UNIVERSIDADE FEDERAL DO ABC
Curso de Pós-Graduação em Nanociências e Materiais Avançados

Dissertação de Mestrado

Ekaterina A. Filatova

Simulating DNA sequencing in graphene nanopores: a QM/MM study to include dynamical and environmental effects

Santo André, SP, Brasil
2014

Universidade Federal do ABC

Simulating DNA sequencing in graphene nanopores: a QM/MM study to include dynamical and environmental effects

Ekaterina A. Filatova

Trabalho apresentado como parte dos requisitos para obtenção do título de Mestre em Nanociências e Materiais Avançados.

Orientador:
Prof. Dr. Alexandre Reily Rocha

Banca Examinadora:
Prof. Dr. Alexandre Reily Rocha - UNESP
Prof. Dr. Maurício Domingues Coutinho Neto - UFABC
Prof. Dr. Roberto Rivelino de Melo Moreno - UFBA

Santo André, SP, Brasil
2014

Abstract

Determining the DNA sequence in the human genome in a fast, cheap and reliable way can be the key to predicting a number of genetic disorders. The cost of sequencing a whole genome, from scratch, has been dramatically reduced over the years. The National Institutes of Health proposed the "US$1000 genome challenge" in 2004, which would be the target cost for sequencing one entire human genome. The technologies inspired by this challenge are expected to be a new step in modern medicine. One of the promissing technologies is DNA sequencing using graphene nanopores. In such a device the transverse current is measured while the DNA strand passes through a graphene nanopore. In principle, this device could significantly reduce the time and the cost of DNA sequencing in comparison with the existing methods, as one could envision a single-step process where a full strand is sequenced in one go.

In our project we proposed to simulate such a device. We were interested to see whereas different kind of pores help in stabilizing the bases. The second problem that we have adressed were if the transverse current changes when different nucleobases are inside the pore. In order to calculate electronic transport the system needed to be treated quantum mechanically. At the same time we were interested in dynamics of the DNA inside the pore and the effects of the water and ions. For this purporses we used a combination of density functional theory coupled to classical molecular dynamics to simulate the electronic structure of a graphene sheet as a DNA strand passes through a nanopore drilled in the sheet. The effects of the solvent and of counter-ions were included in the simulation via the classical part of the simulation (using QM/MM approach) whereas the rest is simulated quantum mechanically (using DFT). From different snapshots of the simulation we used the Non-equilibrium Greens functions methodology to calculate the electronic transport properties of this system.

Our results show that pore termination and the presence of water tend to significantly alter the results. In particular, different pores, either hydrogen or nitrogen lead to significantly different charge transfer effects. Water molecules, either classically or quantum-mechanically described tend to screen this charge transfer.

From the transport point of view we also observe significant changes due to pore termination, but we also note that there is not enough distinction between bases in either case.

For my parents,
Ludmila and *Alexandr Filatovi.*

Acknowledgements

I wish to thank all those who helped me. Without them, I could not have completed this project.

First of all I would like to offer my special thanks to my advisor Professor Alexandre Reily Rocha for his patient guidance and valuable support of this research work. Professor Alexandre Reily Rocha is the person who introduced me to exciting world of nanotechnology and I am deeply thankful to him for it.

I am also thankful to the Professor Gustavo Troiano Feliciano for his support in this project. Support given by Professor Gustavo has been a great help in performing my calculations.

The TransSim group members supported me during my stay in Brazil. These people not only helped me with my research, but also made my stay in Brazil very pleasant. Special thanks to Wudmir, who was helping me a lot in the beginning of my stay in Brazil, Jeconias, for his support in my studing and discussions of results of the calculations, Rodrigo, Pedro, Cesar and Claudio for the group discussions and coffee breaks.

This project would not happen without the financial support from the Brazilian funding agency CNPq and from the Universidade Federal do ABC. The calculations were performed in the HPC clusters at the Universidade de Sao Paulo and Universidade Federal do ABC.

I am thankful to Marcos who provided me unconditional support and encouragement during my Masters.

I am adressing my special thanks to my parents Ludmila and Alexandr Filatovi for their support throughout my studies. I know it was not easy for them to be away from me for such a long period of time, but I am very thankful that they always believed in me and in my undertakings.

Finally, I would like to thank my friend Ksenia, who was always supporting me even being many thousands of kilometers apart.

Contents

3 Results 47

List of Figures

Chapter 1

Introduction

1.1 Nanoscience and applications in biology

Nanoscience is arguably one of the most important research areas in modern science [10]. It is the field of research where at least one of the dimension of the material of interest is in the nanometer scale. These small systems draw our attention, because many of their properties might change when the size of the device becomes nanoscopic. Applications of nanoscopic systems have been shown in areas of diverse as microelectronics, coating, paints and biology [11]. The biological sciences might be one of the largest beneficiaries of nanoscience. The main reason is because biological molecules have dimensions commensurate with nanoscale materials.

Applications in biology come in a variety of forms. An example can be drug delivery [12]. The application of the new technologies in the drug delivery can benefit from the reduction of the amount of the used component, reducing the chances of missing in a dose and localyzing the delivery of potent compounds. In addition, there may be no need to use carrier, flavoring agents, binders and coating, that are used in the drugs nowadays [12]. The current nanotechnologies in the drug delivery include the nanoparticles of application. Nanoparticles are usually smaller than $100nm$, they are made by forming nanocrystals or drug-polymerer complexes or by creating nanoscale shells that entrap drug molecules [12]. They are often taken by cells where larger particles would be excluded. Small molecules can be loaded into nanoparticle that are not recognized by the immune system [13] and that can be targered to particular tissue types.

Another possibility is to use nanoparticles, in this case metallic, as the drug itself. It comes with the success of platinum anticancer drugs [14], that used as a component of nearly 50% of all cancer treatment. One is now looking for new metallic compounds which at the nanoscale result in lower side-effects [14].

An alternative for the nanoparticles can be nanoscopic tubular structures or nanotubes. While nanoparticles can be regarded or zero-dimensional nanostructures (in the sence that all dimensions are nanoscopic), nanotubes are called one-dimensional (since all but one dimension are nanoscopic). Nanotubes have some advantages relative to nanoparticles for biological applications [15]. First, they have larger inner volumes (relative to the dimension of the tube), that can be filled with desired chemical species. In addition, they have distinct inner and outer surfaces, which can be modified for biochemical functionalization. Nanotubes can have an open "mouth", which makes the inner surface accessible and the incorporation within the tubes particularly easy [11]. There exist several types of nanotubes, such as carbon nanotubes, peptide nanotubes and template-synthesized nanotubes [15].

Carbon nanotubes (tubular structures obtained by rolling a sheet of hexagonal carbon atoms) can also be used for development of biosensors. For example, Dai and colleagues [16] attached a reactive molecule to the sidewalls of single-walled nanotubes, that further can be used to attach proteins to their walls. Lieber's group used single-walled nanotubes as probe tips for atomic-force microscopy imaging of biomacromolecules [17].

The working principle of these biological sensors is the idea that the electronic current flowing through the device can be significantly altered by the presence of an external agent on the system's surface. This is in fact the working principle of many kinds of sensors. What makes nanoscopic sensors so exciting is the fact that their surface to volume ratio tends to be extremely high. This means that small changes to the surface will yield significant changes to the electronic properties and enhance sensitivity.

As we will show, a newly found material called graphene [5] can also be used as a kind of sensor with high selectivity, i.e., not only it is able to detect single molecules, it is also able to differentiate between slightly different molecules [18].

1.2 DNA. Why is DNA so important?

DNA, or deoxyribonucleic acid, is the hereditary material in humans and almost all other organisms. It encodes the genetic information of almost any living body [19]. The information in DNA is stored as a code made up of the sequential arrangement or combination of four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T) (Figure 1.1), which are connected to each other via a phosphate forming a strand of "beads". In a DNA molecule, each type of nucleobase on one strand bonds with just one specific type of nucleobase on the other strand (called complementary base pairing); adenine bonds only to thymine through two hydrogen bonds, and cytosine bonds only to guanine via three hydrogen bonds. A hydrogen bond is the electromagnetic attractive interaction between polar molecules in which hydrogen is bound to a highly electronegative atom, such as nitrogen, oxygen or fluorine [10]. In case of DNA hydrogen from one nucleobase bonds to nitrogen or oxigen from another base. Each base is attached to a sugar molecule and a phosphate molecule. Together, a base, sugar, and phosphate are called a nucleotide. Nucleotides are arranged in two long strands that form a spiral called a double helix. This arrangement of two nucleotides binding together across the strands is called a base pair (as shown in Figure 1.1). Furthermore the π interaction between adjacent base pairs favours the formation of helical structure [20].

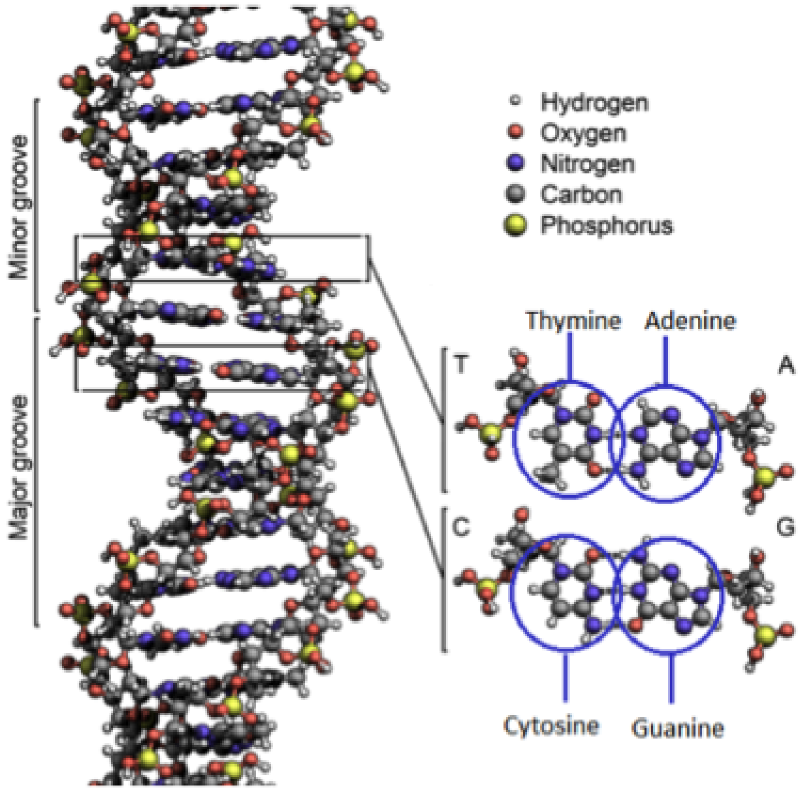


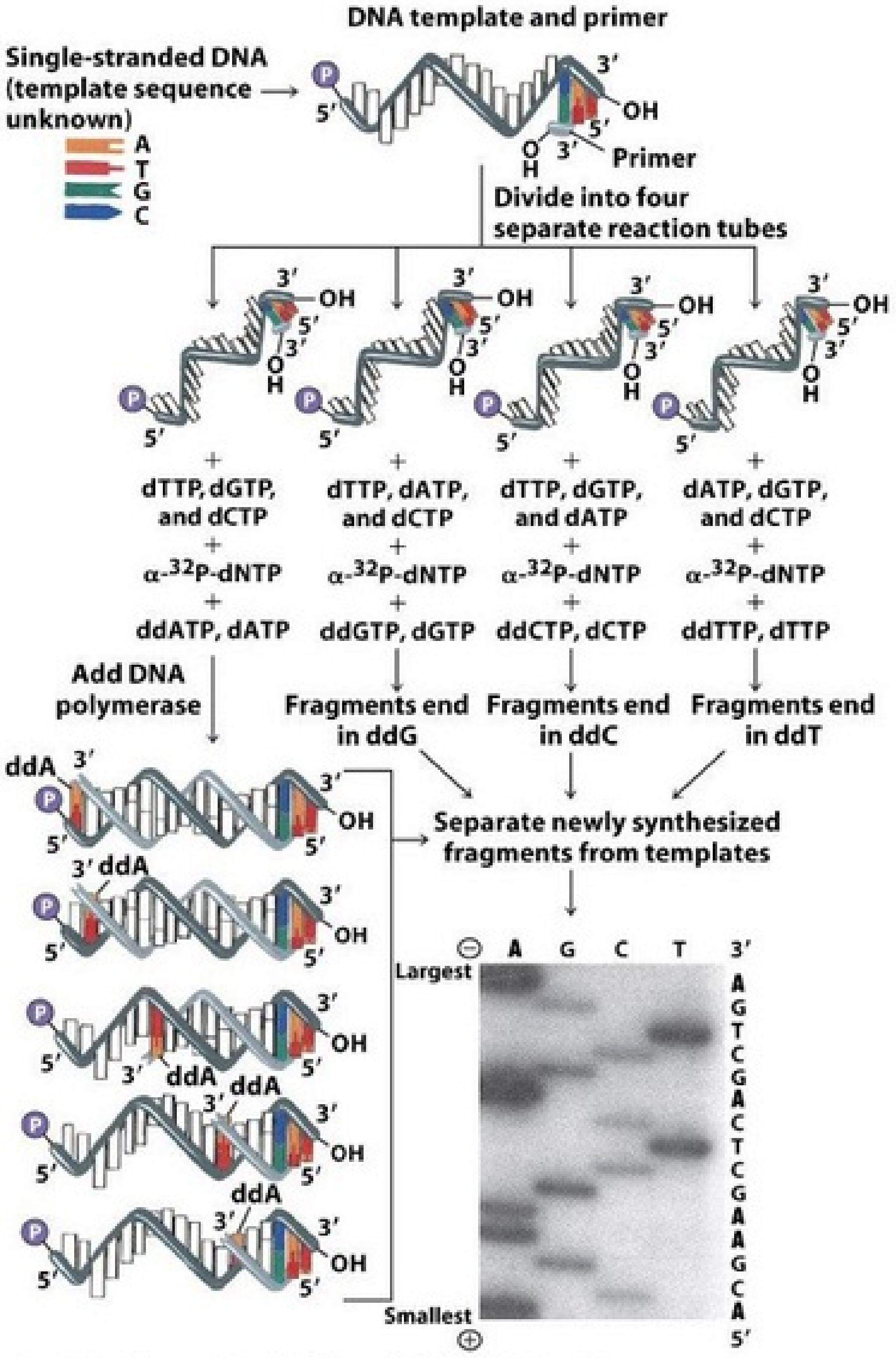
Figure 1.1: Representation of the chemical structure of the DNA. We can note the complimentary pairing between C and G, and A and T. We can also see the double stranded helical nature of the molecule.

As hydrogen bonds are not covalent, they can be broken and rejoined relatively easily. The two strands of DNA in a double helix can therefore be pulled apart like a zipper, either by a mechanical force or high temperature. As a result of this complementarity, all the information in the double-stranded sequence of a DNA helix is duplicated on each strand, which is vital in DNA replication. Indeed, this reversible and specific interaction between complementary base pairs is critical for all the functions of DNA in living organisms.

Human DNA consists of about 3 billion bases, and more than 99 percent of those bases are the same in all people [1]. The order, or sequence, of these bases determines the information available for building and maintaining an organism. Knowing the DNA sequence we can compare a plethora of individual human genomes, enabling us to locate sequences that cause hereditary diseases and to discover mutated ones. Besides, DNA sequencing can be used in the study of other areas of biology. For example, it can be used for comparing genomes between species to find common properties and functions [21]. In essence, once the sequence of DNA has been determined, an endless number of research areas open up.

1.3 Problem of DNA sequencing

The DNA sequencing procedure, that is used nowadays was named after Frederic Sanger [22]. In the current section I will give a brief overview of the Sanger sequencing method and describe it's main disadvantages.



Box 20-1 figure 2 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Figure 1.2: Main steps of the DNA sequencing, using Sanger method. Figure is extracted from [1].

First, the DNA is purified from the cell. Then very large pieces of DNA, such as all chromosomes or genomes, are cut into smaller pieces and stored in vectors, which are larger pieces of DNA with the ability to be reproduced when placed in a host cell such as bacteria. Bacterias containing a vector are placed in a culture medium, where they multiply a million-fold or more. Each time the bacteria divides, the DNA vector placed inside is also copied. Subsequently, a sequencing reaction is performed, which contains four steps. First, the double-stranded DNA is separated into single strands. Then a new DNA strand is made complementary to the template one. The termination step is the key to the sequencing reaction. Strand extension is halted by the incorporation of a dye-labeled nucleotide, which identifies the base at the position where strand extension stopped. When many reactions are performed together, each of the bases in a DNA strand can be identified. The newly synthesized DNA strands, each labeled with one of the four dyes, are now sorted by length using capillary electrophoresis. The results are then displayed on a computer. The main steps of the Sanger method are shown in the Figure 1.2. One then needs to reassemble the information by performing this procedure multiple times and using various computational tools to virtually reconect the different segments.

The cost of sequencing a whole genome, from scratch, has been dramatically reduced over the years mostly due to parallelizing the procedue (see Figure 1.3).

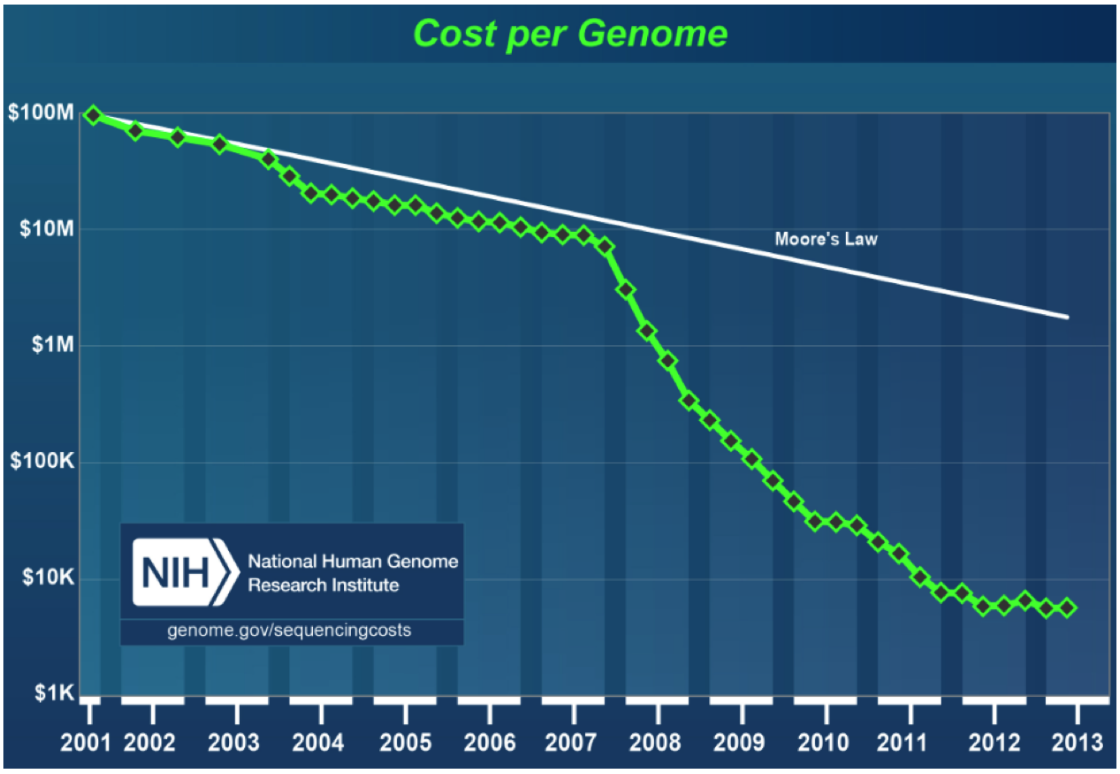


Figure 1.3: Cost of the DNA sequencing by National Human Genome Research Institute. Figure is extracted from [2].

However, the Sanger method has some disadvantages and it is now widely belived it will not overcome the US$1000,00 per genome barrier. First there are occasional artifacts, likely due to contaminant fragments. Also band pile-ups can occur due to loop formation under gel conditions, and is usually depicted as numerous bands in the same position, or very close together on the gel. These pile-ups appear at different positions when the priming is done in opposite directions. The method is not completely accurate and therefore repeating the sequencing more than once, for comparing the result is important. Finally, as stated before, to perform the DNA sequencing with the Sanger method one needs to reproduce millions of copies of a DNA strand. That is extremaly expensive and time consuming. Therefore there is a need for a new DNA sequencing method; one that is fast, reliable and possibly uses only one DNA molecule.

1.4 DNA sequencing using nanopores

The growing need for cheaper and faster genome sequencing resulted in the proposal of the technologies. The National Institutes of Health proposed the "US$1000 genome challenge" in 2004, which would be the target cost for sequencing one entire human genome [2]. The technologies inspired by this challenge are expected to be a new step in modern medicine. One of the promissing technologies is DNA sequencing using nanopore sensors. Nanopore is a tiny hole through a membrane or film, usually, big enough to allow a single DNA molecule to pass. It can be biological or solid state (Figure 1.4).

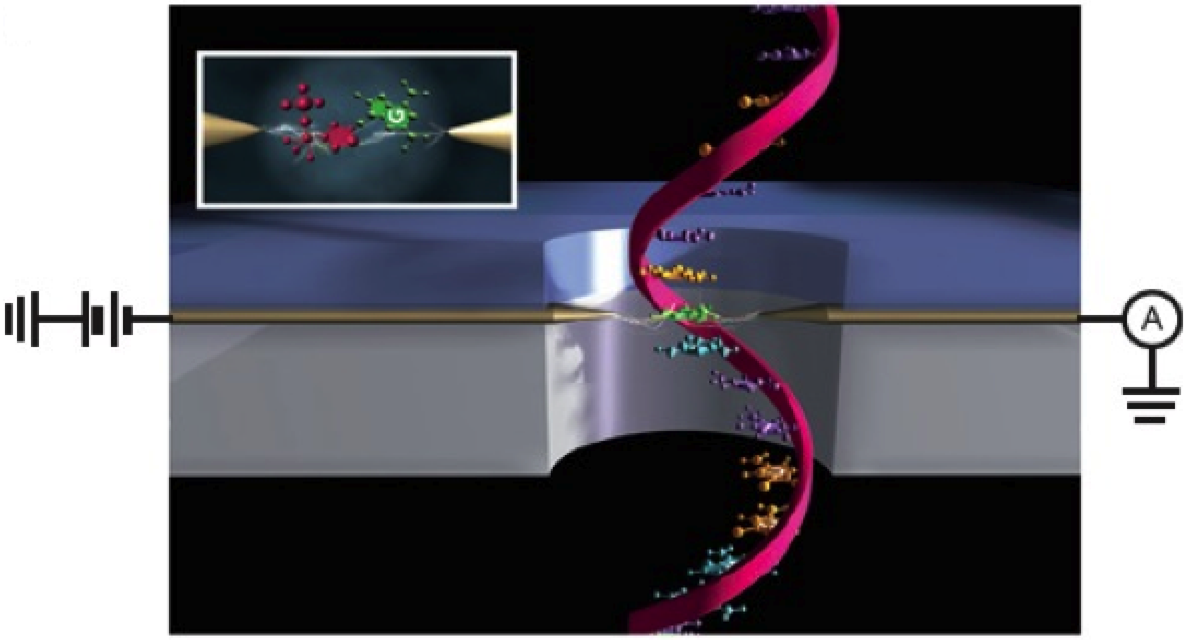


Figure 1.4: Solid state nanopore for DNA sequencing.The current passing along the membrane can be used to differentiate the bases. Figure is extracted from [3].

While passing through the nanopore, different nucleobases result in the changes to the charge transport. There are two types of current that can be measured: ionic current and transverse current. The idea of the ionic current measurements is that when each nucleobase passes through the pore it blocks the movement of ions present in the solution in a manner that reflectes both the polymer size and conformation [23]. So, in essence, different bases give rise to different proportions of blocked ions. By measuring this ionic current flowing through the nanopore one could differentiate the bases. Transverse current is the current that appears in the direction perpendicular to the backbone axis [24]. This idea is demonstrated in Figure 1.4. There the pore is not only a sieve for the DNA but also an electronic gap. The DNA nucleobase bridges that gap allowing electrons to flow. Either nanopore-based sensing technique is attractive for DNA sequencing applications becaurse it is a label-free, amplification-free, single-molecule approach that can be scaled for high-throughput DNA analysis [25]. Finally, the read length of DNA fragments is larger than the length of DNA fragments in the Sanger method. This is an important technological advantage. It excludes the necessity of assembling the separate reads to determine the human genome. Furthermore, large read lengths also provide information on insertions, deletions and gene duplicates that are sometimes difficult to determine from assembling short reads [7].

1.4.1 Biological nanopores

The analysis of the DNA sequencing using nanopore started from the usage of biological nanopores [26]. The idea of using nanopores for DNA sequencing was proposed more than 20 years ago [27]. In the basic experiment DNA, as a charged molecule, is driven through a biological nanopore by an electric field (Figure 1.5). The advantages of using biological nanopore for single-molecule DNA analysis are a follows:

- cells can produce large numbers of biological nanopores with atomic level resolution;
- X-ray crystallography has provided information about the nanopore structure at Ångsrtom resolution;
- physical and chemical properties of the biological nanopores can be tailored by changing the environment or the pore itself;
- remarkable heterogeneity is observed among biological nanopores in terms of size and composition [25].

Example of the biological nanopore include polymer α-haemolysin [28]. Kasianowicz *at al.* first demonstrated the electrophoretic transport through α-haemolysin [28]. The motivation to develop translocation of DNA through α-haemolysin was the ionic current signals, which were found to be different when comparing a homopolymer made up of 100

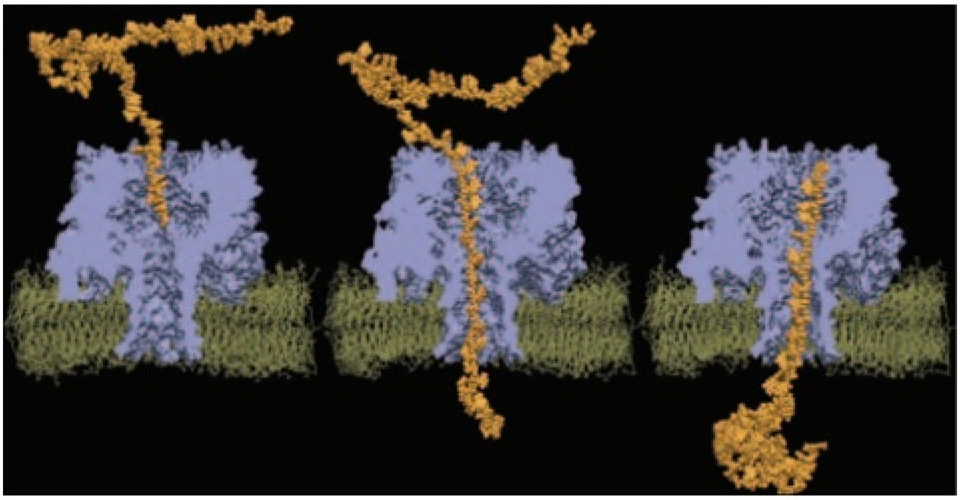


Figure 1.5: Snapshorts of a ssDNA molecule passing through biological nanopore. Figure is extracted from [4].

identical bases, with another one of different type. However, the thickness of the membrane does not allow one to identify each nucleobase, because at one given moment more than one nucleobase is located inside the pore. At the same time a great advantage of protein nanopores is that they can be chemically engineered with advanced molecular biology techniques, that can develope a variety of α-haemolysin [4]. At the February's Advances in Genome Biology and Technology conference in Marco Island, Florida company Oxford Nanopore announced that they would be able to launch a biological nanopore sequencing device by the end of the year 2012[1] [29].

1.4.2 Solid-state nanopores

Biological nanopores have some disadvantages such as fixed size and limited stability . Biological pores can become unstable if changes occur in external parameters such as pH, salt concentration, temperature, mechanical stress, and so on. An alternative approach to nanopore sequencing may involve solid-state nanopores. The advantages of the solid-state pores include very high stability, control of diameter and channel length, adjustable surface properties and the potential for integration into devices. These issues make them preferable in comparison with their biological counter-parts. A solid-state nanopore is typically a nanometer sized hole formed in a synthetic membrane (usually SiN or SiO_2). Solid-state nanopores are fast becoming an inexpensive alternative to biological nanopores [4].

---

[1] That turned out to be an oversell as a comercial device is yet to become available. Nevertheless, the company has a number of prototypes working in partner research institutes. Nature magazine has predicted that in 2014 a portable sequencing device using nanopore technology will become available.

1.4.3 Hybrid biological-solid-state nanopores

The main disadvantage of the solid-state nanopore technology is that it cannot reliably differentiate between analytes of approximately the same size. This can be overcome however, by attaching specific recognition sequences and receptors to the nanopore to create a hybrid structure. This structure has the potential to uniquely identify nucleotides in sequencing applications [25]. Another possibility is to reduce the length of the nanopore channel to try to enhance the signal-to-noise ratio. Another approach is try to design entirely new materials where sensitivity is enhanced.

1.5 Graphene: properties of suspended graphene membranes

A potential candidate for nanopore sequencing is graphene [30]. Graphene is a single layer of graphite - one atomic layer of carbon arranged in honeycomb lattice as shown in Figure 1.6(a). Since its experimental isolation on Si/SiO_2 substrates in 2004 and experimental observation of the anomalous "half-integer" quantum Hall effect, graphene has attracted a great deal of attention [31]. Graphene is the material of interest mostly due to its unique band structure, which consists of "cones" (around the Fermi level) located at the two inequivalent Brillouin zone corners, where the conduction and valence bands touch at a single point (Figure 1.6)

Due to this unique band structure graphene has such characteristics as high mobility, unparalleled thermal conductivity, 98% optical transparency and extraordinary current-carrying capacity [5]. Graphene, differently from traditional 2D electron systems based on $GaAs$ or Si, is truly 2D, having a thickness of one atomic layer. This cone-like band structure results in electrons behaving as massless relativistic particles [5]. In comparison with the Si/SiO_2 substrates, graphene's mobility increases by more than two orders of magnitude; it is stronger than steel but more supple than a Sanger wrap; it is transparent yet very conductive; it conducts heat better than any other material [5]. These characteristics make graphene the material of interest in many areas. In what follows I will highlight some of these properties - the ones that are of interest for this work.

1.5.1 Mechanical properties

Graphene has unique mechanical properties, it is the naturally thinnest elastic membrane. Suspended graphene sheets act as sensitive nano-electromechanical resonators and impermeable membranes with extremely high elastic constant and strength [5]. Most of the graphene membranes are not flat, but form periodic ripples, although the electronic properties of this system are robust enough to prevent them from deviating from the ideal graphene structure. This means that, in principle, one can obtain a reliable suspended

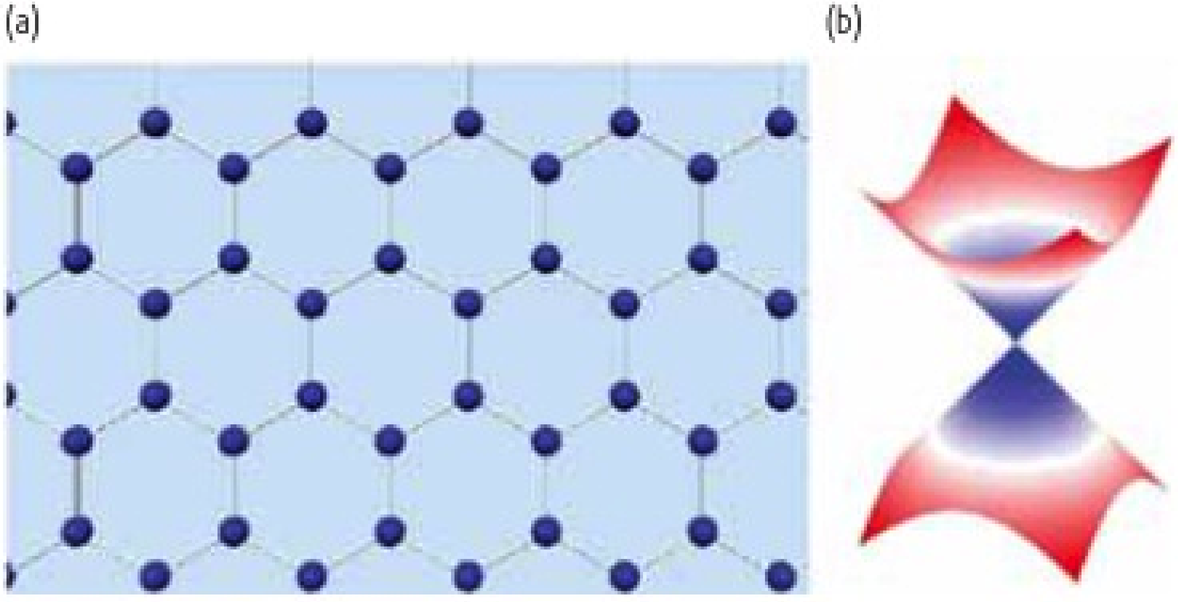


Figure 1.6: (a) Lattice and (b) band structure of graphene. Figure is extracted from [5].

sheet that can withstand the formation of defects (such as a small hole) without crumpling and without loosing most of its electronic properties.

