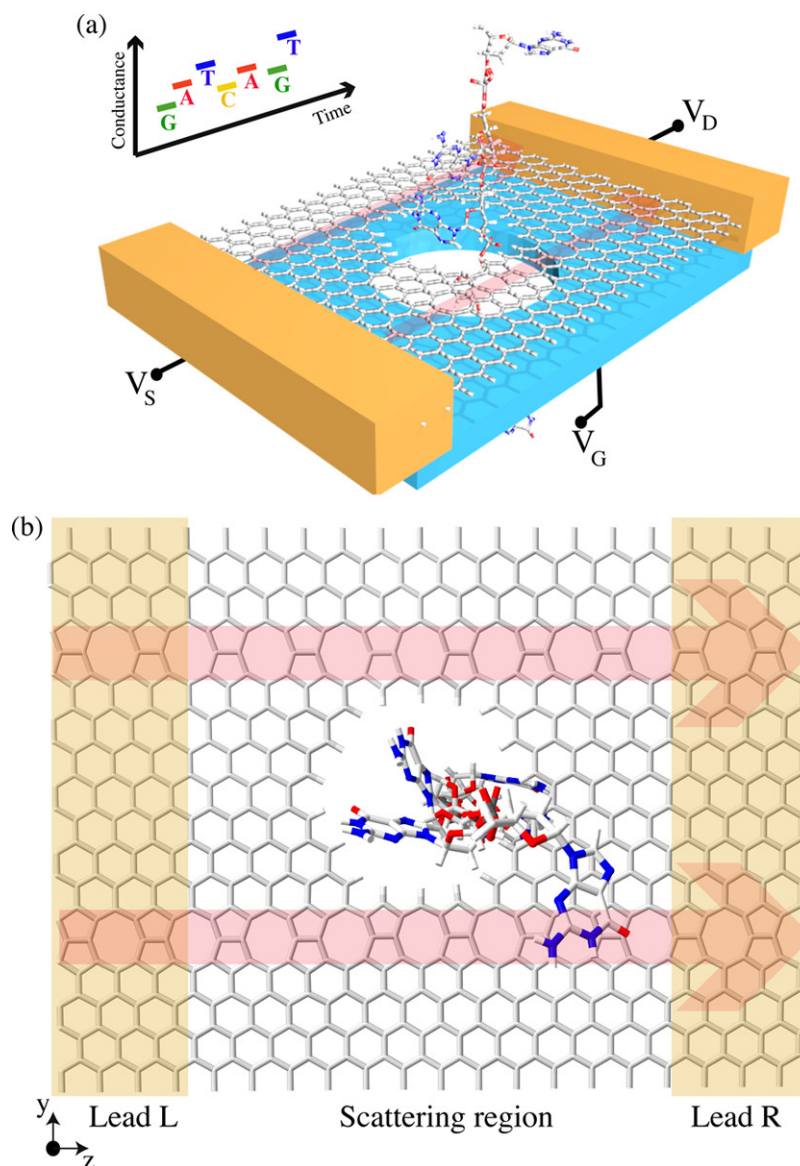


Figure 4.2: (a) Schematic representation for DNA sequencing device using a nanopore between topological line defects in graphene. (b) Setup used in our calculations showing the graphene sheet with a DNA strand inside the nanopore. Reprinted with permission from reference [99]. Copyright © (2018) American Chemical Society.

The idea behind this DNA sequencing device is shown in Figure 4.2. The system is connected to a source and drain (yellow) and to a gate electrode (blue). The idea is apply a bias voltage in the direction of the line defects during the process of translocating a DNA strand. We can also apply a gate voltage in such a way that the chemical potential can be tuned to achieve maximum sensitivity. When the DNA strand is translocated through the pore we can detect each nucleotide by measuring different conductances [3, 86, 89].

Figure 4.2b shows the actual device setup used in our calculations. It contains two line defects of the type OP (octagons and pentagons) and a nanopore saturated with Hydrogen

in between, with its edges close to the defects. The nanopore diameter is ≈ 12 Å. Here we do not simulate the translocation process and in our model the sequencing device is studied using isolated nucleotides one at a time inside the nanopore [92].

Our setup for the system with Hydrogen-terminated nanopore in between line defects in graphene, showed in the Methodology chapter (Figure 4.4) and in Figure 4.2b, was obtained from a geometry optimization. In our molecular dynamics simulations for these systems the sheet is also kept fixed. Thus we have a planar structure for the sheet.

We simulated our DNA sequencing device in the same way as before: placing isolated nucleotides inside the nanopore with a restraining potential applied in a atom close to the center of mass to keep them inside the nanopore but allowing conformational fluctuations due to the interactions between the nucleotide and the remaining atoms of the system (water, ions and graphene sheet). In this way, we also prevent π - π stacking interactions between the nucleotide and the sheet [100].

Here we include the effect of the solvent using a solution of water with 0.2 *M* of NaCl. We studied only one QM/MM setup the whole graphene sheet + DNA nucleotide + one Na atom in the QM part and the liquid (water + Cl atoms + remaining Na atoms) in the MM part.

In Figure 4.3 we can see a top view, zoomed in to the pore, of representative snapshots for all nucleotides inside the nanopore. In the same figure, we provide a side view of the entire simulation box. We use the same pattern of the other chapters to differentiate the atoms in the QM/MM partition, discussed above.

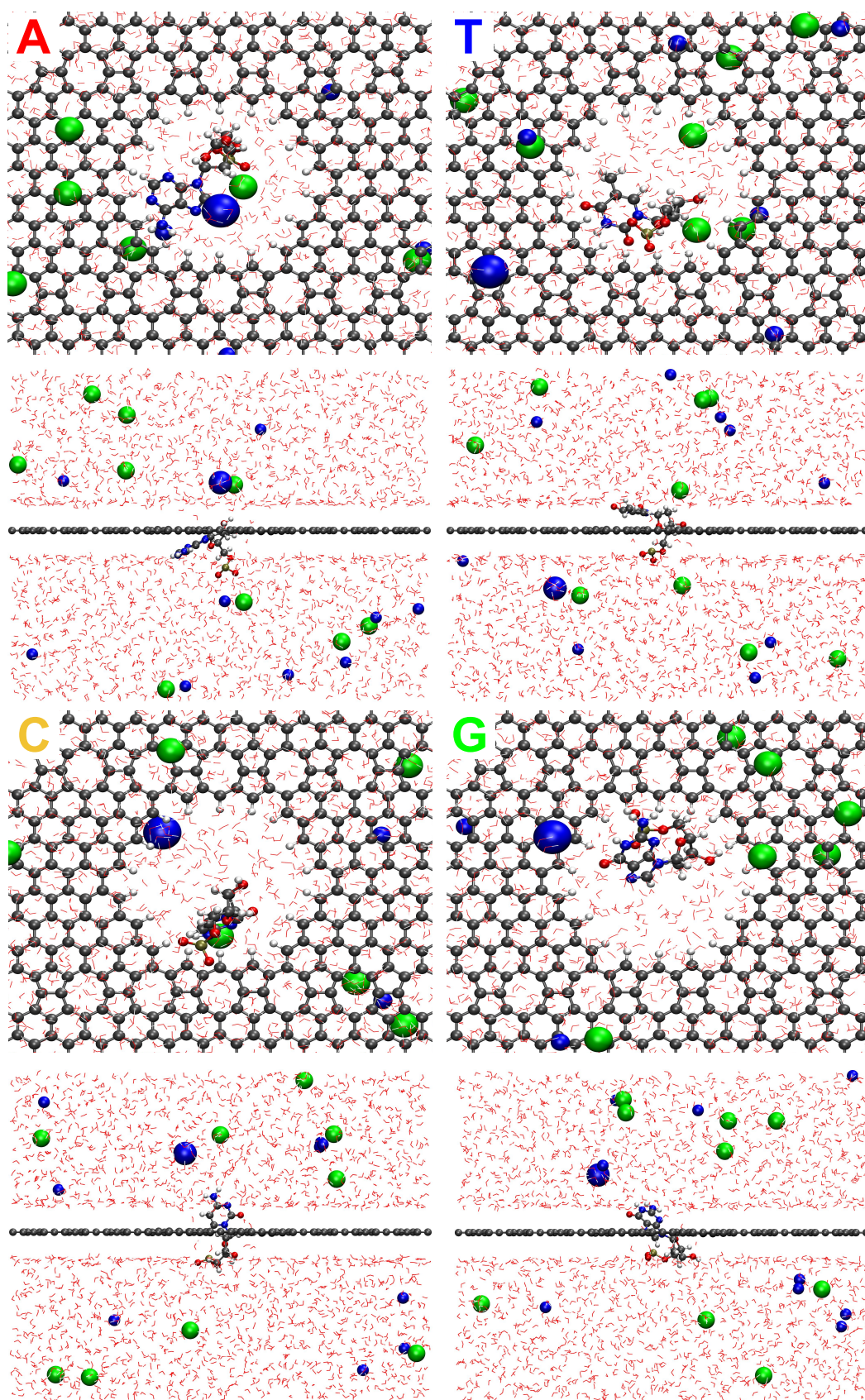


Figure 4.3: Top (zoomed at the nanopore) and side views for representative snapshots of all nucleotides used in the line defect DNA sequencing setup. The *ball-and-sticks* representation and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the *line* representations make up the MM part.

As can be observed in Figure 4.3 we have many different structural conformations for the nucleotides inside the pore. The conformations in which the nucleotide is located close to the line defects are supposed to enhance the electronic transport. Thus, a full dynamical treatment allow us to sample these structures that can contribute to enhance (or not) the electronic transport and, in average, we can estimate the contribution of all of them. The simulation details related with the systems presented in this chapter are shown in Tables 4.1-4.3.

MD calculations		
NPT	Berendsen barostat	20 ns
NVT	Nose-Hoover thermostat	20 ns
Time step	2 fs	
Force Field	AMBER99SB and SPC for water	
Snapshots and QM/MM partition		
50 snapshots	Equally spaced by 0.4 ns	
QM part	Nanopore sheet + DNA + n Na ions	
MM part	Cl ions and remaining water and Na ions	
DFT calculations		
Functional	GGA PBE	
Basis set	SZP for graphene sheet and DZP for the remaining atoms	
k-points grid (x y z)	1 1 2	
Electronic transport calculations		
Number of energy points	400	
Number of k-points	1 (Γ point)	
Energy range	From -0.5 eV to 0.5 eV around Fermi Energy	

Table 4.1: Simulation details for the study of DNA sequencing using line defects in graphene nanopore.

Adenine	
Box dimensions (x, y, z)	(36.927, 38.900, 49.245)
N_{ions}	10 Na and 9 Cl
$N_{water\ molecules}$	2127
N_{DNA}	33
$N_{graphene}$	626 C and 18 H
Thymine	
Box dimensions (x, y, z)	(36.927, 38.922, 49.245)
N_{ions}	10 Na and 9 Cl
$N_{water\ molecules}$	2127
N_{DNA}	33
$N_{graphene}$	626 C and 18 H
Cytosine	
Box dimensions (x, y, z)	(36.927, 38.997, 49.245)
N_{ions}	10 Na and 9 Cl
$N_{water\ molecules}$	2132
N_{DNA}	31
$N_{graphene}$	626 C and 18 H
Guanine	
Box dimensions (x, y, z)	(36.927, 38.972, 49.245)
N_{ions}	10 Na and 9 Cl
$N_{water\ molecules}$	2130
N_{DNA}	34
$N_{graphene}$	626 C and 18 H

Table 4.2: Details of the simulation box for the nucleotides in the study of DNA sequencing using nanopore. N is the number of the specified species. Length unit: Å.

Empty pore	
Box dimensions (x, y, z)	(36.927, 38.995, 49.245)
N_{ions}	9 Na and 9 Cl
$N_{water\ molecules}$	2143
N_{DNA}	0
$N_{graphene}$	626 C and 18 H

Table 4.3: Details of the simulation box for the empty pore in the study of DNA sequencing using nanopore. N is the number of the specified species. Length unit: Å.

The setup for electronic transport calculations for DNA sequencing using line defect in graphene nanopore is shown in Figure 4.4.

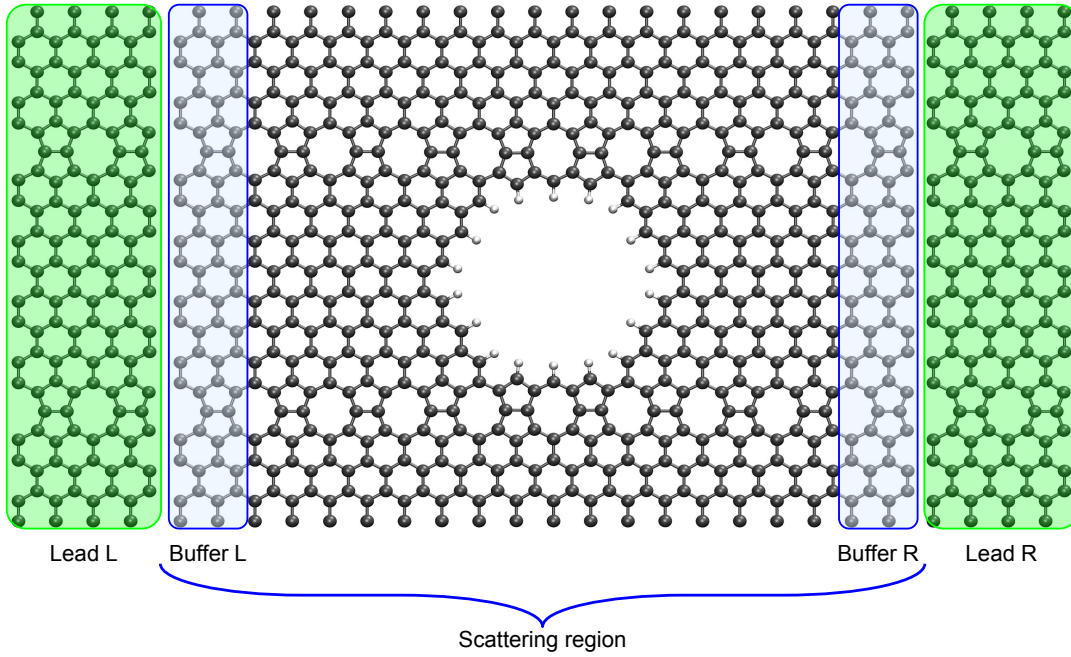


Figure 4.4: Electronic transport setup for the study of DNA sequencing using line defects in graphene nanopore. The green areas delimit the leads and the blue ones delimit the buffers. The scattering region is also shown. The line defects shown in this figure will be further explained.

4.3 Electronic transport results

All electronic transport calculations performed by us for all systems (A, C, T, G and liquid) are shown in Figure 4.5. The black curve shows the average over the other curves in the plot. The average curves are shown in Figure 4.6 together with a gray shaded region. In this figure we can estimate the range of electronic transport curves presented by these systems. We

note that there is significant fluctuations in $T(E)$ depending on the conformation. As done in Chapter 3 we also carry out a convergence analysis of the transmission. Those graphs are shown in Figure 4.7. We observe that despite the large fluctuations the average of the transmission converges rather quickly. In this figure we show the electronic transport averages over 10, 25, 40 and 50 equally spaced snapshots.

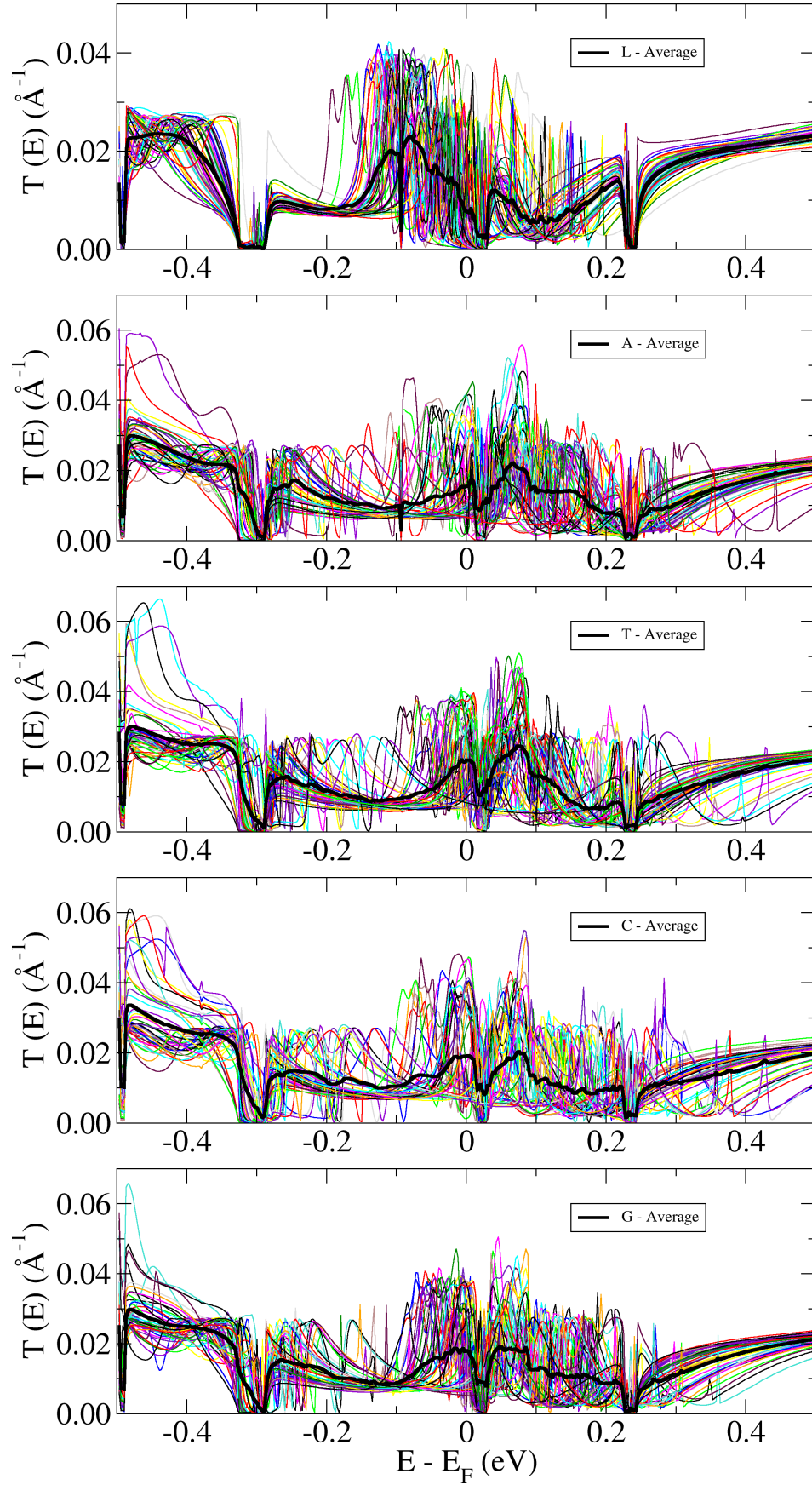


Figure 4.5: Electronic transport calculations for the liquid system (nanopore without nucleotide inside - L) and for all frames of each nucleotide defined by the letters used to refer each one. The black curve is the average over the presented curves.

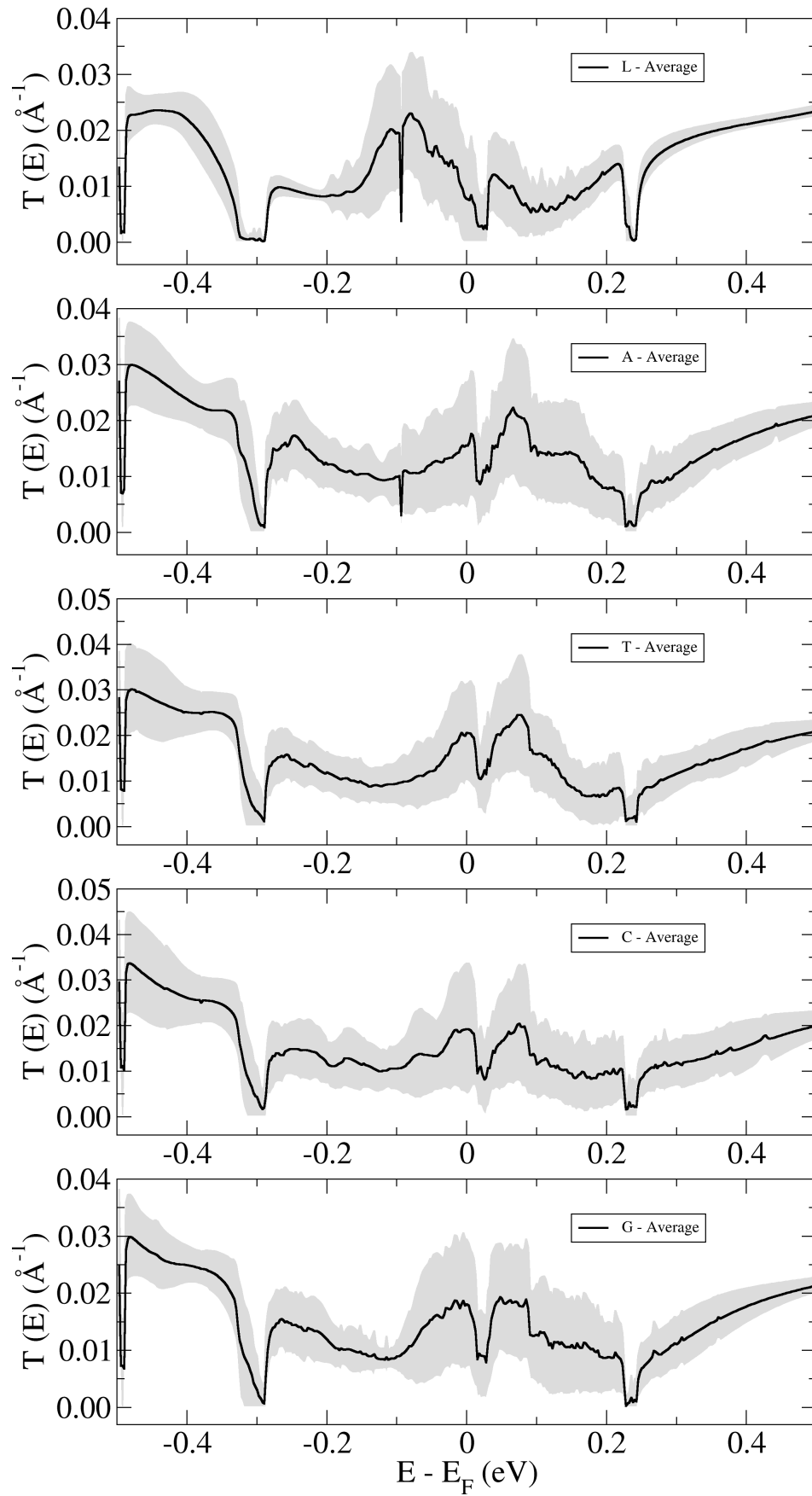


Figure 4.6: Electronic transport calculations for the system (L) system and for all nucleotides, defined by the letters used to refer each one. The gray shaded region comprises the addition and subtraction of 1 SD.

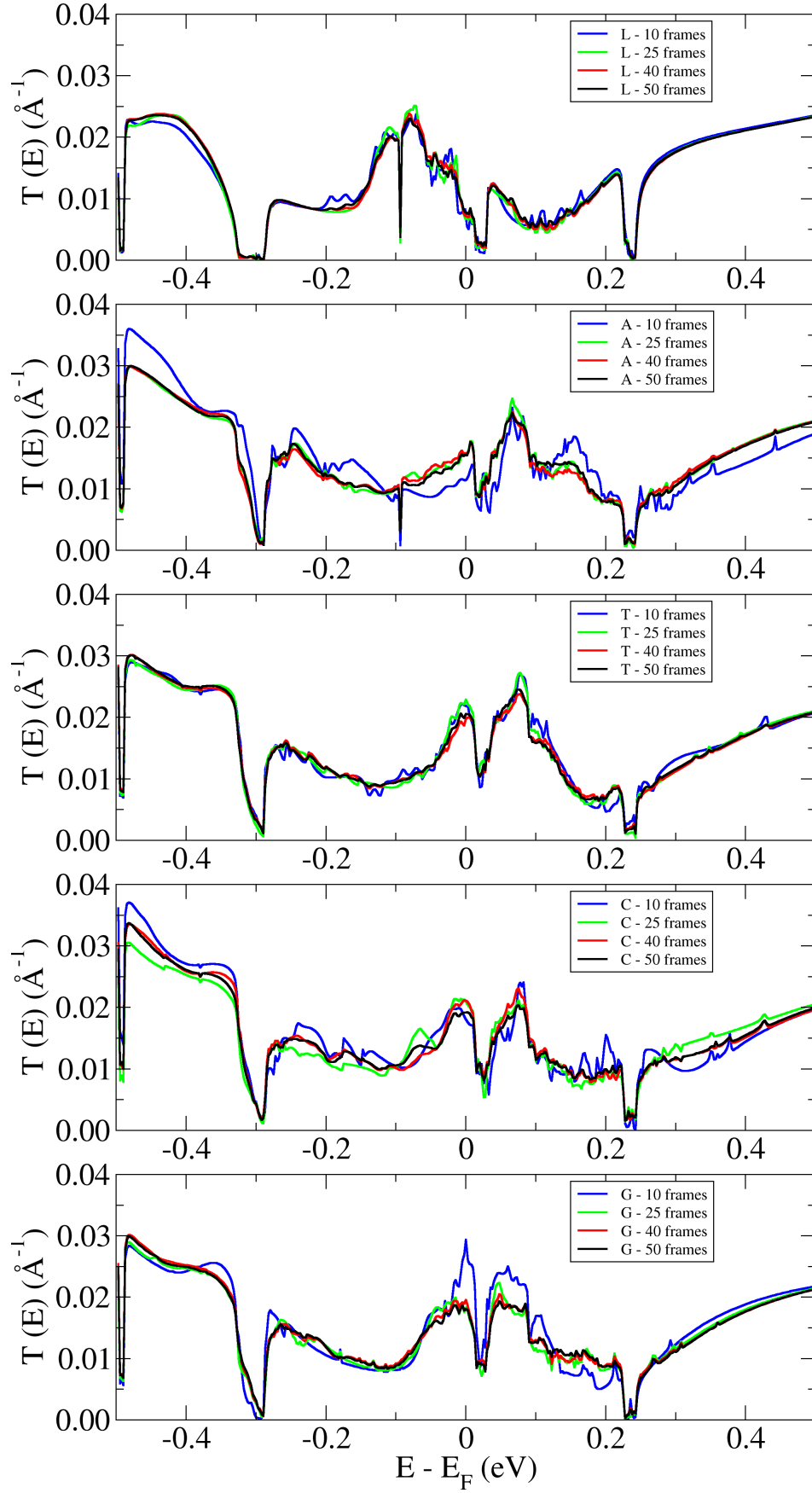


Figure 4.7: Electronic transport averages over different number of frames for the liquid system (L) and for all nucleotides, defined by the letters used to refer each one.

As a first analysis we can understand the effect of the solvent in the electronic transport properties of the empty nanopore (no nucleotide inside the nanopore). In this case we can estimate how the fluctuation of the liquid, included via the calculation of V_{ext} , affects the electronic transport of these systems. Thus, Figure 4.8 shows the comparison between the electronic transports for the empty nanopore in vacuum and the average in liquid (water with 0.2 M of NaCl).

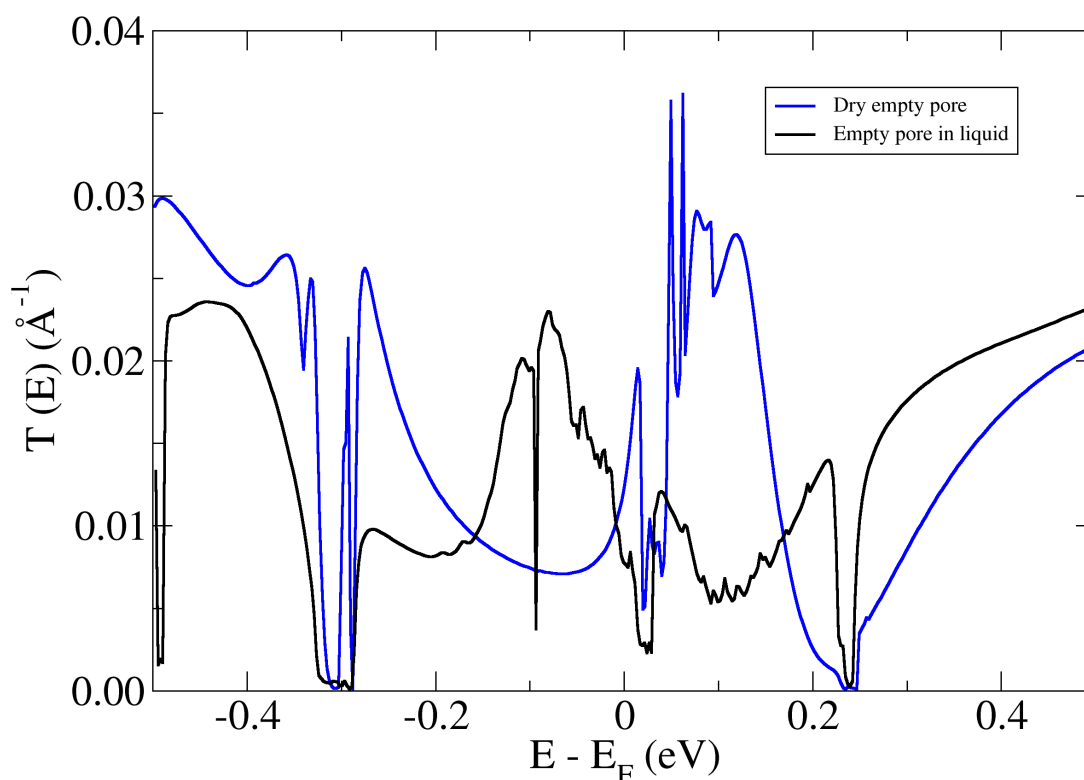


Figure 4.8: Electronic transport calculation for dry empty pore and the same immersed in liquid (average).

As can be observed in Figure 4.8 the inclusion of the solvent decreases the transmission, especially around the Fermi energy ($E = 0$ eV). As already discussed in Chapter 3, the main effect of water as a solvent is shifting downwards the curve thus the sheet contains the pore and the topological defects tends to become n-doped as pristine graphene. The overall trend for the black curve is qualitatively the same for the blue curve.

The averaged electronic transports calculated for the four nucleotides are shown in Figure 4.9. Interestingly, the presence of the nucleotides shifts the curves upwards. The main reason for this is steric as the molecule shields the pore from the effects of the external potential. It also comes from the fact that an electronegative QM object will present charge rearrangements as also seen in Chapter 3. Finally, as we can observe, for energies close to the Fermi energy, especially up to 0.1 eV, the four nucleotides present different average transport curves.

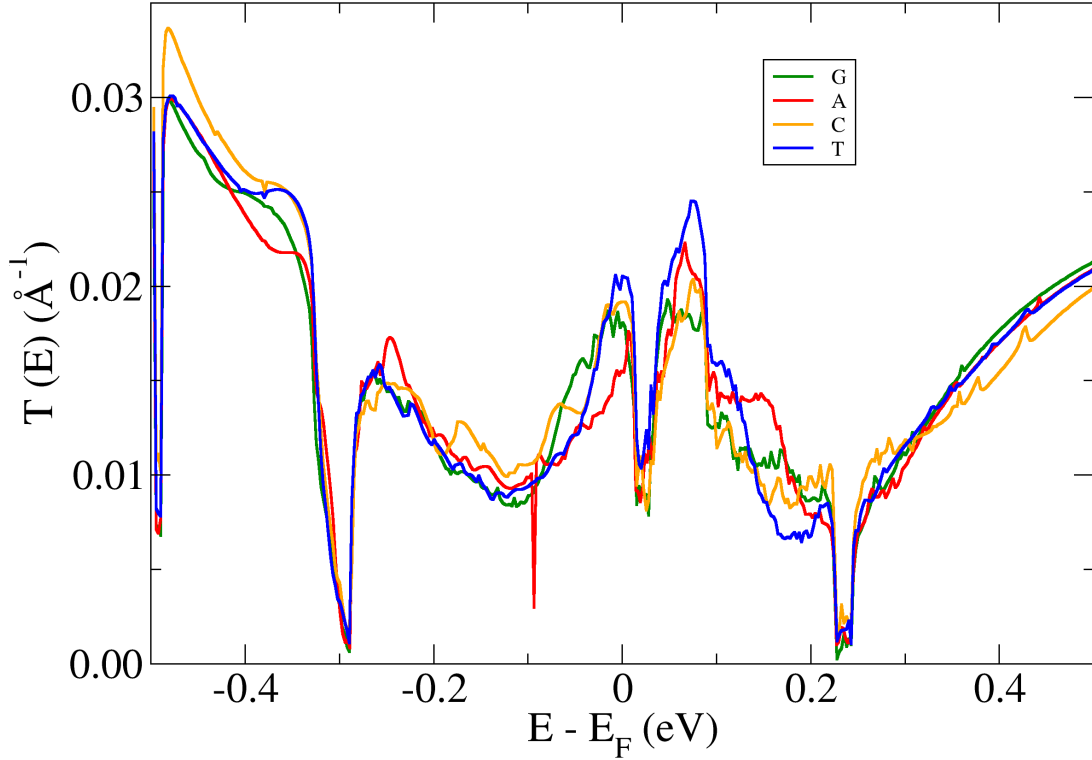


Figure 4.9: Comparison between the average electronic transport calculated for all nucleotides.

