Modeling and Simulation of Charge Transport Through Nanoscale Nucleic Acid Structures

Hashem Mohammad

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Manjeri P. Anantram, Chair

Robert B. Darling

Mo Li

Program Authorized to Offer Degree:

Electrical and Computer Engineering

University of Washington

Abstract

Modeling and Simulation of Charge Transport Through Nanoscale Nucleic Acid Structures

Hashem Mohammad

Chair of the Supervisory Committee:
Manjeri P. Anantram
Department of Electrical and Computer Engineering

Understanding and controlling the electrical conductivity of nucleic acids has gained more interest in the past decade. Measuring DNA conductance for sensing biological processes, developing new sequencing techniques, and future molecular device applications have led to an interest in its electrical properties. Further, DNA Origami exploits the self-assembly property of DNA to create complex three-dimensional architectures. This technique helps build nanoscale structures bottom-up instead of the top-down approach currently used in nanoelectronics. Therefore, understanding how charge transports through nucleic acids can help engineer a new class of biosensors and nanoelectronics. The difficulties in explaining experiments arise because the system is at the nanoscale, exists in a solvent environment and is floppy. Therefore, the atomic details of the molecule have a substantial impact on the results.

This thesis focuses on the theory and modeling of quantum charge transport through nucleic acid structures using Green's function method. First, we discuss the development of the model for elastic and inelastic electron scattering. We show that the weak coupling between the DNA bases is untreatable with first-order perturbation approximation. Hence, the phenomenological Büttiker probe method is developed to be used on DNA structures instead of the commonly used self-consistent NEGF method to include scattering. Next, we discuss the energy-dependent decoherence model, where the decoherence decays from the molecular orbitals of the system. We demonstrate that using the real and imaginary parts of the self-energy in this model is critical to obtaining the correct integration of the density of states. With the developed models, we study the following electrical-based applications in the metal-DNA-metal junction setup: (1) detection of single-base mismatch; (2) DNA doping through intercalation to modulate the conductance; (3) using the sequence to engineer DNA heterostructures; (4) understanding perpendicular charge transport in DNA lying on a gold substrate. This thesis demonstrates that the model can explain experiments and help simulate unexplored paths. However, further work is required to model large DNA Origami structures and include contact atoms to the ab-initio methods to help study the current-voltage characteristics, expanding the scope of the DNA applications.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The past seven years have been an incredible journey. First, I would like to express my highest gratitude to my advisor, professor M. P. Anantram, who not only helped me as an academic advisor but as a great mentor and teacher. I joined Prof. Anantram's group as a Master's student working my way through grad school to develop my knowledge in the field of nanotechnology. Even though I had minimal experience in this field, prof. Anantram noticed my enthusiasm to learn new and exciting topics and has welcomed me into his group. He has helped me build and acquire the necessary skills and knowledge to excel in this field. It has been a pleasure working with such a caring, helpful, and patient advisor who always had an open-door policy (and open-phone policy as well!). Whenever I needed to discuss new ideas, help with understanding specific topics, or just wanted some general guidance, he was always willing to listen and help as much as possible. One of the distinguishing traits of Prof. Anantram is that he focuses on the quality of the work and always pushes and challenges me to tackle difficult questions. This helped me develop and understand the theory and explanation of the systems we model and study. His way of thinking and analyzing the problem has been engraved in me as part of my academic progress.

I want to thank our international collaborators, starting with Prof. Ersin Emre Oren and his fantastic group at TOBB University of Economics and Technology (in Ankara, Turkey), especially his grad students Busra Demir and Cagalanaz Gultakti. Our fruitful discussions and weekly meetings have led to several publications. They have also helped me understand and appreciate the importance of running MD simulations to generate the atomic coordinates prior to studying the electronic properties of molecular structures.

Throughout the collaborative work, in Chapter 6 of this thesis, Busra Demir has carried out the MD simulations for the DNA:RNA system with anthraquinone and Anthraquinone-Neomycine intercalators to generate the structures prior to the DFT step. Also, in Chapter 8, Busra Demir helped me with this work by running the MD simulations for the DNA lying on a gold substrate.

I would also like to thank our industry liaison for the NSF-SRC SemiSynBio project, Dr. Binquan Luan (Computational Biological Center, IBM Thomas J. Watson Research, NY, USA), his experience and insights have immensely helped me with the DNA+intercalator work discussed in Chapter 6. In addition to his perceptive input, he has carried out the MD simulations for the ds-DNA system (long and short strands) with the different anthraquinone intercalation locations.

I would also like to extend my gratitude to Prof. Sunil R. Patil from the Institute of Science in India. While working with our group as a visiting professor, I interacted with Prof. Sunil daily. The discussions on charge transport through DNA heterostructures helped formulate Chapter 7, leading to a publication. In this chapter, to generate the structures prior to the DFT step, Prof. Sunil first carried out energy minimization calculations of the structures.

My gratitude also extends to our experimental collaborator Prof. Josh Hihath and his group, for his continuous interaction and for providing us with exciting ideas and systems to model. Having an experimentalist collaborator has been essential to my understanding of DNA conductance measurements and the difficulties faced in interpreting the data. As modeling and theory researcher, this has motivated me to help shed light on the different experimental setups and how they affect the system properties.

I would also like to thank my committee members, Prof. Robert B. Darling and Prof. Mo Li from the electrical and computer engineering department, and Prof. James M. Carothers from the chemical engineering department, who was the acting GSR for my defense.

My most incredible gratitude goes to my dear family for their tremendous support and always believing in me. To my wife and our lovely child, thank you for being there, especially during the difficult coronavirus times.

Finally, I would like to thank Kuwait University Fellowship for financially supporting me and giving me the chance to pursue my Ph.D. degree.

Chapter 1. INTRODUCTION

The development of nanoscale devices has a continuous demand of reducing the size of the system. Evidently, Silicon-based electronics are limited by the resolution of lithography techniques since they employ the top-down approach for fabrication. On the other hand, molecular electronics offer a bottom-up approach based on chemical reactions which provide a higher resolution for fabricating at the nanoscale. In this domain, nucleic acids have become a prominent component by providing important properties for electronic applications: long range charge transport and self-assembly. This has led to extensive theoretical and experimental work to study their electronic properties. However, different electronic behavior results are obtained experimentally, which shows the sensitivity of the structure to the details of the experimental set up. This calls for a better understanding of the factors affecting charge transport through nucleic acids as applied to future electronic devices. In this thesis, I work on modeling charge transport through nucleic acid structures in the nanoscale, developing methods for calculations and applying them to analyze their potential applications in electronics.

1.1 NUCLEIC ACIDS STRUCTURE

In all living organisms, nucleic acids are essential polymers present in the cell's nucleus. There are two main types of nucleic acids [1], deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA has the genetic information essential to make functional molecules (proteins) that help establish the organisms, whereas the RNA acts as the linking agent for biological processes such as the protein synthesis. In this process, RNA gets the information from DNA in order to create the protein necessary for a specific cell function. Both structures consist of three main groups:

pentose (5-carbon-atom) sugar, organic bases, and inorganic phosphate as shown in Figure 1.1, these groups form the building block of the nucleic acid.

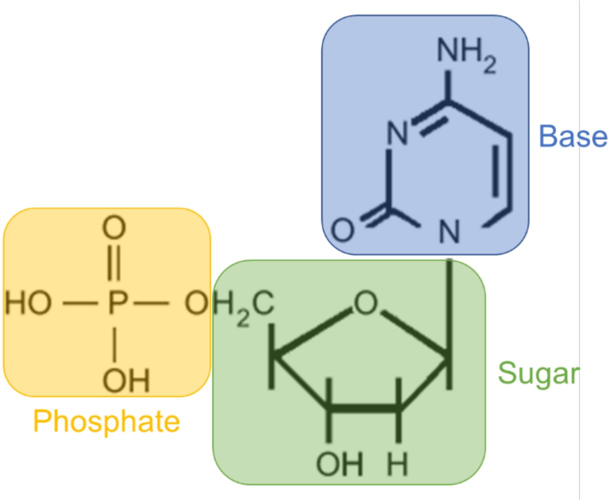


Figure 1.1 Nucleic acid structure consists of three main groups: base, sugar group, and phosphate.

The two different names arise from the two different kinds of sugar groups found in each type: Figure 1.2 shows that deoxyribose (DNA) sugar is missing oxygen from one of its carbon atoms, as opposed to the ribose (RNA) sugar group. Note that the unnamed vertices in the figures are carbon atoms.

Figure 1.2 (left) Ribose sugar found in RNA (Right) Deoxyribose sugar found in DNA (from [1]).

Moving to the bases, there are two kinds of bases in the nucleic acid: purines and pyrimidines. Looking at Figure 1.3 left, a purine base consists of two carbon rings, whereas a pyrimidine has only one ring. DNA is made up of four bases: adenine (A), guanine (G), cytosine (C), and thymine

(T). Adenine and guanine are purines, while cytosine and thymine belong to pyrimidines. Another difference between DNA and RNA is that in RNA, thymine is replaced with uracil (U). Uracil is also a pyrimidine base which forms the complementary pairing with Adenine. As for G, C, and A, they are present in both DNA and RNA strands. The chemical structures of these bases can be seen in Figure 1.3.

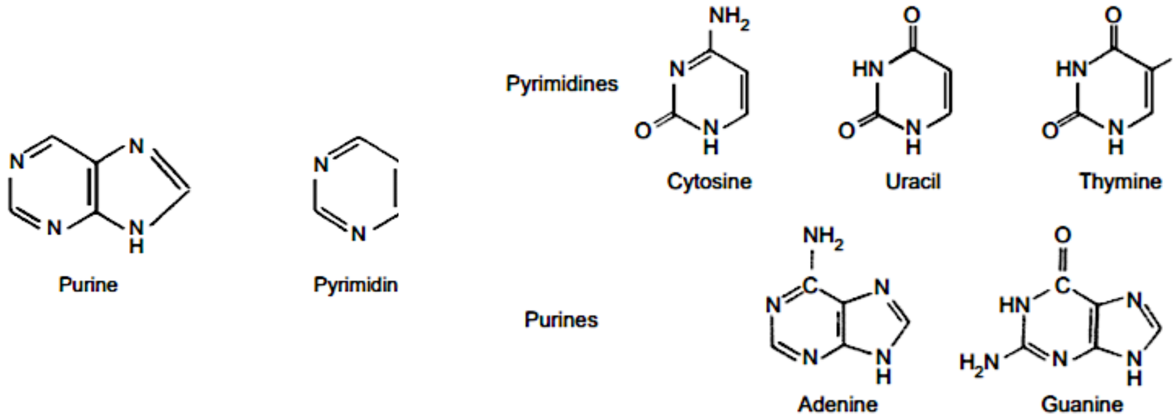
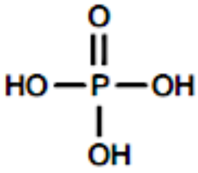


Figure 1.3 (left) Purines and Pyrimidines found in DNA and RNA (Right) The bases found in DNA and RNA, with Uracil replacing Thymine in RNA (from [1]).

The third group within the building block of the nucleic acid is the phosphate group. It is derived from the phosphoric acid, and has the structure shown below.



$$\mathrm{HO-\underset{\displaystyle OH}{\overset{\displaystyle O}{\overset{\|}{\underset{|}{P}}}}-OH}$$

Figure 1.4 Phosphate group, the third part of the building block found in nucleic acids (from [1]).

The building block of the nucleic acid, the nucleotide, consists of a base chemically bonded to the backbone. The backbone is the sugar-phosphate group, Figure 1.5 shows a single strand DNA connected via the backbone. The base is connected to the sugar group by a glycosidic bond, which is a covalent bond that joins a sugar molecule to another group. A strand is formed when multiple blocks are connected via the phosphodiester bond, which involves covalent bonds linking the 5’ end of one nucleotide and a 3’ end of another. DNA is usually present in a medium with water and salt (a polar solvent). In such an environment, each phosphate group has a charge of -1 with the corresponding positive charge of +1 being in the polar solvent. This positive charge of +1 may be in the form of Na+ (in case of the salt being NaCl), K+ (in case of the salt being KCl), or other counterions (e.g., Cs+, Mg2+). Since the dielectric constant of water is relatively large (78.4 at room temperature), the electrostatic interaction between the negatively charged groups of the phosphate backbone and positive counterions is kept low such that they do not form stable ionic bonds. In contrast, by changing the salt concentration and lowering the solvent dielectric constant, these ionic bonds become more stable and cause the DNA to precipitate out of the solution (dry state). Relevant information on the solvent effect on DNA conformation is presented in the following section.

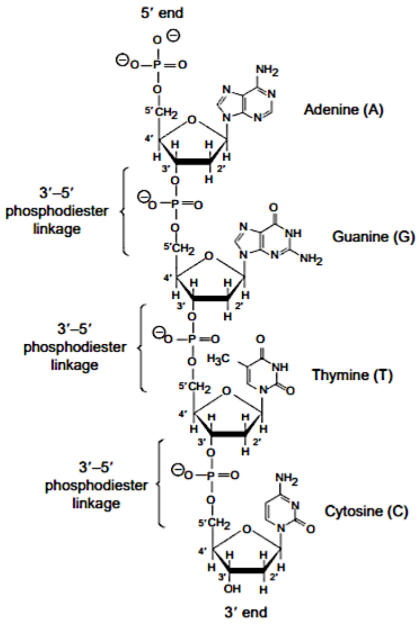


Figure 1.5 A sample of a single strand DNA molecule (from [1]).

Both DNA and RNA consist of multiple building blocks, with different conformations and different base sequences. DNA was found to be in the form of a double helix structure by Watson and Crick in 1953 [2]. The structure, called double-strand (ds), consists of two strands of sugar-phosphate backbones connected through their complementary bases (Figure 1.6). Each base of the nucleotide is connected to its exclusive complementary base via hydrogen bonds, adenine to thymine, and guanine to cytosine form the two exclusive complementary pairings, respectively. This exclusiveness is due to their chemical affinities, thus having molecular recognition capabilities.

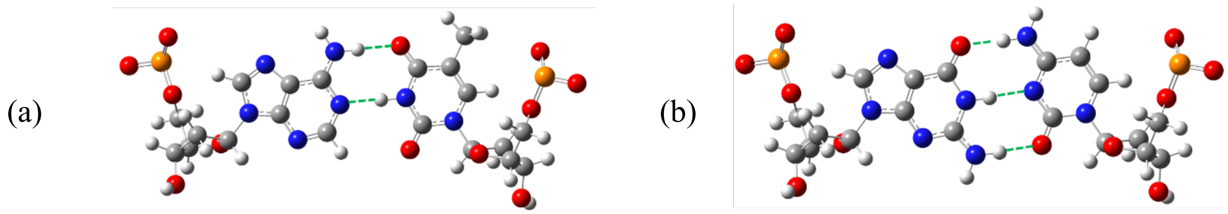
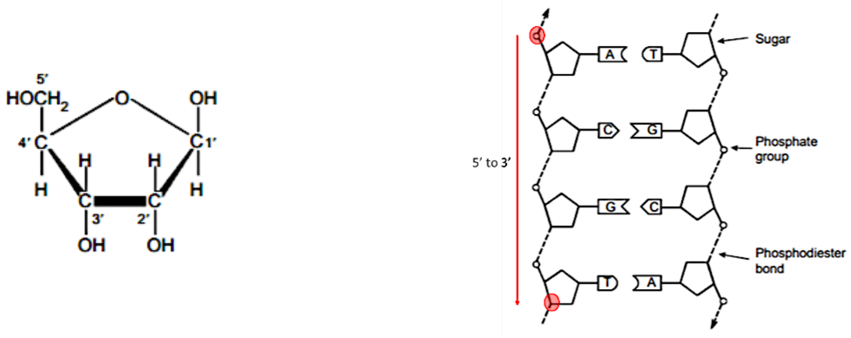


Figure 1.6 The complementary base-pairs of the DNA. (a) A-T complementary base-pair with two hydrogen bonds (green dashed line). (b) G-C complementary base-pair with three hydrogen bonds.

The two DNA strands are of opposite direction, hence called antiparallel. This can be seen from the ends of the DNA strand, which are either 5’ end or 3’ end. The 5’ end has a phosphate group terminated with the 5’ sugar carbon, whereas the 3’ end has the 3’ sugar carbon terminated with a hydroxyl (X-OH) group. Figure 1.7 (a) shows the numbering of the carbon atoms in the sugar group. As for RNA, although it is usually found in nature as a single-strand structure, it can form a hybrid DNA-RNA double helix structure. This changes the complementary AT pairing to AU: uracil on the RNA side is now paired with an adenine on the DNA side.



(a) (b)

Figure 1.7 Illustrations taken from [1]. (a) Numbering of carbon atoms in the sugar backbone. (b) Definition of 5' and 3' ends when reading a DNA sequence.

Due to the helical structure, two grooves are formed in the ds-DNA, major groove and minor groove. The major grooves form access points for binding sites of other molecules such as transcription factors, which are proteins that bind to specific DNA sequences for the transcription process. As can be seen from Figure 1.8, the major groove has a larger opening in which the base-pairs are more exposed for binding, while the minor groove has less accessible bases.

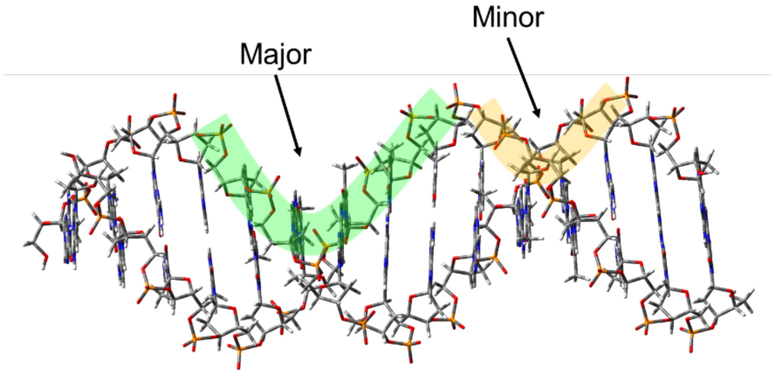


Figure 1.8 Major and minor grooves in double strand nucleic acid structures.

1.2 DNA CONFORMATION

The backbone of the nucleic acids is negatively charged due to the phosphate groups, and thus the counterions surround these groups to maintain charge neutrality [3]. However as mentioned earlier, the salt concentration and solvent dielectric constant affect the electrostatic interaction between the counterions and the phosphate groups. As a result, this changes the geometric arrangement of the atoms, yielding different conformations. The ds-DNA adopts three major forms: A-form, B-form, and Z-form. A-DNA is considered the ideal case for a ‘dry’ DNA, while B-DNA is considered as the ideal ‘wet’ DNA. In nature, DNA does not exist in the ideal forms discussed next, but rather adopts to conformations that are close to them (in structural parameters), as it fluctuates while interacting with the environment.

B-DNA is a right-handed helical structure with an average distance between two adjacently stacked base-pairs of 3.4 Å with respect to the helical axis, and of a diameter of 20 Å. As for the A-form DNA, it is also a right-handed structure but with an average adjacent base-pair distance of 2.3 Å with respect to the helical axis, and a diameter of 23 Å [4]. A transformation from B-DNA to A-DNA could occur in a ‘dry’ environment where low water concentration is present around the DNA. The change from B- to A- DNA can be imposed by increasing salt concentration and adding an organic polar solvent like ethanol. The smaller distance between the stacked base-pairs in A-DNA is due to the lowered separation between the neighboring phosphate groups. The high salt concentration (> 4 mM for 5-150 nM range of DNA concentration) and the addition of ethanol to the solution which reduces the dielectric constant yield a strong ionic bond between the counterions and the phosphate groups in the A-DNA backbone. As a result, the electrostatic repulsion between the phosphate groups due to their negative charge is reduced, thus, lowering their separation. The opposite picture is present in the B-DNA, as the higher separation between the base-pairs is attributed to the lowered ionic bonds between the counterions and the phosphate groups, yielding larger electrostatic repulsion between the neighboring phosphate groups that increases their separation. Furthermore, a major structural difference that can be seen between the conformations is the base-pairs distribution within the duplex. B-DNA has its base-pairs around the center of the helical axis, whereas A-DNA has its base-pairs more spread out of the center and closer the major groove. As for the Z-form DNA, it is quite different and less probable to occur than the other two forms. It is a left-handed helical structure and forms when the DNA has an alternating purine-pyrimidine sequence d(CG)n in an environment of high salt concentration [5], [6]. A unit cell can be defined as the number of base-pairs found per turn in the double-strand structure. A turn is the 360° rotation, corresponding to the top view of the structure in Figure 1.9. B-DNA has 10 base-pairs per turn, A-form has 11, and Z-DNA has 12 base-pairs per turn. Table 1-1 lists the geometrical differences between the three forms of DNA [7].

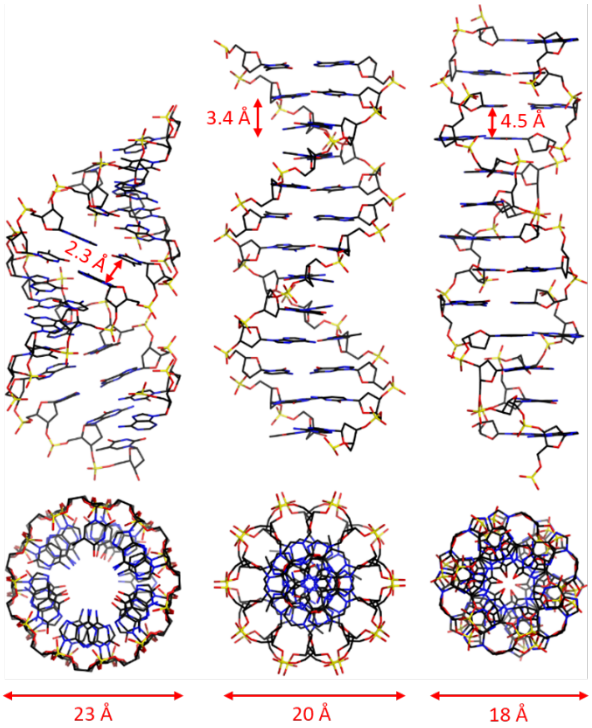


Figure 1.9 A-DNA, B-DNA, and Z-DNA, respectively (from[4]). Top figures correspond to side-view, and bottom figures correspond to top-view.

Table 1-1 Comparison of different helical parameters for A-, B-, and Z-DNA

Parameter	**A-DNA**	**B-DNA**	**Z-DNA**
Helix sense	Right	Right	Left
Base pairs per turn	11	10	12
Axial rise (Å)	2.3	3.4	4.5
Base pair tilt (°)	20	-6	7
Twist angle (°)	33	36	-30
Diameter of helix (Å)	23	20	18

Furthermore, another difference arising from the conformations is the amount of hydration at the phosphate groups in the DNA backbone. The stacking of DNA base-pairs is hydrophobic; hence water molecules can only interact and bind to DNA via accessible areas [8]. Hydration of DNA occurs in the major grooves, minor grooves, and phosphate groups [8]–[10]. In the major and minor grooves, water molecules can bind via hydrogen bonds to nitrogen and oxygen atoms of the DNA bases as shown in Figure 1.10. As for the phosphate groups, the large separation between the neighboring groups in B-DNA (6.6-8.1 Å) creates a larger solvent-accessible area. This yields a high hydration of oxygen in the phosphate, with an average of 4.5 water molecules per phosphate group. However, in the A- and Z-DNA, the average separation of the phosphate groups is less than 5.7 Å, yielding an average hydration of 1.5 water molecules per phosphate group [10], [11].

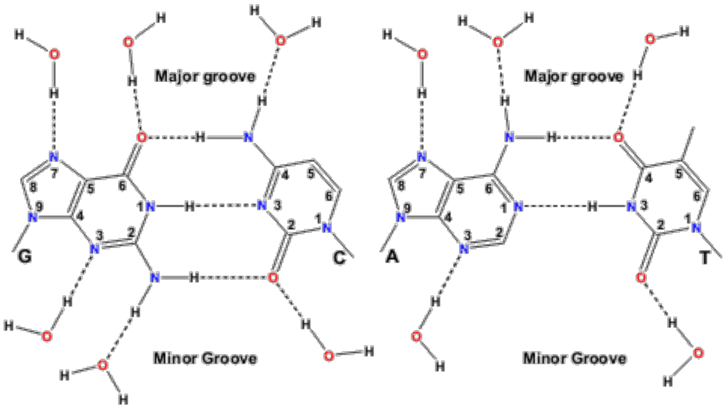


Figure 1.10 DNA base-pairs showing hydration sites, from [12].

For us, an important question is if the conductivity depends on the conformation of the DNA (for a fixed sequence). The initial intuition in the community was A-form would not conduct as well as B-form because the bases are not aligned as well. However, conductance measurements reported in [13] show that the A-form yielded 10 times higher conductance than the B-form. Although the more detailed theoretical calculations explained the experiments, it could not fully explain the ten-times increase in conductance [9]. Theory showed that the inclusion of backbone into the simulation was essential to determine the distribution of the energy levels, and the conductance.

1.3 ELECTRONIC PROPERTIES OF DNA

Electronic properties of DNA have been studied since 1962 [14], and it has since been suggested that the noncovalent π-π interactions of the stacked bases' P_z orbitals are of importance (see Figure 1.11). This is because in the ds-DNA or the DNA-RNA hybrid, the bases are planar and the entire structure is of helical shapes; this can form an effective coupling between neighboring bases, which leads to conducting behavior.

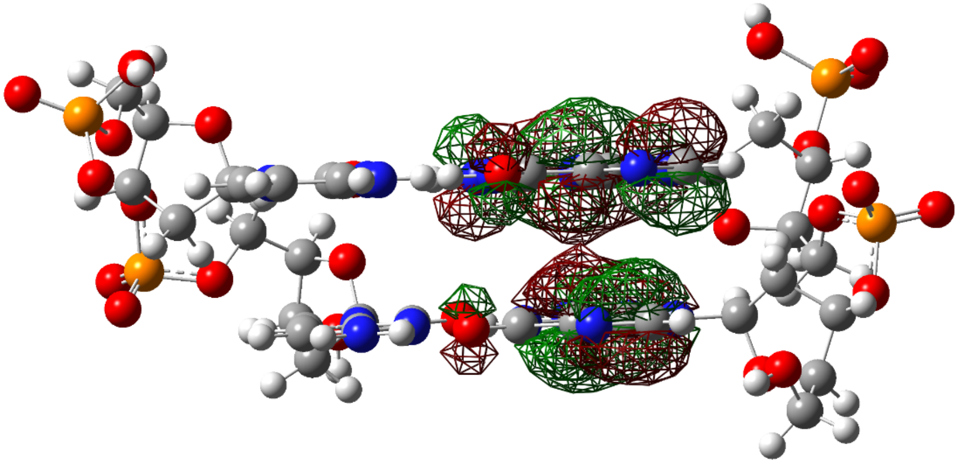


Figure 1.11 Illustration of the π-π interactions resulting from the stacked P_z orbitals of the bases (isosurface). The isosurface resembles the spatial distribution of the molecular orbitals, i.e. the electron wavefunction. The two colors are for positive and negative signs of the wave function (phases). The parallel stacking causes π-π coupling between the molecular orbitals.

What makes the DNA an interesting subject for electrical engineers is the electronic properties of its bases, and the ability to create any desired sequence of bases through the oligonucleotide synthesis process. One of the main properties is the ionization potential (IP) of the bases. IP is the amount of energy required to remove an electron from the outermost shell of a molecule to vacuum, also known as the highest occupied molecular orbital (HOMO). The study of IP is important in understanding the properties of the molecular devices because it is akin to the concept of Work Function that is extensively used in understanding the charge transport in electronic devices. There are two types of IP in molecules: adiabatic ionization potential and vertical ionization potential. Adiabatic IP corresponds to the energy difference between the cation and its neutral molecule, both in their respective ground-state geometry. Thus, it accounts for the relaxation of atoms in the formed ion after the removal of an electron [15]. Vertical IP (VIP) corresponds to the energy of removing an electron without the reorganization of atoms, the ion is thus in the same geometry as the neutral molecule [16]. In experiments, the ionization potentials are measured through photo-induced ionization, by targeting the DNA base with a laser pulse and measuring fluorescence emissions [17], either in gaseous state or in a solvent. In solvent, the IP measured is either fitted with adiabatic or vertical IP models, however, further simulations show that using VIP yields closer values to experiment, supporting the latter [18]. Although DNA tends to be in an aqueous state in nature, to have an intrinsic sense of the IP properties, it is also of interest to study it in a dry state, minimizing the effect of the surrounding environment. Several papers have studied the IP of bases, bases+sugar (nucleoside) and base+backbone (nucleotide) in gaseous (dry state) and aqueous phases, both experimentally and theoretically [19]–[21]. Qualitatively, there is an overall agreement that the IP of the bases are ordered as follows: U (7.01 eV)>T (6.71 eV)>C (6.62 eV)>A (6.35 eV)>G (6.01 eV). This shows that the IP of guanine is the lowest, hence, it has the highest oxidation probability out of all bases.

Initially, Eley and Spivey have studied the electronic properties of DNA back in 1962 [22], suggesting that the π-π coupling between the stacked bases could lead to conductive behavior. About thirty years later, charge transfer through DNA had gained traction and was studied by several groups between 1999-2003 [23]–[27]. These studies targeted holes as the possible charge carriers in the nucleic acids. In these experiments, the charge migrates between donor and acceptor molecules connected to the two ends of a double strand DNA, as shown in Figure 1.12. They were able to demonstrate through transient-absorption spectroscopy that charge transfer in DNA can be initiated via formation of a guanine radical cation, where G is chosen due to having the lowest IP. The cation would then attract an electron from the next most oxidizable base (nearest G), creating a hole movement throughout the DNA. Adenine and thymine act as barriers for hole movement between the G bases as they have higher ionization potentials. Following the extensive experimental and theoretical work in the years after, the transfer mechanism was found to depend on the distance between the G bases. As the length of the AT base-pair (bp) barrier between the G bases becomes > 3 bps, the transport mechanism was observed to change from coherent tunneling into sequential hopping. Coherent tunneling is exponentially dependent on the length (distance between donor and acceptor), while sequential hopping probability linearly decays as length increases.

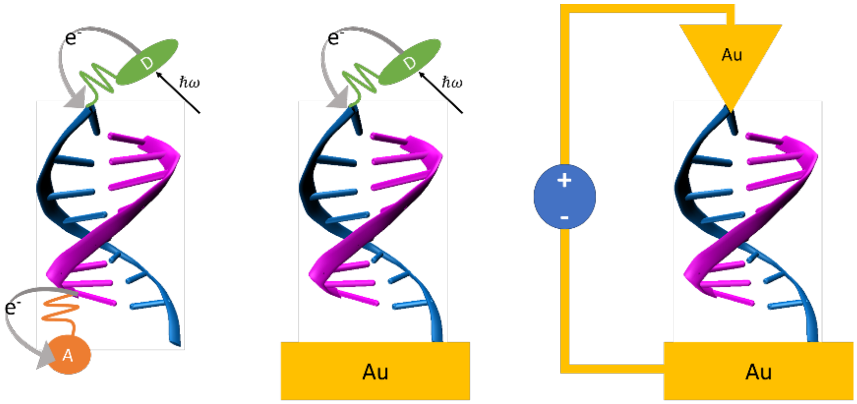


Figure 1.12 The charge transport investigation schemes: **Left & Middle** Charge transfer, **Right** is Charge transport. **Left** shows the electrochemical setup where a small donor molecule is attached to the terminal of the strand, and the other end is attached to an acceptor. A laser pulse initiates charge movement; charge transfer rates are measured using time-resolved spectroscopy. **Middle** displays a similar setup; however, the acceptor is now a gold surface to measure the current and is curve-fitted to yield charge transfer rate. **Right** Charge transport: conductance measurement setup, the molecule is attached to two metal contacts on its two ends, and a bias is applied to measure the current, conductance value is the output of this method.

These results have prompted scientists and engineers to study charge transport of nucleic acids from their perspective, using single molecule conductance measurements. In this setup, the nucleic acid is bridging the gap between two electrodes, creating a metal-molecule-metal configuration (Figure 1.12 right). A small bias is applied and the current flowing between the two electrodes via the nucleic acid is measured accordingly. At the small bias regime, the conductance is linearly proportional to the current, $G = I/V$. This method was initiated by Xu and Tao back in 2003 [28], where they used the scanning tunneling microscopy break-junction (STM-BJ) technique to measure conductance of a single molecule. The method involves a metal tip being brought into contact with the molecule that is lying on a metal surface. The molecule is terminated on both of its ends with linker groups that can covalently bond with both the metal surface and the tip. Hence, bridging the gap between the tip and the surface and forming the metal-molecule-metal junction. Conductance is then measured as the tip is slowly pulled away from the surface, which results in a single molecule contact. The process is repeated thousands of times yielding conductance histograms from which the molecule conductance may be extracted and analyzed. There are different variations on how to form the metal-molecule-metal junction other than the STM-BJ, such as conductive probe atomic force microscope break-junction (CPAFM-BJ)[29] which uses a metal coated AFM tip instead of fully metallic tip. Another method is the mechanically controlled break-junction (MCBJ)[30], [31] which has the metal electrodes on an elastic substrate, and by bending the substrate the junction can either be brought into contact or broken, yielding single molecule conductance accordingly. A nice summary of the break-junction methods can be found in [32], displayed in Figure 1.13 (a-e). Furthermore, a different technique from the break-junction was reported in [33], where Scanning Tunneling Spectroscopy (STS) was utilized to measure the tunneling current flowing through nucleic acid bases. As shown in Figure 1.13 (f), the tip is brought within close proximity to the molecule, and the quantum tunneling current is measured via scanning tunneling spectroscopy.

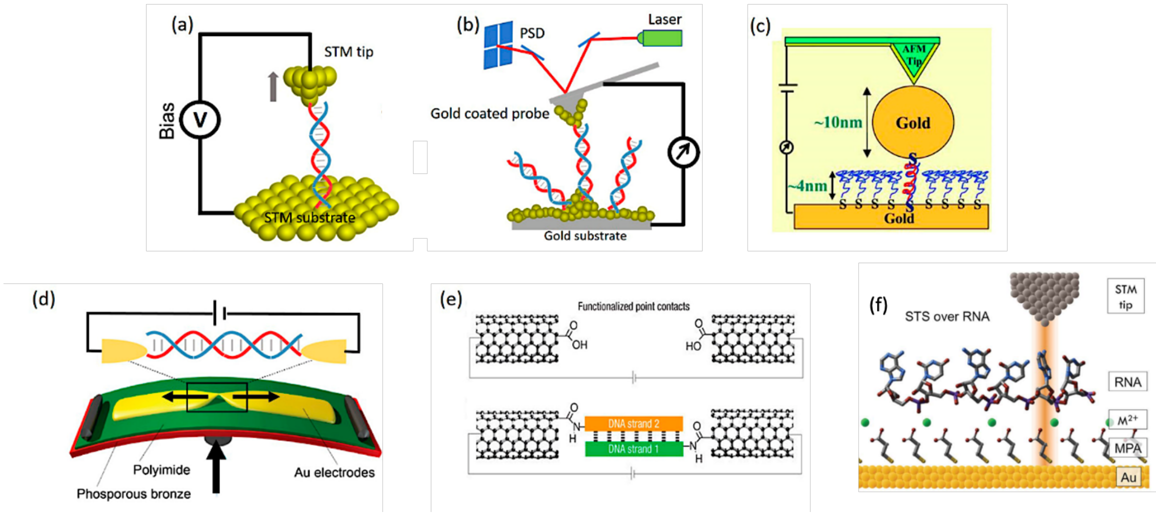


Figure 1.13 Single-molecule break-junction techniques. (a) Scanning tunneling microscope break-junction (STM-BJ) technique; (b) Conductive probe atomic force microscope break-junction (CPAFM-BJ) technique; (c) CPAFM-BJ involving nanoparticle capped DNA molecule. (d) Mechanically controlled break-junction (MCBJ) technique; (e) Single-Walled carbon nanotube (SWCNT) molecular junction. From references [32], [33].

1.4 CHARGE TRANSFER RATE VS CONDUCTANCE MEASUREMENTS

The transport properties of DNA are studied using charge transfer in isolated molecules in solution (with metal contact to carry current) and charge transport where a DNA lies between metal contact and a battery is connected between the metal contacts (Figure 1.12). The former is directly related to the electrochemistry associated with the donor-DNA-acceptor setup that measures charge transfer rates. The latter is due to the relation between charge transport and the zero bias conductance as will be shown later in the methodology chapter. However, one may ask: what is the difference between the measured transfer rates and conductance? And can we map the transfer rates to estimate conductance?

Venkatramani et al. [34] explored in depth the relationship between the electrochemical-based charge transfer rates and conductance-based charge transport. It was found to follow a power-law

$$\sigma = \sigma_{L=0} \times \left(\frac{k^0}{k^0_{L=0}}\right)^{\frac{\beta_\sigma}{\beta_{k^0}}} \tag{1.1}$$

where σ is the conductance, k^0 is the charge transfer rate, $\sigma_{L=0}(k^0_{L=0})$ is the conductance (charge transfer) at zero bridge length (i.e., no DNA is present between the donor and acceptor) , $\beta_\sigma(\beta_{k^0})$ is the decay rate for conductance (charge transfer) with bridge length. The decay ratio may cause nonlinearity in the relationship when $\beta_\sigma \neq \beta_{k^0}$, and that is due to different factors summarized as follows:

1) Changing the terminals from Donor-DNA-Acceptor into Contact-DNA-Contact electrodes changes the energy barrier height for the charge to enter/exit the system as the metal work functions may differ from the energy levels of the donor/acceptor molecules. Furthermore, the electrodes provide a continuum of states which induces larger broadening into the DNA (in the electronic device area, this is similar to Schottky barrier device physics).

2) Different experimental setups yield different geometry of the DNA, which affects the energy levels distribution and the decoherence rates.

These factors contribute to making it difficult to directly map DNA conductance from charge transfer rates. This thesis is primarily concerned with charge transport (metal-DNA-metal system).

Electrical measurements on DNA have been performed since 2000, yet it is still difficult to have unified models that yield results close to experiment and are able to predict them. The modeling of solid-state electronics has reached such maturity that nowadays many modeling tools are available for explaining experiments and predicting device behavior. However, for the nucleic acids, there are several challenges that arise in the development of unified models.

1.5 CHALLENGES

Developing theoretical and computational models for charge transport through a metal-molecule-metal configuration is essential in understanding the electrical properties of nucleic acids, explaining experiments, and designing sequences/origami nanostructures with designer current-voltage characteristic.

The basic setup for DNA-based electronics consists of two metal electrodes linked to a DNA through linker groups. The linker groups are added to the two ends of a DNA strand to bind to the metal contacts; they bridge the gap between the DNA molecule and the two electrodes. Also surrounding the DNA molecule are the counterions as well as the solvent (water molecules).

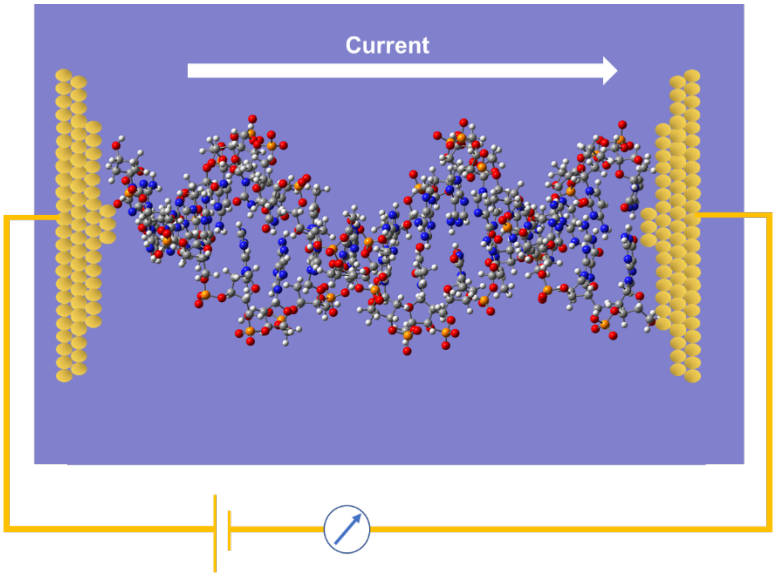


Figure 1.14 The metal-DNA-metal system, the double-strand structure is connected to two metal contacts, the background color resembles the solvent environment.

In order to model the conductance in a metal-DNA-metal structure, many variables affect the process, and these are discussed below:

1) DNA molecule

DNA is a soft molecule, where sequence, length, and geometrical configurations have a significant effect on the charge transport properties. There are a number of factors that differ in modeling DNA's electronic properties when compared to other material systems that we study, such as, carbon nanotubes, graphene sheets, or nanoscale silicon nanowires, where the structure is stiff and at least quasi-periodic. DNA sequence is a major factor, as there are in principle 4^N different sequences for N-long DNA strands, causing major differences in the resulting conductance values. From a device perspective, this is due to the different Ionization Potentials of nucleotides (G, C, A, T, and U); the spatial distribution of energy levels of a non-periodic molecule is also significantly affected by the sequence. Conformation changes also modify the energy landscape of the DNA nanostructure, hence, changing the conductance.

2) The Environment

The DNA molecule is susceptible to the environment as a result of the large surface to volume ratio[14]. This can cause (i) shift of energy levels and (ii) addition of new energy levels and (iii) possibly doping, as a result of preparation protocols. In addition, conformation changes occur due to the specifics of water and counterion components present around DNA (salt concentration, ethanol to water ratio). A more general effect is also the melting temperature, which is the temperature at which the ds-DNA is split into two single strands (a single strand may become a random coil). The melting temperature of short double strands is a function of length and sequence, and it is between 50-70 Celsius [35], [36].

3) DNA-Metal junction

The metal-molecule junction is created by a linker group. A linker group is a chemical group that binds to the terminals of the molecule, DNA in this case. This helps in optimizing the bridge between the DNA and the metal contacts by working as site specific anchoring groups while providing efficient electronic coupling[37]. Although different metals can be used such as Au, Ag, Pd, or Pt, gold contacts are the most commonly used in experiments due to their inertness[38], [39]. Two linkers that are commonly used in DNA experiments are: amine groups ($-NH_2$) and thiol groups ($-SH$), with the latter being the most used linker for gold contacts as it forms a strong S-Au covalent bond [40]. In molecular systems, the details of the metal-molecule interface affect where the Fermi energy of the contact lies with respect to the energy levels of the molecule. Details such as the geometry and orientation of the molecule affect the number of atoms adsorbed to the contacts and modify the number of contact atoms at the metal-molecule interface. The materials of the contact, the linker group, and the detailed configuration between them may affect the coupling strength between the molecule and the electrodes.

The factors described contribute to the uncertainty surrounding experiments. Despite these uncertainties, measuring DNA conductance yields worthy trends from a device perspective. These trends have shown that different sequences, strand lengths, and conformations yield different conductance values, varying the resistivity of the nucleic acids and classifying them as tunneling barriers[41], [42], ohmic molecular wires[43], conductors [44], [45], or even wide bandgap semiconductors[46].

1.6 POTENTIAL APPLICATIONS FOR DNA

DNA has three important features: molecular recognition, self-assembly, and long charge transport. Molecular recognition is the ability for a molecule to selectively bond with another molecule. Self-assembly is the ability of a group of molecules to arrange themselves in a specific order to form a supermolecule (molecules bonding together through non-covalent bonds). Furthermore, experiments have found that ds-DNA exhibits charge transfer over large distances, exceeding 34 nm [36]. Oligonucleotide synthesis, which is the process of synthesizing strands of nucleic acids with designer sequences [47] is very well established and has high precision. These important properties are suitable for creating molecular electronics using the bottom-up approach, which is to create a functioning system composed of small building blocks.

DNA Origami: The molecular recognition and self-assembly properties exhibited by DNA have been exploited by scientists to form complex structures known as DNA origami. It is the process of using the selective bonding of DNA bases to connect multiple helices via crossovers, to construct 2D or 3D nanostructures. Initially, Paul Rothemund constructed 2D origami structures using a long single strand (ss) DNA as a scaffold [48], that folded at designated locations using small complementary strands known as staple strands. An example is shown in Figure 1.15 (left) where a ~900 bases ss-DNA is designed to fold within a specific path to create a 2D structure (black line). The folding locations determine the crossover points, these crossovers connect the helices by which a strand running along one helix switches to an adjacent helix and continues there. The staple strands (colored lines) are designed to form exclusive complementary base-pairing with the scaffold and create the crossovers, yielding the full 2D structure. Examples of different DNA origami structures are displayed in Figure 1.15 (right).

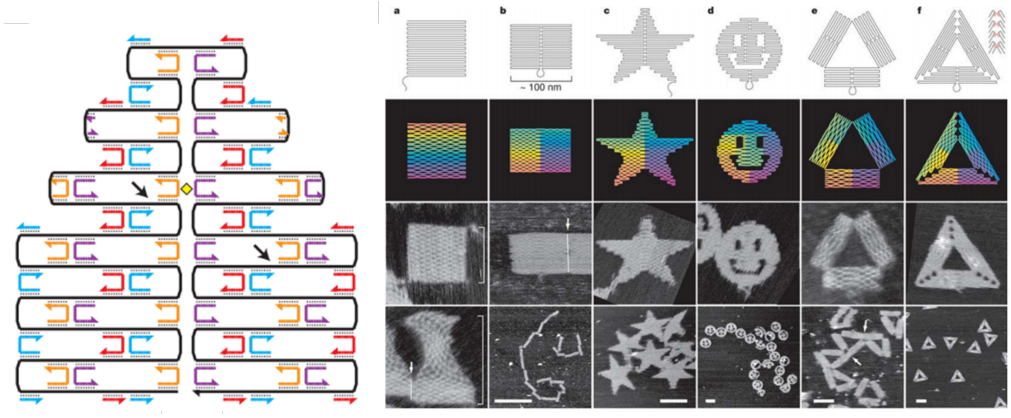


Figure 1.15 **Left** A sample 2D Origami design showing the scaffold (black line) and the staple strands (colored) forming the crossovers and the folding locations. **Right** Examples of constructed structures, top row shows the scaffolds, second row displays the bases, red is the first base, purple is the 7000th base. Bottom two rows are AFM images. From reference [48].

Another DNA origami technique has been developed by Yonggang Ke et al. [49], [50]. This technique utilizes self-assembly of short ss-DNA to form complex 3D structures. Here, DNA bricks of 32 bases with four 8-base binding domains were created (see Figure 1.16 A). Each brick contains a distinct base sequence with head and tail binding domains. The distinct sequences help match DNA bricks via targeted head-to-tail complementary pairing. Each pairing between two bricks forms a 90° angle as shown in Figure 1.16 B, thus creating LEGO-like blocks. Each block is of 2.5x2.5x2.7 nm^3 in size. Almost any arbitrary 3D structure can be formed by this method as shown in Figure 1.16 C where 3D shapes of (10x10x10) blocks are built. DNA origami shows a promising bottom-up approach for building complex electrical devices and systems.

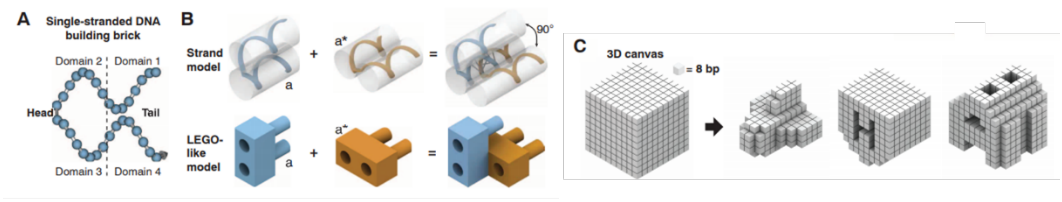


Figure 1.16 3D DNA Origami. **A** Shows the building brick consisting of a single strand DNA with head and tail binding domains. **B** Formation of 90° angle due to attachment of two bricks like a LEGO block. **C** Examples of 3D structures that can be built using this technique. From reference [50].

As an example, Figure 1.17 (left) shows a schematic for an electrically readable memory (ROM) device formed with DNA strands as active components. The side view in Figure 1.17 (right) shows DNA helix bundles (orange and green) that can be utilized as DNA nanowires. The two nanowires read the conductance value of a double-strand (blue strands in Figure 1.17 right) in between the orange and green wires at the *crosspoints*. To achieve a functioning device, the development of tools to design and model the electrical and structural properties of these elements are essential. For instance, engineering and controlling the conductance of the DNA strands, as well as understanding the DNA-electrode and substrate interface will be crucial.

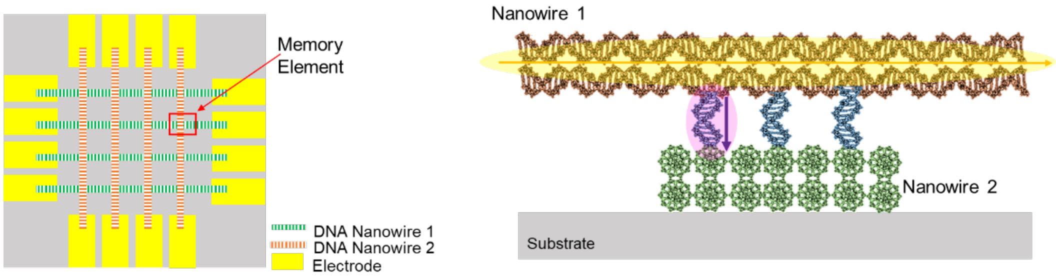


Figure 1.17 DNA ROM device. **Left** Top view of the ROM device formed by DNA nanowires interfacing with conventional semiconductor technology. **Right** Side view of the DNA memory element, the two nanowires are helix bundles, with the memory value saved on the short double-strands (blue) crossing them. By engineering the conductance value of the double-strand, different values can be read, forming the bits of the memory device. Unpublished figure provided by M. P. Anantram, used with permission in this document [PCT/US2019/041396].

Spanning the last 15 years, experiments have demonstrated that the effect of small changes in DNA can be captured through the electrical conductance measurements. DNA base mismatches [51], chemical modifications in DNA bases [52]–[54], and solvent environment effects on structure conformation can all modify the electronic structure and coupling between DNA bases which affects the conductance [13], [55], [56]. Therefore, by tuning and controlling these factors, applications in electronics, sensors, and sequencing techniques are being studied. In terms of electronics: Spin specific electron conductor [57], [58], spin filter [59], double-barrier resonant tunneling structure [56], negative differential resistance [60], DNA doping by intercalation [61], [62], and gate-controlled switch [63][41] are some of the structures studied. Furthermore, DNA in FET structures has been used as an insulator [64], charge injection layer [65], or charge trapping layer [66]. Related to this work is the creation of nanostructures where DNA serves as a template on which semiconductor and metal particles adsorb [67].

It is well known that several diseases including cancer have been attributed to nucleotide mutations [68]. This causes a change in the electronic properties of the mutated DNA strand [54], [69], [70]. Thus, disease detection through electrical conductance measurements of short DNA strands with mutated nucleotides might become viable one day.