1.5.2 Electrical properties

Graphene has a bipolar field effect transistor (FET) behavior, where a nearby electrostatic gate can be used to tune the density and polarity of charge carriers in graphene, thereby modulating its conductance. For devices on Si substrates with 300 nm of SiO_2 1 V in gate voltage V_g induces charge densities $n \sim 7.2 \times 10^{14}\, m^{-2}$. For typical suspended devices, the n/V_g ratio is $\sim 2 \times 10^{14} - 4 \times 10^{14}\, m^{-2}V^{-1}$. Figure 1.7(d) displays the standard curve of device conductivity σ vs. V_g, which is V - shaped, with a finite minimum of conductivity value σ_{min} at the charge neutrality point, (where graphene is undoped with nominally zero charge) [5].

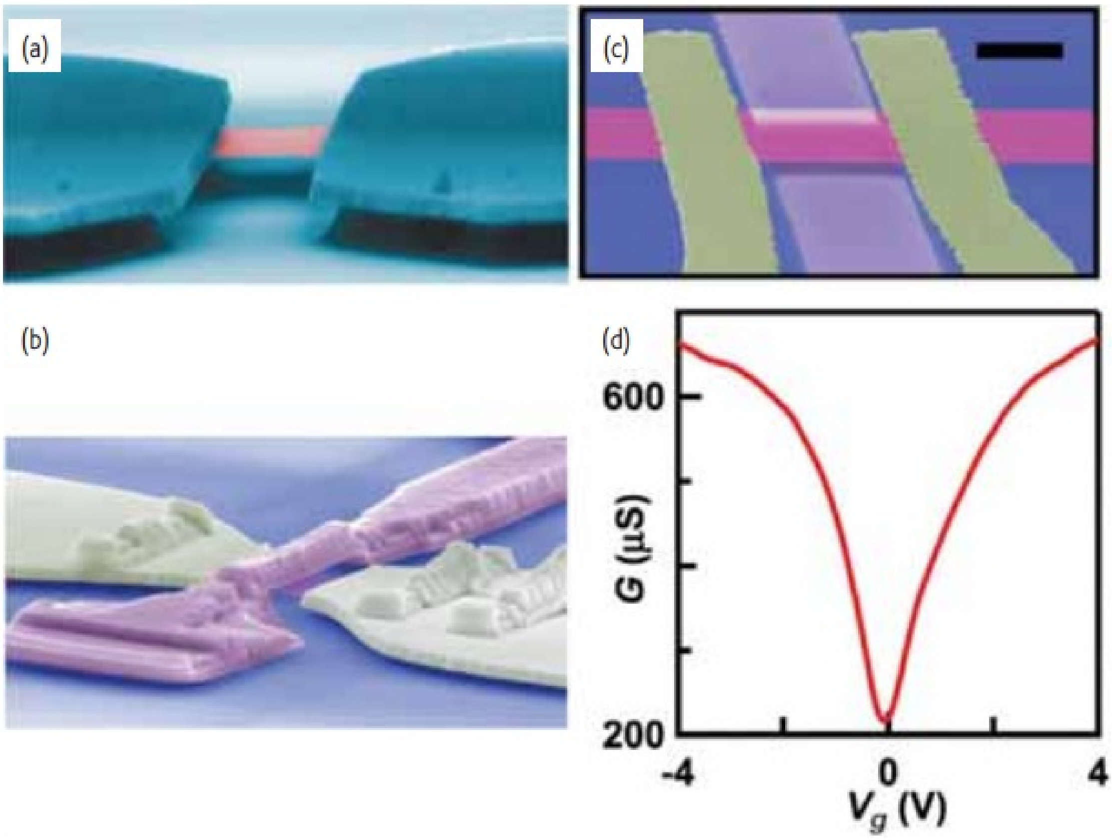


Figure 1.7: (a,b) SEM images of singly - and dual - gate suspended graphene fabricated via acid etching of SiO_2 layer. (c) SEM imade of a graphene device fabricated using the tithography-free technique. Scale bar: 2 μm. (d) Conductance vs gate voltage of a suspended bilayer graphene device, with mobility $\sim$ 40000 cm^2/Vs. Figure is extracted from [5].

1.6 Graphene nanopores for DNA sequencing

Common membrane materials currently used for nanopore device fabrication are insulators such as silicon nitride SiN, aluminum oxide, and silicon oxide SiO_2. In my project I used graphene nanopores. As was shown in the previous sections, graphene is a thin, flexible material with good electronic conductivity and robust mechanical properties. Graphene nanopores are structurally robust and their shape is stable over time. Use of graphene as a nanopore membrane material could permit sensing and control of the electric potential directly at the nanopore. The first proposal of this kind of sensors was done by Postma [18], where atomically thin graphene nanoelectrode have been considered for DNA sequencing based on DNA's transverse conductance [18]. Later there were simulations on 1D graphene nanoribbons containing a pore [25]. In all cases, the basic idea is that the atomically thin membrane only interacts with one base at a time and little or no tunneling occurs between graphene and base which are above or bellow the one in the pore.

Figure 1.8 in [6] presents a cartoon of a typical graphene nanopore device. In the prototypical device, individual ssDNA polymer are electrophoretically driven through a pore and the resulting current is measured. It is expected that ionic current varies when each nucleobase passes through the pore due to the different sizes of the nucleobases and different interaction of this bases with the pore. Another possibility would be to measure the transverse current.

Graphene nanopores have been fabricated and it has been shown that one can detect DNA passing through, albeit no bases differention has been achived (See Figure 1.9) [7]. The following two requirements for device performance must be met: (1) there must be a sufficient difference in the conductance spectra $G(\nu)$ of individual nucleobases to allow them to be distinguished from one another, and (2) the measured conductance for a single nucleobase must be somewhat independent of its orientation as it passes through the pore (or at least variations in the conductance must still allow for differentiation). Thus far, experiments on graphene nanopores have been unable to differentiate between bases. One has only been able to determine whether a strand was present within the pore.

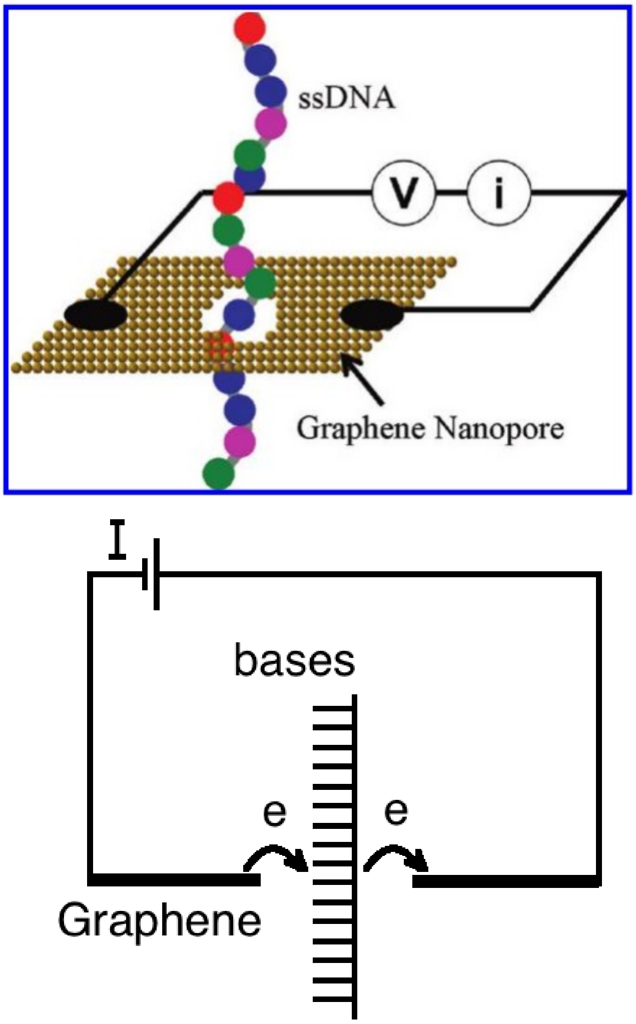


Figure 1.8: Nanopore sensor consists of a graphene nanopore embedded in SiO_2. Single-stranded DNA translocate through the nanopore. Figure is extracted from [6].

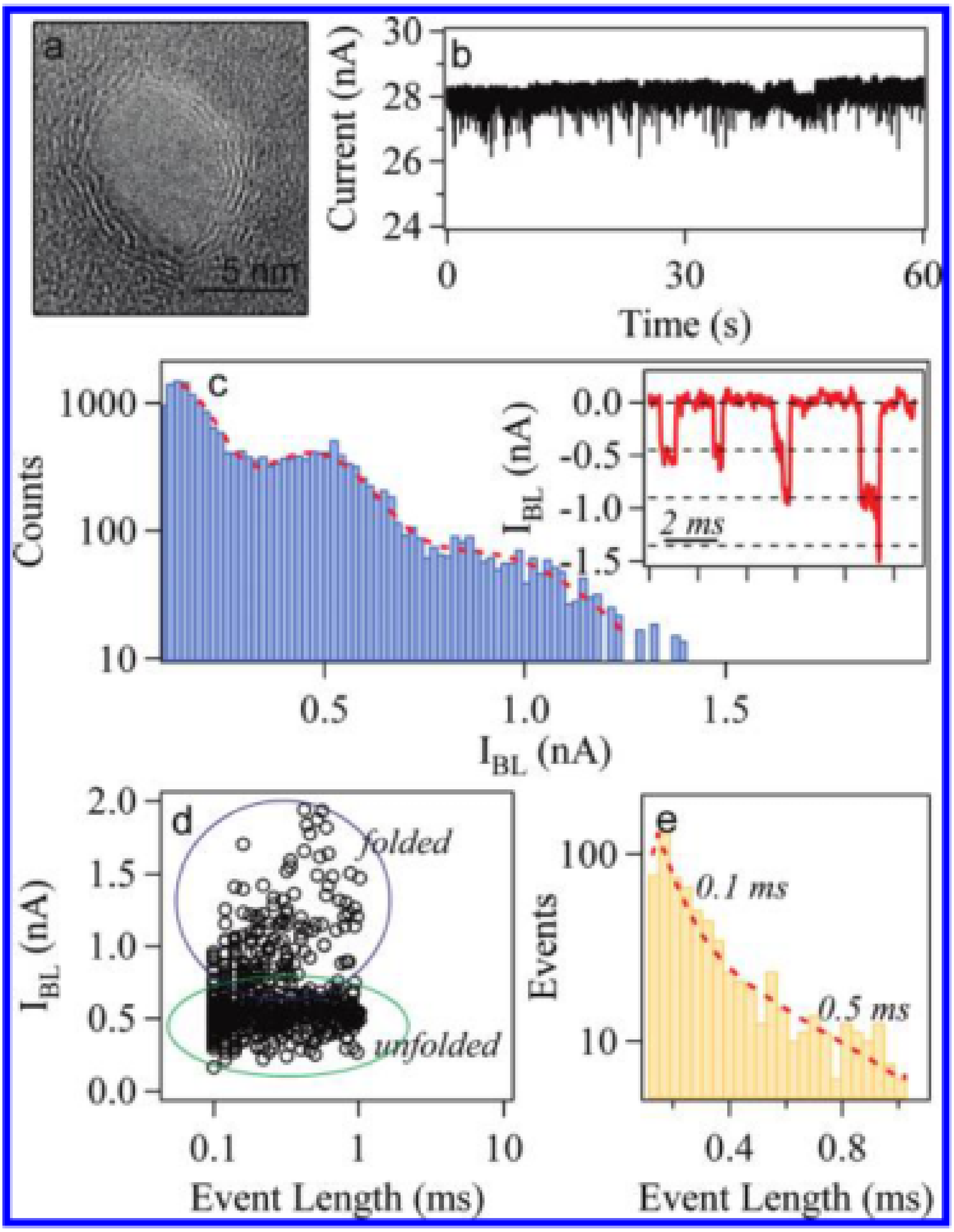


Figure 1.9: (a) TEM image of an $8nm$ graphene nanopore. (b) Time trace of events for nanopore device. (c) Histogram of blocked current. (d) Scatter plot of event length. (e) Histogram of event lengthes. Figure is extracted from [7].

1.7 The challenges for theory and simulations

One can use theoretical models to fully understand experimental results and further still, make usefull predictions about how to optimize new experimental setups. As we are dealing with the nanoscale, the atomic positions are very important. As we will later see, although the nucleotides are similar, the different atomic arrangements give rise to significantly different electronic distributions (and transport properties). With that in mind theoretical tools are of extreme importance. Here we want to calculate the electronic transport. Theory for the electronic transport for calculating transverse current deals with the electrons, so we need to use quantum mechanics. At the same time DNA must be surrounded with water and counter ions (Na and Cl). That signifficantly incrises the number of atoms in the system. Finally, as we are analyzing DNA translocation through the nanopore, we have to deal with the dynamics of the system.

The challenge, thus is to calculate the electronic transport in a system containing tens of thousands of atoms in a dynamical form - in principle - fully quantum mechanically in a realistic fashion, i.e. with atoms positioned in their correct places. This means that model calculations are not an option. We must do this using computer simulation which gives accurate predictions about the electronic structure.

1.8 Objectives

In my project I examine the possibility of performing DNA sequencing using graphene nanopores. Four different graphene nanopore are analyzed: a graphene nanopore passivated with H atoms, a graphene nanopore passivated with N atoms. In both cases we consider nanopores with $10\AA$ and $20\AA$ in diameter correspondingly (Figure 1.11). I analyze four different DNA nucleobases separately in each of this pores and calculate the electronic transport of the system. For this purposes I use a combination of density functional theory coupled to classical molecular dynamics (QM/MM method). The effects of the solvent and of counter-ions are included in the simulation via the classical part of the simulation whereas the rest is simulated quantum mechanically (Figure 1.10). From different snapshots of the simulation I use the Non-equilibrium Green's function methodology to calculate the electronic transport properties of this system.

Thus this text is orginized as follows: chapter 2 describes the methodology, that we used in our simulations, chapter 3 describes results of the simulations and chapter 4 gives conclusions and future perspectives.

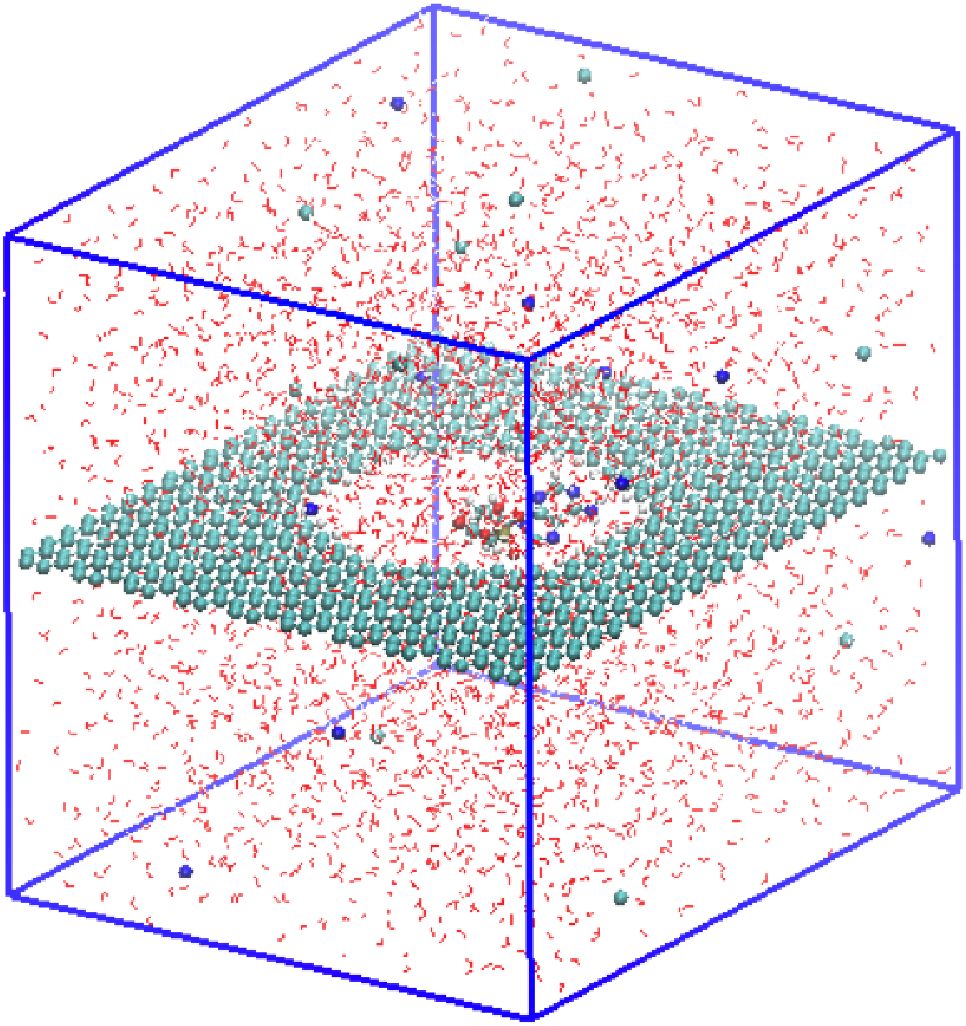


Figure 1.10: Typical system we would like to consider in our simulations of electronic transport.

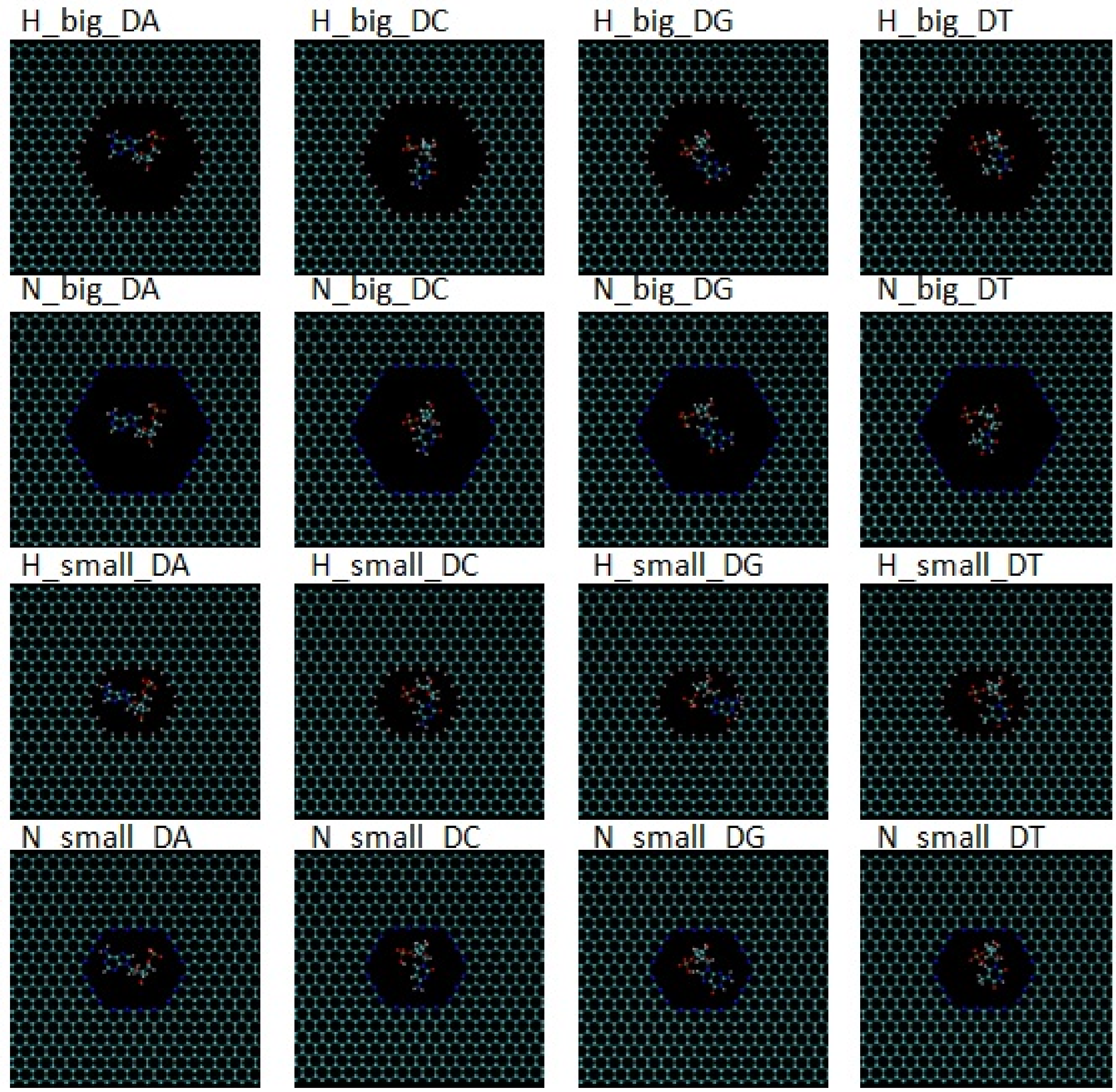


Figure 1.11: Analyzed graphene nanopore with different DNA nucleobases: Adenine (DA), Thymine (DT), Guanine (DG) and Cytosine (DC).

Chapter 2

Methodology

As described in the previous chapter we want to simulate the translocation of a DNA strand through a nanopore. During that translocation process we would like to calculate the electronic transport properties. In order to do that a number of pre-requisites must be met:

- consider the positions of the atoms;
- treat water and counter ions;
- treat the electrons quantum mechanically;
- treat the dynamical effects.

In principle, one would need to treat the system as a collection of electrons and nuclei that interact with each other via Coulomb interactions. This is a quantum mechanical problem that requires one to find the many-body wave function for the electrons and nuclei $\psi(\overrightarrow{r_i}, \overrightarrow{R_j})$, wich is the solution of a Hamiltonian of the form (without spin):

$$\hat{H} = \hat{T} + \hat{V} \tag{2.1}$$

where $\hat{T}$ is the kinetic energy operator and $\hat{V}$ is the potential energy operator. The solution follows the time-dependent Schrödinger Equation:

$$H|\psi\rangle = -i\hbar \frac{\partial |\psi\rangle}{\partial t} [32] \tag{2.2}$$

The solution of this problem is extremely complicated for a large number of particles. The first approximation that can be made to simplify this is the so-called Born-Oppenheimer approximation. The general idea is that given their significantly different masses - electrons and nuclei evolve on different time frames. It is possible to assume that electrons can almost instantaneously accomodate to a specific configuration of nuclei and one need only worry about the time evolution of the nuclei.

Still, this problem requires the solution of the time evolution equation for the nuclei as quantum objects with the effects of the electrons included. In most cases, however, one can consider them as classical objects that ”feel” a potential given by the other nuclei and the electrons. In that case, the evolution of the system follows Newton’s laws where for each new configuration of the nuclei a new electron density can be obtained and therefore a new potential.

We have now two paths to follow. In the first one we calculate the electrons quantum mechanically in what is known as *ab-initio* molecular dynamics. Another possibility is to go one step further and parametrize the effects of the electrons on the different nuclei and on its neighbors. In general, given an arrangement of atoms, an equation known as a force field describes the potential felt by the *i-th* atom due to the different interactions with the remaining atoms in the system.

We have thus two competing approaches. One is fully classical and much faster. It allows one to easily simulate the dynamics of thousands or, perhaps, millions of atoms. The other one is more accurate and includes the electrons explicitly, but is much more computationally demanding. In this project we used both methods. Initially we used them separately to describe different aspects of the problem. At the end I will describe a methodology designed to get the best of both worlds.

2.1 Molecular dynamics

Studying the motion of atoms and molecules is important to the understanding of many important concepts in physics, chemistry and biology. The size and complexity of a typical molecular system, together with the timescales that must be reached, require the use of classical molecular dynamics for the nuclear degrees of freedom. In molecular dynamics simulations Newton’s equations of motion are solved numerically to obtain a trajectory of the dynamics of a collection of atoms over a period of time [9].

2.1.1 Molecular mechanics force field

A force field is a mathematical expression describing the dependence of the energy of a system on the coordinates of its particles. It consists of an analytical form of the interatomic potential energy, $U(r_1, r_2, ..., r_N)$, and a set of parameters entering into this form. The parameters are typically obtained either from *ab initio* or semi-emphirical quantum mechanical calculations; or by fitting to experimental data such as neutron, X-ray and electron difraction, Raman and neutron spectroscopy, etc. Molecules are simply defined as a set of atoms that are held together by simple elastic (harmonic) forces and the force field replaces the true potential with a simplified model hopefully valid in the region being simulated [33].

Atoms basically interact with each other through van der Waals forces and electrostatic forces. When they are covalently bonded to others, a strong force holds them together as a stable chemical group. A widely used mathematical model for the potential energy of a molecular system consists of six types of interactions:

$$U = V_{LJ} + V_{EL} + V_{BS} + V_{AB} + V_{PT} + V_{IT} \tag{2.3}$$

The first type, V_{LJ}, is the Lennard-Jones potential that has an attractive part representing the van der Waals energy and a repulsive part representing the Pauli repulsion:

$$V_{LJ} = \frac{1}{2} \sum_{i,j,i \neq j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^{6} \right] \tag{2.4}$$

where R_{ij} is the distance between the i-th and j-th atom, ε_{ij} is called the van der Waals dissociation energy, and σ_{ij} is called the collision diameter. The dissociation energy is equal to the amount of energy needed to pull a pair of atoms in the strongest van der Waals binding state apart. The collision diameter is approximately the distance at which a pair of atoms bounces off from each other in a normal, non-reacting condensed state.

V_{EL} is the electrostatic potential energy according to Coulomb's Law:

$$V_{EL} = \frac{1}{2} \sum_{i,j,i \neq j} \frac{q_i q_j}{R_{ij}} \tag{2.5}$$

where q_i is the effective charge of the i-th atom. Compared with van der Waals potential, the electrostatic potential is a stronger, more long-range interaction.

V_{BS} is the bond-stretching energy standing for the elastic interaction between a pair of atoms connected by a covalent bond (see Figure 2.1(A)):

$$V_{BS} = \frac{1}{2} \sum_{m \in bonds} k_m^l (l_m - l_m^0)^2 \tag{2.6}$$

where l_m is the distance between the two atoms of the m-th bond, l_m^0 is the equilibrium bond length, k_m^l is the bond strength. The values of l_m^0 can be obtained from the X-ray diffraction experiments, values of k_m - from Raman spectra.

V_{AB} the angle-bending energy standing for the interaction among three covalently-bonded atoms that form a stable angle (see Figure 2.1(B)):

$$V_{AB} = \frac{1}{2} \sum_{m \in angles} k_m^\theta (\theta_m - \theta_m^0)^2 \tag{2.7}$$

where θ_m is the m-th angle between the two adjacent bonds that share a common atom, θ_m^0 is the equilibrium bond angle, and k_m^θ is the strength. The values of the θ_m^0 and k_m^θ can also be obtained from experiments.

V_{PT} and V_{IT} are the proper and improper torsional energies standing for the interactions among four covalently-bonded atoms that form a stable proper and improper dihedral angle (see Figure 2.1(C-D)):

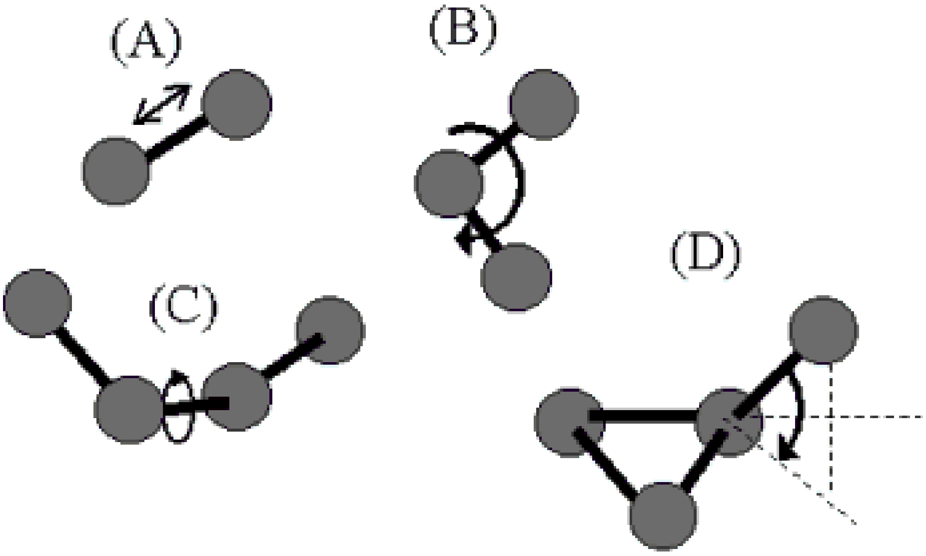


Figure 2.1: Schematic illustration of the interactions that model covalent bonding: A) Bond-stretching force, B) Angle-bending force, C) Proper torsional force, D) Improper tortional force [8].

$$V_{PT} = \frac{1}{2} \sum_{m \in torsions} V_m [1 + \cos(n_m \omega_m - \gamma_m)] \tag{2.8}$$

$$V_{IT} = \frac{1}{2} \sum_{m \in torsions} k_m^{\xi} (\xi_m - \xi_m^0)^2 \tag{2.9}$$

where ω_m is the m-th dihedral angle between the two adjacent angles that share a common bond, n_m is the periodicity factor which determines the number of equilibrium dihedral angles in a 360^0 rotation, γ_m is the phase shift, V_m is the amplitude, ξ_m is the m-th improper dihedral angle among four atoms that are not bonded successively to one another, ξ_m^0 is the equilibrium improper dihedral angle, and k_m^{ξ} is the strength. The tortional parameters are usually derived from the *ab initio* calculations [34].

The sum of equations 2.6, 2.7, 2.8 and 2.9 are called the *bonded interactions*, which maintain the bond lengths, the bond angles and the dihedral angles so that chemical groups will remain stable in an MD simulation[2] [36].

---

[2] Another many-body classical potential, which is not used in this work, is the Brenner empirical potential [35]. The Brenner potential for every atom in the system depends on the nearest neighbors of this atom. The total potential energy of the system interacting via the Brenner potential is:

$$U_{tot} = \frac{1}{2} \sum_{i=1}^{N} \sum_{\substack{i=1 \\ j \neq i}}^{N} U_{ij} = \frac{1}{2} \sum_{i=1} \sum_{j \neq i} f_{cut}(r_{ij}) [U^{(R)}(r_{ij}) - B_{ij} U^{(A)}(r_{ij})] \tag{2.10}$$

where $f_{cut}(r_{ij})$ is the cut-off function, which limits the interaction of an atom to its nearest neighbors, R_1 and R_2 are parameters, that determine the range of the potential, fuctions $U^{(R)}(r_{ij})$ and $U^{(A)}(r_{ij})$ are the repulsive and the attractive energy terms of the potential, respectively. This way it is possible to simulate bond breaking and thus reactions.

The first force fields appeared in the 1960's. Their goal was to predict molecular structures, vibrational spectra and enthalpies of isolated molecules. Allinger's group developed the potentials [37]: MM2, MM3, MM4 which were aimed at treating small organic molecules.

Since then, more widly applicable force field appeared, such as Dreiding and Universal (UFF) force fields, that contain all parameters for all the atoms in the periodic table [33]. Other popular force fields are CHARMM [38], AMBER [39], GROMOS [14], OPLS [40] and COMPASS [41]. All of them are quite general, but the first three are often employed in simulations of biomolecules, while OPLS and COMPASS were originally developed to simulate condensed matter.