As we can observe, for energies close to the Fermi energy, especially up to 0.1 eV, the four nucleotides have different average transports. As showed in the paper (Figure 4.10), full relaxation calculations were performed¹ for the four nucleobases (instead of nucleotides) inside the pore. In this case, the liquid was not taken into account and therefore the results can be analyzed using optimal structures that comes from DFT relaxation.

As in the case of Figure 4.9, the results for dry system using the most favorable structures, shown in Figure 4.10, present the same overall trend. We have done these calculations using the same parameters and the electronic transport results were also obtained at zero-bias.

¹The calculations not taking into account the effect of the solvent (gas phase) were performed by our corroborators.

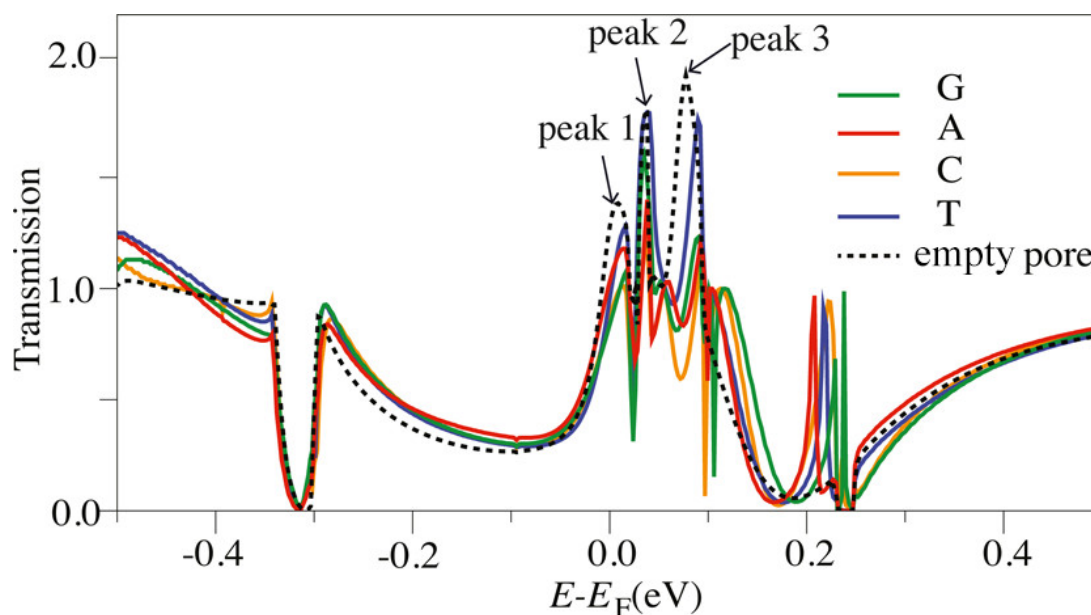


Figure 4.10: Comparison between the single electronic transport calculated for all nucleotides in gas phase using the most favorable structures. Reprinted with permission from reference [99]. Copyright © (2018) American Chemical Society.

4.4 Sensitivity and selectivity of the device

We can estimate the sensitivity and selectivity for this device using our averaged electronic transport calculations for specific energies. Thus, in Figure 4.11 we present the results for sensitivity (results for empty pore in liquid as reference) and selectivity (results for G as reference, not shown in the graphic). Then, Equations 4.1 and 4.2 show how we can calculate the sensitivity and selectivity using the average electronic transport for these systems:

$$Sensitivity(E') = \frac{\overline{T}_x(E') - \overline{T}_L(E')}{\overline{T}_L(E')} \times 100\%, \quad (4.1)$$

$$Selectivity(E') = \frac{\overline{T}_x(E')}{\overline{T}_G(E')} \times 100\%, \quad (4.2)$$

where $x = A, T, C$ or G .

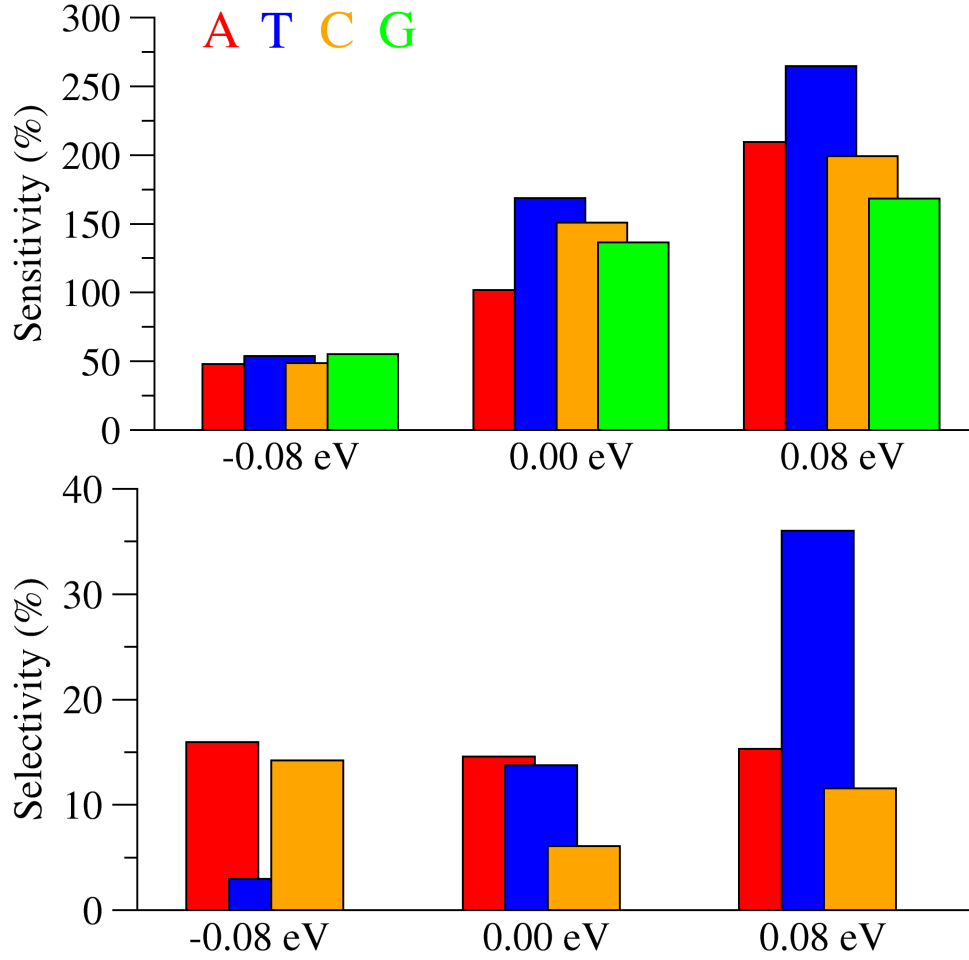


Figure 4.11: Sensitivity (upper panel) for the device using nanopore in between line defects in graphene for specific energies using the liquid system as reference (Equation 4.1) and selectivity (lower panel) for the same energies (gate voltages) using G as reference (Equation 4.2).

As we can observe, the largest sensitivity is obtained for $E - E_F = 0.08 \text{ eV}$ and at the Fermi energy. This result shows that the signals for the four nucleotides, in those energies, are different from the signal for the liquid. In other words, the device is extremely sensitive to the presence of the nucleotides at these energies. For selectivity we can observe that Adenine presents similar results for all energies. A and T presented low selectivities for $E = -0.08 \text{ eV}$ and at the Fermi energy. For $E = 0.08 \text{ eV}$, however, the selectivities are about 10 %, 15 % and 35 % for C, A and T, respectively.

4.5 Conclusions

Our results showed that for the chosen energies the signals for the four nucleotides are different between them, which indicates that by applying a gate voltage the device will be able to distinguish the DNA nucleotides. Our results also showed that sequencing DNA using a nanopore in between topological line defects in graphene appears possible, as we can see in the results for the overall profile for the electronic transport calculations, for selectivity

and for sensitivity calculation. Our results for these systems immersed in liquid corroborate with the results for dry systems.

The results presented in this chapter were done in collaboration with other authors and can be found in reference [\[99\]](#).

CHAPTER 5

Graphene-based DNA sequencing: nanogap

5.1 Introduction

Still focusing on DNA sequencing devices based on graphene, one can also envision the possibility of using nanogaps on the graphene sheet, as illustrated in Figure 5.1. Nanogaps are constructed placing two graphene sheets with saturated edges (with Nitrogen in the case of this figure) spaced by a gap. The idea is also to translocate DNA but, in this case, through the nanogap. Thus, for each DNA nucleotide we can have a different signal based on the electronic transport through the device.

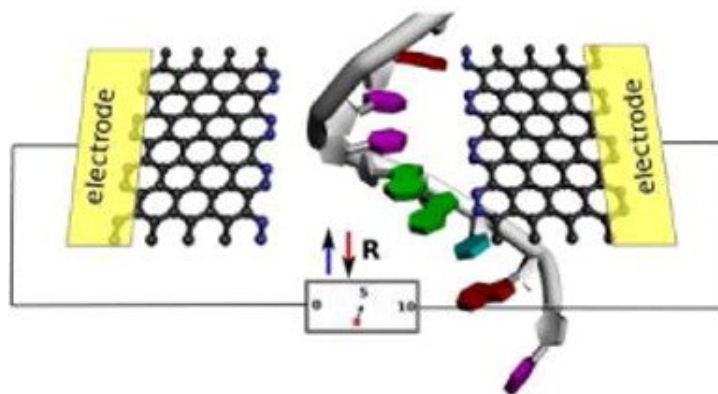


Figure 5.1: Illustration of the DNA sequencing device using Nitrogen-terminated nanogaps in graphene, showing the DNA strand translocating through the gap. Adapted with permission from reference [17]. Copyright © (2016) American Chemical Society.

The main idea behind a nanogap sequencing device is to measure the conductance differences across the nanogap while DNA strands pass through the nanodevice. The physical mechanism consists in electrons flowing from the left electrode to the right one (passing through the DNA and water molecules). The tunneling conductance may depend on the distance between the nucleotide and nanogap, on the electronic structure of each nucleotide and how its eigenstates are coupling with the nanogap. Therefore, the conductance decays exponentially with the distance, and the graphene nanogap [19, 101] device is a potential candidate for applications in this field due to its sensitivity and selectivity.

The tunneling process

Figure 5.2 provides a cartoon for a side view of the Nitrogen-terminated nanogap to explain the idea of DNA sequencing devices based on tunneling transport. In this figure Guanine is used as example, e^- represents the electron and the blue area represents the liquid (pure water with 0.2 M of NaCl).

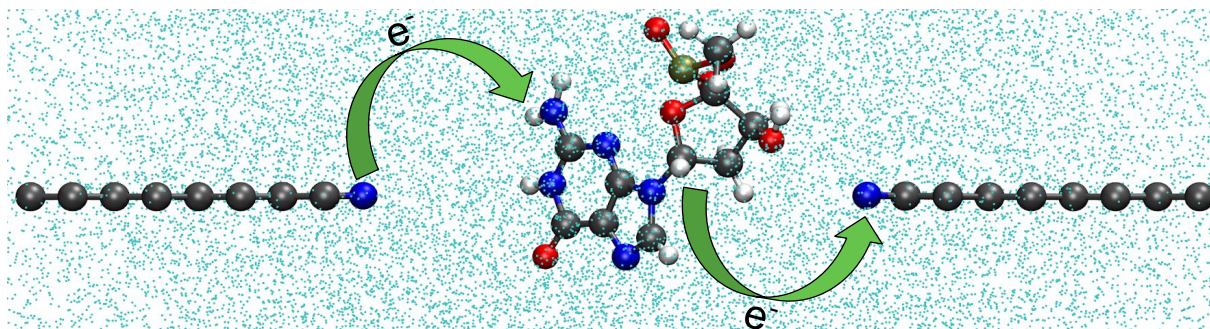


Figure 5.2: Scheme for tunneling process in DNA sequencing using nanogaps. The blue area represents the liquid (water and NaCl) and e^- represents the electron. The Guanine structure was used as example.

In the electronic transport picture, the electrons come from the left side of the device (left side in Figure 5.2), tunnel from the sheet to the liquid or to the nucleotide, are transmitted inside the nucleotide and then tunnel to the right side of the device. In this kind of device, the role played by water is crucial, but is frequently neglected on theoretical works [19, 101].

For dry systems, on the one hand, an optimal structure is found optimizing the system via DFT calculations, for example. In our work, on the other hand, we sample these systems using MD calculations and choosing uncorrelated snapshots which are equally spaced in time, as we discussed in the Methodology chapter (Chapter 2). In this way, making an average over all electronic transport curves that come from these chosen structures, up to a convergence, leads us to a good and reliable sampling.

Amorim [17] et al. have already investigated this DNA sequencing device using Nitrogen and Hydrogen-terminated nanogap in graphene. There the authors also simulated the translocation process by varying the position of the nucleobases inside the nanogap, showing that the transmittance changes orders of magnitude during the translocation. In this paper the authors have showed that Nitrogen-terminated nanogap in graphene enhances the device's sensitivity compared to a Hydrogen-terminated nanogap. In our simulations we have used the same setup used in the referred paper for the Nitrogen-terminated nanogap in order to investigate the effects of the solvent in this DNA sequencing device, using the DNA nucleotides instead of nucleobases.

5.2 The systems

In Figure 5.3 we present a top view for a representative snapshot of the four DNA nucleotides inside the gap and in solution. The isolated DNA nucleotides are positioned inside the gap and can move in a restrained potential. The force constant applied was $10 \text{ kJmol}^{-1} \text{ nm}^{-2}$. In this way we are able to sample the structures in which the nucleotides are placed inside the gap, but still allowing conformational changes in the structures and positions (rotations and smooth translations) of the nucleotides due to the liquid.

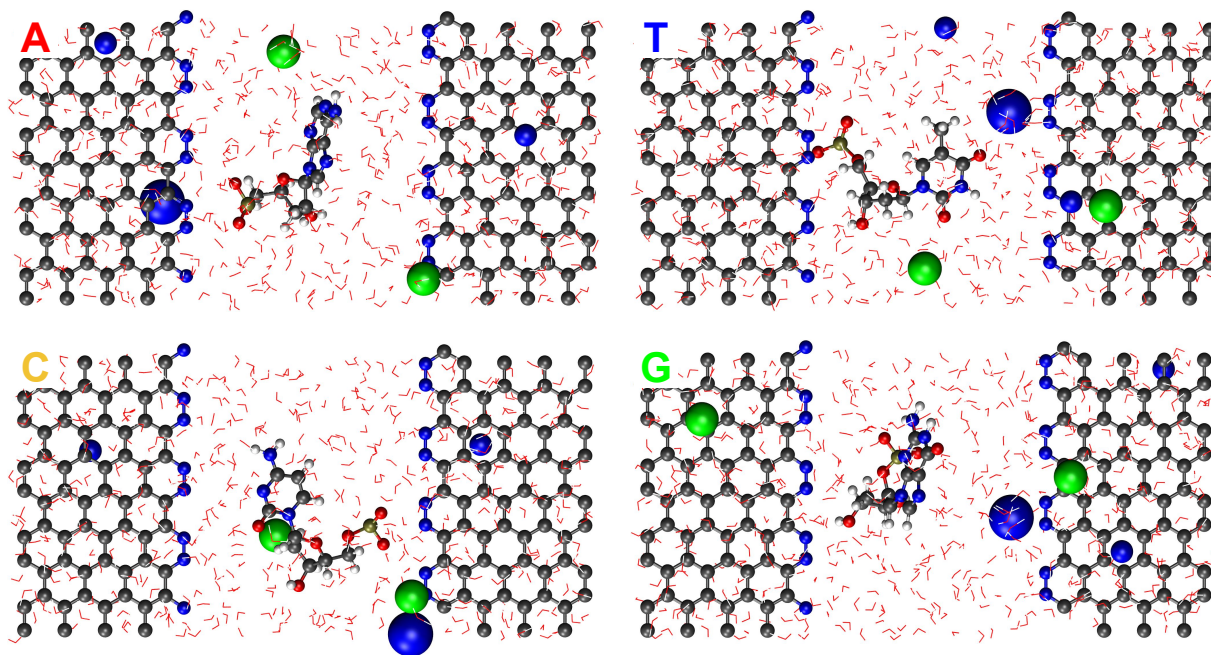


Figure 5.3: Top views for representative snapshots of all nucleotides (A, T, C and G) in the DNA sequencing device using nanogap. In this snapshot the setup 2 for QM/MM partition is shown. The large radius blue spheres and the *ball-and-sticks* representation make up the QM part while the small radius blue spheres, the green spheres and the *line* representations make up the MM part.

We also studied the empty nanogap (no nucleotide inside) immersed in the solution, as a reference system. This case, shown in Figure 5.4, was studied in only one setup: setup 3 in Figure 5.6. This is the unique setup in which we have transmission, as we will explain further in the text, because only in this case there are states for electron transport. There we include water molecules inside the nanogap in the QM part of the system. Since the distance between the two sides of the nanogap is ≈ 17 Å, the transmittance would be, by construction, zero as there will be no overlap between states on the left and on the right.

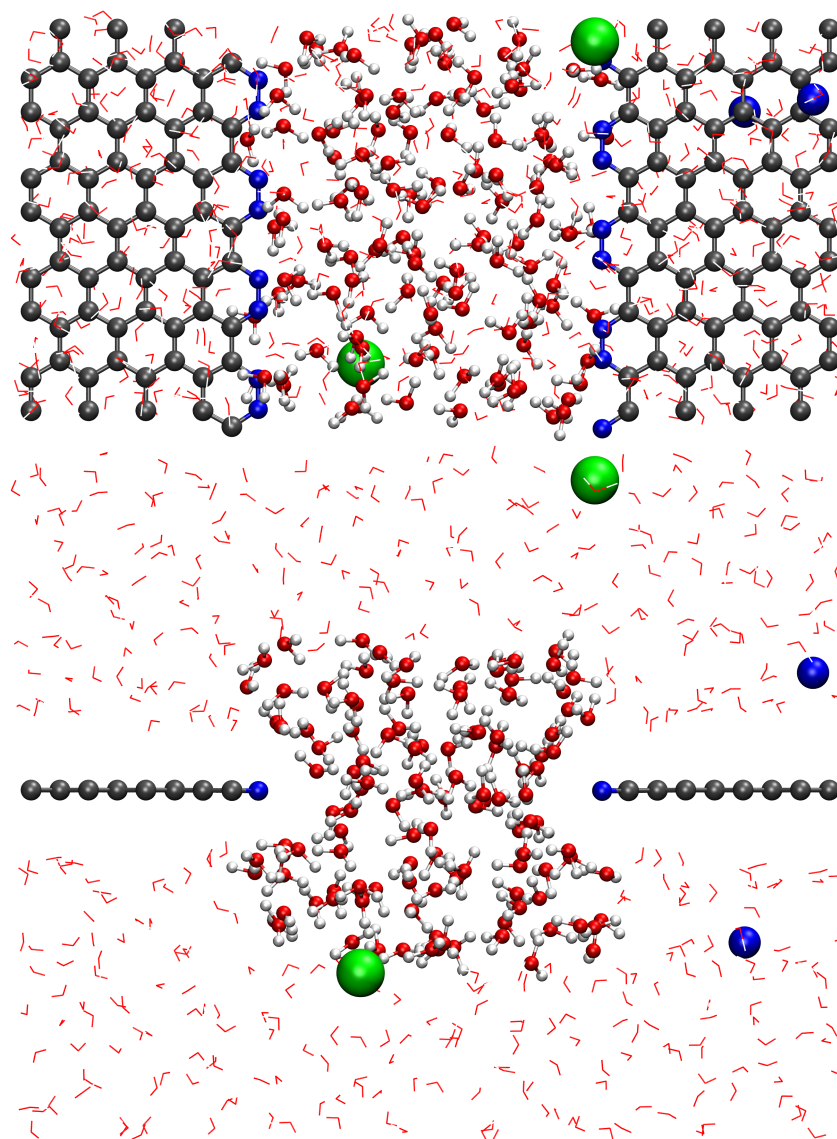


Figure 5.4: Top (upper panel) and side (lower panel) view for a representative snapshot for liquid system (nanogap immersed in water and NaCl without nucleotide) showing the unique QM/MM partition setup for this system (the same as setup 3 for the nucleotide systems). The water molecules in *ball-and-sticks* representation and the whole nanogap sheet make up the QM part, while the remaining atoms make up the MM part.

As we mentioned, we have studied this problem in different setups for the QM/MM partition. More specifically, we studied three different setups shown in Figures 5.5 and 5.6 namely: (1) system in gas phase, (2) solvated system with liquid totally included in the MM part and (3) solvated system with water partially included in the MM part. In all cases we used the same cutoff to select the water molecules to be included in the QM part: 5 Å from the sheet and nucleotide, since we have showed that this distance cutoff is enough to include the first solvation shell (Chapter 3).

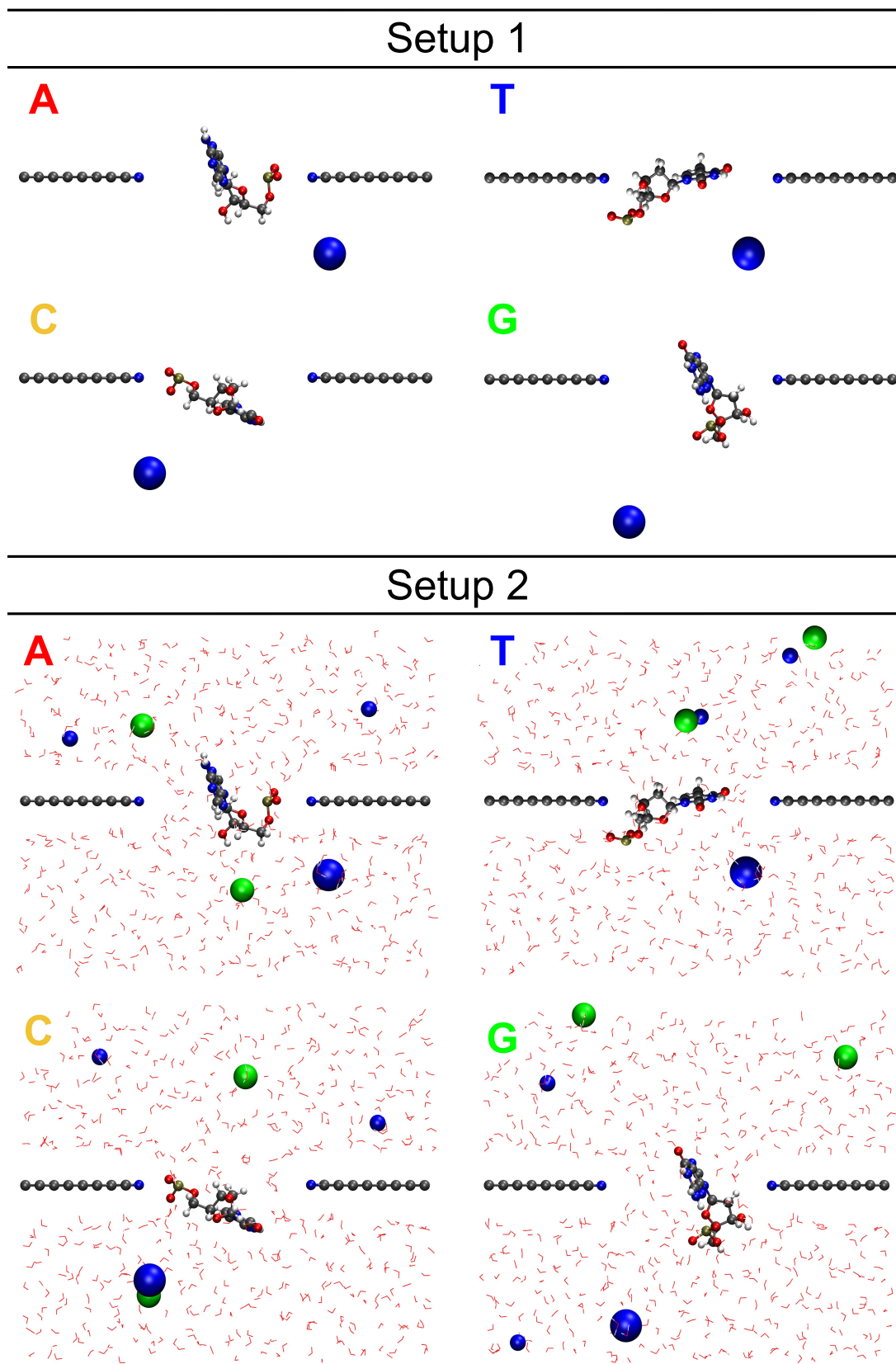


Figure 5.5: Side view of the first 2 setups for QM/MM partitions for all nucleotides. The *ball-and-sticks* representation and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the *line* representations make up the MM part.

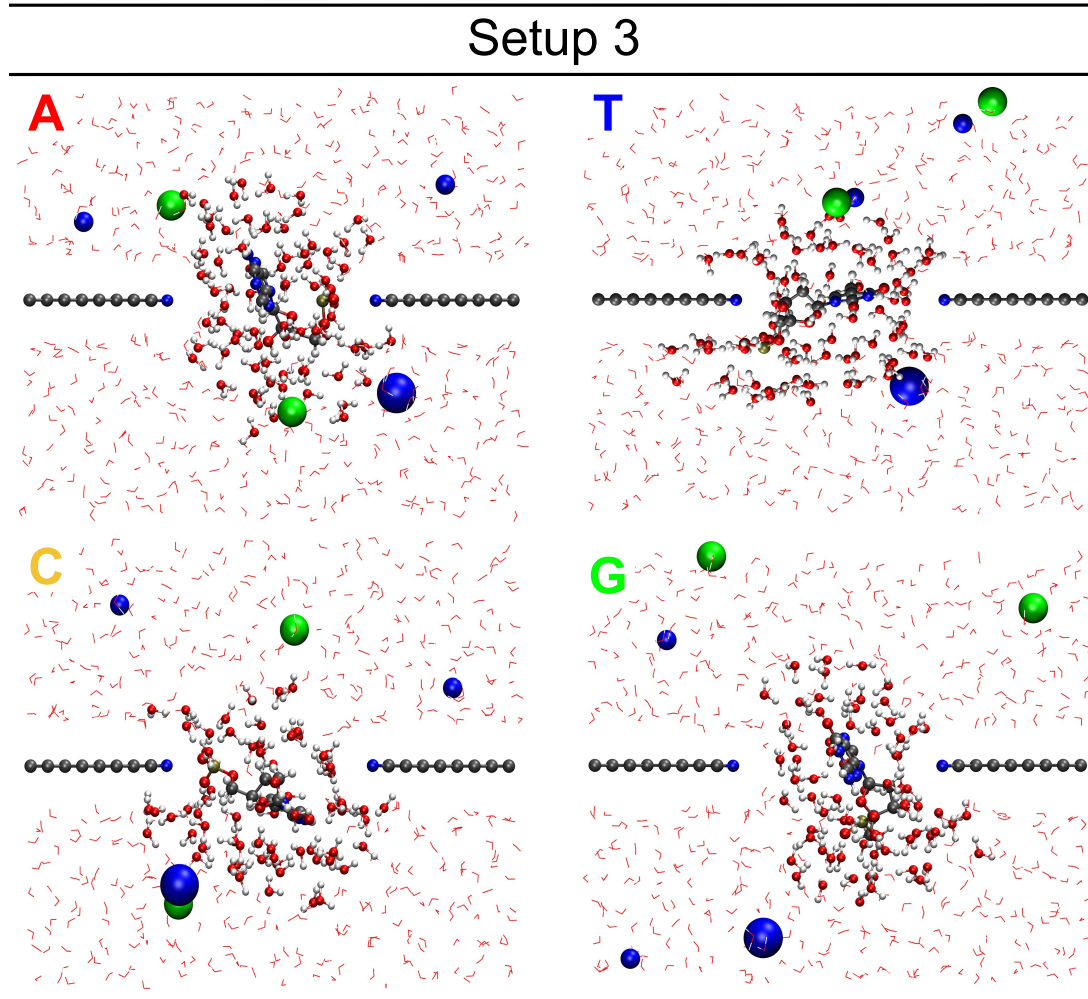


Figure 5.6: Side view of the third setup for QM/MM partitions for all nucleotides. The *ball-and-sticks* representation and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the *line* representations make up the MM part.

The simulation details related with the systems presented in this chapter are shown in Tables 5.1-5.3.

MD calculations		
NPT	Berendsen barostat	10 <i>ns</i>
NVT	Nose-Hoover thermostat	10 <i>ns</i>
Time step	2 <i>fs</i>	
Force Field	AMBER99SB and SPC for water	
Snapshots and QM/MM partition		
50 snapshots	Equally spaced by 0.2 <i>ns</i>	
QM part I	Nanogap sheet + DNA + <i>n</i> Na ions	
QM part II	QM part I + first layer of water	
MM part	Cl ions and remaining water and Na ions	
DFT calculations		
Functional	GGA PBE	
Basis set	DZ for C atoms of nanogap sheet and DZP for the remaining atoms	
k-points grid (x y z)	1 2 2	
Electronic transport calculations		
Number of energy points	800	
Number of k-points	1 (Γ point)	
Energy range	From -2 <i>eV</i> to 2 <i>eV</i> around Fermi Energy	

Table 5.1: Simulation details for the study of DNA sequencing using nanogap.

Adenine	
Box dimensions (x, y, z)	(17.252, 29.478, 35.872)
N_{ions}	3 Na and 2 Cl
$N_{water\ molecules}$	524
N_{DNA}	33
$N_{graphene}$	128 C and 16 N
Thymine	
Box dimensions (x, y, z)	(17.252, 29.548, 35.872)
N_{ions}	3 Na and 2 Cl
$N_{water\ molecules}$	525
N_{DNA}	33
$N_{graphene}$	128 C and 16 N
Cytosine	
Box dimensions (x, y, z)	(17.252, 29.486, 35.872)
N_{ions}	3 Na and 2 Cl
$N_{water\ molecules}$	525
N_{DNA}	31
$N_{graphene}$	128 C and 16 N
Guanine	
Box dimensions (x, y, z)	(17.252, 29.450, 35.872)
N_{ions}	3 Na and 2 Cl
$N_{water\ molecules}$	523
$N_{DNAatoms}$	34
$N_{graphene}$	128 C and 16 N

Table 5.2: Details of the simulation box for the nucleotides in the study of DNA sequencing using nanogap. N is the number of the specified species. Length unit: Å.

Empty gap	
Box dimensions (x, y, z)	(17.252, 29.709, 35.872)
N_{ions}	2 Na and 2 Cl
$N_{water\ molecules}$	539
N_{DNA}	0
$N_{graphene}$	128 C and 16 N

Table 5.3: Details of the simulation box for the empty gap in the study of DNA sequencing using nanogap. N is the number of the specified species. Length unit: Å.

The setup for electronic transport calculations for DNA sequencing using line defect in graphene nanopore is shown in Figure 5.7.

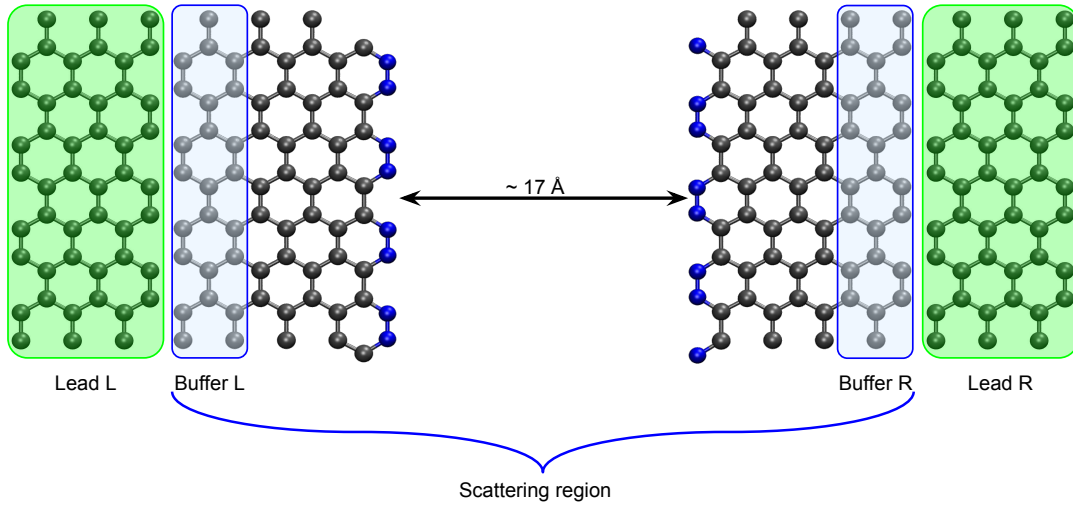


Figure 5.7: Electronic transport setup for the study of DNA sequencing using nanogap. The green areas delimit the leads and the blue ones delimit the buffers. The scattering region is also shown.