DNA sequencing [71] is another area where emerging techniques using conductance have been proposed [72]–[74]. Sequencing of DNA has gone through three main phases; the first phase consisted of using Sanger's method in which the DNA segments are amplified (multiplied) via bacterial multiplication and sent through different tubes to interact with different identifier bases, such that when the segment passes through the gel, it gets separated based on size. Next generation techniques focused on methods to multiplex so that the sequencing process can be parallelized. In addition, the amplification technique has gone in vitro (does not use living organisms), such as the well-established polymerase chain reaction (PCR). Third generation sequencing, which is of interest here, consists of several methods: optical measurements (Pacific Bioscience-PacBio), ionic current measurements, and transverse electrical current measurements via nanopores [72] (Oxford Nanopore Technology-ONT). When a single-stranded DNA passes through a nanopore, the difference in the electronic properties of the nucleotides passing through can provide a signature ionic/electronic current. Another emerging technology in DNA sequencing is the Ion Torrent which uses an ISFET pH sensor that measures the change in pH level and converts it into voltage [73], [74](ThermoFisher, and Roswell Biotechnologies). Briefly describing the process, the DNA sequence is broken into smaller chunks, each chunk is attached to a bead, and the beads occupy their own small wells. Once a well is filled with a specific nucleotide, the added nucleotide will combine with its exclusive complementary pair found on the bead. Hence, releasing a hydrogen ion in the process and changing the pH level of the well. The pH sensor then converts this change into voltage, signaling the existence of the complementary base. What makes third generation a step up in the sequencing techniques is that it is faster, offers the ability to work on a single-molecule, and does not require PCR amplification as in the previous two generations.

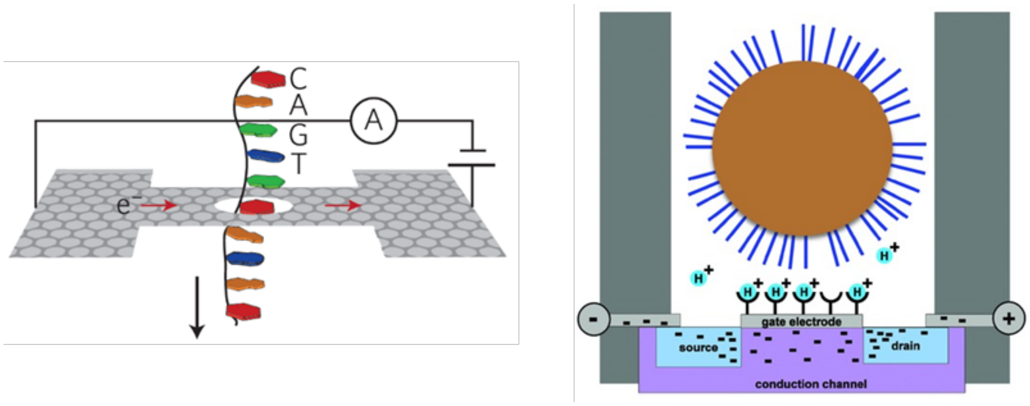


Figure 1.18 Third generation sequencing techniques. **Left** Measuring the electric current fluctuations with respect to the base occupying the nanopore of the surface [72]. **Right** Displaying the bead and the ISFET structure for the Ion Torrent technique [73].

1.7 PREVIOUS WORK

When compared to regular solid-state electronics, the main differences in DNA electronics arise from the large distance between the building blocks (bases) and the fluctuating nature of the underlying Hamiltonian, both due to their interaction with the solvent environment and lack of covalent bonding between bases. This leads to uncertainties when modeling their behavior [75]:

the geometry of the molecule changes with time, and so do the atomic details of the molecule-contact interface. The fluctuations affect the distance between the atoms/molecules and thus change the electronic coupling between the molecular orbitals of the system, which in turn affect their charge transport properties. As for the molecule-contact interface, on the atomic scale, this changes how many atoms are in actual contact with the electrode, and how strong the coupling to the electrode is. In addition, the location of the Fermi energy can vary depending on the details of the atom-molecule contact. Previous work has shown that length and DNA sequence affects charge transport, going from coherent tunneling, to sequential hopping, and even an intermediate regime between coherent and decoherent transport [76], [77]. There still exists differences between theory and experiment, prompting different models to appear which try to incorporate more physics.

Modeling of charge transport through nucleic acids in the contact-molecule-contact system had initially started as coherent based tunneling transport. Within that limit, the intrinsic properties of the DNA were explored under simplifying assumptions, such as:

1) **No Backbone structure**: Only using the bases of the nucleotide, or
2) **Idealized structure**: Neglecting structure fluctuations caused by the environment (temperature, water, counterions)

with the Landauer-Büttiker formalism being the prominent approach to calculate the coherent transport properties [56], [76], [78]–[80]. However, the assumptions made were deemed oversimplifying to study the actual system and explain experiments. Especially, since it was found that as the length of the strand becomes longer, the environmental effects and the fluctuating nature of DNA play an important role. Electron would lose phase information, making the decoherent regime (sequential hopping) more important; thus, is more likely that transport is in an intermediate phase between the coherent and decoherent limits [76], [81], [82]. In modeling the charge transport in DNA, it is important to consider:

1) Decoherent Transport: As the length of the strand increases, the phase-breaking scattering events become significant. Therefore, one cannot disregard the decoherent transport mechanism.
2) Time Evolution of the Structure: The DNA is a soft molecule, constantly fluctuating due to the interaction with the solvent and surrounding counterions. These fluctuations affect the energy levels, hence changing the resulting electronic coupling as well as the energy distribution of the molecular orbitals.

Several approaches are reported in literature to account for these factors in charge transport: dissipative phonon-bath technique [83], [84], solving time-dependent equations that account for the variation in the energy levels of DNA bases with time [76], vibronic dephasing [85], [86], electron-phonon interactions [87] and the phenomenological Büttiker-Probes method [81], [82]. However, most of these models suffer from size limitations as they usually have high computational costs due to the added complexity in theory. These computationally intensive methods to include coupling to a phonon/vibronic bath include only a few energy levels on the DNA bases (having one to three energy levels per molecule). Although these models can qualitatively describe charge transport in DNA, using the full Hamiltonian from quantum mechanical calculations is important to explain the impact of small modifications to the DNA such as methylation [54], single nucleotide mismatches [51], and intercalation [88] on the charge transport. Therefore, adapting these models on relatively large ds-DNA systems of more than ten base pairs -which have more than 800 atoms- is currently unfeasible. One prominent model uses Büttiker probes to capture decoherence. These probes are physically modeled as self-energies in a

Green’s function formalism, which is used to calculate the conductance [81], [89], [90]. While the method is quantum mechanical, it treats decoherence in a phenomenological manner. This model can in principle account for phase breaking processes that are both elastic and inelastic and has been applied successfully in modeling other systems.

As for the time evolution of the structure, molecular dynamics (MD) simulations were introduced as a step prior to the transport calculation to include the effect of the environment on the structure. This yields the time evolution of atomic locations and results in a time varying electronic energy level distribution, thus varying the conductance profile. The MD simulations would produce many structures over the simulation time, which may go up to the microsecond range. Some groups take snapshots at every time step within the femtosecond range and the transmission would then be calculated per snapshot [91], [92]. Others would use statistical methods to extract the averaged energy distribution and the coupling parameters of the DNA strand from MD, and then calculate a conductance value [76], [93]. An interesting observation was made when studying the time evolution of the energy levels of the system. MD results have shown that the molecular orbitals (on-site potential) fluctuate and have a near Gaussian distribution. This is an important result as one of the main effects of the decoherence probe method is to broaden the on-site potentials yielding a Lorentzian shape.

The inclusion of the DNA solvent and counterions into charge transport modeling has started a debate on how the environment affects the transport process. Esther Conwell [94] has been an advocate for having the polarons as the main factor affecting charge transport in solvent. Polaron appears when a charge is moving in a polarizable medium, where two scenarios have been suggested in DNA case: 1) structural changes accompanying charge transfer through the bases, where the change in charge density on one base would create polarization of its neighboring bases, and hence creates changes in the spacing and π-coupling between them. 2) Environmental polarization, where water and counterions are affected by charge transfer. Thus, resulting in the shifting of energy levels surrounding the occupied state. One outcome of this model is that due to polaronic effects, the hole wavefunction is delocalized over several bases, yielding high transmission rates over long DNA strands [95]. On the other hand, Voityuk has published a theoretical analysis showing that polarons due to solvent polarization would cause localization rather than delocalization of the wavefunction [96]. Both Conwell and Voityuk use the same arguments in their favor, which adds to the confusion: The factors affecting localization/delocalization: 1) Polarization of surroundings, 2) internal organization of base-pairs, and 3) dephasing due to geometrical fluctuations of the DNA strands. An interesting comment made by Genereux and Barton in [97] is that the theoretical work in support of Voityuk mainly focused on averaged structures from MD simulations. As such, the delocalization supported by Conwell could be the result of configurations that occur only occasionally. In this thesis, our charge transport model considers the Born-Oppenheimer approximation (BOA). We keep the DNA coordinates fixed while doing the charge transport calculations and do not include the reorganization of the molecule. Our calculations exhibit both cases in which an energy level can be localized at a particular base (or base-pair) or distributed over multiple base-pairs, depending on the DNA sequence and conformation. In addition, the decoherence model induces energy level broadening that extends the overlap of these energy levels with respect to each other, increasing the transmission as a result.

Chapter 2. METHODOLOGY

The overall scheme of our work to simulate charge transport through nucleic acid structures is presented in Figure 2.1. First, MD simulations are carried out to include environment effect on the structure's geometry. It is important to note that only the DNA strand with water and counterions are included, and the contact electrodes are not included. Next, a representative structure is extracted to resemble the most-likely configuration, to be used for the following step. After removing solvent molecules and counterions (usually water and Na^+), a static structure serves as the input to density functional theory (DFT) calculations. We use Gaussian 16 [98] for this step, which produces the system Hamiltonian. Lastly, this Hamiltonian is used as an input for the transport calculation step which yields the transmission, conductance, and current-voltage characteristics of the strand. In this chapter, a brief background including the governing equations for MD and DFT are described first, then the charge transport model is described in full detail. The calculations in this thesis focus on steps 2 and 3 as the MD part is mainly done through collaborators (see acknowledgment section for further details).

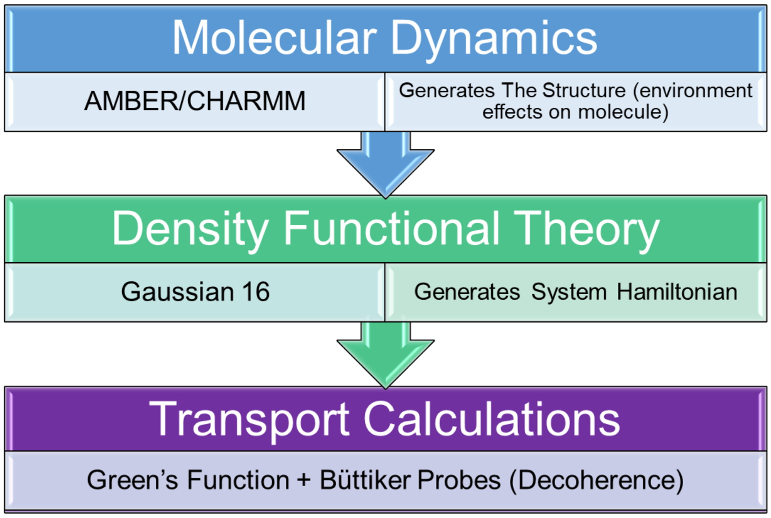


Figure 2.1 Flow chart of the simulation method, which consists of 1) MD simulations, followed by 2) DFT, and 3) Green's function approach with Büttiker probes.

2.1 MD

In MD simulations, the goal is to observe macroscopic variables of a system of particles (atoms/molecules) by obeying microscopic laws. By adding system constraints on some of the observables such as volume (V), temperature (T), pressure (P), total energy (U), or number of particles (N), the system can evolve with time. Simulating over a large period creates a statistical ensemble by which all microstates can be observed, this is known as the Ergodic hypothesis. The main outcome of this simulation is the geometrical rearrangement of atoms with time as they interact with each other.

Force fields (FF) are used to constitute the microscopic laws that capture the dependence of the system energy on the coordinates of its particles [99]. A force field is an analytical expression of the potential energy, which is separated into two categories

$$U = U_{bonded} + U_{nonbonded} \tag{2.1}$$

where U_{bonded} is the bonded potential, and $U_{nonbonded}$ is the nonbonded potential. The bonded potential resembles the covalently bonded atoms, and it consists of 4 terms

$$U_{bonded} = U_{bond} + U_{angles} + U_{torsion} + U_{improper} \tag{2.2}$$

where each term describes the intramolecular contributions to the total energy: bond stretching, angle bending, dihedral angle torsion, and improper inversion. As for the nonbonded term, it consists of two main contributions

$$U_{nonbond} = U_{ee} + U_{VDW} \tag{2.3}$$

where $U_{ee} = \sum \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$ is the Coulomb repulsion representing the electrostatic interaction, and U_{VDW} is the Van der Waals interaction, which is a dispersion force that resembles the effect of instantaneous change in electron position and how it creates a local dipole on the atom, inducing neighboring dipole on its surrounding atoms. Force fields may be generated using ab initio methods, semi-empirical quantum mechanical calculations, or by fitting to experimental data. Thus, they should be carefully used as per the system under study.

The main equation solved in MD is Newton's law of motion per atom i

$$F_i = m_i a_i = m_i \frac{\Delta v_i}{\Delta t} = m_i \frac{\Delta^2 r_i}{\Delta^2 t} \tag{2.4}$$

the force (F_i) can also be defined as the negative of the gradient of the potential,

$$F_i = m_i \frac{\Delta^2 r_i}{\Delta^2 t} = -\nabla U \tag{2.5}$$

with U defined based on the chosen FF. Starting from an initial position and velocity of the atoms, the new position at the next time step is solved for accordingly. The simplified treatment of the atomic interactions makes it possible to simulate systems containing up to 3 million atoms and reach the microsecond range of simulation time [100]. However, the choice of FF and system constraints plays an important role in obtaining sensible results.

2.2 DFT

In studying nanoscale systems, the Hamiltonian resembles the essential part for understanding the electronic structure. It contains the electronic energy levels of the system, as well as the coupling between these levels. These two properties are what defines charge transport through the structure. One of the main methods to find the Hamiltonian is an ab initio calculation. Ab initio methods consist of obtaining the electronic structure of a system by solving Schrodinger's equation using some approximations, but without parameter fitting to the extent seen in the MD approach. The time-independent Schrodinger's equation is

$$H\Psi(\mathrm{R},\mathrm{r}) = \mathrm{E}\Psi(\mathrm{R},\mathrm{r}) \tag{2.6}$$

where H is the Hamiltonian of the many-body system, Ψ is the many-body wavefunction, R is the position of the nuclei, r is the position of electrons, and E is the total energy of the system. It is exceedingly difficult to solve the coupled equation. Hence, the Born-Oppenheimer approximation is applied where it states that the relative nuclei motion is much slower than the electrons in the system since the protons present in the nuclei make it a much heavier particle than the electron. With this approximation, the wavefunction can be decoupled into a product of nuclear and electronic wavefunctions, and the electronic Hamiltonian can be written as [174]

$$H_e = -\sum_{i=1}^{N} \frac{\hbar^2}{2m}\nabla_i^2 + \frac{1}{2}\sum_{i<j}^{N} \frac{q^2}{|r_i - r_j|} + \sum_{i=1}^{N} V_{ext}(r_i) \tag{2.7}$$

where the first term resembles the kinetic energy, the second term is the electron-electron interaction, and the third term is the external potential, usually related to the interaction of electrons with the surrounding nuclei,

$$V_{ext}(r) = \sum_{A=1}^{M} \frac{Z_A}{|R_A - r|} \tag{2.8}$$

with Z_A the charge of the nucleus A, and the denominator is the distance between the electron and the nucleus.

The electron-electron interaction is the most problematic term when solving the N-particle equation, as it requires to solve a 3N-dimensional differential equations (3 is for the 3D coordinates). The many body problem has been approached by several simplifying methods, with Hartree Fock (HF) and Kohn-Sham theories the most famous ones. The basic idea of these approaches is to choose suitable non-interacting electron wavefunctions (orbitals) that can simulate the behavior of the full interacting system. HF uses the Slater determinants expansion to represent the orbitals, with a correction term added later by post-HF methods such as Møller-Plesset perturbation theory (MP2,MP3…,MP6) and Coupled Cluster (CCSD) to represent the electron-electron interaction, called the correlation term.

On the other hand, the Kohn-Sham approach, also commonly known as density functional theory (DFT), is based on two theorems: 1) the ground state of a system depends on the electron density, and 2) the electron density that minimizes the energy of the system is the exact density of the full N-particle interacting system. Therefore, by finding the non-interacting wavefunctions that can yield a good approximation of the exact density, we can solve the N-particle Schrodinger equation using the 1-particle system approach. Within this formalism, we get the single-particle Kohn-Sham equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{eff}(r)\right]\psi_i(r) = \epsilon_i\psi_i(r) \tag{2.9}$$

with the effective potential being,

$$V_{eff}(r) = V_{ext}(r) + \int \frac{\rho(r')}{|r-r'|}dr' + V_{XC}[\rho(r)] \tag{2.10}$$

the second term is the electron-electron Coulomb repulsion, with the density $\rho(r) = \sum_{i=1}^{N} |\psi_i(r)|^2$, and the third term resembles the exchange-correlation (XC) potential. The XC potential is a functional of electron density, and includes the effect of electron correlation, $V_{XC}[\rho(r)] = \frac{\delta E_{XC}[\rho(r)]}{\delta\rho(r)}$.

DFT helps transform the N-particle wavefunction problem into a single-particle equation, significantly reducing the complexity. Furthermore, DFT adds a modifying term -V_{XC}- to the Hamiltonian to yield the correct density of the ground-state system. However, this term is unknown, and therefore different functionals were derived over the years for different systems. Functionals such as the local-density approximation (LDA), the local spin-density approximation (LSDA), and the generalized gradient approximation (GGA) are widely used in modeling inorganic periodic materials. However, hybrid functionals are more commonly used for organic aperiodic systems such as the well-established Becke, 3-parameter, Lee-Yang-Parr (B3LYP) functional. The "hybrid" term comes from using a mixture (linear combination) of Hartree Fock exchange energy with Kohn-Sham exchange-correlation functionals.

After defining the functional for the system under study, the DFT self-consistently solves for the minimized energy level by first guessing the electron density, this yields the effective potential. Then the eigenvalue problem is solved, and a new density is calculated, yielding a new effective potential. The calculation converges when there are no changes in the output between the cycles. Different packages are available that employ DFT, such as SIESTA[102], VASP[103], Quantum Espresso[104], ADF[105], and Gaussian 16 [98]. In my work, I mainly use the Gaussian 16 package for the DFT calculations.

As one may notice, the ab initio approaches are more complex and computationally expensive than MD simulation, DFT is limited to less than 2500 atoms per run. Thus, in the DFT step of our methodology, the water molecules and counterions are removed from the system under study, only keeping the strand.

2.2.1 *DFT With B3LYP/6-31G(d,p)*

In studying charge transport through DNA, the ionization potential (IP) of the DNA bases or nucleotides is a figure of merit to compare between the quantum-mechanical calculation methods. In this current work, we use B3LYP functional with the 6-31G(d,p) basis set. This functional has been used before to calculate the ionization potential of nucleobases, either as a main calculation method [21], [106], or as part of methods comparison and benchmarking [107], [108]. These calculations demonstrated that B3LYP/6-31G(d,p) yields the correct trend as experiments, with an offset in values of about 300 meV [107]. Other computationally more expensive methods such as CCSD[106]and MP2[18] have been proposed to give more accurate ionization potential results, but what makes B3LYP/6-31G(d,p) a method of choice is its relatively low computational cost and relatively good accuracy. In addition, Felix and Voityuk [109] benchmarked the performance of DFT functionals for hole transfer parameters in DNA. They used the smallest basis set for atomic orbital polarizability (the 6-31G(d) basis set). They found that hybrid functionals such as B3LYP have better accuracy than nonhybrid functionals. We use the 6-31G(d,p) in our calculations, which further adds the p-type polarization to the hydrogen atoms better to describe the hydrogen bonds present in the DNA base-pairs. Moreover, ab-initio studies have shown that the polarizability of the solvent that surrounds the molecule is essential to obtain the correct energy levels and IP [18], [110]–[112]. The Polarizable Continuum Model (PCM) implemented in Gaussian 16 adequately accounts for the dielectric constant of the solvent and the associated screening of charge. For example, Slavicek et al. [18] have shown that modeling water through PCM yields DNA IP results that directly correlate with photoelectron spectroscopy measurements without the need to include explicit water molecules.

To display the performance of B3LYP/6-31G(d,p), I present next a comparison with literature for vertical IP (VIP) calculations and adiabatic IP of DNA bases and base-pairs. We ran calculations for DNA bases (C,G,A, and T) in water solvent using the PCM model, and in gas phase. Furthermore, the bases were studied with and without backbone (terminating the base with a hydrogen atom), using DFT with B3LYP/6-31G(d,p) in Gaussian 16.

VIP is attributed to the experimental photoluminescence results [18]. Theoretically, VIP can be calculated through the following equation:

$$VIP = E_{q-1} - E_{neutral} \tag{2.11}$$

where E_{q-1} is the total energy of the ionized molecule (missing one electron) and the atoms are in the same geometry as the neutral molecule, and $E_{neutral}$ is the total energy of the neutral molecule. The structures were generated using the nucleic acid builder (NAB) tool provided in AMBER package [100]. No structure optimizations were done, as it was single point calculation, only the positions of Na counterions -when added- were optimized with respect to the backbone. The difference in calculating VIP and adiabatic IP in Gaussian 16 revolves around the solvent reaction field in the PCM method used. For VIP calculation, the method is called nonequilibrium PCM (NEPCM) and the calculation protocol can be found in Gaussian documentation [113]. In this approach, the solvent reaction field for the second step (ionized molecule) is loaded from the initial calculation (neutral molecule). This ensures that the solvent remains fixed as vertical ionization takes place. As for the adiabatic IP, the use of regular PCM method yields this value.

The main results are summarized in Table 2-1. The calculated VIP values using the DFT method are lower than the values of the other methods reported in the literature, which is expected as DFT underestimates the HOMO-LUMO gap. Moreover, the gas-phase calculations for DNA bases (no BB) are off by ~500 meV from the range of the experiments, and the trend is T>C>A>G which agrees with the literature. The other interesting result is seen in the VIP (labeled NEPCM), as the trend changes to C>T>A>G. This result shows that VIP of cytosine overtakes thymine in the trend, even by less than 40 meV. PMP2 results found in [18] as well as DFT-BMK/6-31+G* in [111] agree with this finding (the latter is included between brackets in the table). Hence, the B3LYP/6-31G(d,p) is a viable option for studying nucleic acid structures.

Table 2-1 Vertical IP comparison between DFT-B3LYP functional with other methods and experiments reported in literature.

B3LYP/6-31G(d,p) **(DFT-BMK/6-31+G* from** [111]**) (eV)**				Experiment (eV) [106]
Nucleobase (BDNA)	Water (NEPCM) BB	Gas-Phase BB + Na	Gas-Phase No BB	Gas Phase No BB
A	6.91 **(7.53)**	7.2963	8.05 **(8.34)**	8.3-8.5
T	6.97 **(7.77)**	7.599	8.67 **(9.09)**	9.0-9.2
G	6.57 **(7.23)**	6.8439	7.50 **(8.08)**	8.0-8.3
C	7.02 **(7.83)**	7.3466	8.20 **(8.78)**	8.8-9.0
Trend	C>T>A>G	T>C>A>G	T>C>A>G	T>C>A>G

2.3 TRANSPORT MODEL

One of the most studied inelastic mechanisms is the electron-phonon interaction that is usually modeled within the well-established Non-Equilibrium Green's Function (NEGF) formalism [114]–[116]. However, in DNA strands, we find that this method fails because the electron-phonon interaction cannot be treated by first-order approximation perturbation theory as the Hamiltonian has terms comparable in magnitude to $\hbar$ times the scattering rates [78]. The details of the NEGF method testing can be seen in Chapter Chapter 4. For weakly-coupled systems, other methods have been proposed [78], [87], [117] but they are infeasible for structures such as DNA, which have a large number of atoms.

This has led us to use the scattering theory approach to include elastic and inelastic scattering with Büttiker probes [81], [118]. While the method is quantum mechanical, it treats decoherence in a phenomenological manner. Büttiker probes are fictitious probes that mimic the scattering event for an electron by drawing the scenario that an electron enters the probe, gets scattered (elastically or inelastically), and then gets re-injected into the system (see Figure 2.2). The model has been able to qualitatively explain experiments and describe different charge transport mechanisms through DNA [51], [77], [81], [90], [119], [120]. These probes can include the effect of decoherence where the energy of the electron does not change but phase information is lost (elastic scattering). They can also incorporate energy loss/gain (inelastic scattering) while traversing the system due to electron interaction with phonons or photons [121]. This model can capture multiple physical phenomena with an engineering-oriented approach using an appropriate set of decoherence parameters. This parameter is related to the average scattering time in which an electron loses its phase information or energy while residing at an energy level in the ds-DNA system. The shortcoming of this model, however, is that it does not include different phonon modes.

In most of the models mentioned previously, the electrode atoms are not explicitly included in the simulations due to the complex geometry of the electrode tip and the computational cost of including even a small cluster of metal atoms. Thus, the effect of contact atoms is included as a self-energy term that is added to the system Hamiltonian (the Hamiltonian represents the on-site energies and electronic coupling of the molecular orbitals of the system).

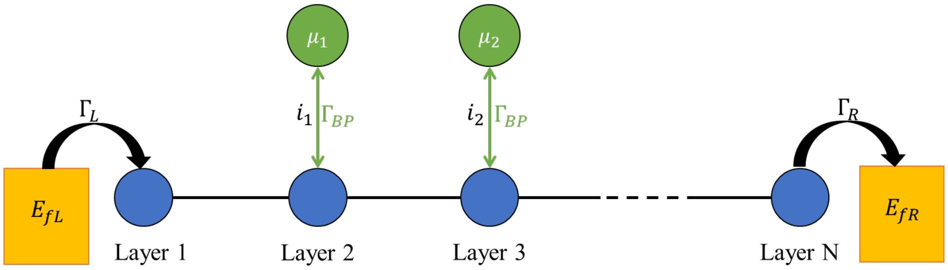


Figure 2.2 The Büttiker probe decoherence model. Each layer can be defined to represent an atom, a molecule, or a molecular group. The fictitious probes (green) are attached to each layer in the system and are represented by chemical potentials and scattering rates. Metal contacts at the terminals are represented by self-energy matrices (see text).

The converged DFT calculations result in the Fock (F) and overlap (S) matrices, which are then used as input in the charge transport calculations. We implement a Löwdin transformation to obtain the Hamiltonian (H_0) in an orthogonal basis set

$$H_0 = S^{-\frac{1}{2}} F S^{-\frac{1}{2}} \tag{2.12}$$

The Hamiltonian of the DNA (bases and backbone) can be expressed as

$$H_0 = H_0^D + H_0^{OD} \tag{2.13}$$

$$H_0^D = \sum_{\substack{\alpha=1\to b_i \\ i=1\to N_{DNA}}} \epsilon_{\alpha,i}\, c_{\alpha,i}^{\dagger}\, c_{\alpha,i} \tag{2.14}$$

$$H_0^{OD} = \sum_{\substack{\alpha=1\to b_i \\ \alpha'=1\to b_j \\ i,j=1\to N_{DNA}}} t_{\alpha,i\to\alpha',j}\, (c_{\alpha,i}^{\dagger}\, c_{\alpha',j} + H.c) \tag{2.15}$$

Here, $\epsilon_{\alpha,i}$ stands for the on-site energy α of the i^{th} atom, $c^{\dagger}$ and c are the creation and annihilation operators, respectively. $t_{\alpha,i\to\alpha',j}$ is the hopping parameter between orbital α at atom i

and orbital α' at atom j, and $H.c$ is the Hermitian conjugate. b_i is the number of basis functions (or states) representing atom i, N_{DNA} is the total number of atoms in the DNA strand.

In terms of matrices, one can think of H_0^D (H_0^{OD}) as a diagonal (off-diagonal) matrix. The diagonal elements of H_0 represent the energy levels at each atomic orbital, and the off-diagonal elements correspond to the hopping parameter between them. The size of the Hamiltonian H_0 corresponds to the total number of basis functions in the system ($\sum_{i=1}^{N_{DNA}} b_i \times \sum_{i=1}^{N_{DNA}} b_i$).

We then partition the DNA into blocks to apply decoherence at each block. The idea behind the partitioning is to how we perceive the electron traveling through the DNA system as it interacts with the environment and loses phase information or energy. DNA is not a crystalline material with identical repeating unit cells. The ds-DNA typically has a HOMO-LUMO gap of 3~4 eV and hopping parameter between its nearest neighbor base-pairs of 10-100 meV [78], [122], [123]. We can thus consider the DNA as a weakly-coupled wide bandgap semiconductor. The DNA sequence, the unique molecular orbitals, and the weak coupling properties give rise to energy barriers for charge transport. In this thesis, the partitioning scheme is considered nucleotide partitioning, unless indicated otherwise.

We show an example of a nucleotide (backbone + base) partitioning scheme in Figure 2.3a. We apply a unitary transformation using the eigenvectors of each partition. This partitioning transforms the original Hamiltonian H_0 into a block diagonal matrix H,

$$H = U^{\dagger} H_0 U \tag{2.16}$$

where the diagonal blocks of H are now diagonal matrices (see the matrix in Figure 2.3b). The diagonal elements of the diagonal blocks represent the eigenvalues of that block (or partition). The off-diagonal blocks of H represent the electronic coupling between the molecular orbitals of the blocks. The size of each diagonal block of H is equal to the total number of basis functions (O_m) used to represent the atoms in the block (or partition) m,

$$O_m = \sum_{i=1}^{N_m} b_i \tag{2.17}$$

where N_m is the total number of atoms in partition m. The matrix U to accomplish the unitary transformation is

$$U = \begin{bmatrix} u_1 & 0 & & \cdots & & 0 \\ 0 & u_2 & & & & \\ & & \ddots & & & \vdots \\ \vdots & & & u_m & & \\ & & & & \ddots & 0 \\ 0 & & \cdots & & 0 & u_N \end{bmatrix} \tag{2.18}$$

where, u_m is a diagonal sub-matrix consisting of the eigenvectors of partition m, and it has the size of $(O_m \times O_m)$. N is the total number of partitions in the DNA strand. Note that in the case of *atomic partitioning*, N will be the total number of *atoms* in the DNA, whereas it becomes the total number of *nucleotides* for the *nucleotide partitioning*.

The redefined Hamiltonian H in second quantized notation is given by:

$$H = H^D + H^{OD} \tag{2.19}$$

$$H^D = \sum_{\substack{k=1\to O_m \\ m=1\to N}} \epsilon_{k,m}\, c_{k,m}^{\dagger}\, c_{k,m} \tag{2.20}$$

$$H^{OD} = \sum_{\substack{k=1\to O_m \\ k'=1\to O_{m'} \\ m\neq m'=1\to N}} t_{k,m\to k',m'}\,(c^{\dagger}_{k,m}\,c_{k',m'} + H.c) \tag{2.21}$$

where $\epsilon_{k,m}$ is the k^{th} on-site energy (molecular orbital) of the m^{th} block, and $t_{k,m\to k',m'}$ is the interaction between the k^{th} molecular orbital of the m^{th} block with the k'^{th} molecular orbital of the m'^{th} block.

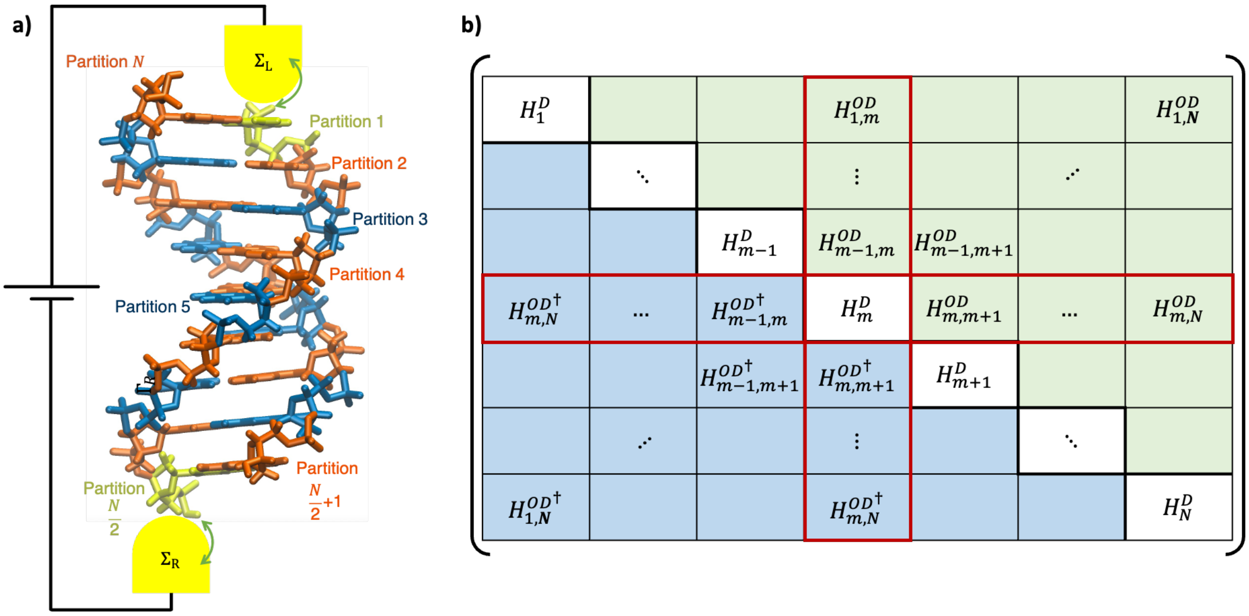


Figure 2.3 a) Schematic of the charge transport calculations setup. The ds-DNA is connected to two electrodes through the 3’- and 5’- ends (nucleotides colored in yellow). The two partitions colored in yellow define the contacts self-energy locations at the 3’- and 5’-end nucleotides. The other two alternating colors of the DNA nucleotides resemble the partitioning scheme considered in the decoherent transport model. b) The Hamiltonian H after the partitioning into blocks. The diagonal blocks contain the on-site potentials of each partition, and the off-diagonal blocks contain the coupling between the partitions (in green and blue, where blue is the transpose of the green). The highlighted partition m in red is to help guide the reader to its row and column. The green (or blue) portion of the mth row and column contains the coupling between partition m with all the other partitions in the system.

The Hamiltonian H defined by equation (2.19) above is used to find the retarded Green’s function (G^r) of the DNA by including both the self-energies due to decoherence and the contacts,

$$[E - (H + \mathit{\Sigma}_{\mathrm{L}} + \mathit{\Sigma}_{\mathrm{R}} + \mathit{\Sigma}_{\mathrm{B}})]G^r = I \qquad (2.22)$$

where E is the energy. $\mathit{\Sigma}_{\mathrm{L(R)}}$is the left (right) contact self-energy, which represents the coupling of the DNA to the left (right) contacts through which electrons enter and leave the DNA (Figure 2.3a). The self-energy for decoherence is represented by $\mathit{\Sigma}_{\mathrm{B}}$. Both contacts and decoherence self-energies are diagonal matrices, which means they are only coupled to H^D(the molecular orbitals) in equation (2.19). We implement the self-energies in accordance with the partitioning scheme; each decohrence probe is connected to a block, while leaving two blocks to be connected to the contacts self-energies. This implementation creates $N_B = N - 2$ decoherence probes. The matrix representation of the decoherence self-energy is

$$\begin{gathered} \mathit{\Sigma}_{\mathrm{B}_{mm'}}(k,k') = \ \mathit{\Sigma}_k \, \delta_{kk'} \, \delta_{mm'} \\ \mathit{\Sigma}_{\mathrm{B}_{L,L}} = 0 \\ \mathit{\Sigma}_{\mathrm{B}_{R,R}} = 0 \\ \\ \{m, m' \in \mathbb{N}: 1 \le m, m' \le N_B\} \\ \{k, k' \in \mathbb{N}: 1 \le k, k' \le O_m\} \end{gathered} \qquad (2.23)$$

where $\mathit{\Sigma}_{\mathrm{B}_{mm'}}$ represents the subblock of $\mathit{\Sigma}_{\mathrm{B}}$ as per the partitioning scheme. At each diagonal block $\mathit{\Sigma}_{\mathrm{B}_{mm}} = \mathit{\Sigma}_k \, \delta_{kk'}$, the diagonal element $\mathit{\Sigma}_k$ represents the decoherence self-energy associated with molecular orbital k of block m. Here, $\delta_{mm'}$ and $\delta_{kk'}$ create nonzero diagonal elements only at the molecular orbitals of block m. Hence, only the diagonal blocks $\mathit{\Sigma}_{\mathrm{B}_{mm}}$ have nonzero diagonal matrices of size $(B_m \times B_m)$. $\mathit{\Sigma}_{\mathrm{B}_{L,L(R,R)}}$ resemble the two contact blocks that do not have decoherence probes attached to them.

To model decoherence, we consider both the real and imaginary parts of the self-energy,

$$\Sigma_{\mathrm{B}}(E) = Re[\Sigma_{\mathrm{B}}(E)] + i \times Im[\Sigma_{\mathrm{B}}(E)] \tag{2.24}$$

where i is the imaginary unit $\sqrt(-1)$. The real part is a measure of an energy-dependent shift in the on-site potential (molecular orbital) of the DNA. The imaginary part of the self-energy is a measure of electron flow between the DNA and decoherence probe. Both parts are related to each other by the Kramers-Kronig relation [124],

$$Re[\Sigma_{\mathrm{B}}(E)] = \frac{1}{\pi} P \int \frac{Im[\Sigma_{\mathrm{B}}(E')]}{E' - E} dE' \tag{2.25}$$

where P is the Cauchy principal value. We define the imaginary part at each diagonal element as

$$Im[\Sigma_B(E)] = -\frac{\Gamma_{k,m}(E)}{2}\, \delta_{kk'}\, \delta_{mm'} \tag{2.26}$$

where $\Gamma_{k,m}$ represents the coupling strength between the probe at block m and the molecular orbital k in that block. We note that we omitted the indices of the left-hand side for clarity.

In the *energy-dependent* (*E-dep*) treatment, we define Γ_k to have an exponential decay factor as a function of energy (E) at each molecular orbital $\epsilon_{k,m}$

$$\Gamma_{k,m}(E) = \Gamma_B \times \exp\left[-\frac{|E-\epsilon_{k,m}|}{\lambda}\right] \tag{2.27}$$

where λ is an exponential decay parameter.

The *energy-independent* (*E-indep*) decoherence model corresponds to having $\lambda = \infty$ in equation (2.27), which makes $\Gamma_{k,m}(E) = \Gamma_B$, energy independent. Substituting this into (2.25) gives,

$$Re[\Sigma_{\mathrm{B}}(E)] = \frac{\Gamma_B}{\pi} P \int \frac{1}{E' - E} dE' \tag{2.28}$$

That is, the real part of the self-energy is zero if the imaginary part is a constant independent of energy.

Regarding the self-energy of the contacts, we used the wide-band limit (WBL), which ignores the real part of the matrices and treats them as energy independent parameters. This approximation stands when the DOS is almost a constant throughout the metal, which is true for gold [125]. Thus, it is treated as an energy-independent coupling parameter and is defined as,

$$\Sigma_{L(R)} = -\frac{i\Gamma_{L(R)}}{2} \delta_{kk'} \delta_{mm'} , \qquad m, m' = L(R) \tag{2.29}$$

After setting up the Green's function equation (2.22), the density of states at block m and energy E can be computed by extracting the corresponding diagonal elements of the retarded Green's function

$$DOS(m, E) = -\frac{Im\left[diag\left(G_m^r(E)\right)\right]}{\pi} \tag{2.30}$$

We will now discuss some detail of the model for decoherence following reference [118], [126]. In the low-bias regime, the current at the m^{th} probe is

$$I_m = \frac{2q}{h} \sum_{n=1}^{N} \int_{-\infty}^{+\infty} T_{mn}(E)[f_m(E) - f_n(E)]dE = \int_{-\infty}^{+\infty} J_m(E)dE \tag{2.31}$$

$$I_m = \frac{2q}{h} \sum_{n=1}^{N} T_{mn}[\mu_m - \mu_n], \quad m: 1 \to N \tag{2.32}$$

$$T_{mn}(E) = \mathrm{Tr}\left[\left(-2 \times Im[\Sigma_B(E)]_{m,m}\right) \times G^r_{m,n} \times \left(-2 \times Im[\Sigma_B(E)]_{n,n}\right) \times G^a_{n,m}\right] \quad (2.33)$$

where q is the elementary charge, h is the Planck's constant, T_{mn} is the transmission between the m^{th} and n^{th} probes, and $G^a = (G^r)^\dagger$ is the advanced Green's function. $f_m(E) = \left[1 + \exp\left(\frac{E-E_{fm}}{kT}\right)\right]^{-1}$ is the Fermi distribution, and $J_m(E)$ is the current at probe m and energy E. In equation (2.33), the terms in parentheses correspond to the coupling strength to the contact or the decoherence probes (as per equations (2.26) and (2.29)). In this thesis, we calculate the transmission at room temperature, thus kT = 0.0259 eV.

2.3.1 *Elastic (Dephasing) Decoherence Probe*

The main impact of including decoherence probes is to broaden the energy levels as illustrated in Figure 2.4. The elastic model has been previously developed in our group by Jianqing Qi [81], where it only allows for electrons to scatter and lose phase information. The main condition to implement it here is that the current at each energy of each decoherence probe should be zero, $J_m(E) = 0$. This yields N_B independent equations from which the following relation can be derived

$$\mu_m - \mu_L = \left(\sum_{m=1}^{N_B} W^{-1}_{mn} T_{nR}\right)(\mu_R - \mu_L), \;\; n = 1 \rightarrow N_B \quad (2.34)$$

here, W^{-1}_{mn} is the inverse of $W_{mn} = (1 - R_{mm})\delta_{mn} - T_{mn}(1 - \delta_{mn})$, where R_{mm} is the reflection probability at probe m, and is given by $R_{mm} = 1 - \sum_{m\neq n}^{N} T_{mn}$. The currents at the left (I_L) and right (I_R) obey the relation $I_L + I_R = 0$. This yields the equation for the current at the left contact as

$$I_L = \frac{2q}{h} T_{eff}(\mu_L - \mu_R) \quad (2.35)$$

Comparing equations (2.32) to (2.35) yields the effective transmission term as

$$T_{eff} = T_{LR} + \sum_{m,n}^{N_B} T_{Lm} W_{mn}^{-1} T_{nR} \tag{2.36}$$

where the first term is the coherent transmission from the left electrode to the right electrode. The second term is the decoherence contribution into the transmission via the decoherence probes. The low-bias conductance as a function of Fermi energy (E_f) is calculated as

$$G(E_f) = \frac{2q^2}{h} \int dE\, T_{eff}(E) \frac{\partial f(E - E_f)}{\partial E} \tag{2.37}$$

2.3.2 *Inelastic (Energy-Exchange) Decoherence Probe*

In the inelastic model, the electron can scatter and lose/gain energy in the process. This results in electrons entering and exiting the system with different current distributions $J_m(E)$. The main difference in implementing inelastic scattering lies in the current condition. Here, the total integrated current at each decoherence probe must equal to zero, but not at each energy point (compared to elastic condition of $J_m(E) = 0$), that is

$$I_m = \frac{2q}{h} \sum_{n=1}^{N} \int_{-\infty}^{+\infty} T_{mn}(E)[f_m(E) - f_n(E)]dE = \int_{-\infty}^{+\infty} J_m(E)dE = 0 \tag{2.38}$$
$$m = 1,2,3,\dots,N_B$$

where the Fermi distribution functions at each probe m are defined by the chemical potentials μ_i, which we now solve for iteratively to satisfy the condition in equation (2.38). In this model, the condition yields N_B number of nonlinear equations, which are solved via the Newton-Raphson method [115], [121]

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_{N_b} \end{bmatrix}_{itr+1} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_{N_b} \end{bmatrix}_{itr} - \alpha \begin{bmatrix} \frac{\partial I_1}{\partial \mu_1} & \cdots & \frac{\partial I_1}{\partial \mu_{N_b}} \\ \vdots & \ddots & \vdots \\ \frac{\partial I_{N_b}}{\partial \mu_1} & \cdots & \frac{\partial I_{N_b}}{\partial \mu_{N_b}} \end{bmatrix}_{itr}^{-1} \begin{bmatrix} I_1 \\ \vdots \\ I_{N_b} \end{bmatrix}_{itr} \tag{2.39}$$

$$\frac{\partial I_m}{\partial \mu_m} = \frac{2q}{h} \int \sum_n T_{mn} \frac{\partial f_m(E)}{\partial \mu_m} dE\,, \qquad \frac{\partial I_m}{\partial \mu_n} = \frac{2q}{h} \int T_{mn} \frac{\partial f_n(E)}{\partial \mu_n} dE \tag{2.40}$$

where itr is the iteration number, and $0 < \alpha < 1$ is a pre-factor that helps with achieving convergence if the default updating step was too large. Convergence can be achieved using an initial guess of linear ramp for each chemical potential within the bias range $E_{fR} < \mu_i < E_{fL}$, where $E_{fR} < E_{fL}$. The convergence criteria are met when the leakage current to each probe is small ($|I_i|/|I_L| < 10^\wedge - 5$) and μ_m does not change after the itr^{th} iteration.

After reaching convergence, the current injected into the system I_L can then be calculated via

$$I_L = \frac{2q}{h} \sum_{n=1}^{N} \int_{-\infty}^{+\infty} T_{nL}(E)[f_L(E) - f_n(E)]dE = \int_{-\infty}^{+\infty} J_L(E)dE \tag{2.41}$$

where J_L is the injected current density, and even though the ejected current $I_R = -I_L$, the ejected current $J_R(E) \neq J_L(E)$ due to inelastic scattering.

In the inelastic scattering model, a nonzero bias is applied in the calculations to provide an appreciable energy window for electrons to scatter. The effect of applying a nonzero bias to the molecule is included as a linear onsite potential shift, via the relation

$$\begin{gathered} \epsilon_{k,m}(V_{bias}) = \epsilon_{k,m}(V_{bias} = 0) - \left(V_{bias} \cdot \frac{\mathrm{m}-1}{N-1}\right), \\ \{m, m' \in \mathbb{N}: 1 \leq m, m' \leq N\} \\ \{k, k' \in \mathbb{N}: 1 \leq k, k' \leq O_m\} \end{gathered} \tag{2.42}$$

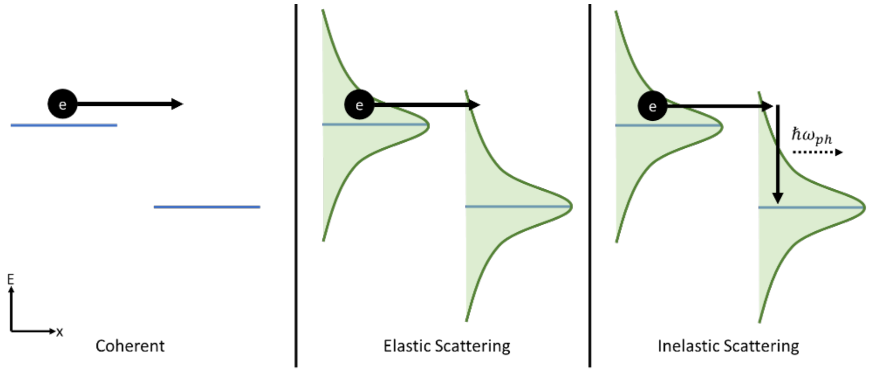


Figure 2.4 Different charge transport mechanisms. Coherent tunneling transport which deals with the actual sharp onsite potentials of the system, yielding larger barriers to overcome. Elastic scattering due to dephasing induces broadening to the onsite potentials, lowering the energy barrier and yielding higher transmission. Inelastic scattering allows for vertical movement (energy) along with the broadening effect.

2.3.3 *Inelastic Scattering Testing and Analysis*

The initial testing on the inelastic decoherence probe consists of applying it on a model Hamiltonian shown in Figure 2.5 (top). The Hamiltonian consists of two onsite potential values $\epsilon_{1,2}$=[0.25 eV or 0.75 eV] distributed randomly on two atoms per layer, with a total number of layers set to 10, t_1=1.5 eV and t_2=0.1 eV resembling strong and weak coupling, respectively. This would mimic the Hamiltonian of a DNA as it yields a wide bandgap with fluctuating conduction and valence bands as seen in Figure 2.5 (bottom). Contact self-energies are attached to the first and last layers with $\Gamma_{L(R)}$=200 meV, while *E-indep* decoherence probes are applied to the rest with Γ_B=10 meV. The simulations are carried at room temperature (298 K).

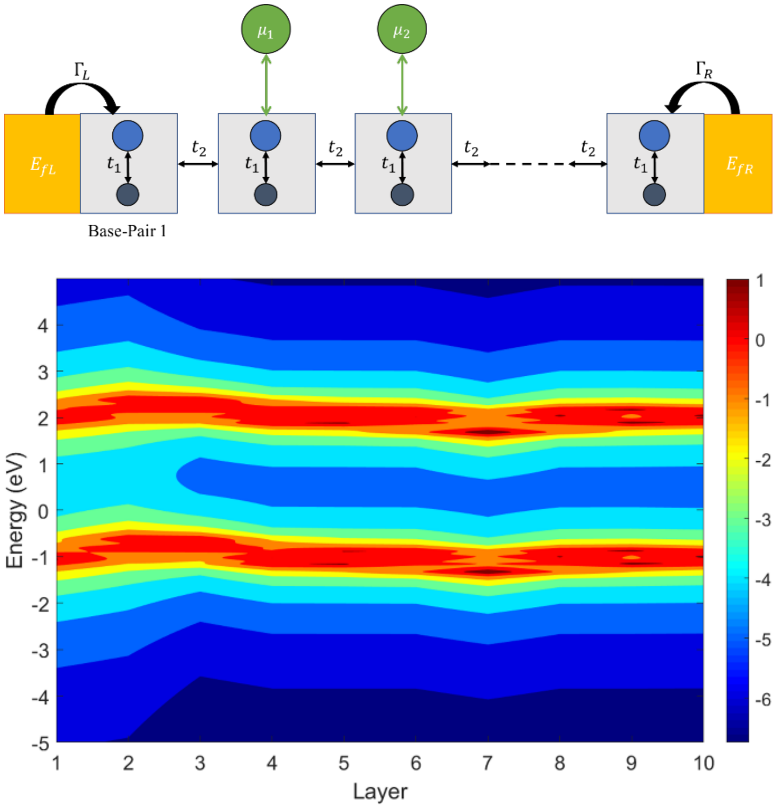


Figure 2.5 **Top** Model Hamiltonian used; each block consists of two atoms with random onsite potentials coupled with t_1. The blocks are coupled to their nearest neighbors by t_2, contact self-energy is applied to the two terminals of the system. **Bottom** Contour DOS plot of the resulting system with each layer representing two atoms coupled by t_1.

First, we investigate the converged chemical potential profile. We set the Left contact Fermi energy (E_{fL}) to 0 eV and the right contact $E_{fR} = E_{fL} - qV_{bias}$, where V_{bias} =100 meV. The choice of E_{fL} puts the bias energy window to be in the bandgap (BG) close to the valence band. The converged calculation yields the chemical potentials of the probes, plotted in Figure 2.6. We notice that the high resistance seen for operating in the BG drops the chemical potential across the device, which is expected [115].