In our project we use AMBER force-field:

$$U = \frac{1}{2}\sum_{i,j,i\neq j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{R_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}}\right)^{6}\right] + \frac{1}{2}\sum_{i,j,i\neq j}\frac{q_i q_j}{R_{ij}} + \frac{1}{2}\sum_{m\in bonds} k_m^l (l_m - l_m^0)^2 + + \frac{1}{2}\sum_{m\in angles} k_m^\theta(\theta_m - \theta_m^0)^2 + \frac{1}{2}\sum_{m\in torsions} V_m[1+\cos(n_m\omega_m - \gamma_m)] \quad (2.11)$$

2.1.2 Molecular dynamics

Once a force field has been chosen and the corresponding parameters obtained the positions, velocity and acceleration of each atom are calculated using a numerical method to solve Newton's equations of motion. A set of initial conditions is needed (positions and velocities of each particle), and a good model to represent the forces acting between the particles (either from electronic structure calculations or using the empirical force fields presented in the previous sections), and to define the boundary conditions to be employed. After this the classical equation of motion need to be solved:

$$m_i \frac{d^2\overrightarrow{r_i}}{dt^2} = f_i = -\frac{d}{dr}U(\overrightarrow{r_1}, \overrightarrow{r_2}, ..., \overrightarrow{r_N}), \quad (2.12)$$

where $U(R_1, R_2, ..., R_N)$ is the potential energy depending on the coordinates of the N particles. This is a system of N coupled second order non linear differential equations that cannot be solved exactly, it has to be integrated numerically step by step using an appropriate integration algorithm such as Verlet algorithm, Leap-frog algorithm, Velocity Verlet and Beemans algorithm[3].

Initial conditions

To define the initial positions and velocities of each particle of the system we need to analize the system. In the case of a crystal the positions will be available in the form of a

---

[3] There are other methods to evolve a system of quantum particles such as Carr Parrinello MD [42] and Path integral MD [43]. Since those are not going to be used they will not be discussed here.

crystallographic file. Typically, the initial velocity of each particle is attributed randomly and then they are adjusted in order to zero the angular momentum and the center of mass velocity of the total system.

Boundary conditions

In case we want to calculate bulk properties of the system, we have to set up boundary conditions. To exclude the surface effects we need to use periodic boundary conditions (PBC). Therefore the simulation box is surrounded by an infinite number of replicas of itself. This is not a problem when simulating a crystal, but in the case of disordered systems an artificial periodicity can be introduced. If the size of the box is large, the effects are not important, but we need to be careful while considering any properties that depends on long-range correlations. PBC are used together with the minimum image convention (only the interactions with the nearest image are considered). But it involves the time to compute the forces to scale as N^2. The potential is not constant on the surface of a cube around a given particle. Therefore the usual approach consists in employing a spherical truncation, such as the only interaction between particles separated by a distance smaller that R_c are taken into account:

$$U_{trunc}(r) = \begin{cases} U^{LJ}(r), r \leq R_c, \\ 0, \qquad\quad r > R_c. \end{cases} \tag{2.13}$$

The cutoff radius R_c cannot be larger than $L/2$.

Integration algorithms

To solve Eq.2.12 numerically we need to discretize the trajectory and use an integrator to advance over small steps:

$$r_i(t_0) \to r_i(t_0 + \Delta t) \to r_i(t_0 + 2\Delta t) \to ...r_i(t_0 + n\Delta t). \tag{2.14}$$

The desirable properties of such an integrator are:

- minimal need to compute the forces,
- good stability if large time steps Δt are used,
- good accuracy,
- good conservation of energy and momentum,
- time-reversibility,
- conservation of the phase space volume [44, 45, 33].

The solution to these problem was proposed by Verlet. If one sums the Taylor expansion of the equations of motion 2.15 by $+\Delta t$ and $-\Delta t$, the terms Δt, Δt^3, etc. cancel and Eq.2.16 is obtained:

$$r_i(t_0 + \Delta t) = -r_i(t_0) + \frac{dr_i(t_0)}{dt}\Delta t + \frac{d^2 r_i(t_0)}{dt^2}\Delta t^2 O(\Delta t^3) \tag{2.15}$$

$$r_i(t_0 + \Delta t) = -r_i(t_0 - \Delta t) + 2r_i(t_0) + a_i(t_0)\Delta t^2 + O(\Delta t^4) \tag{2.16}$$

The velocities are not used in the algorithm, but they can be obtained as:

$$v_i(t_0) = \frac{1}{2\Delta t}[r_i(t_0 + \Delta t) - r_i(t_0 - \Delta t)] \tag{2.17}$$

The other type of algorithms are of the predictor-corrector type. The iteration starts with a prediction of the new positions and all the required derivatives. Then the correct forces at $t+\Delta t$ are evaluated using the new positions. The difference between the predicted and corrected forces is used to correct the predicted values. For small time steps this algorithm is more accurate that the Verlet integrator.

Thermodynamical ensembles

The integration of Eq.2.12 keeps constant the number of particles N, the volume of the cell V, and the total energy of the system E. This implies that the generated trajectory will be in the NVE ensemble (constant number of particles N, volume of the cell V and energy E). However, the integration errors, and force fluctuations may cause slow drifts in the total energy. It is therefore desirable to have some ways to control the temperature. Furthermore, we are interested to set a proper assemble experimentally.

Constant temperature simulations: The temperature of the system is directly connected with the kinetic energy, so T can be changed by modifying the velocities of the atoms. The simplest scheme then to control the temperature is the "velocity scaling". It consists of scaling all the velocities by a factor $\sqrt{T_B/T(t)}$, where T_B is the desired temperature and $T(t)$ is the instantaneous temperature of the system before the scaling.

This method, however, doesn't allow one to sample the true NVT ensemble. A more physical way to control the temperature is to use a thermostat as proposed by Berendsen [46], consisting in a weak coupling of the system to a heat bath. This is done by modifying the equation of motion using:

$$a_i = \frac{F_i}{m_i} + \frac{1}{2\tau_T}\left(\frac{T_B}{T(t)} - 1\right)v_i \tag{2.18}$$

The additional term acts as a frictional force and the coupling constant τ_T determines the strength of the coupling. This method allows us to rapidly achieve the desired temperature. However, it does not allow to sample the correct canonical ensemble and it can affect the system dynamics when small values of τ_T are used.

The most rigourous method to control the temperature is to use the Nose-Hoover thermostat. In this case the heat bath becomes an integral part of the system by adding the artificial variable with an associated effective mass [47]. This is the method we use in our simulations.

Constant pressure simulations: The pressure can be controlled in the same way as the temperature. In this case the instantaneous virial takes the role of the temperature and the pressure is controlled by scaling the system volume. One of the common used barostats is the Parrinello-Rahman barostat. In this approach Parrinello and Rahman suggested that the simulation box can change its volume [48].

The theory described here has been implemented in a variety of computational codes available freely for scientific use. Our simulations were performed on Gromacs [49]. A full description of the protocol used will be given in the next chapter.

2.2 Density-functional theory (DFT)

If one wishes to calculate the electronic wave function of a many-electron system, the Hamiltonian

$$H = -\frac{\hbar^2}{2m}\sum_i \nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\overrightarrow{r_i} - \overrightarrow{r_j}|} - \sum_{i,j}\frac{Z_j e^2}{|\overrightarrow{r_i} - \overrightarrow{R_j}|} \tag{2.19}$$

(where $-\frac{\hbar^2}{2m}\sum_i \nabla_i^2$ is a kinetic energy operator, $\frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\overrightarrow{r_i}-\overrightarrow{r_j}|}$ describes the electron-electron interaction, $\sum_{i,j}\frac{Z_j e^2}{|\overrightarrow{r_i}-\overrightarrow{R_j}|}$ describes the nuclei-electron interaction) needs, in principle, to be solved.

In 1927-1928 L.H. Thomas and E. Fermi suggested that the full electronic density is the variable of interest in a many-body problem, and derived a differential equation for the density. Unfortunately, the approximation they used for the kinetic energy was unable to sustain a bound state. Newertheless, this approximation gave a basis for the later development of Density Functional Theory (DFT), that nowadays is one of the most popular quantum mechanical approaches. Determining the N-body wave function $|\Psi\rangle = |\Psi(\overrightarrow{r_1}, \overrightarrow{r_2}, \overrightarrow{r_3}...\overrightarrow{r_N})\rangle$ by exactly solving Eq.2.19 is a hard job even for a very small number of electrons.

2.2.1 Thomas-Fermi theory

Thomas (1927) [50], and independently Fermi (1928) [51], gave a prescription for calculating the energy of an electronic system exclusively in terms of the electronic density. In their original work Fermi proposed an expression for the total electronic energy, where the kinetic, exchange, and correlation contributions were taken from the homogeneous electron gas, for which good approximations were known. The idea was to construct the same quantities for the homogeneous system as

$$E_\alpha[n] = \int n(r)\varepsilon_\alpha[n(r)]dr \tag{2.20}$$

where $\varepsilon_\alpha[n(r)]$ is the energy density of contribution α (kinetic, exchange, and correlation), calculated locally at the value assumed by the density at every point in space. This was the first time that the *local density approximation* (LDA) was proposed. For a homogeneous electron gas the electronic density is related to the Fermi energy

$$n = \frac{1}{3\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E_F^{3/2}, \tag{2.21}$$

and the kinetic energy is $T = 3nE_F/5$, so that the kinetic energy density is

$$t[n] = \frac{3}{5}\frac{\hbar^2}{2m}(3\pi^2)^{2/3}n^{2/3}. \tag{2.22}$$

Therefore, the LDA kinetic energy is

$$T_{TF} = C_k \int n(r)^{5/3}dr, \tag{2.23}$$

with $C_k = 3(3\pi^2)^{2/3}/10$. Exchange can be introduced considering Slater's expression for the homogeneous electron gas

$$E_X[n] = -C_X \int n(r)^{4/3}dn, \tag{2.24}$$

with $C_k = 3(3/\pi)^{1/3}/4$. Correlation can also be easily included by using any local approximation to the homogeneous electron gas. The one proposed by Winger

$$E_C[n] = -0.056 \int \frac{n(r)^{4/3}}{0.079 + n(r)^{1/3}}dr, \tag{2.25}$$

By replacing the above approximations into the general expression for the energy of an inhomogeneous electronic system, we obtain the $Thomas - Fermi - Dirac$ (TFD) energy expression:

$$E_{TFD}[n] = C_k \int n(r)^{5/3}dr + \int n(r)\vartheta_{ext}(r)dr + \frac{1}{2}\int\int \frac{n(r)n(r')}{|r-r'|}drdr' - C_X \int n(r)^{4/3}dn + E_C[n]. \tag{2.26}$$

The idea of the Thomas-Fermi approach was to write the energy of the many-body system exclusively in terms of the electronic density. Nevertheless, at the time it was seen as an argument which didn't necessarily need to hold everytime. In fact, it seems counter intuitive that one could obtain the properties of a QM system by considering only the total charge density. Later it was prooved that this is indeed the case. In 1964 Hohenberg and Kohn formulated and proved a theorem that showed that the initial idea of Thomas and Fermi approach is valid.

2.2.2 The Hohenberg-Kohn theorem

At the heart of DFT is the Hohenberg-Kohn (HK) theorem. This theorem states that for ground states the calculation of the particle density (Eq.2.27) from the wave function can be inverted: given a ground-state density $n_0(r)$ it is possible to calculate the corresponding ground-state wave function $\Psi_0(r_1, r_2, ..., r_N)$, so Ψ_0 is the functional of n_0. This means, that all ground-state observables are functionals of n_0 too.

$$n(r) = N \int d^3r_2 \int d^3r_3 ... \int d^3r_N \Psi^*(r_1, r_2, ..., r_N)\Psi(r_1, r_2, ..., r_N) \tag{2.27}$$

The ground-state wave function Ψ_0 must not only reproduce the ground-state density, but also minimize the energy

$$E_{\vartheta,0} = \min_{\Psi \rightarrow n_o} \langle \Psi | \hat{T} + \hat{U} + \hat{V} | \Psi \rangle \tag{2.28}$$

where $E_{\vartheta,0}$ denotes the ground-state energy in potential $\vartheta(r)$

First Hohenberg-Kohn theorem

The nondegenerate ground-state wave-function is a unique functional of the GS density:

$$\Psi_0(r_1, r_2, ..., r_N) = \Psi[n_0(r)]. \tag{2.29}$$

As a consequence, the ground-state expectation value of any observable $\hat{O}$ is a functional of $n_0(r)$, too:

$$O_0 = O[n_0] = \langle \Psi[n_0] | \hat{O} | \Psi[n_0] \rangle. \tag{2.30}$$

For example energy:

$$E_{\vartheta,0} = E_{\vartheta}[n_0] = \langle \Psi[n_0] | \hat{H} | \Psi[n_0] \rangle. \tag{2.31}$$

Second Hohenberg-Kohn theorem

If energy $E[n]$ for fixed ϑ_{ext} is evaluated for a density that is not the ground-state density of the system in potential ϑ_{ext}, one never finds a result below the true ground-state energy:

$$E_{\vartheta}[n_0] \leq E_{\vartheta}[n'] \tag{2.32}$$

In practical terms what the two Hohenberg-Kohn theorems state is that if one knows the energy functional there exists a variational principle on the density and upon minimization, the energy obtained is that of the ground state and the corresponding density is the ground state density. Once the ground state density has been found any observable can be calculated. Unfortunately, no prescription is given in the theorems about how one goes obtaining these functional $E[n]$.

2.2.3 The Kohn-Sham (KS) equations

This problem was partially solved by Kohn and Sham one year later [52]. The strategy is to separate the classical electrostatic energy (so-called Hartree term) from the exchange and correlation contributions. This strategy is useful becaurse it devides the generally unknown electron-electron interaction energy into pieces: Hartree, exchange, and correlation. More importantly, the Hartree term, which is the largest contribution, is just the classical electrostatic energy, which is known exactly. The second largest, the exchange term, is also well-known and can be calculated as in Hartree-Fock theory. Eventually, all the ignorance about the electronic many-body problem is displaced to the smallest contribution, namely the correlation term.

The main problem here is with kinetic energy $T = \langle\Psi|\hat{T}|\Psi\rangle$ because its explicit expression in terms of the electronic density is not known. In 1965 Kohn and Sham suggested the approach that starts from the observation, that a system of non-interacting electrons is exactly described by an antisymmetric wave function of the Slater determinant type, made of one-electron orbitals. For example, for a two particle system with opposite spins

$$\Psi(x_1, x_2) \quad = \quad \frac{1}{\sqrt{2}}\left(\chi(x_1)\chi(x_2) - \chi(x_2)\chi(x_1)\right)\alpha(x_1)\alpha(x_2) \tag{2.33}$$

As in Hartree-Fock theory, for such a wave function the kinetic energy can be easily obtained in terms of the one-electron orbitals. In this case the ground state density matrix $n_1(r, r')$ is given by

$$n_1(r, r') = \sum_{i=1}^{\infty} f_i \phi_i(r)\phi_i^*(r'), \tag{2.34}$$

where $\phi_i(r)$ are the one-electron orbitals and f_i are the occupation numbers corresponding to these orbitals. Then the exact expression for the kinetic energy of non-interacting electrons is

$$T = -\frac{\hbar^2}{2m}\sum_{i=1}^{\infty} f_i \langle\phi_i|\nabla^2|\phi_i\rangle. \tag{2.35}$$

Kohn and Sham's idea was, if one can find the system of non-interacting electrons that produces the same electronic density of the interacting system, then the kinetic energy of the non-interacting system can be calculated exactly via Eq.2.35. This is not the exact kinetic energy of the interacting system. The missing fraction is due to the fact that the true many-body wave function is not a Slater determinant. There is then a correlation contribution to the kinetic energy that is not taken into account, which must be included in the correlation energy term.

We assume that the equivalent non-interacting system (a system of non-interacting electrons whose ground state density coincides with that of the interacting system) does exist. This system will be called the *non-interacting reference system* of density n(r), and described by the Hamiltonian

$$\hat{H}_R = \sum_{i=1}^{N} \Big[-\frac{\hbar^2}{2m}\nabla_i^2 + \vartheta_R(r_i)\Big], \tag{2.36}$$

with N the number of electrons. Here, the potential $\vartheta_R(r)$, or *reference potental*, is such that the ground state density of $\hat{H}_R$ equals to n(r).

This Hamiltonian has no electron-electron interactions. Therefore, its eigenstates can be expressed in the form of Slater determinants

$$\Phi(r) = \frac{1}{\sqrt{N_s!}} SD[\phi_1(r_1)\phi_2(r_2)...\phi_{N_s}(r_{N_s})], \tag{2.37}$$

where we have chosen the closed shell situation, so the occupation numbers are 2 for $i \leq N_s$ and 0 for $i > N_s$, with $N_s = N/2$ the number of doubly occupied orbitals. For simplicity we ignored a possible spin dependence.

Within this assumption, the density becomes

$$n(r) = 2\sum_{i=1}^{N_s} |\phi_i(r)|^2, \tag{2.38}$$

while the kinetic term is

$$T_R[n] = -\frac{\hbar^2}{2m}\sum_{i=1}^{N_s}\langle\phi_i|\nabla^2|\phi_i\rangle. \tag{2.39}$$

The single-particle orbitals $\phi_i(r)$ are the N_s lowest-energy eigenfunctions of the one -electron Hamiltonian

$$\hat{H}_{KS} = -\frac{\hbar^2}{2m}\nabla^2 + \vartheta_R(r), \tag{2.40}$$

which are obtained by solving a set of one-electron Schrödinger equations

$$\hat{H}_{KS}\phi_i(r) = \varepsilon_i\phi_i(r). \tag{2.41}$$

Using $T_R[n]$, the universal density functional can be re-written in the following form:

$$F[n] = T_R[n] + \frac{1}{2}\int\int \frac{n(r)n(r')}{|r-r'|}drdr' + E_{XC}[n], \tag{2.42}$$

where E_{XC} is the *exchange-correlation energy.*

By substituting this expression for F into the total energy functional $F_\vartheta[n] = F[n] + \int n(r)\vartheta_{ext}(r)dr$, the Kohn-Sham (KS) functional is obtained:

$$F[n] = T_R[n] + \int n(r)\vartheta_{ext}(r)dr + \frac{1}{2}\int\int \frac{n(r)n(r')}{|r-r'|}drdr' + \tilde{E}_{XC}[n], \tag{2.43}$$

The energy is expressed in terms of the N_s orbitals that minimize the non-interacting electronic kinetic energy under the fixed density constrain. The one-electron orbitals are called *Kohn-Sham orbitals.*

The Knoh-Sham orbitals satisfy the one-electron Kohn-Sham equations 2.41, but it is still not known what the reference potential ϑ_R is. What is known is that ϑ_R is a potential that ensures that the density of the non-interecting reference system is the same as the true density of the interecting density. It should be then possible to detemine it by minimizing the KS functional 2.43 with respect to the density , under the constrain that this density integrates to N particles.

$$\frac{\delta}{\delta n(r)}\left(E_{KS}[n] - \mu\int n(r)dr\right) = 0, \tag{2.44}$$

Using the rules of functional derivation, the equation for minimizing ground state density is obtained:

$$\frac{\delta T_R[n]}{\delta n(r)} + \vartheta_{ext}(r) + \int \frac{n(r')}{|r-r'|}dr' + \frac{\delta \tilde{E}_{XC}[n]}{\delta n(r)} = E_{KS}[\mu]. \tag{2.45}$$

The functional derivative $\delta T_R[n]/\delta n(r)$ is obtained by considering the non-interacting Hamiltonian $\hat{H}_R$ of Eq.2.36. The particles in the reference system do not interact between themselves, they interact with the potential, so this Hamiltonian corresponds to the energy functional

$$E_{\vartheta_R}[\tilde{n}] = T_R[\tilde{n}] + \int [\tilde{n}]\vartheta_R(r)dr, \tag{2.46}$$

with the same ground state energy of the interacting system. Therefore, in general $E_{\vartheta_R}[\tilde{n}] \geq E_0$. Hence, the functional derivative of $E_{\vartheta_R}[\tilde{n}]$ must vanish for the ground-state density. Applying Eq.2.44 - 2.46, we obtain

$$\frac{\delta T_R[n]}{\delta n(r)} + \vartheta_R(r) = \mu_R, \tag{2.47}$$

where μ_R is the potential of the non-interacting system.

By comparing 2.45 and 2.47 the following expression for the reference potential is obtained:

$$\vartheta_R(r) = \vartheta_{ext}(r) + \int \frac{n(r')}{|r-r'|}dr' + \mu_{XC}[n](r), \tag{2.48}$$

where

$$\mu_{XC}[n](r) = \frac{\delta \tilde{E}_{XC}[n]}{\delta n(r)}. \tag{2.49}$$

2.2.4 Approximations

The many-body electronic problem can be solved by deviding the total energy into a number of different contributions, equation $E[n] = T_R + \vartheta_{ext} + E_H + E_X + \tilde{E}_C$. These are the non-interecting energy T_R, the classical electron-electron interaction or Hartree term E_H, the interaction of the electrons with external fields ϑ_{ext}, the exchange energy E_X, and the correlation term E_C. ϑ_{ext} and E_H are known as explicit functionals of the electronic density. T_R and E_X are known as the functionals of the non-interacting orbitals, which are in turn functionals of the density[4]. E_C is the big unknown. To express this term we have to use some approximations. Two the most widely used approaches to the exchange-correlation potential, named local-density (LDA) and generalized gradient approximations (GGA), will be described in the following section.

Local functional: local-density approximation (LDA)

The first approximation to the exchange-correlation (XC) energy is the LDA. The way density-functional theory, in the local-density approximation, deals with the inhomogeneous many-body problem is by decomposing it into two simpler problems: the solution of a spatially homogeneous interacting problem (the homogeneous electron liquid) gives the uniform XC energy $e_{xc}^{hom}(n)$, and the solution of a spatially inhomogeneous noninteracting problem (the inhomogeneous electron gas described by the KS equations) gives the particle density. Both steps are connected by the local-density potential (Eq.2.50) [53]. The exchange-correlation energy can be written as the average of an energy density $\epsilon_{XC}^{LDA}[n]$:

$$E_{XC}^{LDA}[n] = \int n(r)\tilde{\epsilon}_{XC}^{LDA}[n(r)]dr, \tag{2.50}$$

In practice, the exchange-correlation energy within the LDA is calculated via Eq.2.50 using $\tilde{\epsilon}_{XC}^{LDA}[n(r)] = \epsilon_X^{LDA}[n(r)] + \tilde{\epsilon}_C^{LDA}[n(r)]$, where $\epsilon_X^{LDA}[n(r)]$ is the exchange energy density, given by Dirac's expression:

$$\epsilon_X^D[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} n^{1/3} = -\frac{3}{4}\left(\frac{9}{4\pi^2}\right)^{1/3}\frac{1}{r_s} = -\frac{0.458}{r_s}a.u., \tag{2.51}$$

where $r_s = (3/4\pi n)^{1/3}$ is the mean interelectronic distance expressed in atomic units.

---

[4] Althoug E_X is known exactly it is an explicit function of the KS orbital instead of the density, thus we will also provide an approximation for $E_X[n]$.

For the correlation energy $e_c^{hom}(n)$ is not known exactly. The modern approximate expressions for $e_c^{hom}(n)$ are based on adjusted curves obtained from Quantum Monte Carlo calculations for the electron liquid by Ceperley and Alder [54].

Semilocal functional: generalized-gradient approximation (GGA)

In the LDA one uses knowledge of the density at point $\overrightarrow{r}$, but any real system is spatially inhomogeneous (it has a spatially varying density $n(r)$), and it would be useful also to include information on the rate of this variation in the function.

In GGA the functional is written as

$$E_{xc}^{GGA}[n] = \int d^3 f(n(r), \nabla n(r)), \tag{2.52}$$

For many cases GGA's seem to give reliable results for all main types of chemical bonds (covalent, ionic, metallic and hydrogen bridge). For van der Waals interactions, however, common GGA's and LDA's fail, specially for total energy calculations and relaxations. Interestingly, if the positions of the atoms are known *a priori* it doesn't influence the electronic structure. To describe these very weak interactions several more specialized approaches have been developed. Information about these approaches can be found in [55, 56].

More detailed explanations of DFT can be found in the [57, 53, 58].

A number of implementations of DFT exist with various levels of approximations. In my simulations we used SIESTA package [59] since it uses a localized basis set which is particulary suited for treating the electronic transport.

2.3 Hybrid quantum mechanics/molecular mechanics approach (QM/MM)

In previous section I described the first-principles[5] methods for simulating reactions and electronic processes to high accuracy. These are however limited by their computational cost to relatively small systems. When the system of chemical interest consists of many thousands of atoms, the hybrid QM/MM schemes can be used. Scientists, Martin Karplus, Harvard University, USA (also at the University of Strasbourg, France), Michael Levitt, Stanford University, USA and Arieh Warshel, University of Southern California, USA, have been awarded the 2013 Nobel Prize in Chemistry for the development of multiscale models for complex chemical systems - QM/MM scheme [60]. QM/MM scheme is obtained by embedding a quantum mechanics calculation in a classical molecular mechanics model of environment. It is probably both of the worlds as it enables to treat a specific region

---

[5] These are called first principles because they do not introduce any parameters into a calculation except the fundamental constants.

quantum mechanically, but evolve it under the influence of external environment. It was originally developed for dealing with reactions in biological systems.

The first step in QM/MM scheme is to divide the entire system into inner and outer regions. The main concept of QM/MM approach is shown in Figure 2.2.

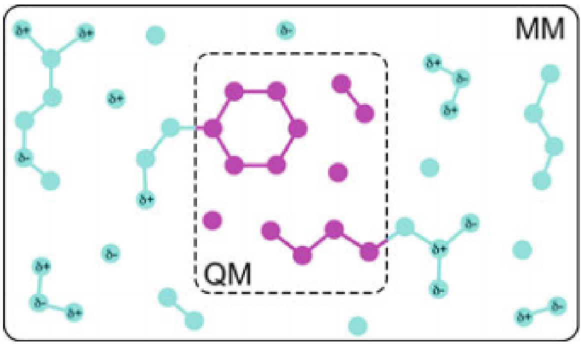


Figure 2.2: Illustration of the QM/MM concept. A small region, in which a chemical reaction occurs, is treated at level of QM theory. The remainder of the system is modelled at the MM level. Figure is extracted from [9].

The objective then is to use QM calculations to model processes in the inner region and to use MM to model the outer region. Therefore, some form of treatment of the boundary is required. To do this several approaches were proposed. This approaches can be divided in two categories: subtractive and additive coupling schemes [9].

2.3.1 Subtractive QM/MM Coupling

In this scheme the QM/MM energy of the system is obtained in three steps. First, the energy of the total system is evaluated at the MM level. Second, QM energy of the inner region is added. Third, the MM energy of the inner system is computed and subtracted. The last step is to correct the inclusion of the interactions within the inner system twice:

$$V_{QM/MM} = V_{MM}(MM + QM) + V_{QM} - V_{MM}(QM), \tag{2.53}$$

The main advantage of this scheme is that no communication is required between the QM and MM. The disadvantage of the method is that the force-field is required for the inner system, that sometimes is not available.

2.3.2 Additive QM/MM Coupling

In the additive scheme the potential energy is the sum of the QM inner system term, MM outer system term and coupling QM/MM terms.

$$V_{QM/MM} = V_{QM} + V_{MM}(MM) + V_{QM+MM}(QM + MM), \tag{2.54}$$

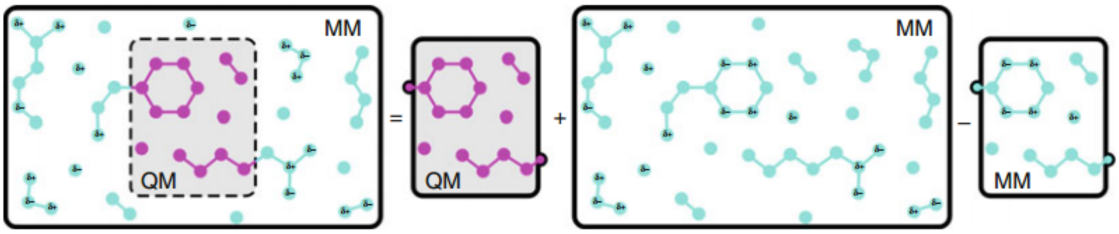


Figure 2.3: Subtractive QM/MM coupling. Figure is extracted from [9].

There are three basic approaches in additive scheme: mechanical embedding, electrostatic embedding and polarization embedding [9].

Mechanical Embedding

In mechanical embedding approach all forces between inner and outter system are described in force-field level. Chemical bonds between atom from inner system and atom from outer system described by harmonic potential V_{bonds}, angles between one atom of inner system and two atoms of outer system are described by harmonic potential V_{angles}, torsion involving two atoms of inner atoms - by V_{tors}. Non-bonded interactions are described in force-field as well. Van der Waals interactions are described by Lennard-Jones potential V_{LG}, electrostatic interactions - by Coulomb potential V_{Coul}.

Electrostatic Embedding

Electrostatic embedding approach is an improvement of the Mechanical embedding approach. In this approach the polarization effects are included. The charged MM atoms are presented in QM Hamiltonian as one-electron operators:

$$h_i^{QM-MM} = h_i^{QM} - \sum_J^M \frac{e^2 Q_J}{4\pi\epsilon_0 |r_i - R_J|}, \tag{2.55}$$

where r_i and R_J are the positions of electron i and MM atom j, h_i^{QM-MM} is original one-electron operator for the kinetic and nuclear attraction energy of electron i, M is the number of MM atoms that have a particle charge Q_J.

Polarization Embedding

Including polarization of the MM atoms is the next approach, that make an improvement of the previous approaches. It is possible that both regions can mutually polarize each other. The most popular methods in this approach are charge-on-a-spring model, the induced dipole model and the fluctuating charge model [61].

2.3.3 Capping Bonds at the QM/MM Boundary

The choice between what is classical and what is quantum is somewhat orbitrary and is usually a balance between the minimum size of the quantum region, that is required to capture an effect and the computational effort required for a calculation.

While choosing the QM and MM region the cut through the QM/MM bond creates artificially one or more unpaired electrons in the QM region. There are exist a number of approaches to solve this problem. Two main approaches are: link atom and localized orbitals [9].

Link Atoms

One of the solutions is to add a monovalent link atom along the bond between the QM and MM regions. The type of the link atom can be chosen, but usually it is a hydrogen. In practice, the link atom is placed at a fixed position along the bong on each step of the simulation, and in the corresponding charge distribution is corrected to account for this.

Localized Orbitals

An alternative method for the link atom is to replace the chemical bond between QM and MM regions by a double-occupied molecular orbital. Two most used approaches are: localized hybrid orbital method, which introduces the molecular orbital in QM part, and generalized hybrid approach, which introduces the molecular orbital in the MM part. In the localized hybrid orbital method the atomic orbitals on QM atom of the broken bond are localized and hybridized. The hybrid orbital pointing towards the MM atom is occupied by two electrons. The other orbitals are occupied by one electron.During the optimization of the QM wave function, the double-occupied orbital is kept frozen, while other orbitals are optimized with all orbitals in the QM region. In the generalized hybrid approach hybrid orbitals are placed on the MM atom of the broken bond. The orbital pointing to the QM atom is optimized while others kept frozen.

2.4 Electronic transport calculations

As a pictorical view of transport and DNA sequencing we start from a scattering problem for a very simple one dimentional system with an infinite number of identical atoms with the same on-site potential ε and state $|j\rangle$ as shown in Figure 2.4. This simple problem can be solved exactly if we assume that only first nearest neighbor interactions exist and that all the couplings are the same. We thus write a Hamiltonian for this system:

$$H = \sum_{j=-\infty}^{\infty} \varepsilon |j\rangle\langle j| + \sum_{j=-\infty}^{\infty} \gamma |j\rangle\langle j+1| + \gamma^* |j\rangle\langle j-1| \tag{2.56}$$

By proposing a solution of the form:

$$|\Psi_k\rangle = \frac{1}{\sqrt{N}} \sum_{j'} e^{ij'k} |j'\rangle \tag{2.57}$$

with $E = \varepsilon + 2\gamma \cos k$ it is possible to show that $|\Psi\rangle$ is an eigenstate. We will make an approximation that each atom only interacts with its neighbor. The coupling term is γ.

Now we note that the state $|0\rangle$ have a different on site potential ε_0 and the coupling term between this atom and it's neighbors is γ_1 (see Figure 2.4).