5.3 Electronic transport results

The electronic transport calculations averaged over different numbers of snapshots, for setups 1, 2 and 3, are shown in Figures 5.8 - 5.10, respectively. Here we used up to 50 snapshots to perform the calculations. As this system presents tunneling transport we expect an exponential dependence of transmission with respect to the length of the system in the transport direction. This means that small fluctuations in structure will yield very large fluctuations in transmission. As such a better quantity that describes the transport properties of these systems is $T_{ln}^i = \ln T_i$.

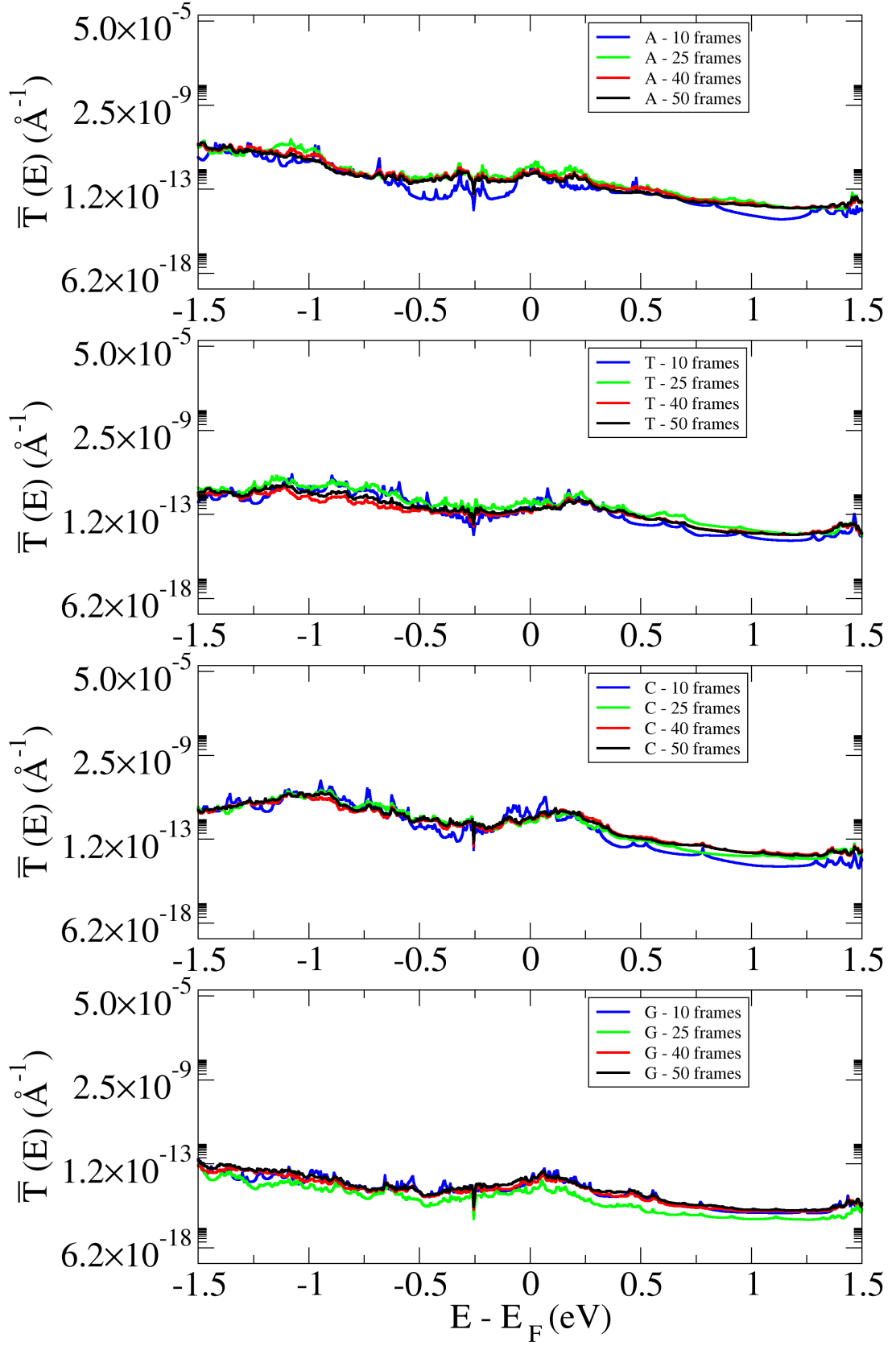


Figure 5.8: Electronic transport averages over different number of frames for all nucleotides in setup 1. The letters defined before identifies each system.

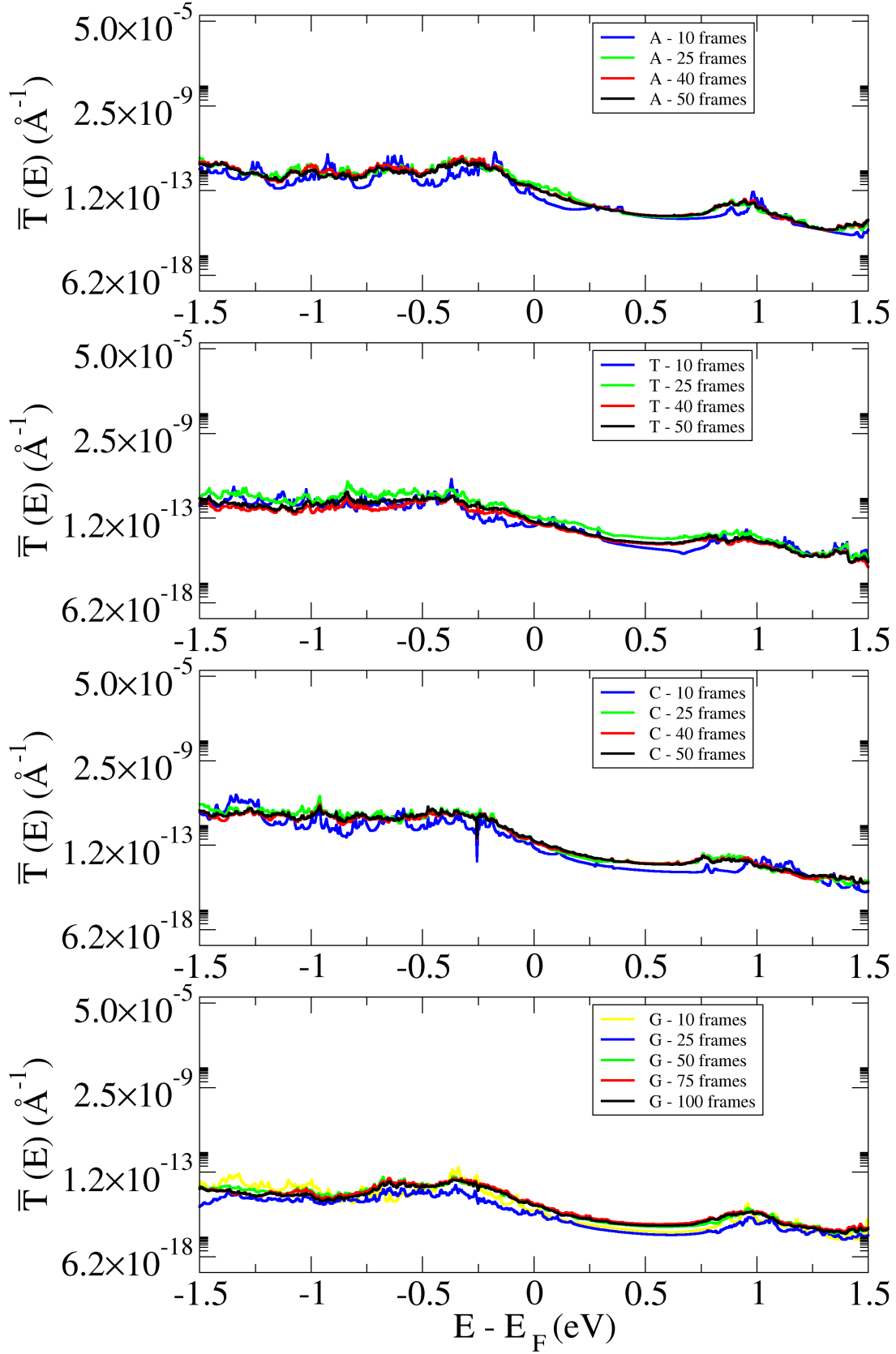


Figure 5.9: Electronic transport averages over different number of frames for all nucleotides in setup 2. The letters defined before identifies each system. For nucleotide G we used 100 frames because the convergence with 50 frames was not conclusive.

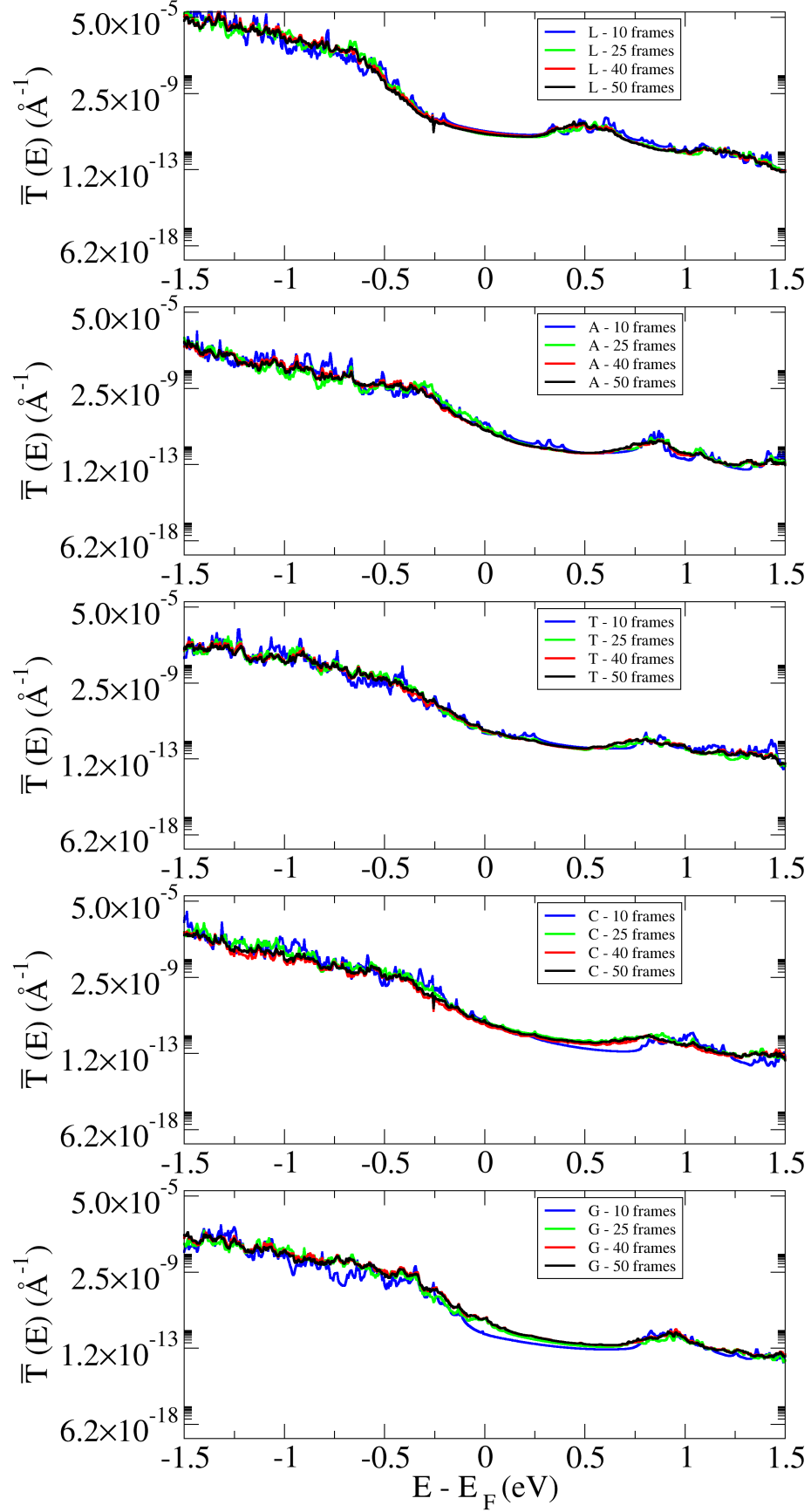


Figure 5.10: Electronic transport averages over different number of frames for all nucleotides in the setup 3. The letters defined before identifies each system. This is the setup in which we have the liquid (L) system used as reference in our calculations.

As can be observed in Figures 5.8 - 5.10, 50 snapshots are enough to converge the electronic transport for these systems. One exception is the case of Guanine (shown in Figure 5.9), where we needed an average over up to 100 snapshots. The comparison between the three setups for all systems is presented in Figure 5.11, as well as the reference transport for the empty gap in liquid (L).

The effect of the inclusion of water molecules in the QM part (setup 3) can be seen in the dashed lines of Figure 5.11. As we can observe, the transmission increases several orders of magnitude mostly as a rigid shift in the transmission for setup 2 (filled lines). Firstly, the states attributed to water are far from E_F and secondly previous works that looked at the effects of water on transport pointed towards an electrostatic effect at most [16]. Although we are increasing the size of the Hilbert space and including several new electronic states coming from water, this result is somewhat unexpected. Here we show the very important role played by water as an active part in this sort of device.

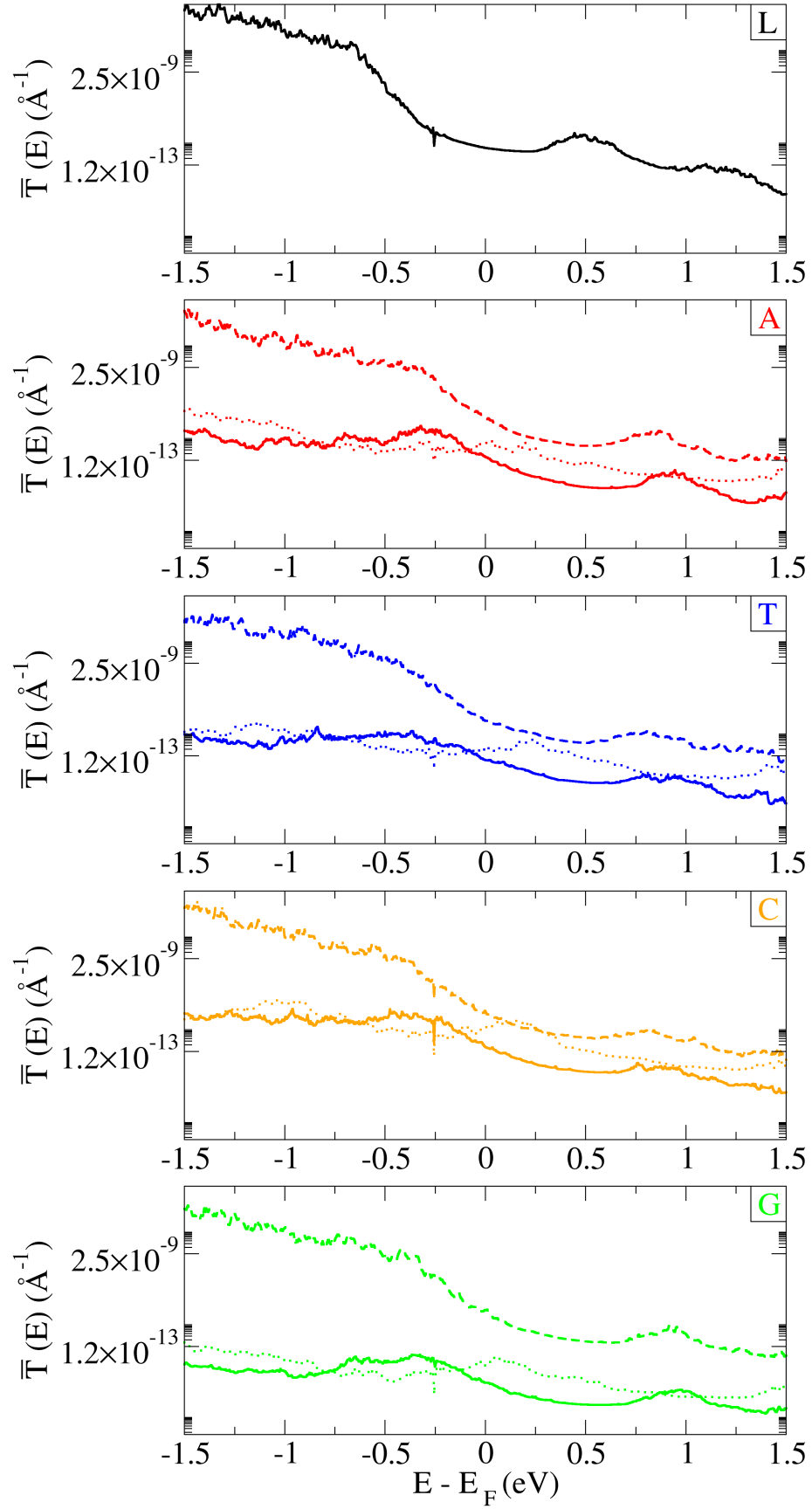


Figure 5.11: Averaged electronic transport calculations in setup 3 for L (filled black line) system and for A, T, C and G in setups 1 (dotted line), 2 (filled line) and 3 (dashed line). The letters and the color-code defined before identifies each system.

A comparison between the four nucleotides and the liquid in setup 3 is provided by Figure 5.12, where we summarize the results presented before. As Figure 5.12 shows, the total transmission for liquid (L) is the largest one for energies below -0.5 eV. The effect of including nucleotides can be seen for the other energies, in particular for $E - E_F \approx 1$ eV and -0.5 eV $< E - E_F < 0.25$ eV, where we observe an increase of total transmission.

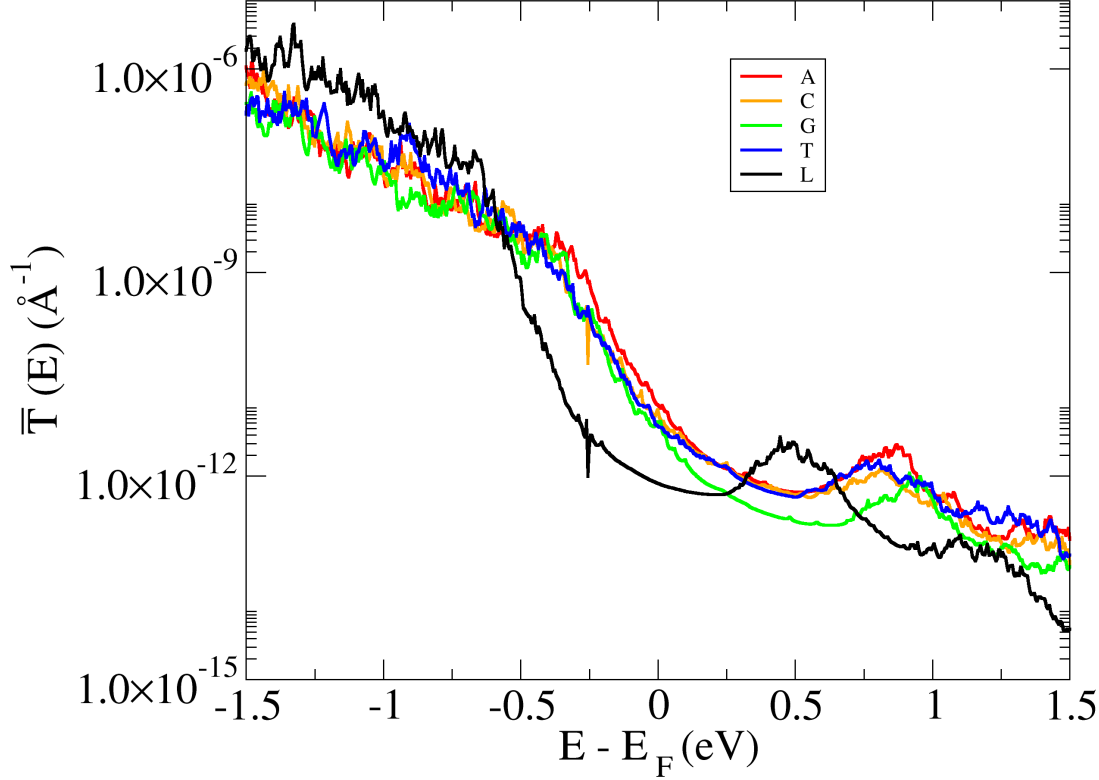


Figure 5.12: Comparison between the average electronic transport calculated in setup 3 for A, T, C, G and L.

It is important to note that the curves for both water and the nucleotides in solution are qualitatively similar indicating that many of the features in the transport are electrostatic in nature and own much to the interaction between the solvent and the edge of the gap.

Another important aspect in Figure 5.11 is the fact that the order of magnitude for setups 1 and 2 (dotted and solid lines) are similar in some energies. Setup 1 is not realistic since its structures are obtained from the simulation inside liquid and this liquid is removed (DFT step not reading the V_{ext}) to perform the next calculations. However, we can observe that for some range of energies, the inclusion of liquid via the V_{ext} calculation has the effect of decrease the transmission. This is an expected result for the other kind of systems that we present in this thesis (DNA detection, see Chapter 6), but for nanogap devices this effect is not expected.

This occurs because the inclusion of solvent effects via an external potential in these nanogap systems, or in other words, via a change in the Hamiltonian without increasing the Hilbert space, will not have the effect of increasing the transmission. As the space is

increasing no new electronic states, no new transmission channels are added. This effect can be better understood in the local current analysis, in which we can see the path of the current in the device (next section).

As we discussed before for the case of an empty gap inside liquid, electron transmission in nanogap devices only occurs if there are available electronic states. However, even in the case with the nucleotide inside the gap we can have low transmission if we don't include water molecules in the QM part. This effect can occur because in certain snapshots the coupling between the nucleotide and the nanogap edges can be low. In these cases, the inclusion of the solvent effects for the water molecules placed in between the nucleotide and the nanogap via an V_{ext} is not enough, demonstrating the need to include explicit water molecules in the QM part of the system.

5.4 Local currents

One way of determining the effects of water states on these devices is the local currents analysis. For this purpose we selected a frame with transmission similar to the average and calculated the local currents from left to right around the Fermi Level (Equation 2.88). In Figure 5.14 we can see top and side views of the four devices with the cylindrical arrows illustrating these local currents. All currents are taken into account in the calculation but for clarity to plot this result and construct Figure 5.14 we divide the system into 6 bins in the z direction (electronic transport direction), as shown in Figure 5.13:

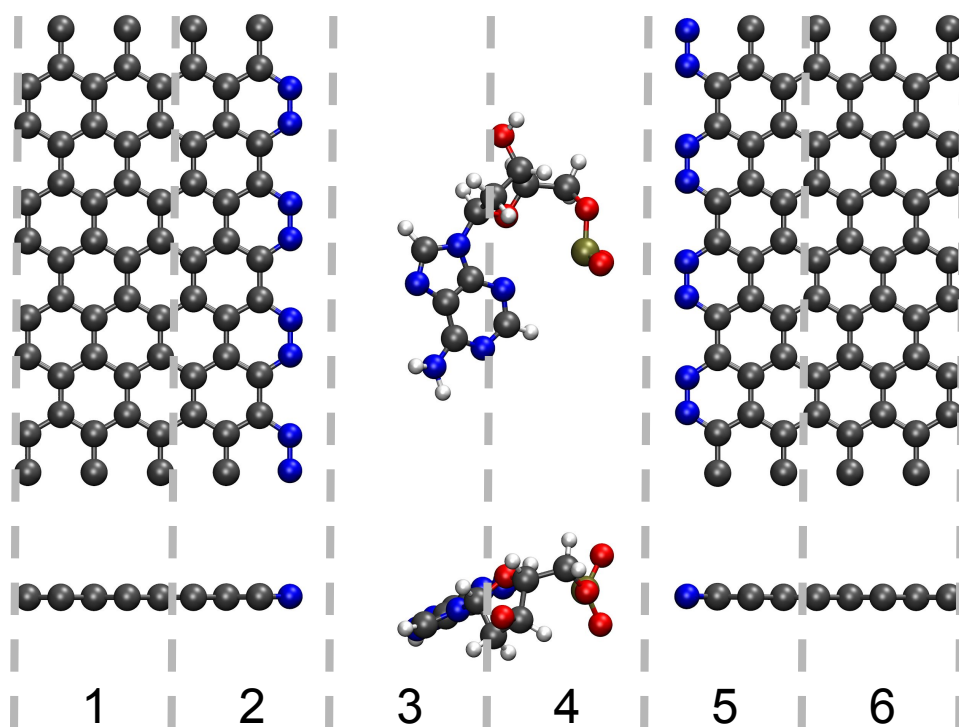


Figure 5.13: Representation of the six bins used to calculate and normalize the local currents. The system for nucleotide A in top and side views without QM water is used as example.

We then normalize the current using the largest current on each bin (Figure 5.14). Thus, although the current decay exponentially in z direction, we still can see the thick arrows in all the device extension. Therefore, the arrow's thickness in Figure 5.14 illustrates the value of the local current in that bin and are proportional to its normalized value. In this figure we only show the largest currents in the transport direction (with positive z -component).

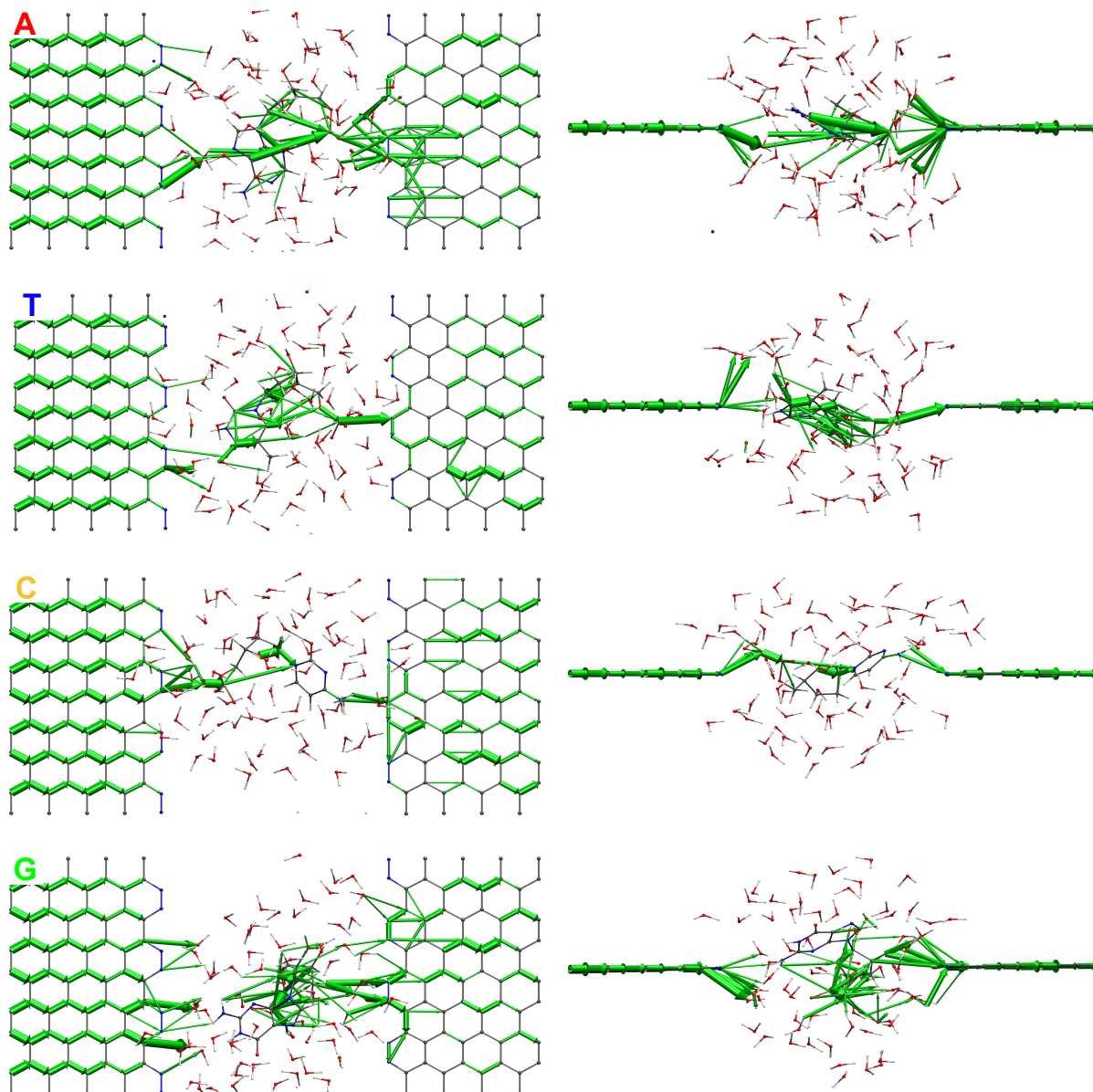


Figure 5.14: Top and side views for the left-to-right local currents on all nucleotides in QM water. The arrows' thickness are proportional to the value of the current normalized by the largest current in each bin. A cutoff has been introduced for clarity.

As we can observe in Figure 5.14, the current goes through the carbon atoms of graphene sheet and flow from the Nitrogen atoms to either the water molecules or directly to the nucleotides, in some cases.

The maximum currents in Figure 5.14, for each bin, can be identified by the ticker arrows. In the regions of graphene sheet, the maximum currents goes in the zigzag direction

following the nearest neighbor structure. Then, the current flows from the sheet to the water molecules close to the nucleotide. Inside each nucleotide we also identify maximum currents and in the right side of the device we observe a similar behavior.

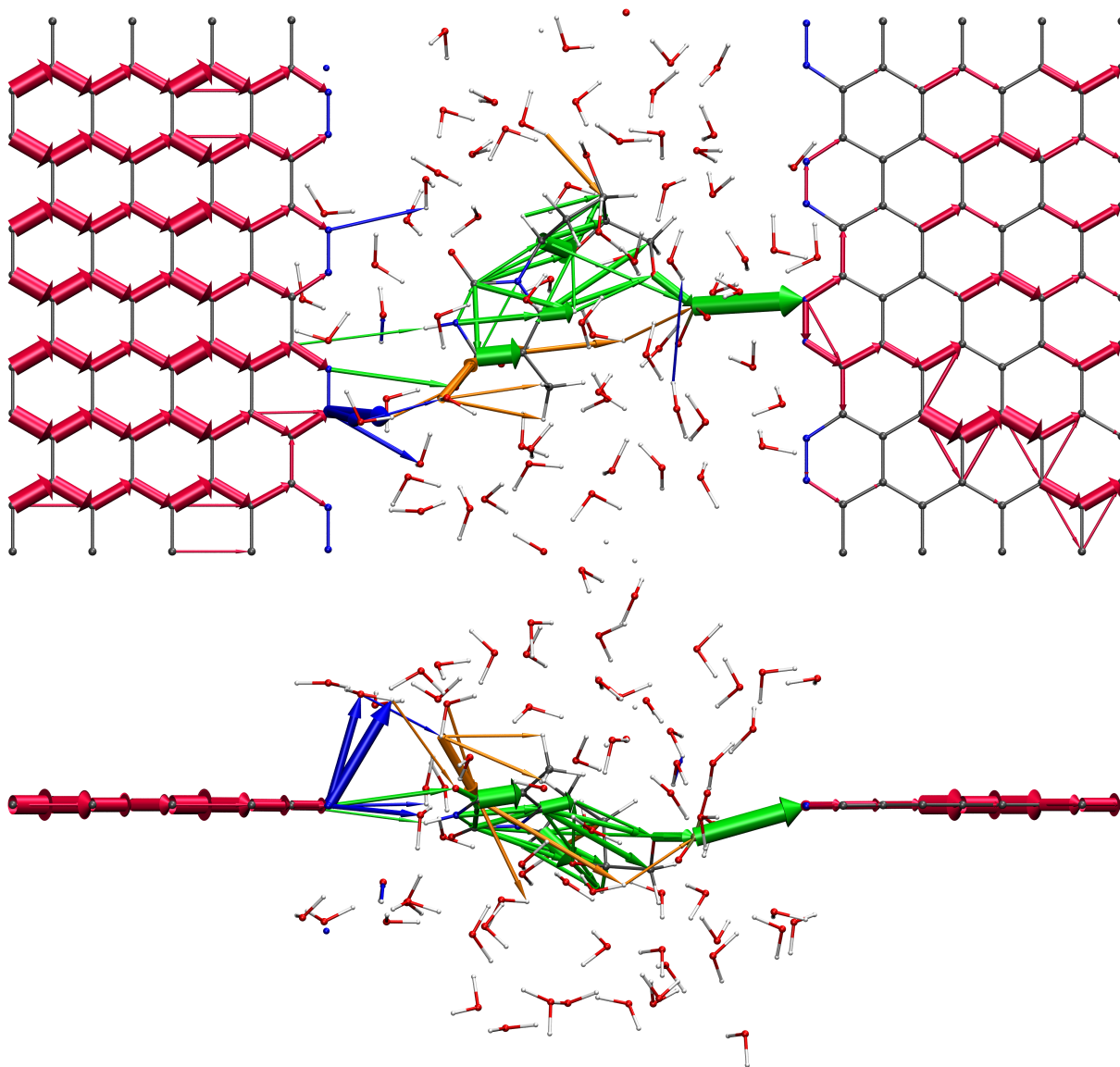


Figure 5.15: Top and side view for the currents of nucleotide T. The arrows' colors identify each subsystem: pink for local currents inside the sheet, blue for the local current from the sheet to water, yellow for the local currents from (to) the water to (from) the nucleotide and green for the local currents from (to) the sheet to (from) the nucleotide. A cutoff has been introduced for clarity.