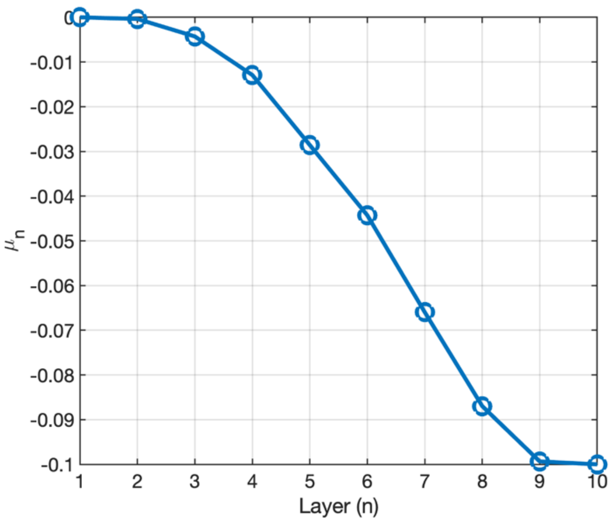


Figure 2.6 Converged chemical potentials of the decoherence probes (from layers 2-9, layers 1 and 10 are the contact potentials).

Next, we look at the applied bias impact on the elastic and inelastic current values. For this test, we fix E_{fL} at 0 eV, with $E_{fR} = E_{fL} - qV_{bias}$ at varying bias values, and the results are summarized in Table 2-2. At a very low bias of 10 mV, the system can be considered to be operating close to the equilibrium regime as the onsite potential shift between consecutive bases is small at 10 mV bias. The inelastic current is higher than the elastic current by ~3 orders of magnitude. However, as we apply higher bias (500 mV), we shift into the non-equilibrium regime where potential drop across the device has a higher shift on the onsite potentials (see equation (2.42). In this case, the inelastic current becomes comparable to the elastic current. At an extremely large bias (2000 mV), surprisingly, the elastic current is now higher than the inelastic current.

Table 2-2 Elastic and inelastic current comparison under different bias.

Applied Bias (mV)	I_{inel} **(nA)**	I_{el} **(nA)**	**Trend**
10	0.112	7.60E-5	$I_{inel} > I_{el}$
500	0.476	0.170	$I_{inel} > I_{el}$
2000	2.14	2.98	$I_{inel} < I_{el}$

To understand this trend, we find that the operating window (E_{fL} location) plays an essential role for inelastic scattering. Figure 2.7 shows a contour plot of the DOS as a function of both position and energy with the region of operation highlighted in the black box. The density of states is low at the lower energy levels, especially as we move to layers 7-10. For inelastic scattering, there is a competition to occupy the lower states between electrons injected from the contacts at higher energies and electrons entering at lower energies. It becomes difficult for the electrons at higher states that scatter and lose energy to occupy the lower states and travel to the right contact. Pauli's exclusion principle prevents electrons from occupying an already occupied state, hence, lowering the conductance. This can be seen from the middle plot in Figure 2.7 as it shows the inelastic current at the inject (J_L) and extract (J_R) layers. The inelastic J_L is localized at ~0 eV even though the window goes to -2 eV, and J_R is mostly localized around -2 eV. In contrast, the elastic current in Figure 2.7 (right) is more distributed as electrons are not allowed to lose/gain energy. The peak in the elastic current seen at approximately -1.6 eV corresponds to the longest spatial distribution of high DOS values.

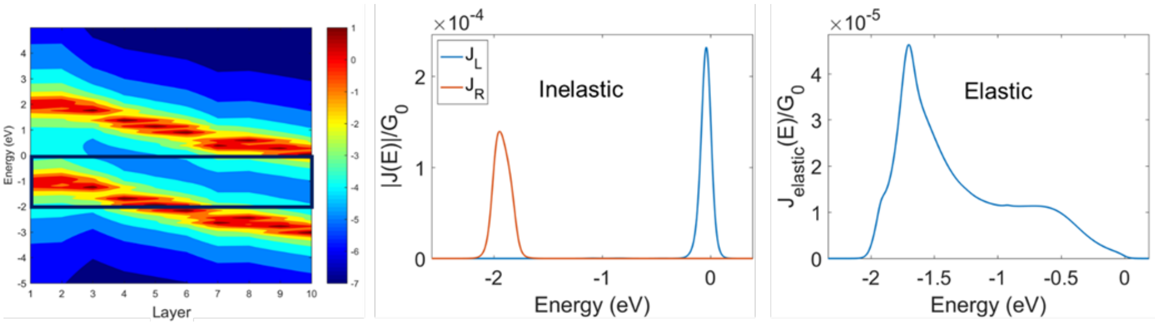


Figure 2.7 Contour DOS plot is shown on the left, the window highlights region of operation for E_{fL}=0 eV and V_{bias} = 2000 mV. Elastic current as a function of energy is shown in the middle, displaying higher delocalization than inelastic current in the right figure.

The picture becomes clear as we shift E_{fL} to -0.75 eV in Figure 2.8. This would allow for higher density of states to be present at lower energies as the electron traverses the structure. Here,

J_L and J_R have higher energy distribution and the inelastic current is several orders of magnitude higher than the elastic current.

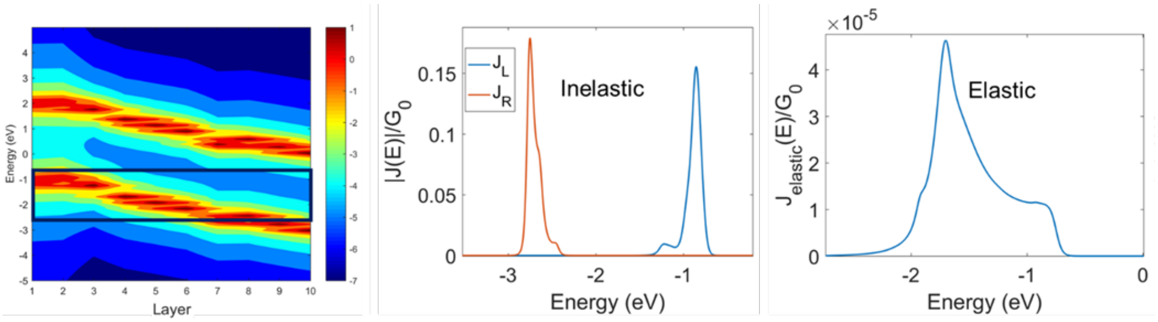


Figure 2.8 Contour DOS plot, window highlights region of operation for E_{fL}=-0.75 eV and bias = 2000 mV. Elastic current as a function of energy is shown in the middle. Inelastic current on the right shows higher delocalization at the new Fermi position, yielding larger inelastic current.

Next, we apply the inelastic scattering model on a real DNA Hamiltonian and compare the calculated conductance with the elastic scattering model. We used DNA with sequence 3’-$G_3A_9G_3$-5’ from **Chapter 6** and partitioned the system into nucleotides to apply the inelastic scattering but kept the contact self-energy value and decoherence rate the same as the main calculation setup in section 6.2. Therefore, each layer is now defined as a base+backbone.

We calculate the effective voltage sweep to find the conductance at different Fermi energies. This helps sweep the energy spectrum and find the conductance within different energy regions (HOMO, Bandgap, and LUMO regions). The effective gate voltage sweep calculation is done by changing left contact Fermi energy (E_{f_L}) to V_{gate} value per run, and fixing the applied bias (V_{bias}), with the right contact Fermi energy defined as $E_{f_R} = E_{f_L} - qV_{bias}$. Phase decoherence (elastic) and inelastic scattering are calculated at 50 mV bias. In Figure 2.9, the results show that when either operating in the HOMO or LUMO region, the inelastic scattering conductance is larger than elastic conductance, where it increases the current by 3-4 times. However, within the bandgap region, inelastic scattering does not show a considerable increase in conductance from the elastic model. Hence, inelastic scattering does not show significant change on the transmission of the system within the HOMO-to-bandgap region. The differences between elastic and inelastic scattering conductance values are seen lying deep in the HOMO and LUMO regions where the DOS is high.

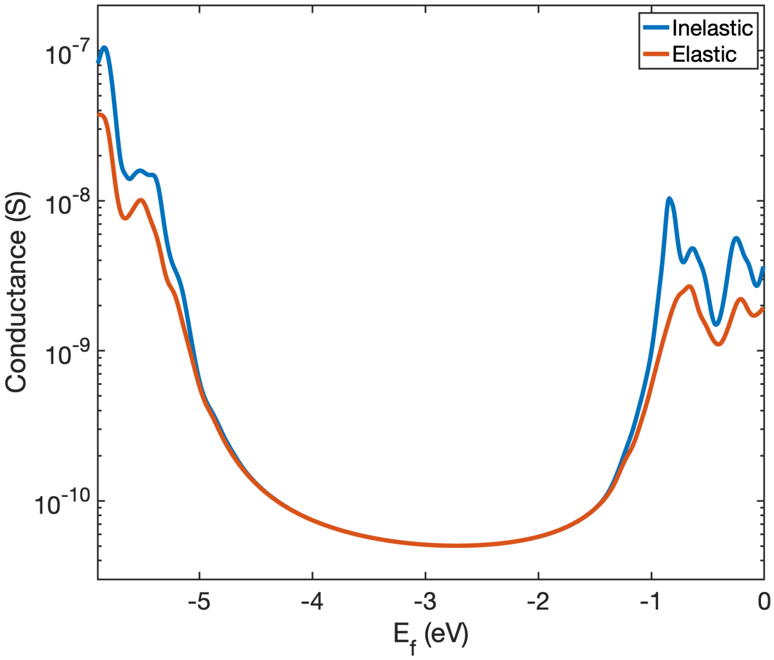


Figure 2.9 Effective gate voltage sweep results of ds-DNA at V_{bias} = 50 mV. The difference between the elastic and inelastic scattering models appears inside high DOS regions (deep in HOMO and LUMO regions).

The main outcome from this analysis is that as electron traverses from left contact (higher energy) to right contact (lower energy), inelastic scattering enhances the current as long as there are states at lower energies for the right-moving electrons to scatter into.

Chapter 3. ENERGY-DEPENDENT DECOHERENCE MODEL

3.1 INTRODUCTION

Understanding and controlling the electrical conductivity of nucleic acids (DNA) has gained more interest in the past decade. Measuring DNA conductance for sensing biological processes [51], [54], [69], [127], developing new sequencing techniques[33], [71], [72], and future molecular device applications [13], [60]–[63], [128]–[132] have led to an interest in its electrical properties. DNA origami exploits the self-assembly property of DNA to create complex three-dimensional architectures [50]. This technique helps build nanoscale structures bottom-up instead of the top-down approach currently used in nanoelectronics. Further, the ability to build heterostructures [133] and dope [61]–[63], [132] DNA makes it a potential candidate for nanoelectronics beyond lithography. Therefore, understanding how charge transports through nucleic acids can help engineer a new class of biosensors and nanoelectronics [13], [54], [58], [88], [123]. From a fundamental viewpoint, the precise mechanism of charge transport in DNA is not fully understood. The electronic hopping between consecutive bases along a ds-DNA varies from 10-100 meV [78], [122]. Theoretical methods that probe realistic DNA strands with many hundreds of atoms by accounting for the (i) backbone, (ii) sequence, (iii) transport of electrons between the complementary strands, and (iv) environment consisting of a solvent are essential.

In probing the electrical transport, we consider a system that consists of two metal contacts with DNA in between them (Figure 2.3a). It has been shown previously that coherent transport calculations yield conductance values that are orders of magnitude smaller than experimental values [81]. The calculated conductance becomes comparable to experiments by including decoherence to account for the realistic DNA effects mentioned above. Decoherence has been modeled at various levels of theory. Gutiérrez et al. used a dissipative phonon-bath technique [83], [84], Kubar et al. model decoherence by solving time-dependent equations that account for the variation in the energy levels of DNA bases with time [76], while Karasch et al. model decoherence due to including vibronic dephasing [85]. These computationally intensive methods to include coupling to a phonon/vibronic bath include only one to three energy levels on the DNA bases. Although these models can qualitatively describe the small hopping parameter between DNA bases, using the full Hamiltonian is important to explain the impact of the backbone, methylation of bases [54], single nucleotide mismatches [51], and intercalation [88].

One prominent model uses Landauer-Büttiker approach to capture decoherence. In this approach, the Hamiltonian accounts for all coherent scattering events while the decoherence is treated at a phenomenological level using *decoherence probes* (also known as *Büttiker probes*). These probes are physically modeled as self-energies in a Green's function formalism, which is used to calculate the conductance [81], [89], [90]. This model can in principle account for phase breaking processes that are both elastic and inelastic and has been applied successfully in modeling other systems. Central to this model are spatially dependent decoherence parameters that represent the average scattering time in which an electron loses its phase information. The model is adept at including the effect of density of states (DOS) broadening. In the coherent limit, the DOS peaks along the length of the DNA have almost no energetic overlap, making transport difficult. Including the decoherence via Büttiker probes broadens the DOS peaks significantly and hence increases the transmission probability across the DNA length [81]. The Büttiker probes model has qualitatively explained experiments and described different charge transport mechanisms through DNA such as direct tunneling, sequential hopping, and intermediate coherent-decoherent regime [51], [77], [81], [90], [119], [120].

A major drawback of prior work is that the decoherence rates are *energy-independent* -an unphysical assumption-. While this assumption simplifies the problem considerably, allowing one to neglect the real part of the self-energy because it is precisely zero, it yields some qualitatively incorrect results. First, the transmission inside the bandgap is found to be unphysically large. As a result, there is only a weak dependence of transmission as a function of length in the bandgap. Second, interesting features in the transmission are washed out within the conduction and valence bands. When the *energy-independent* decoherence model is used in a crystalline semiconductor, to make the transmission small in the bandgap, the decoherence rate is set to zero (or a small value). This treatment is unphysical because it does not conserve the norm of the DOS integrated over all energies.

To summarize, in modeling transport through DNA, we must carefully consider the impact of the decoherence model on both the transmission and the integral of the DOS. Notably, the model should allow for exponential decay of transmission with the length within the bandgap and maintain features in the transmission spectra within the bands. In this chapter, we modify the phenomenological decoherence model to make the decoherence rate energy-dependent. The role of the real part of the self-energy on both the transmission and the integral of DOS are also discussed.

We follow the methodology for elastic decoherence mentioned in section 2.3, we compare the *energy-dependent* and *energy-independent* decoherence models (see equations (2.23 to (2.28) with a model Hamiltonian for the DNA. We then discuss the role of the real part of the self-energy. Finally, we discuss the transport results with the full Hamiltonian of the double stranded DNA and compare the results of our model to experiments.

3.2 EXPONENTIAL DECAY WITH LENGTH

The ds-DNA typically has a HOMO-LUMO gap of 3~4 eV and hopping parameter between its nearest neighbor base-pairs of 10-100 meV [78], [122], [123]. We can thus consider the DNA as a weakly-coupled wide bandgap semiconductor. With this large energy gap, we expect the transmission near the middle of the gap (i.e., the midgap) to drop exponentially with an increase in strand length. To demonstrate the importance of the *E-dep* model, we use a model Hamiltonian for the DNA where each block is a base-pair (Figure 3.1) to calculate the transmission as a function of length. Each base-pair of the model Hamiltonian has two identical energy levels to represent HOMO (ϵ_H = -4 eV) and LUMO (ϵ_L = 0 eV). We set the hopping parameters between the nearest neighbor blocks (t_1) and the inter-strand hopping parameter (t_2) to 100 meV and 1 meV, respectively. For simplicity, we set the left (right) contacts coupling $\Gamma_{L(R)}$ = 1000 meV at the end terminal blocks of this system.

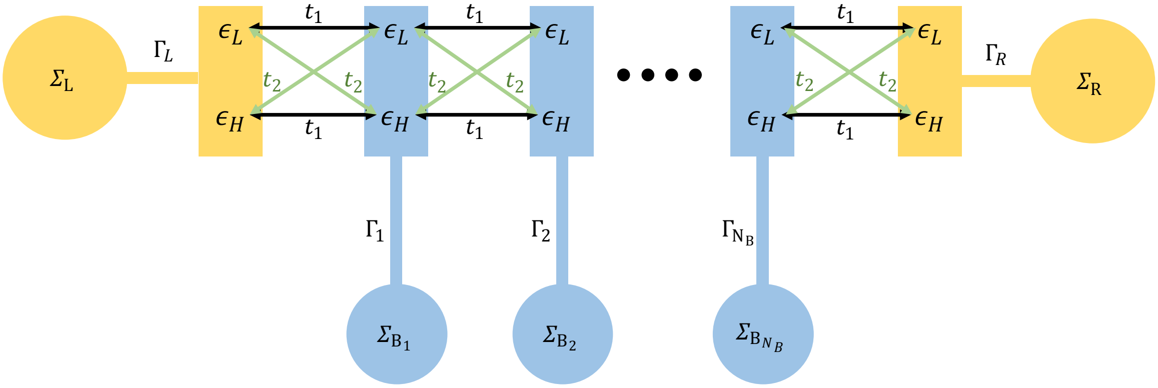


Figure 3.1 Model Hamiltonian representing a ds-DNA with N_B+2 number of blocks. Each block resembles a base-pair with two onsite potentials resembling HOMO (ϵ_H) and LUMO (ϵ_L), the nearest neighbor hopping parameter is t_1 and the inter-base hopping parameter is t_2. N_B number of blocks are connected to the decoherence probes, and the two terminal blocks are connected to the contact self-energies.

We calculate the transmission in the coherent limit and using *E-indep* and *E-dep* decoherent models. For the *E-indep* model, we set the decoherence rate to a relatively low value of Γ_B = 1 meV. As for the *E-dep* model, we increase the decoherence rate by 100 times to Γ_B = 100 meV with varying the energy decay factor λ = [10, 50, 100, 500] meV. We plot the results for a DNA strand with length = 7 base-pairs (bps) in Figure 3.2a, showing that the coherent transmission decays as we enter the bandgap region. Similarly, the *E-dep* model has an exponentially decaying tail into the bandgap (from HOMO or LUMO edges). We also note that increasing both Γ_B and λ can cause the *E-dep* model to deviate from the coherent limit deep in the bandgap. For low values of λ (10 meV), the E-dep model deep in the bandgap follows the coherent limit. With moderate values of λ (50 and 100 meV), the E-dep model starts to vary in the bandgap but follows the coherent limit at midgap. However, for large values of λ (500 meV), the *E-dep* model yields similar values to the *E-indep* model, which is unphysical. This behavior is because increasing λ makes the decoherence rate larger off-resonance (away from ϵ_r) as seen in the exponential-decaying factor of equation (2.27). As for the *E-indep* model, it does not maintain the decay deep in the bandgap. Initially, the transmission drops as we enter the bandgap but then saturates -within 200 meV from HOMO or LUMO- at a relatively high value ($T>10^{-10}$).

Next, we plot the transmission extracted at the midgap as a function of length, shown in Figure 3.2b. At λ = 10, 50 meV, the *E-dep* model and coherent transport have comparable exponential decay in transmission with increasing length. Increasing λ to 100 meV lowers the decay at longer strands. Similar to the impact of large λ values (500 meV), the transmission in the *E-indep* model only decreases by a small amount for length ≥ 7 bps. However, we note that the value of Γ_B for the *E-indep* model is a hundred times smaller than the *E-dep* model. Therefore, even at low decoherence rates, the transmission of the *E-indep* model is unrealistically high at energies deep in the bandgap where no states are available for the electron to occupy.

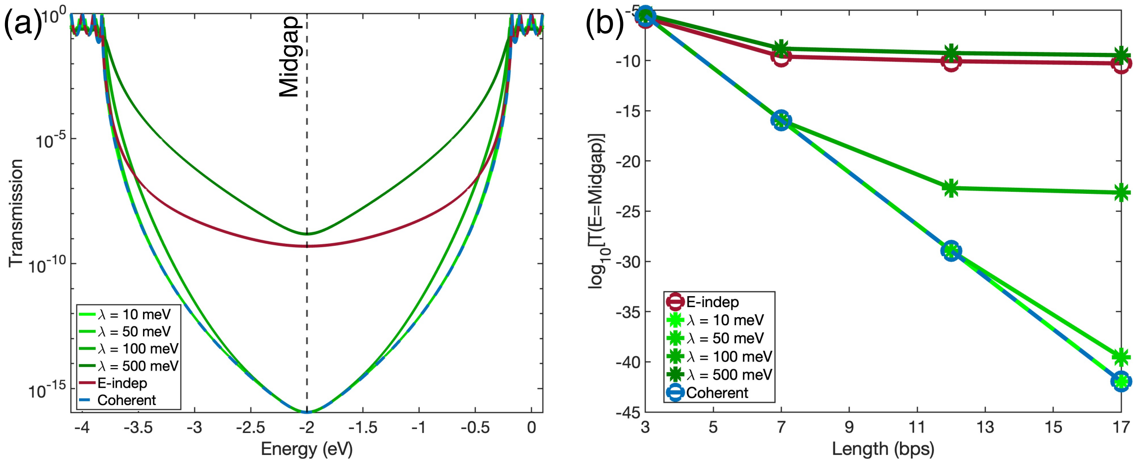


Figure 3.2 (a) Transmission plot for the DNA model Hamiltonian with length = 7 base-pairs. (b) Transmission extracted at the midgap and plotted as a function of the DNA length. Calculation parameters are *E-dep*: Γ_{B} = 100 meV and λ = [10,50,100,500] meV, *E-indep*: Γ_{B} = 1 meV, and $\Gamma_{\mathrm{L(R)}}$ = 1000 meV for all cases.

3.3 THE IMPACT OF THE DECOHERENCE PROBE SELF-ENERGY TREATMENT

The real and the imaginary parts of the self-energy are related by the Kramers-Kronig relation (equations (2.24)-(2.27)). The number of molecular orbitals per nucleotide is approximately 380 orbitals. As a result, there are ~11,400 orbitals for a 15 base-pairs long strand. The mathematical form of the self-energy used is dictated by the physical need to decrease the unphysically large DOS in the bandgap of the DNA induced by the E-indep model. The real part of the self-energy should be calculated from the imaginary part, and this will require the evaluation of a large number of integrals.

To understand the physical implication of including $Re[\Sigma_{\mathrm{B}}(E)]$, we first plot the real and imaginary parts of the self-energy in Figure 3.3 for an onsite potential ϵ_m = 0 eV. The real part of the self-energy $Re[\Sigma_{\mathrm{B}}(E)]$ shifts the onsite potential ϵ_m. The imaginary part is inversely proportional to the scattering due to the decoherence probe, which causes a broadening of the density of states. Notice how the amount of shift and broadening decays at energies away from resonance (ϵ_m = 0 eV). Next, we quantitively compare the density of states of the model Hamiltonian: (i) using the real and imaginary parts of the decoherence self-energy as defined in equation (2.24), and (ii) omitting the real part and defining $\Sigma_{\mathrm{B}}(E) = \mathrm{i} \times Im[\Sigma_{\mathrm{B}}(E)]$.

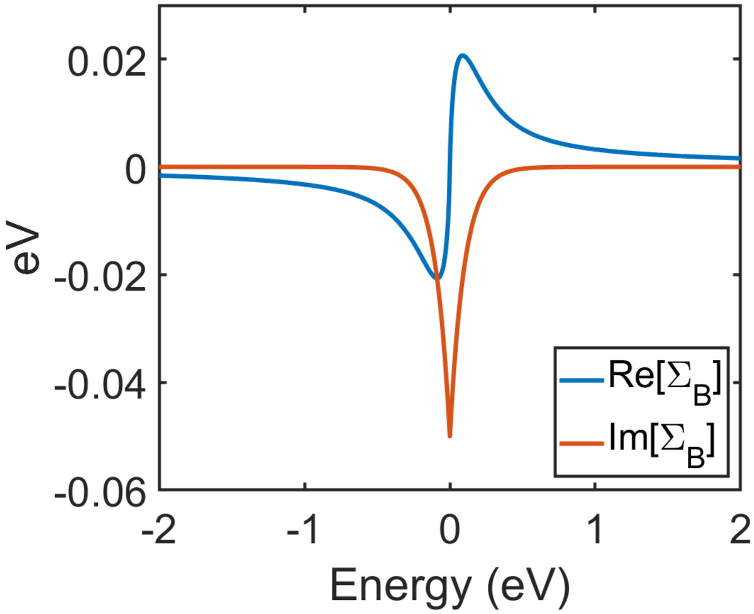


Figure 3.3 The real and imaginary parts of the self-energy obtained from evaluating equations (2.24)-(2.27), with $\Gamma_{\mathrm{B}} = \lambda = 100$ meV.

For this example, we consider the structure defined in Figure 3.1 with 7 bps. In Figure 3.4 a,b, we can see that the impact of including the real part of the self-energy is to shift the DOS peaks by approximately the value of the $Re[\Sigma_{\mathrm{BP}}(E)]$. However, as we enter the HOMO-LUMO gap, the impact of the real part is minimal and both treatments converge. Similarly, the impact on the resulting transmission follows the same trend as the DOS (see Figure 3.4 c,d). We then calculate the integral of DOS from -2000 eV to 1996 eV, with a fine mesh of 1 meV around the onsite potentials $[\epsilon_{H(L)} \pm 1.5$ eV], and a coarse mesh of 10 meV otherwise. The total number of energy levels is 14 (two per block). Thus, the expected integral of DOS is 14 (excluding the electron spin multiplicity). Including both the real and imaginary parts of the self-energy yielded

$\int DOS\, dE$ = 14, as expected. However, including only the imaginary part of the self-energy yielded $\int DOS\, dE$ = 13.461. This result implies that the total density of states is not conserved when scattering is included, which is incorrect [134].

Although the full self-energy treatment is essential to yielding the correct integral of DOS, the qualitative and quantitative differences in energy windows around the HOMO are small. These results show that we can safely neglect the real part of the self-energy for calculations that do not require a self-consistent integration of DOS to achieve convergence. Therefore, in the remainder of this chapter, we neglect the real part and use $\Sigma_{\mathrm{B}}(E) = \mathrm{i} \times Im[\Sigma_{\mathrm{B}}(E)]$ to model the *energy-dependent* decoherence.

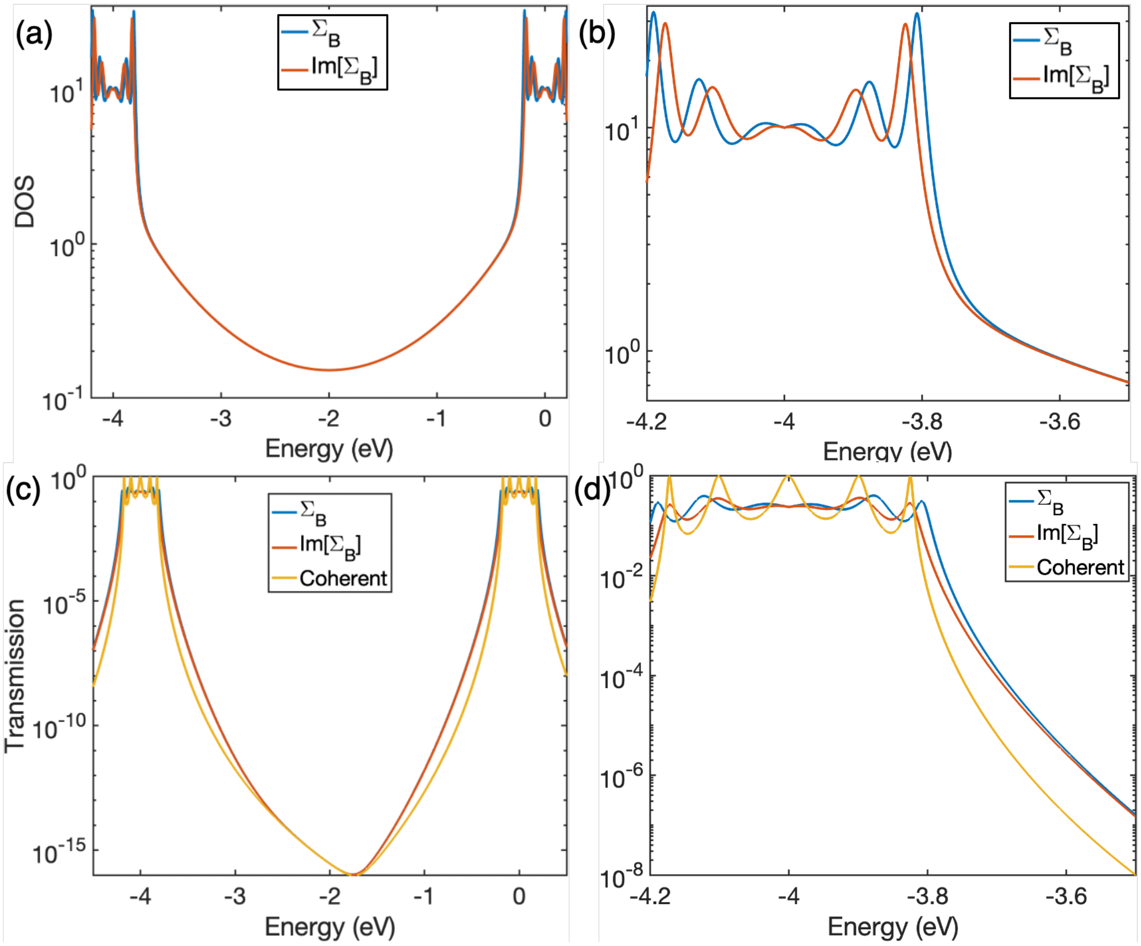


Figure 3.4 a,b Density of states plots using the full complex decoherence probes self-energy and only the imaginary part of the self-energy. c,d Transmission plot using the full complex decoherence probes self-energy and only the imaginary part of the self-energy, in addition to the coherent transmission. (a) The DOS plot over the energy spectrum displaying HOMO, HOMO-LUMO gap, and LUMO regions. (b) Zoom-in to the edge of HOMO region. The calculation parameters are $N_B = 5$, $\Gamma_B = 100$ meV, $\lambda = 100$ meV, and $\Gamma_{L(R)} = 1000$ meV. (c) Transmission over the energy spectrum displaying the coherent transmission for comparison. (d) Zoom-in to the edge of HOMO region. The calculation parameters are $N_B = 5$, $\Gamma_B = 100$ meV, $\lambda = 100$ meV, and $\Gamma_{L(R)} = 1000$ meV.

3.4 THE DNA SYSTEM

The molecular building blocks of π-stacked and conjugated organic semiconductors with fixed atomic coordinates have sharp energy levels. In bulk form, these molecular building blocks are weakly coupled. It is experimentally known that these semiconductors have a broadened density of states with an exponentially decaying tail into their bandgap [135]–[139]. Several factors contribute to the broadening: thermal disorder (electron-phonon interactions), structural disorder, dopant impurities, and randomly distributed defects.

A single DNA strand consists of π-stacked aromatic rings and has a bandgap of 3-4 eV. The distribution of energy levels is sequence-dependent and varies along the length of the strand (see Figure 3.5). The non-rigid nature of the DNA molecule and its ionic/solvent environment results in a broadening of its molecular orbitals. Previous work showed that capturing the effect of this energy level broadening is necessary to explain experiments [81]. The *E-indep* decoherence model used in the previous work missed important aspects (the unphysical increase of transmission in the bandgap) that we discussed in section 3.2.

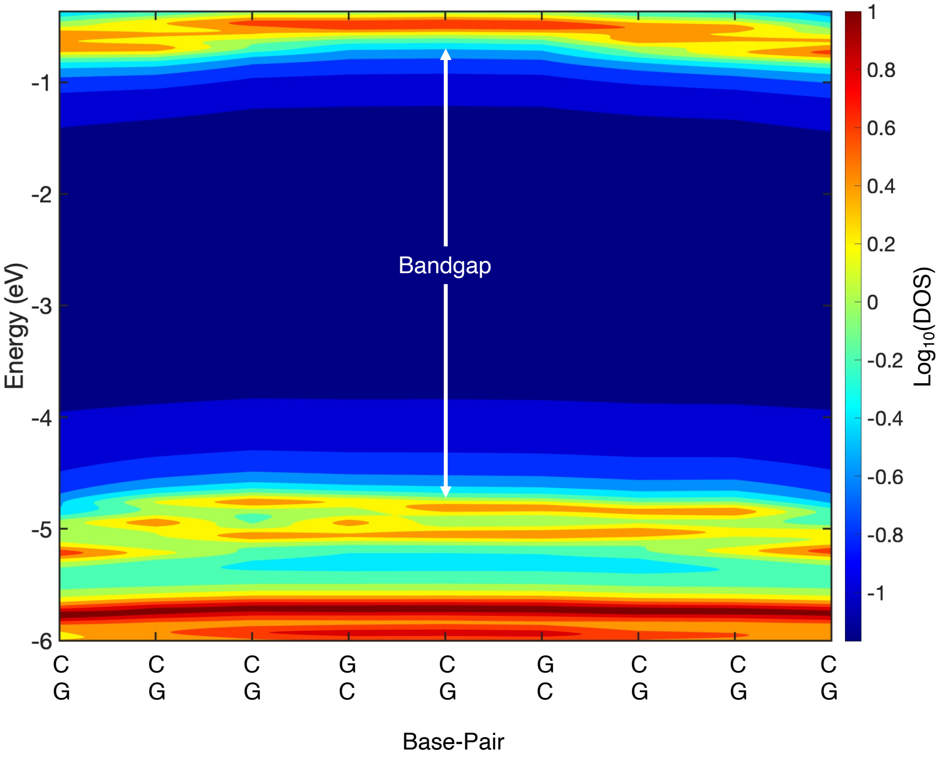


Figure 3.5 Contour DOS plot of 3’-C_3GCGC_3-5’. The plot shows the nonuniform spatial and energy distribution of the molecular orbitals along the length of the strand. The calculation parameters are Γ_B = 100 meV, λ = 50 meV, and $\Gamma_{L(R)}$ = 600 meV.

In applying the E-dep model to DNA, a challenge is determining the values of Γ_B and λ. Therefore, we look at the literature to determine the decoherence rates of other organic molecules. Ref [137] found the DOS broadening of aromatic molecules such as pentacene to reach 300 meV. In addition, Parson's work [140], [141] estimates the decoherence rates for solvated aromatic molecules to be in the 60-130 meV range. Using similar decoherence values in the *E-indep* model heavily broadens the DOS and washes out the energy dependence of transmission in DNA and makes the transmission in the bandgap unphysically large. That is why it is crucial to consider the *E-dep* model.

3.4.1 *System Under Study*

We study the double-strand B-DNA sequence 3'-$C_3G[CG]_lC_3$-5', with l = 1-4 in a contact-DNA-contact setup as shown in Figure 2.3a. The strands are 9-mer, 11-mer, 13-mer, and 15-mer, which correspond to varying l = 1-4.We model the DNA using DFT and calculate transmission and conductance using Green's function method with decoherence probes to account for the decoherence. We obtain the atomic structure of the DNA strands using the nucleic acid builder in Amber [142]. We perform DFT calculations using the B3LYP functional and 6-31G(d,p) basis set. We include the effect of water solvent around the DNA *via* the polarizable continuum model (PCM) using the integral equation formalism (IEFPCM). For the charge transport calculations, we assumed the contacts to be at the cytosine nucleotides at both the 3'- and 5'- ends of the strands (as shown in).

3.4.2 *Choosing the Contact Coupling*

To minimize the impact of the contact coupling value on the results, we calculate the transmission for 3'-C_3GCGC_3-5' under different coupling rates $\Gamma_{L(R)}$ = [100-10,000] meV. The low value of 100 meV resembles weak coupling to the electrodes, where values between 500-1000 meV resemble strong coupling. We also added the extreme values of 5000 and 10,000 meV to see their impact on transmission Figure 3.6 shows that the transmission increases with increasing the coupling value. We concentrated the analysis on the energy window extending from HOMO to HOMO+260 meV.

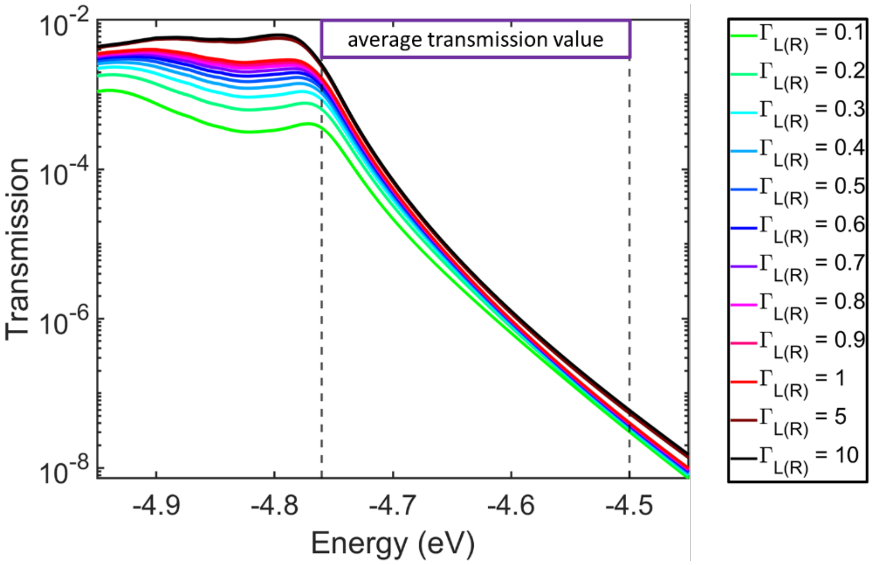


Figure 3.6 Transmission plot for 3’-C_3GCGC_3-5’ at $\Gamma_B = \lambda$ =100 meV with different contacts coupling. The average transmission is taken from the energy window HOMO to HOMO+260 meV. Legend shows the values in eV.

We plot the average transmission as a function of $\Gamma_{L(R)}$ in Figure 3.7. We can see that increasing the coupling by 10 times from the low coupling limit of 100 meV to 1000 meV increases the average transmission by 3.5 times. We also observe that once we have a strong contact coupling ($\Gamma_{L(R)} > 500$ meV), the increase in transmission starts to saturate. For instance, increasing the coupling from 500 meV to 1000 meV increases the transmission by less than 30%. Therefore, using $\Gamma_{L(R)} > 500$ meV can be sufficient to limit the contact coupling impact on the transmission amplitude.

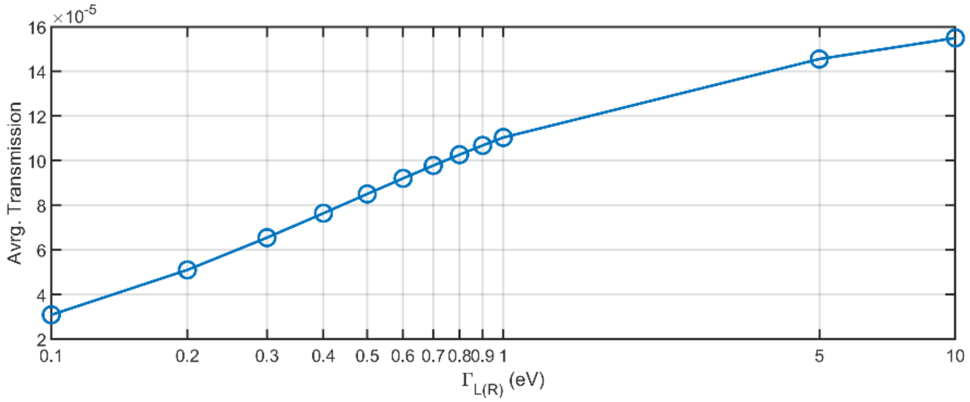


Figure 3.7 Average transmission as a function of contacts coupling. The x-axis is plotted in log scale for clarity.

Focusing on the coupling value of 600 meV, we plot the transmission increase with respect to $\Gamma_{\mathrm{L(R)}} > 600$ meV (see Figure 3.8). The average increase in transmission (blue curve) is only 20% when we increase the coupling from 600 to 1000 meV (66.67%). In addition, the average increase in transmission is less than 70% when increasing the coupling to 10,000 meV (1600%).

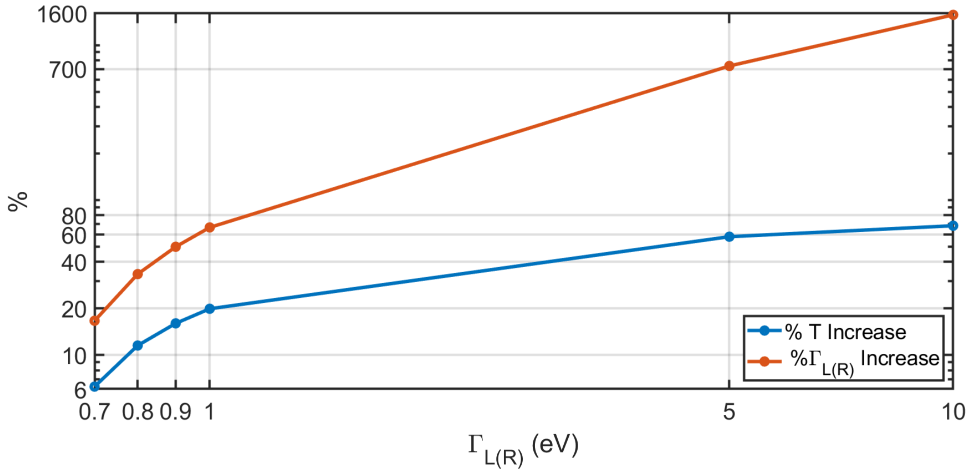


Figure 3.8 Percentage of Transmission (blue) and contact coupling (red) increase with respect to $\Gamma_{L(R)}$= 600 meV.

To summarize, we vary $\Gamma_{L(R)}$ from 100 meV to 10,000 meV and find that our results do not change significantly for $\Gamma_{L(R)} > 600$ meV. We are interested in a physical scenario where the contact resistance is smaller than the intrinsic resistance of the DNA. Hence, we set the contact coupling to be 600 meV in the following calculations.

3.4.3 *DOS and Transmission*

To highlight the main differences between the *E-dep* and *E-indep* models, we start with examining the density of states of l =1 case and vary the exponential decay term λ = [50, 100, 150, and 200 meV] while keeping the decoherence parameter fixed at Γ_B = 100 meV. We plot the DOS in Figure 3.9, and it shows that the *E-indep* model heavily broadens and washes out the variation between high (peaks) and low (valleys) DOS regions. We also see that for the *E-dep* model, as λ increases, the DOS peak broadening increases. As a result, the nearby peaks start to merge. The low-DOS regions (valleys) deep inside the occupied energy levels (E~-5.5 eV) and in the bandgap (E~-4.5 eV) are higher than their *E-dep* counterparts by more than an order of magnitude. The transmission results are consistent with these observations. Figure 3.10 shows the transmission plot when Γ_B = 100 meV while changing λ. We notice that the *E-dep* model has relatively the distinct peaks and valleys.

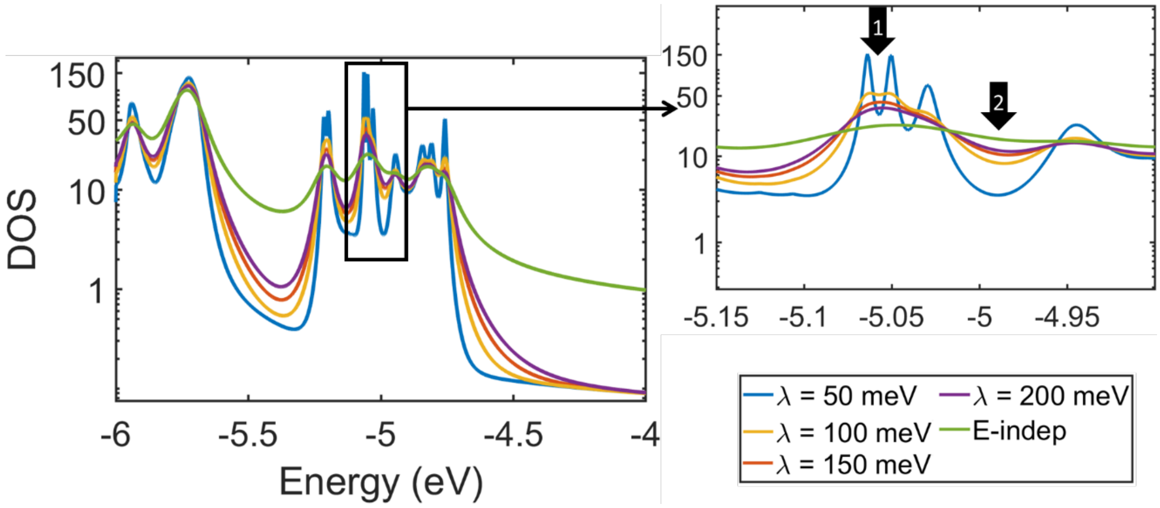


Figure 3.9 Total DOS for l=1 (3’-C$_3$GCGC$_3$-5’). The inset shows the highlighted region with arrows pointing at the two main observations: 1) increasing λ increases the broadening and nearby peaks start to merge, 2) Increasing λ increases low-DOS regions found between the high-DOS peaks due to the higher broadening.

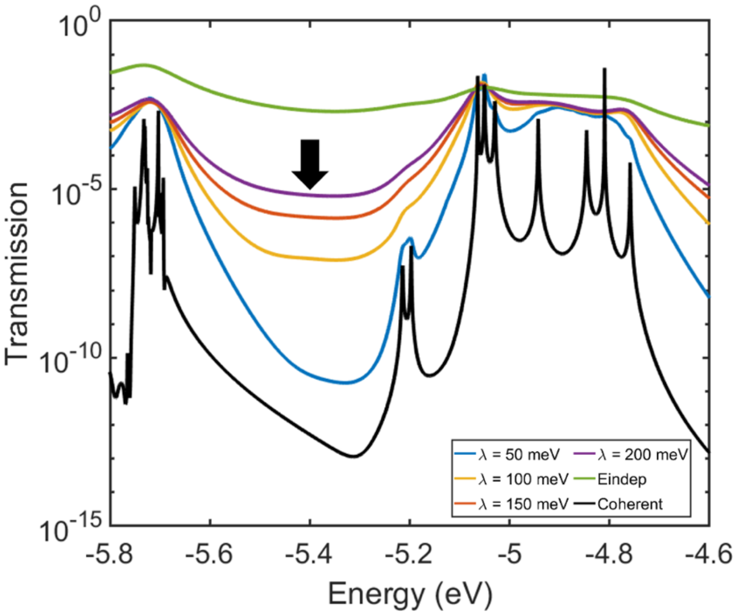


Figure 3.10 Transmission vs energy for 3’-C$_3$GCGC$_3$-5’. For both *E-dep* and *E-indep* models Γ_B = 100 meV. The arrow is pointing at the transmission in an energy gap between two mini-bands (~-5.7 eV and -5.2 eV) within the HOMO region.

We plot the DOS for a molecular orbital localized at the first three CG base-pairs at the 3’-end of the strand in Figure 3.11a. The calculated full-width-half-maximum (FWHM) of the DOS peak as a function of λ is shown in Figure 3.11b. We find that the peak width increases monotonically with increasing λ and approaches the line shape of the *E-indep* model.

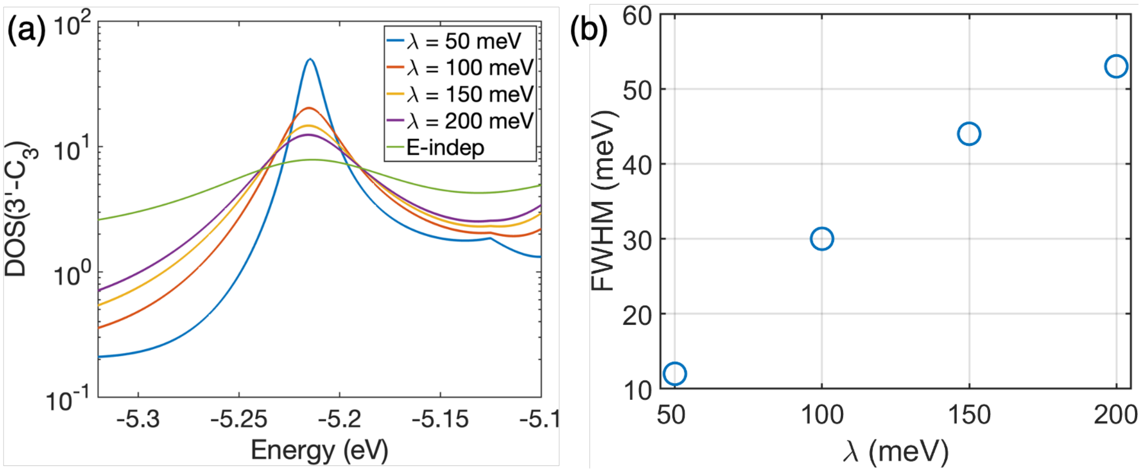


Figure 3.11 DOS of molecular orbital localized at the first three GC base-pairs within the strand 3'-C_3GCGC_3-5' at . (a) DOS of Molecular orbital at -5.225 eV. (b) Full-Width-Half-Maximum of the DOS peak under different λ values, with Γ_B = 100 meV.

In our DNA calculations, we find that if λ is set to 10 meV or less, the transmission spectrum would be closer to the coherent transport limit (see Figure 3.12). In addition, as we have discussed above, significantly increasing λ can cause the transmission near the midgap to become unphysically large, acting more like the *E-indep* model. Therefore, in the following section, we limit λ to be within 10-150 meV.

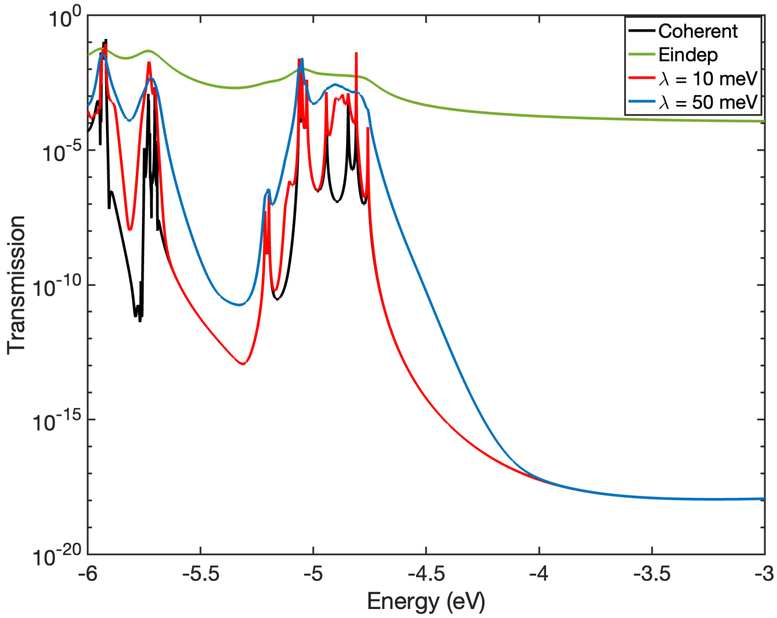


Figure 3.12 Transmission of 3'-C3GCGC3-5' strand comparing λ=10 meV with the coherent transport, λ=50 meV, and the *E-indep* model.