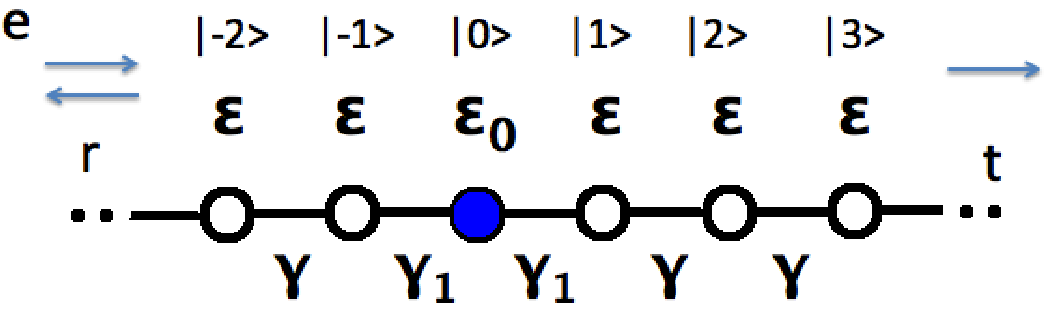


Figure 2.4: Pictorical representation of the electronic transport problem. We consider a single level per site and first - nearest neighbor interaction only.

The total Hamiltonian of this system can now be written as:

$$H = \sum_{j \neq 0} \varepsilon |j\rangle\langle j| + \varepsilon |0\rangle\langle 0| + \sum_{j\neq 0, j\neq -1} \gamma |j\rangle\langle j+1| +$$

$$+ \sum_{j\neq 0, j\neq -1} \gamma |j+1\rangle\langle j| + \gamma_1 |-1\rangle\langle 0| + \gamma_1 |0\rangle\langle -1| + \gamma_1 |0\rangle\langle 1| + \gamma_1 |1\rangle\langle 0| \tag{2.58}$$

Let's now imagine one electron e entering the system (say from the left). There exists a probability that, meeting the potential barrier, the electron can be reflected or transmitted. We call the reflection term r and transmission term t.

The proposed solution of this problem is then an liniar combination of the wave function of the incoming electron, wave function of the reflected electron (multiplied by reflection coefficient r), wave fuction of the scattering region (multiplied by the coefficient a) and wave function of the transmitted electron (multiplied by transmission coefficient t):

$$|\Psi\rangle = \frac{1}{\sqrt{N}} \sum_{j'=-\infty}^{-1} e^{ij'k} |j'\rangle + \frac{r}{\sqrt{N}} \sum_{j'=-\infty}^{-1} e^{-ij'k} |j'\rangle + a|0\rangle + \frac{t}{\sqrt{N}} \sum_{j'=1}^{\infty} e^{ij'k} |j'\rangle \tag{2.59}$$

By solving Schrödinger Equation (Eq.2.60) we can find the transmission and reflection coefficients.

$$H|\psi\rangle = E|\psi\rangle \tag{2.60}$$

We do this by projecting basis $|-1\rangle$, $|0\rangle$ and $|1\rangle$ onto the Schrödinger Equation to obtain the system of three equations with three unknowns. Reorganizing the system we build the following linear system:

$$\begin{pmatrix} (\varepsilon + \gamma e^{ik} - E)e^{ik} & \gamma_1 & 0 \\ \gamma_1 e^{ik} & \varepsilon_0 - E & \gamma_1 e^{ik} \\ 0 & \gamma_1 & (\varepsilon + \gamma e^{ik} - E)e^{ik} \end{pmatrix} \begin{pmatrix} r \\ a \\ t \end{pmatrix} =$$

$$= \begin{pmatrix} (E - \varepsilon - \gamma e^{-ik})e^{-ik} \\ -\gamma_1 e^{-ik} \\ 0 \end{pmatrix} \tag{2.61}$$

From this the probability of the electron to be reflected R and transmitted T will be:

$$\begin{aligned} R &= |r|^2 \\ T &= |t|^2 \end{aligned} \tag{2.62}$$

The solution of the matrix is presented in the Figure 2.5. For an appropriate choice of parameters (see table 2.1) we can note that the transmission is low for a wide range of energies except for narrow region around $-0,5eV$. The width of the peak of the curves depends on the γ_1 value (here we choose $\gamma_1 << \gamma$).

Table 2.1: Values of the parameters of the linear system of equations 2.61

Parmeter	Value
γ	1
γ_1	0.3
ε	0
ε_0	-0.5

Basically, the smaller the coupling γ_1 between the two electrodes and the central state, the sharper the resonance. Furthermore, the position of the resonance is $-0,5eV = \varepsilon_0$. This can be understood by the following picture: electrons came in "freely" from the left electrode, but reach a region where there is low probability of hopping (γ_1). In fact the solution presented in Eq.2.57 is what is known as a Bloch state a propagating state, the equivalent in a solid, to a free electron. However, when the energy is approximately equal to ε_0, there will be a resonance and the electron will tunnel to the other side. The width of this pick is related to γ_1. Thus, if we have a "molecule" in the middle of two electrodes it is possible to obtain information about its electronic structure and coupling to the electrodes.

As mentioned, the electrons require a quantum mechanical description. In 1987 Landauer proposed [62] that the problem of electronic transport could be thought of as a scattering problem. In other words, the propogation of electrons in a material could be described by the same formalism used for free electrons being scattered by some potential barrier [63].

From this perspective, the idea of conductance can be associated to the probability that an incoming electron might be transmitted across the barrier in the same.

We will map this simple system to the more complicated one. We can compare ε with the potential of the carbon atoms of graphene, ε_0 with the potential of the base, γ with the coupling term of the carbons in graphene and γ_1 with coupling term between base and pore.

In principle, the methodology presented above can be generalized for a more realistic Hamiltonian (one obtained from DFT, for example). The proposed wavefunction (Eq.2.59) would have the same form but would have a much larger number of available "channels".

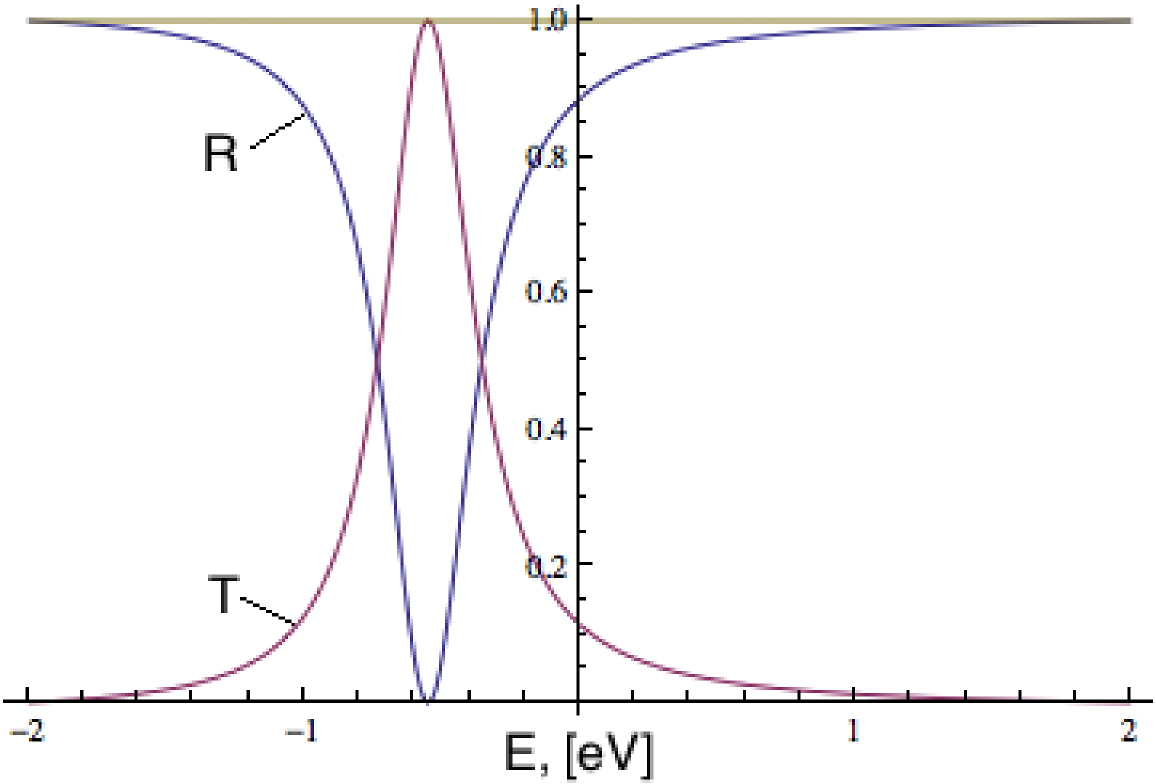


Figure 2.5: Probability of the electron to be reflected and transmitted. Blue curve R - reflection. Red curve T - transmission

2.5 Non equilibrium Green's Functions (NEGF)

The alternative way is to solve Schrödinger's Equation using a Green's function instead a wave function. Non equlibrium Green's function method is used to calculate transmission and charge densities in nanoscale devices, mainly for ballistic transport calculations [64], [65]. To define the Green's functions let's write the Schrödinger Equation:

$$H|n\rangle = E|n\rangle \tag{2.63}$$

We devide the Hamiltonian and wavefunction of the system into contact ($H_{L,R}$, $|\psi_{L,R}\rangle$) and device ((H_M, $|\psi_M\rangle$):

$$\begin{pmatrix} H_L & \tau_L & 0 \\ \tau_L^\intercal & H_M & \tau_R^\intercal \\ 0 & \tau_R & H_R \end{pmatrix} \begin{pmatrix} |\psi_L\rangle \\ |\psi_M\rangle \\ |\psi_R\rangle \end{pmatrix} = E \begin{pmatrix} |\psi_L\rangle \\ |\psi_M\rangle \\ |\psi_R\rangle \end{pmatrix} \tag{2.64}$$

where $\tau_{L,R}$ is interaction between device and contacts (we assume, that the contacts are independent and there is no cross term τ between the different contacts)[6].

We now can define the Green's function:

$$(E - H)G(E) = I \tag{2.65}$$

where I is the identity matrix.

For the infinite system it is easier to calculate the Green's function than to solve the whole eigenvalue problem. Most of the properties of the system can be obtained from the Green's function. For example, the wavefunction of the contact $|\psi_R\rangle$ can be calculated if we know the wavefunction on the device $|\psi_M\rangle$ (from the equation 2.64):

$$H_R|\psi_R\rangle + \tau_R|\psi_M\rangle = E|\psi_R\rangle \rightarrow \tag{2.66}$$

$$(E - H_R)|\psi_R\rangle = \tau_R|\psi_M\rangle \rightarrow \tag{2.67}$$

$$|\psi_R\rangle = g_R(E)\tau_R|\psi_M\rangle \tag{2.68}$$

where g_R is the Green function of the isolated contact R ($(E - H_R)g_R = I$).

To include the effects of the contacts on the device we have to introduce the idea of self-energies. From the definition of the Green's function we have:

$$\begin{pmatrix} E - H_L & \tau_L & 0 \\ -\tau_L^\intercal & E - H_M & -\tau_R^\intercal \\ 0 & \tau_R & E - H_R \end{pmatrix} \begin{pmatrix} G_L & G_{LM} & G_{LM} \\ G_{ML} & G_M & G_{MR} \\ G_{RL} & G_{RM} & G_R \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \tag{2.69}$$

Selecting the three equations in the second column:

---

[6] We note here that the contacts represent a system with a very large ($\rightarrow \infty$) number of degrees of freedom. The dence on the other hand is finite.

$$(E - H_L)G_{LM} - \tau_L G_g = 0 \tag{2.70}$$

$$-\tau_L^{\intercal} G_{LM} + (E - H_M)G_M - \tau_R^{\intercal} G_{RM} = I \tag{2.71}$$

$$(E - H_R)G_{RM} - \tau_R G_M = 0 \tag{2.72}$$

.

We solve Eq. 2.70 and 2.72 for G_{LM} and G_{RM} and substituting G_{LM} and G_{RM} in the equation 2.71 obtain expression for the Green's function of the device Eq. 2.76:

$$G_{LM} = g_L \tau_L G_M \tag{2.73}$$

$$G_{RM} = g_R \tau_R G_M \tag{2.74}$$

$$-\tau_L^{\intercal} g_L \tau_L G_M + (E - H_M)G_M - \tau_R^{\intercal} g_R \tau_R G_M = I \tag{2.75}$$

,

$$G_g = (E - H_M - \Sigma_L - \Sigma_R)^{-1}, \tag{2.76}$$

where $\Sigma_L = \tau_L^{\intercal} g_L \tau_L$ and $\Sigma_R = \tau_R^{\intercal} g_R \tau_R$ are the so called self-energies.

Basically, to include the effects of the contacts on the device we need to add self-energies to the device Hamiltonian. This way it is possible to focus on the device region while factoring out the degrees of freedom in the contacts. When we calculate Green's functions for the device we calculate them only for the effective Hamiltonian $H_{effective} = H_M + \Sigma_L + \Sigma_R$ [66].

Schematic representation of the scattering problem at nanoscale of the studied system is shown in the Figure 2.6. It is a central region, called scattering region, sandwiched by electrodes, called left (right) lead. The leads are defect-free periodic structures. We assume that each lead cell interacts only with it neighbour. The central scattering region is described by Hamiltonian H_M, coupling between scattering region and leads is described by the Hamiltonian H_{RM} (H_{LM}). The unit cell of the lead is descibed by Hamiltonian H_0 and coupling between two lead cells - by Hamiltonian H_1.

As we are dealing with an infinite-dimensional non-periodic Hermitial problem[7], the problem can be written in terms of retarded[8] Green's function G^R for the system by solving the Green's function equation [64]:

---

[7] The Hamiltonian matrix is Hermitian, it is a square matrix with complex entries that is equal to its own conjugate transpose - that is, the element in the i-th row and j-th column is equal to the complex conjugate of the element in the j-th row and i-th column, for all indices i and j

[8] Since we have infinite system, we obtain two types of solutions for the Green's functions, the retarded and the advanced, corresponding to outgoing and incoming waves in contacts.

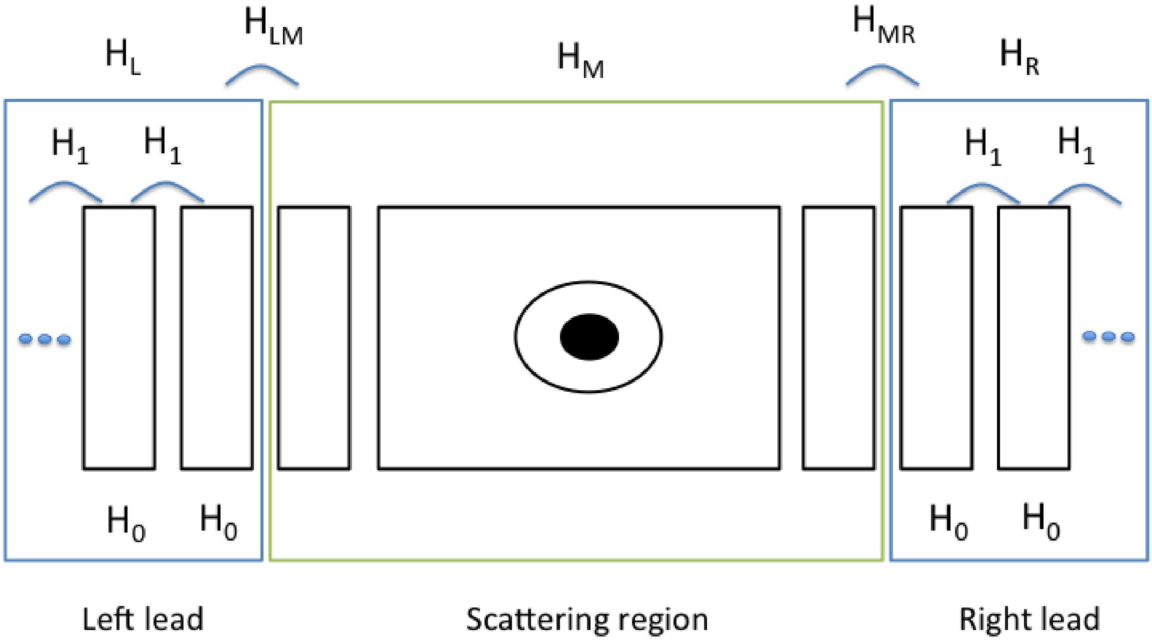


Figure 2.6: Schematic representation for transport problem at nanoscale. The central scattering region is descibed by the Hamiltonian H_M, coupling between scattering region and leads is described by H_{RM} (H_{LM}). H_0 corresponds to unit cell of lead and H_1 - to coupling between two lead cells.

$$[\epsilon^{+}S - H]G^{R}(E) = I, \tag{2.77}$$

where I is an infinitely-dimensional identity matrix, $\epsilon^{+} = \lim_{\delta\to 0^{+}} E + i\delta$ and E is the energy.

The same equation can be written in matrix form:

$$\begin{pmatrix} \epsilon^{+}S_L - H_L & \epsilon^{+}S_{LM} - H_{LM} & 0 \\ \epsilon^{+}S_{ML} - H_{ML} & \epsilon^{+}S_M - H_M & \epsilon^{+}S_{MR} - H_{MR} \\ 0 & \epsilon^{+}S_{RM} - H_{RM} & \epsilon^{+}S_R - H_R \end{pmatrix} \begin{pmatrix} G_L^R & G_{LM}^R & G_{LR}^R \\ G_{ML}^R & G_M^R & G_{MR}^R \\ G_{RL}^R & G_{RM}^R & G_R^R \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I_M & 0 \\ 0 & 0 & I \end{pmatrix} \tag{2.78}$$

where G_L^R and G_R^R describe left lead and right lead respectively, G_{LM}^R and G_{RM}^R describes interaction between the molecule and the leads, G_{LR}^R describes the direct scattering between the leads and G_M^R describes a molecule.

H_M is an MXM matrix and I_M is the MXM unit matrix. The infinite matrices H_L and H_R have the following block-diagonal form:

$$H_L = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & H_{-1} & H_0 & H_1 & 0 \\ \dots & 0 & H_{-1} & H_0 & H_1 \\ \dots & \dots & 0 & H_{-1} & H_0 \end{pmatrix} \tag{2.79}$$

$$H_R = \begin{pmatrix} H_0 & H_1 & 0 & \dots & \dots \\ H_{-1} & H_0 & H_1 & 0 & \dots \\ 0 & H_{-1} & H_0 & H_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.80}$$

Overlap matrix S has the same expression. Infinite coupling matrix H_{LM} (H_{RM}), that describes the interaction between the molecule and the closest unit cell of the lead, is the matrix whose elements are all zero except for a rectangular block coupling of the unit cell of the lead and molecule:

$$H_{LM} = \begin{pmatrix} \vdots \\ 0 \\ 0 \\ H_{LM} \end{pmatrix} \tag{2.81}$$

$$H_{RM} = \begin{pmatrix} H_{RM} \\ 0 \\ 0 \\ \vdots \end{pmatrix} \tag{2.82}$$

In our calculations we obtain Hamiltonian matrices H_{LM}, H_{RM} and H_M from the DFT calculations.

Considering that there is no change in electronic structure of the leads we can treat the effect of the electrodes in terms of an effective interaction. The final expression for G_M^R associated to the effective Hamiltonian matrix $H_{eff} = H_M + \Sigma_L^R(E) + \Sigma_R^R(E)$ can be than obtained:[9]

$$G_M^R(E) = [\epsilon^+ S_M - H_M - \Sigma_L^R(E) - \Sigma_R^R(E)]^{-1}, \tag{2.83}$$

where $\Sigma_L^R(E)$ and $\Sigma_L^R(E)$ are self-energies for the left and right lead, respectively (Eq.2.84, 2.85). G_M^R is the retarded Green's function

$$\Sigma_L^R(E) = (\epsilon^+ S_{ML} - H_{ML}) g_L^{0R}(E)(\epsilon^+ S_{LM} - H_{LM}), \tag{2.84}$$

$$\Sigma_R^R(E) = (\epsilon^+ S_{MR} - H_{MR}) g_L^{0R}(E)(\epsilon^+ S_{RM} - H_{RM}), \tag{2.85}$$

---

[9] The degrees of freedom of the electrodes are eliminated one by one from deep into the leads all the way to the interface with a molecule.

where g_L^{0R} and g_R^{0R} (Eq.2.86, 2.87) are the Green's functions for the isolated leads, while g_L^R and g_R^R - for the leads attached to the molecule.

$$g_L^{OR}(E) = [\epsilon^+ S_L - H_L]^{-1}, \tag{2.86}$$

$$g_R^{OR}(E) = [\epsilon^+ S_R - H_R]^{-1}. \tag{2.87}$$

Knowing G_M^R we can find the total transmission coefficient $T(E)$ using the Buttiker-Landauer formula [62]:

$$T(E) = T_r[\Gamma_L G_M^{R\intercal} \Gamma_R G_M^R], \tag{2.88}$$

where Γ_L and Γ_R are the coupling matrices.

$$\Gamma_L(E) = i[\Sigma_L^R(E) - \Sigma_L^R(E)^\intercal], \tag{2.89}$$

$$\Gamma_R(E) = i[\Sigma_R^R(E) - \Sigma_R^R(E)^\intercal], \tag{2.90}$$

2.6 Returning to the simple problem

The methodology presented above is very general, but we can return to our simple problem.

Let's first calculate the self-energies for each of the electrodes. In that case the Hamiltonian can be writen as:

$$H = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \dots & \varepsilon & \gamma & 0 & 0 & 0 & \dots \\ \dots & \gamma & \varepsilon & \gamma_1 & 0 & 0 & \dots \\ \dots & 0 & \gamma_1 & \varepsilon_0 & \gamma_1 & 0 & \dots \\ \dots & 0 & 0 & \gamma_1 & \varepsilon & \gamma & \dots \\ \dots & 0 & 0 & 0 & \gamma & \varepsilon & \dots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.91}$$

Overlap matrix has expression:

$$S = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \dots & 1 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & 1 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 1 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 1 & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 1 & \dots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.92}$$

Thus

$$H_{LM} = H_{ML} = \gamma_1, \tag{2.93}$$

$$S_{LM} = S_{ML} = 0. \tag{2.94}$$

Consequently

$$\Sigma_{L/R}(E) = \gamma_1^2 g^0_{L/R}(E). \tag{2.95}$$

If we consider that the Gaussian-substitution procedure presented from Eq.2.70 Eq.2.73 can be performed recursively also in the contacts, we can take

$$ES_L - H_L = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots \\ \dots & E-\varepsilon & \gamma & 0 & 0 \\ \dots & \gamma & E-\varepsilon & \gamma & 0 \\ \dots & 0 & \gamma & E-\varepsilon & \gamma \\ \dots & 0 & 0 & \gamma & E-\varepsilon \\ , & & & & \end{pmatrix} \tag{2.96}$$

$$g_0^L = (ES_L - H_L)^{-1}. \tag{2.97}$$

But adding an extra layer will not change the surface. Then using the Green's function equation we find expression for the self-energy $\Sigma_{L/R}$:

$$\begin{pmatrix} (g_0^L)^{-1} & \gamma \\ \gamma & E-\epsilon \end{pmatrix} \quad \begin{pmatrix} g_L & g_{12} \\ g_{21} & g_L \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{2.98}$$

$$(g_0^L)^{-1} g_{12} + \gamma g_L = 0 \tag{2.99}$$

$$\gamma g_{12} + (E-\varepsilon) g_L = I \tag{2.100}$$

$$g_L = \frac{(E-\varepsilon) \pm \sqrt{(E-\varepsilon)^2 - 4\gamma^2}}{2\gamma^2} \tag{2.101}$$

Keeping in mind that $E = \varepsilon + 2\gamma \cos k$ we obtain:

$$\frac{2\gamma \cos k \pm 2\gamma\sqrt{\cos^2 k - 1}}{2\gamma^2} = \frac{\cos k \pm \imath \sin k}{\gamma} = \frac{e^{\pm \imath k}}{\gamma} \tag{2.102}$$

$$\Sigma_{L/R} = \gamma_1^2 \frac{e^{\imath k}}{\gamma} \tag{2.103}$$

and

$$\Gamma_{L/R} = \frac{\gamma_1^2}{\gamma}(e^{\imath k} - e^{-\imath k}) = -\frac{2\gamma_1^2}{\gamma} \sin k \tag{2.104}$$

Using equations above, we can find the total transmission coefficient:

$$T(E) = \frac{4\gamma_1^4}{\gamma^2}\sin^2 k \frac{1}{\left(E - \varepsilon_0 - \frac{\gamma_1^2 e^{\imath k}}{\gamma} - \frac{\gamma_1^2 e^{\imath k}}{\gamma}\right)^2} \tag{2.105}$$

Comparing Figure 2.5 with Figure 2.7 we can see that both of the models show the same results.

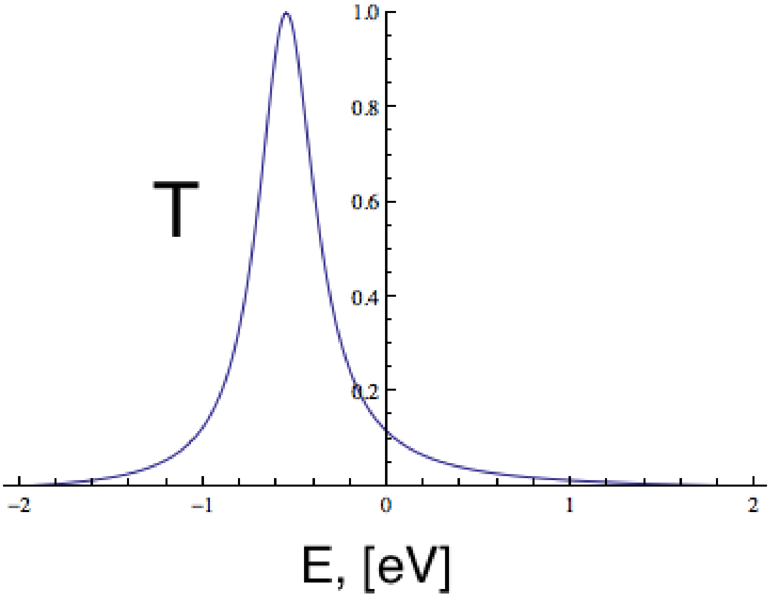


Figure 2.7: Probability of the electron to be transmitted obtained by Non equlibrium Green's function method.

As mentioned previously we are interested in simulating a sequencing device. In order to do that we want to simulate the electronic transport through a graphene nanopore containing DNA molecule (such as the one depicted in Figure 2.8). By applying a potential difference between the left and right extremes of the graphene sheet, electrons will flow across the pore and through the molecular levels.

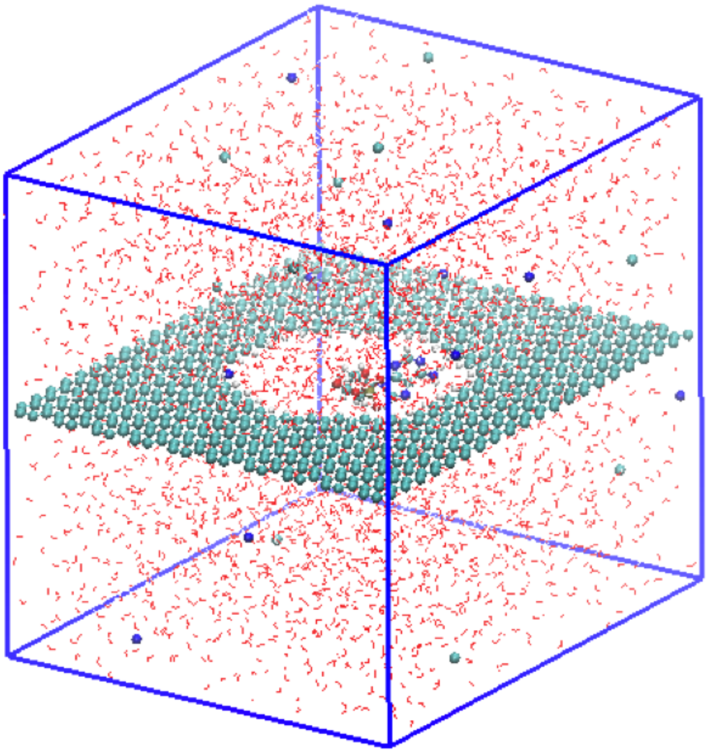


Figure 2.8: Model of the sequencing device to simulate the electronic transport through a graphene nanopore containing DNA molecule.

Chapter 3

Results

3.1 Simulation protocol

In this dissertation we did not simulate the actual translocation of a DNA molecule through the nano pore. The time scale of experiment and simulation is quite different. We studied an isolated bases inside the pore but allowed to move inside in order to probe as many configurations as possible.

The studied system consists of the graphene nanopore (either a large nanopore with H on edge of $22\mathring{A}$ in diameter, a large nanopore with N on edge of $25\mathring{A}$ in diameter, small pore with H on edge of $15\mathring{A}$ in diameter or small pore with N on edge of $17\mathring{A}$ in diameter) as shown in Figure 3.5, DNA nucleobase: Adenine, Guanine, Thymine or Cytosine (see Figure 3.2), inserted one by one inside the pore, counter ions Na and Cl to compensate the negative charges from the DNA bases and also to have a salt concentration commensurate with physyological conditions, and water molecules. In all cases the graphene sheet onto which the nanopore was drilled was $51.9\mathring{A}$ in length and $42.6\mathring{A}$ wide.

In our simulations we use classical molecule dynamics first to analyze the behaviour of different nucleobases inside the different nanopores. The system was simulated using periodic boundary conditions (PBC) with a box size $50X42.6X51.9\mathring{A}$. All simulations were carried out using GROMACS package. To describe the interatomic potentials we are using the AMBER99fb force-field.

The classical simulations were performed in three steps: equilibration, thermalization and production. Thermalization procedure was initially performed for $500ps$ at $300K$ using an NVT ensemble. The equilibration was done for $1000ps$ using NPT ensemble at $300K$ and $1bar$ with Nose-Hoover thermostat and Parrinello-Rahman barostat. During the thermalization and equilibration procedures the movement of bases and graphene nanopore were restricted. In the NPT ensemble the size of the box is allowed to change. Given that the graphene sheet has its position fixed the box size only changes in the perpendicular direction. In the production step the MD calculations were performed for $2000ps$ at a temperature $300K$ using NVT ensemble. Charges of hidrogen and nitrogen of the pore were chosen from organic molecules that were analogue to the N and H pores.

Finally, during the production step the base was allowed to move in the pore, only one atom of the base was fixed in the YZ plane of the pore. 6 snapshots from each trajectory were randomly selected to perform the QM calculations[10].

To determine the electronic structure of the system the QM calculations were performed, using first-principle calculations based on density functional theory (DFT) as implemented in the SIESTA package. For the exchange-correlation potential generalized gradient approximation (GGA) was used in its PBE form. Few different simulations were performed: pure graphene, isolated graphene nanopores, graphene nanopores with inserted nucleobases without water and graphene nanopore with inserted nucleobases surrounded with one layer of water (see Figure 3.1). Core electrons were replaced by norm-conserving pseudopotentials. The Kohn-Sham wave-functions were expanded to a double-ζ polarized for edge atoms of the pore and nucleobase atoms, and a double-ζ basis set for the graphene.

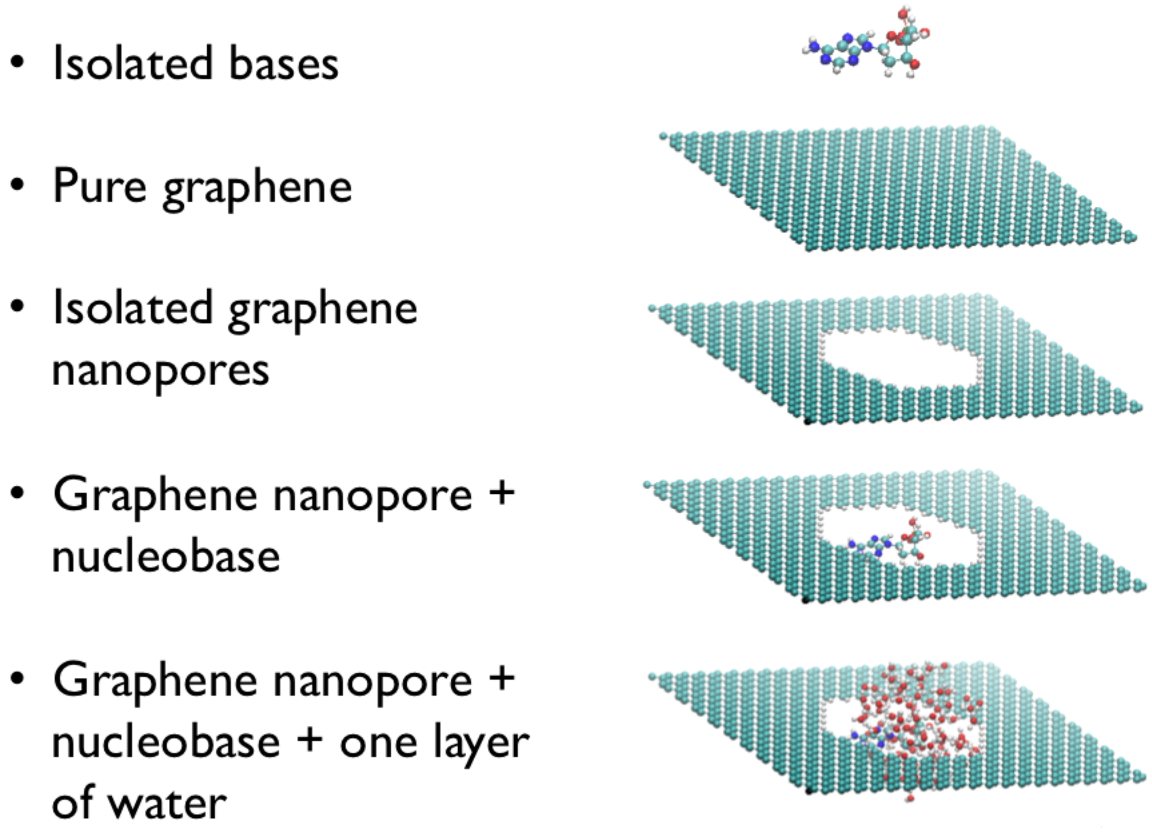


Figure 3.1: Examples of the systems for DFT calculations.