Figure 5.15 shows different views for the same results presented in Figure 5.14 for T but using different colors to identify the currents of the different subsystems: (1) nanogap sheet, (2) water and (3) nucleotide. The pink arrows represent the current flowing between all the atoms of the sheet, blue arrows represent the current from the sheet to the water molecules (and between the water molecules), yellow ones represent the current from (to) the water molecules to (from) the nucleotide and the green arrows represent the current from (to) the sheet to (from) the nucleotide (and between the nucleotide atoms).

As can be observed, the largest currents comes from the graphene sheet (pink arrows) and are transmitted from the sheet to the water molecules (blue arrows). Then this current goes from the water molecules to the nucleotide (yellow arrows). After that the current flows inside the nucleotide (green arrows) and is transmitted in to the right electrode.

Figure 5.16 shows only the currents flowing through the water molecules and the ones from the sheet to the water molecules. In this case the normalization was done using the largest current in this subsystem for each system. In the previous cases the normalization was done using the largest current of the bin independently of the subsystem. As we can observe in the figure the current is distributed through the water molecules and the largest currents appears in the vicinity of the Nitrogen edges, showing the role played by these molecules to transmit the current through the device.

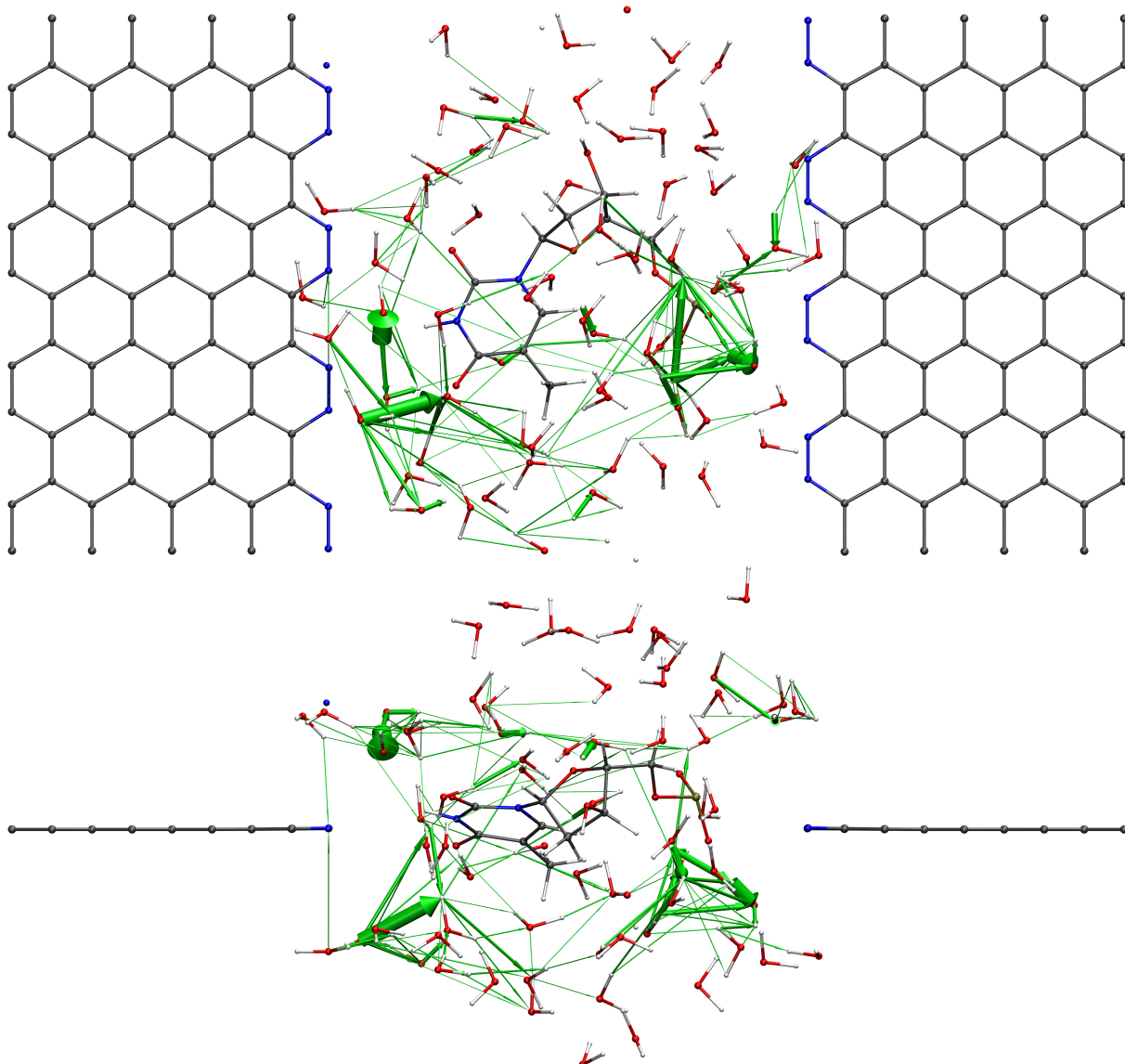


Figure 5.16: Top and side view for the currents of nucleotide T. The arrows shows the current flowing from the sheet to the water molecules and through the water molecules. In this case the currents were normalized by the largest current in the subsystem for each bin. A cutoff has been introduced for clarity.

A 2D analysis of these currents can be done to understand its exponential behavior in the device. To do that we made the graphics presented in Figure 5.17 with only z-component of the current versus the z direction of the device. Figure 5.17a shows this analysis taking into account all currents in the system and Figures 5.17b and 5.17c show this analysis for two subsystems. The subsystem 1 comprises the currents from the sheet to the nucleotide (and between the atoms of the nucleotide) and the subsystem 2 comprises the currents from the sheet to the water molecules (and between water molecules). The currents between the atoms inside the sheet can be seen for $z < 10 \text{ \AA}$ and for $z > 25 \text{ \AA}$ and are included in the three graphics.

To construct Figure 5.17 we also divided the system into 6 bins. The presented currents are the summation of all currents (for the whole system or subsystem) for that bin and the position of each point was defined as the average position inside the bin. Thus, the distance between each 2 points coincides with the size of the bin.

The general profile showed in Figure 5.17a is a decay, as we expect due to the tunneling process. The black curve (L) show that the case of empty gap immersed in liquid has a larger current in the left side (small z) of the device and that this current decrease several orders of magnitude in the right side (large z). In the same plot, the current for C, G and T are larger than the current for liquid (L) at $z \approx 20 \text{ \AA}$, which is the region where the nucleotides are found in the device. Adenine presents an unexpected behavior in the same region, since its current is smaller than the current of L.

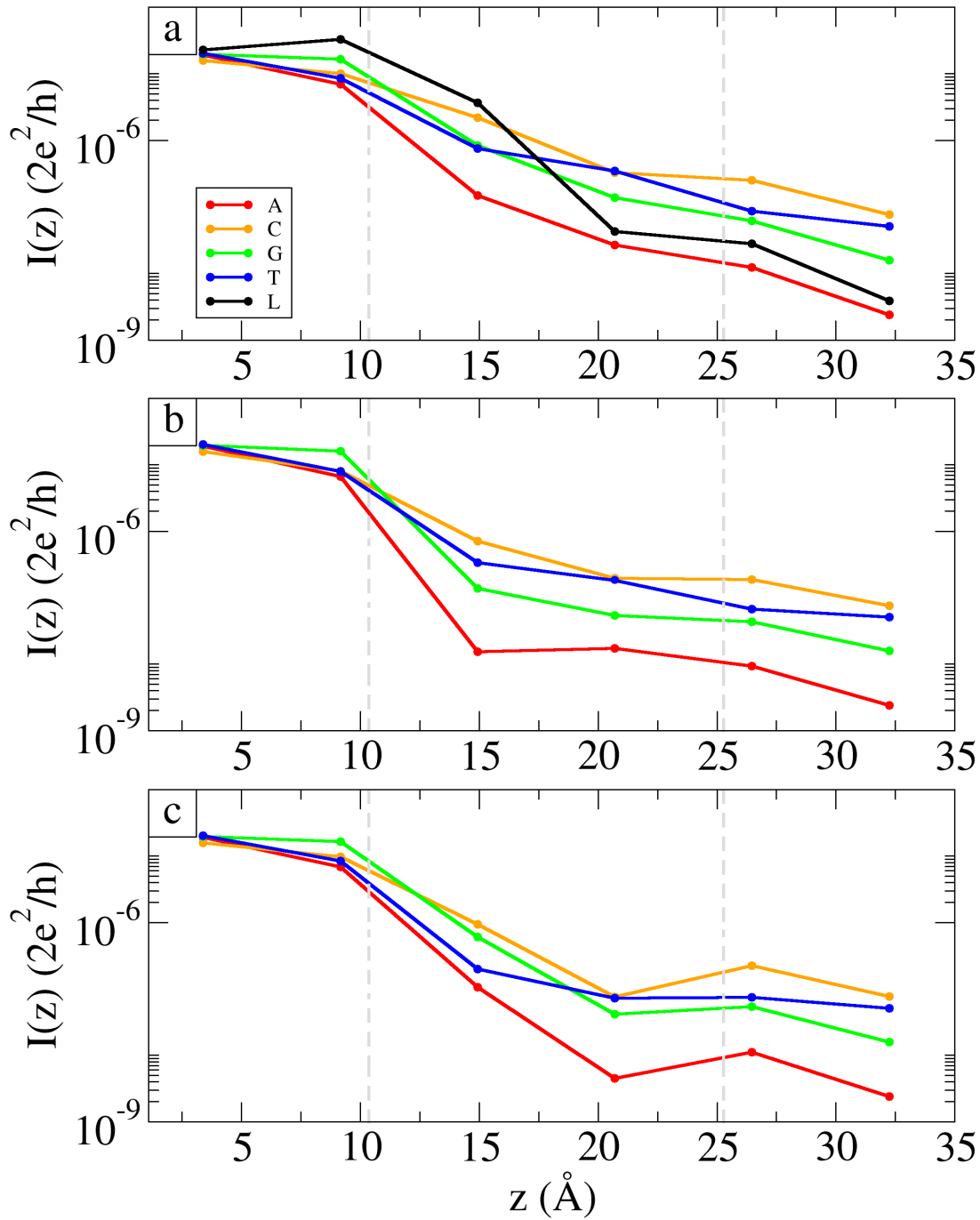


Figure 5.17: (a) z component of all currents for all nucleotides, (b) the same for subsystem 1: currents from the sheet to the nucleotide and (c) for subsystem 2: from the sheet to the water molecules. The dashed lines show the positions of the gaps.

This different behavior for Adenine in comparison to the other nucleotides can be better understood by Figure 5.18, which shows the same structures presented in Figure 5.14, but hidden the water molecules, the Na atom and the current arrows. As we can observe for this snapshot, A is placed a little farther from the nanogap edge in comparison with the other nucleotides. Due to this, its contribution to current showed in Figure 5.17b is not high like in the case of the other nucleotides. This contribution is exactly the one from the sheet to the nucleotide and since we have the nucleotide A farther (from the edges) than the others, this

snapshot will not present these contributions to the current.

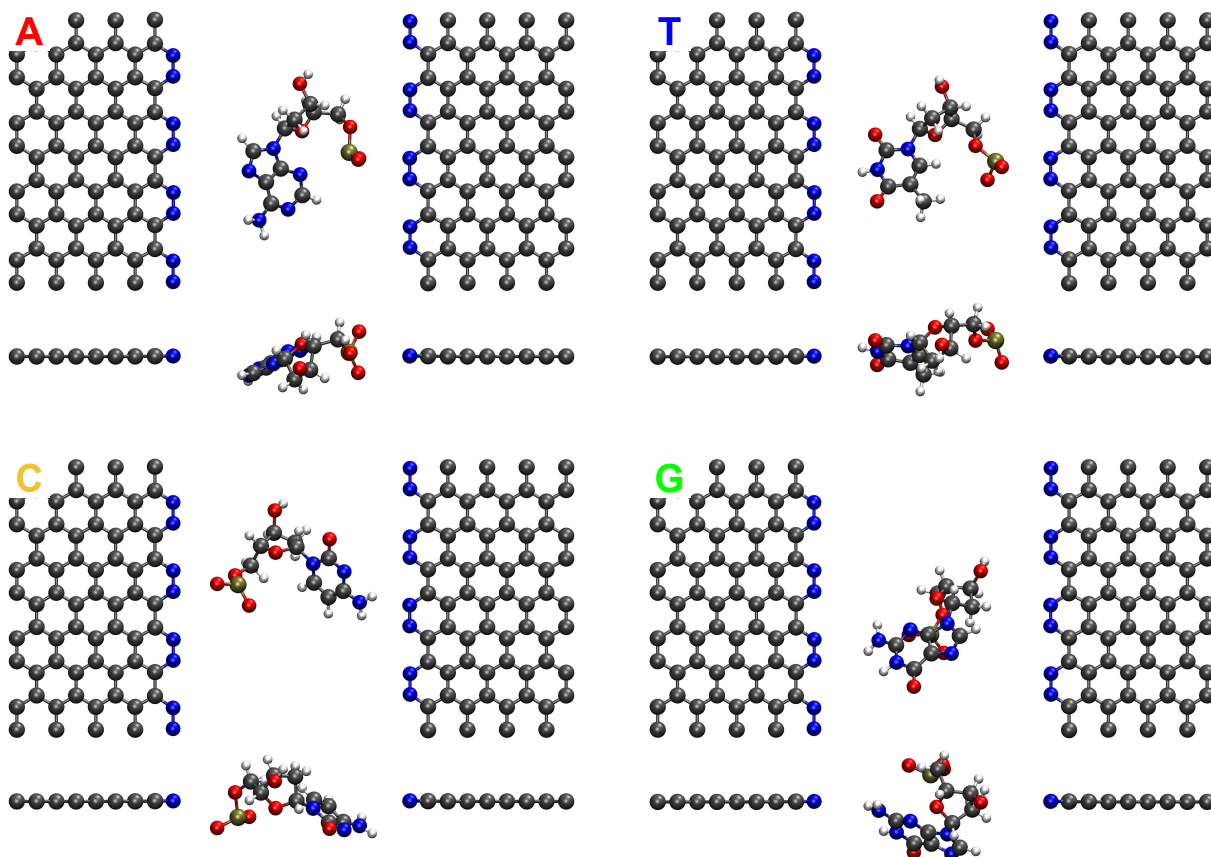


Figure 5.18: Top and side view for the snapshots used in the current calculations showed in Figure 5.14 for all nucleotides (A, T, C and G). The water molecules and the Na atom are hidden to better visualize the geometry.

Comparing Figures 5.17b and 5.17c, we can see that in the region inside the gap, the current that flows from the sheet to the nucleotides (and inside the nucleotides), showed in Figure 5.17b, is larger than the case showed in Figure 5.17c, for the currents from the sheet to the water (and between water molecules). This is explained by the positions of the nucleotides (inside the gap), because the current inside the nucleotides should be larger than the current between the water molecules in that position range.

One aspect that is worth to mention is that the general behavior for electronic transport calculations showed in Figure 5.12 should not be necessarily the same of the currents, in Figure 5.17. The current calculations are done using only one snapshot to represent an average behavior while the electronic transport calculation are done using 50 snapshots. When we choose a representative snapshot for current calculation we are analyzing the current flow for that specific geometry to understand better the magnitude and path of the current for this device.

The results presented in Figures 5.14, 5.15 and 5.17 show why include water molecules in the QM part of the system has the effect of increase the transmission, as we showed in Figure 5.11. As we could observe, the current goes through the water molecules, especially the ones

close to the nucleotide. Smaller currents pass through the other water molecules but are not shown in this picture to make it clean.

5.5 Sensitivity and selectivity of the device

Since we have the average electronic transport calculations for the four nucleotides and for the liquid case, we can use the setup 3 to estimate the sensitivity and selectivity of this device for DNA sequencing. For both sensitivity and selectivity calculations we used the data for setup 3, summarized in Figure 5.12. For sensitivity we used the results for empty gap in liquid as reference and for selectivity we used the results for Guanine, since it has the smallest transmission in the chosen energies. We calculated both sensitivity and selectivity in the same way we did in Chapter 4 for the nanogap system, using Equations 4.1 and 4.2.

The sensitivity calculations, on one hand, can estimate how sensitive to the presence of nucleotides the device is, since our reference is the nanogap inside liquid but with no nucleotide inside the nanogap. The selectivity calculations, on the other hand, give an estimation of how distinguishable the nucleotides are, compared to one specific nucleotide (Guanine). Both sensitivity and selectivity results are shown in Figure 5.19.

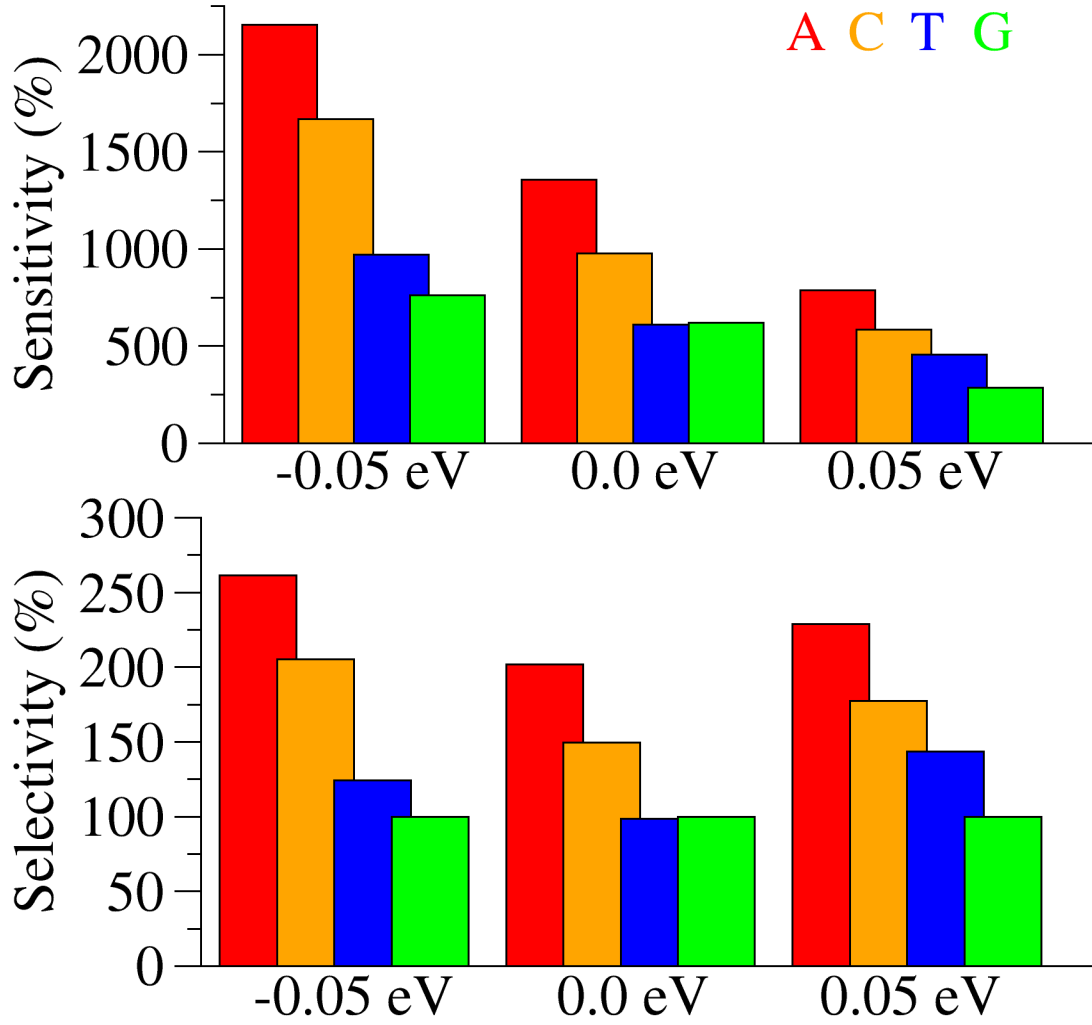


Figure 5.19: Sensitivity (upper panel) for the device using nanogap in graphene for specific energies using the liquid system as reference (Equation 4.1) and selectivity (lower panel) for the same energies using G as reference (Equation 4.2).

In the chosen energies (below , above and the Fermi energy itself) we can observe that the device is highly sensitive to the presence of nucleotides inside the nanogap, especially for $E = -0.05 \text{ eV}$, where the sensitivities are bigger than 500 % for all nucleotides. This result is expected because the electronic transports are orders of magnitude different between them, as we have discussed. For the dry system, presented in reference [17], the authors also found high sensitivities.

The selectivity for energies below and above the Fermi energy is also high. For $E = -0.05 \text{ eV}$ the signal for A is about 150 % bigger than the reference (Guanine - 100 %) and the signals for the other nucleotides are also bigger than the reference signal. For the Fermi energy the nucleotides T and G cannot be properly distinguished.

5.6 Conclusions

The results presented in this chapter showed that the proposed device is highly sensitive and selective, confirming the previous results [17] for gas phase system. We also showed that the water plays a central role in the electronic transport properties of this system, increasing the transmittance in several orders of magnitude when it is explicitly included in the QM part.

The results for local current calculations showed that the current pathway in the device is along the water molecules close to the nucleotide, also showing the role played by water in these tunneling devices. The results also showed that the z-component of the current decays as the transport direction (z direction) increase.

The results presented in this chapter were obtained in collaboration with Rodrigo Garcia Amorim^I, who performed the local currents calculations.

^IDepartamento de Física, ICEx, Universidade Federal Fluminense-UFF, Volta Redonda/RJ, Brazil

CHAPTER 6

Graphene-based DNA Chips

6.1 Introduction

In the previous chapters we presented our results for DNA sequencing devices, in which the main idea is detect the nucleotides sequence by detecting the nucleotides one by one. In that cases, we simulated the devices using isolated nucleotides. Here the proposal is DNA detection, where we address the detection of DNA strands, instead of nucleotides, using pristine graphene. In this manner, here we use the term DNA detection instead of DNA sequencing.

As we discussed in Chapter 1, the detection of specific DNA strands is used to develop devices which are able to detect genetic disorders. Here the idea is detecting DNA using all-electronic devices based on pristine graphene (before we studied devices using modified graphene: nanopores and nanogaps). In these devices, a ssDNA may be attached to graphene, which locally modifies its electronic environment, also affecting the electronic transport through the graphene sheet. The addition of a blood sample having the corresponding strand of this ssDNA which was attached to graphene, can further modify the electronic transport of the system. Thus, we would have a graphene-based biosensor capable of detecting specific sequences of DNA such as those corresponding to certain genetic disorders.

Our proposal for DNA detection using graphene is based on physical adsorption of DNA systems on a graphene sheet. To adsorb DNA on graphene we can use an anchor molecule or put the DNA directly in contact with graphene. In all these cases, the forces that keep the DNA interacting with graphene are vdW and Coulomb interactions. Therefore, there is no chemical bond between DNA and graphene.

All used setups for this end that we will present here are based on the detection of DNA hybridization or, in other words, the detection of double strand DNA formation. The hybridization event is illustrated in Figure 6.1. Since we have a target DNA strand in the device, the complementary strand (probe) will form the dsDNA. If the ssDNA in the probe do not match with the target, the hybridization event do not occur. The hybridization process itself was not studied, since we simulated DNA strands already formed.

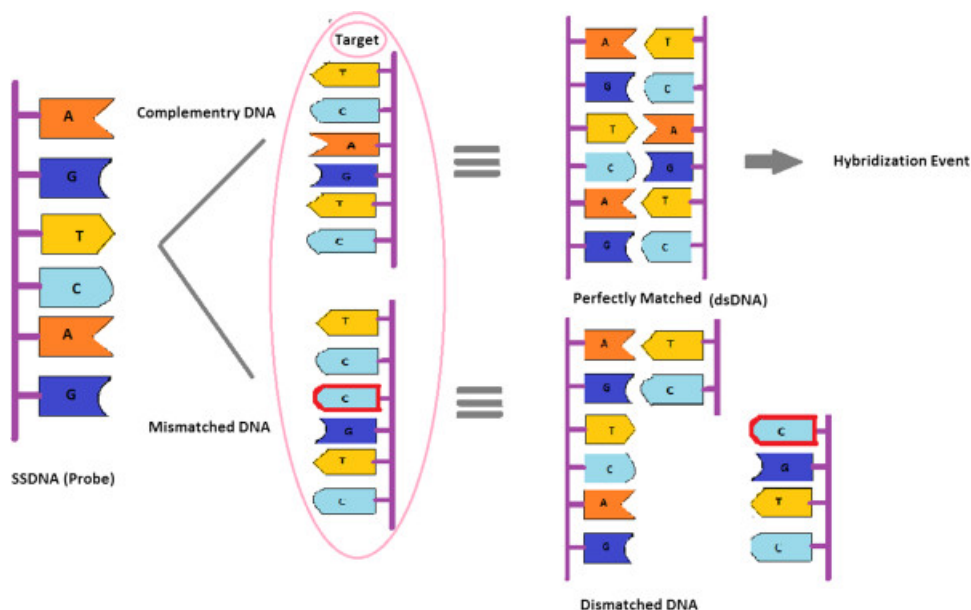


Figure 6.1: Schematic representation of the DNA hybridization process. When the complementary DNA strand is present the hybridization process occurs. Reprinted from reference [102]. Open access.

6.2 DNA detection with anchor molecule

For DNA detection we have investigated three different setups: Section 6.2: vertical DNA using anchor molecule, Section 6.3: vertical DNA not using anchor molecule and Section 6.4: horizontal DNA. In the next sections the results for each of these setups will be presented in details.

6.2.1 The systems

In our first setup for DNA detection we used of an anchor molecule to intermediate the interaction between graphene and DNA. We have chosen pyrene (Figure 6.2a) as anchor molecule due to the possibility of π - π stacking interaction with graphene. Pyrene is a molecule with structure very similar to graphene, composed by four benzene rings saturated by hydrogen. As a test, our ssDNA (Figure 6.2b) is composed by four adenines and one extra phosphate group to allow the connection (chemical bond) with pyrene. The dsDNA (Figure 6.2c) is composed by the same ssDNA and 4 more thymines (the complementary to adenine).

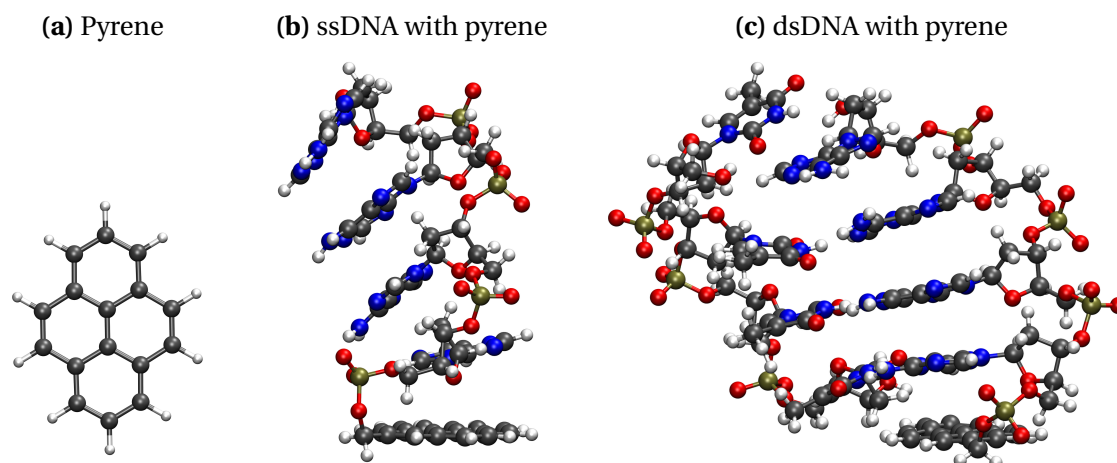


Figure 6.2: Representations of vertical DNA systems using pyrene as anchor molecule. The pyrene molecule is shown in (a), the ssDNA in (b) and the dsDNA in (c). Sequences: 4A + pyrene and 4A + pyrene + 4T.

The protocol to simulate the DNA strands showed in Figure 6.2 is the same for all systems used in DNA sequencing. This protocol, as well as the details for each system were presented in the Chapter 2. To prevent structures containing DNA on top of the electrodes (we require the molecule to remain in the scattering region) we applied a harmonic potential to confine the movement of the DNA to the center of the graphene sheet. The restraints were also applied to prevent more π - π stacking interactions that can appear due to rotations in the DNA strands, since we want to sample structures with this kind of interaction only between the anchor and graphene. This potential was applied in the pyrene and in the first nucleotide, in the ssDNA only. The Figure 6.3 shows the restrained atoms in yellow as well as the axes on which the potential was applied, in the same color. In the direction of the blue axis there was no restraint. The force constant applied was $10 \text{ kJmol}^{-1} \text{ nm}^{-2}$.

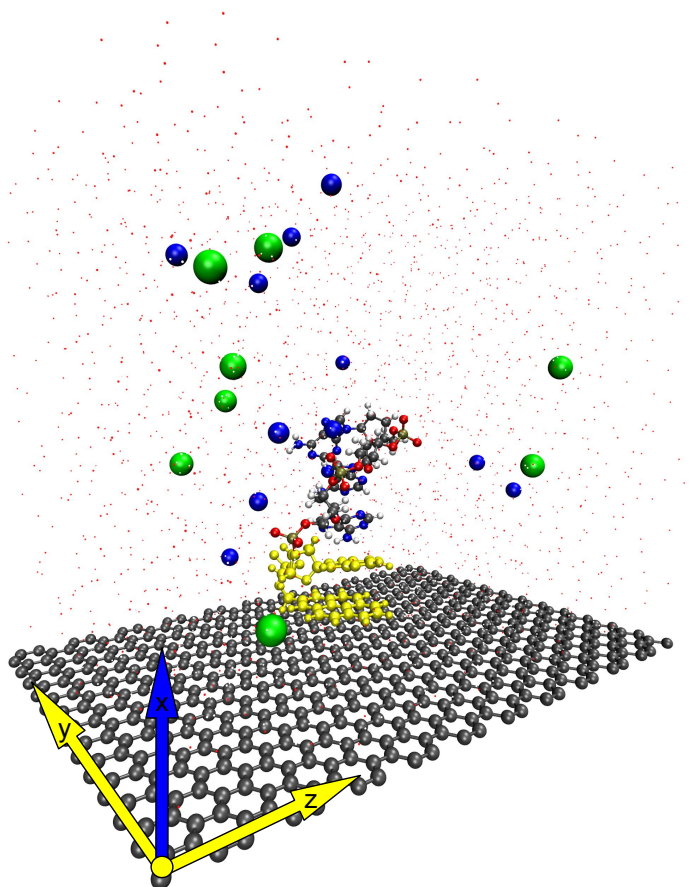


Figure 6.3: Snapshot for a ssDNA on graphene highlighting the restraint applied for ssDNA with anchor molecule. Atoms in yellow are restrained in the direction of the yellow axis. There is no restraint in the direction of the blue axes (x direction).

Since the simulation box is defined we can perform a classical MD calculation for each system (ssDNA and dsDNA). A top and side view of the box for this system is shown in Figure 6.4. After the MD we divide the system into a QM and a MM part. Figure 6.4 shows the simulation with different representations for the atoms to differentiate these two partitions. The blue and green spheres represent the Na and Cl ions, respectively. All atoms presented in *ball-and-sticks* representation and the large radius blue spheres make up the QM part of the system. All atoms in *line* representation, the green spheres and the small radius blue spheres constitute the MM part of the system. This representation pattern to distinguish the QM and MM partition will be used in all systems throughout this chapter.