3.4.4 *DNA Conductance versus Length*

As a case study for the *E-dep* model, we investigate the relationship between the DNA strand length and the conductance. We consider four B-DNA strands 3'-C3G$[CG]_l$C3-5'(l = 1-4) and study the conductance as a function of length. Experimentally, this family of strands shows an exponential decay in conductance with length, but modeling yielded a much weaker dependence of conductance with length [77]. A likely reason for this inadequacy is the large density of states in the bandgap as a result of using the *E-indep* decoherence model. We showed in Figure 3.2 that the *E-indep* model indeed leads to an incorrect dependence of transmission versus length.

We model the strands by fixing the decoherence rate at Γ_B =100 meV and let of λ range from 10 to 150 meV. We calculate the conductance using equation (2.37). We assume that 1) the comparison between the gold work function (5.3 eV) and the ionization potential of the DNA nucleotides puts the expected contact Fermi energy (E_f) closer to the HOMO in the HOMO-

LUMO gap [123], and 2) we are operating at the low-bias regime; hence, it is unlikely that E_f will go below the HOMO (i.e., $E_f \geq \mathrm{HOMO}$). The Fermi energy is difficult to determine because it depends on the details of the contact geometry. As a result, we are guided by the approach in Ref [143], which involves a) sweeping the Fermi energy to gain an understanding of the experimental system being modeled and b) determining the Fermi energy when the calculated conductance is equal to the experimental value (see Figure 3.13). The results display low sensitivity to λ except at 10 meV, where for l = 1 and 2, the conductance is lower than the experimental value over the entire HOMO-LUMO gap (Figure 3.13 a,b). From this analysis, we plot the extracted conductance vs. strand length as shown in Figure 3.14a. The length is calculated using the relation 3.4 Å × number of base-pairs, where 3.4 Å is the axial rise in B-DNA (the shortest distance between neighboring bases along the helical axis of a strand). The motivation is to compare with the experimentally reported exponential decay of conductance with strand length, having a decay factor (β) of 0.196 Å^{-1} [77].

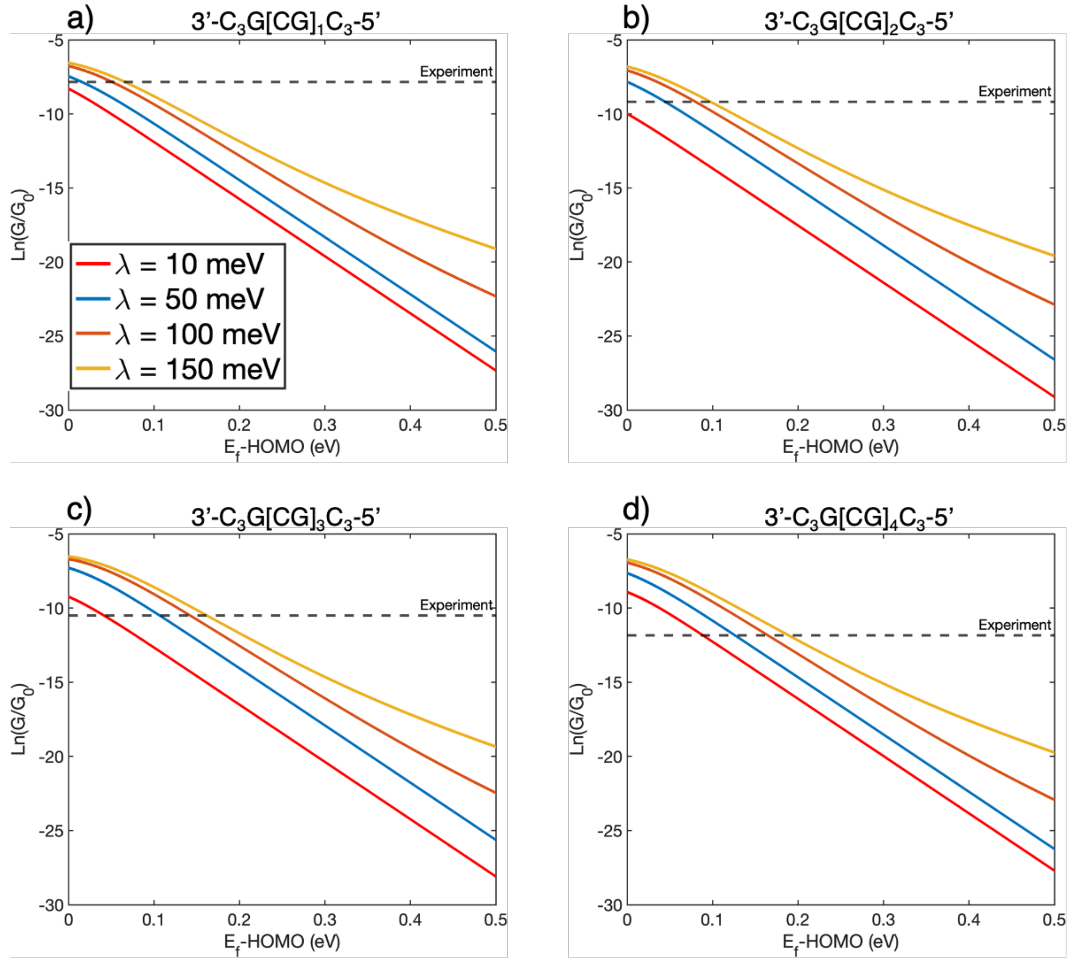


Figure 3.13 Conductance of the four strands (l=1-4) at different λ values. The dashed line corresponds to the targeted experimental conductance value. The trend shows λ impact on the conductance decay rate as we enter the HOMO-LUMO gap.

Our calculations further reveal that the modeled value of conductance matches experiments at higher values for the Fermi energies as λ increases (Figure 3.14b). This trend can be explained by Figure 3.13, which shows the tail of the conductance increasing in amplitude with the increase in λ in the HOMO-LUMO gap. We have shown through equation (2.27) Figure 3.2 that λ dictates the exponential decay into the bandgap. Therefore, the conductance at the edge of the bandgap increases with increasing λ.

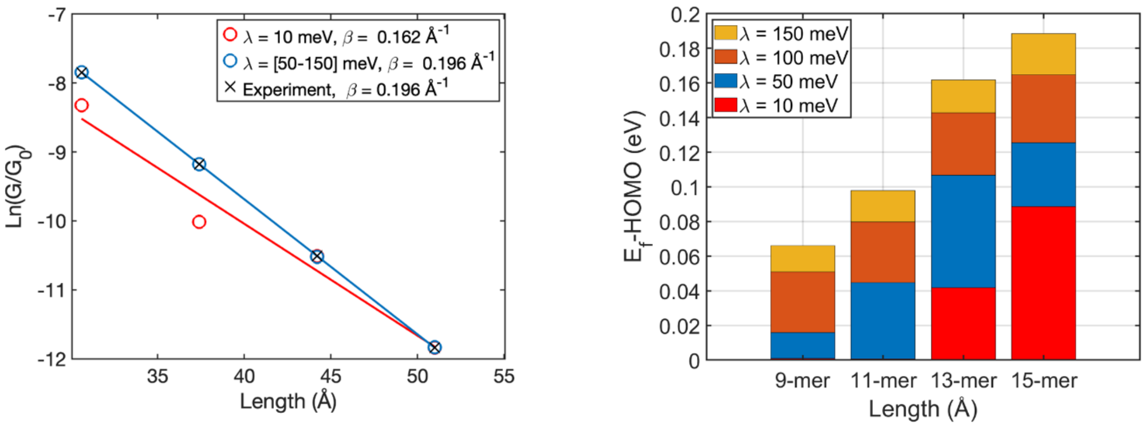


Figure 3.14 Conductance as a function of strand length. **a** Conductance trend at $\Gamma_B = 100$ meV, $\lambda = [10,50,100,150]$ meV compared with experiment. The lines correspond to the curve fitting of the data points. The blue curve represents the results for $\lambda > 10$ meV. **b** E_f of the extracted conductance values with respect to HOMO of the strand.

The second trend we observe from Figure 3.14b is that as the strand length increases, the Fermi energy at which the modeled conductance matches experiments moves farther away from the HOMO (for a fixed value of λ). To understand this, next, we discuss the effect of the contact-molecule junction on the expected Fermi energy location and how it is sensitive to the DNA length.

For a contact-DNA-contact system, partial charge transfer (δn) occurs from the contacts to the DNA due to the energy level broadening caused by the contacts and the misalignment between the Fermi energy of the contact (E_f) and the DNA Fermi energy ($\mathrm{E}_{f_{DNA_0}}$)[143]. In turn, the molecular orbitals of the DNA shift to higher energies (the HOMO gets closer to the E_f of the contacts). The relationship between the amount of shift in the molecular orbitals and E_f due to partial charge transfer has been described in [143], [144] as

$$E_f = \mathrm{E}_{f_{DNA_0}} + \frac{\delta \mathrm{n}}{\mathrm{n}'} + \mathrm{U}\delta \mathrm{n} \tag{3.1}$$

where $\mathrm{E}_{f_{DNA_0}}$ is the Fermi energy of the DNA before forming the junction with the contacts, δn is the amount of partial charge transferred, and n′ is the derivative of the electron number with respect to Fermi energy (dn/dE_f), where $n = \int DOS(E)f\left(E - E_f\right)dE$. U is the charging energy, which is a material property that describes the change in the potential of the molecule per one added electron. In equation (3.1), the second term describes the impact of the density of states in the DNA on the Fermi energy shift (or the molecular orbitals shift). The last term describes the electrostatic interaction due to the charge transfer.

To find the exact Fermi energy of the molecule after the contact-molecule-contact junction is formed, we require a self-consistent calculation that accounts for the contact details (atoms geometry and orientation), integral of DOS, and Poisson's equation. Although previous work shows the utilization of such a method on small molecules, the results yielded a wide range of Fermi energy locations extending from HOMO to LUMO [143]. This outcome is due to the sensitivity of the calculations to the DOS variation in the bandgap, the contacts geometry, and the choice of DFT exchange-correlation functionals. The difficulties faced in precisely calculating the Fermi energy of the molecule have shifted our focus into quantifying another critical parameter, the rate of the shift in energy per partial charge transfer.

We can utilize the second term in equation (3.1) to estimate the amount of molecular orbitals shift per strand. To achieve this, we first extract the DOS from our calculations using equation (2.30). Next, we integrate the DOS in the HOMO-LUMO gap as a function of Fermi energy,

$$n(E_i) = \int \int DOS(m, E)\, f(E - E_i)dmdE \tag{3.2}$$

where m is the nucleotide number, and $f(E - E_i)$ is the probability of electron occupancy with respect to energy when $E_f = E_i$. Taking the inverse of the derivative $[dn(E_i)/dE_i]^{-1}$ will yield the rate of Fermi energy change as a function of electron number dE_i/dn ($1/\mathrm{n}'$ in equation (3.1)). Therefore, we can estimate the change in the DNA Fermi energy -which is a direct indicator to the amount of shift in molecular orbitals- as a function of electron number in the DNA.

This rate gives us another viewpoint of the expected trend in $E_f -$ HOMO. The higher the dE_i/dn value, the higher the shift in the molecular orbitals, and the closer E_f becomes to the HOMO of the DNA. The rate dE_i/dn in the HOMO-LUMO gap is shown in Figure 3.15 and we listed the average values in Table 3-1. The results show that the expected molecular orbitals shift decreases with respect to the strand length. This means that the longer the strand, the farther E_f is expected to be from HOMO, which is consistent with the obtained trend seen in Figure 3.14b.

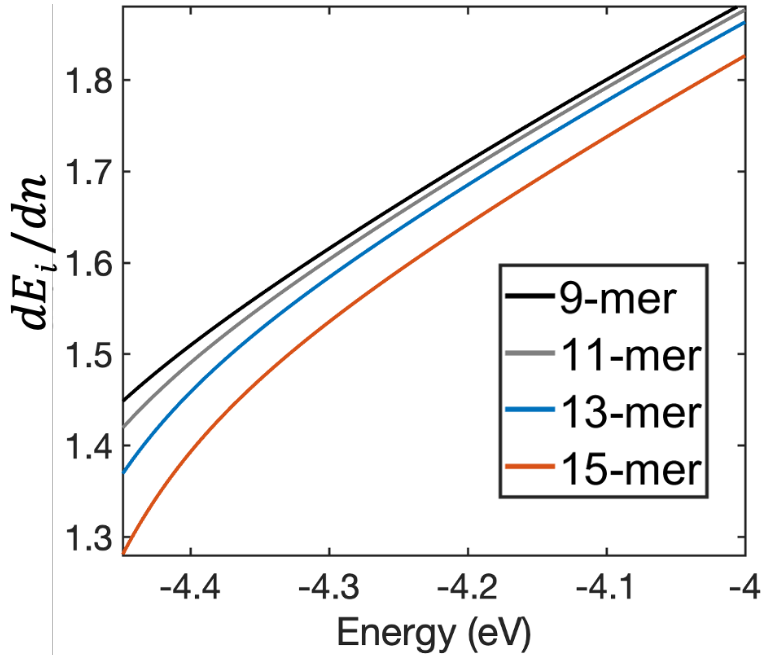


Figure 3.15 The resulting dE_i/dn extracted from the HOMO-LUMO gap window for each strand, calculated at λ = 100 meV.

We also note that the trend is maintained for different λ values. Table 3-1 shows a comparison between λ = 50 and 100 meV. We notice that dE_i/dn increases with decreasing λ. This can be understood as follows: as λ decreases, the energy levels broadening decreases, and the DOS decreases in the bandgap. As a result, $n(E_i)$ decreases in the bandgap as well (n is the integral of DOS) and its rate of change in energy (dn/dE_i) decreases. The reciprocal (dE_i/dn), however, increases.

Table 3-1 Comparison between the average rate of change in energy with respect to electron number (dE_i/dN) for 3'-C$_3$G[CG]$_l$C$_3$-5'(l = 1-4), λ = 50 and 100 meV

Strand	**dE_i/dn (eV/electron) λ = 50 meV**	**dE_i/dn (eV/electron) λ = 100 meV**
(N=1) 9-mer	1.9773	1.9750
(N=2) 11-mer	1.9697	1.9658
(N=3) 13-mer	1.9570	1.9503
(N=4) 15-mer	1.9211	1.9109

3.5 SUMMARY

In this chapter, we use first-principal calculations coupled with Green's function for charge transport calculations. One prominent model includes decoherence through fictitious probes that depend on spatially variant scattering rates. However, the built-in energy-independence of the decoherence (*E-indep*) model overestimates the transmission in the bandgap and washes out distinct and interesting features inside the valence or conduction bands. We modify the model to have the decoherence rate energy-dependent (*E-dep*). This model decoherence rate is maximum at the system's energy levels and decays from these energies. We establish that for self-consistent calculations, it is critical to include both the real and the imaginary parts of the decoherence self-energy to yield the correct integration of the density of states. Our results show that the *E-dep* model allows for exponential transmission decay with the DNA length and maintains features within the bands' transmission spectra. We further demonstrate that we can obtain DNA conductance values within the experimental range without rigorously curve fitting the decoherence parameters. Of course, the ultimate goal for us is to derive a direct relation between the exponential decay term and the amount of broadening or decoherence in the system. However, this requires further work in both experiment and theory. From the given information in the literature about DOS broadening in molecular systems and decoherence rates reaching a few hundred meV, the *E-dep* model is essential to apply such decoherence or broadening rates without yielding unphysical results or losing important transmission features. The new model can help study and design nanoelectronics devices that utilize weakly-coupled molecular structures such as DNA.

Chapter 4. LIMITATIONS OF NEGF ON THE DNA SYSTEM

Non-Equilibrium Green's Function (NEGF) formalism is a powerful tool that has been applied to study charge transport through nanoscale devices [116], [145]. It can incorporate scattering effects such as electron-phonon interactions. Devices such as carbon nanotube, nanoscale MOSFTs, and nanowires are all applicable to the method. In recent years, DNA has emerged as a strong candidate to form nanomolecular devices due to its commercially available precise structure engineering. To better understand its electronic properties, several methods have been developed to study charge transport through DNA structures. Such methods look at a metal-molecule-metal system and try to incorporate the fluctuating nature of DNA and its effect on charge transport. In this chapter, we investigate the use of the recursive NEGF method [116], [145] to study the effect of phonons on charge transport through short double-strand nucleic acid structures (ds-DNA). A fundamental understanding of the NEGF method is required to expand its usage to molecular devices where atomic fluctuations (phonons) may have a dominant effect on charge transport. This chapter is organized as follows: First, we present the governing equations and flowchart of the method. Then we test the NEGF code on a 1D conductor (wire) as a test case for validation. Next, we apply the NEGF method on a 1D strongly coupled semiconductor. Finally, we test it on a 1D model Hamiltonian representing the DNA, which is a weakly coupled system (10-100 meV hopping parameter between the bases).

4.1 METHODOLOGY

The NEGF formalism is described in detail in [116], [145], here, an overview of the governing equations will be presented. The electronic Hamiltonian of a layered structure is represented as follows:

$$H = \begin{bmatrix} H_1 & T_{1,2} & & & & \\ T_{1,2}^{\dagger} & H_2 & T_{2,3} & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & T_{N-2,N-1} & \\ & & & T_{N-2,N-1}^{\dagger} & H_{N-1} & T_{N-1,N} \\ & & & & T_{N-1,N}^{\dagger} & H_N \end{bmatrix} \tag{4.1}$$

where the diagonal blocks H_q represent the Hamiltonian of layer q, and off-diagonal blocks $T_{q,q+1}$ represent interaction between layers $q$and $q+1$. The retarded Green's function equation for the system with electron-phonon scattering is defined as

$$[EI - H - {\Sigma^r}_L - {\Sigma^r}_R - {\Sigma^r}_{ph}]G^r = I \tag{4.2}$$

where E is the energy, H is the Hamiltonian, and $\Sigma^r_{L(R)}$is the left (right) contact self-energy which couples the left (right) contact leads to the device and determines the rate at which the charge enters(leaves) the system. The electron-phonon scattering self-energy is Σ^r_{ph}, and G^r is the retarded Green's function. When bias is applied, the system will be away from equilibrium, and the Green's function equation is split into in-scattering and out-scattering Green's functions

$$[EI - H - \Sigma^r_L - \Sigma^r_R - \Sigma^r_{ph}]G^< = \Sigma^< G^a \tag{4.3}$$

$$[EI - H - \Sigma^r_L - \Sigma^r_R - \Sigma^r_{ph}]G^> = \Sigma^> G^a \tag{4.4}$$

where the $\Sigma^{<(>)}$ self-energy matrices represent the electrons (holes) in-scattering (out-scattering). Also, the advanced Green's function is defined as $G^a = [G^r]^{\dagger}$. The self-energy $\Sigma^{<(>)}$ is composed of both contact leads and electron-phonon interactions

$$\Sigma^{<(>)} = \Sigma_L^{<(>)} + \Sigma_R^{<(>)} + \Sigma_{ph}^{<(>)} \tag{4.5}$$

$$\Sigma_{L(R)}^{<} = -i2Im[\Sigma_{L(R)}^{r}]f_{L(R)}(E) = i\Gamma_{L(R)}(E)f_{L(R)}(E) \tag{4.6}$$

$$\Sigma_{L(R)}^{>} = i2Im[\Sigma_{L(R)}^{r}](1 - f_{L(R)}(E)) = -i\Gamma_{L(R)}(E)(1 - f_{L(R)}(E)) \tag{4.7}$$

where $f_{L(R)}(E)$ is the Fermi-Dirac distribution function of the left (right) contacts, and $1 - f_{L(R)}(E)$ is the probability of finding an unoccupied state in the left (right) contact. Unless mentioned otherwise, $\Gamma_{L(R)}$ represents energy broadening caused by lead-device interaction and is treated here within the wide band limit (WBL). In the wide band limit, $\Gamma_{L(R)}$ is taken to be energy independent and the real parts of the contact self-energy matrices are neglected, thus $\Sigma^{r}{}_{L(R)}$ are defined as

$$\Sigma_{L(R)}^{r} = -\frac{i\Gamma_{L(R)}}{2} \tag{4.8}$$

As for the phonon self-energy Σ_{ph}^{α}, where $(\alpha = r, <, or\ >)$, it is defined as

$$\Sigma_{ph}^{\alpha}(E) = \Sigma_{el}^{\alpha}(E) + \Sigma_{inel}^{\alpha}(E) \tag{4.9}$$

$$\Sigma_{el}^{\alpha}(E) = D_a G_{q,q}^{\alpha}(E) \tag{4.10}$$

$$\Sigma_{inel}^{r}(E) = \frac{1}{2}[\Sigma_{inel}^{>}(E) - \Sigma_{inel}^{<}(E)] \tag{4.11}$$

$$\Sigma_{inel}^{<}(E) = D_o[n_B(E_{ph})G_{q,q}^{<}(E - E_{ph}) + (n_B(E_{ph}) + 1)G_{q,q}^{<}(E + E_{ph})] \tag{4.12}$$

$$\Sigma_{inel}^{>}(E) = D_o[n_B(E_{ph})G_{q,q}^{>}(E + E_{ph}) + (n_B(E_{ph}) + 1)G_{q,q}^{>}(E - E_{ph})] \qquad (4.13)$$

where $\Sigma_{el(inel)}^{\alpha}$ is the elastic (inelastic) electron-phonon scattering self-energy. D_a and D_o are the acoustic and optical deformation potentials, respectively, and they are assumed to be energy independent parameters. $G_{q,q}^{\alpha}$ is the diagonal element of the Green's function at layer q. The probability of finding phonons at temperature T and at energy $E_{ph} = \hbar\,\omega_{ph}$ is assumed to be the equilibrium Bose-Einstein factor $n_B = 1/exp(E_{ph}/k_BT) - 1$. The inelastic in-scattering in (4.12) has $G_{q,q}^{<}(E - E_{ph})$ and $G_{q,q}^{<}(E + E_{ph})$ which are the density of occupied electronic states at $E - E_{ph}$ and $E + E_{ph}$, respectively. The first term of (4.12) represents in-scattering of electrons from $E - E_{ph}$ to E (phonon absorption), and the second term is the in-scattering from $E + E_{ph}$ to E (phonon emission). The opposite can be said about (4.13), as $G_{q,q}^{>}(E - E_{ph})$ and $G_{q,q}^{>}(E + E_{ph})$ are the density of unoccupied electronic states at $E - E_{ph}$ and $E + E_{ph}$, respectively. Here, the first term in (4.13) represents the out-scattering of an electron from E to $E + E_{ph}$ (phonon absorption), and the second term is the out-scattering of an electron from E to $E - E_{ph}$ (phonon emission).

The recursive algorithm for solving the NEGF equations and obtaining G^r,$G^<$, and $G^>$ is described in [116]. The algorithm initially calculates the left-connected Green's function and then uses it to obtain G^{α}, where ($\alpha = r, <, or\ >$). The flowchart seen in Figure 4.1 explains the method of including the phonon self-energy in the calculations. The first iteration calculates only the ballistic transmission, i.e. no electron-phonon scattering. Once the ballistic G^{α} is calculated, it is then used as input for the phonon loops. The current density flowing from layer q to $q + 1$ can then be calculated by

$$J_{q,q+1}(E) = (2e/\hbar)\, Tr[T_{q,q+1} G^{<(>)}_{q+1,q}(E) - T_{q+1,q} G^{<(>)}_{q,q+1}(E)] \tag{4.14}$$

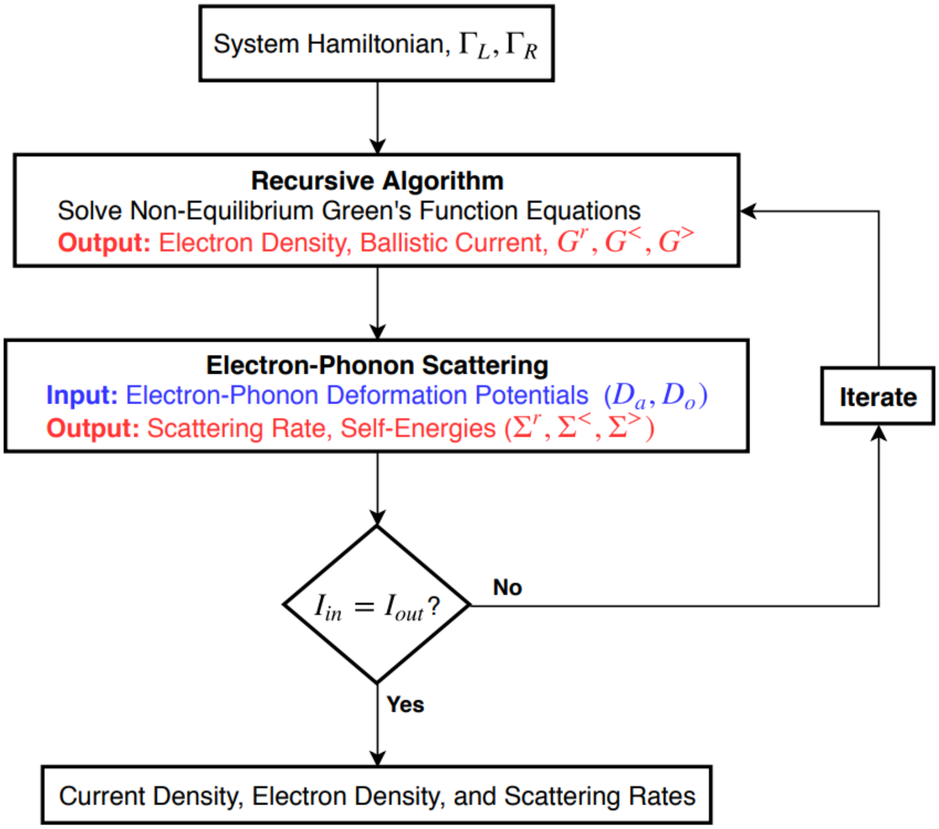


Figure 4.1 Flowchart of the recursive NEGF code.

4.2 TESTING 1D CONDUCTOR WIRE

First, we test the decoherence effect of electron-phonon coupling by evaluating if Ohm's law is obtained (The resistance is constant under increasing voltage, yielding the relationship $V = IR$, where I is the current). Furthermore, the resistance of the wire is proportional to its length. A simple 1D conductor is used as a test case. The structure is a single atom unit cell repeated N_{atom} times, connected to semi-infinite leads at its left and right ends.

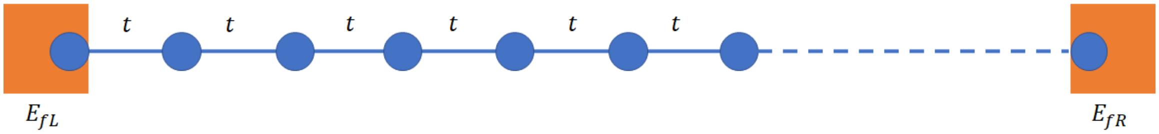


Figure 4.2 Schematic of the 1D conductor wire, orange blocks represent contact locations.

The system Hamiltonian seen in (4.1) is of size $N_{atom} \times N_{atom}$, composed of only off-diagonal elements, $T_{i,j} = t = 1\ eV$. Therefore, the Hamiltonian has zero diagonal elements. This represents a 1D conductor with an onsite potential of 0 eV. The transmission profile yields a band in the range of $-2t$ to $+2t$. Table 4-1 lists all the simulation parameters, where $E_{f_{L(R)}}$ is the left (right) contact Fermi energy, and $E_{f_R} = E_{f_L} - qV_{bias}$. The energy range is taken to be from the lower contact Fermi energy to the higher Fermi energy, $-/+10kT$, respectively. The acoustic and optical deformation potentials, D_a and D_o respectively, were set as arbitrary values since this an investigation of the overall NEGF code behavior. The phonon energy, E_{ph}, was also chosen to be within the biasing window. The number of phonon loops, N_{ph}, required to reach current convergence is dependent on the system size, but we find that 50 loops were sufficient for all cases. It is important to note that for this section only, the leads are treated in terms of the exact surface Green's function g_L and g_R of the semi-finite 1D leads at a constant potential [145], instead of the WBL approximation. This is found to be necessary for testing 1D conductors to mimic Ohm's law. The rest of the sections use the WBL approximation to replicate metal-DNA-metal types of calculations.

Table 4-1 Simulation parameters for 1D conductor

Parameter (units)	Value
E_{f_L}(eV)	0
V_{bias} (mV)	10-500
kT (meV)	25.86
D_a (meV)	40
D_o (meV)	60
E_{ph} (meV)	20
Energy Range (eV)	$E_{f_R} - 10kT : E_{f_L} + 10kT$
N_{ph}	50
Number of Atoms	8, 40, 160, 320

To calculate the resistance, we apply a small bias and the derivative of the voltage with respect to the calculated current is taken to yield the total differential resistance, as follows

$$R_{tot} = dV/dI \tag{4.15}$$

$$R_{tot} = R_{ballistic} + R_{device} \tag{4.16}$$

where $R_{ballistic}$ is found from quantum conductance

$$R_{ballistic} = 1/G_0 = 1/(2e^2/h) = 12.941\ \text{k}\Omega \tag{4.17}$$

Therefore, $R_{ballistic}$ is subtracted from R_{tot} to get the resistance of the device. The average R_{device} value of each length over the biasing window is listed in Table 4-2. The ratio between the different length cases is listed in Table 4-3 in comparison with Ohm's law. It is noticeable that for the current bias window of 100-500 mV, as the ratio value increases so does its deviation from the theoretical estimate. This is later demonstrated to be an artifact of the inelastic scattering, D_o. The next step was to apply lower biases, as high biasing seems to cause higher fluctuation in the resistance curve. When calculating the resistance for smaller biases of 10-50 mV range, the ratio almost reaches the theoretical value.

Table 4-2 Average device resistance values

Device Length (Atoms)	$\bar{R}_{device}(k\Omega) = \bar{R}_{tot} - R_{ballistic}$
8	3.0778
40	13.4213
160	49.5498
320	92.8952

Table 4-3 Average device resistance ratios for different bias windows

$R_{device1}/R_{device2}$ **(theoretical value)**	**R_{320}/R_{160} (2)**	**R_{160}/R_{40} (4)**	**R_{40}/R_8 (5)**	**R_{320}/R_{40} (8)**	**R_{160}/R_8 (20)**	**R_{320}/R_8 (40)**
Calculated Ratio (V_{bias} = 100-500 mV)	1.87	3.69	4.36	6.92	16.10	30.18
Calculated Ratio (V_{bias} = 10-50 mV)	1.98	3.92	4.95	7.76	19.38	38.39

A closer look into the biasing effect on the calculations shows that as the bias increases, fluctuations increase in the resistance plot (Figure 4.13). Moreover, the fluctuations occur with different trends per wire length. These different behaviors of R_{device} as a function of V_{bias} for each wire yield different averages, hence, lowering the calculated ratio with respect to the theoretical values.

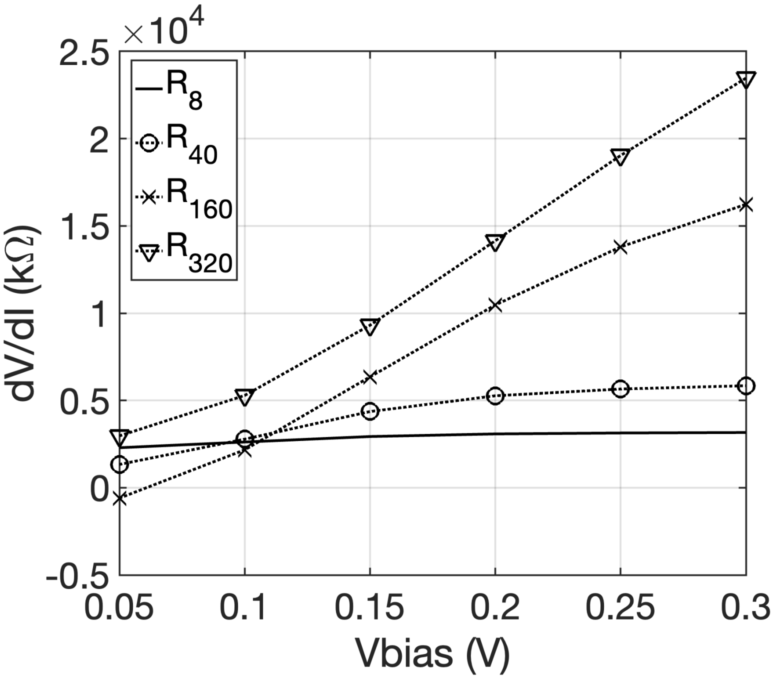


Figure 4.3 Different R_{device} behavior per wire length, resistance values were shifted for better comparison.

The biasing window appears to take effect due to optical phonons (D_o is in effect), as when simulating the system with D_o=0, it would yield the exact ratio irrespective of V_{bias} value (see Figure 4.4 left). And once D_o is nonzero, the ratio starts deviating as V_{bias} increases. Increasing phonon loops does not remove the deviation and that proves that current convergence is not the issue here.

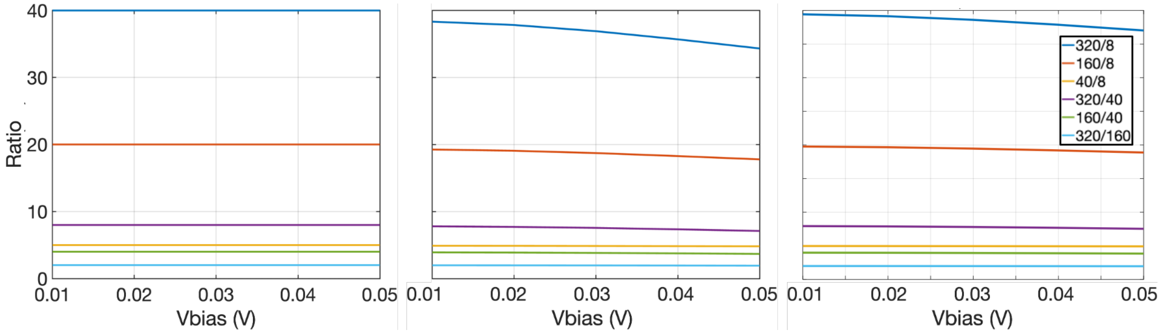


Figure 4.4 D_o and $D_a$effects on calculated ratios of the resistances. Left: D_o is zero and D_a is nonzero, constant ratio values over the whole biasing window. Middle: $D_o$is nonzero and D_a is zero, deviation from theoretical ratio values is apparent and continuous decay as V_{bias} increases. Right: D_o and D_a are both nonzero, deviation in ratio values still exists but at lower rate than the middle case. The highest deviations appear for the highest ratios.

The inelastic electron-phonon scattering is represented in D_o, which affects both electron absorption and emission of phonons. Looking further into this, one main factor of this deviation is the value of E_{ph} within the biasing window. Therefore, in the following test I keep the simulation parameters fixed, and only change E_{ph}= [160 meV, 50 meV, or 10 meV] and the results are plotted in Figure 4.5. At E_{ph}=10 meV, it is easier for electron-phonon interactions to occur, i.e. to emit/absorb phonons, within the biasing window. That is due to the increase in available energy states for electrons to transfer into, which are multiples of E_{ph}. Thus, the R_{device} value would resemble a textbook Ohmic conductor. With E_{ph} increasing, it becomes more difficult for electrons to emit/absorb phonons within the same energy window, yielding unphysically low R_{device} values, hence higher deviation. The fluctuations are more apparent when comparing large ratios.

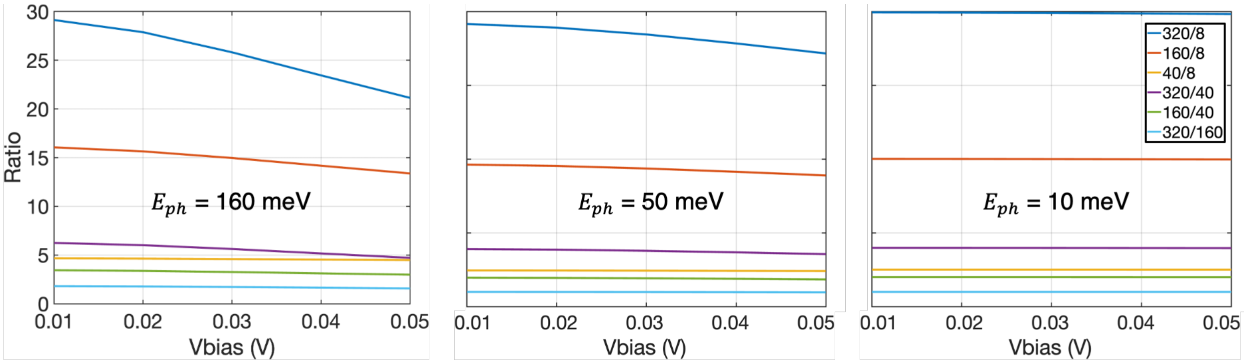


Figure 4.5 E_{ph} effect on resistance values. As E_{ph} increases under the same bias window, the inelastic scattering is insufficient to yield consistent results with Ohm's law.

4.3 TESTING 1D SEMICONDUCTOR

Moving on from the 1D conductor, the next step was to test the NEGF code on a 1D semiconductor. The system consists of a 2-atom unit cell with two coupling values t_1 and t_2, t_1 is intra-coupling and t_2 is inter-coupling, with $t_1 > t_2$. Based on the value of t_2, one may create a semiconductor, or a weakly-coupled system.

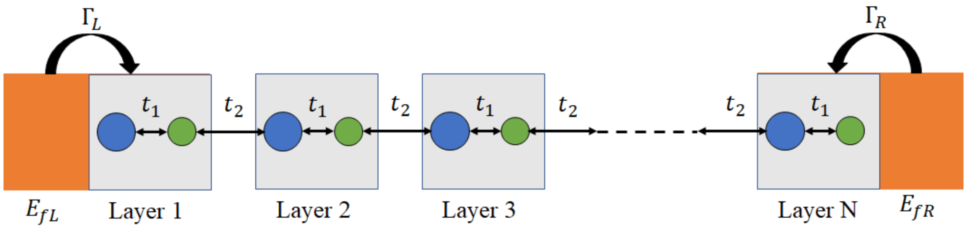


Figure 4.6 Schematic of the 1D semiconductor, orange blocks represent contact locations.

The coupling is defined as t_1=1.5 eV (strong coupling), and t_2= 0.9 eV (moderate coupling), the onsite energy is kept at 0 eV, this results in the diagonal and off-diagonal blocks to be

$$H_q = \begin{bmatrix} 0 & 1.5 \\ 1.5 & 0 \end{bmatrix}, T_{q,q+1} = \begin{bmatrix} 0 & 0 \\ 0.9 & 0 \end{bmatrix} \tag{4.18}$$

Length of the device is 50 unit cells (100 atoms), the contacts are located at left and right ends, respectively. Contact effects are now included through the WBL approximation to better mimic

DNA-type calculations. This system will create two minibands for electron transport. The energy window was chosen at the first miniband (Figure 4.7), between -1.4 eV to -1.5 eV, E_{f_L}= -1.4 eV, $E_{f_R} = E_{f_L} - qV_{bias}$. Simulation parameters are included in Table 4-4, note that D_a is zero in this test because as seen from the previous section, optical phonons have the major effect on NEGF behavior, and thus will be the focus of this investigation.

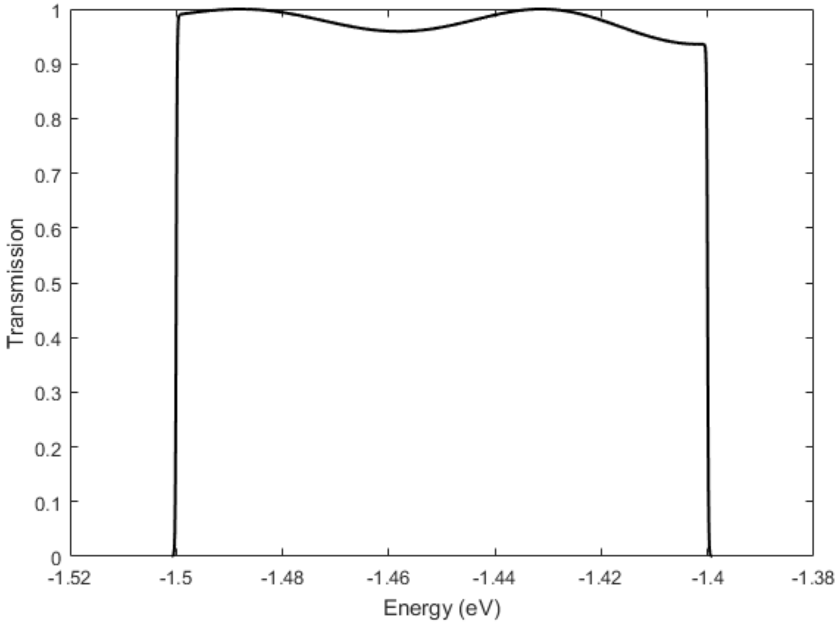


Figure 4.7 First miniband chosen as the bias window, the fluctuating transmission is due to using the wide band limit.

Table 4-4 Simulation parameters for 1D semiconductor

Parameter (units)	Value
E_{f_L}(eV)	-1.4
V_{bias} (mV)	100
kT (meV)	25.86
D_a (meV)	0
D_o (meV)	100
E_{ph} (meV)	20
Energy Range (eV)	E_{f_R}- 10*kT: E_{f_L}+ 10*kT
N_{ph}	50
Number of Atoms	100

The initial calculations converge after N_{ph} = 50 loops, and the current density profile plotted in Figure 4.8 (right) shows the current flowing from layer 1 to layer 2, J_{12}, as well as the current flowing through the right-end layer, $J_{N-1,N}$. When comparing between J_{12} and $J_{N-1,N}$, the inelastic scattering causes the energy-resolved current to shift its peak. However, the interesting part is seen in the behavior of J_{12} curve, as the current at energy close to E_{f_R}= -1.5 eV is almost zero. Since the left contact Fermi energy E_{f_L}= -1.4 eV, one would expect the electron current flowing from layer 1 -which is connected to left contact- to have high values at all energy ranges between $E_{f_L}$and E_{f_R}. However, the current starts decaying once the energy shifts away from E_{f_L}value.

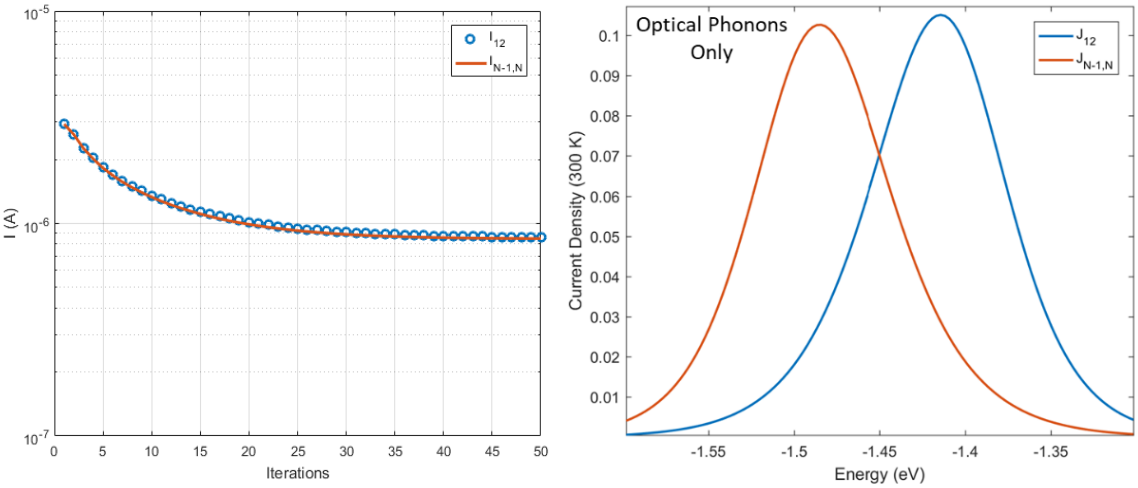


Figure 4.8 **Left** Current converges after 50 iterations. **Right** Current density plot flowing through leftmost layer (inject) and rightmost layer (extract) at room temperature, optical phonons only.

To better understand the results, we calculate the current at zero temperature. At very low temperature (temperature ~ 0K), the resulting current is equivalent to transmission which is a fundamental property of any system. The calculation converges, and the current density plot shows two features (see Figure 4.9):

1) Electrons enter and exit the device at different energies. This is due to energy dissipation from inelastic scattering.
2) Looking at electron transmission from layer 1 to layer 2, J_{12}, the electrons do not travel between the two layers at energy < -1.44 eV, and the opposite picture is seen for $J_{N-1,N}$, which is at the rightmost end layer, where electrons only exit the system at energy < -1.46 eV.

The second observation implies that electrons entering the system from layer 1 at E_{f_L}= -1.4 eV tend to scatter and get reflected in the opposite direction, losing energy in the process via spontaneous phonon emission for 0K temp (or in general, electron-phonon interaction for room temp). The reflected electrons now occupy lower states, preventing further injection of electrons from the left contact reservoir at that energy level. Hence, with Pauli's exclusion principle, J_{12} at energies lower than -1.44 eV almost depletes.

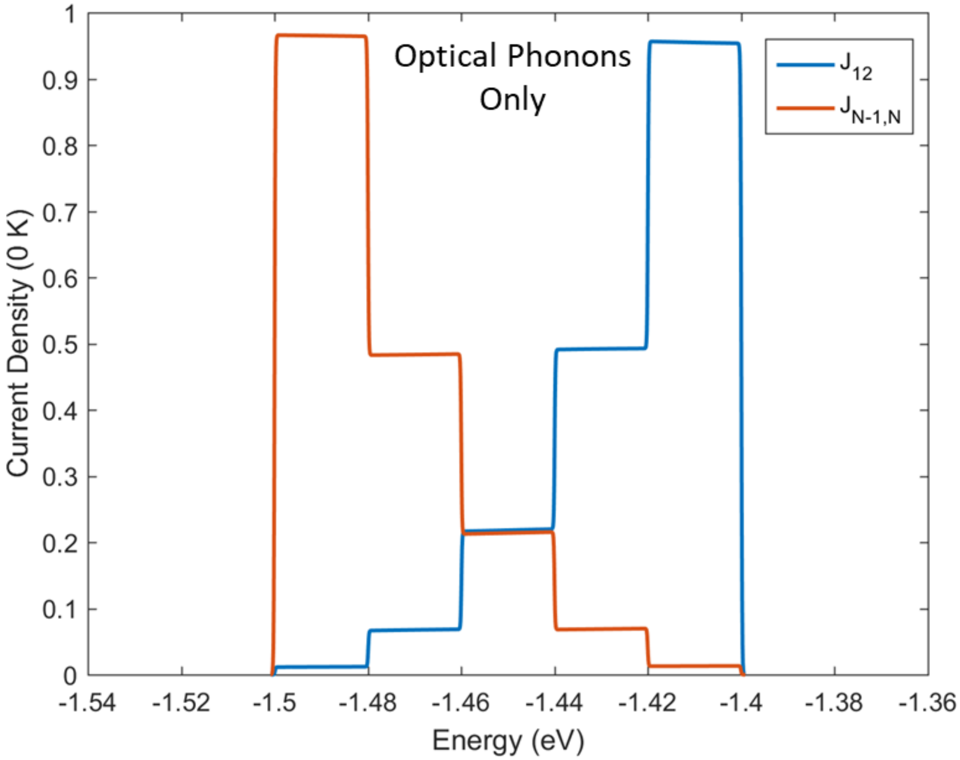


Figure 4.9 Current density plot for the 50 layers semiconductor. The ladder-effect represents the electron-phonon emission as the steps are equal to E_{ph} (only optical phonons).

The following subsections investigate different parameters such as acoustic (elastic) phonons, and phonon energy effects on the current profile.

4.3.1 *Acoustic Phonons*

The initial transport calculations only included optical phonons. By setting $\mathrm{D_a}$ to be nonzero, the acoustic phonons effect can be studied as well. Looking at Figure 4.10, it further reduces the current by ~50% even at the first E_{ph} window (-1.42 eV < energy < -1.4 eV). This is because acoustic phonon corresponds to elastic scattering, which now allows for electrons to get scattered at the same energy level, further limiting electron entrance from left contact at E_{f_L} = -1.4 eV.

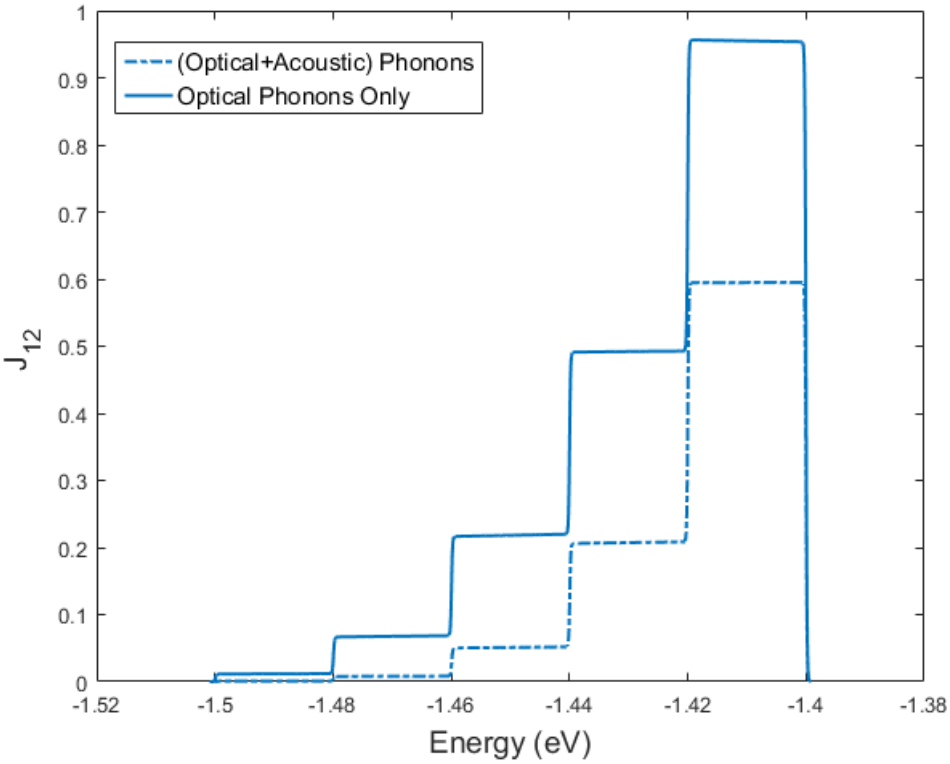


Figure 4.10 Current density J_{12} for the 50-layer device, for both systems of (acoustic + optical) phonons and optical phonons only.

4.3.2 *Phonon Energy*

The decrease in current with respect to energy is shown to be dependent on E_{ph} steps. Therefore, at lower $E_{ph} = 10$ meV, the decay becomes more spread out with respect to the energy window when compared to $E_{ph} =$ 20 meV (see Figure 4.11). Reflected electrons have more states to scatter into and move in the opposite direction.