Finally, to include the effect of all the water and ions in our system the QM/MM calculations were performed in the SIESTA package. For the QM region it was chosen the

---

[10] In the true QM/MM approach one would perform a coupled classical/quantum evolution of the system. Since we are not interested in reaction paths we assume that a classical molecular dynamics simulation will yield proper structural information.

pore and base, for the MM region it was chosen water and ions. The inner cutoff radius for the electrostatic and LJ QM-MM interaction (the potential is truncated for distances lower than this cutoff) was chosen $1.10^{-6}\mathring{A}$. The external cutoff radius for the electrostatic and LJ QM-MM interaction (the potential equal zero for distances longer than this cutoff) was chosen as $25\mathring{A}$. The cutoff radius for the electrostatic and LJ interaction between MM atoms was chosen as $25\mathring{A}$. The external radius where the MM atoms beyond which the total force is equal to zero (the MM atoms outside this cutoff are fixed in the calculation) was chosen as $10\mathring{A}$. Here a link atom was not included, because there were no bonds that require cutting between QM and MM region.

Finally, once the Hamiltonian for each snapshot is obtained, zero-bias transmission coefficients were calculated using non-equilibrium Green's functions as implemented in SMEAGOL package. For the electrode it was taken 120 atoms of carbon from the left and from the right of the graphene sheet.

3.2 Electronic structure of the isolated bases

The nucleotide consists of sugar, phosphate and nucleobase. The difference in them is only in the nucleobases. There exist four types of nucleobases: Adenine (DA), Guanine (DG), Cytosine (DC) and Thymine (DT). Adenine from one DNA strand bonds with Guanine on other strand, forming compound Purine. The same way Cytosine bonds only with Thymine, forming Pyrimidine. The chemical composition of the nucleobases is presented in Figure 3.2

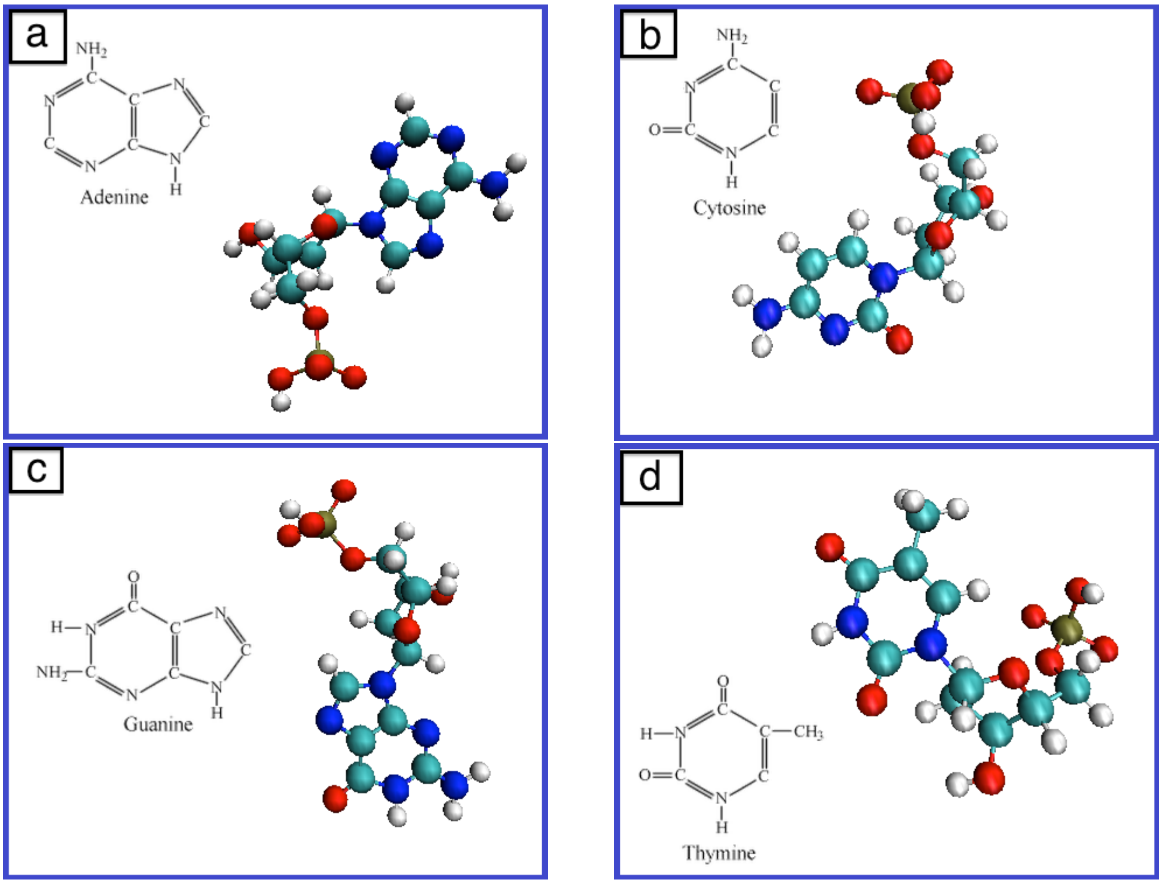


Figure 3.2: Chemical composition of the nucleobases. a) Adenine (DA), b) Guanine (DG), c) Cytosine (DC), d) Thymine (DT). Color code: nitrogen - blue, carbon - light-blue, hydrogen - white, oxigen - red, phosphorus - gold.

We performed DFT calculations of the isolated bases (see Figure 3.3). For all the bases we can see that the largest single-atom contribution to the HOMO states comes from oxigen from the backbone[11] whereas the LUMO - is mostly carbon from the nucleobase. We also note that all of them show a very similar HOMO - LUMO gap.

---

[11] We also note a strong contribution from the base itself. In otherwords as seen in Figure 3.3, the HOMO is mostly delocalized over the entire molecule.

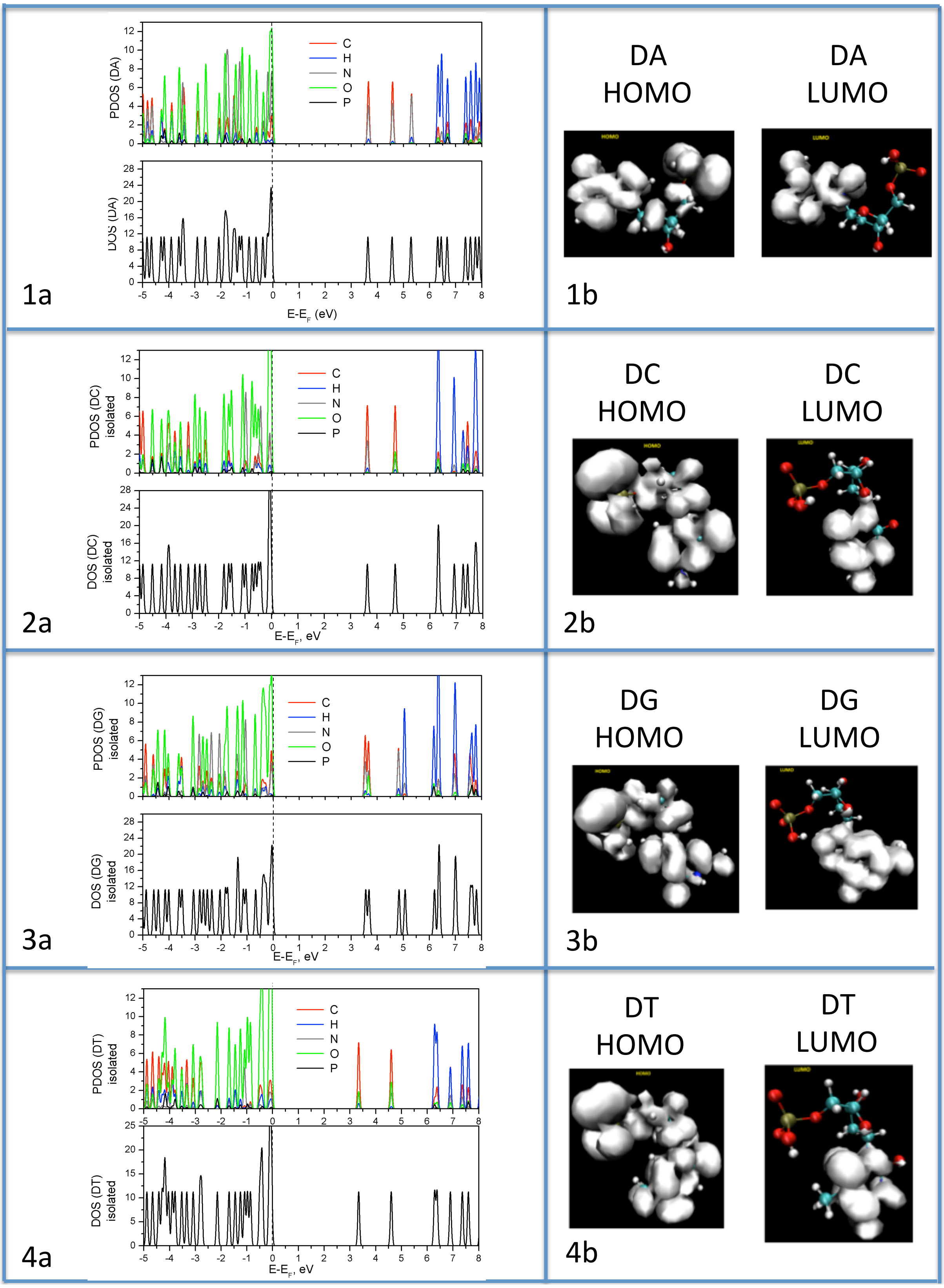


Figure 3.3: Projected Density of States on each atom (PDOS) and Total Density of State (DOS) of the nucleobases: 1a) DA base, 2a) DC base, 3a) DG base, 4a) DT base. Electronic Density, HOMO and LUMO, of the nucleobases: 1b) DA base, 2b) DC base, 3b) DG base, 4b) DT base. Color code: nitrogen - blue, carbon - light-blue, hydrogen - white, oxigen - red, phosphorus - gold. Value of the isosurface used: 0.0016.

3.3 Electronic structure of graphene

We performed DFT calculations for the graphene sheet without a nanopore. The size of graphene sheet was $51.9\AA$ in length and $42.6\AA$ wide, box size $50X42.6X51.9\AA$. The number of κ-points chosen 4. We plot total DOS of pure graphene sheet (see Figure 3.4). We can see that the graph shows the V-shape feature expected for graphene [5].

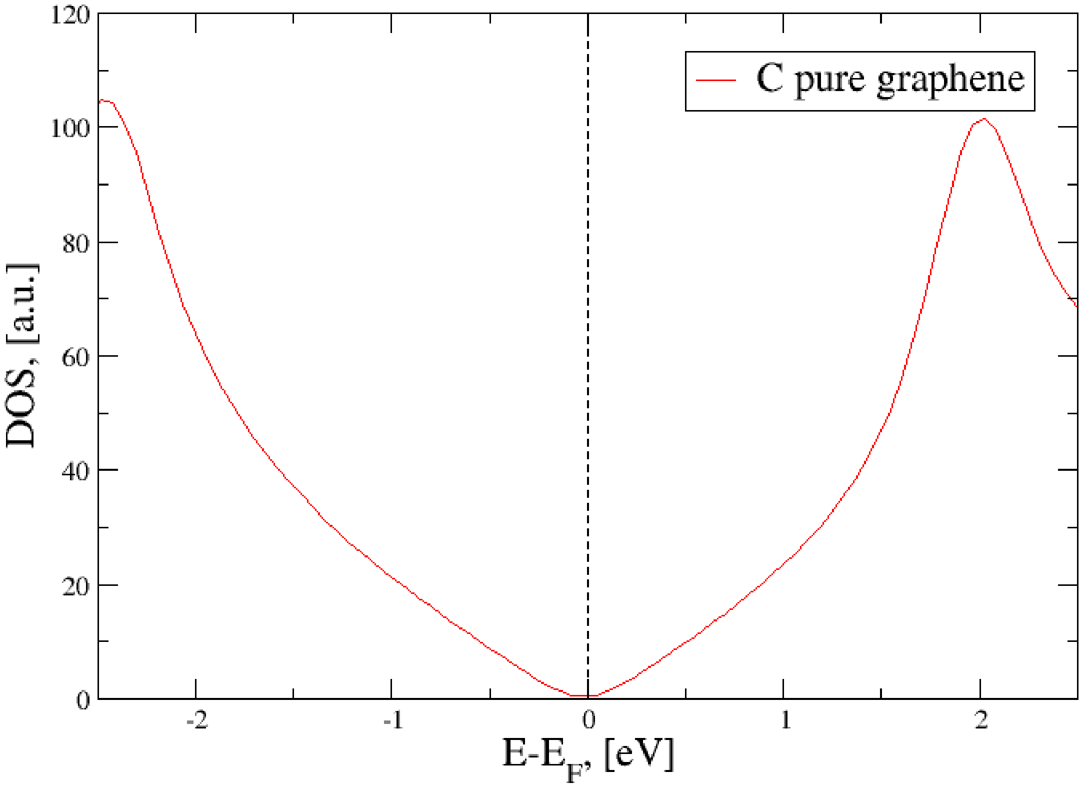


Figure 3.4: Electronic structure of the pure graphene sheet Total Density of States (DOS).

3.4 Electronic structure of the pores

We consider nanopores in isolated graphene sheet $51.9X42.6\AA$ in size with different terminations: hydrogen or nitrogen. In both cases we consider zigzag edge as those are the lowest energy edge terminations in graphene structures. The nanopores that we used for our simulations are shown in the Figure 3.5. In experiments the typical nanopore is approx. $20\AA$ due to the specificity of the nanopore fabrication (one particular fabrication method is described in [7]). There graphene is grown by chemical vapor deposition of methane over polished copper foils. The copper foils are etched in solution so that bare graphene sheets, approximately 1-5 nm thick, float on the surface of the liquid. Suitable sized graphene is then scooped onto prepatterned silicon nitride membranes. Nanopores are then drilled through the suspended graphene membranes by transmission electron beam ablation lithography [7].

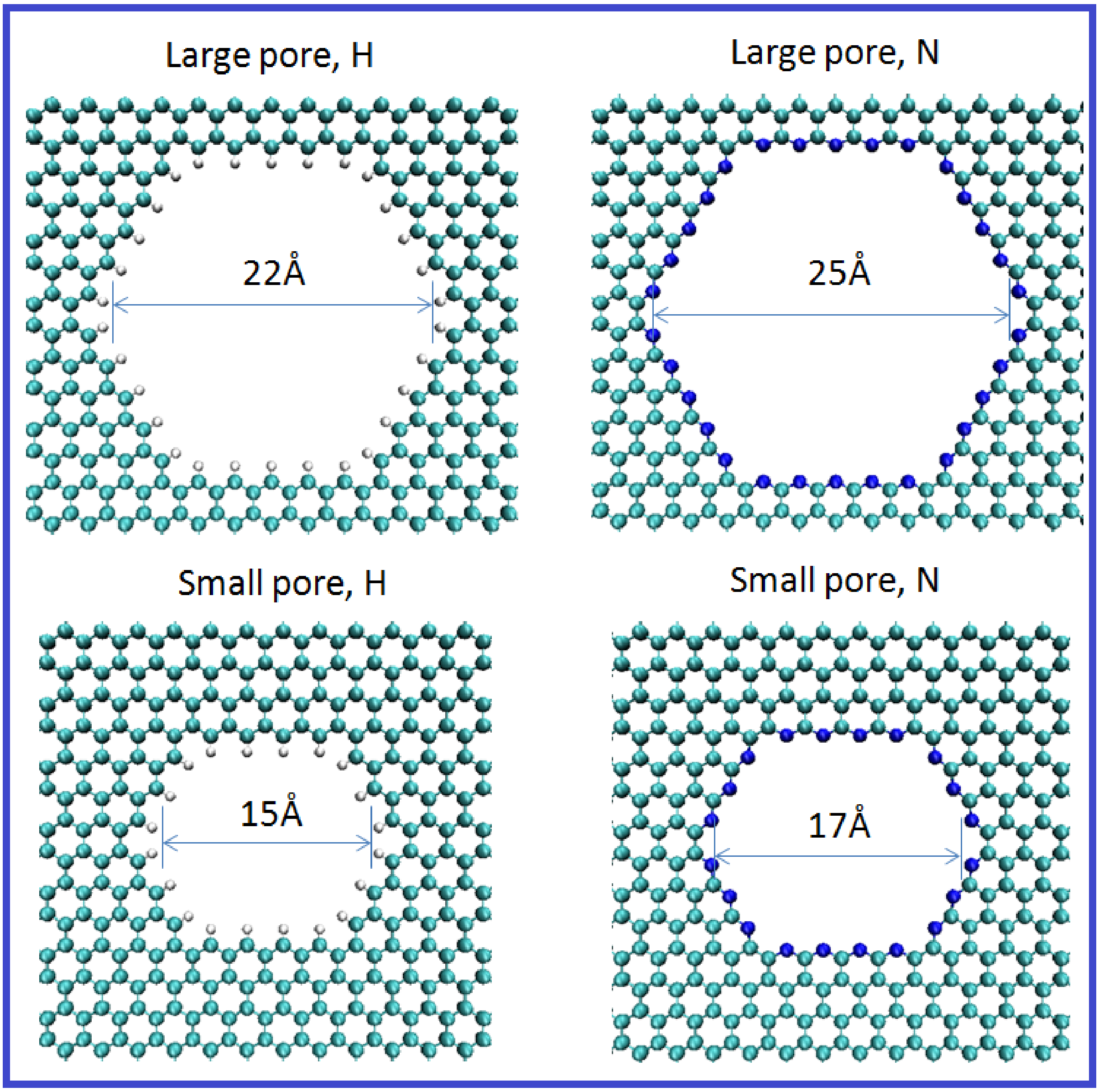


Figure 3.5: Types of pores, used for calculations. Color code: nitrogen - blue, carbon - light-blue, hydrogen - white.

Figure 3.6 shows the DOS resulting from our DFT calculations. We see that the system shows the V-shape features of graphene DOS. At the same time there are a number of resonances coming from defect states (both the absence of C atoms and the presence of either N or H). The hydrogen states are far from the Fermi level. Nitrogen states on the other hand are close to the Fermi level. It can be explained by the fact that in case of H-pore main contributaion to the Fermi level comes from the p-orbitals of the carbon (contribution from hidrogen is due to the s-orbitals, that is weaker). While in case of

N-pore p-orbitals of nitrogen interact with p-orbitals of carbon both contributing to the Fermi level. Hence, we might infer that when we put bases in the pore, they should interfere more with nitrogen than with hydrogen.

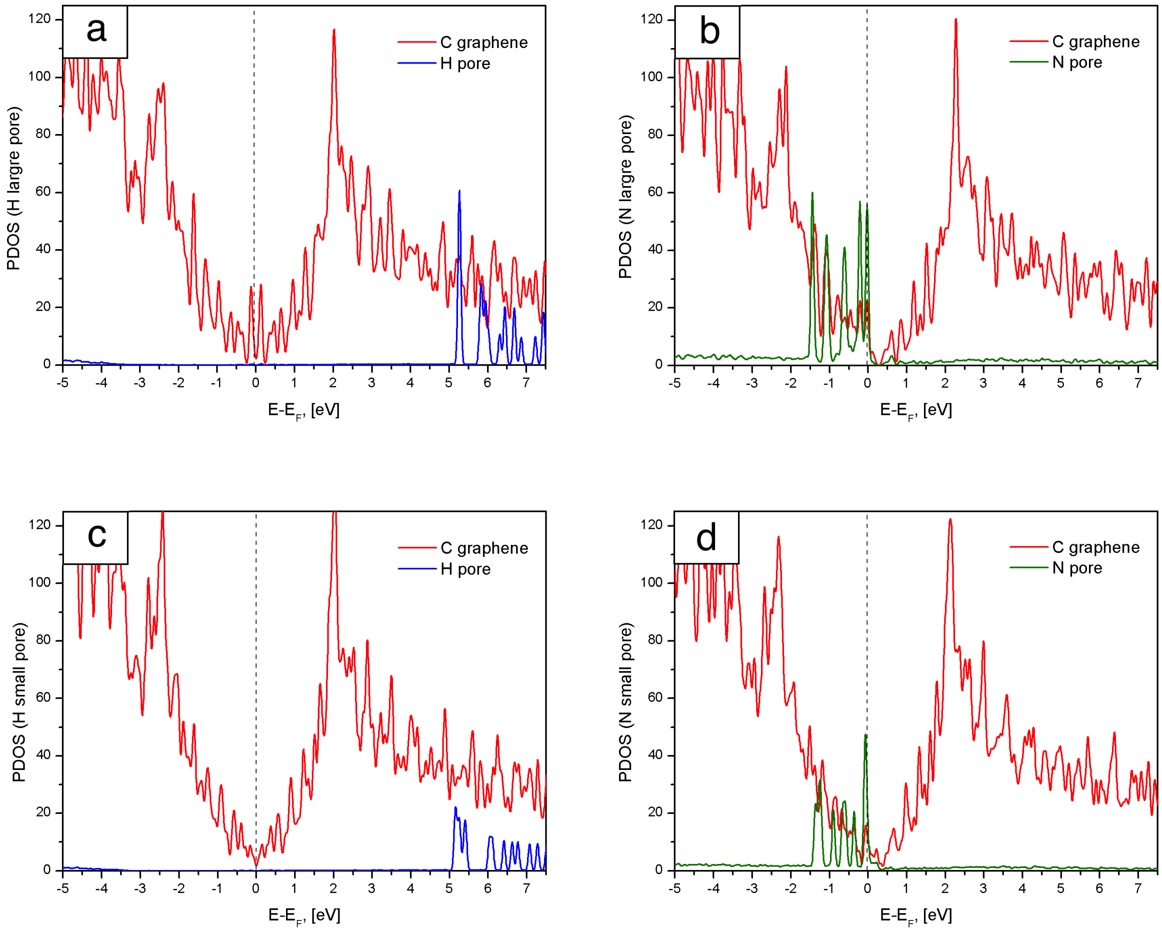


Figure 3.6: Electronic structure of the nanopores (Projected Density of States (PDOS)). a) large pore with H termination, b) large pore with N termination, c) small pore with H termination, d) small pore with N termination.

3.5 Pore + base: molecular dynamics and electronic structure

Following this initial study we placed bases inside pores and, using our simulation protocol, performed molecular dynamics calculations. The first thing we wanted to adress was wether the different pore terminations and different pore sizes might have some effect on the dynamics of the base. Most importantly wether one of them might help stabilize the base. Typical system for molecular dynamic simulations is shown in the Figure 3.7.

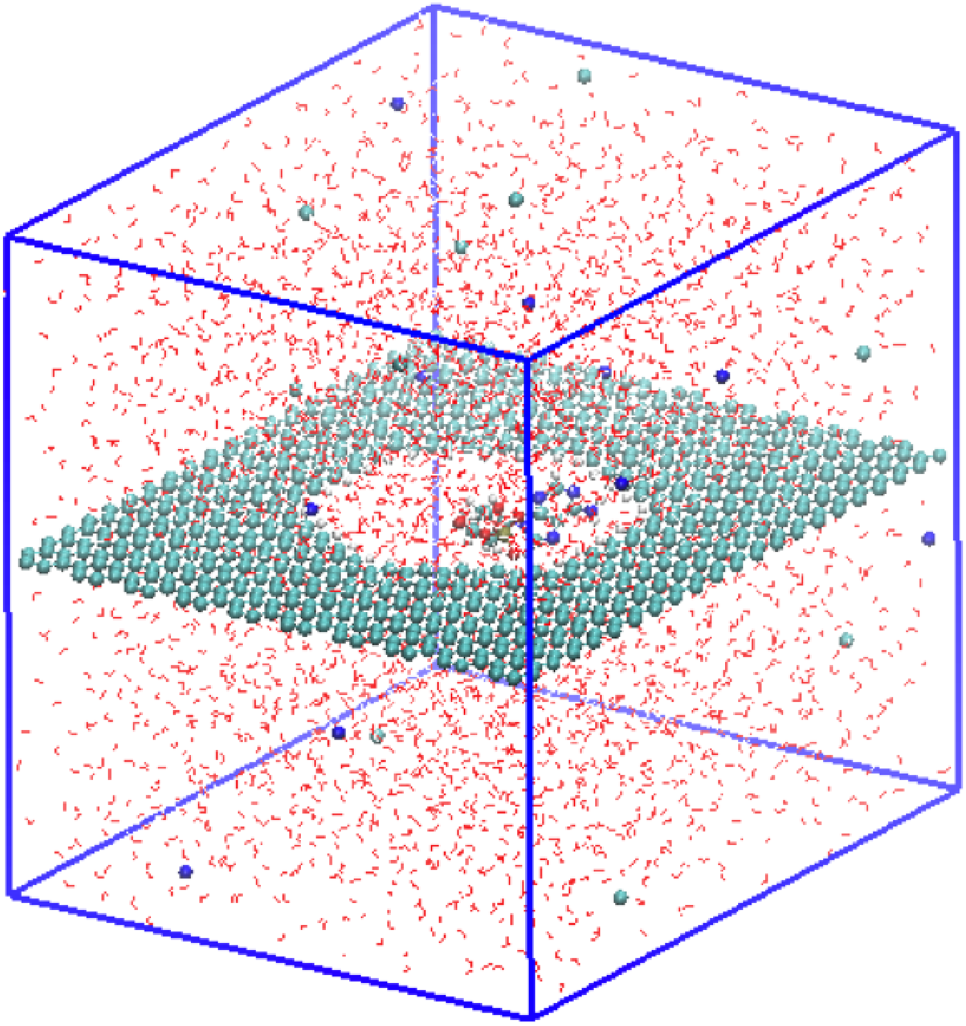


Figure 3.7: Typical system for molecular dynamics simulations.

3.5.1 Bases behaviour in the different types of pore

To analyze the bases behaviour inside the pores the angle between the base and pore plane were measured (Figure 3.8). Results for each base and different pores are presented in the Figure 3.9.

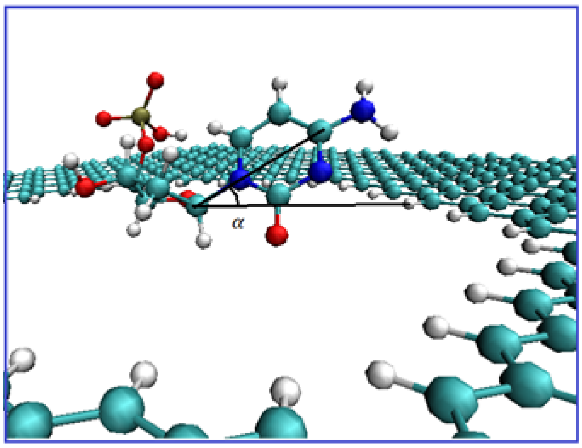


Figure 3.8: Angle between the base and pore plane. Color code: nitrogen - blue, carbon - light-blue, hydrogen - white, oxigen - red, phosphorus - gold.

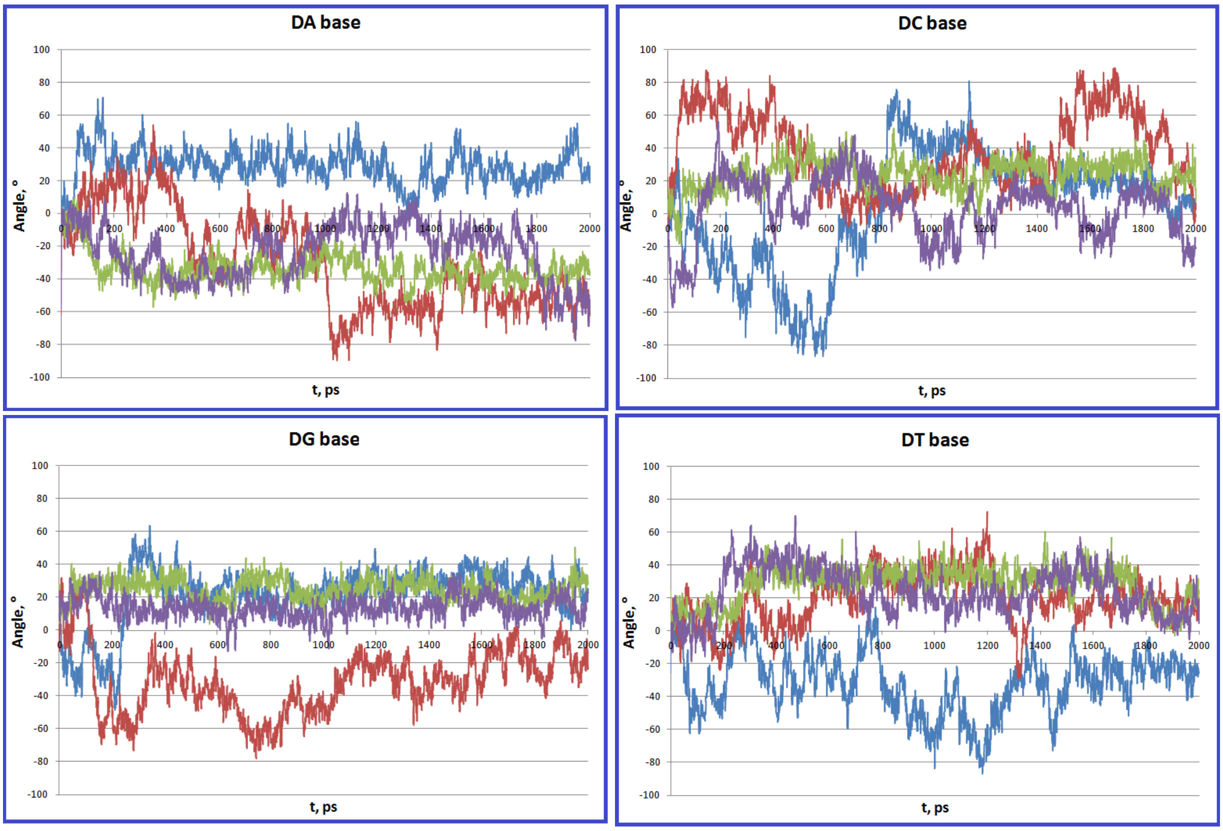


Figure 3.9: Angle between the base and pore plane for different pores and bases. Blue line - large pore H, red line - large pore N, green line - small pore H, purple - small pore N

From Figure 3.9 we can come to the conclusion that, in the large pores, bases have a tendency towards stronger angular movements while in the small pores the bases behaviour is more stable. That is expected as the interaction with the edges is smaller the bases have more freedom of movement within the pore. Even though nitrogen has a larger charge, the pore don't seems to improve the stabilization of the base. It must be noted that if we had the backbone of the DNA in our simulations, the freedom of the different bases would be smaller.

3.6 Electronic structure of the system: pore and base

In order to understand the effects of the dynamics on the electronic structure of the system we selected at random six snapshots from the classical dynamics. Those snapshots are shown in the Figures 3.10, 3.11, 3.12, 3.13 for each of our choices of pore termination and pore size.