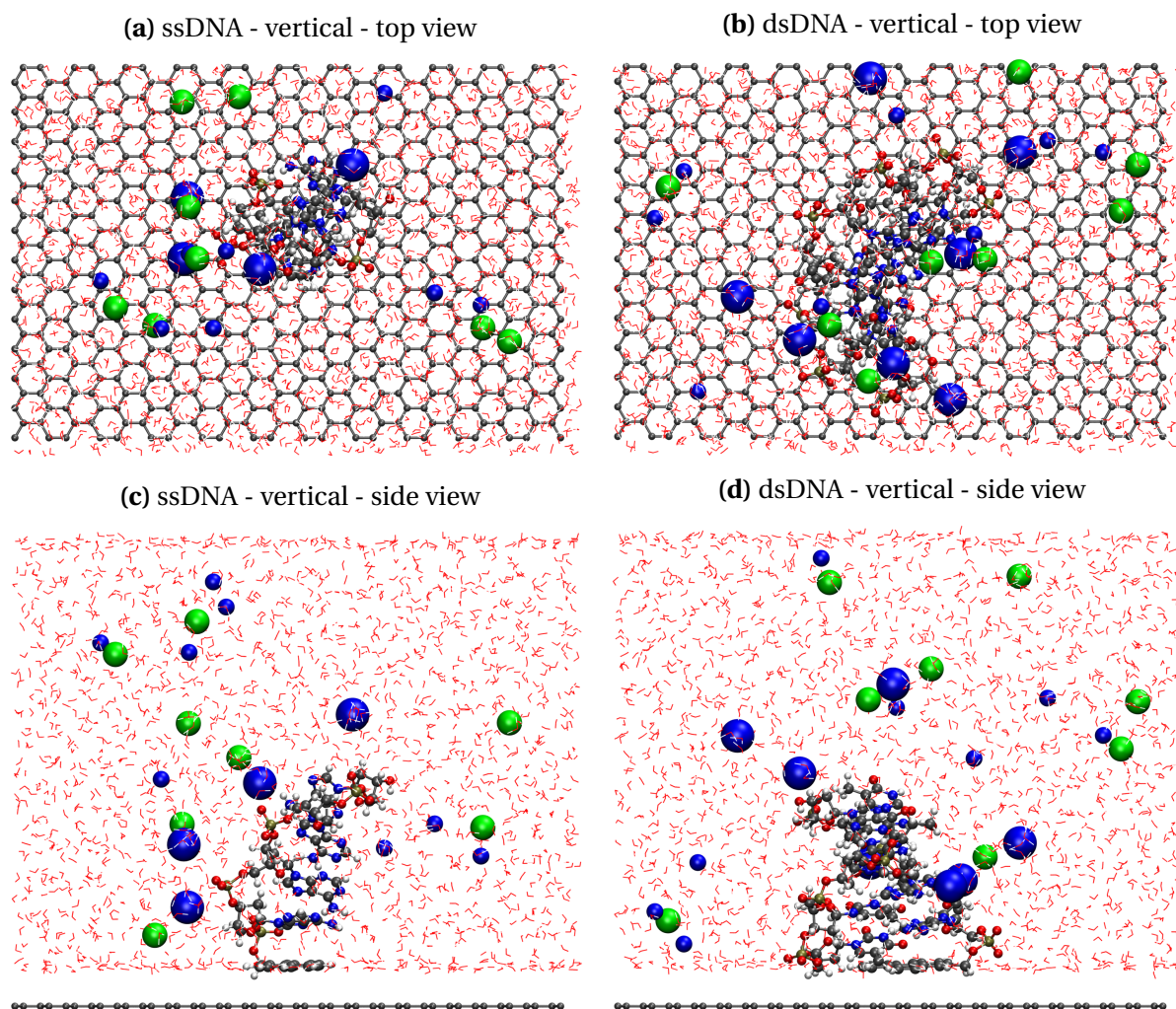


Figure 6.4: Top and side views for a representative snapshot of vertical ssDNA (a and c) and dsDNA (b and d) without QM water using pyrene as anchor molecule. 4 Na atoms in the QM part for ssDNA and 7 Na atoms in the QM part for dsDNA.

A schematic representation of the QM/MM partition idea is shown in Figure 6.6. Therefore, our QM partition is composed by a graphene sheet, all DNA atoms and the number of Na ions (positive ions) needed to equilibrate the negative charge of the phosphate groups. The MM partition is composed by the liquid (water, Cl and remaining Na). We select the nearest Na atoms to include in the QM part¹. For this system, we also studied another QM/MM partition: we included one layer of water in the QM part. Thus, this new QM/MM partition is shown in Figure 6.5, in which the QM water is also shown in *ball-and-sticks* and the MM water is shown in *line* representation.

¹This criterion is the same in all results that we will present throughout this chapter.

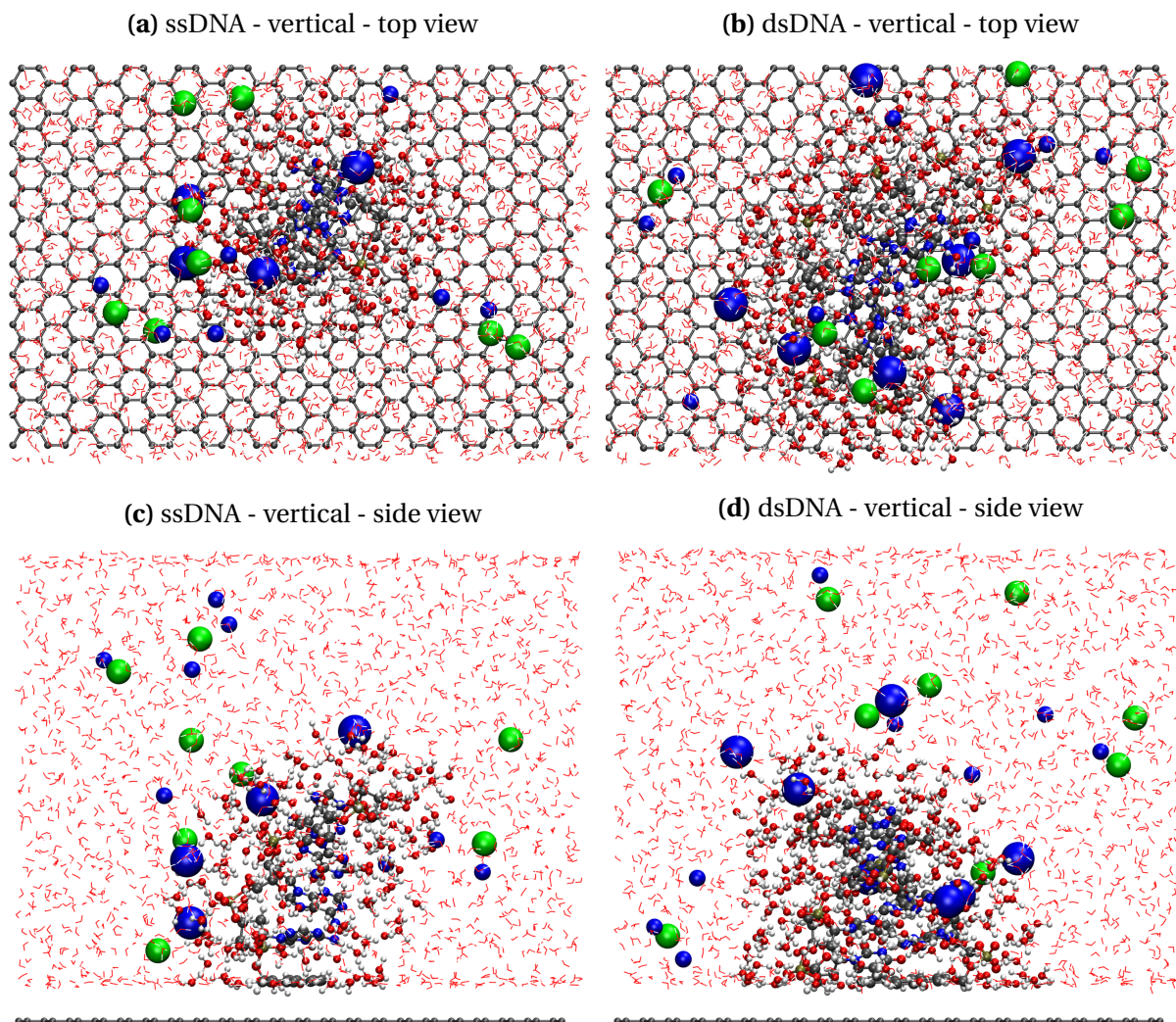


Figure 6.5: Top and side view for a representative snapshot of vertical ssDNA (a and c) and dsDNA (b and d) with QM water using pyrene as anchor molecule.

Thus, we included one layer of QM water around the DNA strands to investigate the effect in the electronic transport calculations. These water molecules are supposed to have configurations affected by the presence of DNA and since we include these configurations in the QM part, the electronic transport should be altered, since now we have more scattering states. We choose the water molecules up to 5 Å farther from the DNA. This distance criterion was defined and explained in Chapter 3, Section 3.4.1.

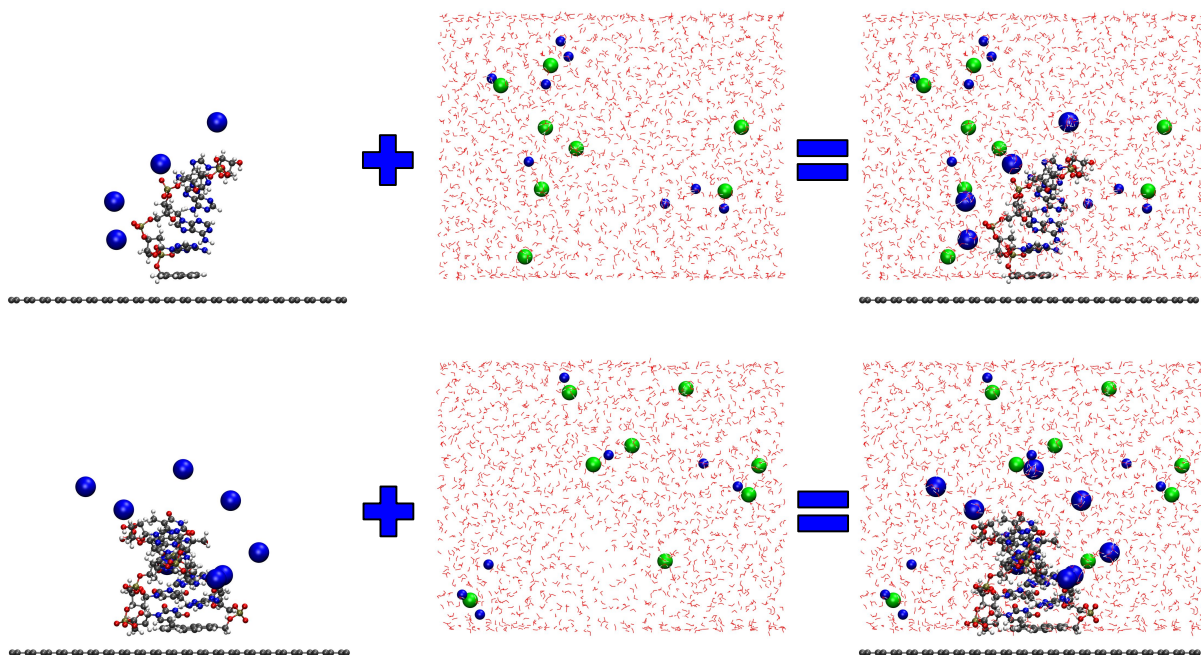


Figure 6.6: Side view for vertical ssDNA (upper panel) and dsDNA (lower panel) without QM water using pyrene as anchor molecule to represent the QM/MM partition idea.

The simulation details related with the systems^{II} presented in this section are shown in Tables 6.1 and 6.2.

^{II}The details for the systems used as background were already showed in Chapter 3.

MD calculations			
NPT	Parrinello-Rahman barostat	10 <i>ns</i>	
NVT	Nose-Hoover thermostat	100 <i>ns</i>	
Time step	2 <i>fs</i>		
Force Field	AMBER99SB and SPC for water		
Snapshots and QM/MM partition			
50 ^{III} snapshots	Equally spaced by 2 <i>ns</i>		
QM part I	Graphene sheet + DNA + <i>n</i> ^{IV} Na ions		
QM part II ^V	QM part I + first layer of water		
MM part	Cl ions and remaining water and Na ions		
DFT calculations			
Functional	VdW DRSL		
Basis set	DZ for graphene and DZP for the remaining atoms		
k-points grid (x y z)	1 2 2		
Electronic transport calculations			
Number of energy points	1500		
Number of k-points	16		
Energy range	From −4 <i>eV</i> to 3 <i>eV</i> around Fermi Energy		

Table 6.1: Simulation details for the study of DNA detection.

^{III}For QM/MM partition test we used 10 frames instead.^{IV}*n* is the number of Na ions needed to equilibrate the charge of the system.^VWe tested more QM/MM partitions for ssDNA using pyrene as anchor molecule.

ssDNA	
Box dimensions (x, y, z)	(42.042, 31.980, 46.860)
N_{ions}	12 Na and 8 Cl
$N_{water\ molecules}$	1896
N_{DNA}	158
$N_{graphene}$	572 C
dsDNA	
Box dimensions (x, y, z)	(42.400, 31.980, 46.860)
N_{ions}	15 Na and 8 Cl
$N_{water\ molecules}$	1843
N_{DNA}	285
$N_{graphene}$	572 C

Table 6.2: Details of the simulation box for the study of vertical DNA detection using anchor molecule. N is the number of the specified species. Length unit: Å.

The setup for electronic transport calculation in the systems presented in this chapter is the same as the one presented in Figure 3.2 for the study of graphene in solution.

6.2.2 Electronic transport results

For all the systems described above we extracted 50 structures from the MD calculation equally spaced by 2 ns. These structures are divided into a QM and a MM part, as shown before. Then, we obtain external classical potential (V_{ext}) for the MM part and introduce that effect into our DFT calculation of the QM partition. Finally an electronic transport calculation over the generated Hamiltonian is performed.

Figure 6.7 shows the electronic transport calculations for all frames as well as the average over all curves. The results for ssDNA and dsDNA with and without QM water are also shown in this figure (Figure 6.7). As we can see in Figure 6.7 the total transmission ($T(E)$) varies, as it would be expected, around the average curve for the systems with and without QM water. In order to address these structural fluctuations, in Figure 6.8 we show the average electronic transport properties for all systems with one standard deviation (SD) added and subtracted (gray shaded region).

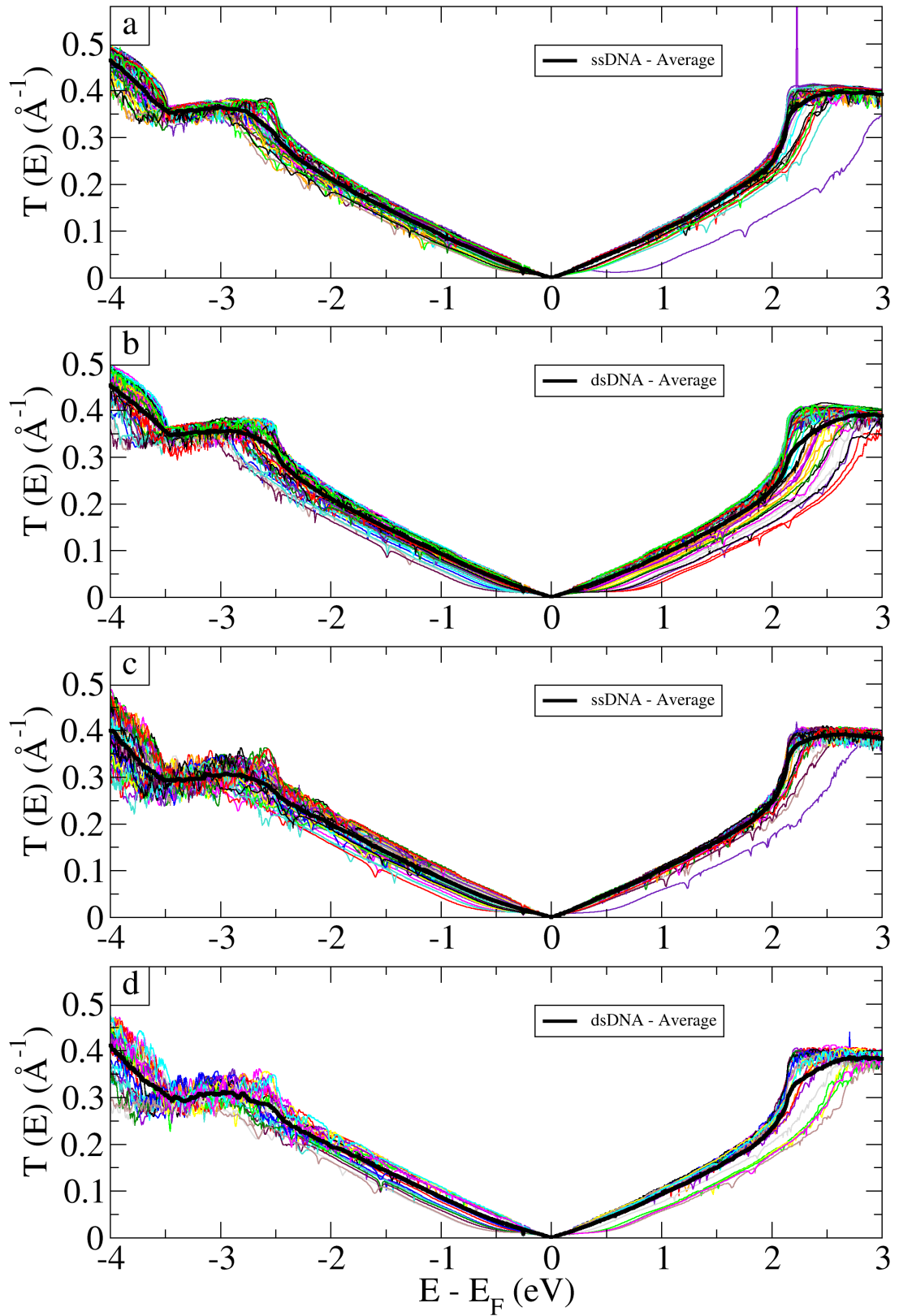


Figure 6.7: Electronic transport calculations and the averages over the curves for vertical DNA systems using anchor molecule: (a) ssDNA and (b) dsDNA without QM water and (c) ssDNA and (d) dsDNA with QM water.

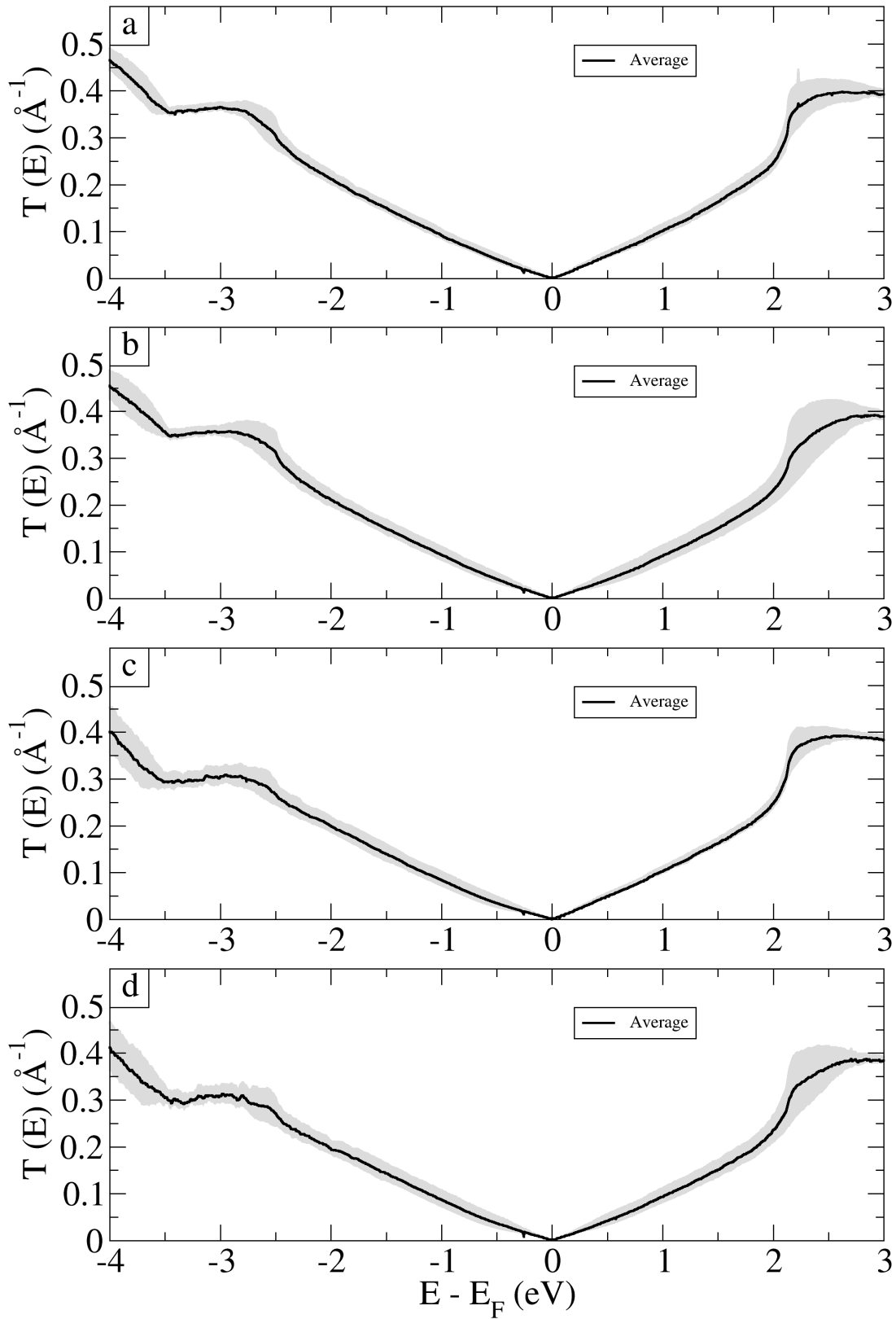


Figure 6.8: Electronic transport calculations for vertical DNA systems using anchor molecule: (a) ssDNA and (b) dsDNA without QM water and (c) ssDNA and (d) dsDNA with QM water. The gray shaded region comprises the addition and subtraction of 1 SD.

Both Figures 6.7 and 6.8 shows that the SD from the average for the case with QM water

is larger than the deviation for the case without QM water, but not significantly so. A larger fluctuation is expected since now we have more scattering channels. Nonetheless our results indicate that QM water does not play a significant role in transport.

For the systems presented here we have used 50 frames, but the number of frames needed to converge the average of the electronic transport can vary. As was done in previous chapters, an analysis that allows us to identify the number of frames that are needed to converge the electronic transport is the evolution of the average and SD as a function of the number of frames. To do so we can choose a certain energy and plot the average electronic transport for that number of frames. As we can see, the electronic transport for different structures seem to be rigid displacements from the average electronic transport^{VI}. Therefore, the chosen energy E' to analyze the convergence is somewhat arbitrary, and will lead to similar results. Thus in all cases we choose $E \simeq 0.0$ eV, close to the Fermi Level.

^{VI}This is in line with our results from Chapter 3 where the effects of the solvent is purely electrostatic.

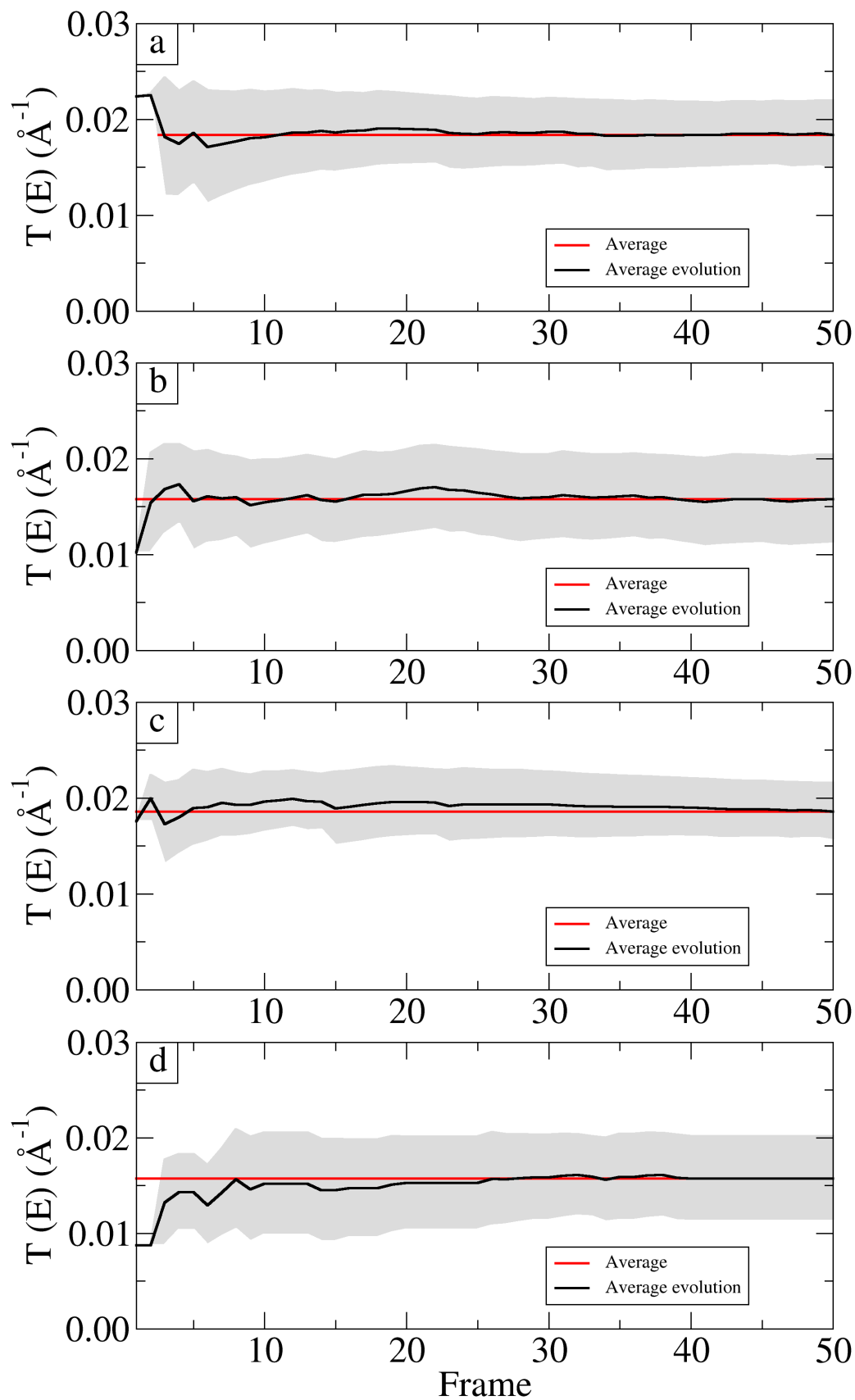


Figure 6.9: Evolution of the average and SD for all vertical DNA systems using anchor molecule: (a) ssDNA and (b) dsDNA without QM water and (c) ssDNA and (d) dsDNA with QM water. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \simeq 0.0 \text{ eV}$.

In this way, from Figure 6.9, we can see that 50 frames are enough to converge the electronic transport for these systems. In the same figure we can see that the SD evolution is also converged.

Figure 6.10 provides a comparison between the averages of pristine graphene (dry graphene), graphene in solution (same solution conditions of the DNA systems, presented in Chapter 3), ssDNA and dsDNA for all systems. The red curve in Figure 6.10 shows the electronic transport for pristine graphene. This is our reference and is expected to be the largest transmittance because the effect of immerse the system in solution (black curve) or adsorb molecules on it (blue and green curves) is always decrease the transmittance, for this kind of system. When we have pores or nanogaps this is not necessarily true.

Compared to pristine graphene (red curve) both ssDNA and dsDNA have smaller transmittance. However, from -2 eV to 2 eV the transmittance for ssDNA and dsDNA are very close to the transmittance for the system in liquid (black curve). This means that, in average, the transmittances for the systems with DNA and for the systems with only liquid are basically the same. For the purpose of a device this is obviously a problem, since it would not be possible to differentiate between the DNA and the background signal from the solvent alone.

The comparison between the blue and green curves in Figures 6.10a and 6.10b show that the effect of including water in the QM part of the system is to decrease the total transmission, but this effect can be better seen for energies below -2 eV , far from the Fermi energy (0.0 eV). The main analysis for a device with the purpose of detecting DNA is the difference between the two DNA strand signals.

As shown in Figure 6.10, for both systems with and without QM water, the signals for ssDNA and dsDNA are virtually the same. This means that the sensibility of a device with this setup is not ideal. In this setup for DNA detection the charged groups of the nucleotides are far from graphene, what can be the reason why the signals for ssDNA and dsDNA are similar. Furthermore the choice of anchor molecule can also affect the signals. One possibility is use the DNA nucleotides itself as anchor molecules since we can also have π - π stacking interactions in this case.

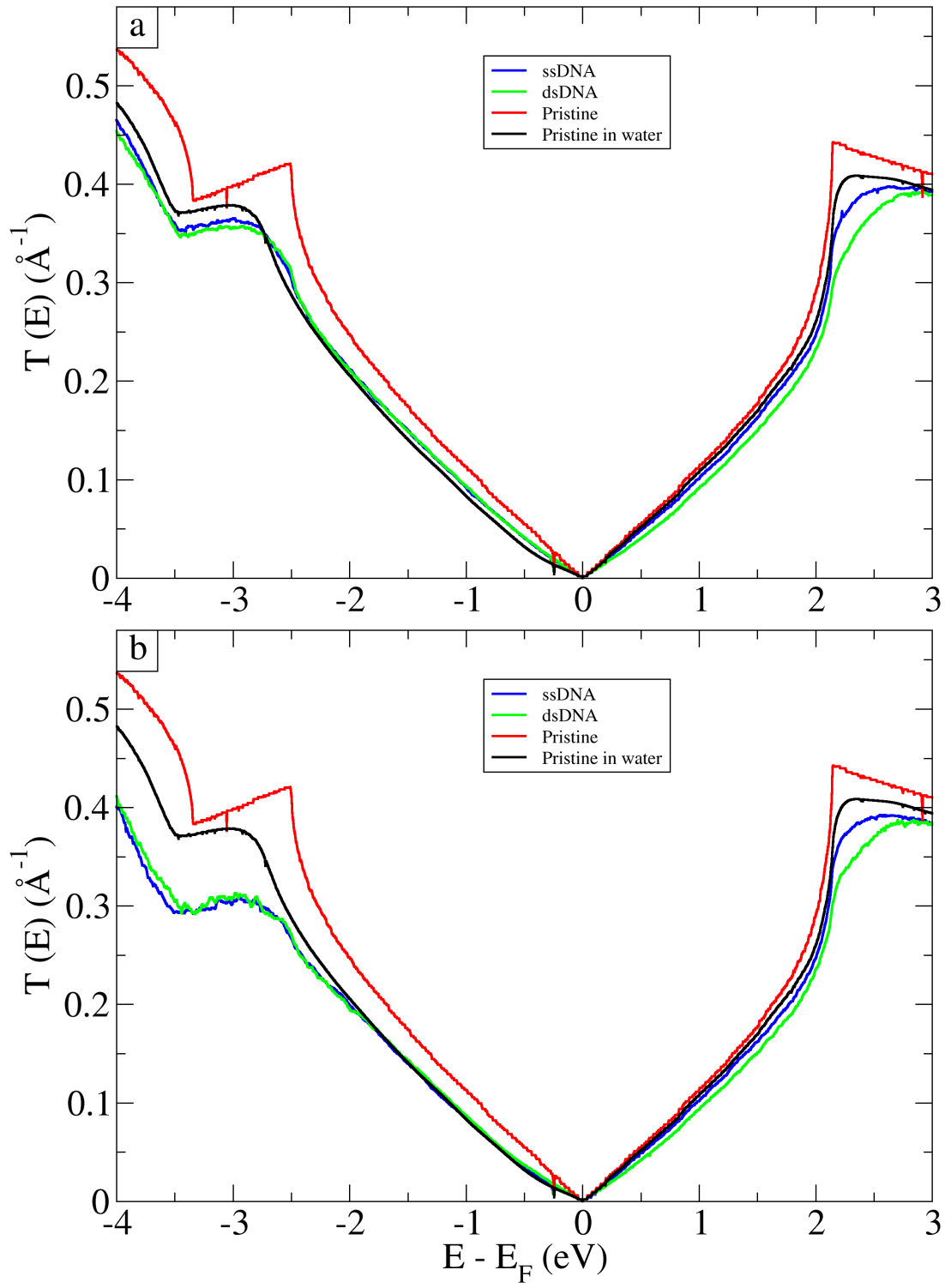


Figure 6.10: Comparison between the averages over 50 electronic transport calculations for vertical ssDNA and dsDNA with (a) and without (b) QM water.

6.3 DNA detection without anchor molecule - vertical case

6.3.1 The systems

As mentioned at the end of the previous section a possible way to remedy to be signal difference is to bring the DNA closer to graphene by directly adsorbing it on the graphene sheet. Thus, the second setup for DNA detection is based on the use of the nucleotide itself as anchor molecule instead of pyrene^{VII}. In this case the DNA strands are also adsorbed vertically on the graphene sheet. The ssDNA (Figure 6.11a) is composed by 6 guanines and the dsDNA (Figure 6.11b) is composed by the same ssDNA and 4 complementary cytosines.