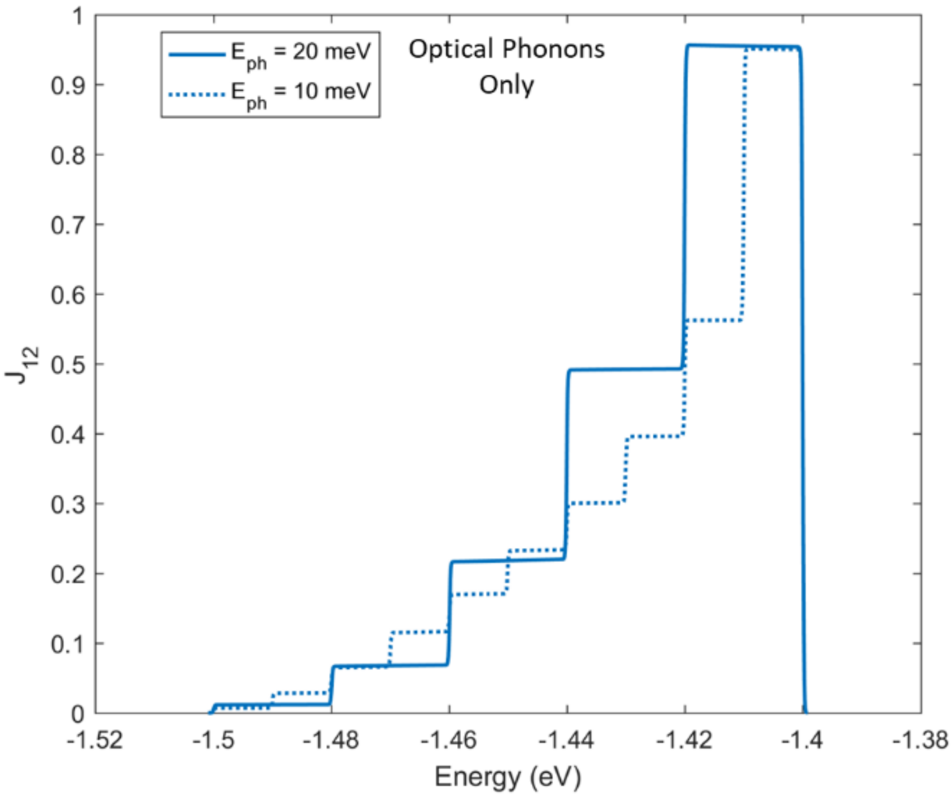


Figure 4.11 Change in phonon energy yields different ladder-like current decay rate.

4.4 TESTING 1D WEAKLY-COUPLED SYSTEM

After testing the NEGF code on an ohmic conductor, where the results follow Ohm's law, and on a semiconductor to test out the effects of acoustic and optical phonons. Here, we used a model Hamiltonian involving a 1D weakly-coupled system to represent a DNA structure, and the results of inelastic scattering are discussed in this section. It is important to note that we should be able to handle transport calculations in this simple system before moving over to DNA.

A 15-layer device with a unit cell containing two atoms is used as shown in Figure 4.12. The coupling is defined as $t_1 = 1.5\,eV$ (strong intra-coupling), $t_2 = 0.1\,eV$ (weak inter-coupling). This system will mimic ds-DNA, where the intra-coupling (between a base-pair) is much higher than inter-coupling (neighboring bases) [78]. The onsite energies of either 250 meV or 700 meV were randomly assigned in the Hamiltonian, this results in the diagonal and off-diagonal blocks to be

$$H_q = \begin{bmatrix} 0.7(0.25) & 1.5 \\ 1.5 & 0.7(0.25) \end{bmatrix}, T_{q,q+1} = \begin{bmatrix} 0 & 0 \\ 0.1 & 0 \end{bmatrix} \tag{4.19}$$

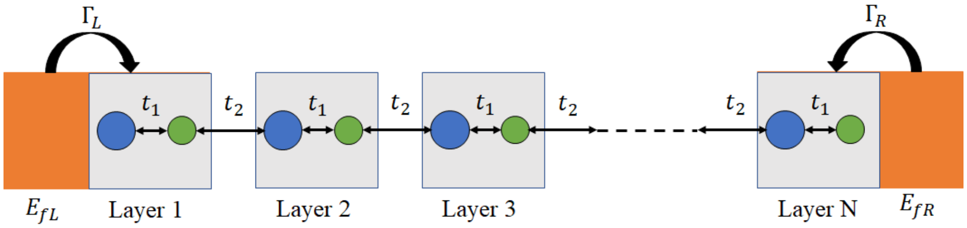


Figure 4.12 Schematic of the 1D semiconductor, orange blocks represent contact locations.

For a 15-layered system, the transmission minibands are formed around -1 eV or 2 eV (Figure 4.13, left). The energy window was chosen to be the first miniband, which is between -1.05 eV and -0.5 eV. The right plot in Figure 4.13 shows three distinct characteristics: high transmission value (T> 10^{-5}), low value (T < 10^{-12}), and the zeros/negative values (T<10^{-18}). The zeros/negative values are shown as discontinuities in the plot and are due to finite machine precision.

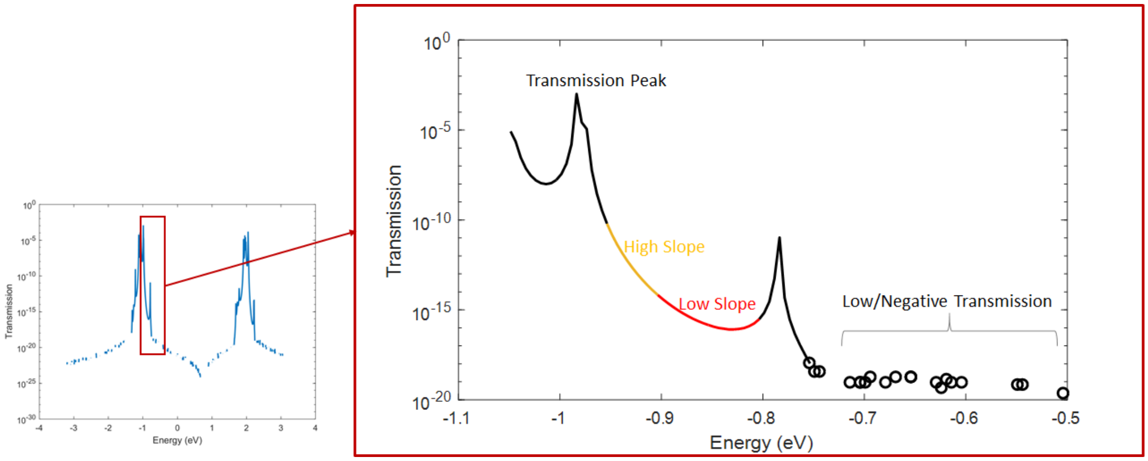


Figure 4.13 Ballistic transmission plots. left: minibands formed by the 15-layered system. right: the chosen energy window sweep, highlighted regions have different characteristic effects on current convergence as explained further in text.

A small bias of 50 meV was chosen for the calculations, with $E_{f_R} = E_{f_L} - eV_{bias}$, the left contact Fermi energy E_{f_L} was swept to go from peak to low transmission regions (we can think of this as applying a gate voltage). To study inelastic scattering, we apply a deformation potential corresponding to energetic phonons, D_o, with D_a set to zero. The other parameters used are listed in Table 4-5.

Table 4-5 Simulation parameters for 1D weakly-coupled system

Parameter (units)	Value
V_{bias} (mV)	50
kT (meV)	25.86
D_a (meV)	0
D_o (meV)	10
E_{ph} (meV)	10
N_{ph}	1000
Device Length	15 layers

The convergence criterion is to have $I_{N-1,N} = I_{12}$, where $I_{N-1,N}$ (I_{12}) is the exiting (entering) current. The results in Figure 4.14 show poor convergence for three energy windows. The first window, Figure 4.14 (a), corresponds to the current calculated at the transmission peak region shown in Figure 4.13. The second window, Figure 4.14 (b), corresponds to the current at the high slope region (yellow curve in Figure 4.13), and the third window is for the low slope region (red curve in Figure 4.13). In the first window, the calculations come close to convergence but only after 1000 phonon loops, which is unreasonably large for a 30-atom system. The divergence in the second and third windows is higher in Figure 4.14 (b) and (c). Furthermore, at the low/negative transmission region (bandgap region), the current yields negative unphysical values (plot not shown).

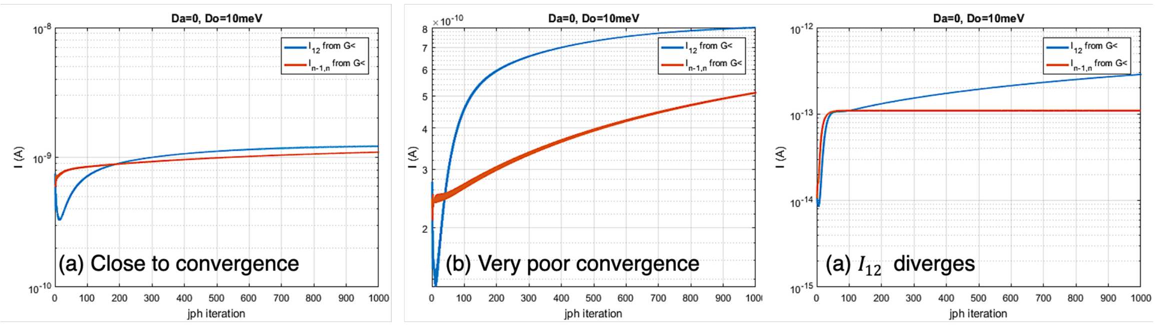


Figure 4.14 Divergence issues at different energy windows. (a) Energy window from -1.0 to -0.95 eV. (b) Energy window from -0.95 to -0.90 eV. (c) Energy window from -0.80 to -0.75 eV.

For the weakly coupled system, the calculations do not converge and the reason for this is that the decoherence rate due to scattering is not weak compared to the inter-base hopping.

After exploring the weakly-coupled system, it is evident that the weak inter-coupling between the unit cells, which is equivalent to neighboring bases in ds-DNA, is causing current divergence. The electron-phonon interaction cannot be treated by first-order approximation of perturbation theory as it is comparable in magnitude to the inter-base hopping term in the Hamiltonian. Therefore, the first-order approximation of the phonons self-energy used in NEGF does not hold due to weak electronic coupling between the DNA bases comprising the strand [78]. Hence, we conclude that the current NEGF method cannot be applied to the DNA system.

4.5 Conclusion

Weakly-Coupled systems have been studied by researchers in the field of organic crystals. Their treatment has taken the main approach of using the polaron hopping model [146], in which the deformation of the molecule caused by the charge transfer is accounted for. In this model, the time-dependent electronic Hamiltonian is used to capture electron-phonon interaction via polaron hopping to study charge transfer. This method is well developed for crystals [147], [148], in which Bloch's theorem is utilized with periodic boundary conditions, however, it cannot be applied directly for the aperiodic DNA strands. Surprisingly, other groups[85], [86], [119] were able to use NEGF on DNA to account for elastic acoustic phonons. However, they have used a simple tight-binding model for the DNA Hamiltonian with electronic couplings that are an order of magnitude higher than the electron-phonon coupling parameter. This assumption is unphysical in modeling realistic DNA system as shown by studies estimating the hopping parameter between bases [78], [122]. Therefore, in this thesis, the decoherence probe method is used to model charge transport in DNA structures.

Chapter 5. SINGLE NUCLEOTIDE MISMATCH DETECTION THROUGH DIRECT ELECTRICAL MEASUREMENTS

5.1 INTRODUCTION

Single nucleotide mismatch detection has been an ongoing challenge for researchers. The ability to detect minuscule changes within a nucleic acid strand in an efficient and rapid manner is of interest, especially for genetic mutation diagnosis. Chemistry intensive methods have been widely used for mutation and mismatch detection, but they usually require amplification via polymerase chain reaction (PCR), and labeling of target sequences prior to sequence processing [71]. Thus, it is of interest to develop new detection schemes. Here, we demonstrate a direct electrical measurement method for the detection of a single nucleotide mismatch through conductance measurements of a 15-base-pair strand [51], via a contact-molecule-contact scheme. This method would allow for a new perspective of mismatch detection using an electrical technique.

5.2 SYSTEM UNDER STUDY

The system under study is a DNA:RNA hybrid of 15 bps. Figure 5.1 shows the sequence, with the location of the three mismatches denoted as X, Y, and Z. In the case of no mismatch, the RNA strand is 5’-CGACCCUCUUGAAC-3’. For the first mismatch case, the X base location has the G substituted for an A. In the second mismatch, Y location has the C substituted for the T base. And the third case, Z location, has A substituted for G. Biologically, each mismatch represents a different strain found in the pathogenic species of E. Coli [51].

In a metal-molecule-metal structure, the DNA strand is terminated with a linker group, to enhance the coupling to the two measuring probes. A commonly used linker for gold-type contacts is the thiol group -CH2 - CH2 - CH2 – SH [149]. In this simulation, to reduce system complexity and computational costs, neither the gold contacts nor the linker groups are explicitly included in the structure, but the coupling of contact effect is included as a self-energy.

The method is composed of the following steps: ab initio DFT calculations to determine the Hamiltonian, and the Green's function (GF) based transport calculations. We perform the ab initio DFT calculations using Gaussian 16 with the B3LYP/6-31G(d,p) functional and basis set to obtain the Hamiltonian (see chapter Chapter 2 for further details on the methodology). For all the structures, the counterions are not included, and so the charge of the DNA is taken as -28. The number -28 corresponds to a negative charge of -1.6E-19 C on each of the 28 phosphate groups in the DNA structure. Note that there is one phosphate group per nucleotide, except at the backbone of the left and right contact sites. Water solvent is included in the calculations via the polarizable continuum model (PCM). This DFT calculations give us the system Hamiltonian, which will then be used in the transport calculations.

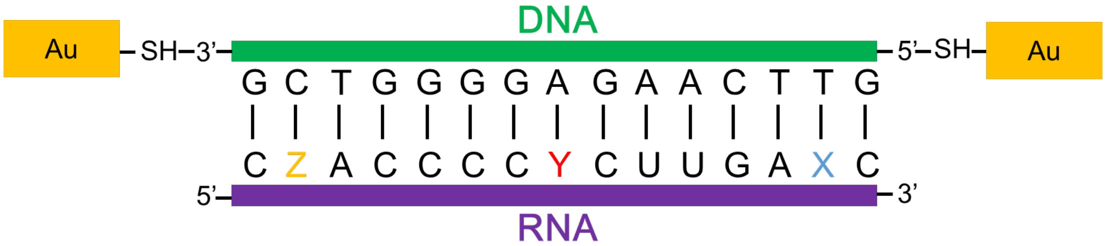


Figure 5.1 Schematic of the structure under study. The green side is the DNA, terminated with the thiol groups which are connected to gold contacts. The purple side represents the RNA sequence, without mismatches X=A, Y=U, and Z=G. The three mismatch cases are defined as X=G, Y=C, and Z=A, respectively.

Elastic (dephasing) transport is used to model charge transport through the 15-bp strands. The decoherence probes are attached to each nucleotide (backbone + base) as shown in Figure 5.2. In a structure consisting of $N = 30$ bases, the total number of decoherence probes is calculated as $N_b = N - 2 = 28$. We used the E-indep model with uniform coupling value of 10 meV for all the energy levels. The contact self-energy is considered as a fitting parameter, it usually ranges between 50 meV and 1 eV [150]. These values would represent going from weak to strong coupling, respectively. We tested 100 meV and 600 meV values, and the overall trend was the same, mainly showing an increase in conductance magnitude with an increase in contact coupling. Therefore, only the results of 600 meV are presented here.

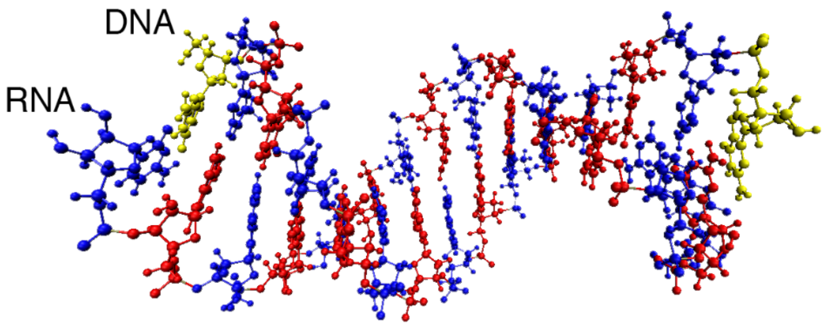


Figure 5.2 The blocking scheme for applying decoherence probes in transport calculations. The alternating colors (red and blue) represent the 28 blocks consisting of the backbone and base. The two yellow blocks correspond to the contact locations, both of which include the two guanines at the two ends of the DNA strand with their backbones.

5.3 RESULTS AND DISCUSSION

It is difficult to determine the exact Fermi level of the metal-molecule-metal system, due to its sensitivity to the contact details (the orientation and the geometry of the atoms that are in contact), as well the metal work function, thus shifting in between the HOMO-LUMO gap [150]. Therefore, the Fermi level is treated as a fitting parameter, where the transmission plots are focused between -5.5 and -4.5 eV region, which is the region close to the HOMO level of the different structures.

We first look at the coherent transmission results in Figure 5.3a. The coherent transport does not show a clear trend. At energies lower than -4.95 eV, X-mismatch is dominant, followed by Z-mismatch, original structure, and then the Y-mismatch. However, when the energy is between -4.95 eV and -4.7 eV, the dominance is not conclusive as within each region a different structure has the highest transmission probability. The conductance plot further illustrates these observations (Figure 5.3c).

Looking at the decoherent transmission case (Figure 5.3b), the first thing to notice is that the amplitude is multiple orders of magnitude higher than the coherent case. Decoherence due to the phase-losing environmental effects tends to broaden the energy levels of the system [118] that are primarily located at different regions of the DNA, thus, increasing the transmission probability. At energy below -4.95 eV, the plot shows X-mismatch to be dominant with transmission probability of 10^{-4}, followed by Z-mismatch, the original structure, and ending with Y- mismatch as the lowest transmission of less than 10^{-5}. However, at energies closer to the HOMO levels, which is the region of energy above -4.95 eV, the Y-mismatch begins to show dominance, and it specifically maintains its dominance at -4.8 eV and onwards. Since molecular structures tend to be charge neutral, it is reasonable to assume that the Fermi level lies above the HOMO level of the structures. Within this energy range from -4.8 eV to -4.5 eV, the conductance plot shows the trend of Y-mismatch >Original structure>Z-mismatch >X-mismatch (Figure 5.3d).

An interesting result shows up when comparing the conductance plots in Figure 5.3 c and d. The coherent conductance shows X-mismatch dominating along the whole energy range. This is due to the large regions of transmission peaks within the HOMO region (at energies lower than -4.95 eV), which are several orders of magnitude higher than the other structures. Moreover, the transmission in the HOMO-LUMO gap (above -4.95 eV) for all cases is ten orders of magnitude lower than their HOMO region. Thus, the transmission band in the HOMO region for X-mismatch makes the conductance -which is the integral of transmission with respect to energy (see equation (2.37)) dominate when E_f > -5.3 eV. As for the decoherent case, the plot of conductance shows a similar trend to its transmission counterpart, which is due to the transmission amplitude not changing drastically when going from HOMO into the HOMO-LUMO gap, thus having a moderate contribution in the conductance calculation.

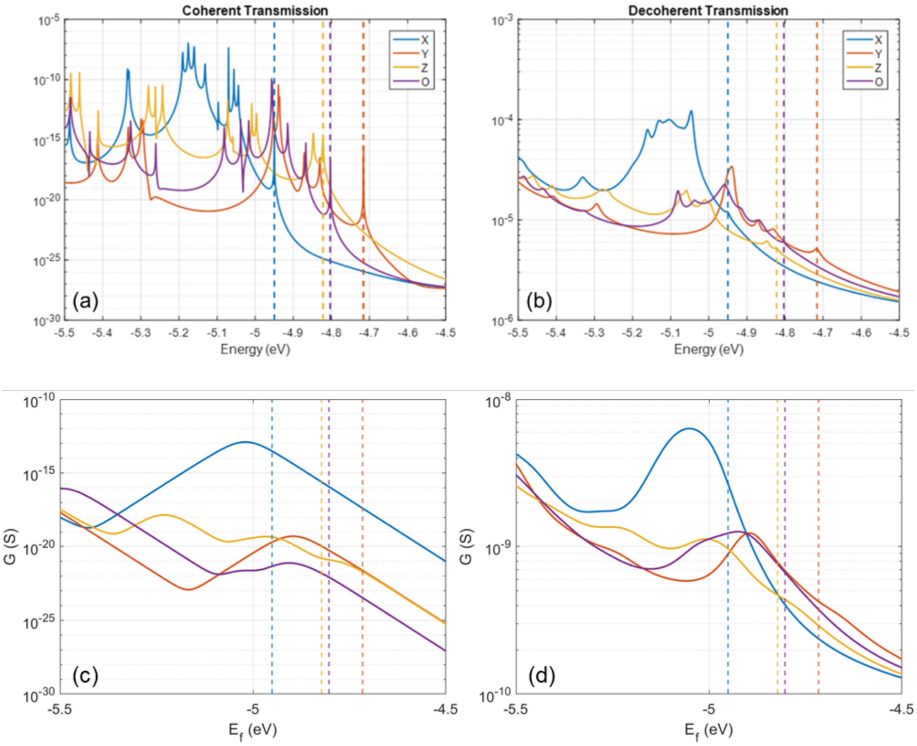


Figure 5.3 Transmission plots, of (a) coherent and (b) decoherent transmissions. (c) and (d) are the corresponding zero-bias conductance plots for coherent and decoherent transmissions, respectively. The dashed vertical lines represent the HOMO level of each structure. X, Y, and Z represent the three mismatch cases, where O represents the original strand.

To understand the reason for the decoherent transmission trend (Y-mismatch >Original structure>Z-mismatch >X-mismatch), we look at the contour DOS plots of the different structures (Figure 5.4). The grey dashed lines in Figure 5.4 represent the HOMO level of each structure. And the narrow peaks along the horizontal direction in the DOS plot show that the DOS is delocalized among the neighboring bases. We compare between the cases by focusing on the region closest to the HOMO levels of the structures, and several observations can be made. First, the Y-mismatch has the highest HOMO level, which makes it the easiest candidate for ionization. Second, the Y-mismatch case shows the highest delocalized states above -5 eV when compared with the others, where an almost fully delocalized level can be seen around -4.93 eV. Third, in Y-mismatch, the highest HOMO-1 to HOMO-N levels (high DOS) are at higher energy levels than X- and Z-mismatches (see regions -4.9 eV to -4.7 eV). As a result, it is easier for charge to move through the Y-mismatch molecule, thus resulting in the highest transmission. The original structure comes second in terms of energy distribution and spatial delocalization. As for X- and Z-mismatch cases, even though X-mismatch has a better HOMO region delocalization, for energy > -4.9 eV, the Z-mismatch has higher DOS values both spatially and energetically and it has a higher HOMO level as well. These features make the Z-mismatch case have a larger transmission than the X-mismatch.

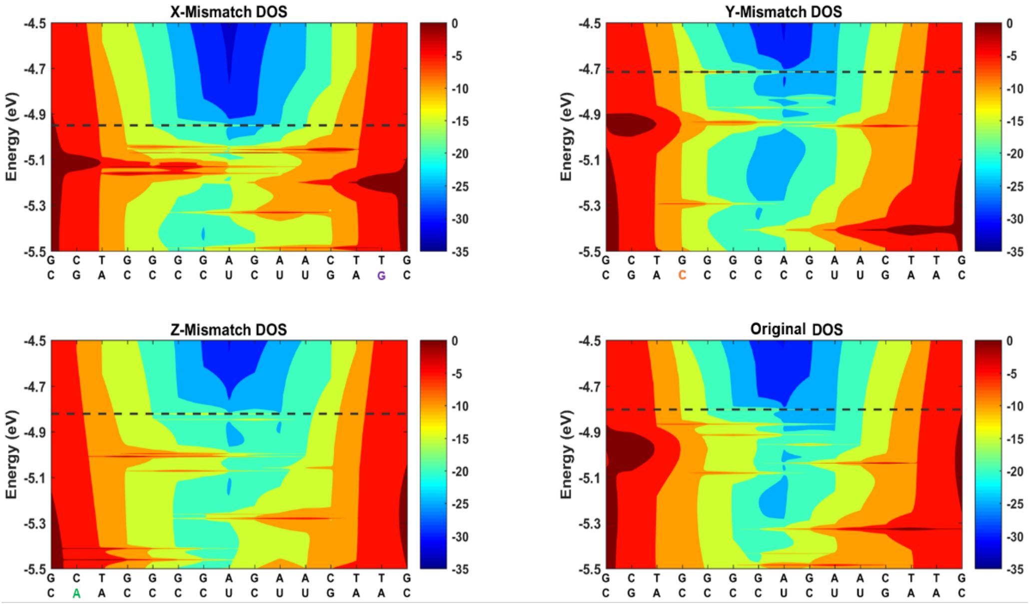


Figure 5.4 Contour DOS plots for the four DNA:RNA structures. The x-axis represents the base-pairs, the y-axis is the energy range. The dashed horizontal line represents the HOMO level of each structure.

5.4 SUMMARY

In this chapter, we investigated the applicability of electrical measurements to detect and differentiate between strands with single nucleotide mismatches. The decoherent results, both transmission and conductance, agree in showing that there is a hierarchy of transmission values between the four studied structures. This is also confirmed by the DOS analysis, thus showing the possibility of differentiating between the different mismatch cases through conductance measurements. Our calculations agree qualitatively with the experimental results in [51], which show that (Y-mismatch > Original > X-mismatch). Experiments done by our collaborator [51] show the conductance of Y-mismatch to be ~40% higher than the original case, and the conductance of X-mismatch is ¼ of the original case. However, conductance of Z-mismatch was below the detectable range (less than 1 pA).

Chapter 6. ROLE OF ANTHRAQUINONE INTERCALATION ON ELECTRICAL CONDUCTANCE OF DNA

6.1 INTRODUCTION

Electronic properties of DNA have attracted particular interest over the past 20 years [14] due to its long-range charge transport and self-assembly properties, making it a desirable nanoelectronics component. DNA consists of four main bases: guanine (G), cytosine (C), adenine (A), and thymine (T), and the stacking and pairing of these bases form a helical structure. The π-π interactions of the stacked bases lead to long-range charge transport. Simultaneously, due to its molecular recognition (exclusive base-pairing), sophisticated geometries can be precisely fabricated, overcoming the scaling challenges with solid-state/silicon-based electronics. Thus, the understanding of the electrical properties of DNA, theoretically and experimentally, becomes a key factor for future device developments.

The last decade has witnessed some claims of the sensitivity of short oligonucleotides' electrical conductance (5 nm range) to chemical modification of nucleotide [151]–[153], conformational change [154]–[157], and single base mismatch [158]. The ability to tune and control these factors has wide implications in electronics, sensors [78], [159]–[167] and sequencing [168]–[170].

Intercalation is an insertion of small molecules between bases of a nucleic acid structure, triggering structural changes such as local unwinding and elongation [171]. Small molecules called intercalators are often used as anticancer drugs to inhibit nucleic acid replication [172] or as fluorescent markers to visualize the structure [173]. The interaction between intercalators and nucleic acids has been studied both experimentally and computationally [174]–[179].

Consequently, there has been a growing interest in controlling nucleic acid conductance by adding intercalators.

Recently, a graphene quantum dot intercalator in a 59-basepair long abasic-DNA was found to enhance the charge transfer rate [132]. An 11-basepair DNA intercalated with coralyne showed three times larger conductance than the native DNA at negative biases [180]. However, other studies have shown a contrasting impact of intercalation on DNA conductance. Harashima et al. [62] demonstrated that intercalation using ethidium bromide increased conductance by four times; they also showed that using Hoechst-33258 as a minor groove binder did not induce an increase in conductance on the same ds-DNA having 8 base-pairs. On the other hand, an earlier experiment conducted by Wang et al. [181] demonstrated that the same intercalator (ethidium bromide) decreased the conductance of 20 base-pairs long ds-DNA. Recently, Xiang et al. [182] reported that an anthraquinone (Aq) molecule directly attached to the end of the backbone of ds-DNA via a uracil (U) base could modify the conductance of a ds-DNA based on its redox state. In their experiment, they showed that the Aq moiety intercalated the complementary strand between the guanine bases, and the reduced state Aq yielded higher conductance, while the oxidized state Aq yielded lower conductance than the unmodified DNA. Further, there have been some recent observations that anthraquinone intercalation can change the conductance by 600 times (unpublished). The experiments do not resolve the underlying mechanisms in a unified manner as exemplified from the literature. A recent modelling paper by Aggarwal et al. [183] reported the effect of ethidium and daunomycin intercalators on the conductance of 8 and 12 base-pair long DNA strands. They used ab initio calculations and showed that upon intercalation, the conductance increases by one order of magnitude mainly due to reducing the twist angle between the intercalated base pairs, enhancing their structural alignment.

In this chapter, we explore the fundamental pathways to alter the conductance of DNA upon intercalation, which is relevant to molecular electronics. We employ density functional theory (DFT), and Green's function-based charge transport calculations. We focus on two different nucleic acids, the ds-DNA and the DNA:RNA hybrid, both intercalated with Aq. We study the impact of redox state, location, and quantity of intercalators in altering ds-DNA's electronic properties considering different measurement scenarios. We report a new approach to analyze the effect of intercalations on the Fermi energy using partial charge transfer concept. Subsequently, we demonstrate that Aq intercalation in ds-DNA can increase or decrease the conductance, depending on the oxidation state, due to the following two reasons: 1) Intercalation decreases the bandgap energy by inducing new energy levels within the bandgap region of the bare ds-DNA, either close to the HOMO region (reduced Aq case) or in the middle of the bandgap (oxidized Aq case), and 2) intercalation shifts the Fermi energy with respect to the HOMO of the intercalated ds-DNA. We also report that the effects of Aq intercalation on both ds-DNA and DNA:RNA hybrids are alike.

We additionally model intercalation by the anthraquinone-neomycin conjugate (AqNEO) in DNA:RNA hybrids. Neomycin has been shown to increase the thermal stability of RNA duplexes [184] and the binding affinity of the conjugated intercalator to DNA:RNA duplexes[185], [186]. We show that the neomycin part of the conjugate, acting as a groove binder, has a small effect on the conductance, which agrees with the experimental findings [187],[181]. We report that this is due to the indistinctive changes on molecular orbitals around HOMO region of DNA:RNA hybrid.

The chapter is organized as follows: first, we introduce the cases under study, we follow it by DFT calculations to find the distribution of the energy levels and compare between the different cases. Then, we focus on the energy levels induced by the intercalators and their impact on transmission. Afterward, we study the effect of having two intercalators on altering electron transmission through DNA. After this focused analysis on the intercalator induced levels, we shift our attention to the overall conductance. Thus, we analyze the partial charge transfer rates to predict the Fermi energy for conductance comparisons. Finally, we study an additional case of a structure with a shorter AT region and illustrate key differences.

6.2 SYSTEM UNDER STUDY

We study a family of three structures: 1) Aq (both oxidized and reduced states) intercalation into a 15 base-pair ds-DNA with the sequence 3′-$G_3A_9G_3$-5′. 2) Aq and AqNEO intercalation into a DNA:RNA hybrid having the same sequence, with the RNA strand formed by the purines and the DNA formed by the pyrimidines. 3) Aq (both oxidized and reduced states) intercalation into a ds-DNA with a shorter AT region, having the sequence 3′-$G_3A_3G_3$-5′. We first focus on 1 and 2 (Figure 6.1a) to understand the intrinsic effects of the intercalation on the strands. The 15 base-pair structures are expected to be comparatively resistive due to the high number of AT base-pairs, creating a barrier for the charge to traverse the structure to the terminal guanines. Thus, as for the 3rd structure, we lowered the AT-region length from 9 to 3 base-pairs.

The exact location of intercalation is difficult to identify experimentally as Aq and its derivatives can reside between any consecutive base-pairs depending on the molecular interactions between the intercalator molecule and the nucleic acids [188], [189]. To account for this, we consider intercalation between **a)** AT base-pairs at 7 and 8 ($I_{7,8}$) and **b)** GC base-pairs at 13 and 14 ($I_{13,14}$) among other possibilities (Figure 6.1).

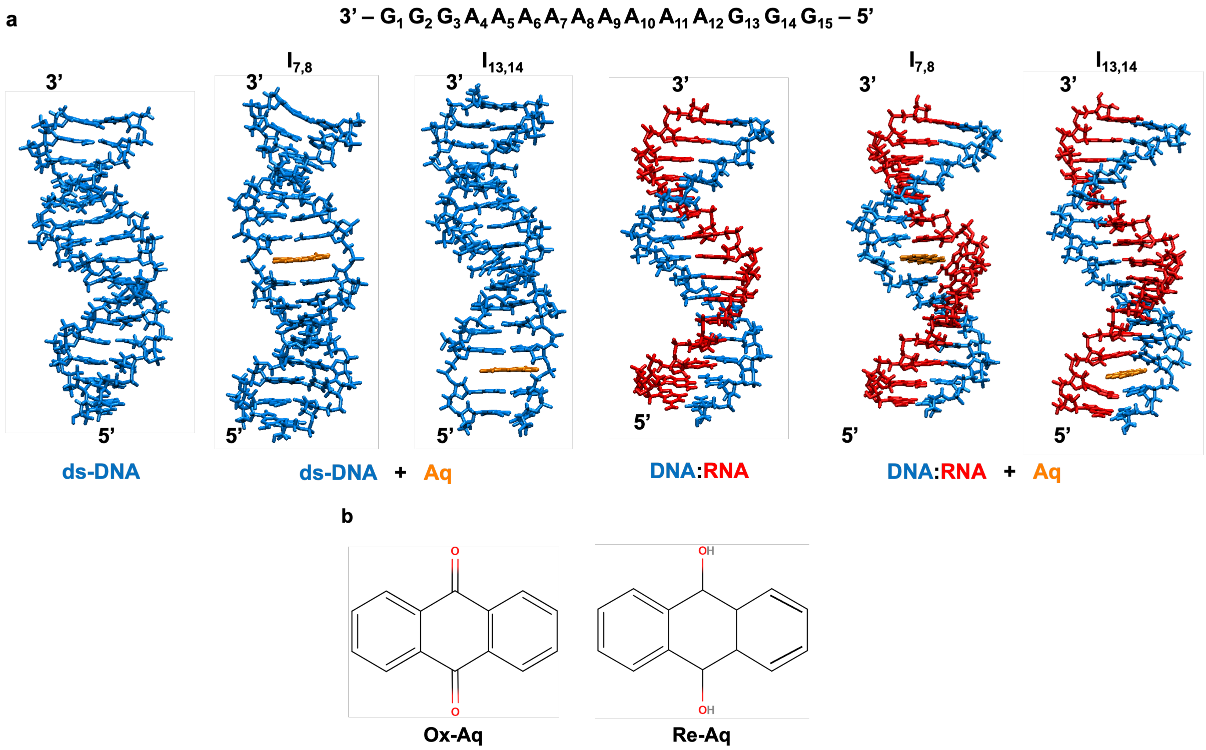


Figure 6.1 Schematic representation of the system: **a,** Simulated intercalation cases for the 15 base-pair long structures. **b,** Oxidized and reduced states of the Aq molecule.

For the DFT step, the water molecules and counterions are removed from the system snapshot (for DFT convergence) and DFT calculations are then carried out using Gaussian 16 [190] with B3LYP/6-31G(d,p) basis set. To account for water solvent effect, we use the polarizable continuum model (PCM). Furthermore, since the counterions that neutralize the DNA backbone are removed from this calculation, the total charge is set equal to the number of phosphate groups in the system, which is -28 (the terminal bases do not include their phosphate groups). The total charge for the ds-DNA system (without AqNEO) is set to -28 for both bare and intercalated cases. For the DNA:RNA system, the total charge is set to -28 for the bare system and for the intercalated case it is set to -22 due to the total charge of +6 on the AqNEO. After reaching convergence, the Hamiltonian from the DFT results is used for charge transport calculations. After obtaining the Hamiltonian from the DFT results, the charge transport calculations are performed for a system consisting of a contact-DNA-contact configuration (Figure 6.2). The calculation is set up as follows: since the DNA strand in STM-BJ methods comes into contact with gold *via* the thiol linker groups at its two ends, we assume contact locations to be at the guanines in 3'-end and 5'-end. The contact self-energy is set at a portion of the backbone of the leftmost and rightmost guanine bases (Figure 6.2). We set the left (right) contact scattering rate Γ_{L} (Γ_{R}) to 100 meV, which resembles moderate contact-molecule coupling. The *E-indep* decoherence scattering rate Γ_{BP} is set at 10 meV. These values are found to be within the acceptable range [191], [192]. The calculations are carried out at room temperature, $k_B T = 0.0259\ eV$. The decoherence probes are applied to each atom in the DNA system. The charge transport modeling is discussed in further detail in chapter 3.

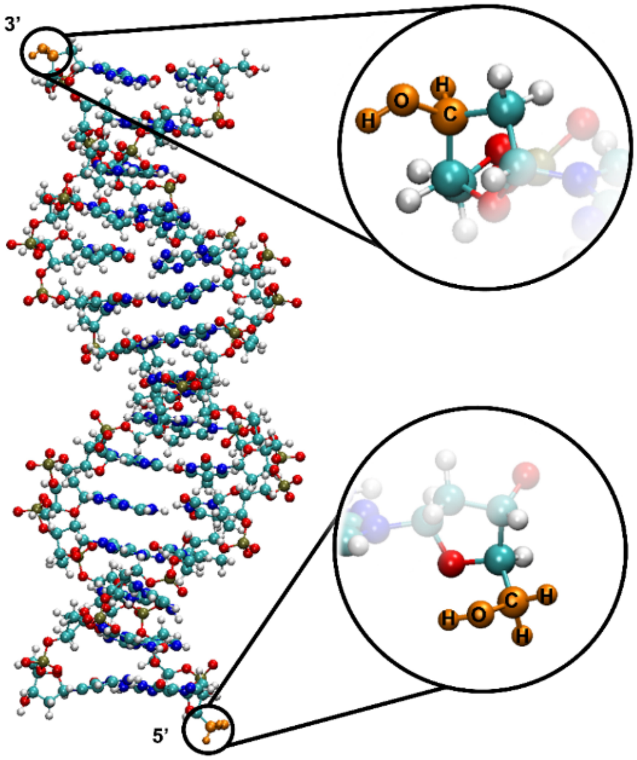


Figure 6.2 The contact atoms (highlighted with orange) are defined at the backbone atoms of the guanine bases located at both 3' and 5' ends of the same strand.

6.3 QUANTUM MECHANICAL MODELING

An essential effect of Aq intercalation on the DNA is in the HOMO and LUMO spatial distribution. Without intercalation, the LUMO is localized at a single Thymine (Figure 6.3a), whereas its HOMO is delocalized on two guanines (Figure 6.3b). However, when Ox-Aq is introduced into the system, its effect is seen as adding a new unoccupied level (LUMO) localized at the Aq. This result is true for both the $I_{7,8}$ and $I_{13,14}$ cases (Figure 6.3c,e). As for the HOMO, it is still localized at the guanines (Figure 6.3d,f). However, in the Re-Aq case, both HOMO and LUMO are localized at the Aq (Figure 6.3g-j). Therefore, it adds two energy levels into the bandgap as opposed to Ox-Aq. The added levels for Re-Aq (Ox-Aq) near HOMO (LUMO) are fully occupied (empty). Hence, this is not like the conventional doping of semiconductors where p-doping (n-doping) has unoccupied (occupied) levels induced near the HOMO (LUMO) band.

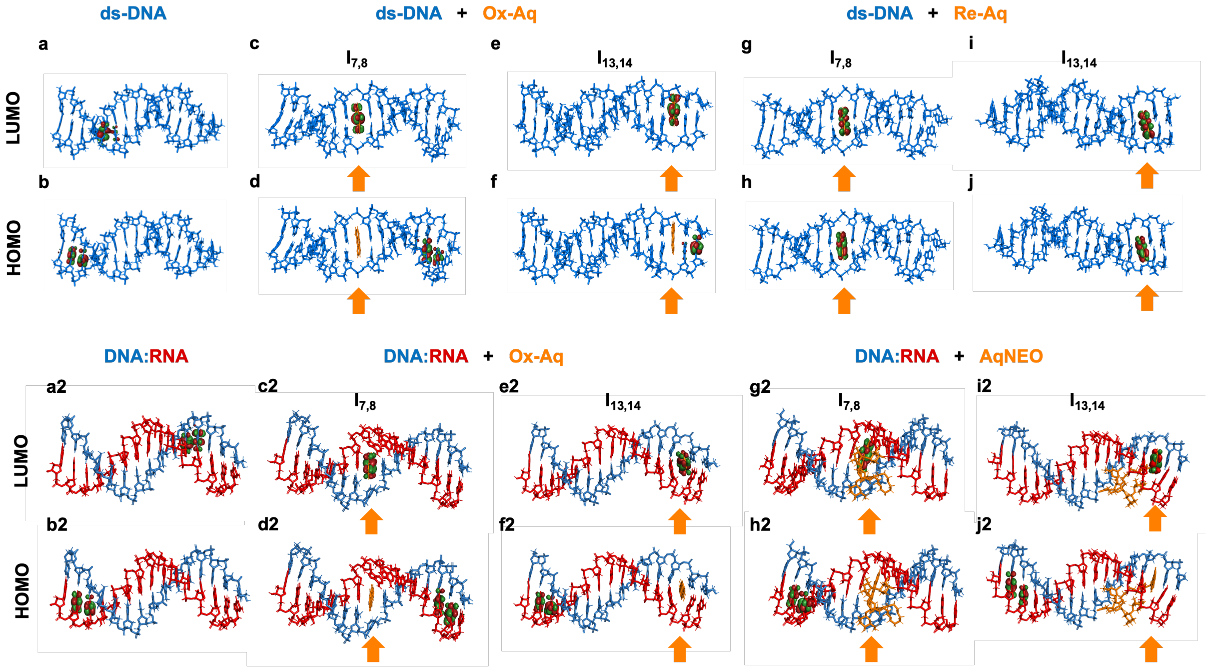


Figure 6.3 LUMO and HOMO plots both without and with Aq (ISO value is 0.02). **a** and **b**, LUMO and HOMO distribution of ds-DNA. **c** and **d**, LUMO and HOMO distribution of ds-DNA intercalated with Ox-Aq between 7th and 8th base pairs. **e** and **f**, LUMO and HOMO distribution of ds-DNA intercalated with Ox-Aq between 13th and 14th base pairs. **g** and **h**, LUMO and HOMO distribution of ds-DNA with same intercalation position with c and d using Re-Aq. **i** and **j**, LUMO and HOMO distribution of ds-DNA with same intercalation position with e and f using Re-Aq. **From a2-j2**: LUMO and HOMO plots both without and with Aq/AqNEO for the DNA:RNA system (ISO value is 0.02). **a2** and **b2**, LUMO and HOMO distribution of DNA:RNA. **c2** and **d2**, LUMO and HOMO distribution of DNA:RNA intercalated with Ox-Aq between 7th and 8th base pairs. **e2** and **f2**, LUMO and HOMO distribution of DNA:RNA intercalated with Ox-Aq between 13th and 14th base pairs. **g2** and **h2**, LUMO and HOMO distribution of DNA:RNA with same intercalation position with **c2** and **d2** using AqNEO. **i2** and **j2**, LUMO and HOMO distribution of DNA:RNA with same intercalation position with **e2** and **f2** using AqNEO.

The bandgap value changes as the intercalator adds new energy levels into the DNA. In the Ox-Aq case, the LUMO induced by Aq is seen in the middle of the bandgap closer to the LUMO region of ds-DNA (Figure 6.4a). The bandgap is reduced by more than 1.2 eV, from 3.67 eV of bare DNA to 2.12 eV for $I_{7,8}$ and 2.45 eV for $I_{13,14}$. The Re-Aq induces two new levels in the bandgap region, a HOMO and a LUMO, which also decrease the bandgap compared with the bare DNA, to 2.97 eV and 2.86 eV for $I_{7,8}$ and $I_{13,14}$ intercalations, respectively. Similar trends are observed for DNA:RNA system (Figure 6.3a2-j2). We observe the LUMO localizing on the intercalating molecule in both Aq and AqNEO cases. As for the HOMO, it remains localized on the guanines. The bandgap is reduced by 1.48 eV for the Aq case and by 1.66 eV with AqNEO (Figure 6.4b). We also notice that the incorporation of neomycin as a groove binding molecule does not significantly alter the molecular orbitals of the system in the HOMO region (AqNEO cases in Figure 6.4b). However, the LUMO region is shifted to lower energies for the AqNEO cases. The average shift in the induced energy between Aq and AqNEO cases is 175 meV. Thus, we find that the groove binder neomycin does not cause a significant effect on the molecular orbitals of the system or the intercalator induced energy levels. In the following section, we explore the transport calculation results for the different cases, specifically at the induced energy levels.

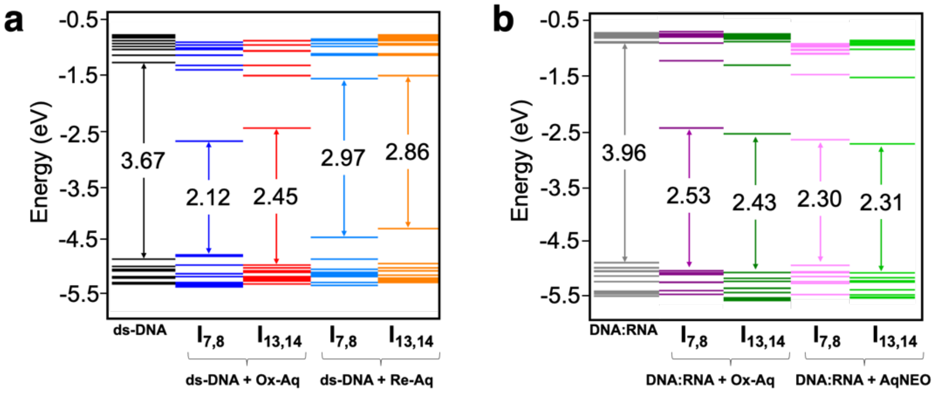


Figure 6.4 Energy levels and bandgap values for different intercalation cases of ds-DNA and DNA:RNA. **a,** Energy levels for ds-DNA system. Ox-Aq induces a LUMO into the bandgap of DNA, while Re-Aq induces both HOMO and LUMO which are close to the energy levels of the DNA, respectively. **b,** Energy levels for DNA:RNA system with Ox-Aq and AqNEO cases. Both Ox-Aq and AqNEO induce LUMO into the bandgap of DNA:RNA.

6.4 TRANSMISSION AT THE INDUCED LEVELS

The direct comparisons between the original states of ds-DNA and DNA:RNA is discussed in prior references [77], [193]. In this work, we focus on the intercalator's effect on charge transport in the ds-DNA and DNA:RNA systems. The transmission is plotted as a function of energy ranging from -6 to 0 eV in Figure 6.5.

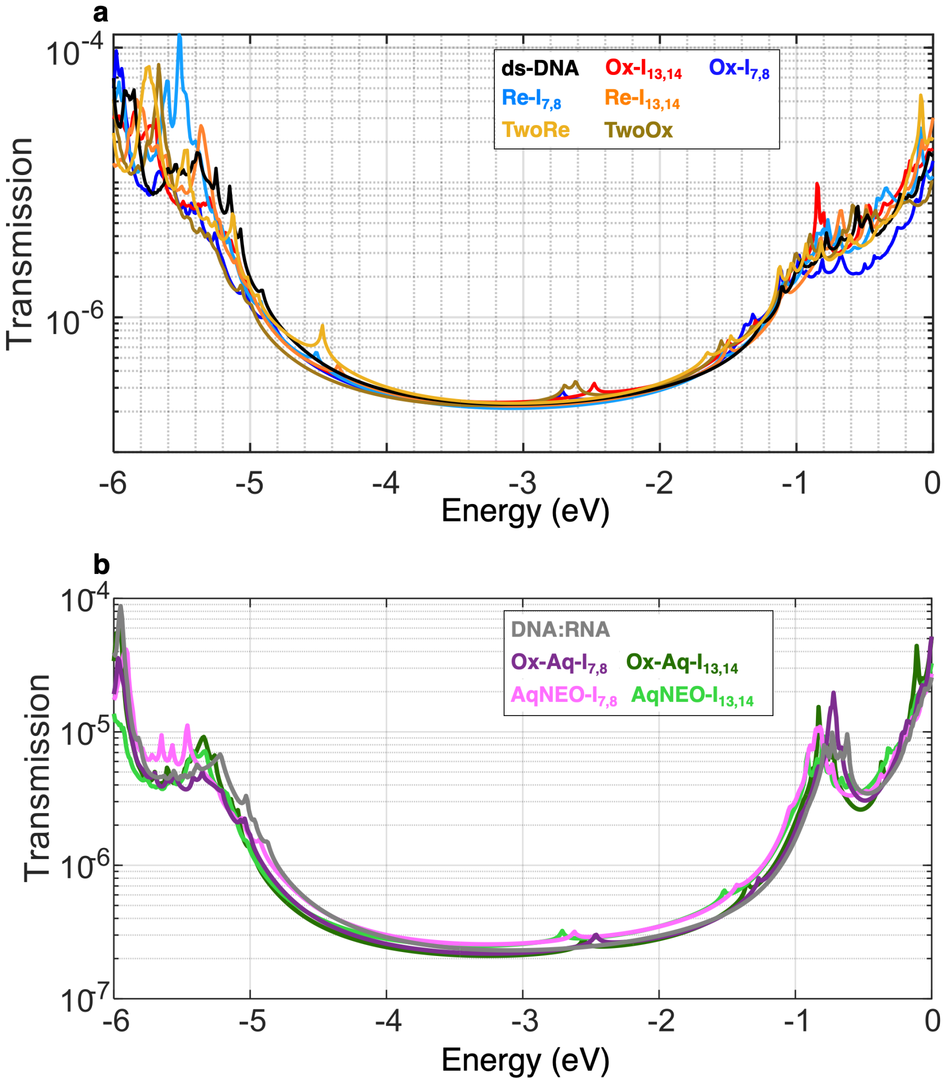


Figure 6.5 Transmission over a large energy window spanning HOMO and LUMO for **a** ds-DNA 3'-G_3A_9G3'5', **b** DNA:RNA.

Comparison of the results for intercalating Aq between the 13th and 14th base-pairs (the $I_{13,14}$ case) of the ds-DNA as well as the DNA:RNA system is given in Figure 6.6. We notice that the induced energy level by the intercalator is very similar for both cases (80 meV shift in energy), and the transmission peaks only vary by less than 15%. The impact of intercalation on either ds-DNA or DNA:RNA systems is similar as we stated previously.