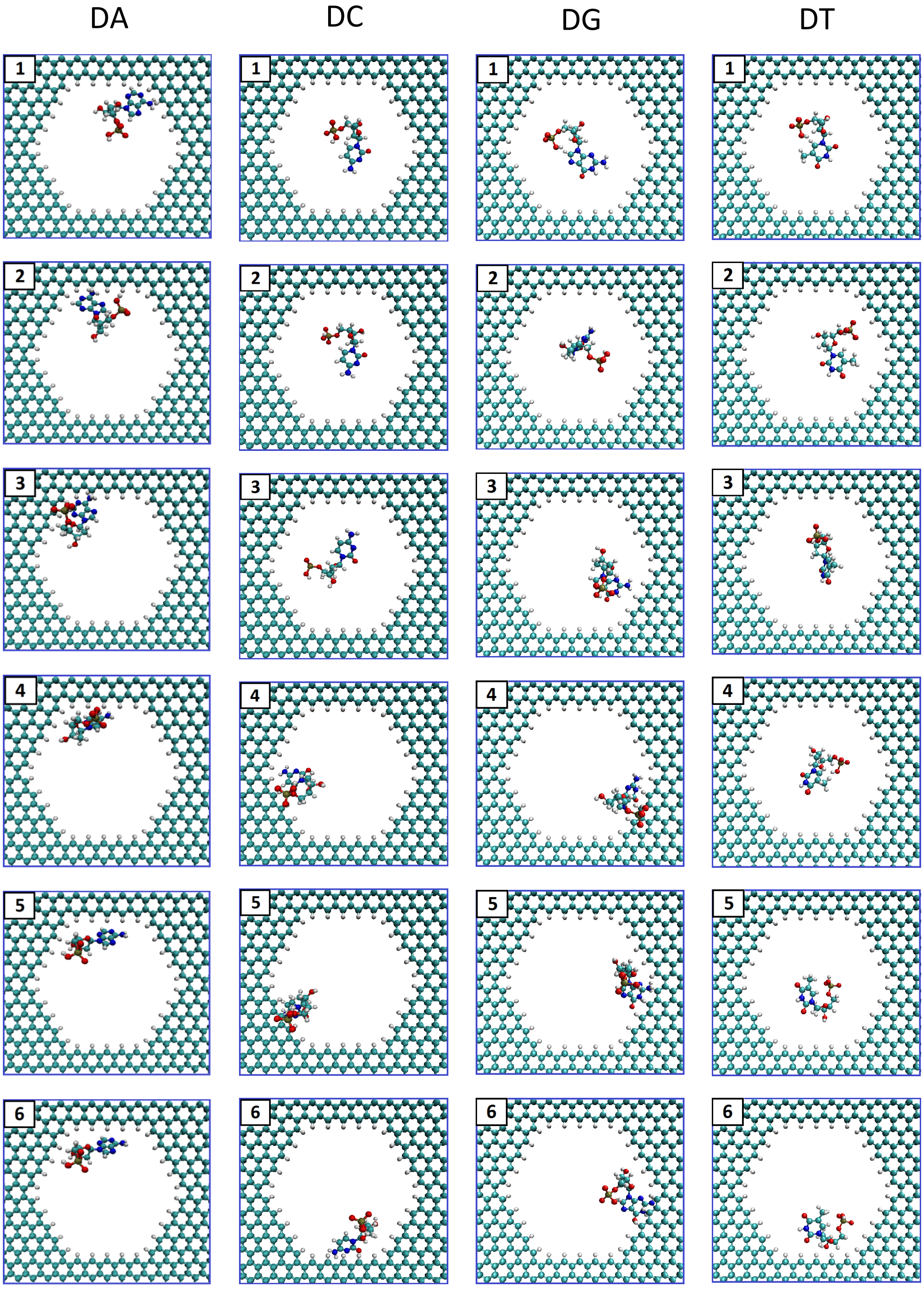


Figure 3.10: Bases orientation correspondes to the randomly chosen snapshorts for the system with large pore, hydrogen termination. Color code: nitrogen - blue, carbon - light-blue, hydrogen - white, oxigen - red, phosphorus - gold.

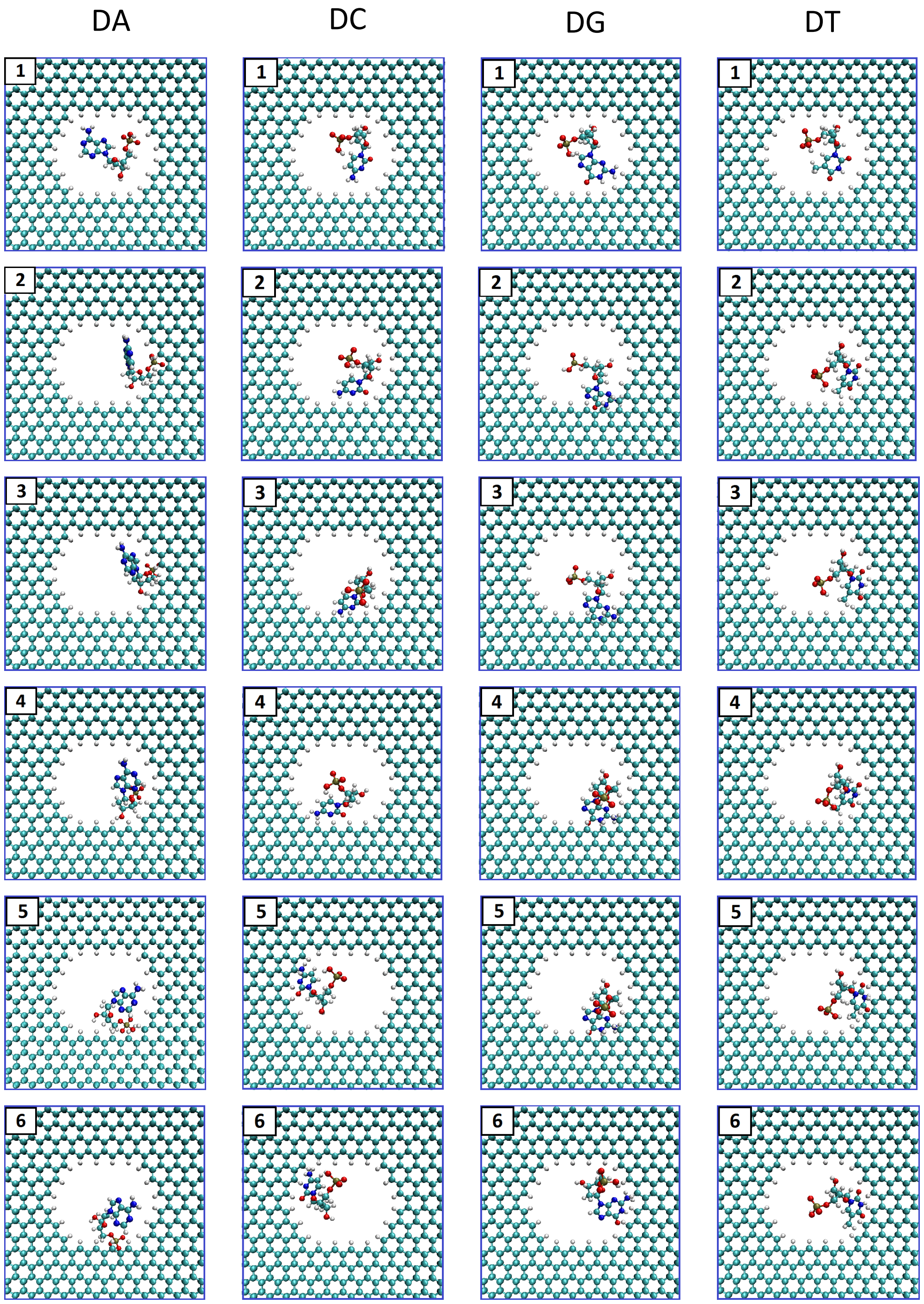


Figure 3.11: Bases orientation correspondes to the randomly chosen snapshorts for the system with small pore, hydrogen termination. Color code: nitrogen - blue, carbon - light-blue, hydrogen - white, oxigen - red, phosphorus - gold.

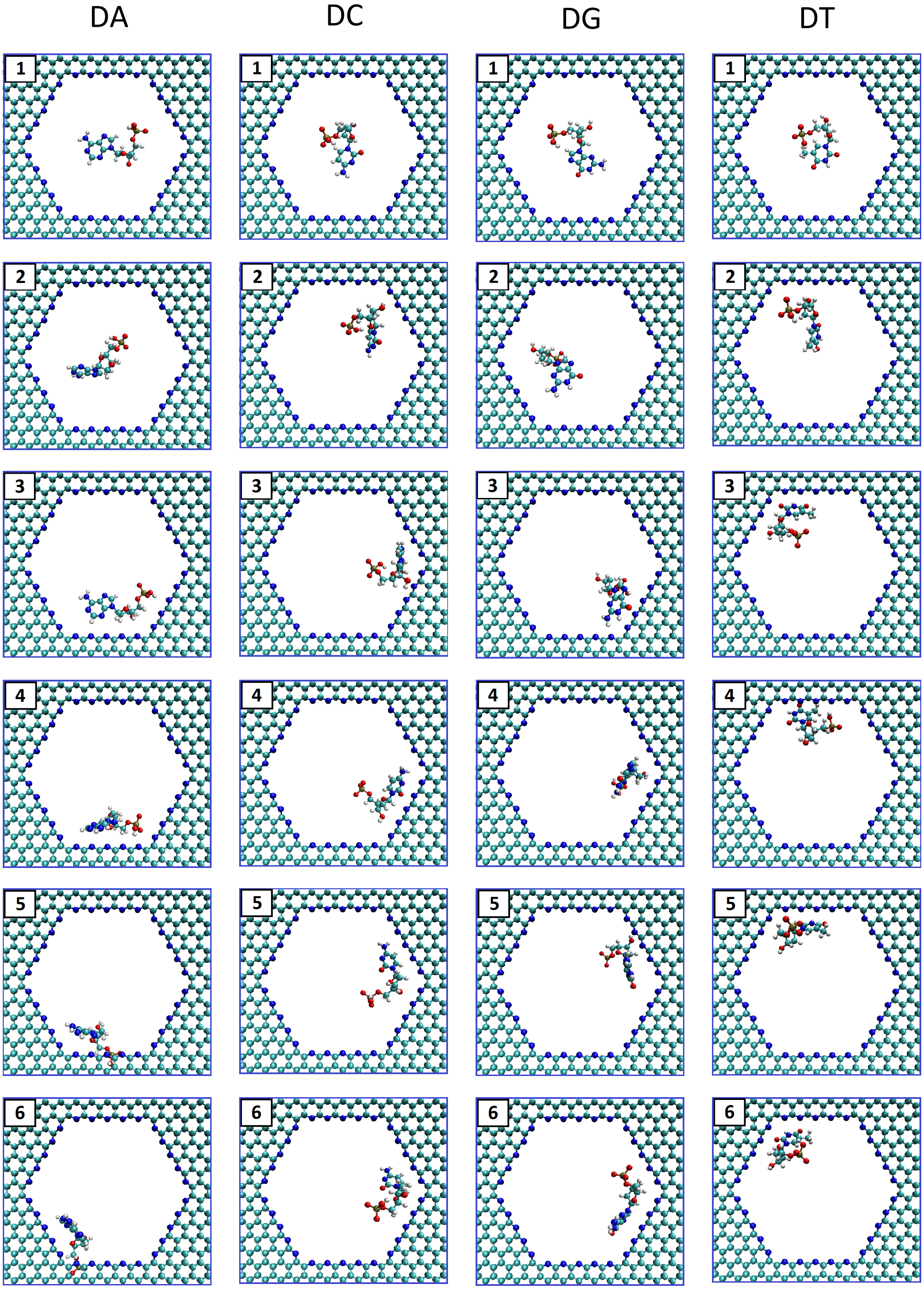


Figure 3.12: Bases orientation correspondes to the randomly chosen snapshorts for the system with large pore, nitrogen termination. Color code: nitrogen - blue, carbon - light-blue, hydrogen - white, oxigen - red, phosphorus - gold.

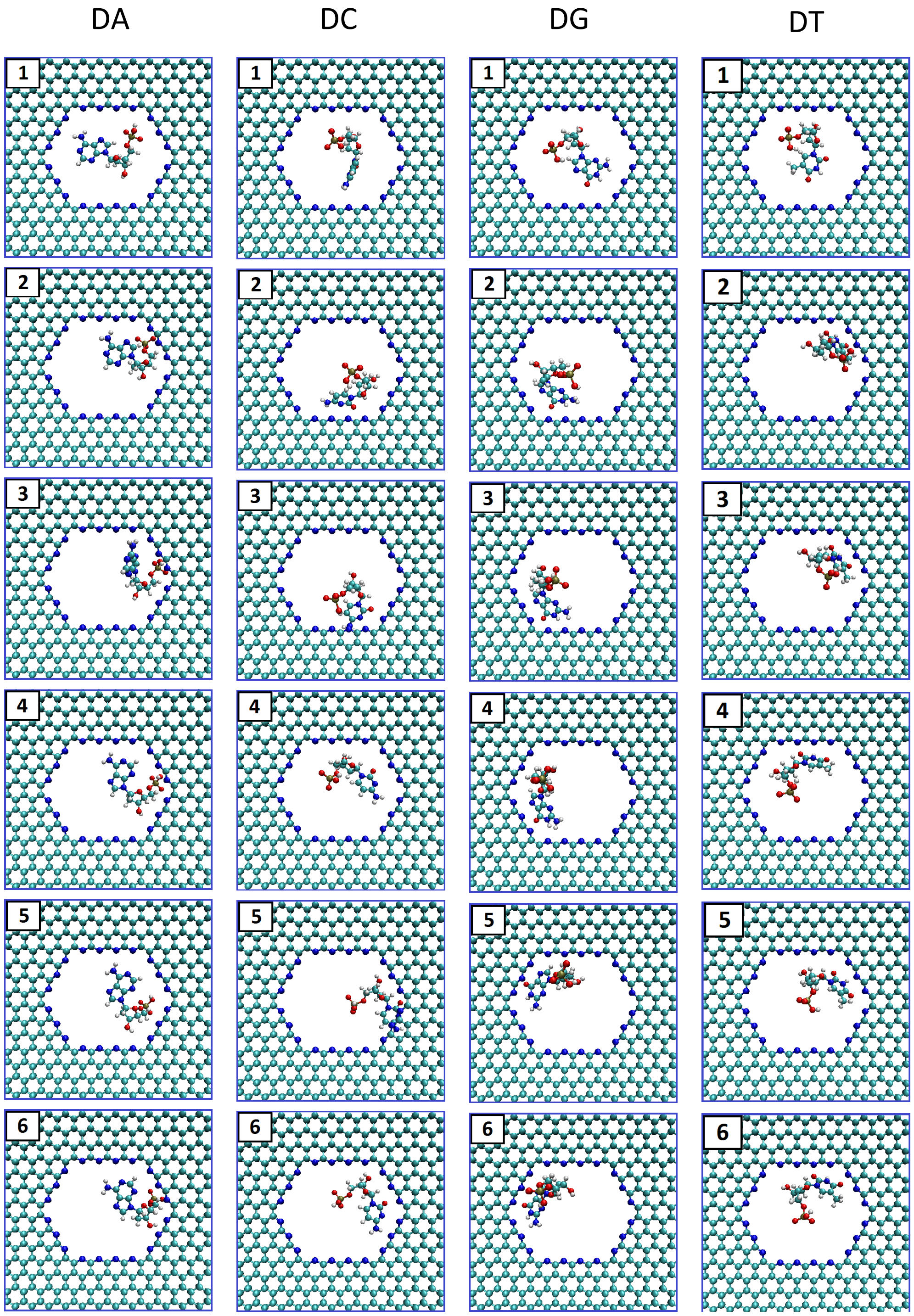


Figure 3.13: Bases orientation correspondes to the randomly chosen snapshorts for the system with small pore, nitrogen termination. Color code: nitrogen - blue, carbon - light-blue, hydrogen - white, oxigen - red, phosphorus - gold.

We have performed the DFT calculations for the system not including water. This was done by taking the classical dynamics structures and removing all water molecules and counter-ions, but retaining the structure of the base. Figures 3.14, 3.15, 3.16, 3.17 show density of states of the bases in the large pore with hydrogen termination, small pore with hydrogen termination, large pore with nitrogen termination and small pore with nitrogen termination.

PDOS for all the systems for the bases show the smaller energy gap ($\approx 1eV$ difference) in comparison with the isolated bases.

Let's first compare large pores with the small pores. Figures 3.14 and 3.15 for the pores with H terminations show us that the energy gap is the same for the large pore with H termination and small pore with H termination. While the energy states are more splitted and spread for the system with small pore. The same picture is observed for the systems with pores with N termination (Figures 3.16 and 3.17).

Now let's compare pores with H termination to the pores with N terminaions. For all the systems we can see that the energy gap is smaller for the pores with H termination. HOMO remains the same, while LUMO is shifted. Energy states for the pores with H termination are splitted more than for the pores with N termination.

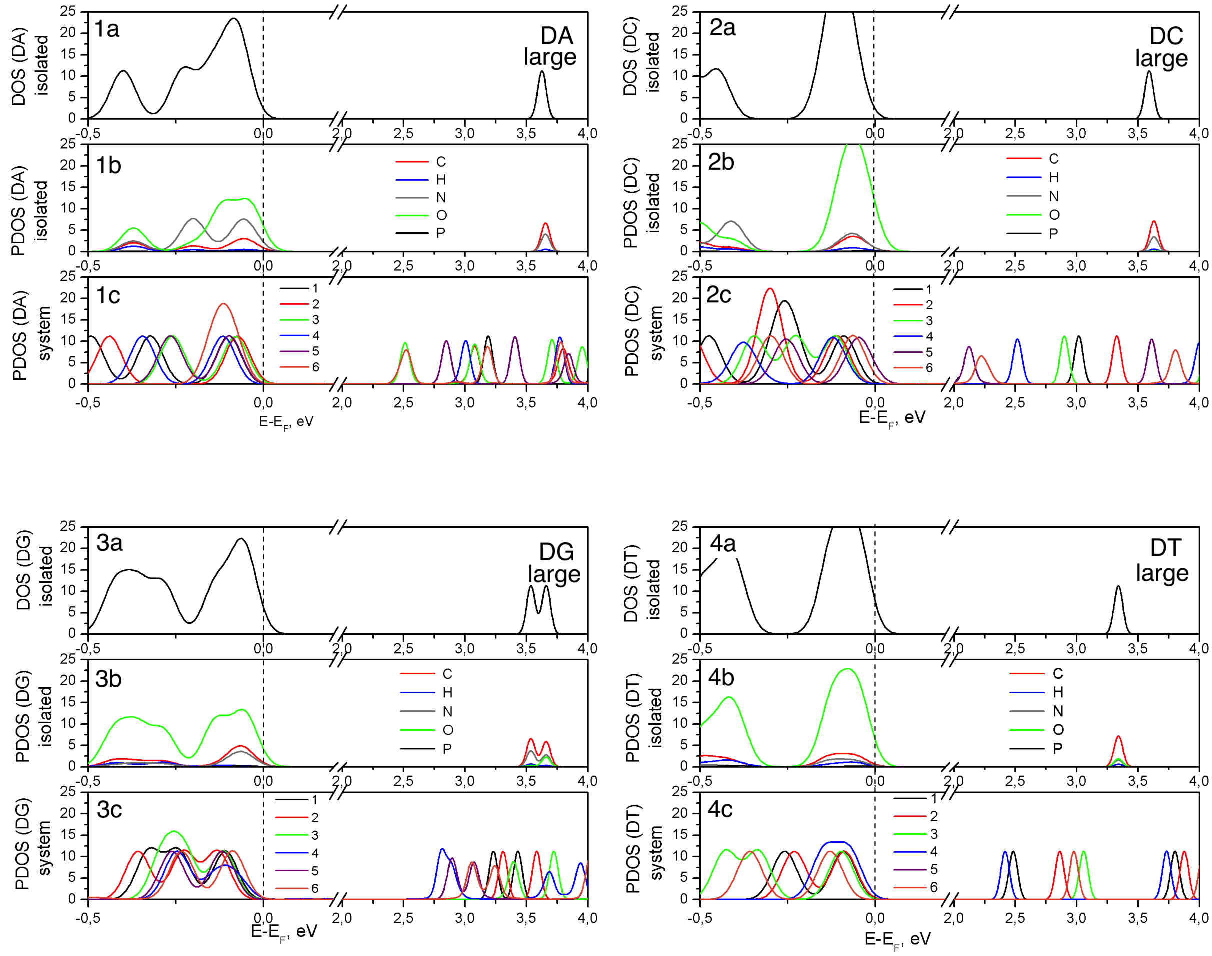


Figure 3.14: Electronic structure of the bases in the system (H, large pore, bases) comparing to the electronic structure of the isolated bases. 1) DA, 2) DC, 3) DG, 4) DT. a) Total Density of States of the isolated bases, b) Projected Density of States to each atom of the isolated bases, c) Projected density of states of the system to the bases.

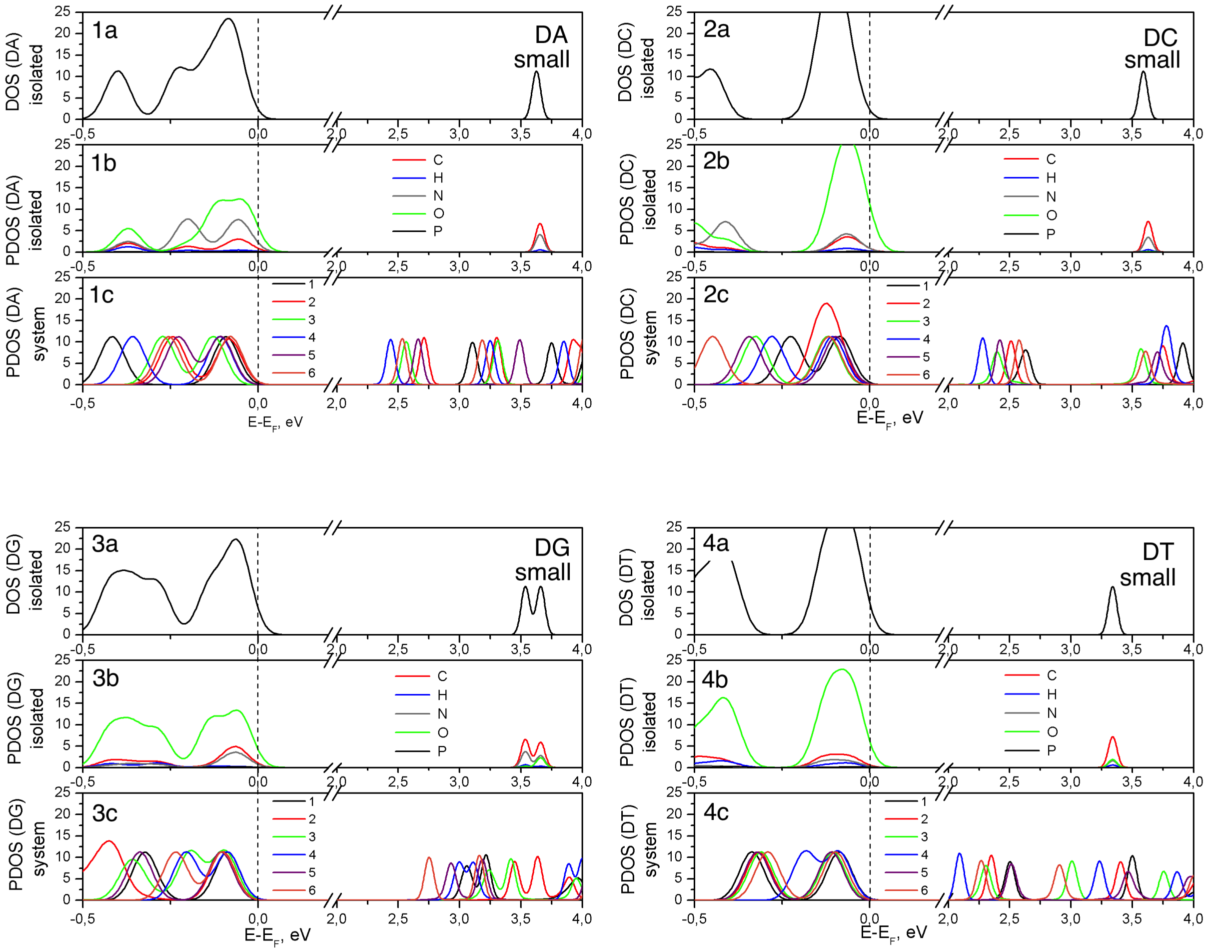


Figure 3.15: Electronic structure of the bases in the system (H, small pore, bases) comparing to the electronic structure of the isolated bases. 1) DA, 2) DC, 3) DG, 4) DT. a) Total Density of States of the isolated bases, b) Projected Density of States to each atom of the isolated bases, c) Projected density of states of the system to the bases.

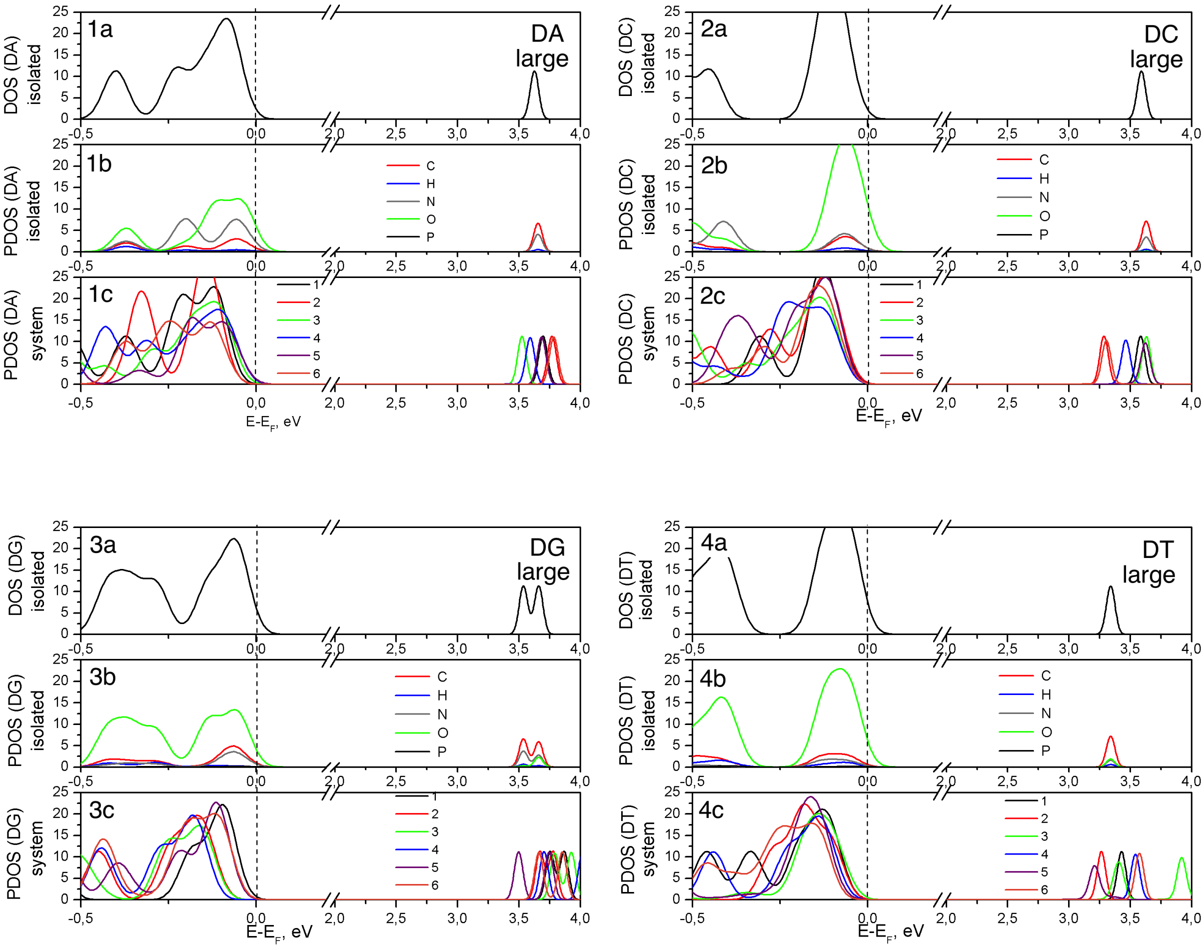


Figure 3.16: Electronic structure of the bases in the system (N, large pore, bases) comparing to the electronic structure of the isolated bases. 1) DA, 2) DC, 3) DG, 4) DT. a) Total Density of States of the isolated bases, b) Projected Density of States to each atom of the isolated bases, c) Projected density of states of the system to the bases.

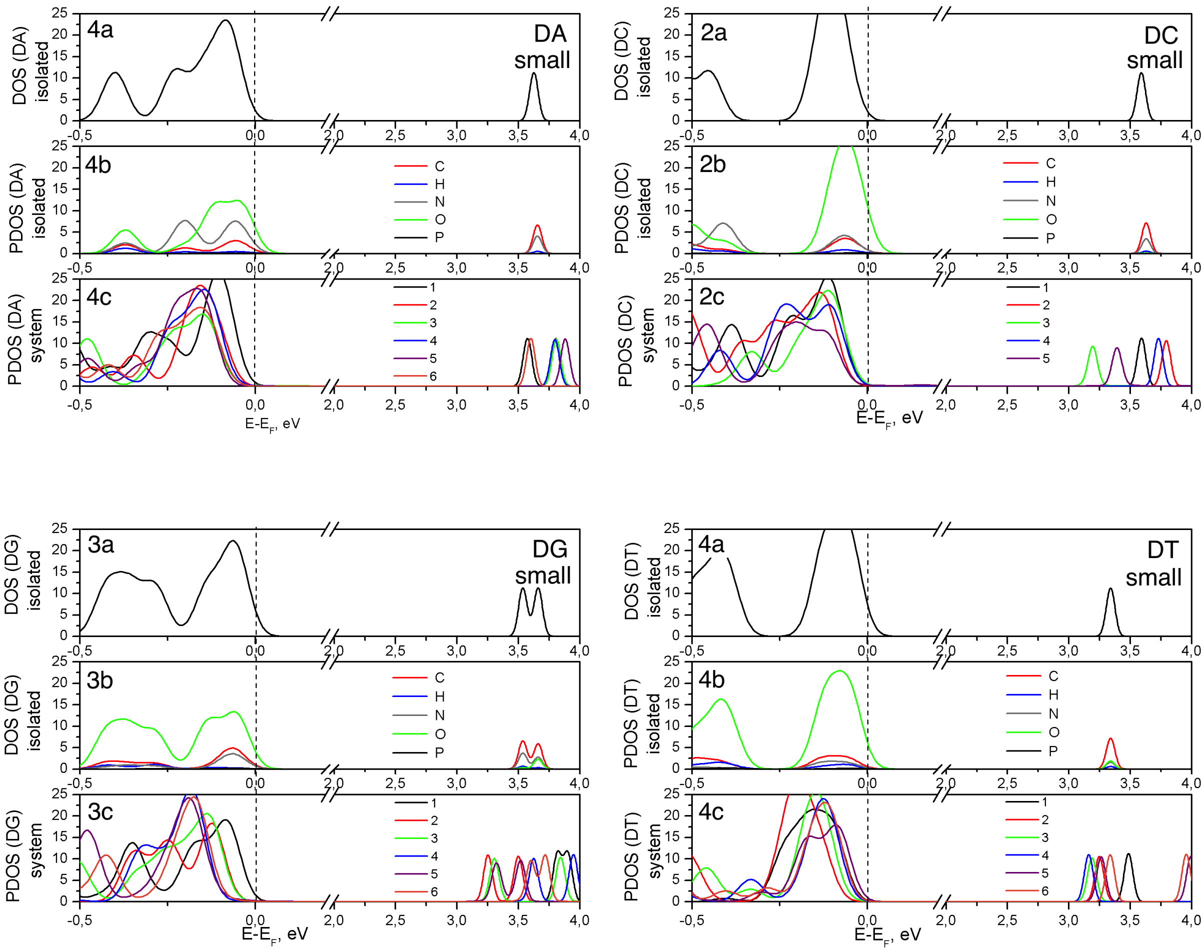


Figure 3.17: Electronic structure of the bases in the system (N, small pore, bases) comparing to the electronic structure of the isolated bases. 1) DA, 2) DC, 3) DG, 4) DT. a) Total Density of States of the isolated bases, b) Projected Density of States to each atom of the isolated bases, c) Projected density of states of the system to the bases.