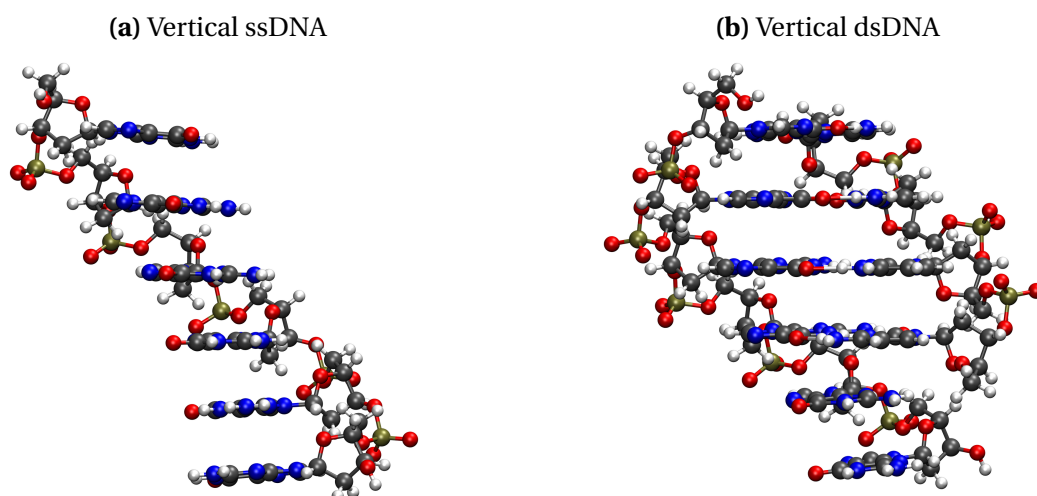


Figure 6.11: Representations of vertical ssDNA (a) and dsDNA (b) systems not using pyrene as anchor molecule. Sequences: 6G and 6G + 4C.

The protocol to simulate the DNA strands showed in Figure 6.11 is the same one used before (Section 6.2), but here the restrains to prevent structures with the DNA systems above the electrodes are applied in the first three nucleotides of the ssDNA only. Figure 6.12 shows those restrains. The force constant and the directions of restrains are also the same of the previous case.

^{VII} Here after we will use the term "with anchor molecule" to refer the case with pyrene

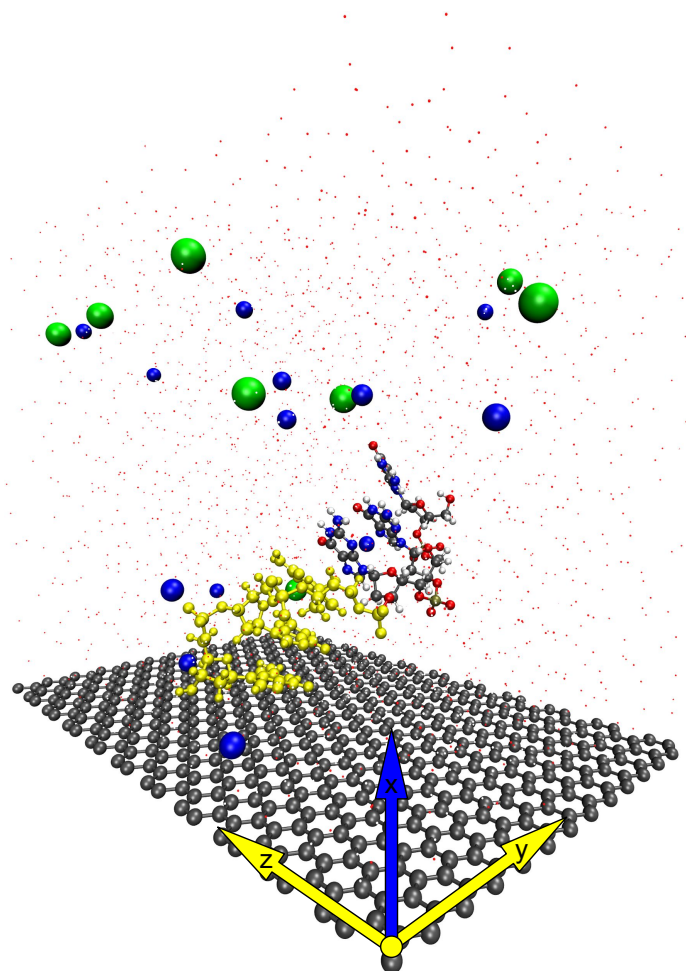


Figure 6.12: Restraint applied for ssDNA with nucleotide as anchor molecule. Atoms in yellow are restrained in the direction of the yellow axes. There is no restraint in the direction of the blue axis (x direction).

The Figure 6.13 shows top (a and b) and side (c and d) views of representative snapshots for both ssDNA and dsDNA used in this setup. In this case, the number of Na ions in the QM part for both cases is different of the previous case as the number of phosphate groups is larger. For this system we only studied the case without QM water. The representation pattern is the same as before.

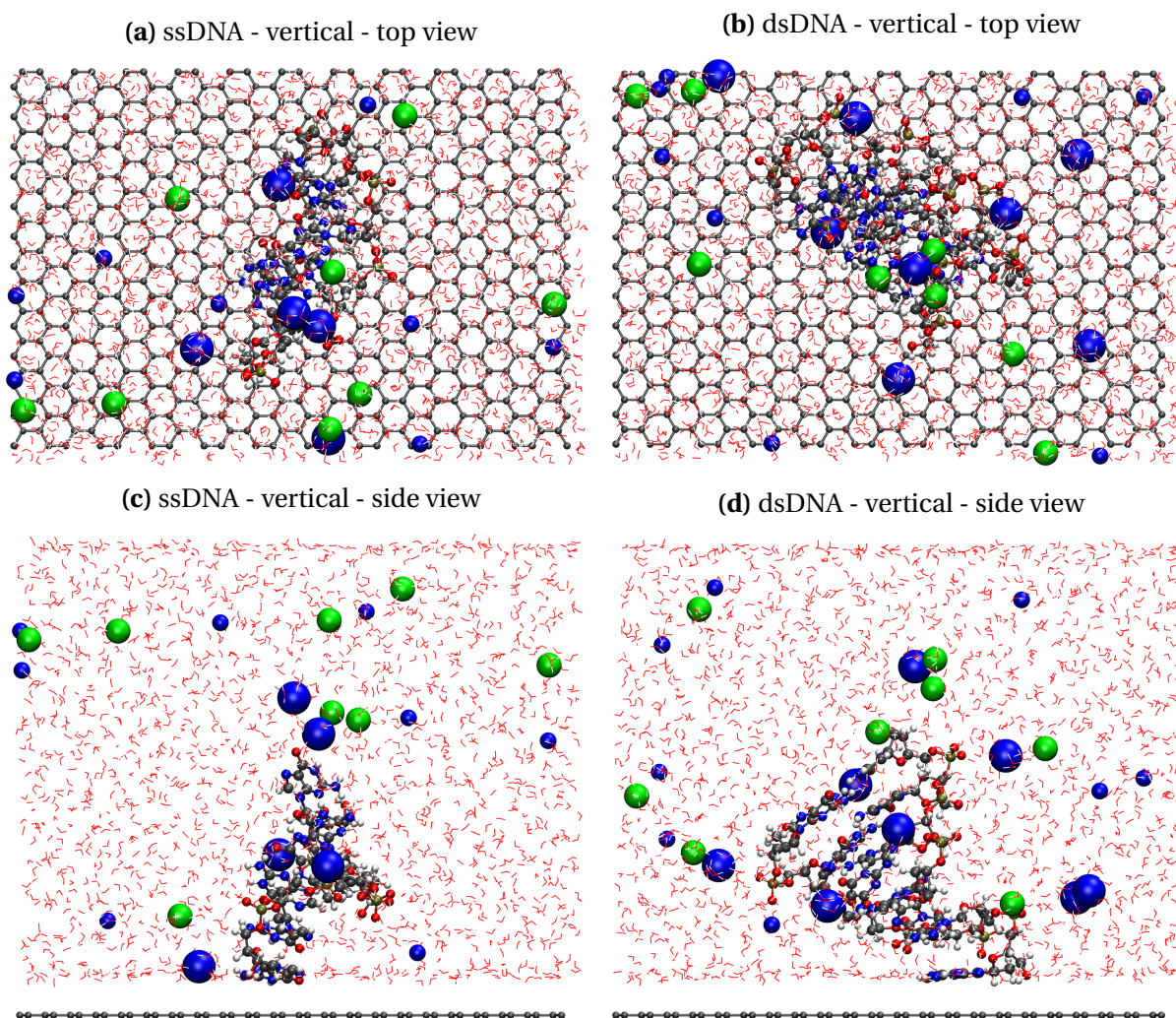


Figure 6.13: Top and side views for a representative snapshot of vertical ssDNA (a and c) and dsDNA (b and d) without QM water and not using anchor molecule. 5 Na atoms in the QM part for ssDNA and 8 Na atoms in the QM part for dsDNA.

The simulation details related with the systems presented in this section are shown in Table 6.3.

ssDNA	
Box dimensions (x, y, z)	(42.602, 31.980, 46.860)
N_{ions}	13 Na and 8 Cl
$N_{water\ molecules}$	1883
N_{DNA}	197
$N_{graphene}$	572 C
dsDNA	
Box dimensions (x, y, z)	(42.178, 31.980, 46.860)
N_{ions}	16 Na and 8 Cl
$N_{water\ molecules}$	1828
N_{DNA}	316
$N_{graphene}$	572 C

Table 6.3: Details of the simulation box for the study of vertical DNA detection not using anchor molecule. N is the number of the specified species. Length unit: Å.

6.3.2 Electronic transport results

As done in previous cases we first look at the convergence of our results. The Figure 6.14 shows all electronic transport curves for vertical ssDNA and dsDNA without an anchor molecule as well as the average over these curves (black curve). These averages with one SD added and subtracted can be seen in Figure 6.15. The Figure 6.16 presents the evolution of the average and SD for these systems, showing that the chosen number of frames is enough to converge the electronic transport.

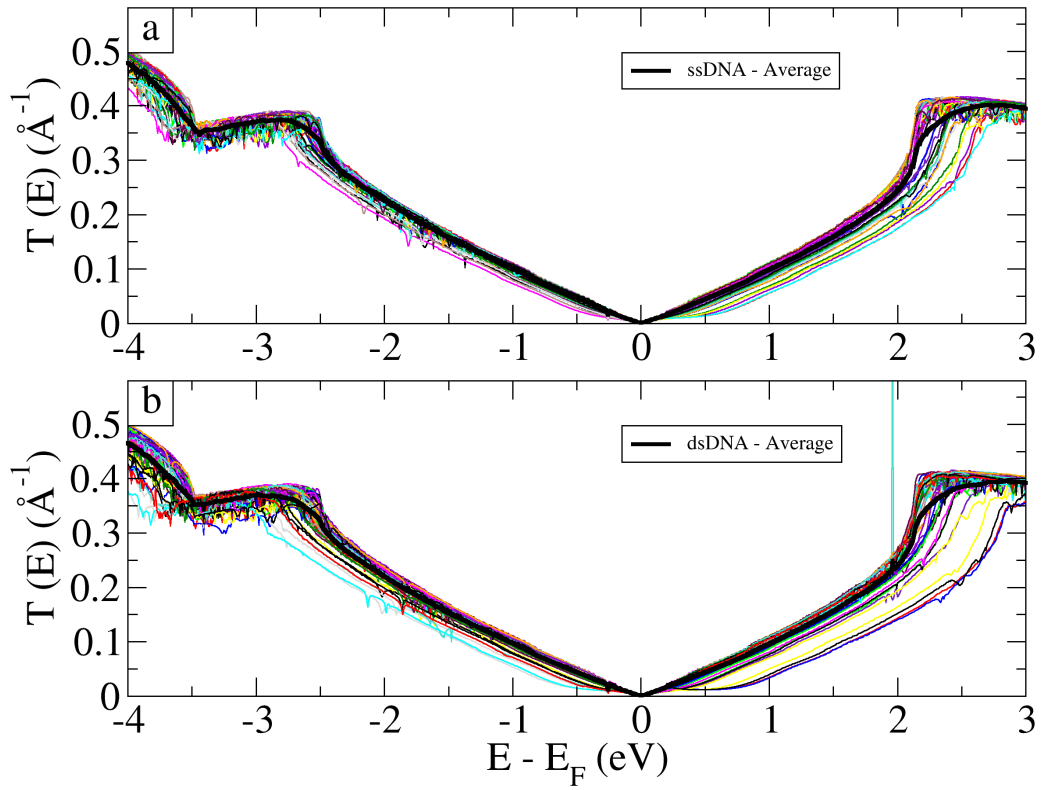


Figure 6.14: Electronic transport calculations for all frames of vertical ssDNA (a) and dsDNA (b) systems without anchor molecule and the average over these curves.

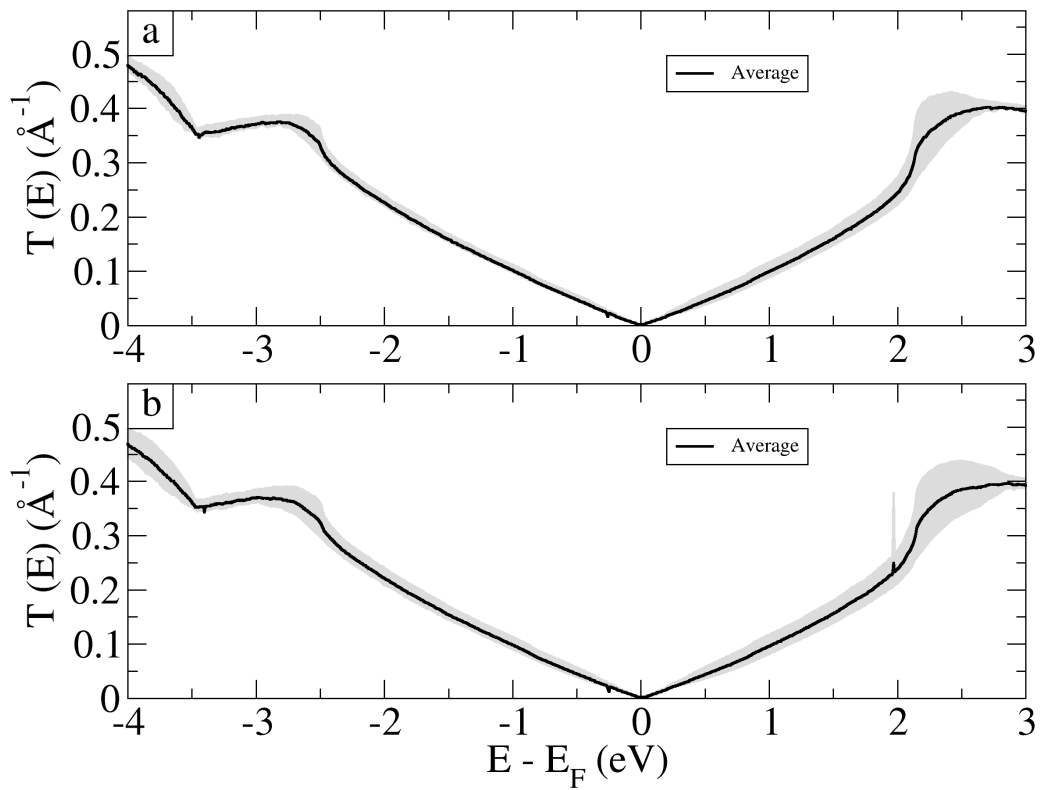


Figure 6.15: Electronic transport calculations for vertical ssDNA (a) and dsDNA (b) systems without anchor molecule. The gray shaded region comprises the addition and subtraction of 1 SD.

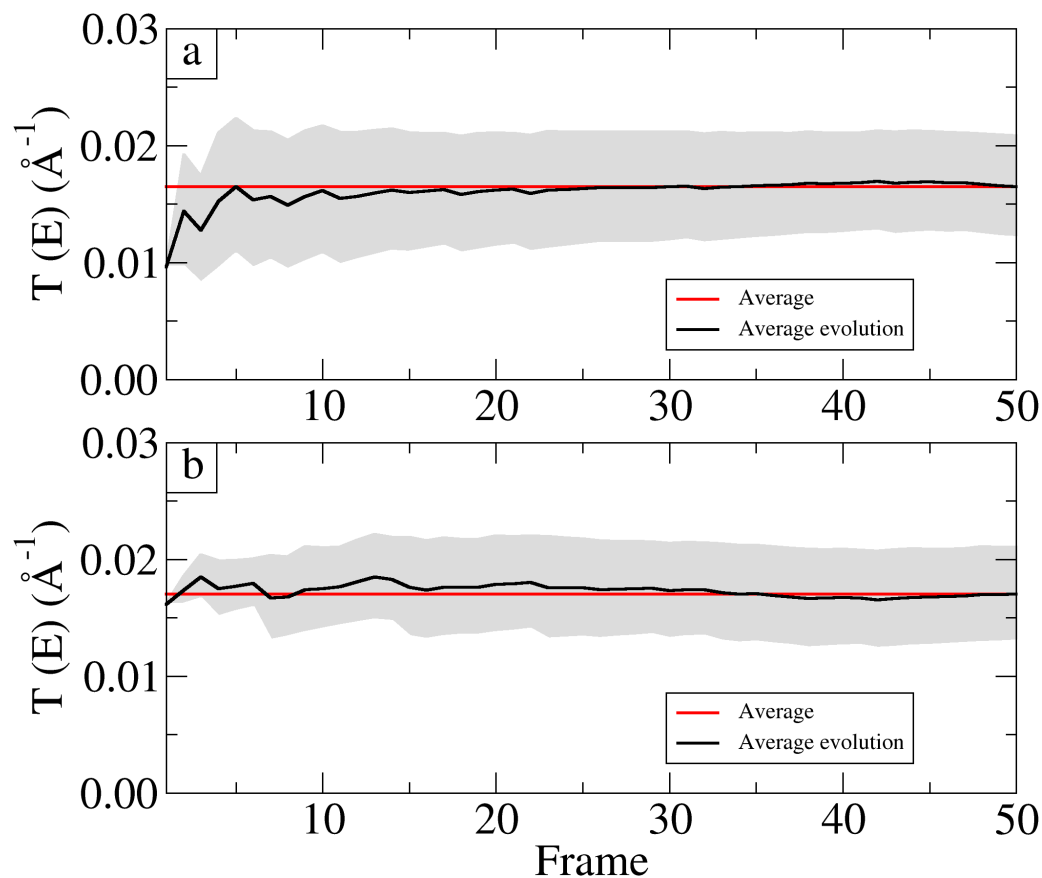


Figure 6.16: Evolution of the average and SD for vertical ssDNA (a) and dsDNA (b) systems without anchor molecule. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \approx 0.0$ eV.

The Figure 6.17 shows the averaged electronic transport for pristine graphene, pristine graphene in liquid, ssDNA and dsDNA. Differently from the previous cases (Section 6.2) the signals for ssDNA and dsDNA are different from the one of pristine graphene in solution. This result show that in this setup the device is more sensitive to the presence of DNA. However, the difference between the signals for ssDNA and dsDNA is still too small for a device.

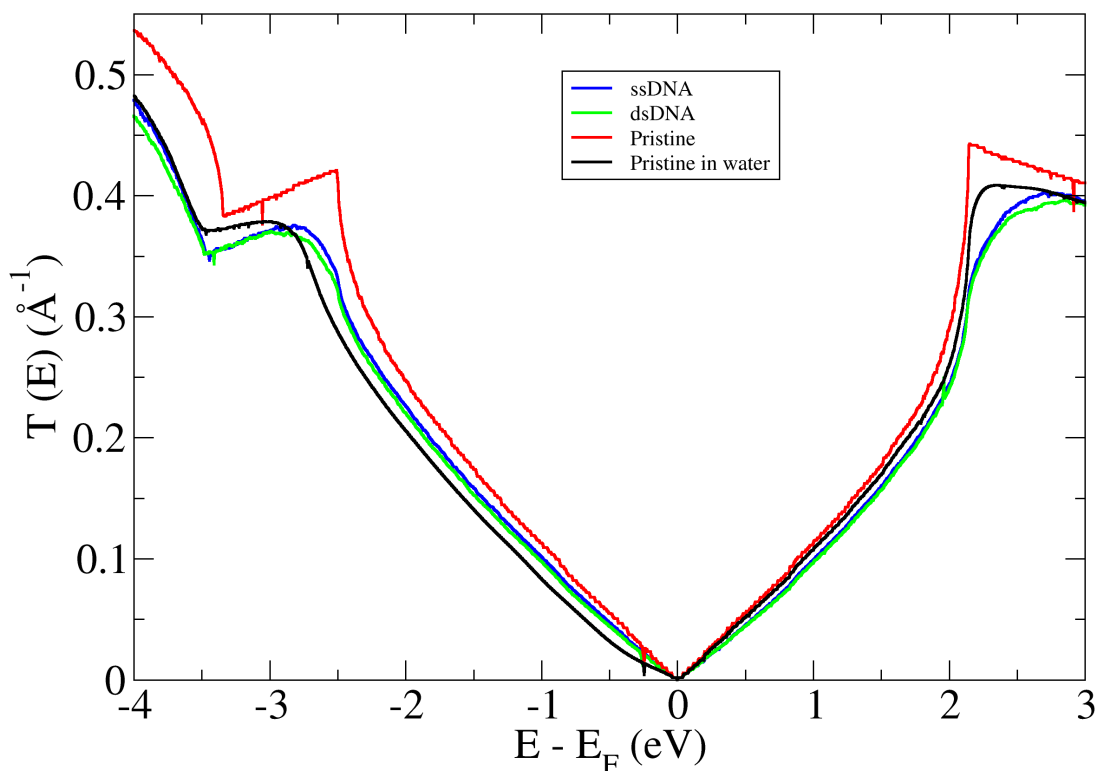


Figure 6.17: Comparison between the average over 50 electronic transport calculations for vertical ssDNA and dsDNA without anchor molecule.

In this new setup, as before, the reason to this similar signals for ssDNA and dsDNA is most likely the fact that the charged groups of the nucleotides are far from graphene. The presence of water quickly screens the charge. Therefore, when the ssDNA hybridizes forming the dsDNA, these new charges that come from the new nucleotides do not alter the electronic properties of graphene.

In summary, this new setup for DNA detection using the nucleotides itself as anchor molecules is more sensitive to the presence of DNA in comparison to the previous setup, since we observe that the signals for DNA strands are different from the signal of graphene in solution. The selectivity (ability to distinguish between ssDNA and dsDNA), however, is still not ideal in this setup.

6.4 DNA detection without anchor molecule - horizontal case

Finally, a last structural configuration that was considered is one in which there is no anchor molecule to intermediate the interaction between DNA and graphene, as in the case of Section 6.3, but here the DNA is initially placed horizontally on graphene. The horizontal placement allows the charged groups of DNA to be closer to graphene. A device of this kind can be developed by treating the DNA in another chamber and, then, deposit it onto graphene.

Charged groups close to graphene can, in principle, change the electronic environment of graphene and this change can alter the electronic transport throughout graphene sheet, providing differences in the signals for ssDNA and dsDNA. It is important to point out that DNA tends to strongly interact with graphene, and, as mentioned in the previous section, unless strong restraints are put in place, it tends to fully stack onto the sheet.

6.4.1 The systems

For this purpose we used 4 adenines as ssDNA and 4 adenines + 4 thymines as dsDNA. A side view of these two strands is shown in Figure 6.18. The procedure for structures generation is the same of the previous DNA systems. In this case, the DNA systems were restrained only in the z direction, as we can see in Figure 6.19. In this way, we prevent the structures in which the DNA systems are above the electrodes.

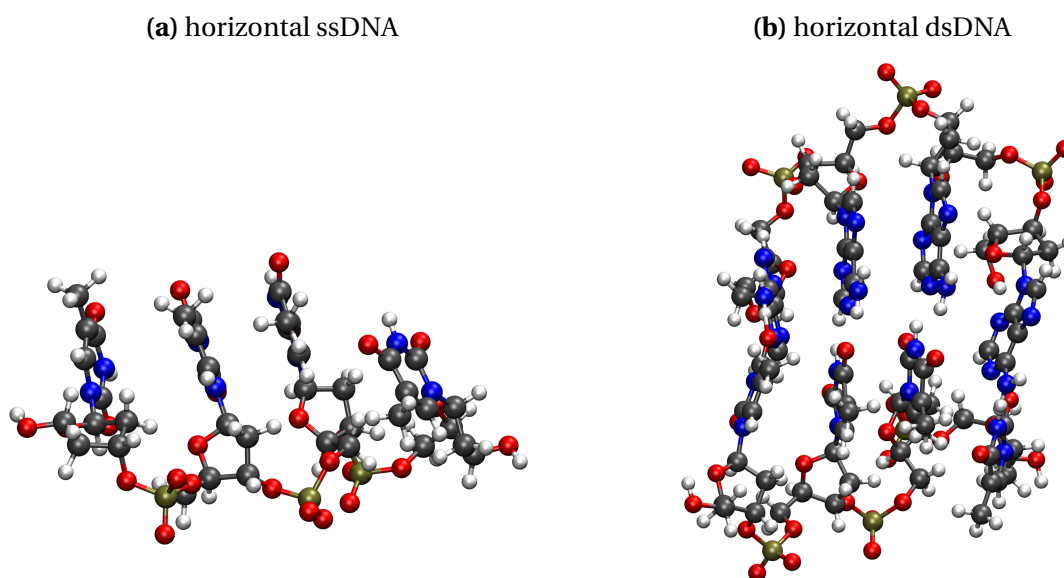


Figure 6.18: Representations of horizontal ssDNA (a) and dsDNA (b) systems. Sequences 4A and 4A + 4T.

The ssDNA is initially positioned with the phosphate groups close to graphene. Since here we don't apply restraints (in y direction) to prevent the π - π stacking interactions, we can observe, for ssDNA, the formation of 3 π - π stacking interactions between the nucleotides

and graphene, whereas the fourth nucleotide is not directly interacting with graphene. A side and top view for a representative snapshot of ssDNA with these π - π stacking interactions is shown in Figures 6.20a and 6.20c.

The dsDNA (side view in Figure 6.18b), on the other hand, is positioned with the nucleobases close to graphene, as shown in Figures 6.20b and 6.20d. In this case we observe π - π stacking for only a few snapshots (in the one showed in the figure we don't observe these interactions). This case of dsDNA is different from all the other cases presented before for the interaction between DNA systems and graphene sheet because now there is no π - π stacking interaction keeping the DNA systems somehow connected to graphene. However, even without π - π stacking interaction and restraints, the dsDNA remains interacting with graphene during all MD simulations.

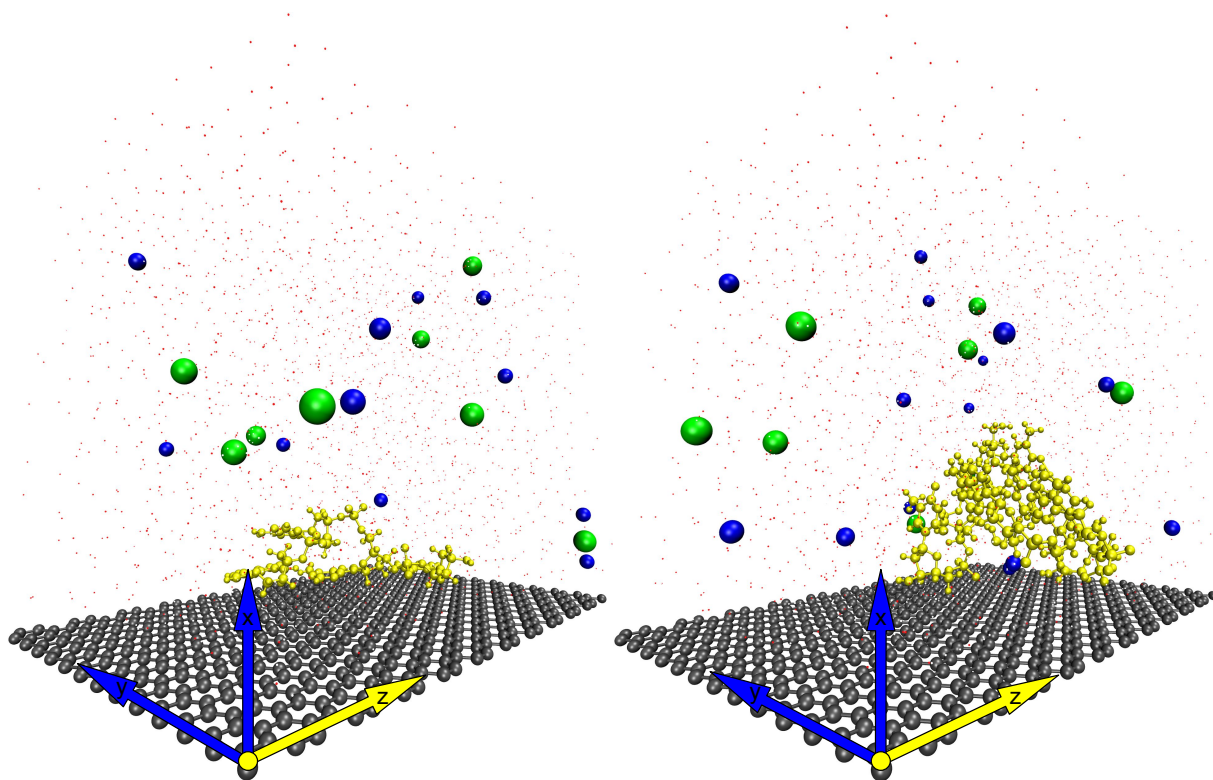


Figure 6.19: Restraint applied for ssDNA (left) and dsDNA (right) for the horizontal case. All atoms are restrained in both cases in the direction of the yellow axis. There is no restraint in the direction of the blue axes (x and y directions).

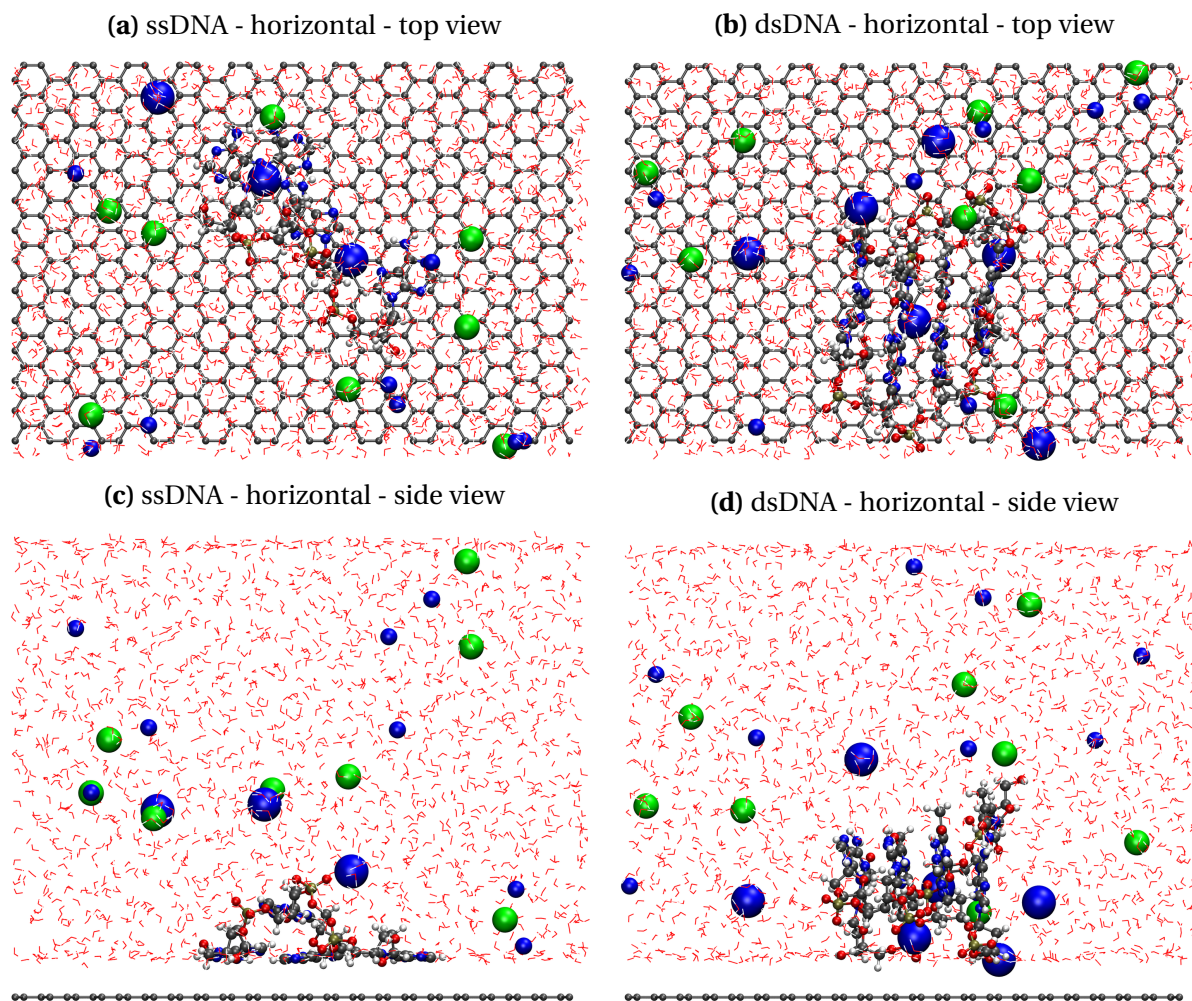


Figure 6.20: Top and side views for a representative snapshot of horizontal ssDNA (a and c) and dsDNA (b and d) without QM water. 3 Na atoms in the QM part for ssDNA and 6 Na atoms in the QM part for dsDNA.