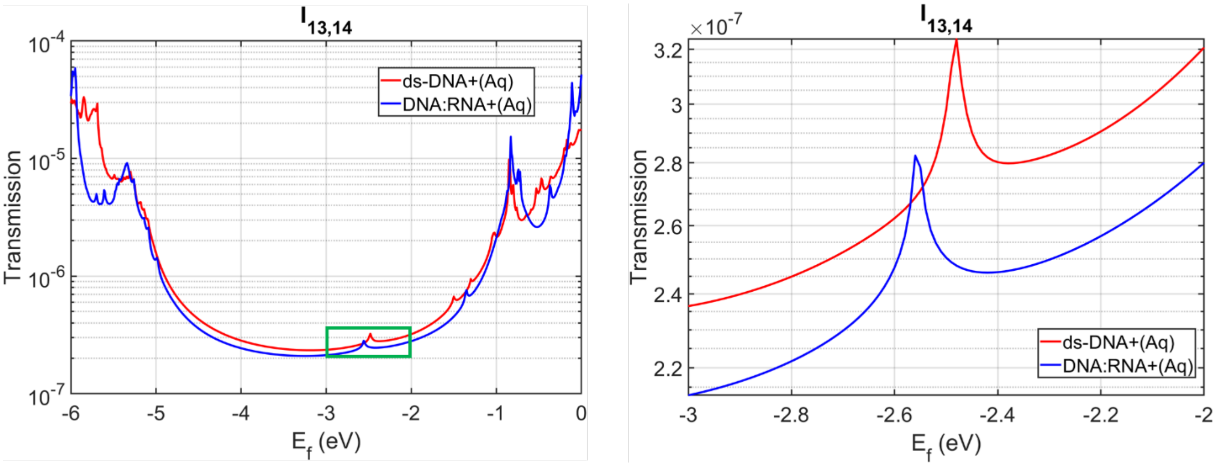


Figure 6.6 (Left)Transmission plot of ds-DNA and DNA:RNA systems with Aq intercalation at the $I_{13,14}$ location. (Right) Transmission peaks of the Aq-induced levels highlighted in the green box in the left figure.

Next, we shift our focus to the transmission near the intercalator induced energy levels (Figure 6.7). The transmission increases at the Aq-induced energy levels by 16-34% compared to bare DNA (see transmission at the starred peaks in Figure 6.7). The Re-Aq cases ($I_{7,8}$ and $I_{13,14}$ in Figure 6.7a) show a ~16% increase, whereas Ox-$I_{7,8}$ and Ox-$I_{13,14}$ increase the transmission by 23% and 34%, respectively (Figure 6.7b). Similarly, the DNA:RNA system with AqNEO (Aq) exhibits the same trend (Figure 6.7c): the intercalator-induced peaks increase the transmission at their respective energies compared to the bare system. It is essential to mention that the induced peaks for Aq (or AqNEO) in its oxidized state are deep in the bandgap region of the no intercalated cases, -2.8 eV < Energy < -2.4 eV (Figure 6.7b,c and Figure 6.5). This range is more than 1.5 eV above the anticipated Fermi energy (E_f) location, which is assumed to be in the HOMO vicinity.

We expect that the closer the Aq-induced levels get to the high density of states (DOS) region, the higher the transmission becomes. This is due to the enhanced coupling of Aq to neighboring bases. The location of intercalation has a small effect on the induced energy levels as shown in Figure 6.4. For instance, changing the Ox-Aq location from $I_{7,8}$ to $I_{13,14}$ induces an energy shift of 245 meV, and the corresponding increase in transmission peak is ~14%. As for the Re-Aq, it induces an occupied level close to the HOMO vicinity. Here, the Re-$I_{7,8}$ has its HOMO energy level closer to the HOMO of Re-$I_{13,14}$ by 163 meV, and results in a ~25% increase in transmission peak (Figure 6.7). As for the DNA:RNA system with Aq, Figure 6.7c shows that by changing the location of intercalation from $I_{7,8}$ to $I_{13,14}$, the induced level shifts by 91 meV closer to the LUMO region, only increasing the transmission peak by 7%. On the other hand, in the neomycin conjugate (AqNEO) case, the induced level shifts by 89 meV away from LUMO. However, this opposite shift yields a change in transmission peak of less than 1%. We find that transmission is not affected by the addition of the groove-binder, which is in qualitative agreement with experiments[187], [194] that finds groove-binding molecules have an insignificant impact on DNA conductance. In the next section, we look at the effect of having double intercalations. We also like to note that since the effect of intercalation yields similar trends regardless of the system (ds-DNA or DNA:RNA), the following sections will focus on the ds-DNA system.

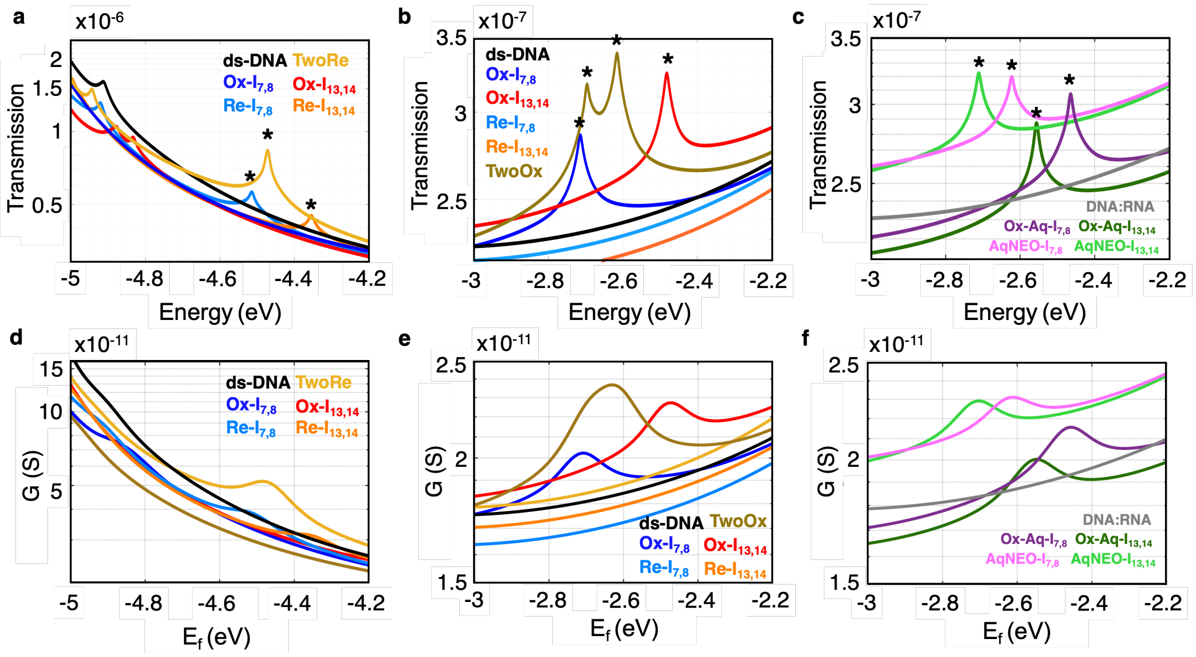


Figure 6.7 Transmission plots for **a**, ds-DNA system in the HOMO region (stars: Re-Aq induced peaks). **b**, ds-DNA system in the Aq-induced region (stars: Ox-Aq induced peaks). **c**, DNA:RNA system in the Aq-induced region (stars: Aq and AqNEO induced peaks). The x-axis is energy because these results do not depend on the occupancy or Fermi factor of the electrons. Conduction plots for **d**, ds-DNA in the HOMO region, **e**, ds-DNA in Aq-induced region and **f**, DNA:RNA in Aq-induced region.

6.5 DOUBLE INTERCALATION

Multiple intercalations may occur in experiments depending on the intercalator molecule concentration and the length of the DNA strand[195], [196]. We intercalate two anthraquinone molecules corresponding to $I_{7,8}$ and $I_{13,14}$ positions.

The transmission results show that in the Re-Aq case, a single transmission peak occurs near the HOMO region for double intercalators (TwoRe), which is ~83% higher than the single intercalation cases (Figure 6.7a). However, in the Ox-Aq case, the double-intercalation (TwoOx) yields two transmission peaks close in amplitude to the single intercalation cases (Figure 6.7b). To understand this, we look at the contour DOS plot. The plot shows the distribution of electrons as a function of both space and energy. Figure 6.8a shows that the two intercalators induce two nearly degenerate levels near the HOMO region of the DNA for the Re-Aq case. Hence, having higher coupling with the neighboring bases of the DNA, they yield a large transmission peak. However, in the Ox-Aq case, the energy levels are deep in the bandgap region (dark blue area in Figure 6.8b). Therefore, they are residing in an almost isolated region with low coupling to the DNA bases. Hence, increasing the number of Re-Aq intercalators can yield higher transmission probability at the induced levels.

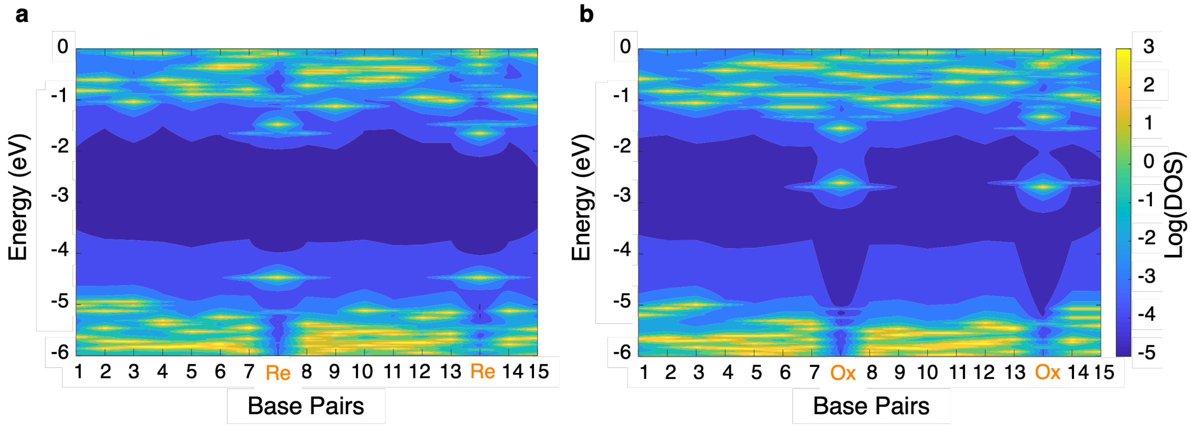


Figure 6.8 Double-Intercalation effect on transport properties of the ds-DNA. **a,** the contour DOS plot of double Re-Aq, **b,** the contour DOS plot of double Ox-Aq, showing the energy levels distributions along the strand.

Although we observe energy shifts and transmission peaks can quantitatively vary, the average percentage of transmission increase is only 25% compared to the no-intercalation case. These results indicate that the spatial location of intercalation -alone- cannot be the reason behind the intercalation induced conductance differences reported in the literature. In the following section, we consider a hopping model, which also supports this conclusion.

6.6 EFFECT OF INTERCALATION LOCATION ON TRANSMISSION

To further investigate the effect of intercalation location on the long AT region structure (ds-DNA), we focused on the charge transport mechanism induced by the intercalators by using two-step hopping scheme shown in Figure 6.9. The aim of this analysis is to see the impact of distance on the transmission between Ox-Aq and the ds-DNA. The first calculation in this scheme sets the left contact location at the leftmost guanine base, while the right contact is set at the Aq. This yields the transmission probability for an electron to enter the DNA and hop into the Aq, T_1. The second calculation, T_2, assumes the left contact is at the Aq, while the right contact is set at the other end of the DNA strand, which is also a guanine. This will yield the transmission probability for an electron present at Aq to traverse the structure and exit the DNA through the other end. Then, we calculated the equivalent transmission for this system using equation (6.1).

$$\frac{1}{T_{eq}} = \frac{1}{T_1} + \frac{1}{T_2} \tag{6.1}$$

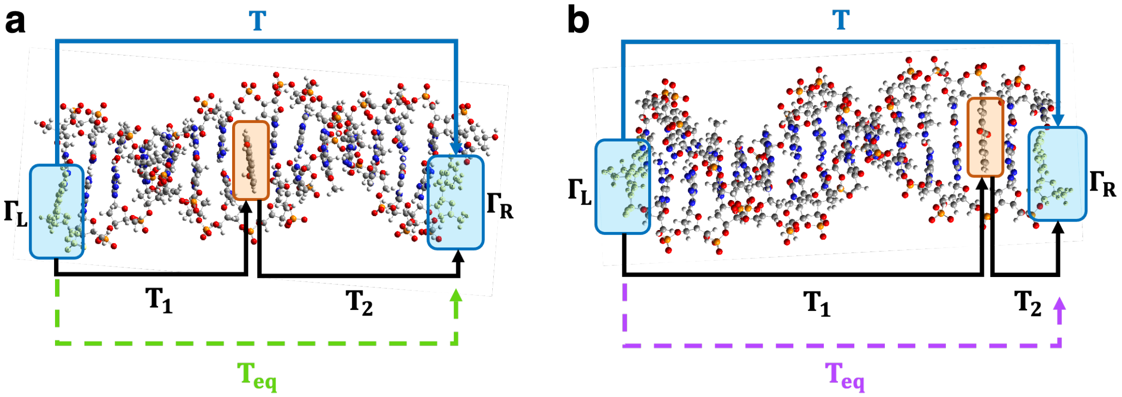


Figure 6.9 The two-step hopping scheme for $I_{7,8}$ (**a**), $I_{13,14}$ (**b**). $T_1$is the transmission from leftmost guanine to Aq. T_2 is the transmission from Aq to rightmost guanine. T is the transmission from leftmost guanine to rightmost guanine as in the total transmission from left to right, which is used in the main text. T_{eq} is the equivalent transmission calculated using equation (6.1).

The results show that the distance between contact points and intercalation molecule (thus, the intercalation location) plays a significant role in transmission unless the strand is too long to show the effect(Figure 6.10). As even though the second hop, T_2, is two orders of magnitude higher than T_1 for $I_{13,14}$, the equivalent transmission is going to be lowered due to the low value of T_1. This is expected since T_2 has the right contact location (Guanine) very close to Aq (0.7 nm), while the first hop requires the electron to travel a large distance of 12 base pairs (3.68 nm) to reach the Aq which dominates the overall transmission along the system. Moreover, a similar effect is seen for the $I_{7,8}$ case where the electron must travel almost the same distance per hop, 2.38 nm and 2.72 nm for T_1 and T_2, respectively. Even though the induced energy levels for $I_{13,14}$ and $I_{7,8}$ are at different spatial locations, the equivalent tunneling through the whole molecule is weak regardless of the location of Aq.

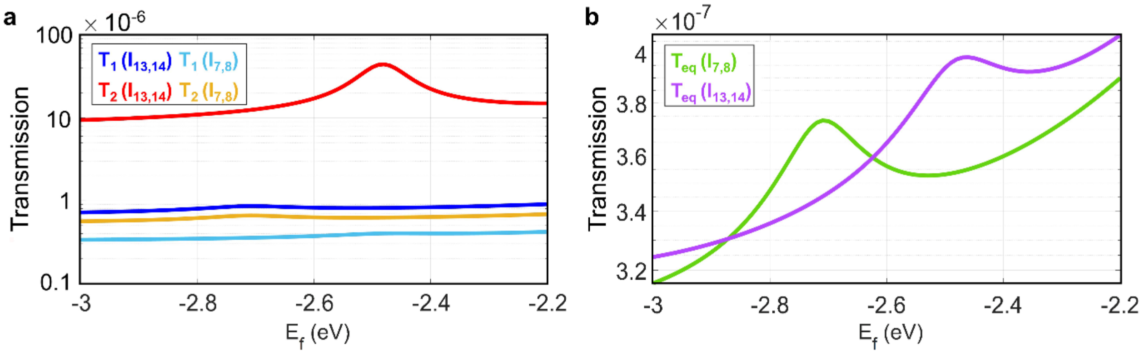


Figure 6.10 **a,** Transmission plots of Ox-$I_{13,14}$ and Ox-$I_{7,8}$, showing T_1 and T_2, with T_2>>T_1. **b,** Comparison between T_{eq}, showing that both cases yield similar transmission regardless of the location of Aq.

The calculation was also carried out for the two Ox-Aq molecules intercalating the DNA. In the sequential model, this involves three hops to go from the left contact to the right contact (Figure 6.11):

$$\frac{1}{T_{eq}} = \frac{1}{T_1} + \frac{1}{T_2} + \frac{1}{T_3} \tag{6.2}$$

The results confirm the length effect on lowering the electronic coupling between the two ox-Aqs, where the induced energy levels are also deep in the BG which isolates them from the molecular orbitals of the DNA. Indeed, as even though $I_{13,14}$ would have a higher coupling to the right contact, the low coupling between the two Aqs lowers the transmission along the structure. Hence, the lower transmission to/from $I_{7,8}$ results in an equivalent transmission that is comparable to single ox-Aq intercalation as discussed in the main text.

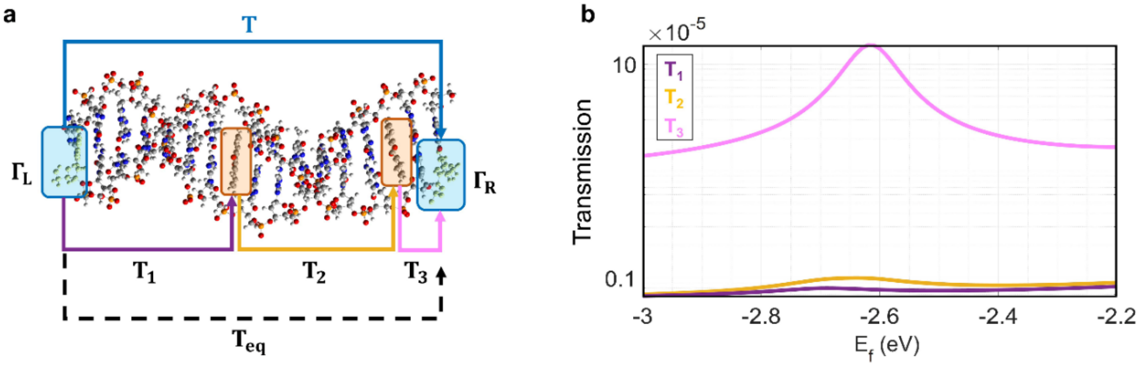


Figure 6.11 **a,** Sequential hopping scheme for the two anthraquinone intercalation case, going from leftmost guanine, to the two anthraquinones and exiting through the rightmost guanine. **b** Comparison between T_1, T_2, and T_3, showing distance effect on transmission which causes lower equivalent transmission.

So far, we have focused our discussion on the Aq-induced levels. The calculations show that the induced energy levels inside the bandgap cause a transmission increase of less than 45% for this 15 base-pair long strand. At the same time, a higher percentage increase is reached with double-intercalators. Compared with literature, in which more than three times increase was reported, we deduce that the induced levels cannot be the only factor in increasing the DNA conductance. To probe other mechanisms that may influence the conductance, in the next section, we look at the overall conductance profile of the ds-DNA cases at different energy regions.

6.7 EFFECT OF CONTACT COUPLING

The transmission trends obtained are similar when the contact self-energies are changed. Within the range tested, the increasing (decreasing) contact self-energy increases (decreases) transmission. Figure 6.12 displays two examples of how the transmission profile is qualitatively maintained, while a y-axis shift is visible depending on the contact scattering rate value. The top plot in Figure 6.12 is for the ds-DNA without any intercalators, while the bottom plot is for a case with an intercalator. The transmission amplitude increases as we increase the contact scattering rate, going from 50 meV to 1000 meV. However, the ratio between the amplitudes is not directly proportional to the rate values. For instance, in HOMO (-6 to -5 eV) or LUMO (-1 to 0 eV) regions, the transmission at 1000 meV is about 3.5 times as much as the transmission at 50 meV. While in the region of interest (bandgap region insets of Figure 6.12), the change is lower than 14%. The reason behind this is that the broadening due to the contacts is mostly effective at the molecular orbitals localized at or closest to the contact locations. The main component of the resistance at many energies of interest arises from the intrinsic resistance of the DNA. Deep inside the HOMO and LUMO bands, the transmission can increase significantly with the contact coupling. This corresponds to the case where states are much more extended throughout the DNA in a small energy window. These deep lying states are however not accessible in experiments and do not affect the observed trends (example is shown in the LUMO region insets of Figure 6.12).

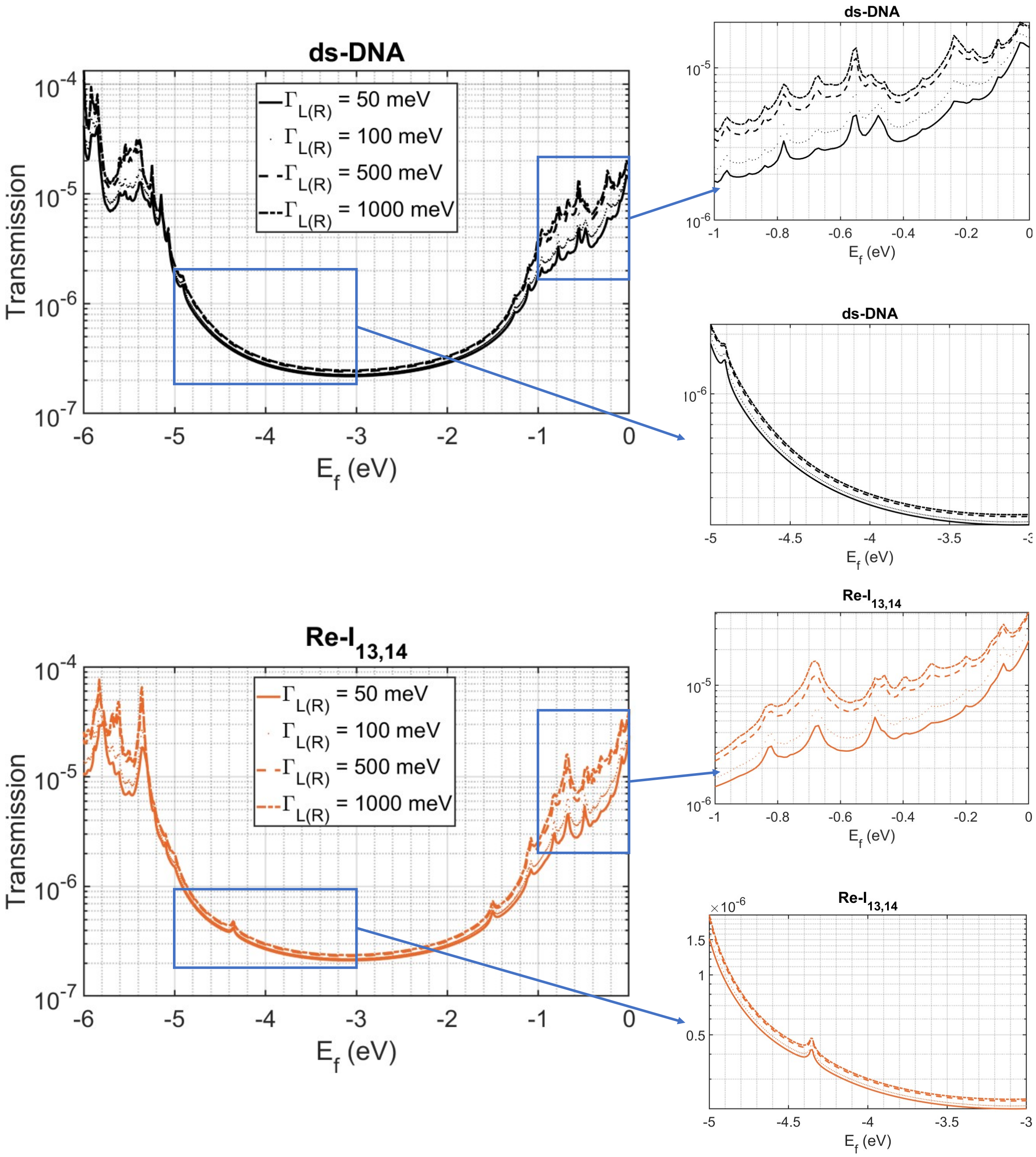


Figure 6.12 Transmission plot for ds-DNA case under different contact coupling rates. The regions with high density of states (molecular orbitals), i.e., deep in HOMO or LUMO regions, can have some peaks affected and become larger due to the higher broadening of these molecular orbitals at higher coupling rates (top inset). However, the general trend seen is maintaining the transmission peaks while observing a complete shift in transmission amplitude (y-axis shift). Bottom figure is for the Re-$I_{13,14}$ case.

6.8 CONDUCTANCE AND FERMI ENERGY LOCATION

In this section, we present results for two types of modelling scenarios: ds-DNA is connected in between two electrodes with and without a gate electrode system, respectively named as the three-terminal and the two-terminal measurement scenarios. In the three-terminal case (Figure 6.13b), the Fermi energy is swept between the HOMO and LUMO energies using a gate electrode. In the two-terminal case (Figure 6.16b) we estimate the Fermi energy location. We assumed that the electrodes are coupled to the DNA ends via thiol linker groups in both experiments, which is the common approach in STM-Break Junction setups.[197], [198] This assumption provides us with the information that for small biases we operate in the linear response regime[153], [192], [199] (where most of the potential drop occurs across the electrode-thiol linker junction). However, we have not explicitly included the thiol linkers or the electrodes due to the enormous computational complexity.

6.8.1 *Three-Terminal Measurement Scenario (Fermi Energy Sweep)*

Here, we assume a fixed small bias applied to the molecule while we sweep the Fermi energy *via* a third gate electrode and calculate the conductance. The conductance plot is shown in Figure 6.13a as a function of E_f, where we shifted the HOMOs of all molecules to match. If the Fermi energy is at the HOMO for each case (or HOMO plus a few hundred meV), it is apparent that the conductance is largest for oxidized Aq (Ox-$I_{13,14}$), followed by bare DNA and the two remaining oxidized cases. In contrast, the Re-Aq cases have the lowest conductance values. The reason behind this is that the HOMO of the Re-Aq cases is localized at the Aq, whereas it is localized at the DNA strand for Ox-Aq (see Figure 6.3). Therefore, we predict that the three-terminal experiments should result in this conductance trend when E_f is close to the HOMO of each structure.

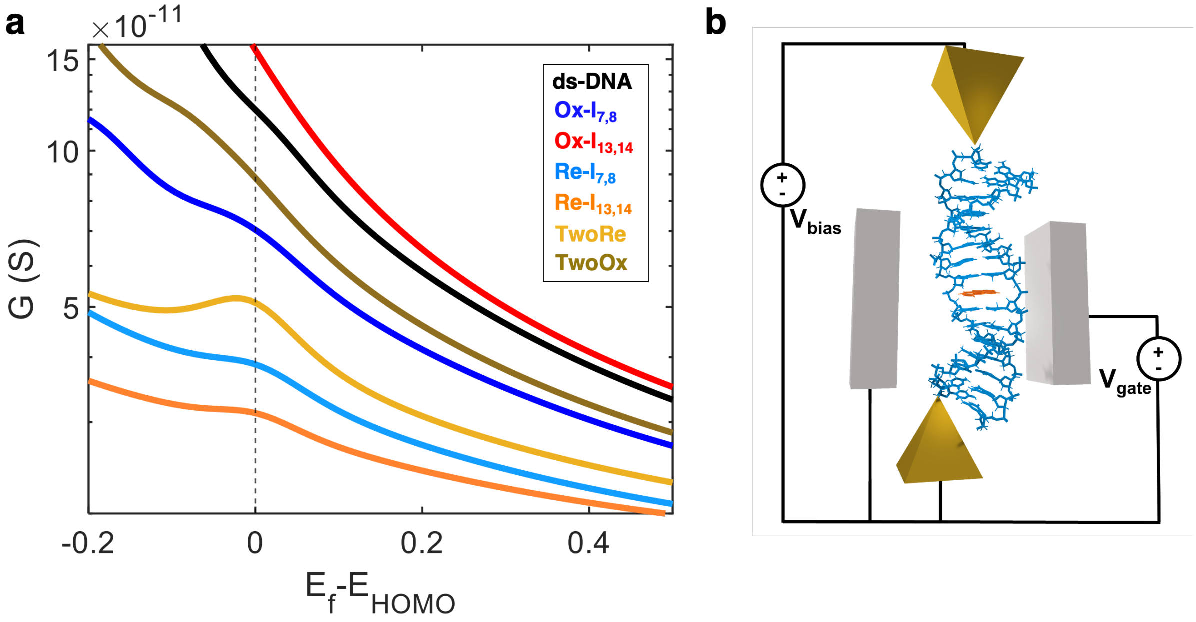


Figure 6.13 **a,** Conductance plot as a function of Fermi energy for ds-DNA intercalation cases with all HOMOs aligned (dashed line). **b,** three-terminal measurement scenario.

In contrast to a sweeping the Fermi energy, in the next section we explore a broader approach by estimating the location of Fermi energy for each case using the principle of partial charge transfer.

6.8.2 *Two-Terminal Measurement Scenario (Fermi Energy Location)*

The second type of measurement scenario involves two-terminal measurements (without a gate). Here, we do not know where the contact's Fermi energy lies *a priori*, but it is determined *via* the partial charge transfer method. In molecular junctions, it is known that the nature of the interface (gold electrode /ds-DNA) affects the Fermi energy of the contact. The adsorption of the molecule to the contacts causes both broadening of the molecular energy levels and partial charge transfer. Both the geometry and the orientation of the gold contacts may affect the amount of broadening and charge transfer within the molecular system under study. Including these effects in the calculations will, in principle, allow the evaluation of the Fermi energy position. This approach has been utilized previously on small molecules resulting in a wide range of Fermi energy locations extending from HOMO to LUMO [143]. Given this and the difficulty of including gold contacts for studying a molecule of more than a thousand atoms [200], we take an alternate approach proposed by Zahid and collaborators [143]. They suggested calculating the conductance at various Fermi energies to understand experimental trends.

In the DNA-contact system, we expect a shift in the molecular energy levels with respect to the Fermi energy of the gold contacts. Partial charge transfer occurs from the contact to the DNA, shifting the MOs to higher energies. The E_f is in general closer to HOMO than to LUMO. Our approach here is to estimate E_f using the rate of change in energy as we vary the number of electrons in the system within the bandgap region. The higher the rate is, the higher the shift in the MOs, and the closer the HOMO gets to E_f. To achieve this, we first estimate the number of electrons in the bandgap region by integrating the local density of states (LDOS) using the Green's function as:

$$N(E_i) = \iint LDOS(x, E) f(E - E_i) dx dE = \frac{1}{\pi} \iint -\mathrm{Im}[diag(G^r)] f(E - E_i) dx dE \qquad (6.3)$$

where x represents the location of the atoms in the system, and $f(E - E_i)$ is the Fermi-function, which describes the probability of electron occupancy at energy E when $E_f = E_i$.

The electron count $N(E_i)$ is calculated at room temperature, using the same charge transport calculations set up. It is important to mention that since the gold contacts are not explicitly included in the model, the resulting Hamiltonian from DFT calculations that is used in the Green's function (GF) does not yield the exact electron number of the system. Thus, the GF approach requires the Hamiltonian to be updated self-consistently while solving Poisson's equation with the appropriate boundary conditions to account for the contact atoms' effect on the electrostatics, which would yield the accurate electron count. However, in our approach, we are not solving explicitly for the total number of electrons of the system, but rather looking at a smaller energy window (bandgap) to investigate the amount of shift caused by partial charge transfer. Electron count for each case is plotted in Figure 6.14, where the flat regions resemble the bandgap. We can then find the number of electrons required to go from HOMO to LUMO by following this flat region. The first thing to notice is that for the Ox-Aq (blue, red, and black), sudden jumps are seen in the electron count as we pass from -3 eV to -2.4 eV. The charge increases by 1 electron (2 electrons for the two Aq case), which is consistent with the location of the LUMO induced by the Aq. We assume the partial charge transfer to be less than one electron, and this limits the energy sweeping for Ox-Aq cases to energies between -4.5 and -3 eV. In this energy window, the electron count increases only by 0.4 electron (0.013 electron/base). As for the Re-Aq cases, they provide a larger bandgap (energy window) for the energy sweep. Therefore, the electron count increases by 0.8 electron (0.027 electron/base) as energy is swept from -4.2 eV to -1.8 eV. As for the bare DNA, 0.73 electron (0.024 electron/base) increase is seen when going from -4.5 eV to -1.5 eV. It is important to note that the values for bare DNA and Re-Aq are to go from HOMO to LUMO, whereas in principal this need not be the case as Fermi energy of the contacts can be assumed to be initially in the middle of the bandgap (around -3 eV). Therefore, we calculate the electron count for energy windows similar to the Ox-Aq cases (from -3 eV to HOMO, per case). The calculations show that it will require a smaller amount of charge to go to the vicinity of HOMO (an average of 0.44

electron for Re-Aq cases, and 0.46 electron for ds-DNA). These values are comparable to the literature [201], where a thiol-conjugated hexathiophene molecule was shown to have ~0.06 electron charge transfer occurring from the electrode to the molecule. Noting that the hexathiophene has 54 atoms, which can correspond to two DNA nucleotides. Thus, the amount of charge required for our system (0.013-0.027 electron/base) can be reasonably supplied by the partial charge transfer occurring from gold contacts into the molecule.

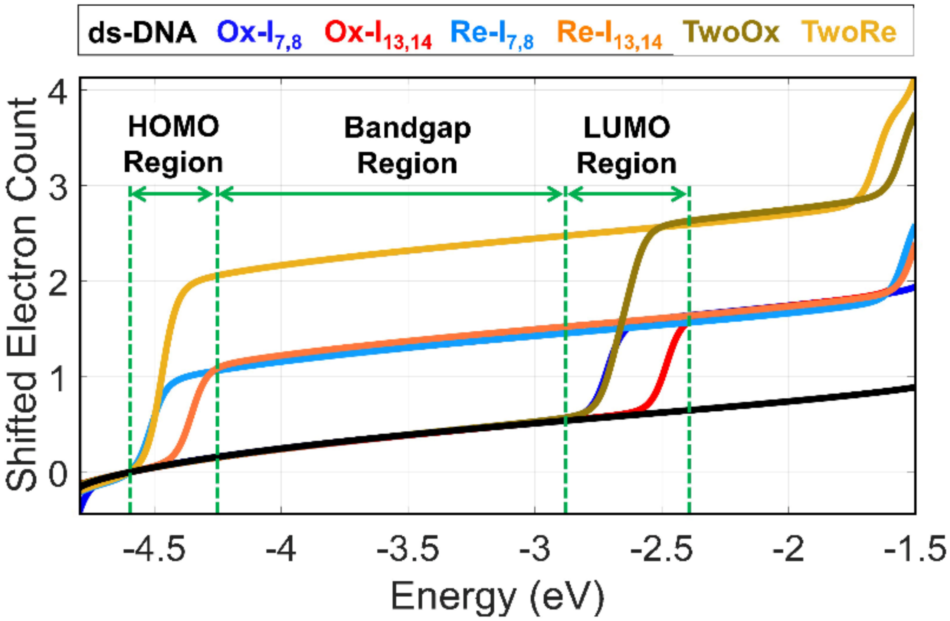


Figure 6.14 Shifted electron count for each case found by integrating DOS, focused on the bandgap region and the edges of HOMO and LUMO regions.

This difference in electron number within the bandgap region tells us that the molecular orbitals shift by a different amount in each of the cases considered. Thus, the Fermi energy location is not necessarily close to HOMO, as will be shown next. Figure 6.15 shows a sample of how dE_i/dN is extracted. The flat region between the two peaks resembles the bandgap in which the average value of the inverse plot is taken to extract dE_i/dN.

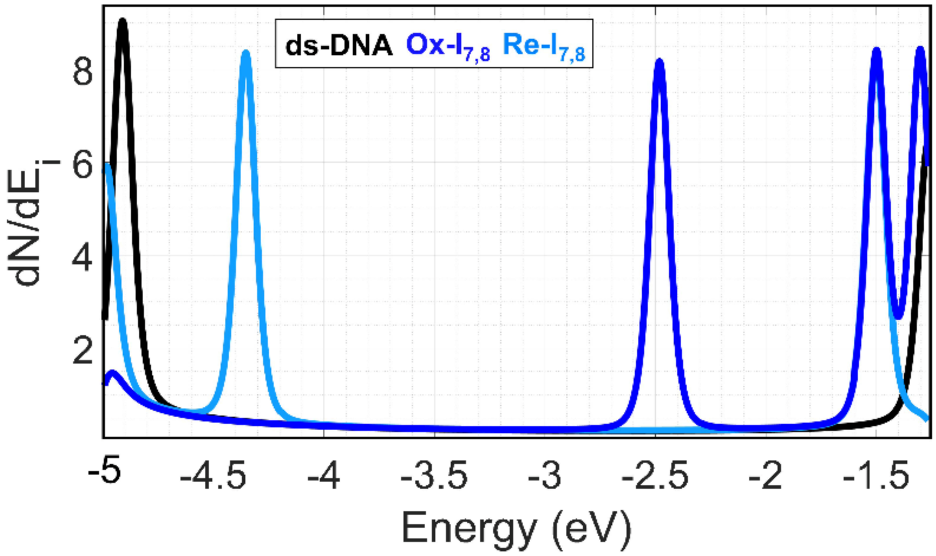


Figure 6.15 The average rate of change is taken from the flat region between the peaks. The peaks at energy < -4 eV (> -3 eV) resemble the HOMO (LUMO) of the different cases. Only three cases are shown for clarity.

Calculating the derivative (dE_i/dN) gives us the average change in energy dE_i for a small change in electron number dN. The average change per one electron for each case is listed in Table 6-1. The values show that Ox-Aq cases have the lowest rates, by an average difference of 0.81 eV/electron from bare DNA. Re-Aq cases have a higher rate than the bare DNA by an average of 0.47 eV/electron. This trend can explain how partial charge transfer affects each case: when the same amount of partial charge is transferred from gold contact to DNA; 1) Ox-Aq has the lowest shift of molecular orbitals, which translates to E_f being the farthest from HOMO region. 2) Re-Aq, has the largest shift, making E_f the closest to HOMO. 3) Bare DNA is between Ox-Aq and Re-Aq cases.

Table 6-1 Average rate of change in energy with respect to number of electrons for ds-DNA and DNA:RNA cases.

ds-DNA			DNA:RNA		
Molecule	dE_i/dN **(eV/electron)**	**Change**	**Molecule**	dE_i/dN **(eV/electron)**	**Change**
TwoOx	2.61	-0.93	Ox-AqNEO $I_{7,8}$	2.57	-0.29
Ox-$I_{7,8}$	2.61	-0.93	Ox-AqNEO $I_{13,14}$	2.62	-0.24
Ox-$I_{13,14}$	2.97	-0.57	Ox-Aq $I_{7,8}$	2.66	-0.20
ds-DNA	3.54	0	Ox-Aq $I_{13,14}$	2.72	-0.14
TwoRe	3.75	0.21	DNA:RNA	2.86	0
Re-$I_{7,8}$	4.13	0.59			

Therefore, we can estimate the location of E_f by incrementing the amount of partial charge transfer (dN), using the following equation:

$$E_f = E_{f_0} - \left(\frac{dE_i}{dN}\right) \times dN \tag{6.4}$$

where E_{f_0} is the Fermi energy of the electrode before making the contact-molecule junction, dE_i/dN is taken from Table 6-1, dN is a sweeping parameter, and E_f is the expected location of the Fermi energy after the MOs shift due to the partial charge transfer. We assumed different starting points for E_{f_0}, and extracted the conductance of each case at their respective E_f. The main condition for dN is to be below 1 electron, and the maximum allowable shift for E_f (from E_{f_0}) is to reach the HOMO of each case. This condition limits dN to have a maximum corresponding the HOMO of each case .

Next, we compared the conductance values with the bare DNA case, and the ratios are plotted in Figure 6.17. For E_{f_0} = -3.5 eV (Figure 6.17a), notice how at dN > 0, conductance ratios for TwoRe and Re-$I_{13,14}$ become greater than 1, while Re-$I_{7,8}$ needs dN > 0.2 electron to reach ratios greater than 1. As for the Ox-Aq cases, the ratios are below 1 for $dN > 0$. The results are consistent for the different E_{f_0} cases (Figure 6.17 a-c) and show a maximum increase in conductance of ~1.6 times for Re-Aq, with an average ratio higher than bare DNA. The Ox-Aq cases have average ratios of less than 1. We also notice that the maximum ratio value decreases at $E_{f_0} = -4$ eV (Figure 6.17d). This result indicates that relatively higher conductance ratios are expected if the initial location of Fermi energy (before the contact-molecule junction is made) was more than 500 meV away from the HOMO region ($E_{f_0} < -4$ eV). The maximum, minimum, and average conductance values for the 3'-$G_3A_9G_3$-5' strand are reported in Figure 6.18. These results are at least qualitatively consistent with the conductance measurements reported in the literature [182], which show that $G(E_f^{\text{Re-Aq}}) > G(E_f^{\text{DNA}}) > G(E_f^{\text{Ox-Aq}})$; however, the structure in the reference is for a different Aq-based strand. An example graph is plotted in Figure 8a for $dN = 0.2$ eV and $E_{f_0} = -3.5$ eV values showing $G(E_f^{\text{Re-Aq}}) > G(E_f^{\text{DNA}}) > G(E_f^{\text{Ox-Aq}})$.

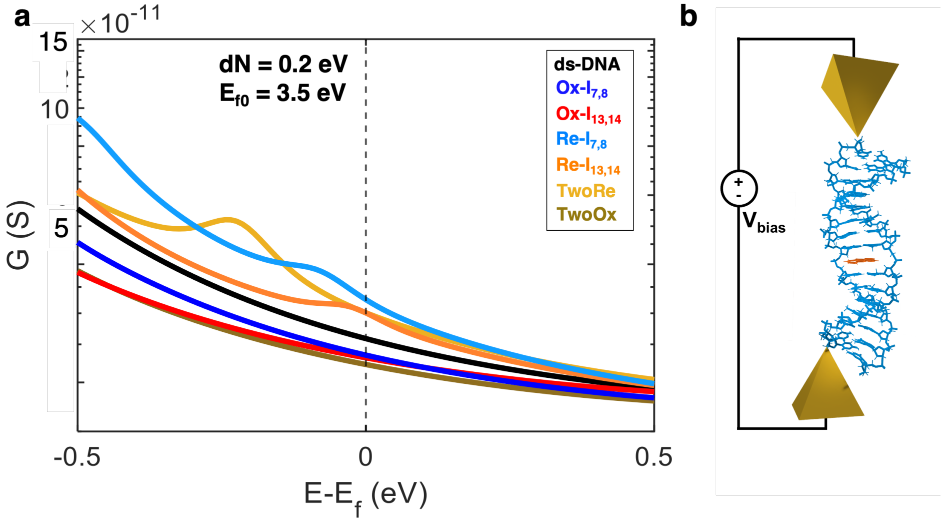


Figure 6.16 **a,** Example of conductance plot as a function of Fermi energy for ds-DNA intercalation cases for a chosen $dN = 0.2$ eV and $E_{f_0} = -3.5$ eV values, where E_f is calculated using Eq. 2. **b,** two-terminal measurement scenario.

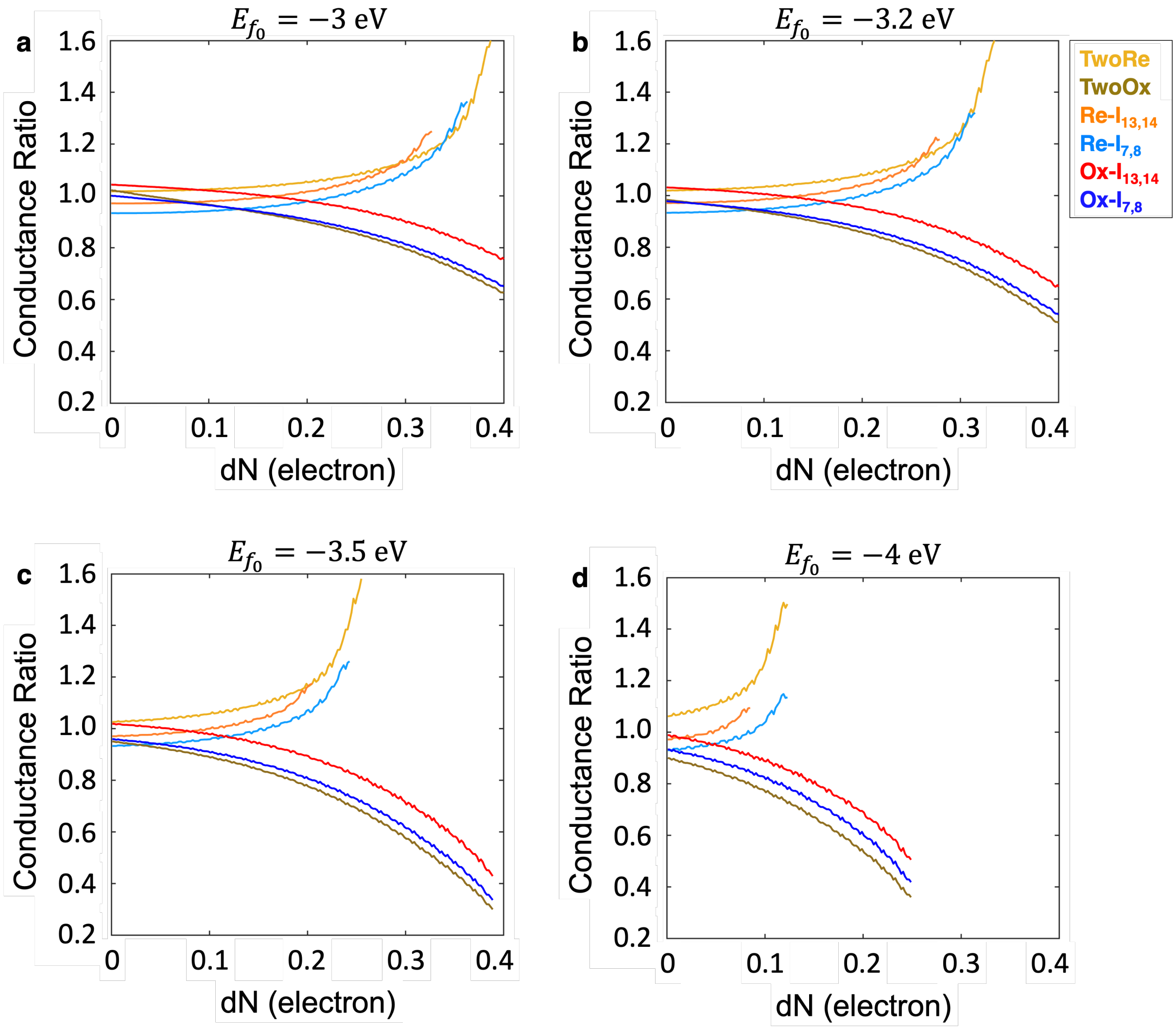


Figure 6.17 Conductance ratios for the ds-DNA structure (3'-$G_3A_9G_3$-5') as a function of partial charge transfer at different initial E_{f_0}. The cutoff seen in the ratios is due to dN causing the maximum allowable shift to E_{f_0} (i.e. E_f = HOMO). The conductance is extracted at each E_f value based on equation (6.4) and compared with the conductance of ds-DNA without intercalator. The cut-off seen in some curves near 0.25 electron (Re-Aq cases) are because E_f has reached their respective HOMO which is set as the cut-off for incrementing dN.

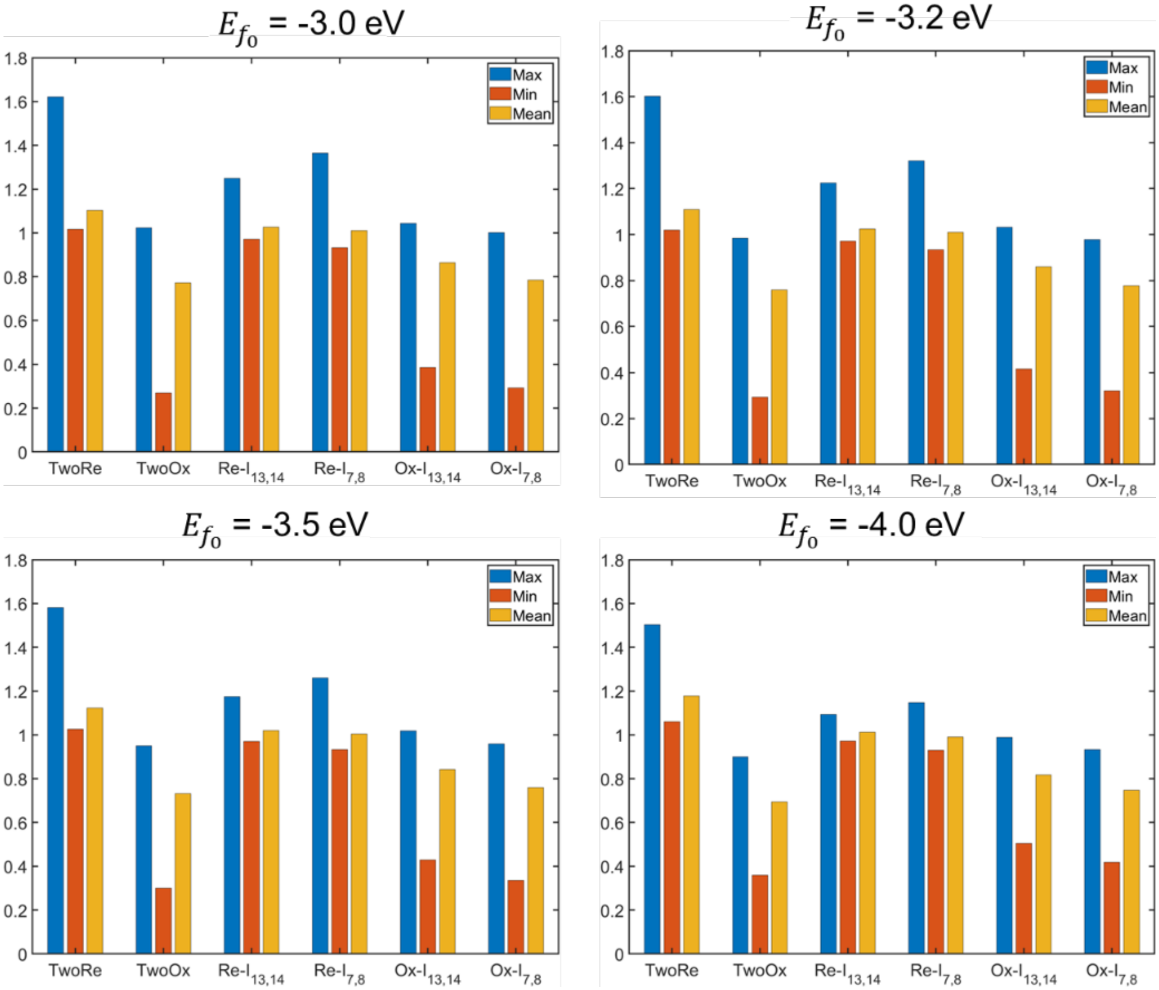


Figure 6.18 The maximum, minimum, and average conductance for the 3’-G_3A9G_3-5’ strand values are taken for dN range going from 0.001 electron to 0.4 electron, starting from different E_{f_0}.

To summarize, the location of E_f for the DNA does impact the conductance ratios between bare DNA and DNA + Aq. The present analysis shows that E_f of Re-Aq intercalation cases are closer to the HOMO region than E_f of bare DNA. As a result, the conductance can be as high as 1.6 times for DNA + TwoRe case. In contrast, we find E_f to be farther away from HOMO in Ox-Aq cases, and the conductance are lower than bare DNA. These results indicates that there are two mechanisms are in play in increasing the conductance of Re-Aq intercalated DNA: 1) a Fermi energy shift, and 2) induced energy levels above the HOMO to help traverse the AT region.