We also plot the charges of the bases for this systems. Comparing Figures 3.18(1a) with 3.18(2a) and Figures 3.18(1b) and 3.18(2b) we can clearly see that the charges of the bases in the pores with H termination are larger than the charges of the bases in the pores with N termination.

Summarize this section, for the systems without the water and ions the bases interact more with the pores with H termination than with the pores with N termination. However, there are no significant differences for large and small pores.

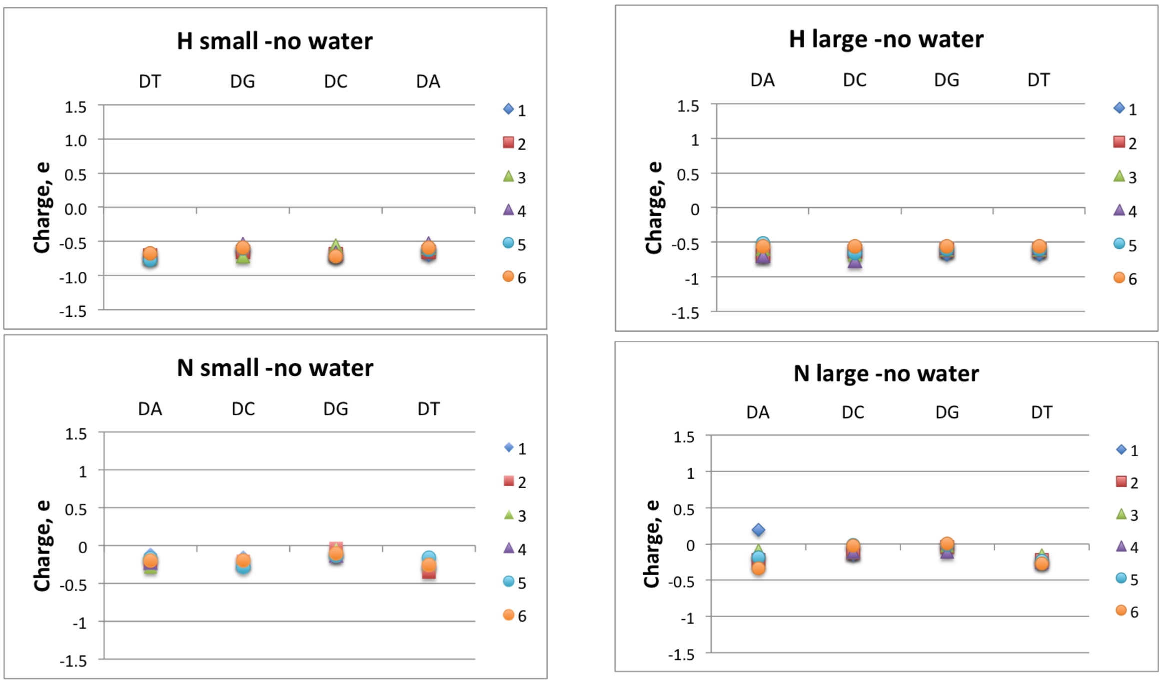


Figure 3.18: Charges on bases in the different systems for two types of pores for six snapshots. 1) Pores with H termination, 2) Pores with N termination. a) Small pores, b) Large pores.

3.7 Electronic structure of the system: small pore, base, one layer of water

Having noticed that there are no significant differences for large or small pores we analyzed the electronic structure of some of the system including one layer of QM water around the base for the small pores for six snapshots (see Figure 3.11 - 3.13). Here we chose ≈ 30 molecules of water around the base, choosing the molecules of water within $5\AA$ around bases. The DFT calculations of thes systems were done and are shown in Figures 3.19 and 3.20. We plot the PDOS of the base in the system without water and compare it to the PDOS of the base in the system including one layer of water.

Analyzing Figures 3.19 we see that presence of one layer of water around the base markedly affects the electronic structure of the base in the pore with hydrogen on the boarder. The states of the base in system including one layer of water are more spread than the states of the base in the system without water. The same time states are shifted in case of the system with water and the energy gap is bigger for each snapshots except of the first snapshot. In the first snapshots (see black curve in the Figure 3.19) the states stay close to the Fermi level and the energy gap decreases. The first snapshot can be different from others because during the molecular dynamics calculations we let the base move in the pore only on the last production step. On the previous steps bases are not allowed to move. First snapshot in our calculations is the very early step of the productions phase. As on the previous step we freeze the base, but allow water to move, the forces between base and atoms of water around the base can be artificially high. We will see further that the charges of the bases for the first snapshots are very different from the rest snapshots (see Figure 3.21). For this reason the results for the latest snapshots must be closer to the realistic bases behaviour.

Analyzing Figures 3.19 we see that presence of one layer of water around the base affects the electronic structure of the bases in the pore with nitrogen on the border. The states are more spread, but the energy gap does not change in contrast with the case of the pore with hydrogen on the border (see Figure 3.19).

In our calculation we notice that the presence of water leads to a significant decrease in the gap for the case of the hydrogen but only minor changes in the gap for nitrogen. On the other hand, the wider dispession of states, specially for occupied ones, indicates that the base hybridizes much more strongly for the N-pore that the case of hydrogen.

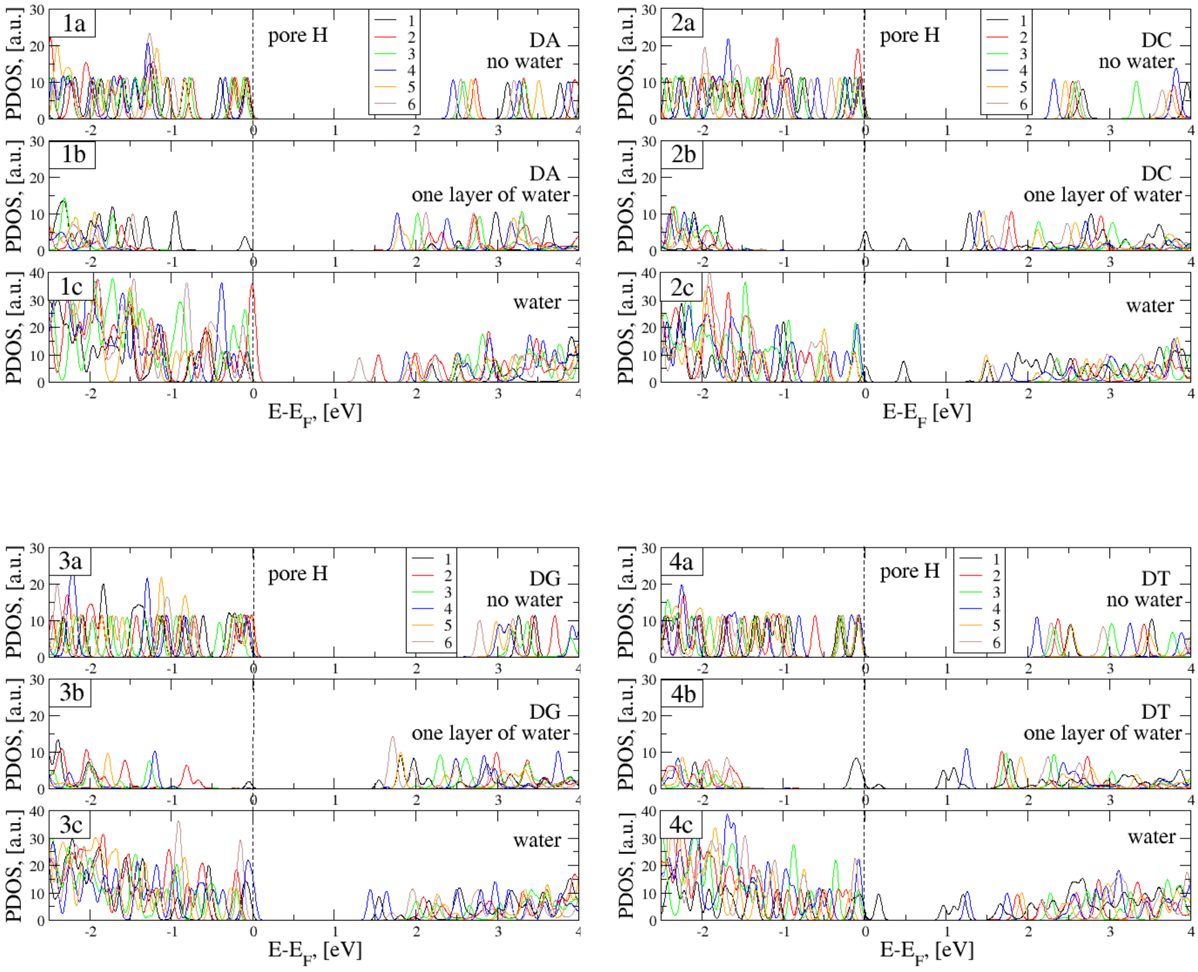


Figure 3.19: Electronic structure of the bases in the system (H small pore, bases, one layer of water) comparing to the electronic structure of the bases in the system (H small pore, bases, no water) and electronic structure of the one layer of water in the system (H small pore, one layer of water). 1) DA, 2) DC, 3) DG, 4) DT. a) Projected Density of States to the bases in the system (H small pore, no water), b) Projected Density of States to the bases in the system (H small pore, bases, one layer of water), c) Projected density of states to the water in the system (H small pore, bases, one layer of water.

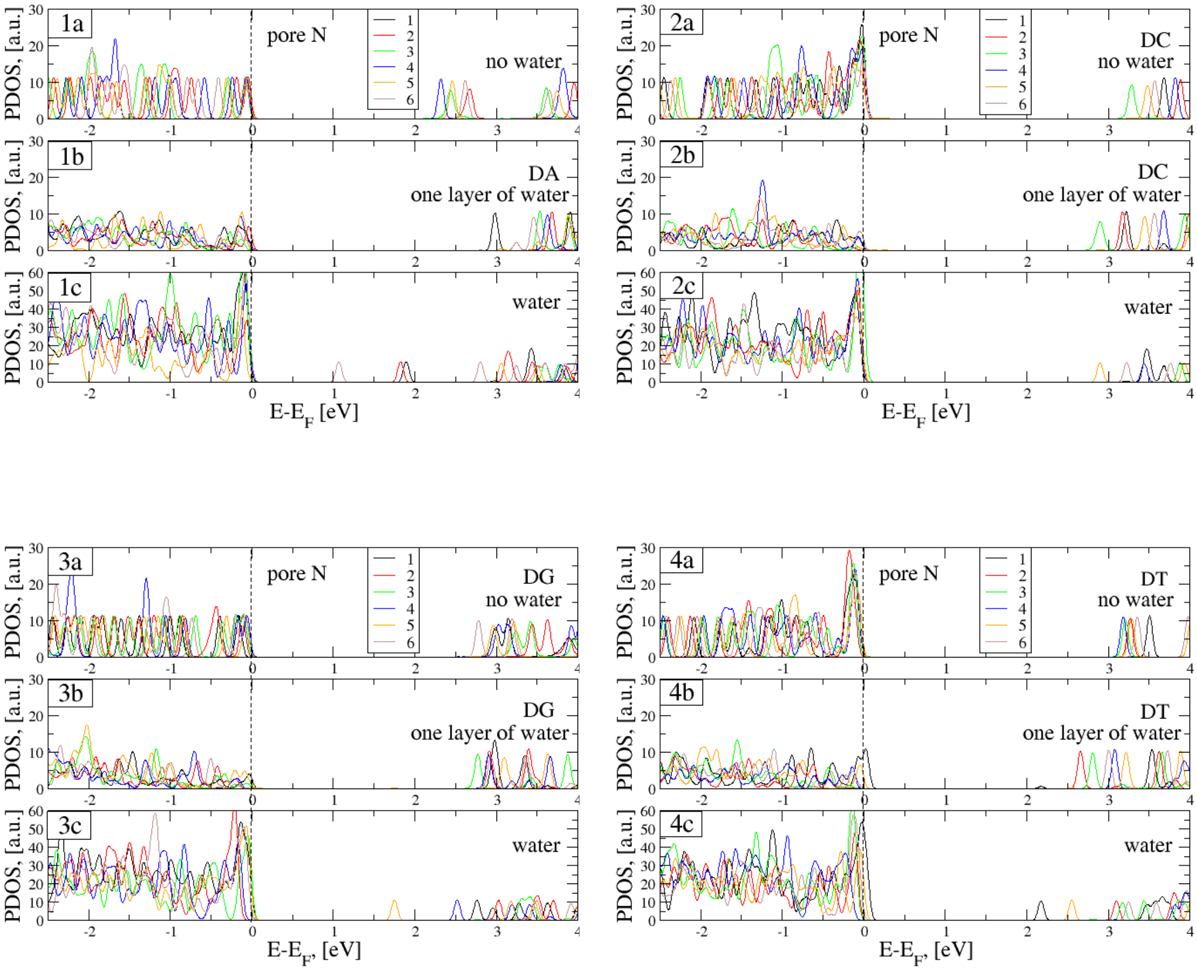


Figure 3.20: Electronic structure of the bases in the system (N small pore, bases, one layer of water) comparing to the electronic structure of the bases in the system (N small pore, bases, no water) and electronic structure of the one layer of water in the system (N small pore, one layer of water). 1) DA, 2) DC, 3) DG, 4) DT. a) Projected Density of States to the bases in the system (N small pore, no water), b) Projected Density of States to the bases in the system (N small pore, bases, one layer of water), c) Projected density of states to the water in the system (N small pore, bases, one layer of water.

We also plot the charges of the bases for this two systems. Figure 3.21 shows that the charge of the bases does not change when we include one layer of water for pore with nitrogen on the border (see Figure 3.21 (2a, 2b)). The nitrogen terminated pore does not favour charge transfer whereas the hydrogen terminated one does. This can be explained by the fact that the base tends to be negatively charged, and hydrogen atoms can easily donate charges to the base. On the other hand nitrogen is very electronegative, so it hinders the transfer of electrons to the base. In case of the pore with hydrogen on the border including one layer of the water decreases the charges of the base (see Figure 3.21 (1a, 1b)). Due to the higher electronegativity nitrogen has a tendency to do not releaze a charge to the water.

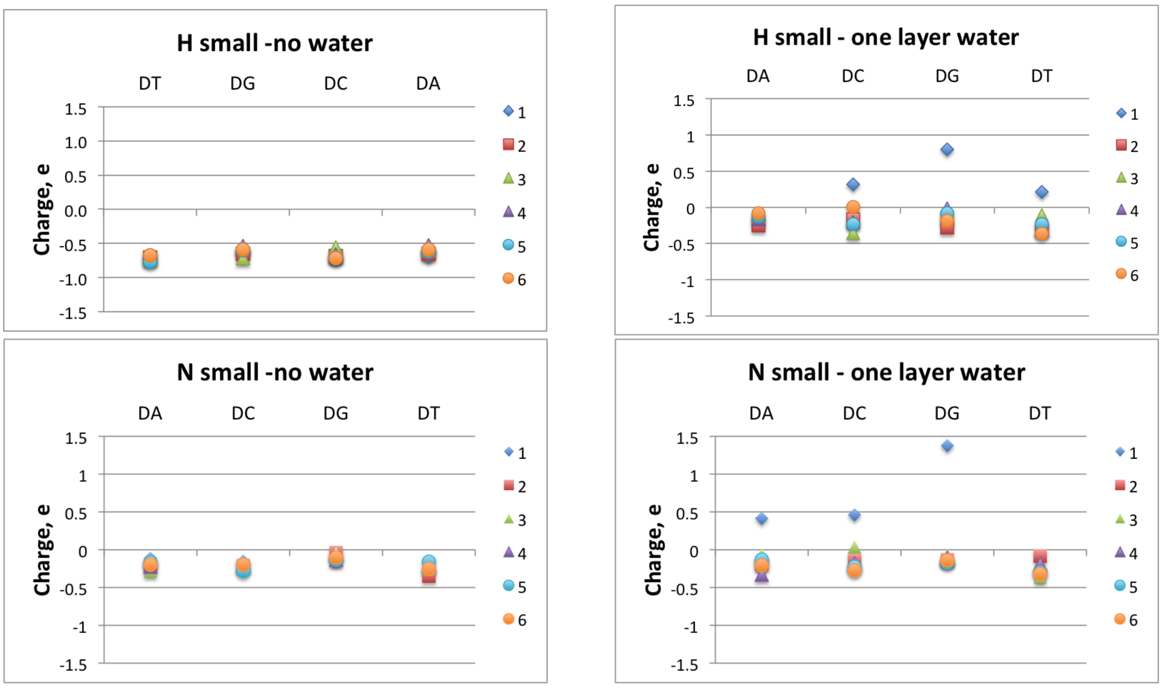


Figure 3.21: Charges of the bases in the different systems for two types of pores for six snapshots. 1) H small pore, 2) N small pore. a) Pore, base, no water, b) Pore, base, one layer of water around the base.

3.8 Electronic structure of the system: small pore, base, including effect of water and ions (QM/MM approach)

The final configuration that was considered was the effect of water and ions on the system using the QM/MM approach. DFT calculations were done for the systems with the small pores for six different snapshots and presented in the Figures 3.22, 3.23. We plot PDOS for the base (DA, DC, DG and DT) for the system without the water (see Figures 3.22a, 3.23a) and for the system including effect of water and ions (see Figures 3.22b, 3.23b).

For all the bases presence of the water and ions in calculations decreases the energy gap for the pore with hydrogen termination. While for the pore with nitrogen termination the energy gap remains the same. This is similar to the results considering a single layer of quantum mechanical water. For cases considered energy states are more spread in energy, including the effect of water and ions.

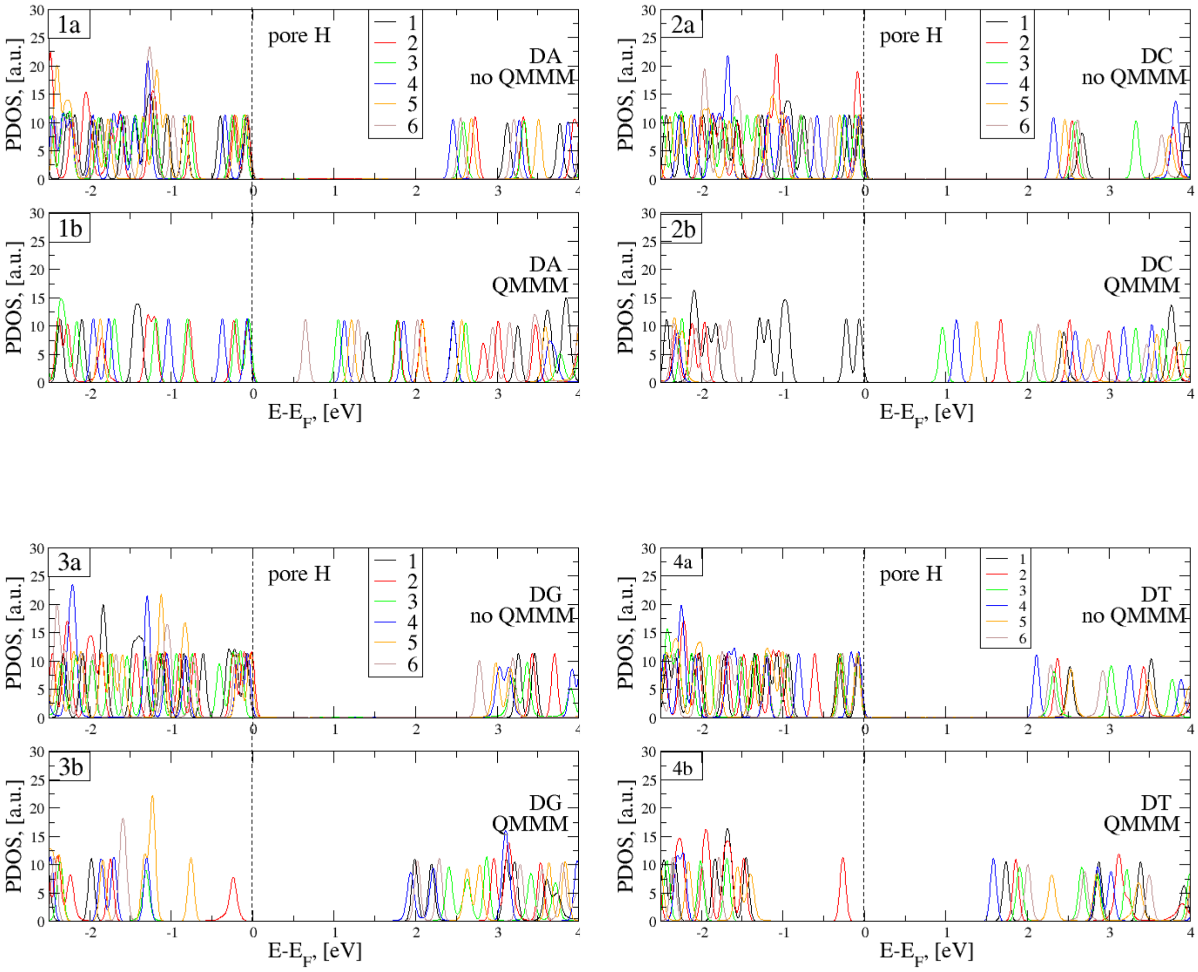


Figure 3.22: Electronic structure of the bases in the system (H, small pore, bases, including V_{ext} from atoms of water and ions) comparing to the electronic structure of the bases in the system (H, small pore, bases, not including V_{ext} from atoms of water and ions). 1) DA, 2) DC, 3) DG, 4) DT. a) Projected Density of States to the bases in the system (H small pore, including V_{ext} from atoms of water and ions), b) Projected Density of States to the bases in the system (H small pore, not including V_{ext} from atoms of water and ions).

We plot the charge on the bases for the system including effect of water and ions as well. Comparison of the charges in the system without water, including one layer of water and including effect of water and ions using the QM/MM approach is presented in the Figure 3.24. Values of the charges are presented in the Figure 3.25. Analyzing

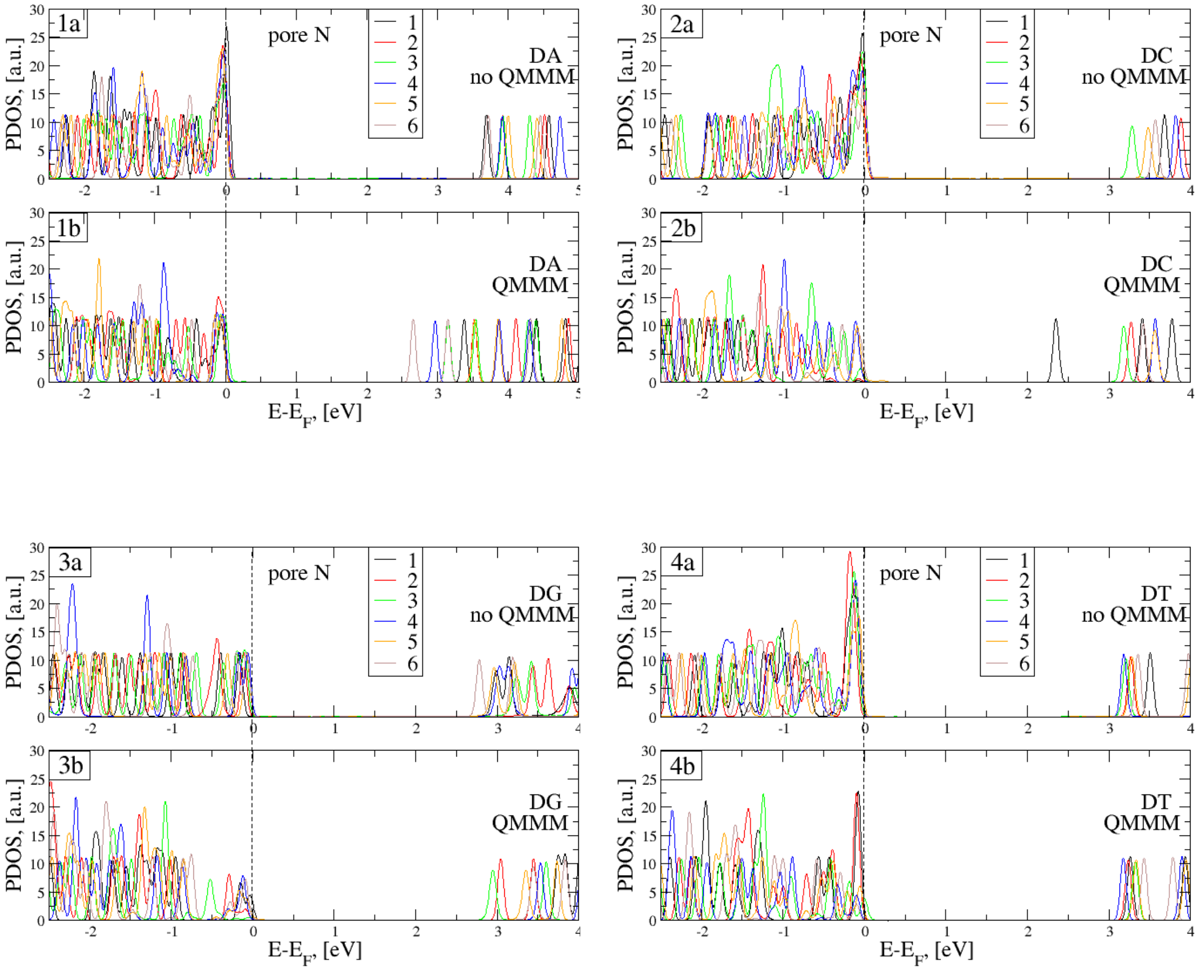


Figure 3.23: Electronic structure of the bases in the system (N small pore, bases, including V_{ext} from atoms of water and ions) comparing to the electronic structure of the bases in the system (N small pore, bases, not including V_{ext} from atoms of water and ions). 1) DA, 2) DC, 3) DG, 4) DT. a) Projected Density of States to the bases in the system (H small pore, including V_{ext} from atoms of water and ions), b) Projected Density of States to the bases in the system (N small pore, not including V_{ext} from atoms of water and ions).

Figures 3.24 and 3.25 we can come to conclusion, that including one layer of water treated quantum mechanically the same as including the effect of the water and ions using QM/MM approach makes significant change in describing electronic structure of the system. Therefore the future steps would be calculating the system: nanopore, base + one layer of water described quantum mechanically, while the rest of the water and ions included via QM/MM approach.

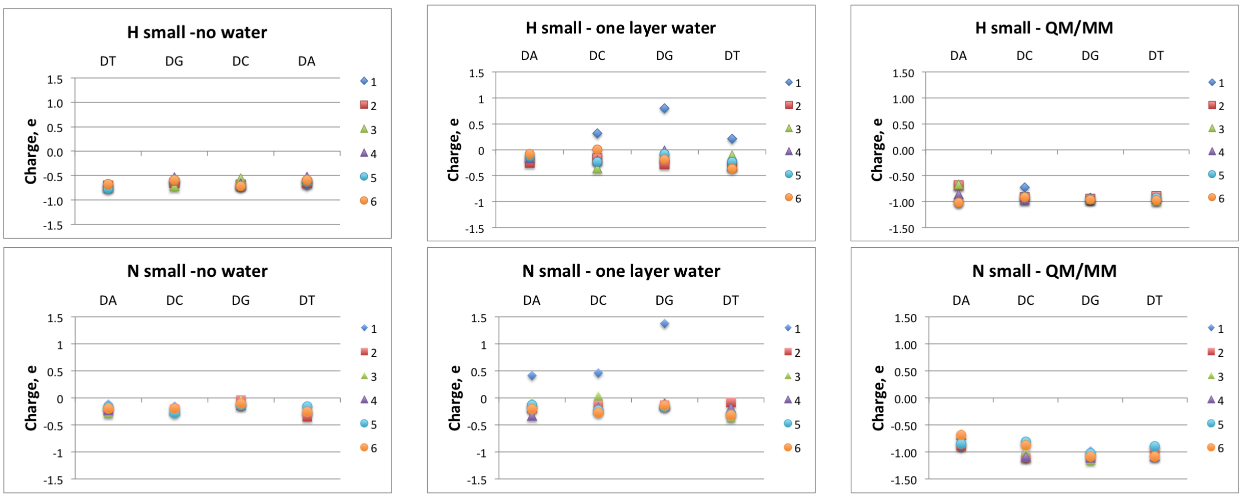


Figure 3.24: Charges of the bases in the different systems for two types of pores. 1) H small pore, 2) N small pore for six snapshots. a) Pore, base, no water, b) Pore, base, one layer of water around the base, c) Pore, base, including V_{ext} from atoms of water and ions (QM/MM approach).

DA	snapshots	no water	one layer water	QM/MM
H small	1	-0.70	-0.17	-0.94
	2	-0.66	-0.25	-0.69
	3	-0.59	-0.13	-0.66
	4	-0.52	-0.17	-0.86
	5	-0.62	-0.12	-1.01
	6	-0.59	-0.08	-1.02
N small	1	-0.11	0.42	-0.91
	2	-0.21	-0.19	-0.89
	3	-0.28	-0.07	-0.78
	4	-0.23	-0.33	-0.67
	5	-0.16	-0.13	-0.85
	6	-0.20	-0.20	-0.69

DC	snapshots	no water	one layer water	QM/MM
H small	1	-0.69	0.32	-0.73
	2	-0.69	-0.16	-0.92
	3	-0.57	-0.35	-0.95
	4	-0.66	-0.21	-0.96
	5	-0.71	-0.23	-0.90
	6	-0.72	0.00	-0.91
N small	1	-0.15	0.47	-1.11
	2	-0.21	-0.14	-1.10
	3	-0.21	0.03	-0.99
	4	-0.20	-0.18	-1.08
	5	-0.28	-0.23	-0.81
	6	-0.20	-0.27	-0.88

DG	snapshots	no water	one layer water	QM/MM
H small	1	-0.67	0.80	-0.92
	2	-0.65	-0.28	-0.94
	3	-0.72	-0.07	-0.95
	4	-0.54	-0.01	-0.94
	5	-0.60	-0.08	-0.95
	6	-0.59	-0.18	-0.96
N small	1	-0.06	1.37	-0.98
	2	-0.03	-0.13	-1.07
	3	-0.05	-0.07	-1.15
	4	-0.13	-0.11	-1.10
	5	-0.13	-0.17	-1.02
	6	-0.10	-0.14	-1.08

DT	snapshots	no water	one layer water	QM/MM
H small	1	-0.72	0.22	-0.94
	2	-0.70	-0.32	-0.90
	3	-0.73	-0.11	-0.98
	4	-0.69	-0.28	-0.95
	5	-0.76	-0.23	-0.92
	6	-0.67	-0.37	-0.97
N small	1	-0.28	-0.25	-0.90
	2	-0.35	-0.08	-0.98
	3	-0.24	-0.36	-1.05
	4	-0.19	-0.19	-1.08
	5	-0.16	-0.31	-0.90
	6	-0.25	-0.32	-1.08

Figure 3.25: Charges of the bases DA, DC, DG, DT in the different systems (Pore, base, no water; pore, base, one layer of water around the base; pore, base, including V_{ext} from atoms of water and ions (QM/MM approach)for two types of pores: H small pore and N small pore.

3.9 Electronic transport

Following our simulation protocol we performed calculations of the zero-bias transmission coefficients using the non-equilibrium Green's function. Figure 3.26 illustrates the conductance of the electrode (pristine graphene). For the electrode it was taken 120 atoms of carbon from the left and from the right of the graphene sheet. The number of κ-points was chosen 10.

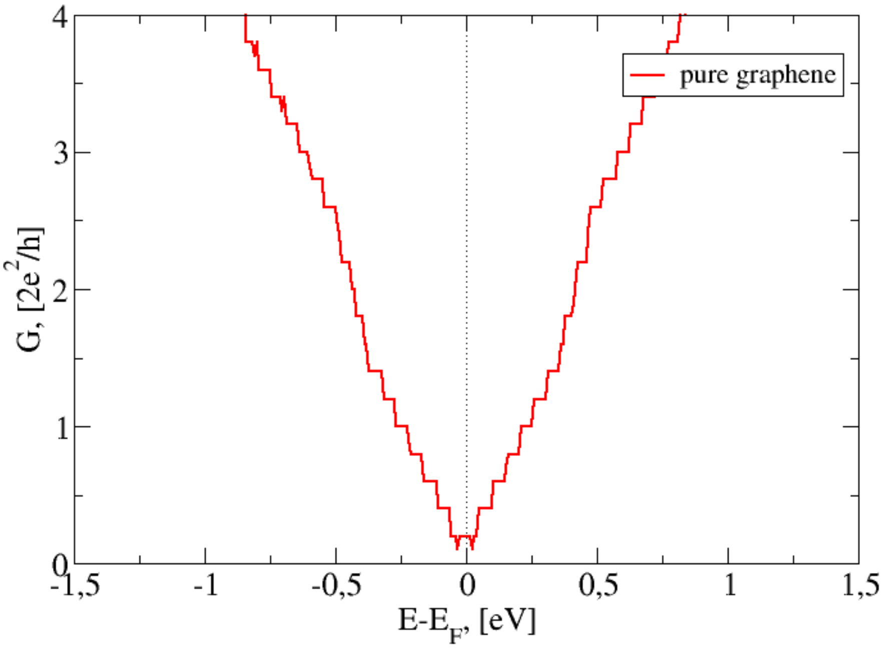


Figure 3.26: Conductance as a function of energy for the electrode (pristine graphene).