The simulation details related with the systems presented in this section are shown in Table 6.4.

ssDNA	
Box dimensions (x, y, z)	(40.414, 31.980, 46.860)
N_{ions}	11 Na and 8 Cl
$N_{water\ molecules}$	1791
N_{DNA}	127
$N_{graphene}$	572 C
dsDNA	
Box dimensions (x, y, z)	(40.211, 31.980, 46.860)
N_{ions}	14 Na and 8 Cl
$N_{water\ molecules}$	1745
N_{DNA}	254
$N_{graphene}$	572 C

Table 6.4: Details of the simulation box for the study of horizontal DNA detection. N is the number of the specified species. Length unit: Å.

6.4.2 Electronic transport results

The Figure 6.21 shows the electronic transport properties calculated for each frame of ssDNA and dsDNA as well as the average over these curves. As we can see, the variation range for dsDNA is larger than the variation for ssDNA, particularly for the conduction band. However, for energies below the Fermi energy we can observe that the SD for dsDNA is small.

The convergence of the electronic transport can be checked using Figure 6.23 in which we present the evolution of the averages (black curve) over different numbers of frames as well as the evolution of the SD (gray shaded region). As we observe, the transmittance and the SD for $E - E_F = E'$ are converged, since they do not change from 25 frames onwards.

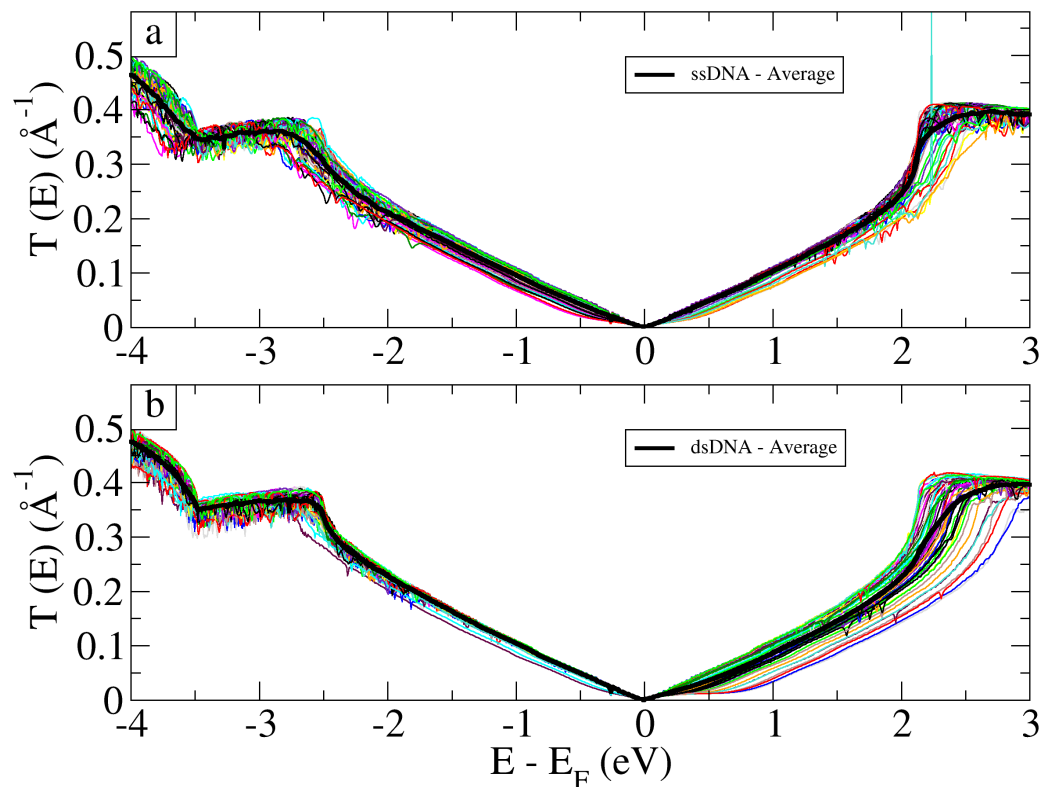


Figure 6.21: Electronic transport calculations for all frames of horizontal ssDNA (a) and dsDNA (b) systems. The average over these curves is shown in black.

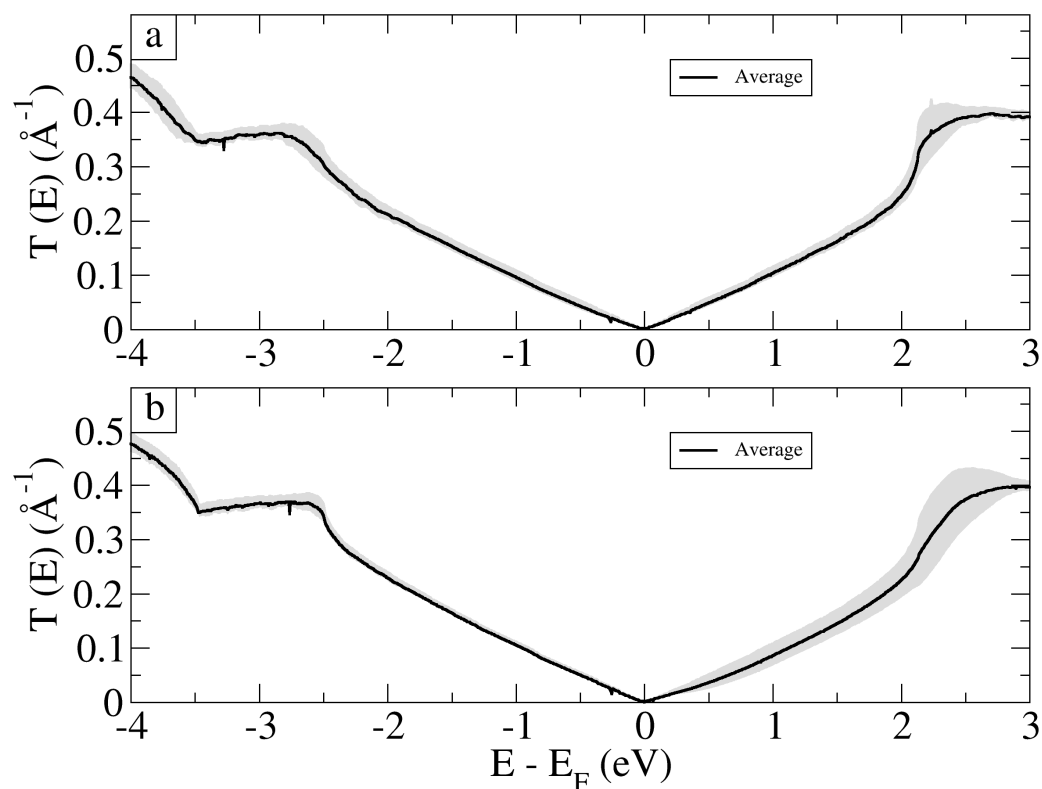


Figure 6.22: Electronic transport calculations for all horizontal DNA systems: (a) ssDNA and (b) dsDNA. The gray shaded region comprises the addition and subtraction of 1 SD.

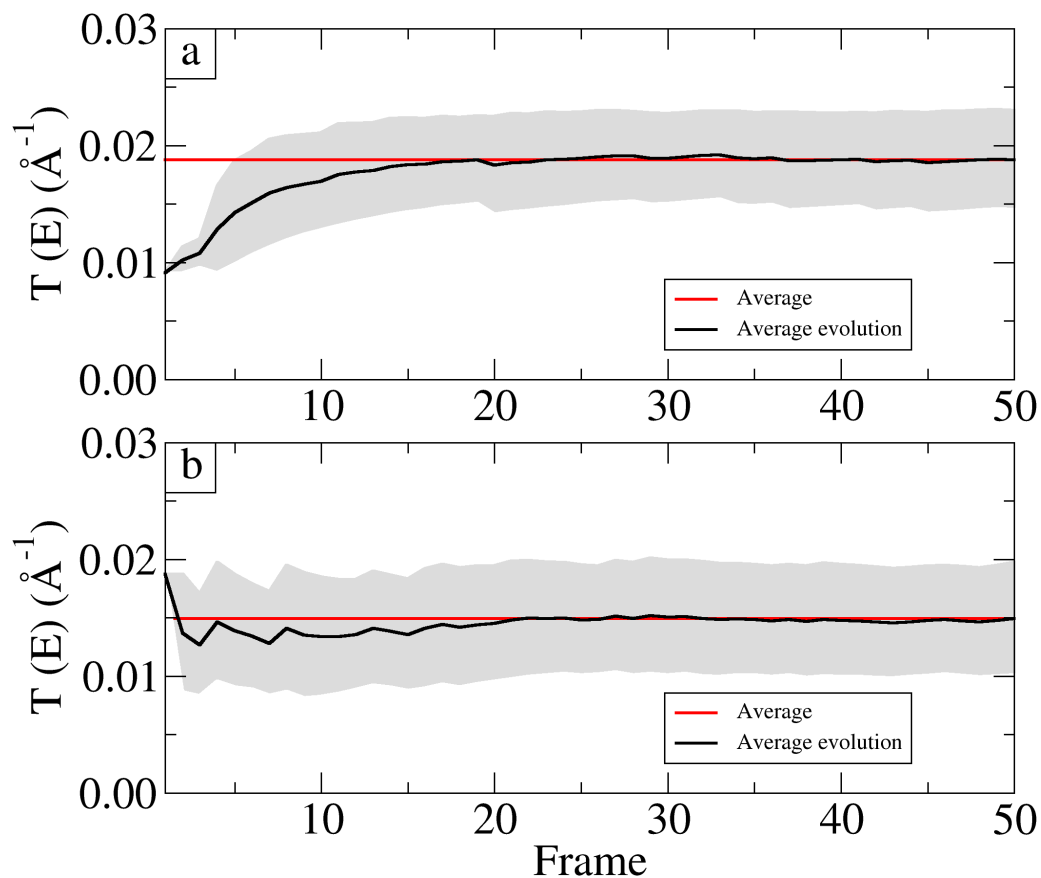


Figure 6.23: Evolution of the average and SD for all horizontal DNA systems: (a) ssDNA and (b) dsDNA. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \simeq 0.0 \text{ eV}$.

The comparison between the electronic transports for ssDNA, dsDNA, pristine graphene and pristine graphene in liquid are shown in Figure 6.24. In this case for horizontal DNA we can see that the signals for DNA are different from the signal of pristine graphene in liquid (same solvation conditions). In the case of vertical DNA using pyrene as anchor molecule (Section 6.2) the signals for DNA and pristine in liquid were similar.

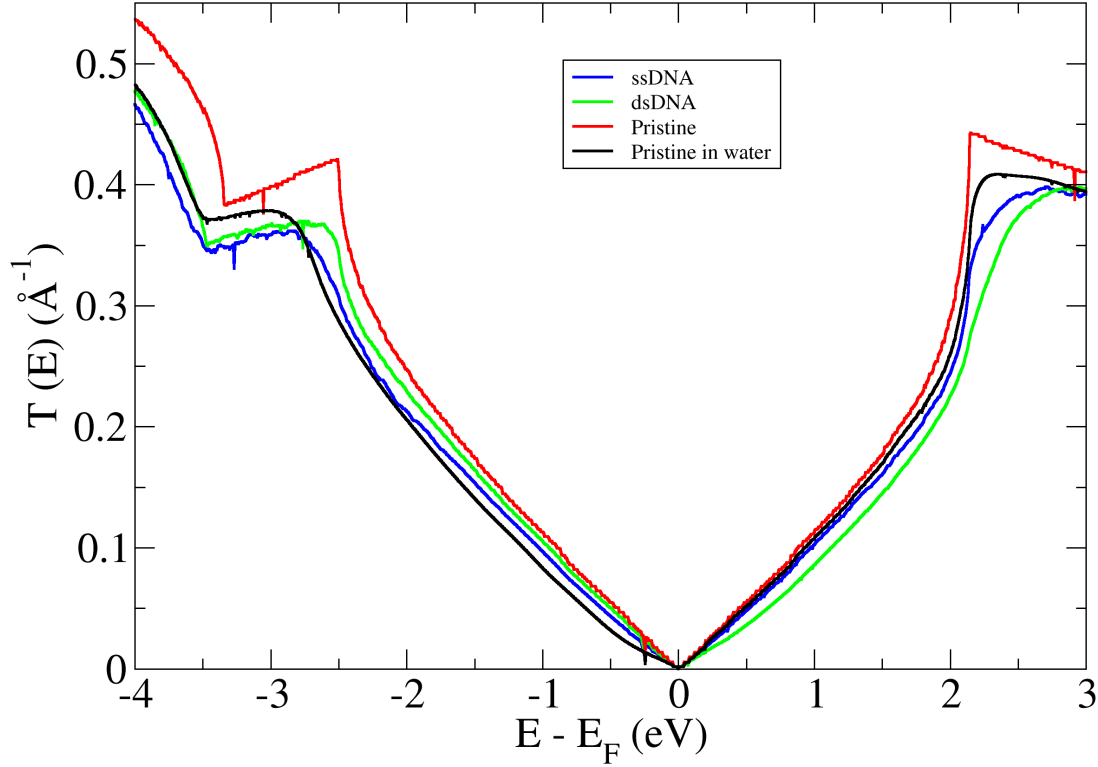


Figure 6.24: Comparison between the average over 50 electronic transport calculations for horizontal ssDNA and dsDNA, average over 200 electronic transport calculations for graphene in liquid and single calculation for dry pristine graphene.

Another aspect present in Figure 6.24 is that the transmittance for DNA is larger than the transmittance of pristine graphene in liquid, for energies below the Fermi Energy, and smaller above it. Most importantly, we note that for this placement of DNA we can clearly see that the averaged transmittances for ssDNA and dsDNA are different. In the previous cases both transports were virtually the same.

6.4.3 Sensitivity and selectivity of the device

For this setup we have done an analysis of the sensitivity and selectivity of the device, given the curves in Figure 6.24. Thus, Figure 6.26 present the results for sensitivity and selectivity for this system for three gate voltages (V_g). The sensitivity and selectivity for the whole energy spectrum is shown in Figure 6.26. By sensitivity calculations we can estimate how sensitive to the presence of ssDNA or dsDNA the device is. By selectivity calculations we can estimate how distinguishable the signals for ssDNA and dsDNA are. Then, Equations 6.1 and 6.2 show how we can calculate the sensitivity and selectivity using the average electronic transport for these systems:

$$\text{Sensitivity}(E') = \frac{T_x(E') - T_L(E')}{\overline{T_L(E')}} \times 100\%, \quad (6.1)$$

$$Selectivity(E') = \left(\frac{T_{dsDNA}(E')}{T_{ssDNA}(E')} - 1 \right) \times 100\%, \quad (6.2)$$

where $x = ssDNA$ or $dsDNA$ and $T_L(E')$ is the average electronic transport for the system in liquid for that energy.

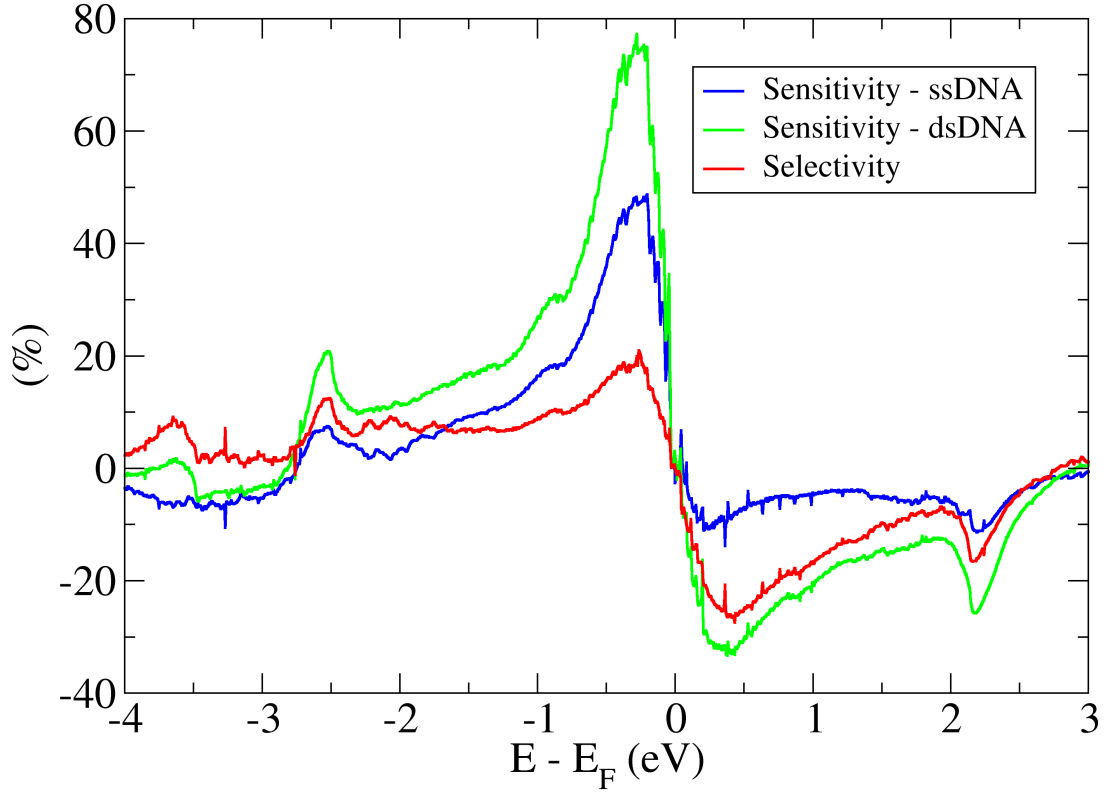


Figure 6.25: Sensitivity for ssDNA and dsDNA using the results for graphene in solution as reference and selectivity for dsDNA using ssDNA as reference, for the whole energy spectrum.

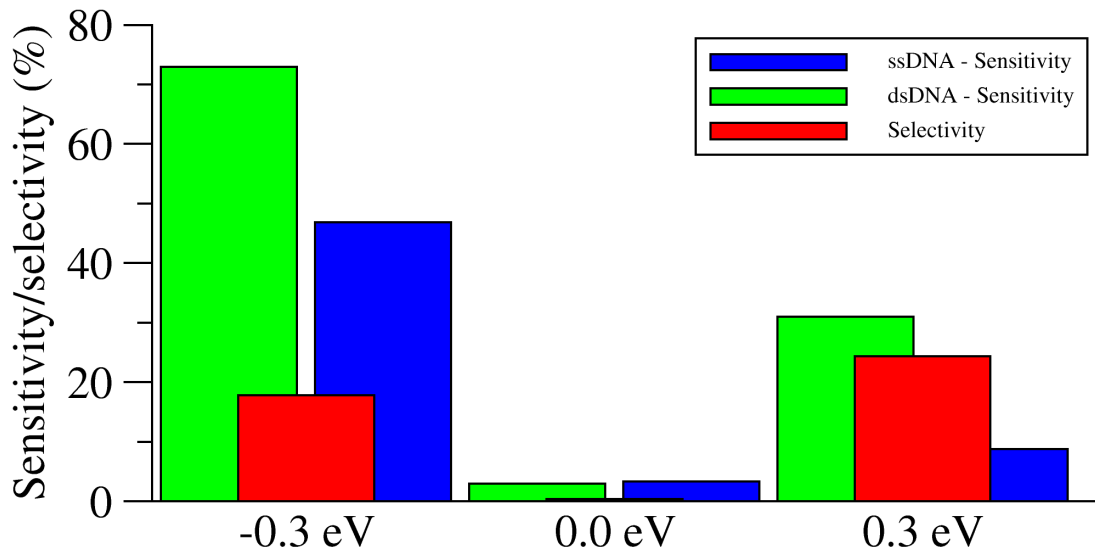


Figure 6.26: Sensitivity for ssDNA and dsDNA using the results for graphene in solution as reference and selectivity for dsDNA using ssDNA as reference, for specif energies.

As we can observe for selectivity at Fermi energy, the signals for both ssDNA and dsDNA are very small and the sensitivity is virtually zero. For energies slightly above and below the Fermi energy we can observe that they present sensitivities up to 80 %, particularly for $V_g < 0$.

This setup for DNA detection present the best results to be used as a device since we have a good sensitivity to the presence of DNA and also a good selectivity, because now we have different signals for ssDNA and dsDNA. These better results can be understood by analyzing the representative structures presented in Figure 6.20. As we can see, the typical structures for ssDNA and dsDNA are quite different. In the ssDNA case we have π - π stacking interactions and in the case of dsDNA we have this kind of interactions for only a few snapshots. Another aspect in these structures is the placement of the charged groups close to graphene sheet. When the other DNA strand is present, forming the dsDNA, the new charges will be close to the sheet, changing its local electronic environment, providing different signals, allowing the detection.

6.5 Conclusions

In conclusion, our results showed that the horizontal setup for this device presented the best results for DNA detection. In this setup differentiating the signals for ssDNA and dsDNA appears possible. The two vertical setups (with and without anchor molecule) presented similar signals, which is not ideal for a device.

We also showed that the inclusion of one layer of water in the QM part to simulate these systems do not increase the difference between the signals. Thus, the water does not play a important role in these setups for DNA detection.

Final conclusions and perspectives

In conclusion, we have presented a full dynamical investigation of electronic transport calculations for DNA sensing devices using QM/MM methods coupled to NEGF. Through the combination of these methodologies we were able to investigate the electronic transport properties of graphene-based devices comprising thousands of atoms, since it is possible to divide the system into a quantum and a classical part, which considerably reduces computational costs. We have investigated systems which had, in some cases, one thousand five hundred atoms in the quantum part (DFT calculation) repeated up to 200 times upon the convergence, in each system. Thus, our calculations can be considered a benchmark in the field of electronic transport calculations of solvated systems.

Firstly, we have investigated the electronic transport properties of graphene in solution. Our results showed that the overall trend of including solvent effects in the electronic transport through graphene is decrease the total transmission. However, include different salt concentration in the solution or taking into account polarization effects in graphene do not significantly change the electronic transport properties of graphene. Therefore, the inclusion of solvent effect is fundamental to study biological systems, such as DNA.

The DNA sequencing devices studied by us presented good sensitivity and selectivity. In the case of nanopore in between line defects in graphene, we showed that the overall trend for electronic transport in our calculations, where the effect of the solvent is taken into account, is similar to the case calculated by our collaborators for gas phase systems. We demonstrated that sequence DNA using this topological defects in graphene appears possible. For the case of nanogap in graphene we showed that the inclusion of explicit water molecules in the QM part increases the total transmission in severals orders of magnitude. We also showed that the current passes through the water molecules close to the nucleotides to be transmitted to the nucleotide and to the right side of the nanogap.

Finally, our results for DNA detection devices showed that the horizontal position is the best conformation for DNA above pristine graphene, because this position allows different average signals for ssDNA and dsDNA. In this position, the charged groups of DNA are closer to the sheet, thus the presence of more nucleotides when we add the other DNA strand, forming the dsDNA, is able to change the electronic transport properties of graphene, allowing the detection.

The results we have presented herein showed that the use of graphene as a DNA sensor can present promising results. Our results can be used as a guide for developing new sensors and for understanding the mechanisms governing its operation.

Further investigations can be done for a deeper understanding of the results presented in Chapter 6 regarding the DNA chip design. More studies can be done in those cases to

understand the effect of the used DFT functional and to investigate the device's ability of detecting another sequences of DNA.

- ¹ M. Kagan K. Kivirand and T. Rinken. *State of the Art in Biosensors - General Aspects*. 2013.
- ² F. Schedin, A. K. Geim, S. V. Morozov, E.W. Hill, P. Blake, M. I. Katsnelson, and K. S. Novoselovi. Detection of individual gas molecules adsorbed on graphene. *Nature Materials*, 6:652, 2007.
- ³ Ji Li, Dapeng Yu, and Qing Zhao. Solid-state nanopore-based DNA single molecule detection and sequencing. *Microchimica Acta*, 183(3):941–953, Mar 2016.
- ⁴ Lewis J. Alberts B., Johnson A. and et al. *Molecular Biology of the Cell*. New York: Garland Science, 4 edition, 2002.
- ⁵ K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov. Electric field effect in atomically thin carbon films. *Science*, 306(5696):666–669, 2004.
- ⁶ P. Lauffer, K. V. Emtsev, R. Graupner, Th. Seyller, L. Ley, S. A. Reshanov, and H. B. Weber. Atomic and electronic structure of few-layer graphene on SiC(0001) studied with scanning tunneling microscopy and spectroscopy. *Phys. Rev. B*, 77:155426, Apr 2008.
- ⁷ A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Rev. Mod. Phys.*, 81:109–162, Jan 2009.
- ⁸ M.I Katsnelson and A.K Geim. Electron scattering on microscopic corrugations in graphene. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 366(1663):195–204, 2008.
- ⁹ Jing Guo, Youngki Yoon, and Yijian Ouyang. Gate electrostatics and quantum capacitance of graphene nanoribbons. *Nano Letters*, 7(7):1935–1940, 2007.
- ¹⁰ Hee Dong Jang, Sun Kyung Kim, Hankwon Chang, and Jeong-Woo Choi. 3D label-free prostate specific antigen (PSA) immunosensor based on graphene-gold composites. *Biosensors and Bioelectronics*, 63(0):546 – 551, 2015.
- ¹¹ Vinod Kumar, Saurabh Srivastava, Sima Umrao, Ram Kumar, Gopal Nath, Gajjala Sumana, Preeti S. Saxena, and Anchal Srivastava. Nanostructured palladium-reduced graphene oxide platform for high sensitive, label free detection of a cancer biomarker. *RSC Adv.*, 4:2267–2273, 2014.

- ¹² Minghui Yang, Alireza Javadi, He Li, and Shaoqin Gong. Ultrasensitive immunosensor for the detection of cancer biomarker based on graphene sheet. *Biosensors and Bioelectronics*, 26(2):560 – 565, 2010.
- ¹³ Martin Pumera. Graphene in biosensing. *Materials Today*, 14(7-8):308–315, July 2011.
- ¹⁴ Yijia Wang, Shiwu Zhang, Ting Xu, Tonglin Zhang, Yuru Mo, Jun Liu, Liquan Yan, and Fei Xing. Ultra-sensitive and ultra-fast detection of whole unlabeled living cancer cell responses to paclitaxel with a graphene-based biosensor. *Sensors and Actuators B: Chemical*, 263:417 – 425, 2018.
- ¹⁵ Jagpreet Singh, Aditi Rathi, Mohit Rawat, and Manoj Gupta. Graphene: from synthesis to engineering to biosensor applications. *Frontiers of Materials Science*, 12(1):1–20, Mar 2018.
- ¹⁶ Gustavo T. Feliciano, Carlos Sanz-Navarro, Mauricio Domingues Coutinho-Neto, Pablo Ordejón, Ralph H. Scheicher, and Alexandre Reily Rocha. Addressing the environment electrostatic effect on ballistic electron transport in large systems: A QM/MM-NEGF approach. *The Journal of Physical Chemistry B*, 122(2):485–492, 2018.
- ¹⁷ Rodrigo G. Amorim, Alexandre R. Rocha, and Ralph H. Scheicher. Boosting DNA recognition sensitivity of graphene nanogaps through nitrogen edge functionalization. *The Journal of Physical Chemistry C*, 120(34):19384–19388, 2016.
- ¹⁸ Johan Lagerqvist, Michael Zwolak, and Massimiliano Di Ventra. Fast DNA sequencing via transverse electronic transport. *Nano Letters*, 6(4):779–782, 2006.
- ¹⁹ Jariyane Prasongkit, Anton Grigoriev, Biswarup Pathak, Rajeev Ahuja, and Ralph H. Scheicher. Transverse conductance of DNA nucleotides in a graphene nanogap from first principles. *Nano Letters*, 11(5):1941–1945, 2011.
- ²⁰ M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas. Generalized many-channel conductance formula with application to small rings. *Phys. Rev. B*, 31:6207–6215, May 1985.
- ²¹ M. I. Abedin and A. Hassan. A novel approach to label-free detection of DNA hybridization based on graphene quantum capacitance dependent frequency/wireless readout. *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, May 2015.
- ²² El Mustapha Bahassi and Peter J. Stambrook. Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis*, 29(5):303–310, 2014.
- ²³ Cost per megabase of DNA sequence. Access in may 28, 2018.

- ²⁴ Stephanie J. Heerema and Cees Dekker. Graphene nanodevices for DNA sequencing. *Nat Nano*, 11(2):127 – 136, 2016.
- ²⁵ Grégory F. Schneider and Cees Dekker. DNA sequencing with nanopores. *Nature Biotechnology*, 30, 2012.
- ²⁶ X. Shen, B. Gu, S. A. Che, and F. S. Zhang. Solvent effects on the conformation of DNA dodecamer segment: A simulation study. *The Journal of Chemical Physics*, 135(3):034509, 2011.
- ²⁷ Yoshihiko Arao, Fumiya Mori, and Masatoshi Kubouchi. Efficient solvent systems for improving production of few-layer graphene in liquid phase exfoliation. *Carbon*, 118:18–24, JUL 2017.
- ²⁸ Xiaodan Xu, Lu Cai, Xiaoli Zheng, and Qun Xu. Molecular dynamics simulations of solvent-exfoliation and stabilization of graphene with the assistance of compressed carbon dioxide and pyrene-polyethylene glycol. *Phys. Chem. Chem. Phys.*, 19:16062–16070, 2017.
- ²⁹ Vahid Shabafrooz, Sudheer Bandla, and Jay C. Hanan. Graphene dispersion in a surfactant-free, polar solvent. *Journal of Materials Science*, 53(1):559–572, JAN 2018.
- ³⁰ Vaishali Arunachalam and Sukumaran Vasudevan. Graphene-Solvent Interactions in Nonaqueous Dispersions: 2D ROESY NMR Measurements and Molecular Dynamics Simulations. *Journal of Physical Chemistry C*, 122(3):1881–1888, JAN 25 2018.
- ³¹ Adam Hardy, James Dix, Christopher D. Williams, Flor R. Siperstein, Paola Carbone, and Henry Bock. Design rules for graphene and carbon nanotube solvents and dispersants. *ACS Nano*, 12(2):1043–1049, 2018.
- ³² H. J. Salavagione, J. Sherwood, M. De bruyn, V. L. Budarin, G. J. Ellis, J. H. Clark, and P. S. Shuttleworth. Identification of high performance solvents for the sustainable processing of graphene. *Green Chem.*, 19:2550–2560, 2017.
- ³³ Jr. Josefredo R. Pliego, , and José M. Riveros. The clustercontinuum model for the calculation of the solvation free energy of ionic species. *The Journal of Physical Chemistry A*, 105(30):7241–7247, 2001.
- ³⁴ Sereina Riniker, Jane R. Allison, and Wilfred F. van Gunsteren. On developing coarse-grained models for biomolecular simulation: a review. *Phys. Chem. Chem. Phys.*, 14:12423–12430, 2012.
- ³⁵ H.J.C. Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1):43 – 56, 1995.

- ³⁶ Klaus Capelle. A bird's-eye view of density-functional theory. *BRAZILIAN JOURNAL OF PHYSICS*, 36(4A):1318–1343, DEC 2006.
- ³⁷ Klaus Capelle and Vivaldo L. Campo Jr. Density functionals and model hamiltonians: Pillars of many-particle physics. *Physics Reports*, 528(3):91 – 159, 2013. Density functionals and model Hamiltonians: Pillars of many-particle physics.
- ³⁸ Carlos F. Sanz-Navarro, Rogeli Grima, Alberto García, Edgar A. Bea, Alejandro Soba, José M. Cela, and Pablo Ordejón. An efficient implementation of a QM-MM method in SIESTA. *Theoretical Chemistry Accounts*, 128(4):825–833, Mar 2011.
- ³⁹ U. Rothlisberger and P. Carloni. *Drug-Target Binding Investigated by Quantum Mechanical/Molecular Mechanical (QM/MM) Methods*, pages 449–479. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
- ⁴⁰ Jorge Kohanoff. *Electronic structure calculations for solids and molecules : theory and computational methods*. Condensed matter physics, nanoscience and mesoscopic physics. Cambridge University Press, 2006.
- ⁴¹ P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.
- ⁴² W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.
- ⁴³ J. M. Haile. *Molecular dynamics simulation: elementary methods*. Monographs in Physical Chemistry. Wiley-Interscience, 1992.
- ⁴⁴ Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725.
- ⁴⁵ H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. *Interaction Models for Water in Relation to Protein Hydration*, pages 331–342. Springer Netherlands, Dordrecht, 1981.
- ⁴⁶ H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *The Journal of Physical Chemistry*, 91(24):6269–6271, 1987.
- ⁴⁷ Batrouni G.G. Le Bellac M., Mortessagne F. *Equilibrium and Non-Equilibrium Statistical Thermodynamics*. CUP, 2004.
- ⁴⁸ McQuarrie D.A. *Statistical Mechanics*. HARPER & ROW, 1976.