6.9 EFFECT OF AT REGION LENGTH

In this chapter, we study a strand with a shorter AT-region to investigate the effect of AT region length on the Aq-induced energy levels and the conductance. Here, we simulate the 3’-GGGAAAGGG-5’ strand, with Aq intercalating at $I_{5,6}$, $I_{7,8}$, and at both locations simultaneously (two-Aq), as shown in Figure 6.19. Transport calculations in Figure 6.20 show that the increase in transmission at the Re-Aq induced levels (-4.75 eV < Energy < -4.8 eV) is now higher compared to the longer AT case (Figure 6.7). Re-$I_{5,6}$, has two transmission peaks with more than 3x increase and the two-Aq case having a peak of more than 5x higher transmission than the bare-DNA. The energy of the induced level shifts by 50 meV for the Re-Aq when going from Re-$I_{5,6}$, to Re-$I_{7,8}$, and the transmission changes by 30%. Whereas the variation in the Ox-Aq cases is 120 meV and the transmission changes by less than 2.5%.

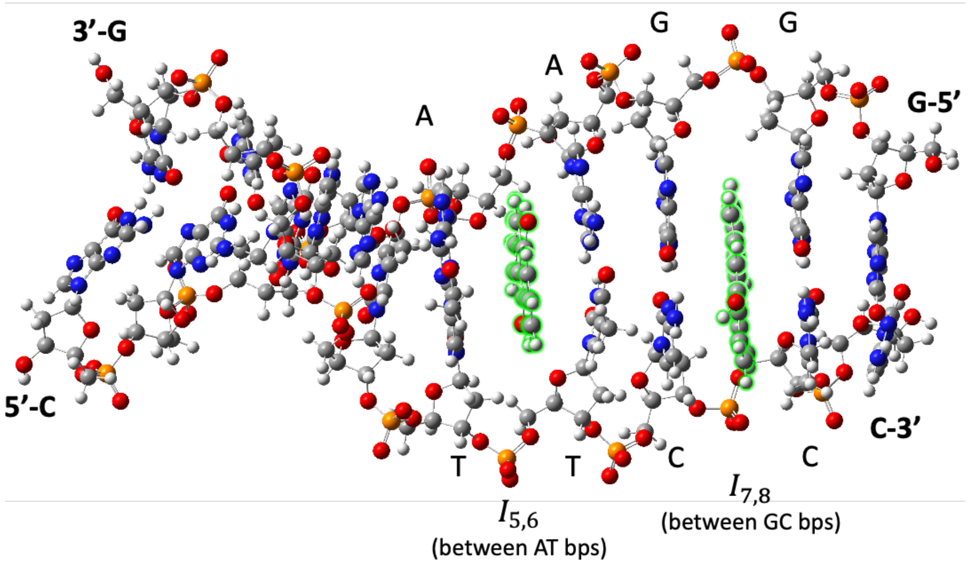


Figure 6.19 the 3’-GGGAAAGGG-5’ strand, with Aq (in green) intercalating at locations $\mathrm{I}_{5,6}$ and $\mathrm{I}_{7,8}$.

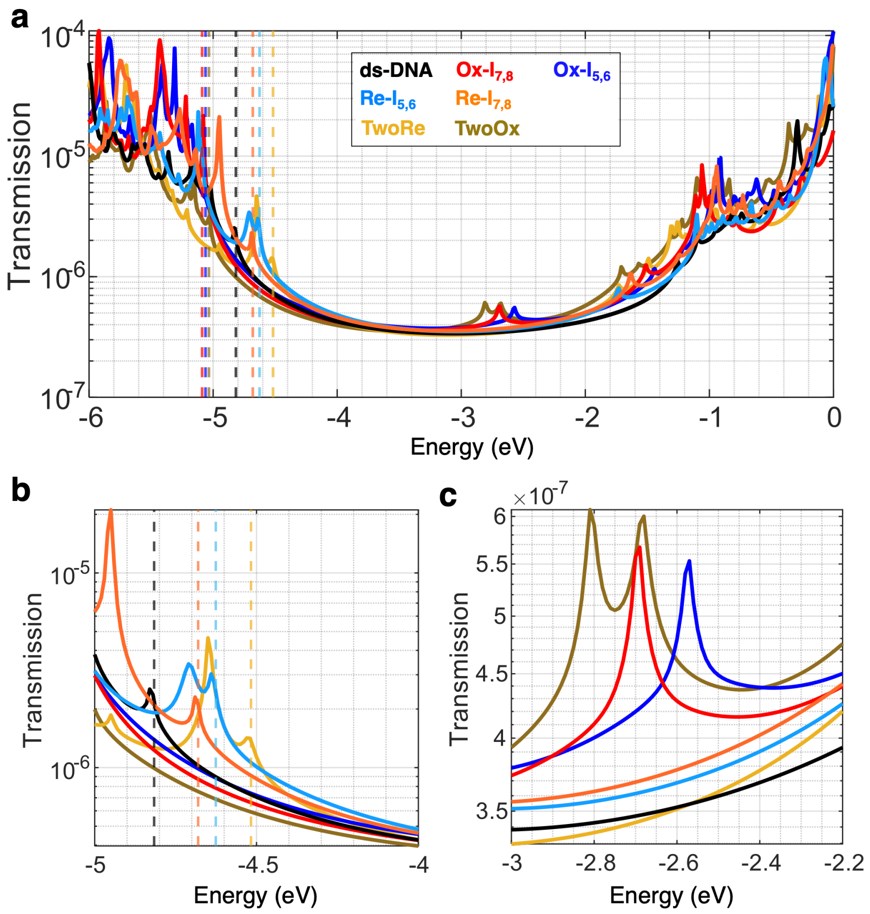


Figure 6.20 **a,** Transmission plot for the 3’-$G_3A_3G_3$-5’ strand. **b** and **c**, show the transmission at the Aq-induced peaks. Re-Aq induced levels increase the transmission by ~3 times compared to DNA at the same energy. Ox-Aq induced levels increase transmission in the bandgap by 56-77% compared to bare DNA, however, the induced levels are unoccupied levels deep in the bandgap region and far away from the expected location of Fermi energy if no gate electrode was applied.

To further understand these results, we look at the contour DOS plots shown in Figure 6.21. We notice that the Re-Aq in $I_{5,6}$ location yields a shift in the HOMO of the DNA on the 3’-G_2 region (see the red arrows in Figure 6.21). The delocalized HOMO in that region is now at a higher energy level (-4.71 eV) as opposed to the original structure (without intercalation) where the HOMO is at the G_3-5’ end of the strand. The Re-$I_{5,6}$ induced level is now closer to the HOMO region of the ds-DNA, lowering the energy separation from 500 meV in 3’-$G_3A_9G_3$-5’ (see Figure

6.4) to less than 70 meV in 3’-$G_3A_3G_3$-5’. We note that this effect is not present in the Re-$I_{7,8}$ of $G_3A_3G_3$ and can be seen at the TwoRe case, where $I_{5,6}$ is present simultaneously with $I_{7,8}$.

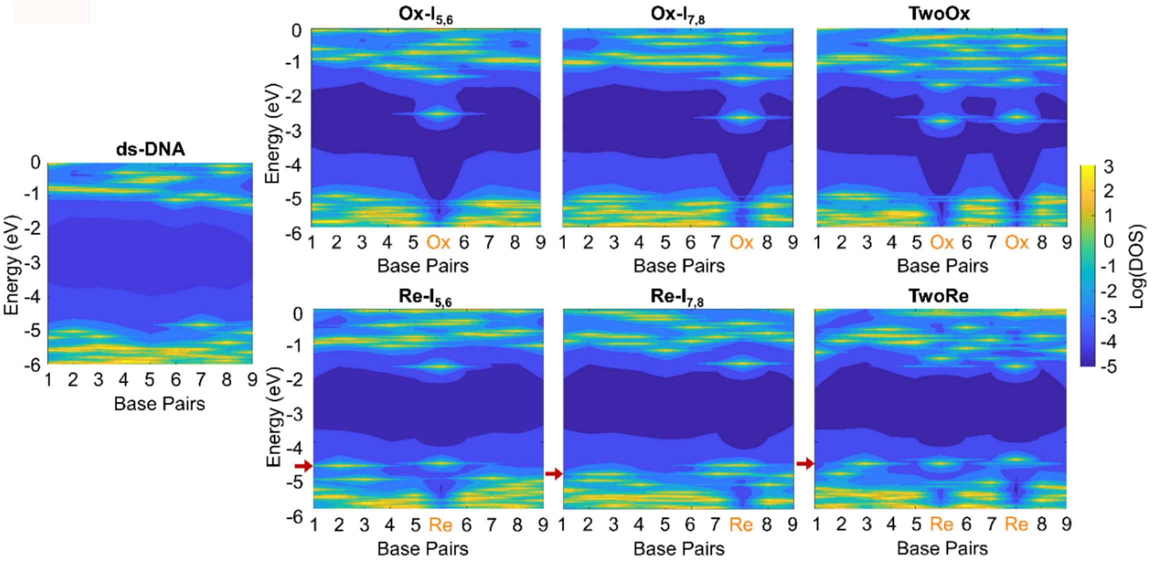


Figure 6.21 Contour DOS plots for shorter AT region. The induced levels are closer to the molecular orbitals of the DNA than the $G_3A_9G_3$ case (see Figure 6.8), with Re-Aq in $I_{5,6}$ causing a noticeable change in HOMO of the DNA (shown with red arrow) on the 3’-GGG side, shifting the delocalized HOMO to be 70 meV from the Aq-induced level. This effect can be seen in the two-Aqs case, but not in the $I_{7,8}$where Aq is on the other side of the strand (5’ end).

To investigate this HOMO shift occurring near the 3’-end, we used X3DNA [202] to analyze the structural properties of the shorter AT structures (in which the HOMO is shifted to higher energy for $I_{5,6}$). Figure 6.22 shows the overlap area between adjacent bases for $I_{5,6}$, $I_{7,8}$, and ds-DNA. The overlap area is an indicator of the electronic coupling strength, which is directly related to the transmission. We notice that both $I_{7,8}$ and ds-DNA have similar trends, whereas $I_{5,6}$ case displays an increase in overlap between the neighboring bases at the 3’-end, C_2C_3 in Figure 6.22a, G_1G_2 and G_3A_1 as seen in Figure 6.22b.

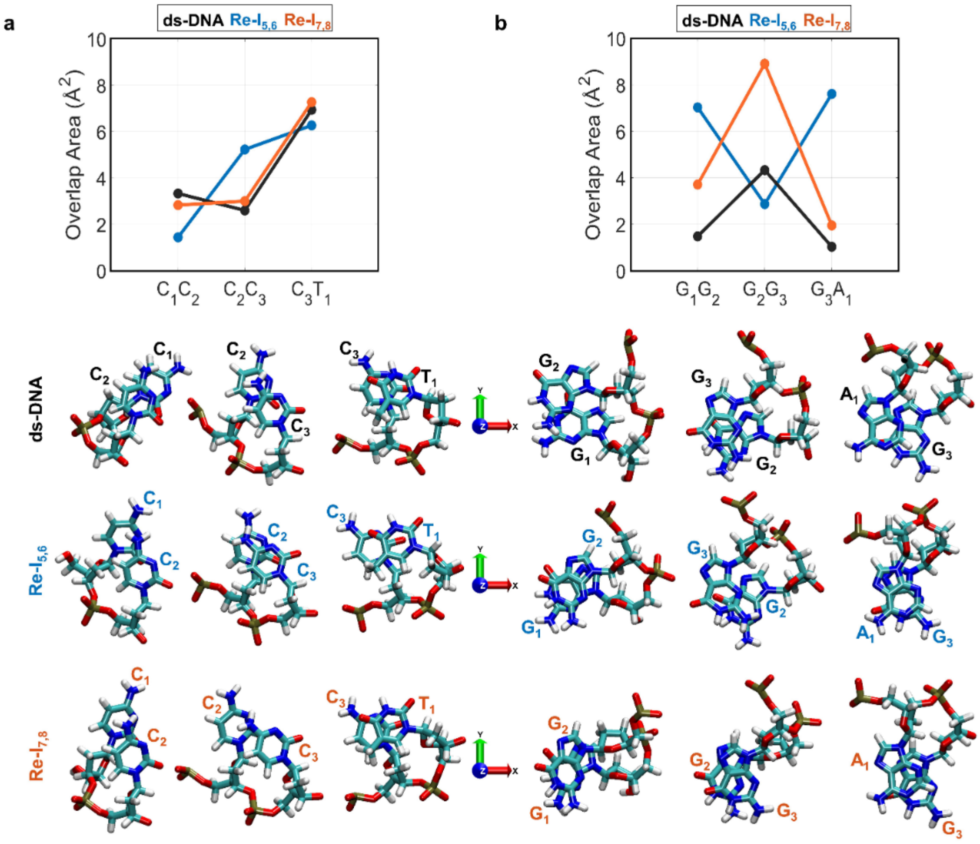


Figure 6.22. a, Overlap area comparison for 5’-CCT and position of the corresponding bases, b, Overlap area comparison for 3’-GGA and position of the corresponding bases.

We also plotted the HOMO coupling for the neighboring bases, further proving that the $I_{5,6}$ causes enhanced electronic coupling (interactions) between the bases of the 3’-end (Figure 6.23). These results show that the location of intercalation does have a unique impact on the structural properties of the shorter strand, hence increasing the coupling between the neighboring bases on the 3’-end. According to this analysis, we expect that the location of intercalation causes the HOMO of the DNA base to shift to a higher energy state, resulting in a near resonant tunneling with the Aq-induced level, yielding higher transmission (conductance) peak.

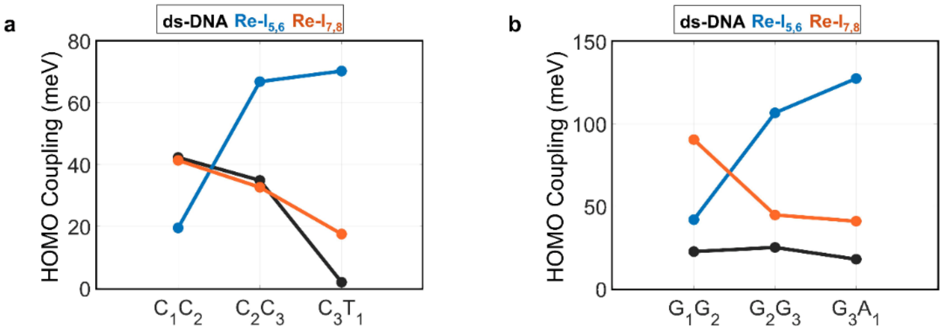


Figure 6.23. Electronic coupling parameters extracted from the Hamiltonian (off-diagonal elements) at **a**, 5’-CCC and **b**, 3’-GGG. The trends show that $I_{5,6}$ displays higher coupling between the HOMO of the bases near the Aq location.

The Aq in $I_{7,8}$ location in the shorter strand ($G_3A_3G_3$) is equivalent to $I_{13,14}$ in the longer strand ($G_3A_9G_3$). In the shorter strand, Aq-induced level is now 250 meV away from the HOMO of the DNA as opposed to the 650 meV in $G_3A_9G_3$ (Figure 6.24).

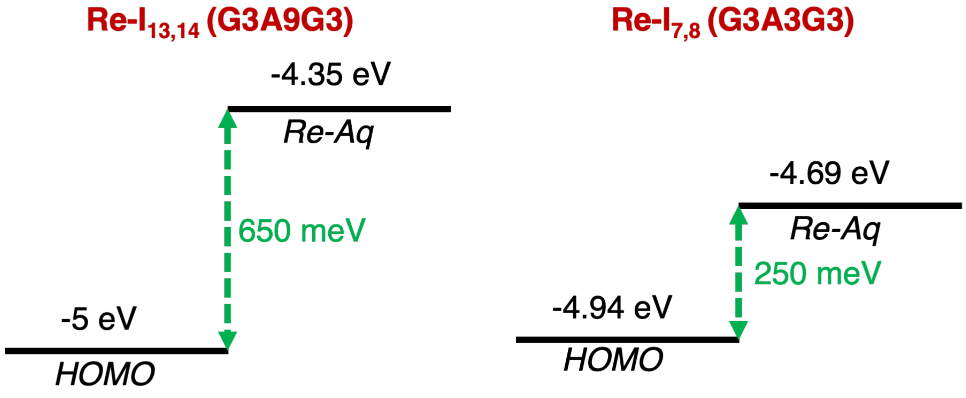


Figure 6.24 Energy separation between Aq-induced level and nearest HOMO of the DNA. **Left** Longer AT region case. **Right** Shorter AT region case, showing the Aq having a closer HOMO.

The structural differences caused by decreasing the length of the helix from 15 base pairs to 9 base pairs play a role in this variation. We also used X3DNA to analyze the overlap area of $I_{7,8}$ ($G_3A_3G_3$) and $I_{13,14}$ ($G_3A_9G_3$) bases. Figure 6.25a shows the overlap area between adjacent bases to the Aq, corresponding to A_6G_7 ($A_{12}G_{13}$ for $G_3A_9G_3$) and G_8G_9 ($G_{14}G_{15}$ for $G_3A_9G_3$). Here, we notice that for $I_{7,8}$, the overlap is ~twice as much as the $I_{13,14}$ case. This is further proven by the

HOMO coupling parameters extracted from the Hamiltonian (Figure 6.25b**)**. These observations demonstrate that for the shorter strand, the overlap (and electronic coupling) between the bases is stronger than the longer strand, thus, it creates a better path for electrons to flow along the DNA.

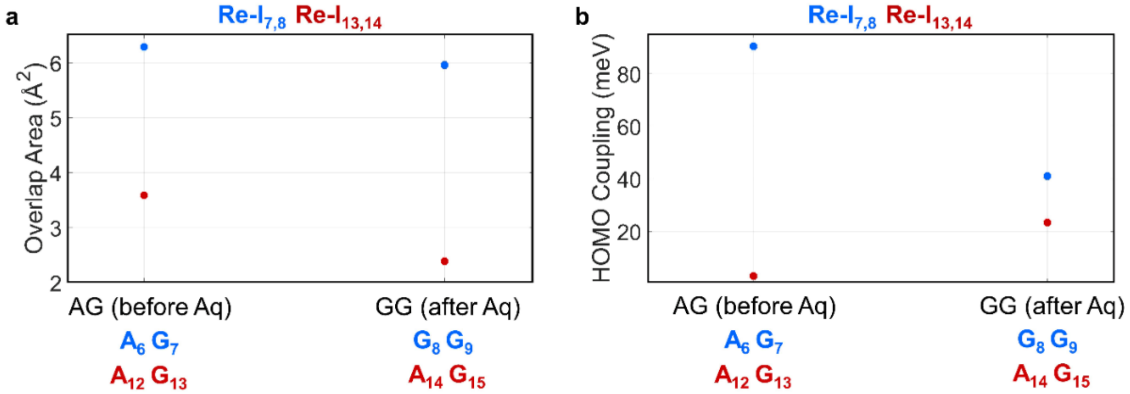


Figure 6.25. **a,** Overlap area comparison for $I_{7,8}$ ($G_3A_3G_3$) (blue) and $I_{13,14}$ ($G_3A_9G_3$) (red). **b,** HOMO coupling extracted from the Hamiltonian calculated with DFT. The adenine-guanine coupling is higher for the shorter structure.

We also carried out the Fermi energy location analysis: 1) Three-terminal scenario where E_f is changed with a gate, and 2) Two-Terminal scenario where we estimate E_f location based on partial charge transfer. For the first scenario, Figure 6.26 shows that the conductance for E_f at HOMO (or at a few hundred meV higher energies), the Ox-Aq cases are higher than bare DNA. Furthermore, Re-$I_{5,6}$ is only marginally greater than the bare DNA at HOMO while it decreases at higher energies, with the remaining two Re-Aq cases displaying lower conductance in general. Again, these findings are similar to the longer strand cases, which is due to the localization of the HOMO at the Re-Aq, while HOMO of the oxidized cases is primarily localized on the DNA bases.

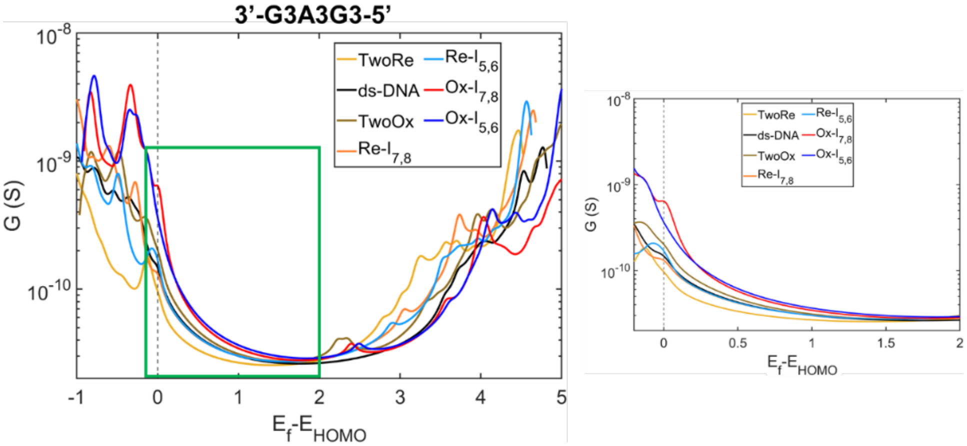


Figure 6.26. Conductance plot as a function of Fermi energy for 3’-G3A3G3-5’ intercalation cases with all HOMOs aligned (dashed line). Inset shows a smaller energy window focused on the HOMO vicinity.

As for the second scenario, the calculated dE_i/dN values yield the same trend as the 3’-$G_3A_9G_3$-5’ structure (Table 6-2). By performing the same conductance ratio analysis, we get a maximum conductance increase of 2.2 for TwoRe, 4.0 for Re-$I_{5,6}$, and 3.0 for Re-$I_{7,8}$, with an overall average ratio greater than 1 (Figure 6.27). The ratio drops below 1.0 for different dN values depending on the intercalation location for the oxidized cases. Figure 6.27a shows the conductance ratios for the E_{f_0} = -3.5 eV case. Here, TwoOx displays ratio less than 1.0 at $dN > 0$, Ox-$I_{7,8}$ drops below 1.0 at $dN > 0.1$ electron, and Ox-$I_{5,6}$ drops below 1 at $dN > 0.2$ electron. These results confirm that for the shorter AT-region strand, location of intercalation, number of intercalations, and partial charge transferred have an important effect on the conductance of the DNA. The general trend, however, is still G($E_f{}^{\text{Re-Aq}}$) > G($E_f{}^{\text{DNA}}$) > G($E_f{}^{\text{Ox-Aq}}$).

Table 6-2. Average rate of change in energy with respect to number of electrons.

Molecule	dE_i/dN **(eV/electron)**	**Change**
TwoOx	3.62	-0.30
Ox-$I_{7,8}$	3.80	-0.12
Ox-$I_{5,6}$	3.84	-0.08
ds-DNA	3.92	0
TwoRe	5.10	1.18
Re-$I_{5,6}$	5.61	1.69
Re-$I_{7,8}$	5.67	1.75

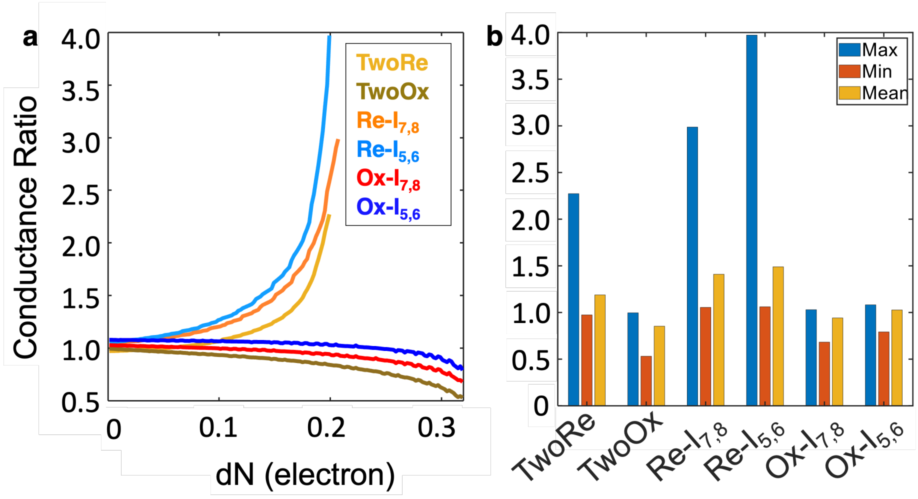


Figure 6.27 Conductance ratio of ds-DNA intercalation cases (3'- $G_3A_3G_3$-5'). **a,** Conductance ratio as a function of partial charge transfer for E_{f_0} = -3.5 eV. The conductance is extracted at each E_f value based on equation (6.4) and compared with the conductance of ds-DNA without intercalator. The cut-off seen in some curves near 0.2 electron (Re-Aq cases) are because E_f has reached their respective HOMO which is set as the cutoff for incrementing dN. **b,** The maximum, minimum, and average values are taken for dN range going from 0.001 electron to 0.31 electron.

In summary, the reduced state of intercalator increases the conductance more significantly as the AT region length decreases. In the shorter case, the intercalator-induced energy levels have lower energy separation from the HOMO level. Therefore, they help increase the coupling between the GCs on the two sides of the DNA. This finding is consistent for any amount of partial charge transfer. The Ox-Aq causes the DNA conductance to decrease, and the decrease is more sensitive to the amount of charge transfer and the intercalation location compared to the Re-Aq.

6.10 SUMMARY

In this chapter, we investigated the effect of Aq-intercalation on the conductance of nucleic acids, ds-DNA and DNA:RNA hybrid. We find the intercalation to yield similar effects on the DNA:RNA hybrid system. Our results show that although Aq induces energy levels in the bandgap of the ds-DNA, this is not like the conventional doping in semiconductors. The added levels for Re-Aq (Ox-Aq) near HOMO (LUMO) are occupied (empty). Therefore, the electrons must travel through these induced energy levels to traverse the structure, as they do not significantly alter the number of charge carriers in the HOMO and LUMO bands in the system. It would be a good scientific problem to study suitable molecular species that can effectively dope DNA to create electrons in the conduction band and holes in the valence band. In addition to the induced energy levels in the bandgap, the contact's Fermi energy with respect to the molecule modifies the conductance. The Fermi energy analysis shows that Aq can either increase or decrease DNA conductance based on its redox state. The Re-Aq was assessed to have the Fermi energy closest to the HOMO region, with Ox-Aq having it the farthest away from HOMO. Hence, yielding the relation $\mathrm{G}({E_f}^{\text{Re-Aq}}) > \mathrm{G}({E_f}^{\text{DNA}}) > \mathrm{G}({E_f}^{\text{Ox-Aq}})$ with maximum conductance 1.6 times higher (TwoRe) than the bare DNA. This relation agrees with the experiment conducted in [182], where they have shown that the DNA conductance can increase or decrease based on the state of the anthraquinone. Furthermore, lowering the AT-region length from 9 base pairs to 3 increases the impact of the Re-Aq, yielding a maximum conductance rise of 2.5 times for two intercalators and more than 4.0 times increase for a single intercalator. The effect of the structural changes on the energy levels distribution of the molecular orbitals due to intercalation are more prominent in the shorter strands. This means, increasing the number of intercalators does not always yield an enhanced conductance, especially for shorter strands. Our results also show that the experimental setup may alter the measured conductance. The three-terminal case, where the gate voltage is swept, the Ox-Aq intercalation yields higher conductance than the Re-Aq intercalation case, when the relative Fermi energy locations with respect to the HOMO levels are the same. Overall, we demonstrated that one can modulate the conductance, by altering the number of intercalators, the AT-regions length, and/or the experimental setup. An interesting application for intercalation would be to utilize multiple intercalators with different redox potentials (such as using both Ethidium Bromide and Anthraquinone as intercalators) to modulate the nucleic acid conductance in a multilevel mode instead of the two levels of high and low conductance. This realization can potentially help create a multilevel molecular device.

Chapter 7. QUANTUM TRANSPORT IN DNA HETEROSTRUCTURES

7.1 INTRODUCTION

The natal function of storing and transmitting information of DNA by pairing and stacking characteristics of its bases kindles the idea that it can also carry electrical signals. DNA offers precise self-assembly and molecular recognition at the nanoscale, which can drive the fabrication of molecular devices based on quantum interference [156], [203], [204]. Therefore, it is appealing to use such features to design devices capable of processing information and signals [205]–[207]. The emergence of 2D and 3D DNA structures will also instigate ideas for the 3D integration of DNA-based electronics. Apart from this, these devices offer biocompatibility [208] and the possibility of electronics beyond lithography limits [203]. Better understanding and prediction of a given DNA molecule's electronic conductivity would further expedite the engineering of DNA-based electronic systems and sensors.

Heterostructures are basic building blocks in electronic devices, therefore DNA heterostructures are key to DNA-based nanoelectronics. The distinct electronic properties of the different DNA bases lead to the idea of nanostructured *wells*, *barriers,* and *superlattices* [156], [203], [204]. Since the AT (GC) base pair has a higher (lower) ionization potential than GC (AT), it can be treated as a 'barrier' ('well') for hole transport [156], [204]. Therefore, one can think of quantum 'wells' and 'barriers' constructed by engineering DNA sequence. We can expect DNA-based heterostructures exhibiting transmission-resonances akin to double barrier resonant tunneling diodes and superlattices built from conventional semiconductor heterostructures. Adessi *et al.* [156] and Qi *et al.* [204] reported resonant tunneling and found a length-independent maximum transmission in the purely coherent regime. Additionally, coherent effects play a key role in charge transport in DNA [209]. However, decoherence may wash out these quantum interference effects. Therefore, quantum interference phenomena in electronic barriers and wells where decoherence plays a role need to be investigated.

Concerning charge transport through DNA (hence through DNA-based heterostructures), most investigators seem to agree that the inter-base π- π coupling in DNA could provide a pathway for charge transport along the bases [14], [156], [203], [204], [210], [211]. Accordingly, B-DNA is expected to conduct better among all the DNA conformations. However, recent studies have challenged this understanding [13], [212] where higher conductance is observed in conformations other than B-DNA, despite having relatively weaker inter-base π- π coupling [14]. These studies indicate distinct and/or additional charge transport pathways in conformations other than B-DNA. Additionally, the backbone and its environment can influence DNAs' electronic behavior, such as solvent and counterions [81], [213], [214]. Therefore, it would be interesting to probe the charge transport mechanism in DNA heterostructures considering the effect of conformation, backbone, and solvent.

Against this backdrop, we present a critical study of quantum transport (where interference is significant) in electronic 'barriers' and 'wells' formed by sequence engineering in DNA. Mainly, we investigate the influence of the width of the 'barrier' and 'well'. Apart from the most common B-conformation, we also study quantum transport in the A-conformation of DNA heterostructures, as dry (relatively dehydrated) DNA exhibits A-conformation [215]. The chapter is organized as follows. We first present an overview of the system under study. Next, we present and discuss the conductance variation for ‘barriers' and 'wells'. We conclude the chapter by discussing the significance of the results and the main findings.

7.2 SYSTEM UNDER STUDY

The sequences 3’-CCCT$_N$CCC-5' *(CT$_N$C)* and 3’-TTTC$_N$TTT-5' *(TC$_N$T)* are considered as 'barrier' and 'well' sequences for hole transport, where *N* defines the width of the 'barrier' and 'well' regions, respectively. We used DFT and charge transport calculation method as reported in detail in chapter Chapter 2. We note that we included the counterions that neutralize the DNA backbone in this chapter. For B-DNA strands, we used the PCM model to include the water solvent effect. As for the A-DNA strands, we used a mixture of ethanol-water as the solvent. The dielectric constant of the ethanol-water mixture (ϵ_m) is determined by considering only the linear dependence on mole fractions in the Jouyban-Acree model [216], i.e. $\epsilon_m = \phi\epsilon_w + (1-\phi)\epsilon_{eth}$. In this equation, ϕ is the mole fraction of water in the mixture, ϵ_w is the dielectric constant of water and ϵ_{eth} is the dielectric constant of ethanol. For the 85% ethanol + 15% water mixture, we used an effective dielectric constant of 32.877. In this study, we employ the atomic partitioning scheme for the decoherence probes that treats each atom in the system as a block. We consider a uniform energy-independent decoherence rate of 10 meV at each site/block. We choose this value because it is consistent with estimates for decoherence from Parsons' quantum molecular dynamics simulations [217] and reference [81]. The contact-DNA self-energy, $\Sigma_{L(R)}$ =100 meV is added to the ends of the DNA, i.e., at 5' (3') end at the backbone-base atoms to inject and extract electrons from the left (right) contact.

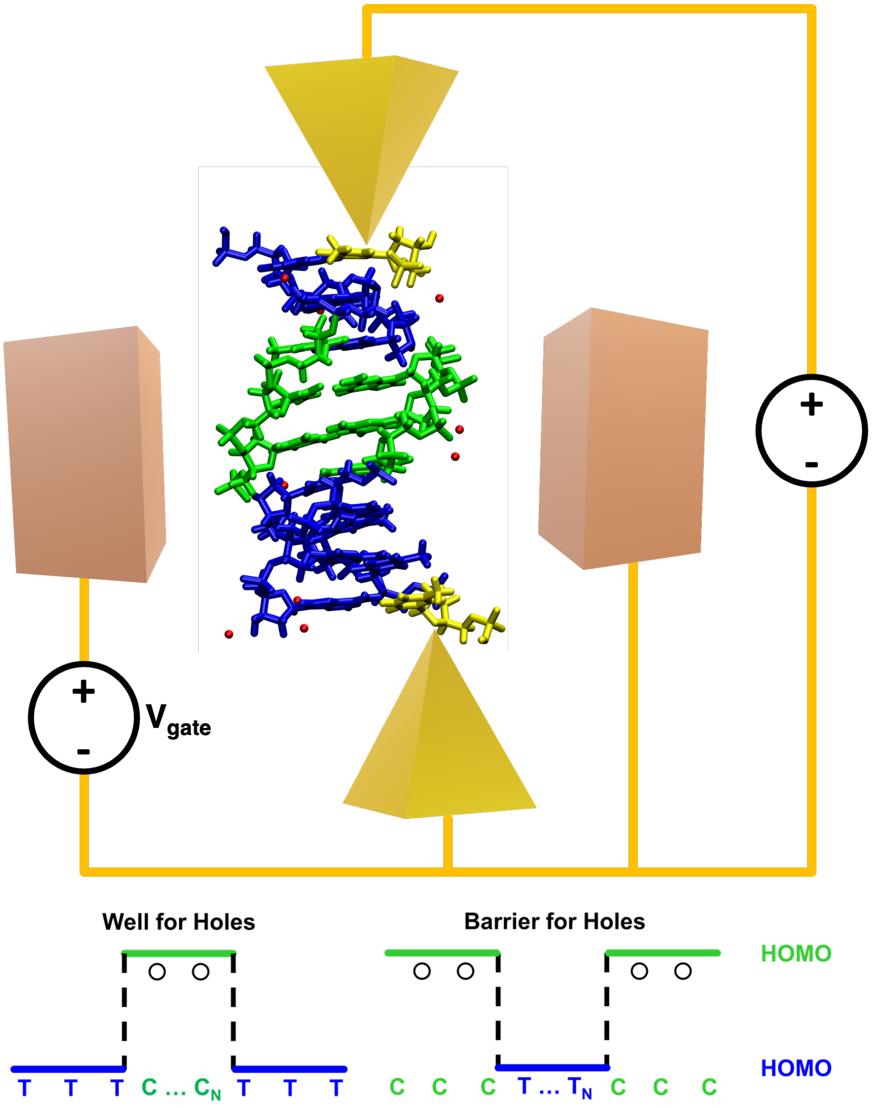


Figure 7.1 The 3-terminal setup having a gate electrode that can sweep the Fermi energy from HOMO to LUMO. The bottom figure is a schematic of how the DNA sequences can create heterostructures that act as wells or barriers.

In this study, we assumed that the Fermi energy shifts similarly for all strands with respect to their respective HOMO energy. This scenario can occur if we have a 3-terminal setup (see Figure 7.1): it consists of two electrodes across which conductance is calculated and the gate electrode that can sweep the Fermi energy from HOMO to LUMO energies.

7.3 RESULTS AND DISCUSSION

In presenting the results in Figure 7.2a-e we assumed that the Fermi energy is gated such that it is at the HOMO of each strand and a small bias is applied across the left and right electrode. The trend in conductance obtained for E_f = HOMO (Figure 7.2e) also holds when E_f = HOMO+100 meV (Figure 7.2f). Figure 7.2a-d shows conductance variation with Fermi energy, while Figure 7.2e shows conductance variation at respective HOMO energies for DNA as a function of the width of the 'barrier' and 'well' regions (N=0 to 5). We also list the HOMO energy levels in Table 7-1.

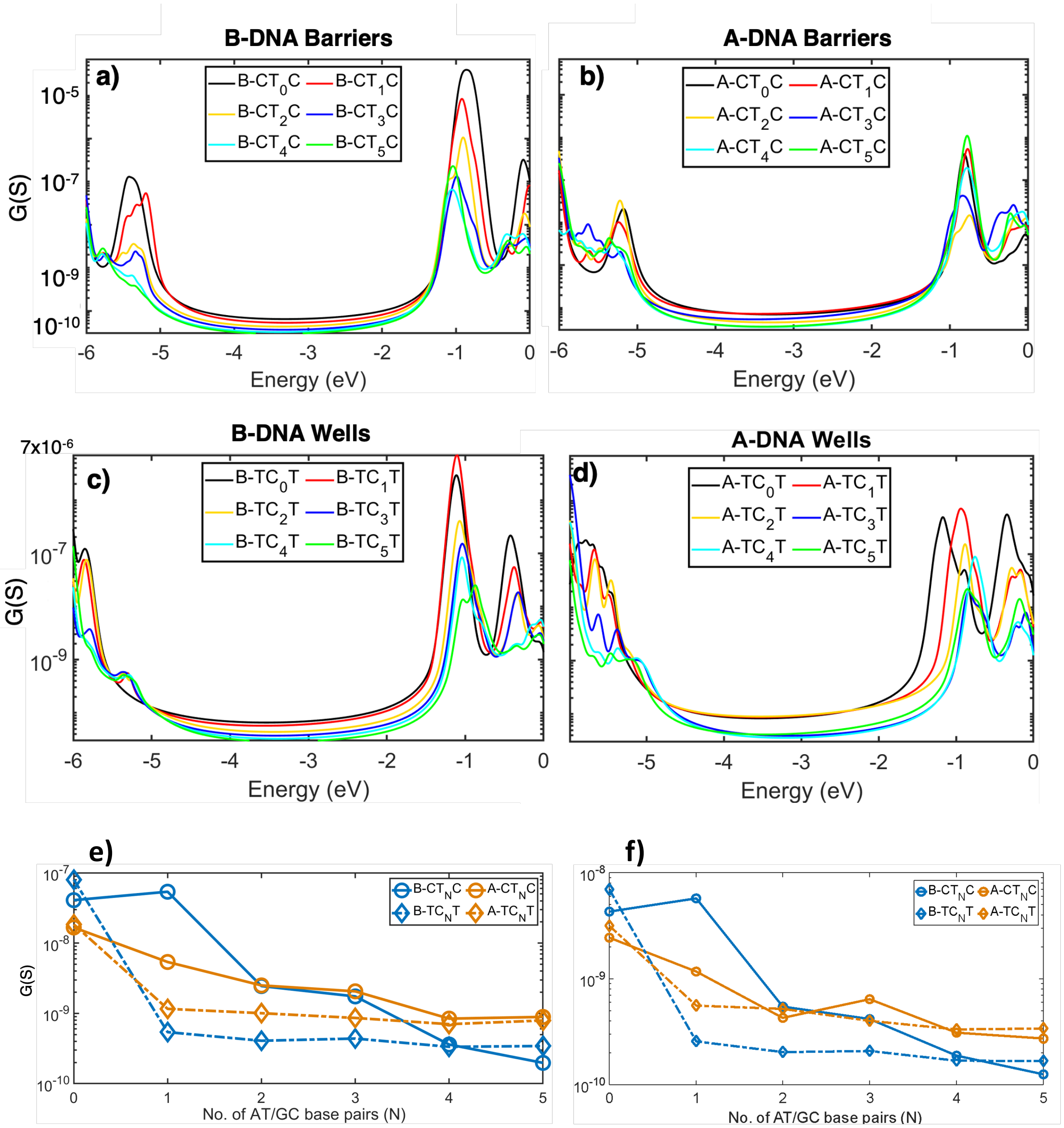


Figure 7.2 Conductance variation with Fermi Energy for a) B-DNA 'barriers', b) A-DNA 'barriers', c) B-DNA 'wells', and d) A-DNA 'wells'. e) Conductance at HOMO of B- (blue) and A- (orange) DNA 'barriers' (solid) and 'wells' (dashed) as a function of 'barrier'/'well' width *N*. Shorter B-DNAs conduct better than that of A-DNA while longer A-DNAs conduct better than B-DNAs. f) Conductance trend at HOMO+100 meV.

These DNA molecules exhibit conductance within the range of 0.2 nS to 80 nS, which lies within the broad range of reported experimental values of 10^{-13} S and 10^{-5} S [41], [63], [133], [210], [218]. A prior modeling study [204] on a DNA strand without a backbone also showed that the conductance decreases monotonically with increasing *N* for 'barriers'. However, the work did not consider the effect of DNA conformation, backbone, solvent, and counterions. Reference [81] considered the backbone but failed to include the impact of the solvent around the DNA. We also note that the experiments with increasing AT base pairs found that the conductance decreases with an increase in *N* [41], [133], [212]. In general, we also find that DNA conductance decreases with increasing *N*. However, there are important exceptions differing from prior studies. We find that for B-CT_NC, the conductance surprisingly increases from 41 nS to 54 nS as *N* changes from *N*=0 to 1. (in Figure 7.2e). That is, the inclusion of a single AT barrier increases conductance. Also, note that the longer A-DNA heterostructures conduct better than B-DNAs. The crossover point occurs upon introducing a GC base pair for 'well' sequences, as against two AT base pairs for 'barrier' sequences. We confirmed that these observations are consistent even in the vicinity of HOMO level to account for variation in the Fermi energy in experiments (see in Figure 7.2f). Next, we explain these observations one by one.

Table 7-1 HOMO energies (E_{HOMO}) of all the DNA molecules under study (in eV)

Sequence/ Conformation	**B-DNA**	**A-DNA**
CT_0C	-5.26	-5.14
CT_1C	-5.19	-5.14
CT_2C	-5.24	-5.09
CT_3C	-5.25	-5.21
CT_4C	-5.26	-5.12
CT_5C	-5.18	-5.08
TC_0T	-5.8	-5.45
TC_1T	-5.31	-5.22
TC_2T	-5.23	-5.17
TC_3T	-5.27	-5.04
TC_4T	-5.2	-5
TC_5T	-5.19	-5.11

7.3.1 *Well Structures*

First, we look at the HOMO-LUMO energy levels and the DOS along the DNA length for the 'well' sequences (Figure 7.3). For brevity, we are restricting our discussion to *N*=0, 1, 3, and 5 cases. In general, the HOMO-LUMO gap for B-DNA (A-DNA) decreases with the increasing number of GC base pairs, from 5.80 eV (5.45 eV) for *N*=0 to 5.19 eV (5.11 eV) for *N*=5, see Figure 7.3a. When we introduce the GC base pairs, additional energy levels are introduced in the HOMO band (-5.1 eV to -5.5 eV), since the HOMO in DNA mainly lies on Guanines [14], [219]–[221]. In Figure 7.3b-i, we plot DOS as a function of energy along the DNA length. Although the HOMO orbitals primarily reside on GC base pairs, i.e., in the middle of these molecules, a smaller orbital component is on the nearby AT base pairs. This localization is true for both conformations, which we can see for *N*=1, 3, and 5. However, in A-DNAs (Figure 7.3c, d and e), the HOMO orbital localization on AT base pairs is considerably higher (Figure 7.3g, h, and i). The HOMO wave functions plotted in Figure 7.4 further support this observation.

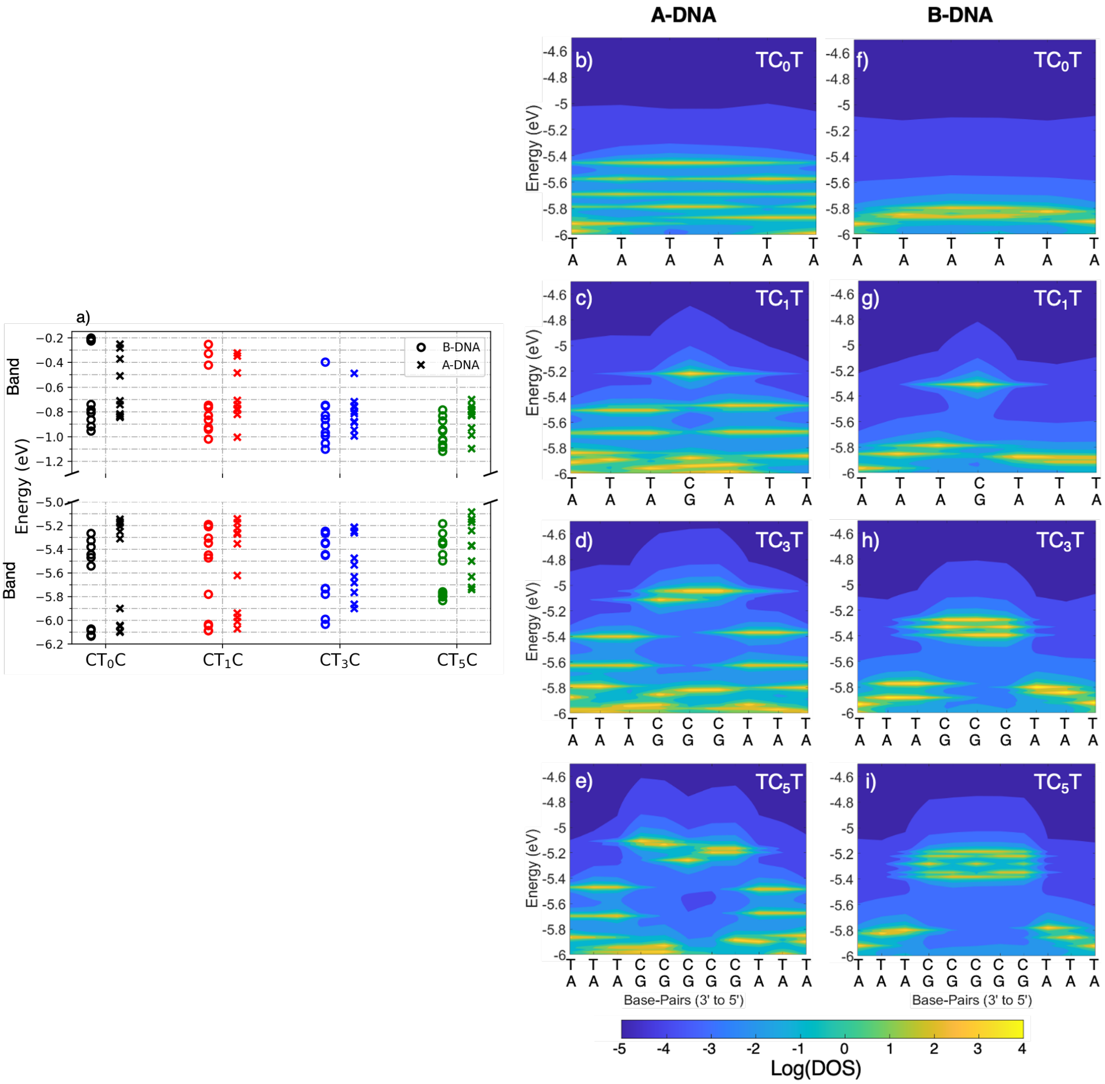


Figure 7.3 a) HOMO-9: HOMO and LUMO: LUMO+9 energy levels for A and B-DNA for 'well' sequences. DOS along the length of b-e) A-DNA and f-i) B-DNA. B-DNA HOMO orbitals are localized mainly on central GC base pairs while A-DNA HOMO orbitals are relatively more delocalized (extended to nearby AT base pairs). The left and right ends of the sequences are 3' and 5' ends respectively.

In Figure 7.4, we see that in B-DNA, the HOMO is mainly localized in the middle of the molecule, whereas in A-DNA HOMO localization extends beyond the central GC base pairs. That is, in A-DNAs HOMO localization on AT base pairs is higher than in B-DNA. For N=1 and 3, the figures clearly show HOMO for A-DNA is more delocalized than for B-DNA (compare Figure 7.4

(a) and (d)) for N=1 and Figure 7.4 (b) and (e) for N=3). For N=5, only HOMO localization is not conclusive enough. Therefore, we plot HOMO, HOMO-1, and HOMO-2 together, which shows higher delocalization for A-DNA.

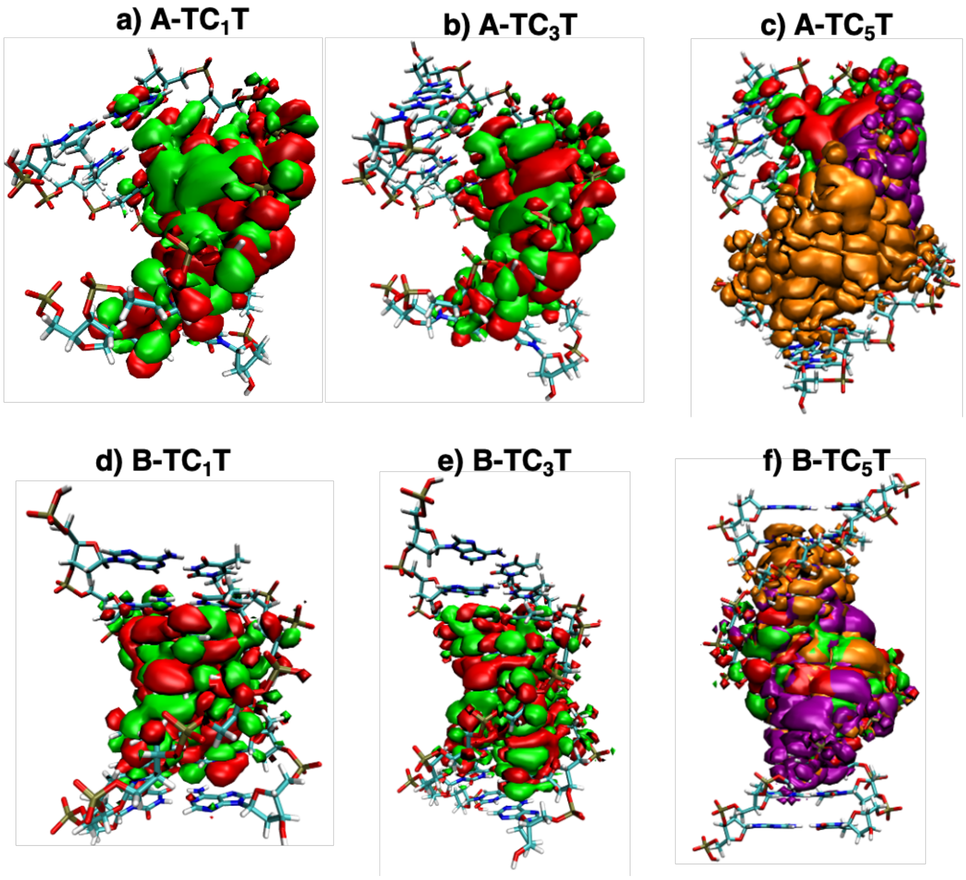


Figure 7.4 HOMO distribution for ‘well’ sequences, a-c) A-DNA and d-f) B-DNA (ISO value = $2\text{x}10^{-4}$. For N=5, i.e., c) and f) additionally HOMO-1 (orange) and HOMO-2 (purple) are also plotted. B-DNA HOMO is mainly localized in the middle of the molecule when compared to the A-DNA for which HOMO extends further along the length of the molecule.