Figure 3.27 illustrates the conductance of the dry empty pore with different terminations a) hydrogen, b) nitrogen comparing them to the conductance of the electrode (pristine graphene). We can clearly see that the conductance for the electrode shows the V-shape feature expected for graphene. Introducing the pore to the graphene sheet significantly changes the conductance. Futhermore, the sharp features near 0.5 eV are present in the curve for the empty hydrogen terminated pore corresponding to the localized states due to the carbon of the pore.

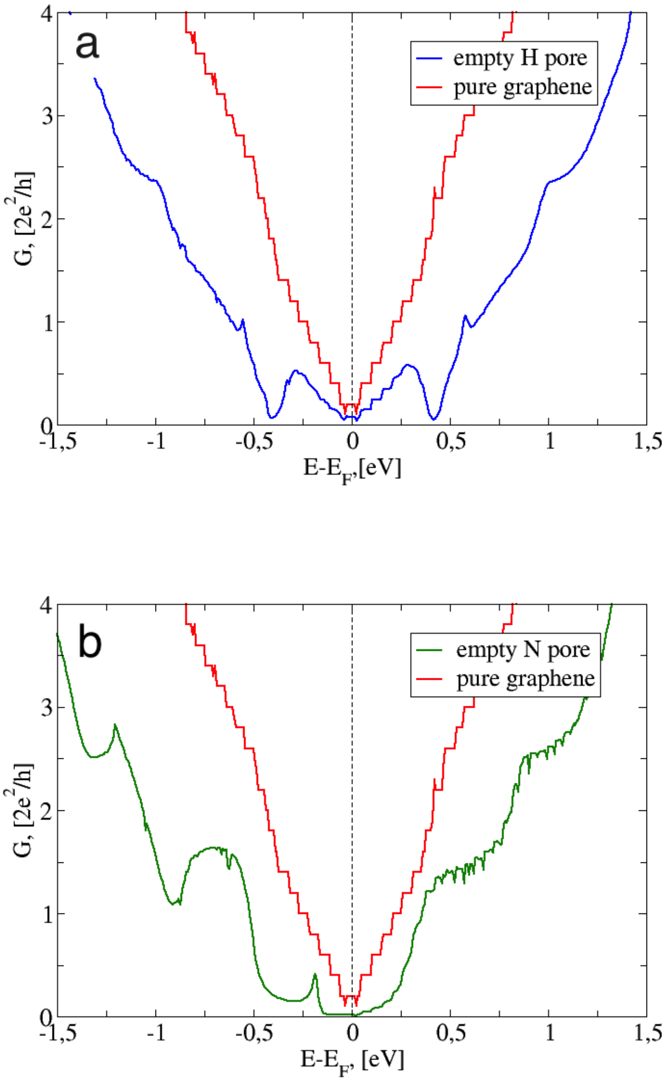


Figure 3.27: Conductance as a function of energy for the dry pore (a) with hydrogen termination, b) with nitrogen termination) comparing to the conductance of the electrode (pristine graphene).

Figures 3.28 and 3.29 show the averaged conductance of the system with bases a) DA, b) DC, c) DG, d) DT inside compared to the electrode conductance (red curve), including the effects of water and ions. The average was taken over the six snapshots that were considered in the previous section. We notice that the curves for each of the DNA bases are rather similar for the hydrogen pore, as well as for the nitrogen pore. Analyzing the standart deviation for the average conductance (yellow area aroung the average conductance curve) we see only small variations in the transmission with respect to the different snapshots for different bases, when considering the range around the Fermi level. The greatest changes are exactly located (in energy) around the dips in conductance associated with the edge states. This indicates that the interaction between the base and the hydrogen pore leads to fluctuations in the conductance.

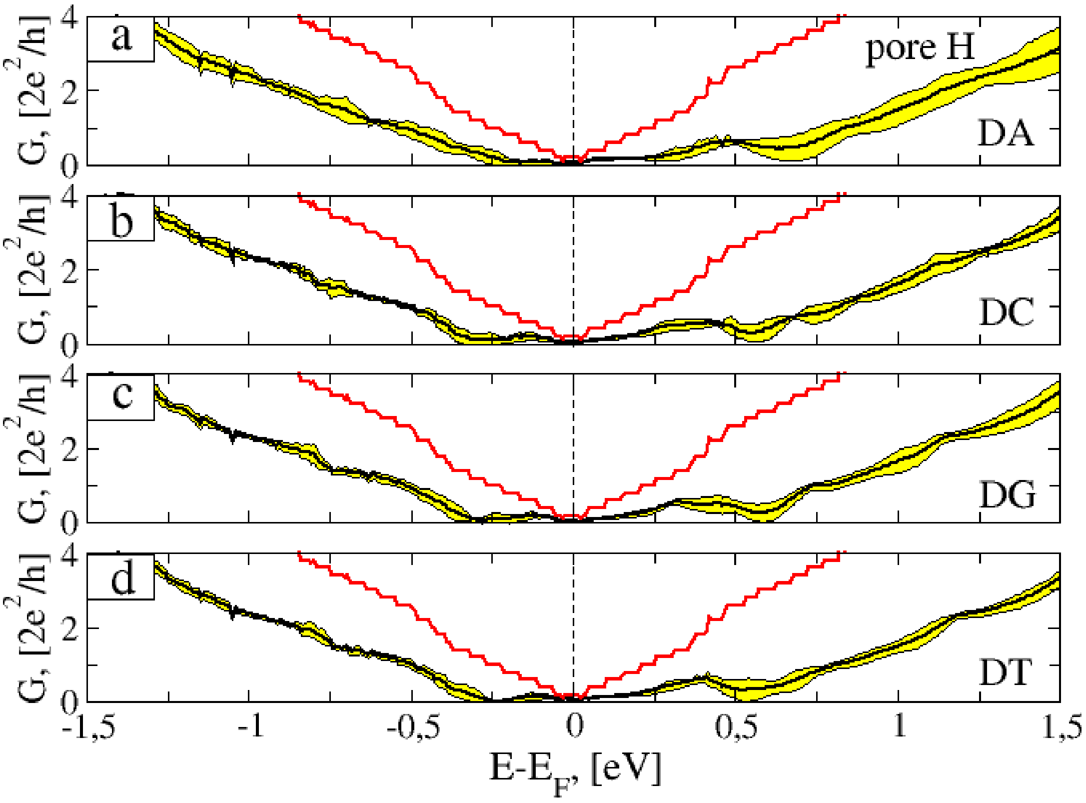


Figure 3.28: The black curve shows the averaged conductance as a function of energy of the pore with hydrogen termination and four different types of bases a) DA, b) DC, c) DG, d) DT inside containing water and counter-ions including via QM/MM approach. The yellow area enveloping the black curve corresponds to one standard deviation. Reg curve shows the conductance of the electrode.

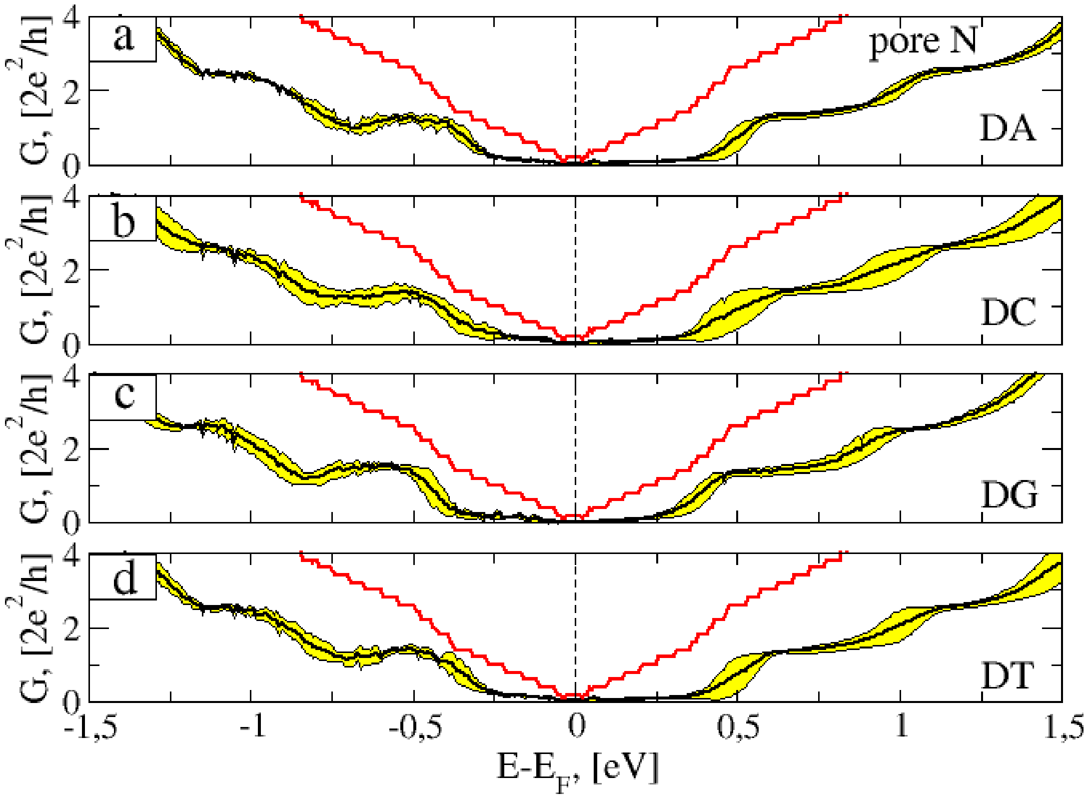


Figure 3.29: The black curve shows the averaged conductance as a function of energy of the pore with nitrogen termination and four different types of bases a) DA, b) DC, c) DG, d) DT inside containing water and counter-ions including via QM/MM approach. The yellow area enveloping the black curve corresponds to one standard deviation.

To compare the conductance of the different bases we plot the averaged conductance of each base in the Figure 3.30a and 3.31a for hydrogen terminated and nitrogen terminated pores respectively. We can see that the localized states of the empty hydrogen pore are softened when we introduce the bases, water and counter ions effects. The same time this peaks are shifted in energy. Figure 3.30b shows the averaged difference in conductance of the dry empty pore and the conductance of the different bases for the pore with hydrogen termination including water and counter-ions via QM/MM approach. We can see that the difference is rather similar for all the ranges of energy. We can distinguish the bases close to the energy 0.75 eV above the Fermi level. While for the nitrogen terminated pore (Figure 3.31b) the difference between the conductances of each of the bases is more significant for high energies. We notice, from the conductance curves that these differences are associated to energy shifts on the slope of the conductance curves (see Figure 3.31a). This is an indication that the main effect of the base is to change the local charge environment of the pore in a slightly different way for each base. This, in turn changes the local chemical potential, shifting the conductance curve. It seems nitrogen gives us better hopes for DNA sequencing compared to hydrogen. The same shift can be observed in hydrogen, but all the bases seem to have a similar effect. We can come to conclusion that the bases are more likely to be diffirenciated in the pore with nitrogen termination.

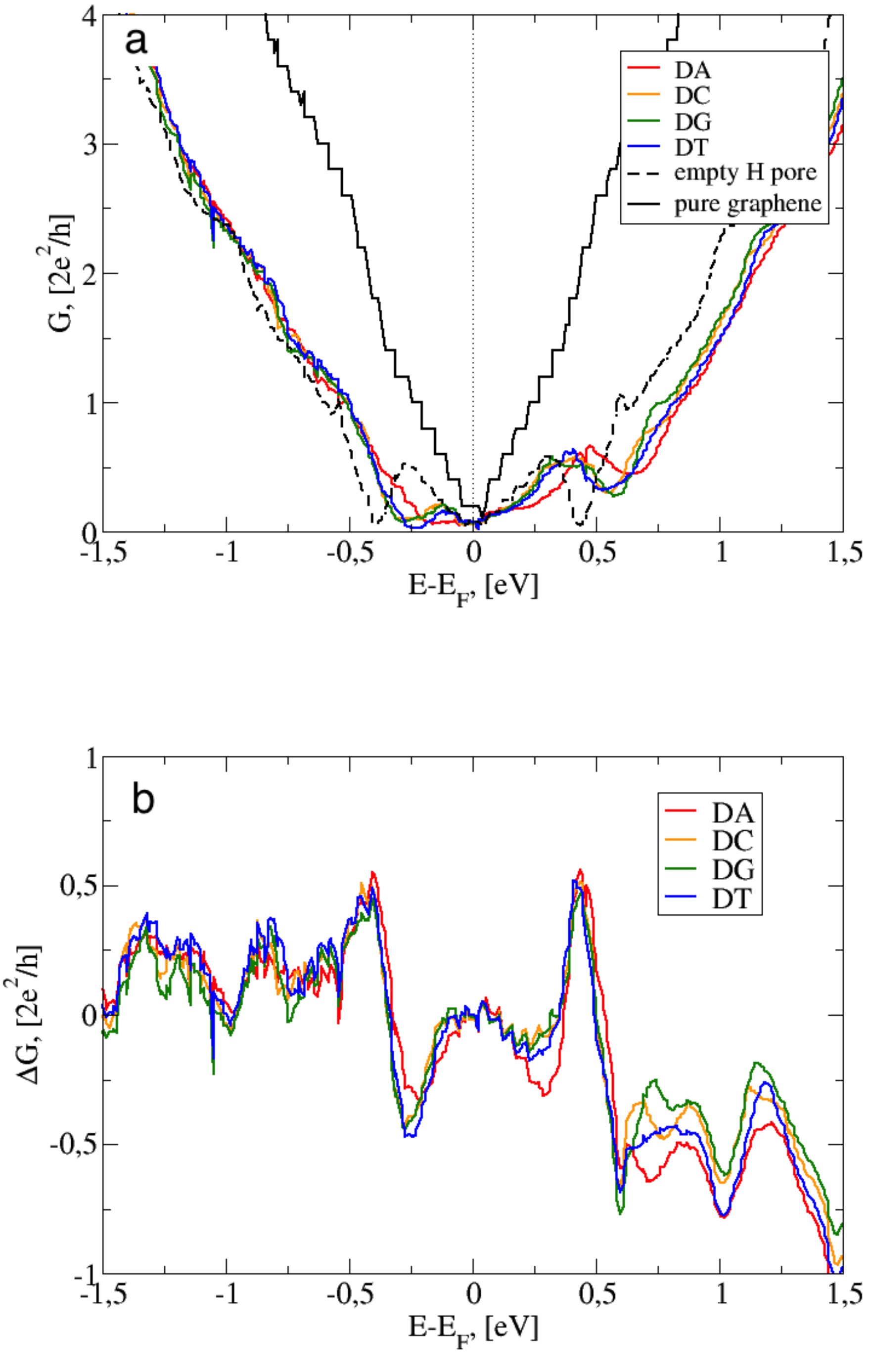


Figure 3.30: a) Comparison of the averaged conductance of the different bases as the function of energy for the H-pore including water and counter-ions including via QM/MM approach to the dry pore and electrode (pure graphene). b) Averaged difference in conductance of the dry empty pore and the conductance of the different bases for the H-pore including water and counter-ions including via QM/MM approach.

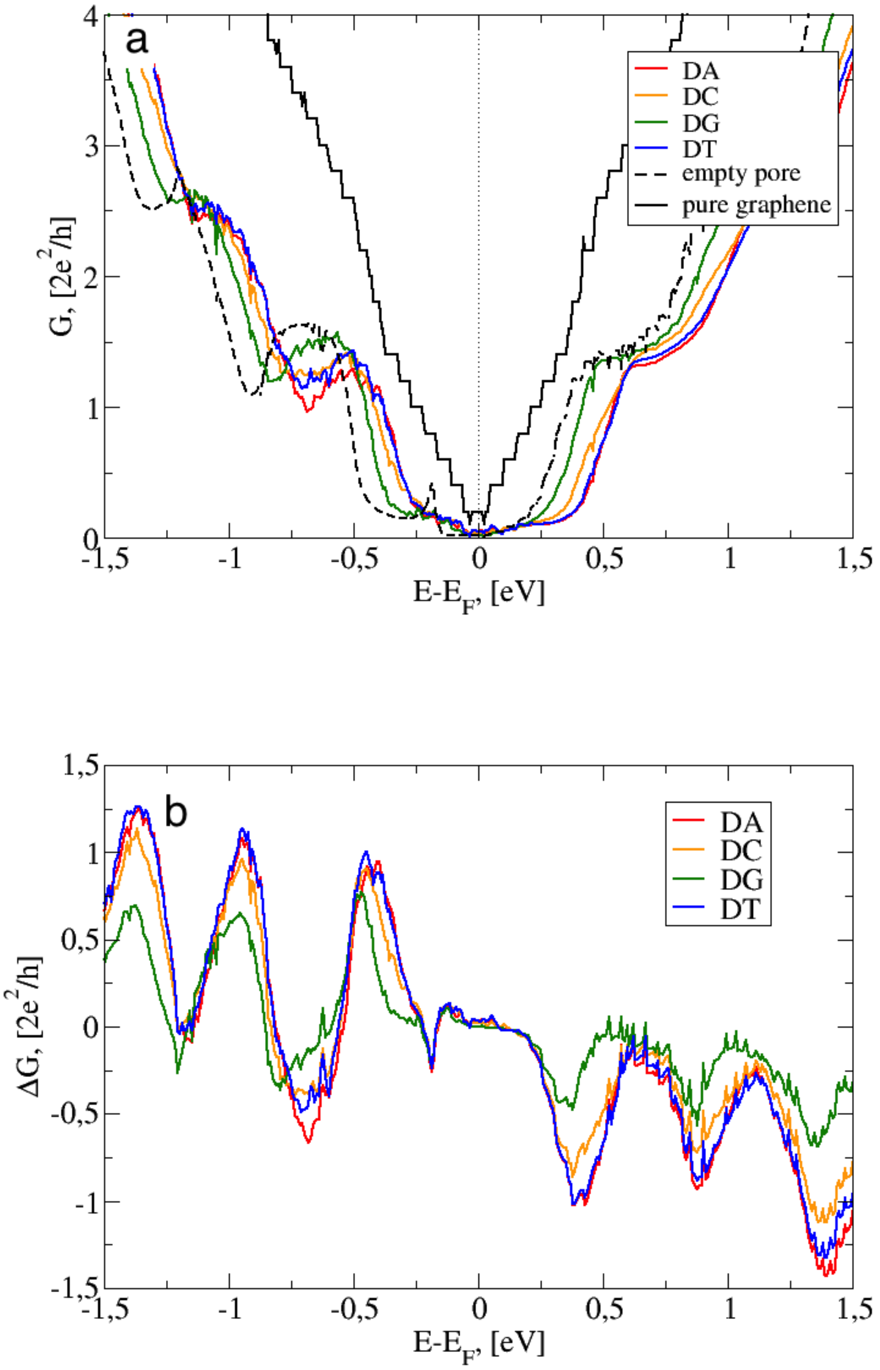


Figure 3.31: Comparison of the averaged conductance of the different bases as the function of energy for the N-pore including water and counter-ions including via QM/MM approach to the dry pore and electrode (pure graphene). b) Averaged difference in conductance of the dry empty pore and the conductance of the different bases for the N-pore including water and counter-ions including via QM/MM approach.

Conclusions

In conclusion, we used the classical molecular dynamics calculations to study dynamical effects of a graphene membrane containing a nanopore and a nucleobase surrounded by water and counter-ions. We chose six random snapshots from the trajectory of the system to calculate electronic structure using DFT. Besides, to analyze the influence of the water we did DFT calculations including one layer of QM water around each of the bases and DFT calculations including effect of all the water and ions, using QM/MM approach for the small pores (containing either H- or N-termination).

For the system with dry pore and bases we can clearly see that the charges of the bases in the pores with H termination are larger than the charges of the bases in the pores with N termination (when comparing with the neutral system). The bases interact more with the pores with H termination than with the pores with N termination. However, there are no significant differences for large and small pores.

The presence of one layer of QM water around the base markedly affects the electronic structure of the base for both types of pores. We also notice that the presence of water leads to a significant decrease in the HOMO-LUMO gap of the nucleobase for the case of the hydrogen but only minor changes in the gap for nitrogen. On the other hand, the wider dispession of states, specially for occupied ones, indicates that the base hybridizes with water much more strongly for the N-pore than the case of hydrogen. The nitrogen terminated pore does not favour charge transfer whereas the hydrogen terminated one does. This can be explained by the fact that the base tends to be negatively charged, and hydrogen atoms can easily donate charges to the base. On the other hand nitrogen is very electronegative, so it hinders the transfer of electrons to the base.

Presence of the water and ions (included via QM/MM approach) in calculations decreases the energy gap for the pore with hydrogen termination. While for the pore with nitrogen termination the energy gap remains the same. This is similar to the results considering a single layer of quantum mechanical water. For cases considered energy states are more spread in energy, including the effect of water and ions.

Comparing three different types of systems we can come to the conclusion that including one layer of water treated quantum mechanically is similar to including the effect of the water and ions using QM/MM approach; both make significant changes to the description of the electronic structure of the system and should thus be included in electronic transport calculations.

From the point of view of transport we have found only small variations in the transmission with respect to the different snapshots but significant differences for the different terminations. In particular, the N-terminated pore suppresses the resonances observed in the H-terminated one. Although, this indicates that sequencing might be hard in this system it also shows that different terminations might be employed to achive this goal.

Thus, although we haven't simulated a full device, we have setup a procedure to do so and we have understood a few key ingredients in the search for a graphene-based third generation sequencing device.

Therefore the future steps would be calculating the system: nanopore, base + one layer of water described quantum mechanically, while the rest of the water and ions included via QM/MM approach. This will give us a better understanding of the effects of the solvent onto the electronic properties of the system. Using this information we can help setup the ideal configuration for the electronic transport studing different pore terminations. In addition, to have a better description of the electronic structure of the system the number of the snapshots must be increased in order to obtain statistically significant results.

Bibliography

[1] R. Horton, L.A. Moran, K. G. Scrimgeour, M.D. Perry, and J.D. Rawn. *Principles of Biochemistry*, volume 832. Peason Prenticer Hall, Inc, 4-th edition, 2006.

[2] National human genome research institute. http://www.genome.gov/12513210, 2013.

[3] Makusu Tsutsui, Masateru Taniguchi, Kazumichi Yokota, and Tomoji Kawai. Identifying single nucleotides by tunnelling current. *Nature nanotechnology*, 5(4):286–90, April 2010.

[4] Cees Dekker. Solid-state nanopores. *Nature, Nanotechnology*, 2:209 – 215, 2007.

[5] C N Lau, D Golberg, and X W Lou. Properties of suspended graphene membranes. *Materials today*, 15(6):238–245, 2012.

[6] Tammie Nelson, Bo Zhang, and Oleg V Prezhdo. Detection of nucleic acids with graphene nanopores: ab initio characterization of a novel sequencing device. *Nano letters*, 10(9):37–42, September 2010.

[7] Christopher a Merchant, Ken Healy, Meni Wanunu, Vishva Ray, Neil Peterman, John Bartel, Michael D Fischbein, Kimberly Venta, Zhengtang Luo, a T Charlie Johnson, and Marija Drndić. DNA translocation through graphene nanopores. *Nano letters*, 10(8):2915–21, August 2010.

[8] Charles Xie and Robert Tinker. Molecular Dynamics Simulations of Chemical Reactions for Use in Education. *J. Chem. Educ.*, 83:77, 2006.

[9] Gerrit Groenhof. *Biomolecular Simulations: Methods and Protocols*, volume 924. Springer Science+Business Media New York, 2013.

[10] Klabunde K.J. *Nanoscience Materials in Chemistry*, volume 804. Wiley-Interscience, New York, 2001.

[11] Charles R Martin and Punit Kohli. The emerging field of nanotube biotechnology. *Nature reviews. Drug discovery*, 2(1):29–37, January 2003.

[12] David a LaVan, Terry McGuire, and Robert Langer. Small-scale systems for in vivo drug delivery. *Nature biotechnology*, 21(10):1184–91, October 2003.

[13] Cohen H. et al. Sustainded delivery and expression of DNA encapsulated in polimeric nanoparticles. *Gene Ther.*, 7:1896–905, 2000.

[14] C. Oostenbrink, A. Villa, A.E. Mark, and W.F. van Gunsteren. A biomolecular force-field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Am. Comput. Chem.*, 25:1656–1676, 2004.

[15] Wendel Andrade Alvesa T. Andrade-Filho, Fabio Furlan Ferreiraa and Alexandre Reily Rocha. The effects of water molecules on the electronic and structural properties of peptide nanotubes. *Chem. Phys.*, 15:7555–7559, 2013.

[16] Chen R.J., Zhang Y., Wang D., and Dai H. Noncovalent sidewall functionalization of single-walled carbon nanotubes for protein immobilization. *J. Am. Chem. Soc.*, 123:3838–3839, 2001.

[17] Wong S.S., Joselevich E., Wooley A.T., Cheung C.L., and Lieber C.M. Covalently functionalized nanotubes as nanometer-size probes in chemistry and biology. *Nature*, 394:52–55, 1998.

[18] Henk W Ch Postma. Rapid sequencing of individual DNA molecules in graphene nanogaps. *Nano Letters*, 10(2):420–425, 2010.

[19] David L. Nelson and Michael M. Cox. *Principles of Biochemistry*, volume 1100. W.H. Freeman and Company, 2008.

[20] J. D. Watson and F. H. C. Crick. A Structure for Deoxyribose Nucleic Acid. *Nature*, 171:737–738, 1953.

[21] Michael Zwolak. Colloquium: Physical approaches to DNA sequencing and detection. *Reviews of modern physics*, 80:141–165, 2008.

[22] F. SANGER, S.NICKLEN, and A. R. COULSON. DNA sequencing with chain-terminating inhibitors. *Proc.Nati.Acad.Sci.USA*, 74:463–5467, 1977.

[23] Li J., Gershow M., Stein D., Brandin E., and Golovchenko J. DNA molecules and configurations in a solid-state nanopore microscope. *Nature mater*, 2:611–615, 2003.

[24] Michael Zwolak. Electronic Signature of DNA Nucleotides via Transverse Transport. *Nano Letters*, 5:421–424, 2004.

[25] Bala Murali Venkatesan and Rashid Bashir. Nanopore sensors for nucleic acid analysis. *Nature, Nanotechnology*, 6:615–624, 2011.

[26] Gregory F. Schneider and Cees Dekker. DNA sequencing with nanopores. *Nature Biotechnology*, 30:326–328, 2012.

[27] David W Deamer and Mark Akeson. Nanopores and nucleic acids: prospects for ultrarapid sequencing. *Trends in Biotechnology*, 18:147–151, 2000.

[28] J. J. et al. Kasianowicz. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences*, 93(24):13770–13773, November 1996.

[29] Oxford nanopore technology. http://www.nanoporetech.com/news/press-releases/view/39, 2012.

[30] A K Geim and K S Novoselov. The Rise of Graphene. *Nature Materials*, 6:183, 2007.

[31] 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.

[32] Shankar R. *Principles of Quantum Mechanics*, volume 694. Springer, 1994.

[33] M.A. Gonzalez. Force Fields and molecular dynamics simulations. *EDP Science*, 12:169–200, 2011.

[34] R. Tinker and Q. Xie. Applying Computational Science to Education: The Molecular Workbench Paradigm. *Computing in Science and Engineering*, 10:24 – 27, 2008.

[35] Brenner D.W. Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. *Phys. Rev.*, 42:9458–9471, 1990.

[36] D. Hestenes. Notes for a Modeling Theory of Science, Cognition and Instruction. *GIREP conference: Modeling in Physics and Physical Education*, 28:1–28, 2006.

[37] N.L. Allinger. Allinger's molecular mechanics research lab. http://europe.chem.uga.edu. http://europe.chem.uga.edu.

[38] A.D. Mackerell et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Phys. Chem.*, 102:3286–3616, 1998.

[39] W.D. Cornell et al. A second generation force field for the simulation of proteins, nucleic acids, and orgnic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.

[40] Jorgensen W.L., Maxwell D.S., and Tirado-Rives J. Development and testing of the OPLS all-atom force-field on conformatinal energetics and properties of prganic liquids. *J. Am. Chem.*, 118:11225–11236, 1996.

[41] Sun H. COMPASS: An *ab initio* force field optimized for condensed-phase application - Overview with details on alkane and benzene compounds. *J. Phys. Chem.*, 102:7338–7364, 1998.

[42] Car R. and Parrinello M. Unified Approach for Molecular Dynamics and Density Functional Theory. *Phys. Rev. Lett.*, 55:2471–2474, 1985.

[43] Marx D. and Parrinello M. Ab initio path integral molecular dynamics: Basic ideas. *J. Chem. Phys.*, 118:166–184, 1996.

[44] M.P. Allen and D.J. Tildesley. *Computer simulations of liquids*, volume 408. Oxford University Press., 1987.

[45] D. Frenkel and B. Smit. *Understanding Molecular Simulation*, volume 664. Academic Press, 2002.

[46] Berendsen H.J.C., Postma J.P.M., Van Gusteren W.F., Dinola A., and Haak J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81:3684–3690, 1984.

[47] S. Nose. A unified formulation of the constant temperature molecular dynamics method. *J. Chem. Phys.*, 81:511–519, 1984.

[48] H.C. Andersen. Molecular dynamics at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384–2393, 1980.

[49] Berendsen H.J.C. and R. van Drunen D. van der Spoel. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91:43–56, 1995.

[50] Thomas L.H. *Proc.Cambridge Phil. Soc.*, 33, 1927.

[51] Fermi E. *Z. Phys.*, 48, 1928.

[52] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[53] Klaus Capelle. A Bird's-Eye View of Density-Functional Theory. *Brazilian Journal of Physics*, 36:318–1343, 2006.

[54] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566–569, Aug 1980.

[55] Walter Kohn, Yigal Meir, and Dmitrii E. Makarov. van der waals energies in density functional theory. *Phys. Rev. Lett.*, 80:4153–4156, May 1998.

[56] E. K. U. Gross M. Lein, J. F. Dobson. Toward the description of van der Waals interactions within density functional theory. *Computational Chamistry*, 20:12–22, 1999.

[57] Jorge Kohanoff. *Electronic structure calculations for solids and molecules*, volume 372, pages 51–70. Cambridge University Press, 2006.

[58] Kieron Burke. Perspective on density functional theory. *Chem. Phys.*, 10(Xc):1–10, 2012.

[59] Jos M Soler, Emilio Artacho, Julian D Gale, Alberto Garca, Javier Junquera, Pablo Ordejn, and Daniel Snchez-Portal. The siesta method for ab initio order- n materials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745, 2002.

[60] Andre Machado. Current science. *Neuromodulation : journal of the International Neuromodulation Society*, 16(6):504–5, November 2013.

[61] Kittusamy Senthilkumar, Jon I Mujika, Kara E Ranaghan, Frederick R Manby, Adrian J Mulholland, and Jeremy N Harvey. Analysis of polarization in QM/MM modelling of biologically relevant hydrogen bonds. *Journal of the Royal Society, Interface / the Royal Society*, 5 Suppl 3:S207–16, December 2008.

[62] M Büttiker, Y Imry, R Landauer, and S Pinhas. Generalized many-channel conductance formula with application to small rings. *Phys. Rev. B*, 31:6207, 1985.

[63] Robert Eisberg and Robert Resnick. *Quantum physics of atoms, molecules, solids, nuclei, and particles*, volume 864. Wiley, 2-d edition, 1985.

[64] Stefano Sanvito Alexandre Reily Rocha. *Electronic transport at the nanoscale*, volume 212. Lambert Academic Publishing, 2010.

[65] S. Datta. *Electronic Transport in Mesoscopic Systems*, volume 390. Cambridge University Press, 1995.

[66] Magnus Paulsson. Non Equilibrium Green ' s Functions for Dummies : Introduction to the One Particle NEGF equations. pages 1–9, 2008.