- ⁴⁹ Ira N. Levine. *Quantum Chemistry*. Prentice Hall, 5 edition, 1999.
- ⁵⁰ J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865, 1996.
- ⁵¹ M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der Waals density functional for general geometries. *Phys. Rev. Lett.*, 92:246401, Jun 2004.
- ⁵² N. Troullier and José Luís Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43:1993–2006, Jan 1991.
- ⁵³ P. Ordejón, E. Artacho, and J. M. Soler. Self-consistent r-N density-functional calculations for very large systems. *Phys. Rev. B*, 53:R10441, 1996.
- ⁵⁴ José M Soler, Emilio Artacho, Julian D Gale, Alberto García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal. The SIESTA method for ab initio order-N materials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745, 2002.
- ⁵⁵ D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler. Density-functional method for very large systems with LCAO basis sets. *Int. J. Quantum Chem.*, 65:453, 1997.
- ⁵⁶ Gerrit Groenhof. *Biomolecular Simulations: Methods and Protocols*, chapter Introduction to QM/MM Simulations, pages 43–66. Humana Press, Totowa, NJ, 2013.
- ⁵⁷ Alejandro Crespo, Damián A. Scherlis, Marcelo A. Martí, Pablo Ordejón, Adrián E. Roitberg, and Darío A. Estrin. A DFT-based QM-MM approach designed for the treatment of large molecular systems: application to chorismate mutase. *The Journal of Physical Chemistry B*, 107(49):13728–13736, 2003.
- ⁵⁸ Mads Brandbyge, José-Luis Mozos, Pablo Ordejón, Jeremy Taylor, and Kurt Stokbro. Density-functional method for nonequilibrium electron transport. *Phys. Rev. B*, 65:165401, Mar 2002.
- ⁵⁹ K.L. Ma, X.H. Yan, Youan Xiao, and Yuan Chen. Electronic transport properties of metallic graphene nanoribbons with two vacancies. 150:1308–1312, 08 2010.
- ⁶⁰ A. R. Rocha, V. M. García-Suárez, S. W. Bailey, C. J. Lambert, J. Ferrer, and S. Sanvito. Towards molecular spintronics. *Nature Materials*, 4:335, 2005.
- ⁶¹ A. R. Rocha, V. M. García-Suárez, S. Bailey, C. Lambert, J. Ferrer, and S. Sanvito. Spin and molecular electronics in atomically generated orbital landscapes. *Phys. Rev. B*, 73:085414, Feb 2006.
- ⁶² Frederico D. Novaes, Antônio J. R. da Silva, and A. Fazzio. Density functional theory method for non-equilibrium charge transport calculations: TRANSAMPA. *Brazilian Journal of Physics*, 36:799 – 807, 09 2006.

- ⁶³ C Caroli, R Combescot, P Nozieres, and D Saint-James. Direct calculation of the tunneling current. *Journal of Physics C: Solid State Physics*, 4(8):916, 1971.
- ⁶⁴ R. Landauer. Spatial variation of currents and fields due to localized scatterers in metallic conduction. *IBM Journal of Research and Development*, 1(3):223–231, July 1957.
- ⁶⁵ Daniel S. Fisher and Patrick A. Lee. Relation between conductivity and transmission matrix. *Phys. Rev. B*, 23:6851–6854, Jun 1981.
- ⁶⁶ T. Frederiksen, M. Paulsson, M. Brandbyge, and A.-P. Jauho. Inelastic transport theory from first principles: Methodology and application to nanoscale devices. *Phys. Rev. B*, 75(20):205413, 2007.
- ⁶⁷ M. Paulsson and M. Brandbyge. Transmission eigenchannels from nonequilibrium Green's functions. *Phys. Rev. B*, 76(11):115117, 2007.
- ⁶⁸ John Stone. *An Efficient Library for Parallel Ray Tracing and Animation*. Master's thesis, Computer Science Department, University of Missouri-Rolla, April 1998.
- ⁶⁹ William Humphrey, Andrew Dalke, and Klaus Schulten. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14:33–38, 1996.
- ⁷⁰ Harold S. Park Wing Kam Liu, Eduard G. Karpov. *Nano Mechanics and Materials: Theory, Multiscale Methods and Applications*, volume 1. John Wiley and Sons, first edition, feb 2006.
- ⁷¹ Ronald E. Miller Ellad B. Tadmor. *Modeling Materials: Continuum, Atomistic and Multiscale Techniques*, volume 1. Cambridge Universtity Press, first edition, nov 2011.
- ⁷² Manoj K. Shukla Jerzy Leszczynski. *Practical Aspects of Computational Chemistry: Methods, Concepts and Applications*, volume 1. Springer Netherlands, first edition, jan 2010.
- ⁷³ Martin Steinhauser. *Computational Multiscale Modeling of Fluids and Solids: Theory and Applications*. Springer-Verlag Berlin Heidelberg, first edition, 2008.
- ⁷⁴ Kim Sneppen and Giovanni Zocchi. *Physics in Molecular Biology*. Cambridge Universtity Press, first edition, 2005.
- ⁷⁵ E Weinan and Bjorn Engquist. Multiscale modeling and computation. *Notices of the AMS*, 50(9):1062–1070, 2003.
- ⁷⁶ Supriyo Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge Studies in Semiconductor Physics and Microelectronic Engineering. Cambridge University Press, 1995.

- ⁷⁷ Eunsu Paek, Alexander J. Pak, and Gyeong S. Hwang. On the influence of polarization effects in predicting the interfacial structure and capacitance of graphene-like electrodes in ionic liquids. *The Journal of Chemical Physics*, 142(2):024701, 2015.
- ⁷⁸ Tuan A. Ho and Alberto Striolo. Polarizability effects in molecular dynamics simulations of the graphene-water interface. *The Journal of Chemical Physics*, 138(5):054117, 2013.
- ⁷⁹ Zak E. Hughes, Susana M. Tomasio, and Tiffany R. Walsh. Efficient simulations of the aqueous bio-interface of graphitic nanostructures with a polarisable model. *Nanoscale*, 6:5438–5448, 2014.
- ⁸⁰ A.J. Stone and D.J. Wales. Theoretical studies of icosahedral C₆₀ and some related species. *Chemical Physics Letters*, 128(5):501 – 503, 1986.
- ⁸¹ Sara Letardi, Massimo Celino, Fabrizio Cleri, and Vittorio Rosato. Atomic hydrogen adsorption on a StoneWales defect in graphite. *Surface Science*, 496(1):33 – 38, 2002.
- ⁸² Wei Chen, Guo-Chang Liu, Jun Ouyang, Meng-Juan Gao, Bo Liu, and Yuan-Di Zhao. Graphene nanopores toward DNA sequencing: a review of experimental aspects. *Science China Chemistry*, 60(6):721–729, Jun 2017.
- ⁸³ Amir Barati Farimani, Payam Dibaeinia, and Narayana R. Aluru. DNA origamigraphene hybrid nanopore for DNA detection. *ACS Applied Materials & Interfaces*, 9(1):92–100, 2017.
- ⁸⁴ Jiapeng Li, Haochen Wang, Yusheng Li, and Kezhen Han. The impact of the number of layers of the graphene nanopores and the electrical field on ssDNA translocation. *Molecular Simulation*, 43(4):320–325, 2017.
- ⁸⁵ Roozbeh Abedini-Nassab. Nanotechnology and nanopore sequencing. *Recent Patents on Nanotechnology*, 11(1):34–41, 2017.
- ⁸⁶ Shaghayegh Agah, Ming Zheng, Matteo Pasquali, and Anatoly B Kolomeisky. DNA sequencing by nanopores: advances and challenges. *Journal of Physics D: Applied Physics*, 49(41):413001, 2016.
- ⁸⁷ Cees Dekker. Solid-state nanopores. *Nature Nanotechnology*, 2, 4 2007.
- ⁸⁸ Daniel Fologea, James Uplinger, Brian Thomas, David S. McNabb, and Jiali Li. Slowing DNA translocation in a solid-state nanopore. *Nano Letters*, 5(9):1734–1737, 2005.
- ⁸⁹ Daniel Fologea, Marc Gershow, Bradley Ledden, David S. McNabb, Jene A. Golovchenko, and Jiali Li. Detecting single stranded DNA with a solid state nanopore. *Nano Letters*, 5(10):1905–1909, 2005.

- ⁹⁰ Stefan Howorka and Zuzanna Siwy. Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.*, 38:2360–2384, 2009.
- ⁹¹ Stefan Howorka, Stephen Cheley, and Hagan Bayley. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nature Biotechnology*, 19, 2001.
- ⁹² Gustavo T. Feliciano, Carlos Sanz-Navarro, Mauricio Domingues Coutinho-Neto, Pablo Ordejón, Ralph H. Scheicher, and Alexandre Reily Rocha. Capacitive DNA detection driven by electronic charge fluctuations in a graphene nanopore. *Phys. Rev. Applied*, 3:034003, Mar 2015.
- ⁹³ Fabio A. L. de Souza, Rodrigo G. Amorim, Wanderla L. Scopel, and Ralph H. Scheicher. Electrical detection of nucleotides via nanopores in a hybrid graphene/h-BN sheet. *Nanoscale*, 9:2207–2212, 2017.
- ⁹⁴ Kwanpyo Kim, Allen Sussman, and A. Zettl. Graphene nanoribbons obtained by electrically unwrapping carbon nanotubes. *ACS Nano*, 4(3):1362–1366, 2010.
- ⁹⁵ Jayeeta Lahiri, You Lin, Pinar Bozkurt, Ivan I. Oleynik, and Matthias Batzill. An extended defect in graphene as a metallic wire. *Nature Nanotechnology*, 5, 2010.
- ⁹⁶ J.-H. Chen, G. Autès, N. Alem, F. Gargiulo, A. Gautam, M. Linck, C. Kisielowski, O. V. Yazyev, S. G. Louie, and A. Zettl. Controlled growth of a line defect in graphene and implications for gate-tunable valley filtering. *Phys. Rev. B*, 89:121407, Mar 2014.
- ⁹⁷ Dongwook Kim, Youngkuk Kim, Jisoon Ihm, Euijoon Yoon, and Gun-Do Lee. Atomic-scale mechanism of grain boundary motion in graphene. *Carbon*, 84:146 – 150, 2015.
- ⁹⁸ Fábio A.L. de Souza, Rodrigo G. Amorim, Jariyanee Prasongkit, Wanderlã L. Scopel, Ralph H. Scheicher, and Alexandre R. Rocha. Topological line defects in graphene for applications in gas sensing. *Carbon*, 129:803 – 808, 2018.
- ⁹⁹ Jariyanee Prasongkit, Ernane de Freitas Martins, Fábio A. L. de Souza, Wanderlã L. Scopel, Rodrigo G. Amorim, Vittaya Amornkitbamrung, Alexandre R. Rocha, and Ralph H. Scheicher. Topological line defects around graphene nanopores for DNA sequencing. *The Journal of Physical Chemistry C*, 122(13):7094–7099, 2018.
- ¹⁰⁰ Grégory F. Schneider, Qiang Xu, Susanne Hage, Stephanie Luik, Johannes N. H. Spoor, Sairam Malladi, Henny Zandbergen, and Cees Dekker. Tailoring the hydrophobicity of graphene for its use as nanopores for dna translocation. *Nature Communications*, 4:2619, Oct 2015.
- ¹⁰¹ Henk W. Ch. Postma. Rapid sequencing of individual DNA molecules in graphene nanogaps. *Nano Letters*, 10(2):420–425, 2010.

- ¹⁰² Hedyeh Karimi, Rubiyah Yusof, Rasoul Rahmani, Hoda Hosseinpour, and Mohammad T. Ahmadi. Development of solution-gated graphene transistor model for biosensors. *Nanoscale Research Letters*, 9, Feb 2014.

Appendices

APPENDIX A

--

QM/MM partition test

In this appendix we will present a QM/MM partition test for the ssDNA system using pyrene as anchor molecule. In all simulations presented in the thesis we included all DNA strands in the QM part. Here we will analyze the effect, on the electronic transport, of including only some nucleotides in the QM part. Thus, the QM/MM partitions that we will test here include chemical bond breaking.

In principle, the best choice for QM/MM partition is to include all DNA strand in the QM part, as we made before. However, with a QM/MM partition test we can conclude if the electronic transport calculation in a Hamiltonian generated from a small QM part is the same of the case with all DNA (or more nucleotides) in the QM part. For the DNA strands that we used in this work it was possible to include the entire DNA strand in the QM part, but for bigger DNA strands the calculation can become unpractical.

A.1 The QM/MM partitions

For this study we used 10 frames equally spaced by 10 *ns* extracted from the dynamics of the system using pyrene as anchor molecule (Section 6.2, Chapter 6). The simulation details related with the QM/MM partitions tests presented in this section are the same as the presented in Table 6.4 (Section 6.2, Chapter 6). The 5 tested QM/MM partitions are schematically shown in Figure A.1. The number of Na ions in the QM part is always choose in such a way that the charge of both partitions is neutral. So, for the QM/MM partition 1 there is no Na ion in the QM part, since there is no phosphate group. For all others partitions we have Na ions in the QM part. All QM partitions include graphene and "liquid" means all water molecules, Cl ions and Na ions not included in the QM part.

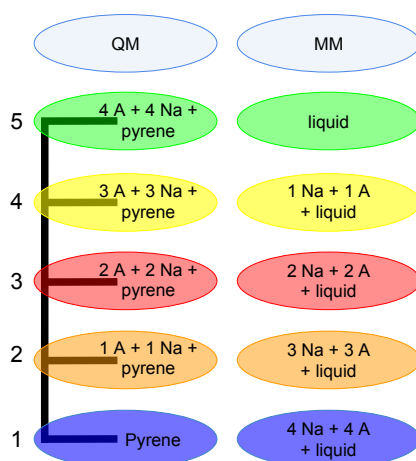


Figure A.1: Schematic representation of the investigated QM/MM partitions. Graphene is always in the QM part and liquid means water, Cl ions and all Na ions that are not included in the QM part.

A side view of the structures (example for one frame) for all QM/MM partitions, schematically showed in Figure A.1, is shown in Figure A.2. The representation pattern is the same as before.

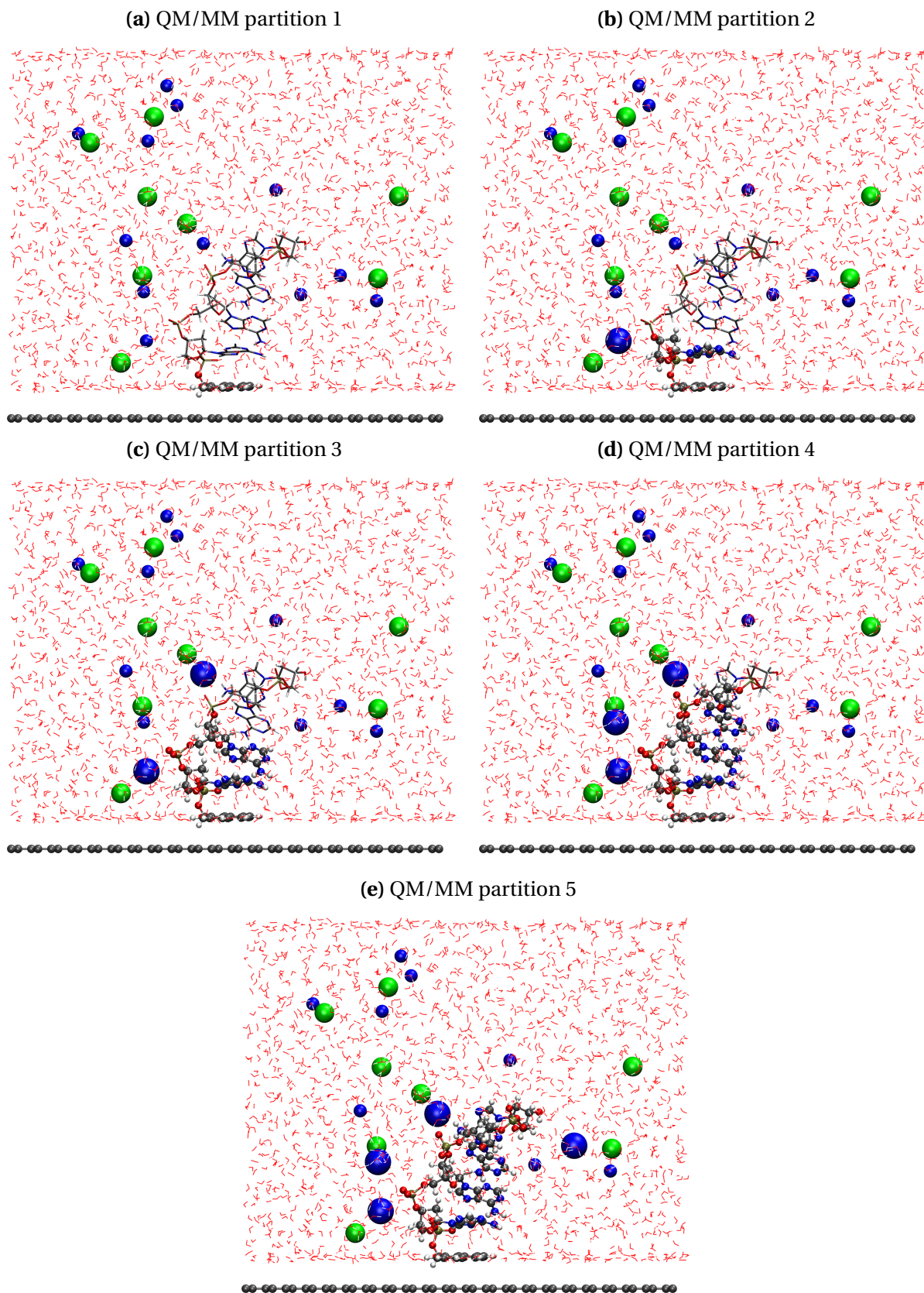


Figure A.2: Side view of the five different QM/MM partitions for a sequence of 4A connected to pyrene. The *ball-and-sticks* representations and large radius blue spheres representations make up the QM part. The small radius blue spheres, the green spheres and the *line* representations make up the MM part.

A.2 Electronic transport results

The average electronic transport for each QM/MM partition with one SD added and subtracted, is shown in Figure A.3. As we can observe, for QM/MM partitions 1 and 2 the SD is small. However, for the others partitions, the SD is a bit larger. For energies above the Fermi energy, the QM/MM partition 5 has the largest SD. This last partition coincides with the setup used before (Section 6.2), in which the entire DNA strand is in the QM part.

A convergence analysis for $E = E'$ is shown in Figure A.4. Despite the small number of frames, we can observe that the electronic transport and the SD are converged. However, for QM/MM partition 5 it is not clear that the electronic transport is converged. Since this partition is the same as before and we have done 50 calculations in Section 6.2, we can compare both averages. Thus, Figure A.5 shows the comparison between the averages over 10 and 50 frames for QM/MM partition 5 and as we can observe, the average over 10 frames (solid blue line) is virtually the same of the average over 50 frames (dotted black line). This shows that the averaged electronic transport over 10 frames for QM/MM partition 5 is also converged.

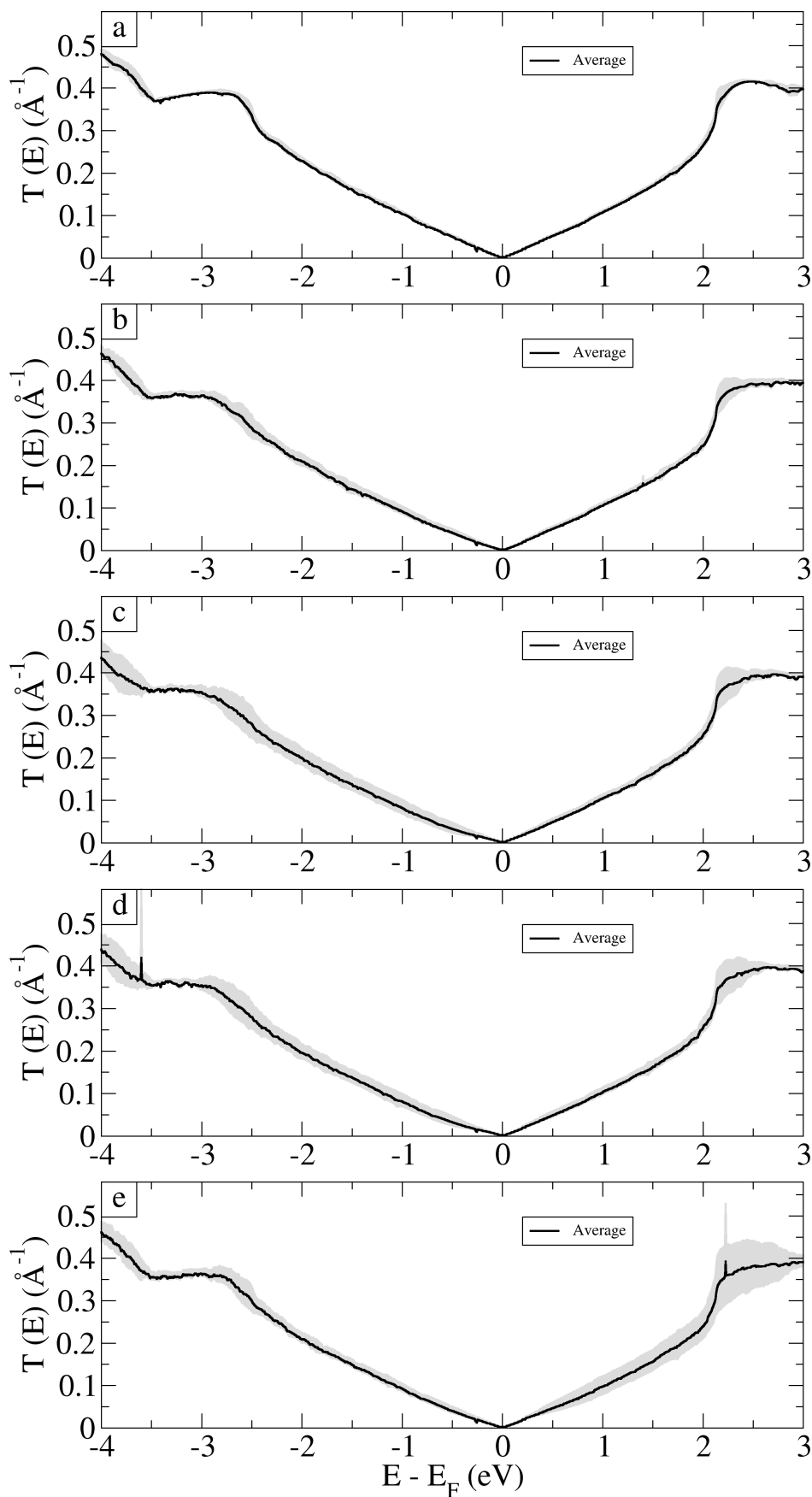


Figure A.3: Electronic transport calculations for QM/MM partitions (a) 1, (b) 2, (c) 3, (d) 4 and (e) 5. The gray shaded region comprises the addition and subtraction of 1 SD.

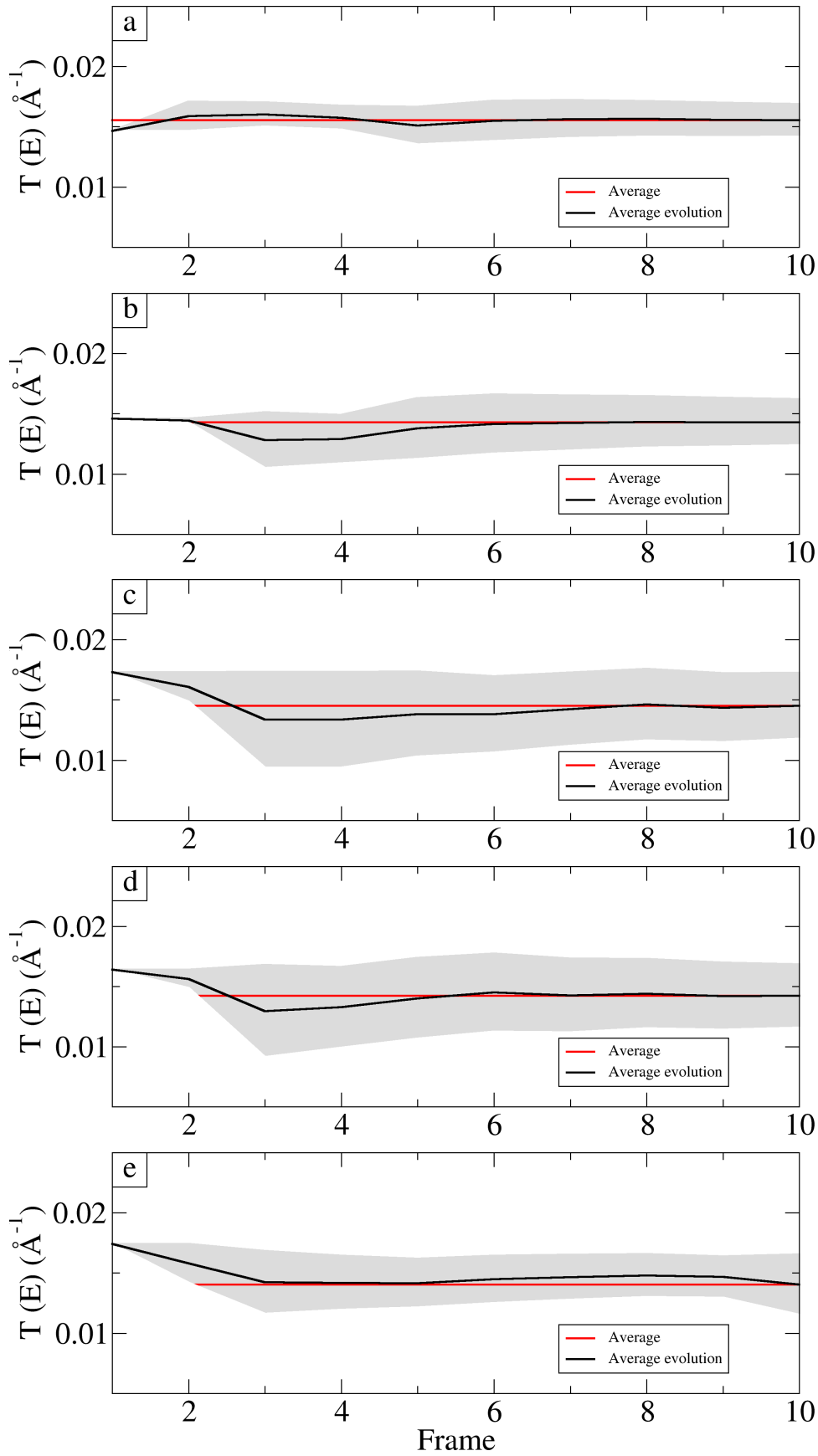


Figure A.4: Evolution of the average and SD for QM/MM partitions (a) 1, (b) 2, (c) 3, (d) 4 and (e) 5. The gray shaded region comprises the addition and subtraction of 1 SD. Chosen energy $E' \simeq 0.0$ eV.

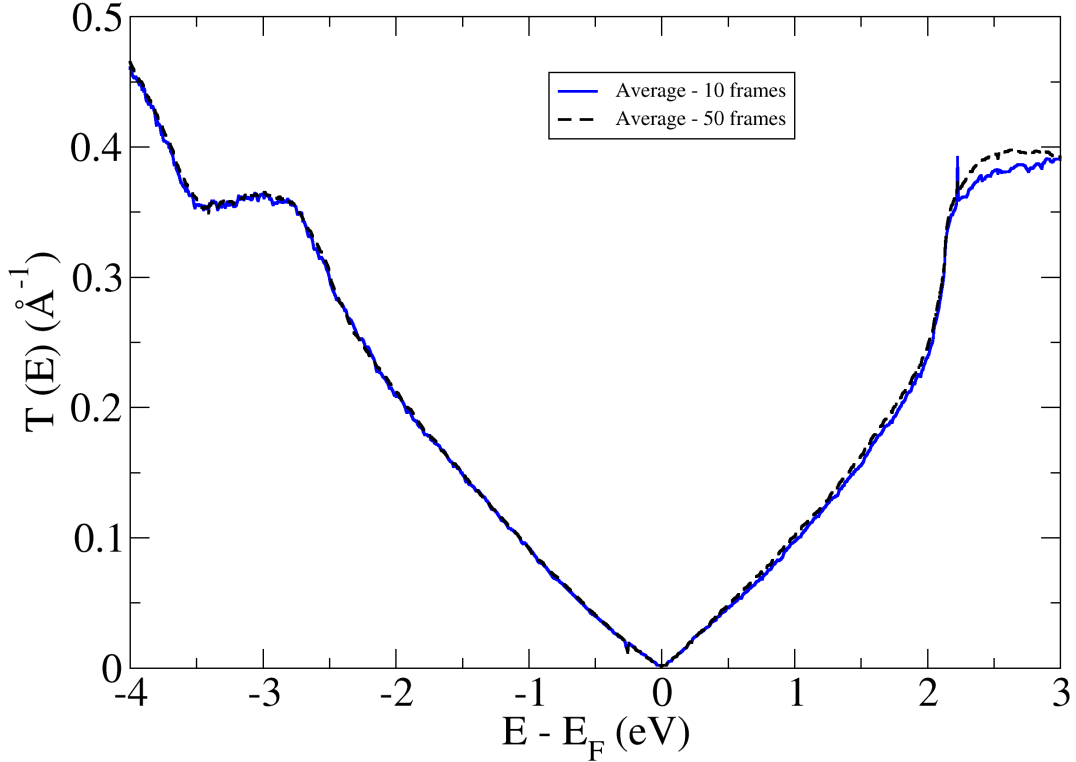


Figure A.5: Comparison of the averages over 10 and 50 electronic transport calculations for QM/MM partition 5.

A comparison between the averaged electronic transport for all QM/MM partitions can be seen in Figure A.6. Since the transmission for energies above the Fermi energy has a large SD, we can analyze the range of energy below the Fermi Energy. Thus, for energies below the Fermi Energy, the QM/MM partition 1, which includes only pyrene in the QM part, has the highest transmission. For QM/MM partitions 2 and 3 we can observe that the transmissions decrease and are virtually the same, indicating that 2 or 3 bases in the QM part should give the same results. These first three results are expected since by increasing the number of atoms the number of scattering centers is also increased. Finally, for QM/MM partition 5 we have the same transmission of QM/MM partition 2, which indicate that the inclusion of only one base in the QM part can be enough to simulate the whole DNA strand.

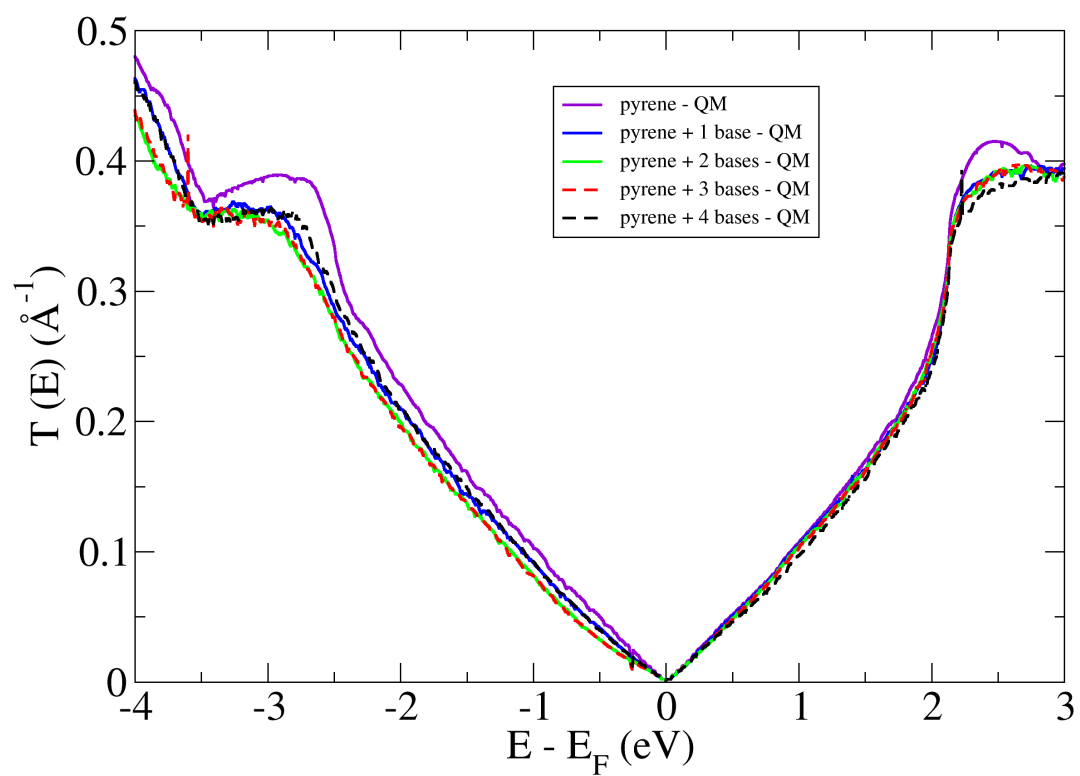


Figure A.6: Comparison of the averages over 10 electronic transport calculations for 5 different QM/MM partitions.