We also explore the HOMO 'band' of 'well' sequences by plotting the decoherent transmission T_{eff} variation in the HOMO band in Figure 7.5. The HOMO 'band' of ‘well' structures shows that the number of transmission peaks is consistent with the number of GC base-pairs between the AT 'barriers'. As the 'well' width increases, the overall structure length increases, increasing scattering and lowering the transmission (conductance) by small amounts. The small amount of variation indicates that the quantum interference phenomena of resonant tunneling remain effective. Another interesting observation is that B-DNA displays a relatively uniform distribution of transmission peaks, whereas A-DNA looks disordered (Figure 7.5a,b). The transmission peaks are better resolved in coherent transmission plots as presented in Figure 7.5c,d, where the number of transmission resonance peaks are due to the guanines. The maximum transmission probability through a system with symmetric barriers is unity. However, the barriers created by the thymine's have an intrinsic asymmetry, as these structures are not geometrically identical due to rotation of bases along the DNA helix.

While it is tempting to think that the transmission spectrum for A-DNA is due to decoherence, we believe that the primary reason is its structure, which makes the energy levels nonuniform when compared to the B-DNA case. We can also see from Figure 7.3a) that the HOMO levels are energetically more closely spaced for A-DNA than B-DNA. For N=5 (3), $E_{HOMO} - E_{HOMO-2}$ is 59 meV (75 meV) for A-DNA but 96 meV (120 meV) for B-DNA. These observations and relatively more delocalized HOMO orbitals support our finding that A-DNA heterostructures conduct better than B-DNA.

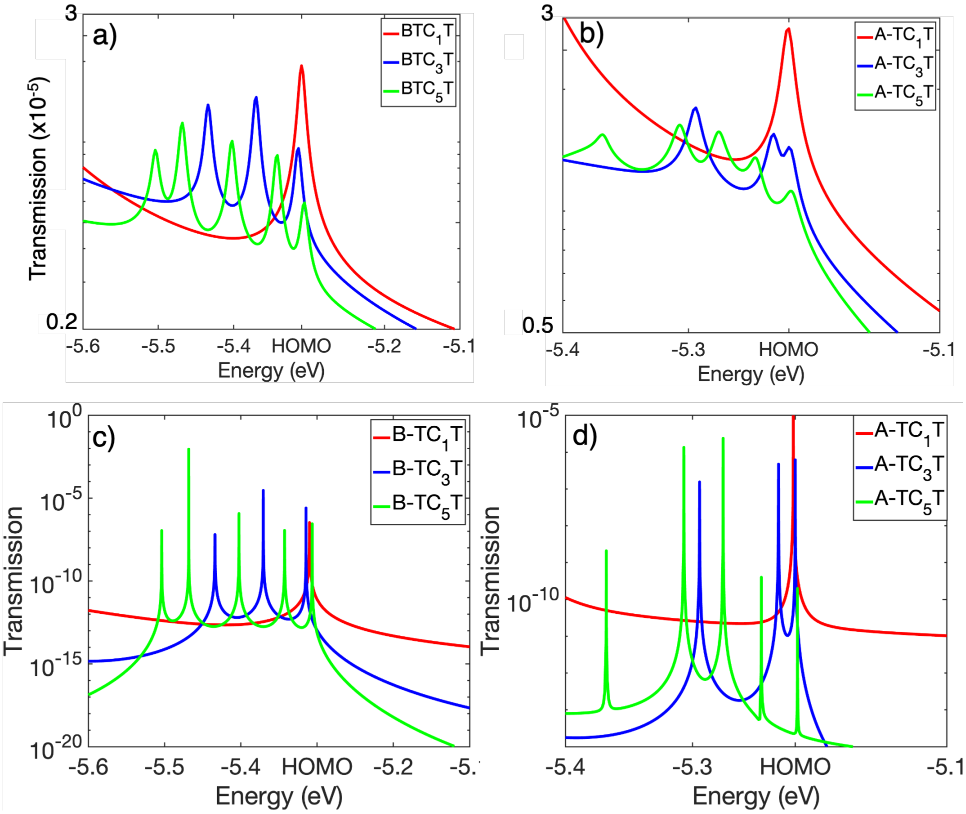


Figure 7.5 Transmission in HOMO band regions for different 'well' widths. a) and c) are B-DNA decoherent and coherent transmission, respectively. b) and d) are A-DNA decoherent and coherent transmission, respectively. HOMOs are aligned to *N=1* case.

7.3.2 *Barrier Structures*

Next, we explain the 'barrier' case. Increasing the number of AT base pairs decreases the conductance (Figure 7.2e) since AT base pairs function as a 'barrier' for holes. However, for the B-DNA, an exception arises at *N=1*, which we will address shortly. Figure 7.6a shows that the HOMO-LUMO gap decreases as *N* changes from 0 to 5. In contrast to the 'well' case, Figure 7.6a shows that when we introduce the AT base pairs, a corresponding number of energy levels are introduced deeper in the HOMO band (between -5.7eV to -5.9eV, also see Figure 7.6b-i).

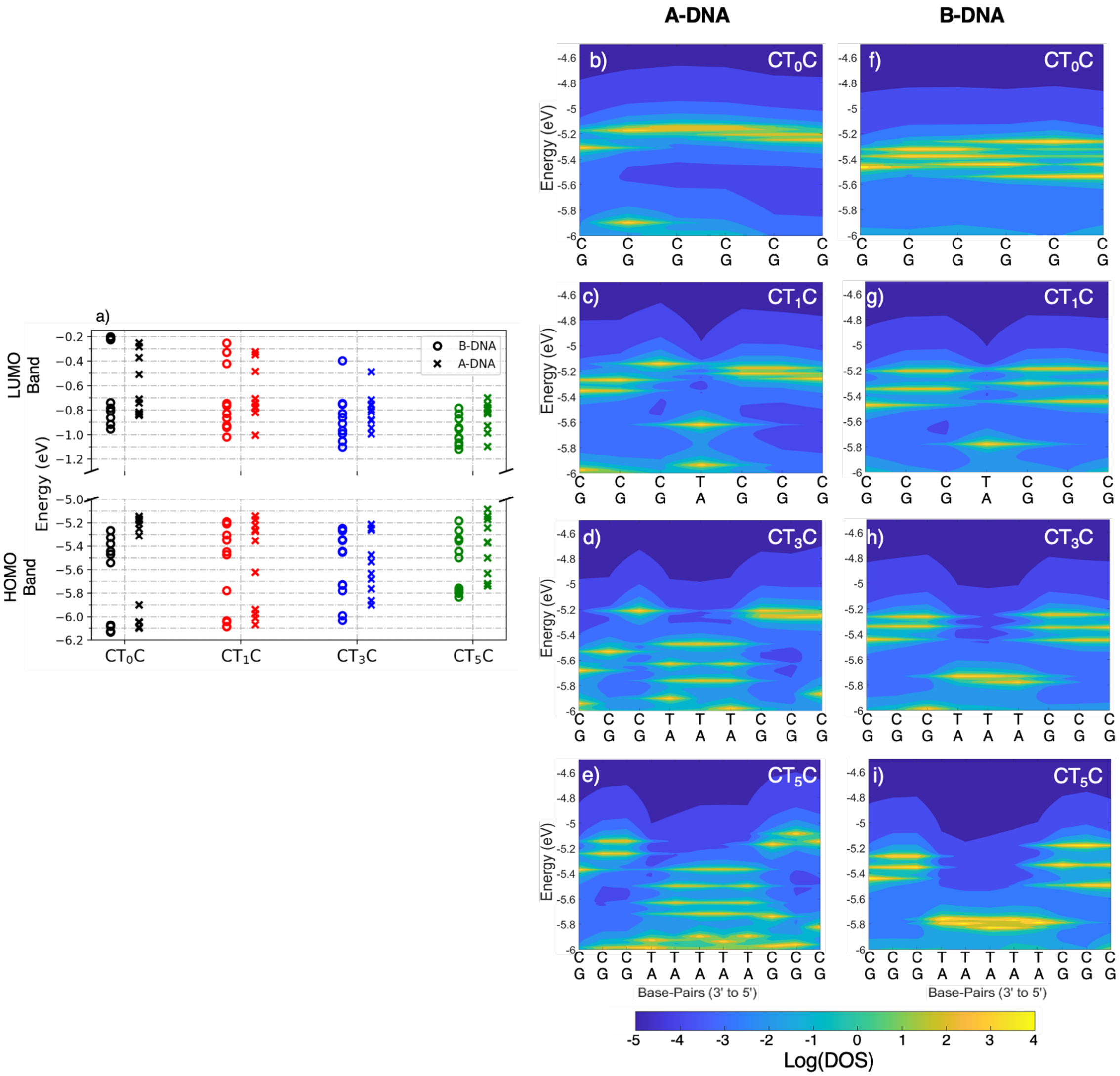


Figure 7.6 a) HOMO-9: HOMO and LUMO: LUMO+9 energy levels for A and B-DNA 'barrier' sequences. DOS along the length for b-e) A-DNA, f-i) B-DNA. The A-DNA HOMO band extends further along the length of the molecule as against B-DNA.

The B-DNA ‘barrier’ with *N=1* conducts better than *N=0*, i.e., longer DNA with a 'barrier' conducts better than shorter DNA without a 'barrier'. Upon introducing a single AT base pair, the HOMO and HOMO-1 energy levels almost become degenerate around -5.2 eV (Figure 7.6a) and the gap between them has reduced from 60 meV to 18 meV. This observation is also reflected in coherent transmission plots for ‘barriers’ in Figure 7.7a. The coherent transmission at HOMO for

N=1 (red) is smaller than that of *N=0* (black) case. Upon the introduction of a single AT base pair, the HOMO and HOMO-1 energy levels almost become degenerate, the gap between them has reduced to just 18 meV. One of the significant effects of decoherence is to broaden energy levels. Hence, applying the 10 meV decoherence helps overcome the 18 meV separation between the two levels [81]. In other words, the presence of decoherence makes these energy levels accessible for holes to traverse through. The DOS plots along the B-DNA length are shown in Figure 7.6f and Figure 7.6g, respectively for *N=0* and *1*. The DOS is more uniformly distributed in the case of *N=1* of B-DNA. Also, at the injection site, DOS is higher for *N=1* than *N=0*. For A-DNA, the relatively smaller DOS with nonuniform distribution for N=1 supports the expected smaller conductance than the *N=0* case. The DOS in the HOMO band as shown in Figure 7.6b and Figure 7.6c also substantiates this.

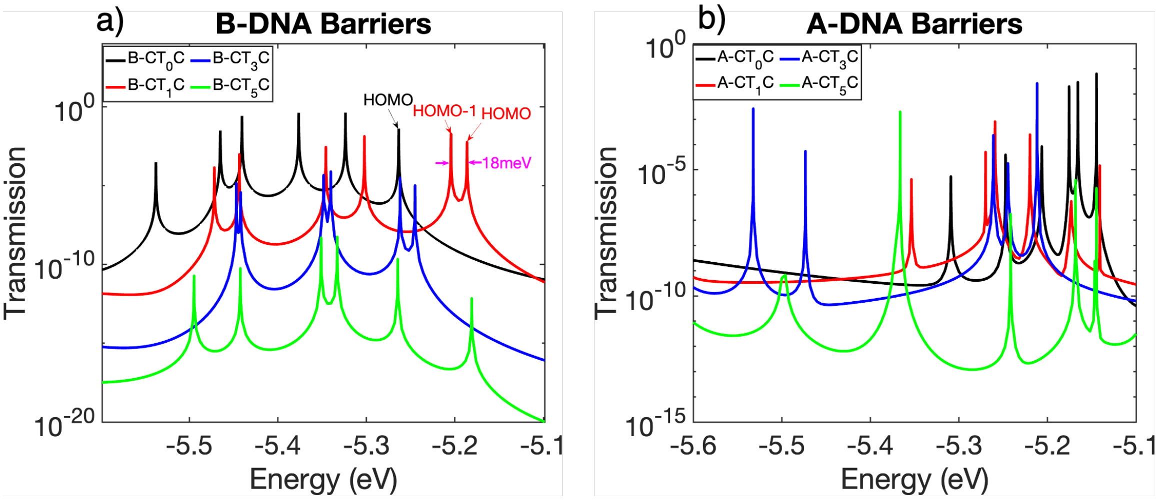


Figure 7.7 Coherent transmission for a) B-DNA barriers b) A-DNA barriers

For *N=3*, the HOMO 'band' of A-DNA (Figure 7.6d) is more delocalized than B-DNA (Figure 7.6h). For *N=5*, we see that the HOMOs are mainly localized on the GC base pairs (Figure 7.6e and i). HOMO distributions presented in Figure 7.8 also support this. The HOMO orbitals are mainly localized on GC base pairs. For *N=1*, clearly B-DNA HOMO is highly delocalized when compared to that of A-DNA. Whereas for *N=3*, HOMO of A-DNA is more delocalized. For *N=5*, HOMO localization looks similar for both the conformations. However, if we extend this further by comparing HOMO, HOMO-1 and HOMO-2 localization together, we find that A-CT_5C and B-CT_5C exhibit comparable delocalization, however these orbitals are energetically closer in A-CT_5C. Like the 'well' case, we find that HOMO levels are energetically more closely spaced in A-DNA than B-DNA. For *N=5* (3), $E_{HOMO} - E_{HOMO-2}$ is 62 meV (33 meV) for A-DNA and is 152 meV (95 meV) for B-DNA.

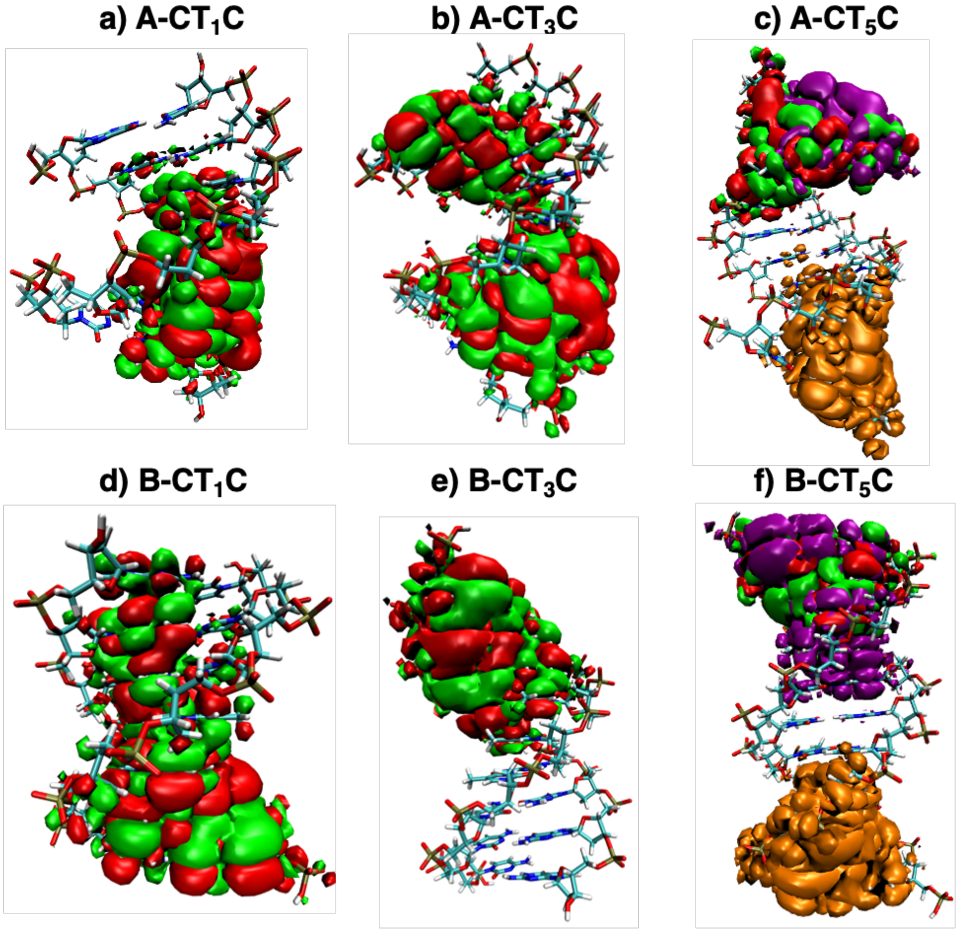


Figure 7.8 HOMO distribution for 'barrier' sequences (ISO value = 8e-4) a-c) for A-DNA, and d-f) for B-DNA. HOMO for B-CT_1C (ACT_3C) is more de-localized as compared to A-CT_1C (B-CT_3C). For N=5 i.e. c) and f) (ISO value = $2x10^{-4}$), additionally HOMO-1 (orange) and HOMO-2 (purple) are also plotted to confirm better localization in A-DNA.

7.4 SUMMARY

In this chapter, we studied quantum transport through DNA-based quantum wells and barriers using a combination of atomistic simulations and Green's function-based charge transport calculations including decoherence. We formed the DNA-based quantum well and barrier with sequences 3′-TTTC$_N$TTT-5′and 3′-CCCT$_N$CCC-5′, respectively, where N varies from 0 to 5. Overall, we find that the strands' resistance is large and varies by over 2 orders of magnitude. We showed that these heterostructures complement their semiconductor-based counterparts. Increasing the well width (N) shows a small decrease in conductance. On the other hand, the decrease in conductance with the increase in barrier width is substantial. Our model, which includes decoherence, shows that these heterostructures exhibit robust quantum interference, as seen by clear peaks in the transmission resonance. It also shows that the conductance depends significantly on barrier width. The role of DNA conformation is also investigated. We find that B-DNA's conductance decreases more sharply with barrier width than does that of A-DNA, which experiments should be able to verify. In deviation from conventional expectation, the smallest barrier (N= 1) shows a conductance higher than that of the no barrier case in B-DNA.B-DNA quantum wells have large electronic coupling between consecutive GC base pairs. As a result, B-DNA quantum wells have a uniform distribution of conductance peaks, whereas they are disordered for A-DNA. In A-DNA, we find that the HOMO levels are energetically more closely spaced, and the HOMO orbitals are spatially more delocalized. Therefore, we conclude that the energetically closer HOMO levels and the larger spatial delocalization of density of states in the conduction energy window may lead to higher conductance in A-DNA heterostructures. These two properties of A-DNA can overcome its relatively weaker $\pi-\pi$ coupling between orbitals on neighboring bases. This observation presents a new understanding of charge transport particularly in DNAs in addition to the idea of charge transport dominated by the π−π interactions of stacked bases. We attribute these properties to the structural differences between DNA conformation in otherwise identical strands. The analysis and computational results demonstrate that DNA-based hetero-structures complement the solid-state semiconductor counterparts and exhibit robust quantum interference, where the DNA conformation can play a significant role.

Chapter 8. PERPENDICULAR CHARGE TRANSPORT THROUGH DNA

8.1 INTRODUCTION

The advances in DNA Origami and sequencing techniques have prompted the study of DNA interaction with inorganic substrates. In DNA sequencing, a Scanning Tunneling Microscopy (STM) [222] technique has been developed to sequence DNA lying on a gold substrate. In this technique, the tip is brought into contact to form a single nucleobase junction as shown in Figure 8.1a. The conductance is then measured and based on the value, the nucleobase is determined. Another sequencing technique uses Scanning Tunneling Spectroscopy (STS) [33] in which they scan for the density of states by taking the first derivative of the I-V curve. In this technique, they relate the HOMO, LUMO, and bandgap information to determine the base.

DNA Origami shows a promising aspect of building complex 3D structures using bottom-up approach. However, several technological and fundamental advances are required to develop nanoelectronics devices beyond the conventional photolithography techniques. Understanding how these DNA structures interact with substrates and how electron transports through them is essential to facilitate the development of this field. Quantum tunneling current which flows both along and perpendicular to the helical axis is essential to describing DNA-based materials. For molecular electronics applications involving double strand DNA (ds-DNA), current flow along the helical axis is important. However, for the emerging sequencing applications and origami-based nanostructures, tunneling both along and perpendicular to the helical axis are crucial.

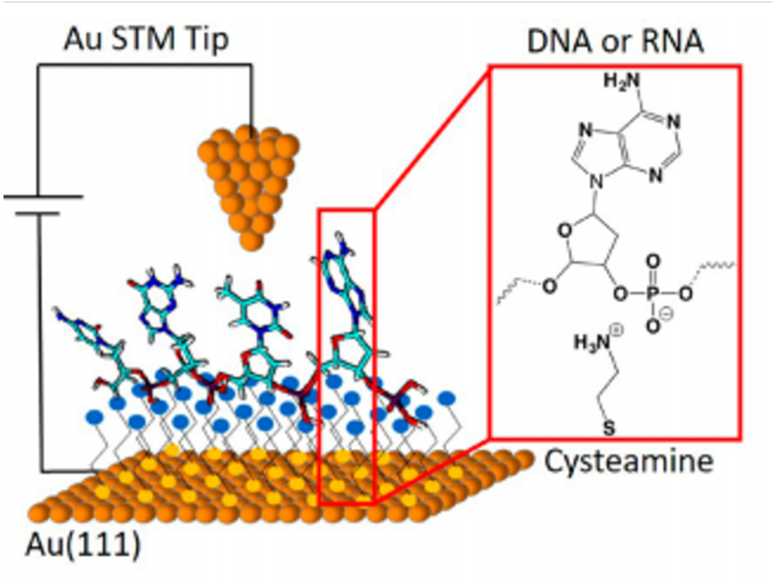


Figure 8.1 Quantum point contact single nucleotide conductance sequencing method [222]. The same setup can be used to apply the STS technique and sequence the strand [33]. The single strand is lying between gold surface and the tip, the tunneling current is measured and sequence is determined based on the reading value.

In this chapter we focus on the molecular interactions between ds-DNA and gold (111) substrate and the transmission perpendicular to the molecule. We modeled three different 9 base-pair long ds-DNAs having additional 3-base single-strand extensions on both ends on top of gold surface using molecular dynamics simulations (as shown in Figure 8.2). We then extract a representative structure from MD and remove the gold atoms to run DFT calculations on the resulting ds-DNA conformation. We use the resulting Hamiltonian for charge transport calculations to model the effect of contact location on the tunneling current. We employ the contact self-energies within our model to simulate DNA lying on a gold surface, which also forms the first electrical contact. The second electrical contact is a gold tip that is perpendicular to the substrate. We vary the contact tip location along the strand and study the electron tunneling path as it travels perpendicular to the DNA.

8.2 METHODS

We model three separate ds-DNA that are 9 base-pair long each. The sequences considered are: ss-$C_3A_3C_3$-ss, ss-$C_3T_3C_3$-ss and ss-$G_3A_3G_3$-ss as shown in Figure 8.2. For simplicity, we will refer to them as CTC, GAG, and CAC, respectively. Three adenines are attached to the ends of the strands at the 3'- and 5'- ends (without complementary base-pairs to help bind the sequences to the gold, see Figure 8.2)substrate. It is known in literature that adenines and guanines have higher affinity to bind to gold surface[223], [224].

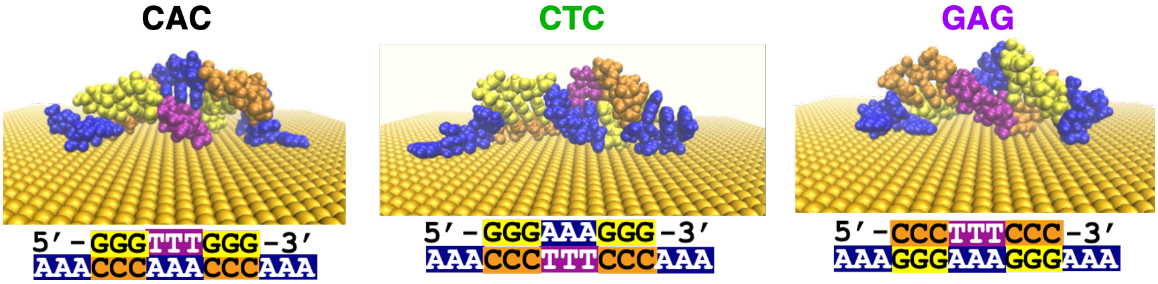


Figure 8.2 Overview of the modeled system, color represents the different nucleobases. Color codes: blue is adenine, yellow is guanine, orange is cytosine, and purple is thymine.

To investigate the differences in electronic properties due to sequence and thus conformational changes in DNA, we extracted representative structures from the >40 ns MD simulations. The gold substrate, water molecules and counterions are removed from the representative structures that come from the MD simulations because DFT calculations at this time are limited by the number of atoms in the system. The DFT calculations are carried out using the B3LYP/6-31G(d,p) basis set with the polarizable continuum model to account for solvent dielectric constant (see Methodology). We use the Green's function method with *E-indep* 10 meV decoherence to calculate the transmission. We setup the calculations to have the electrons flow perpendicular to the DNA helix, flowing from the backbone of one of the nucleotides to the substrate, via the DNA. In this calculation scheme, we only include contact self-energy ($\Sigma_{L(R)}$) to the system Hamiltonian to represent contact coupling. The top contact self-energy location is chosen to be at the backbone of the middle nucleotide of each segment as shown in Figure 8.3 and we choose three different locations for the analysis. The substrate acts as the bottom contact. We set the self-energy for the bottom contact at the DNA atoms closest to the gold substrate within 5 Å range.

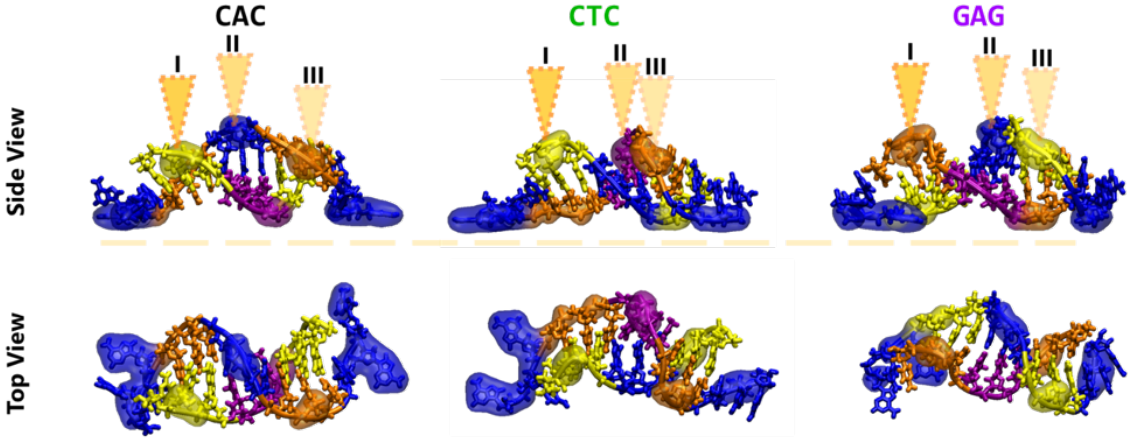


Figure 8.3 Structures used in quantum mechanical calculations. Top row shows molecules from the side view and position of gold substrate represented in dashed yellow line. Bottom row shows the molecule from the top view, gold substrate is removed for the clarity. Highlighted regions with bubbles show the contact self-energy locations chosen for the calculations. While bottom contacts fixed, the top contact location is changed and named as location I, location II and location III as shown at the top row.

8.3 RESULTS AND DISCUSSION

The sequence consists of three GC base-pairs at both the 5'- and 3'-ends, separated by three AT base-pairs in the middle. This causes the highest occupied molecular orbitals (HOMOs) of the structures to be spatially separated as shown in CAC case in Figure 8.4. We varied the sequence by flipping the order of the three base-pair segments to change which bases are interacting with the gold substrate and which bases are accessible to the top contact. In CAC and CTC cases, the guanines at the 5'-end are accessible to the tip, while at the 3'-end, the cytosines are accessible to the tip (see tip locations I and III in Figure 8.3).The middle segment has adenines (thymines) accessible to the tip for CAC (CTC). In GAG, the tip accessible bases are cytosines on the 5'-end, guanines on the 3'-end, and adenine in the middle segment as shown in Figure 8.3. We set top contact location I and location III to be at opposite sides with respect to the the 5'- and 3'-ends of the strand (see Figure 8.3). For location I, we set the contact self-energy to be at the (a) backbone atoms of the middle guanine from the 5'-end in CAC and CTC, and (b) backbone atoms of the middle cytosine from the 5'-end for GAG. As for location III, the contact self-energy is set at the (a) middle cytosine from the 3'-end for CAC and CTC, and (b) middle guanine from the 3'-end for GAG. Location II is set at the backbone atoms of the middle base in the AT segment for all strands. That is, for CAC and GAG, location II is at an adenine, and it is at a thymine for CTC.

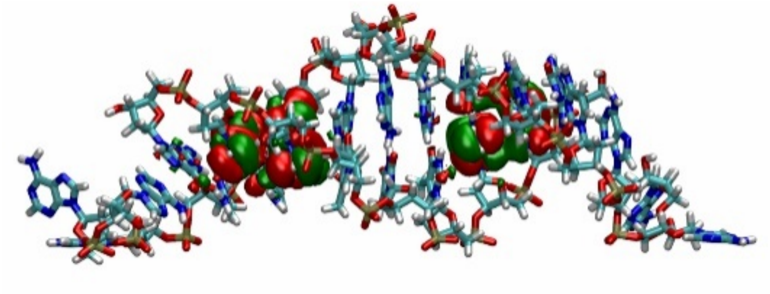


Figure 8.4 The highest and the second highest occupied molecular orbitals (HOMO and HOMO-1, respectively) isosurface plots for CAC case with ISO value = 0.005. The plot shows how the highest occupied energy levels are localized at the GCs of the two ends while the middle AT region is empty.

8.3.1 *Transmission Trends for Top Contact Location*

The transmission versus energy close to the HOMO is shown in Figure 8.5 for the three locations of the contacts considered. We notice that for all strands, the transmission versus energy is different when the tip is at either location I or III (see the darkest and brightest colors in Figure 8.5a-c) . This is because the spatial distribution of the molecular orbitals is playing a role here. For instance, HOMO is localized on the right GC base-pairs, which are located at location III, and HOMO-1 is localized on the left GC base-pairs, which are near location I (vertical dashed lines in Figure 8.5a-c represent the HOMO). Therefore, the injected charge on location I (III) primarily flows through the HOMO-1 (HOMO) energy level, yielding the different transmission profiles. The transmission profile when the tip is located at location II follows location III because: 1) the DOS is more delocalized on location III side, and 2) the closest bottom contact atoms are near location III (see Figure 8.6). The direct distance between location II and the bottom contact atoms nearest to location I and location III is about 1.7 nm and 1.3 nm, respectively. The reason for this asymmetry in the DNA-substrate interface is due to the chirality and helicity of the DNA (rotation of bases along the DNA helix). The helical structure causes major and minor grooves to form in the DNA. We notice that for the middle AT segment, the closest bottom contacts are near the minor groove (see Figure 8.3). These observations show that for each sequence, the location of the top contact, nearest molecular orbitals, and the closest bottom contact points determine the path of the electrons as they travel from the tip to the substrate through the DNA. Another outcome of this analysis is the importance of the lower occupied molecular orbitals when using this setup for sequencing DNA. For instance, the variation between location II and location III appears in lower molecular orbitals (as they have almost the same transmission values at the HOMO). Therefore, focusing on the HOMO alone is not enough to distinguish if the tip was located at an adenine or a guanine.

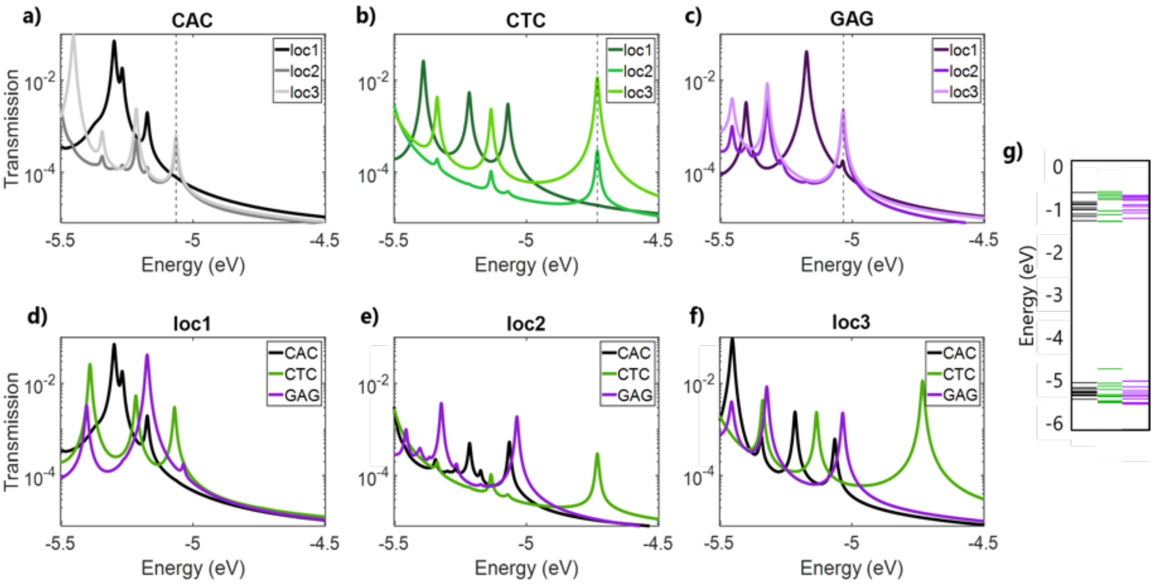


Figure 8.5 Transmission versus energy for the different ds-DNA cases. (a-c) Each case with its different top contact location. (d-f) Replotting the curves based on the contact location to view sequence effect on transmission, g) energy levels diagram, colors indicate corresponding strands.

The results obtained for location II are important because they encompass different aspects of designing DNA nanodevices: **1) Nearest neighbor impact**: location II is at an adenine/thymine, which normally does not have the HOMO localized on it. Therefore, designing the sequence to have nearby GC base-pairs helps maintain large tunneling current to flow, even if the electron was injected at an adenine/thymine. **2) Minor grooves effect**: bottom contact atoms dictate the electron path in the perpendicular charge transport. Designing the strand with the minor grooves location in mind can help control the expected path of the electron injected in the middle region.

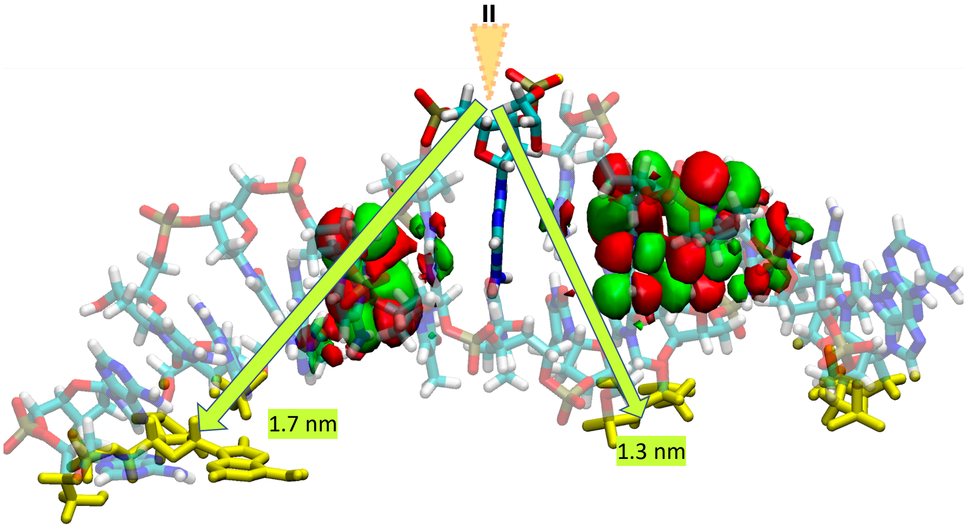


Figure 8.6 CAC strand displaying the ISO surface plot of HOMO and HOMO-1 that are spatially located on the GCs near top contact location I (left side) and III (right side), respectively. The transmission at location II follows location III due to higher delocalization of HOMO-1 and the shorter direct distance to the bottom contacts.

8.3.2 *Sequence Impact on the Transmission*

Our simulations show that changing the sequence by switching the orientation of the bases influences the resulting conformation of the strand lying on the substrate. One possible reason is due to the orientation and order of the purine (adenine and guanine) and pyrimidine (thymine and cytosine) bases at the DNA-substrate interface for each case. We know that purines consist of two carbon rings whereas pyrimidines have one carbon ring. Thus, the surface area of the bases at the interface can change per sequence. We notice that CAC has cytosine-thymine-guanine as the three segments that are at the interface (from left to right, color coded as orange-purple-yellow in Figure 8.2). CTC has cytosine-adenine-guanine (color coded orange-blue-yellow in Figure 8.2), and GAG has guanine-thymine-cytosine (color coded yellow-purple-orange in Figure 8.2) as the three segments at the interface. Therefore, CTC has more purines at the interface than CAC and GAG. Our results show that CTC is more deformed than the others, and its impact on the transmission is discussed next. We notice that changing the top contact location for CTC from location III to location II lowers the transmission at HOMO by 38 times (lighter colors in Figure 8.5b). For GAG and CAC, however, the transmission at the HOMO for location II only decays by 16% with respect to location III (Figure 8.5a,b). To understand this result, we compare between CAC and CTC structures, where only the middle AT segment is flipped. In the CAC case, the minor groove that is at the interface starts forming at the thymines of the middle segment (purple region below location II in Figure 8.3). This interaction constructs a balanced adsorption to the gold substrate by the two single-strand adenines at the ends, which are fully adsorbed onto the substrate (see Figure 8.3). The balance in adsorption makes adjacent guanines near location III have less interaction with the substrate. Thus, the lower interaction reduces the structural deformation, so the DNA maintains its π- π stacking in the middle region. This is not the case for CTC, where the minor groove starts to form at the guanines of the right segment (yellow segment below location III in Figure 8.3). Moreover, in the CTC case, only two of the left side adenines are adsorbed, which causes an additional asymmetry in the structure. We attribute the asymmetry in the DNA-substrate interface to causing an imbalance on the molecule and inducing larger deformation to the middle region. We expect that the deformation of the π- π stacking in the middle region shown in the top view of Figure 8.3 to be the main factor for decreasing the location II transmission at HOMO, as it lowers the hopping parameter between the thymine and nearby CG base pairs. We can further see the impact of the large deformation in CTC on the HOMO energy as it shifts to -4.732 eV, which is higher than the CAC by 332 meV (Figure 8.5g). This new energy can be seen as a transmission peak in Figure 8.5f.

Next, we look at the effect of switching the sides of cytosines and guanines in CAC and GAG on the resulting conformation and transmission. We notice that both CAC and GAG have the same HOMO-1 energy (-5.172 eV, see Figure 8.5d). The transmission peak for location I at this energy is 20 times higher for GAG (Figure 8.5d). The reason for this difference is that a direct tunneling between the guanine of the top contact and the adenine on the interface is possible because they are almost perpendicularly aligned (top view in Figure 8.3) with a direct distance of about 1 nm. However, for the CAC case, such alignment is not observed, and the distance is about 1.8 nm.

From these observations of sequence and contact location impact on the transmission profile, we can highlight the main guidelines for designing DNA - Gold (111) interfaces for use in nanoelectronics applications:

- **Tunneling current does not necessarily diminish if the contact location is shifted:** This gives room for the trade-off between contact location and structure stability when designing DNA origami for nanoelectronics applications (e.g., location II is similar to location III).
- **Sequence and minor groove location in ds-DNA affect structure deformation and bottom contact points**. The careful design of the sequence can help reduce the deformation and control the desired location of minor groove formation within the strand.
- **Current direction can be controlled**. Electrons can be directed to go to the right side (left side) based on where we inject the electrons (contact location) into the strand and by suitably designing nearby molecular orbitals and bottom contacts to be at the right side (left side).
- **DNA sequencing**: The accessible molecular orbitals to the tip location and the nearest neighbor effect show us that we cannot rely on the transmission near HOMO alone to distinguish between the bases. Obtaining more information by studying different permutations of a sequence and applying bias sweeps to access the lower occupied molecular orbitals can help develop the sequencing using this approach.

8.4 CONCLUSIONS AND FUTURE WORK

In this chapter, we modeled 9 base pair long ds-DNAs with different sequences on top of a gold (111) substrate. We examined charge transport properties in a setup where electrons flow perpendicular to the DNA helix and into the substrate. Our results show that the direction of the tunneling current can be controlled based on the location of top contact, nearby molecular orbitals, and the shortest path to the bottom contacts (atoms at the DNA-substrate interface). We further observe that electrons injected in AT base-pairs can maintain large tunneling current at HOMO energy as if they were injected directly into the GC base-pairs where HOMO is localized. We also show that flipping the complementary strands can affect the resulting conformation and the amount of deformation in the strand. Hence, the design requires a careful consideration of the sequence and its orientation with respect to the substrate. We expect this study to help design futuristic DNA nanodevices beyond lithography techniques- like DNA Origami- that exploit the control of charge transport path in three-dimensions.

Chapter 9. CONCLUSIONS AND FUTURE WORK

9.1 SUMMARY

This thesis modeled quantum charge transport through nanoscale nucleic acid structures. The topics are summarized as follows:

In Chapter 1, we go over the structure and composition of the nucleic acids and the properties that make them appeal to nanoelectronics applications. We also review the challenges, potential applications, and previous work in modeling charge transport for such a system.

In Chapter 2, we go through the methodology of our work. We briefly review MD, DFT, and Green's function with Büttiker probes for including decoherence. We also investigate the difference in results between elastic and inelastic scattering in the transport model.

In Chapter 3, we introduce the energy-dependent decoherence model. The decoherence rates of this model are dependent on the molecular orbitals of the system and decay away from resonance. This treatment is more physical as it allows the transmission to decay with strand length near the midgap. Further, it allows us to apply higher decoherence rates without washing out interesting transmission features in HOMO and LUMO bands. We also demonstrate that both the real and the imaginary parts of the decoherence self-energy are required to yield the correct integration of the DOS, which is essential for self-consistent calculations.

In Chapter 4, we investigate the usage of the well-known NEGF approach to include inelastic scattering through phonons. We review the underlying formulas and test the approach on a model Hamiltonian representing a DNA. We find that the calculations do not converge. The approach is not applicable for systems where the phonon scattering rates are comparable to the hopping parameters between the bases in the Hamiltonian. The results of this chapter motivated us to apply the Büttiker probes for modeling decoherence.

In Chapter 5, we explore the possibility of detecting single nucleotide mismatch in a 15 base-pair long DNA:RNA strand through electrical conductance measurements. We studied three different locations for a single mismatch. The results show that there is a hierarchy of transmission values between the original strand and the three mismatch cases. This is also confirmed by the DOS analysis, thus showing the possibility of differentiating between the different mismatch cases through conductance measurements.

In Chapter 6, we investigated the impact of DNA doping through anthraquinone (Aq) intercalation. The modeled strand sequence is 3′-$G_3A_9G_3$-5′. Our results show that although the intercalator induces energy levels in the bandgap of the ds-DNA, this is not like the conventional doping in semiconductors. The added levels near HOMO (LUMO) are occupied (empty). Therefore, the electrons must travel through these induced energy levels to traverse the structure, as they do not significantly change the occupancy of HOMO (valence) and LUMO (conduction) energies that are primarily located at the DNA. It would be a good scientific problem to study suitable molecular species that can effectively dope DNA to create electrons in the conduction band and holes in the valence band. We further find that intercalation changes the expected location of the contact Fermi energy (E_f) with respect to the HOMO of the DNA. The reduced state of the Aq showed E_f to be closest to the HOMO, while the oxidized Aq yielded the farthest E_f value away from HOMO. The intercalation increased the conductance by 1.6 times the bare DNA. This increase can further be enhanced by shortening the length of the strand to 3′-$G_3A_3G_3$-5′ to reach more than four times increase in conductance. Our results further show that the conductance trend can vary depending on the experimental setup. In the three-terminal setup, where a gate voltage is added, the oxidized Aq intercalation yields higher conductance than the reduced Aq intercalation case when the relative Fermi energy locations with respect to the HOMO levels are the same.

In Chapter 7, we investigate modulating the DNA conductance by engineering the sequence to create tunneling wells and barrier structures, with sequences $3'$-TTTC$_N$TTT-$5'$and $3'$-CCCT$_N$CCC-$5'$, respectively (N varies from 0 to 5). We find that the conductance varies by over two orders of magnitude. Our results show that the conductance decreases substantially with increasing barrier width and only by a small amount with increasing well width. We further explore the impact of conformation by comparing the trends in A- and B-DNA forms. We find that B-DNA's conductance decreases more sharply with barrier width than A-DNA's. We attribute this result to A-DNA having energetically closer molecular orbitals that are spatially more delocalized than B-DNA. These properties of A-DNA help overcome the weaker π- π stacking that was originally thought to lower the conductance of A-DNA compared to B-DNA strands of the same sequence.

In Chapter 8, we modeled three 9 base-pair long ds-DNA with additional single strand bases of adenines (ss) on a gold substrate and studied the charge transport perpendicular to the DNA helix. We modeled ss-$C_3A_3C_3$-ss, ss-$C_3T_3C_3$-ss, and ss-$G_3A_3G_3$-ss; the ss adenines are added to enhance the surface adsorption. We investigated the impact of the top contact tip location by varying the contact self-energy location along the strands. Our results show that we can control the expected path of the electron tunneling perpendicularly by carefully designing the sequence, groove binding formation sites, and the location of the contact points. These results can help facilitate the future development of DNA nanoelectronics. For instance, the top contact can be an actual electrode tip, a charge injecting molecule, or even a crossover between DNA Origami structures. The DNA Origami technique can build complex 3D structures interconnected through short bases-pairs known as crossovers. The results of this study may provide design ideas for transport across helices in DNA Origami structures. The controlled path of electron transport paves the way for devices beyond photolithography that can exploit charge transport past the conventional planar motion of electrons.

9.2 FUTURE WORK

The theoretical research proposed here aims to capture and explain the physical principles of the transport mechanism and design devices in the emerging DNA nanostructures. Modeling DNA charge transport has its challenges. The non-rigid nature of the molecule due to the solvent environment adds more uncertainties to realizing the expected conformation, which impacts the energy levels and charge transport. MD simulations have shown that energy levels of the DNA fluctuate and are accompanied by fluctuations of the electronic coupling. The time-averaged behavior of these fluctuations is expected to be similar to the energy broadening effect obtained by the decoherence probes. Since current models in DNA transport do not include off-diagonal electronic coupling fluctuation, incorporating this feature adds more degrees of freedom in modeling the floppy molecule.

In addition, Integrating DNA-based devices with current semiconductor technology requires further study of the interface between the biomolecules and solid-state materials. These contact details alter the expected transmission profile as the contact self-energy needs to be applied carefully. Adding a few layers or atomic clusters of the contact to represent the contact-DNA junction in the DFT calculation can help generate a more accurate contact self-energy, better-simulating experiments.

Furthermore, DNA structures can be synthetically fabricated and chemically manipulated to form complex origami structures. DNA Origami opens endless possibilities in forming 3D structures that can potentially serve as electronic devices. However, to reach this potential, we need to be able to model these large systems. As a result, we need to develop methods to model larger structures, which can be at least 4x the current system size, increasing the computational complexity. A potential solution is to discover new ways to segment the structure into smaller blocks, in which transport calculations are performed.

BIBLIOGRAPHY

[1] C. Bryce and D. Pacini, *The Structure and Function of Nucleic Acids*. Portland Press Ltd, 1998.

[2] J. D. Watson and F. H. C. Crick, "Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid," *Nature*, vol. 171, no. 4356, pp. 737–738, 1953, doi: 10.1038/171737a0.

[3] G. S. Manning, "The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides," *Quarterly Reviews of Biophysics*, vol. 11, no. 2, pp. 179–246, 1978, doi: 10.1017/S0033583500002031.

[4] R. Hardison, "2.5: B-Form, A-Form, and Z-Form of DNA - Biology LibreTexts," 2019. https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA (accessed Apr. 17, 2020).

[5] F. M. Pohl and T. M. Jovin, "Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly(dG-dC)," *Journal of Molecular Biology*, vol. 67, no. 3, pp. 375–396, Jun. 1972, doi: 10.1016/0022-2836(72)90457-3.

[6] A. Herbert and A. Rich, "The biology of left-handed Z-DNA," *Journal of Biological Chemistry*, vol. 271, no. 20. American Society for Biochemistry and Molecular Biology, pp. 11595–11598, May 17, 1996. doi: 10.1074/jbc.271.20.11595.

[7] A. Rich, A. Nordheim, and A. H. J. Wang, "The Chemistry and Biology of Left-Handed Z-DNA," *Annual Review of Biochemistry*, vol. 53, no. 1, pp. 791–846, Jun. 1984, doi: 10.1146/annurev.bi.53.070184.004043.

[8] D. Porschke, "Boundary conditions for free A-DNA in solution and the relation of local to global DNA structures at reduced water activity," *European Biophysics Journal*, vol. 45, no. 5, pp. 413–421, 2016, doi: 10.1007/s00249-015-1110-1.

[9] M. Kulkarni and A. Mukherjee, "Understanding B-DNA to A-DNA transition in the right-handed DNA helix: Perspective from a local to global transition," *Progress in Biophysics and Molecular Biology*, vol. 128, pp. 63–73, 2017, doi: 10.1016/j.pbiomolbio.2017.05.009.

[10] K. M. Knee, S. B. Dixit, C. E. Aitken, S. Ponomarev, D. L. Beveridge, and I. Mukerji, "Spectroscopic and molecular dynamics evidence for a sequential mechanism for the A-to-B transition in DNA," *Biophysical Journal*, vol. 95, no. 1, pp. 257–272, 2008, doi: 10.1529/biophysj.107.117606.

[11] W. Saenger, W. N. Hunter, and O. Kennard, "DNA conformation is determined by economics in the hydration of phosphate groups," *Nature*, vol. 324, no. 6095, pp. 385–388, 1986, doi: 10.1038/324385a0.

[12] M. Chaplin, "Nucleic Acid Hydration," Apr. 02, 2020. http://www1.lsbu.ac.uk/water/nucleic_acid_hydration.html#d (accessed May 14, 2020).

[13] J. M. Artés, Y. Li, J. Qi, M. P. Anantram, and J. Hihath, “Conformational gating of DNA conductance,” *Nature Communications*, vol. 6, no. 1, p. 8870, Dec. 2015, doi: 10.1038/ncomms9870.

[14] R. G. Endres, “Colloquium : The quest for high-conductance DNA,” *Reviews of Modern Physics*, vol. 76, no. January, pp. 195–214, 2004, doi: 10.1103/RevModPhys.76.195.

[15] R. Johnson, “CCCBDB Notes on Ionization Energies.” https://cccbdb.nist.gov/adiabatic.asp (accessed Apr. 18, 2020).

[16] J. Franck, “Elementary processes of photochemical reactions,” *Transactions of the Faraday Society*, vol. 21, no. FEBRUARY, pp. 536–542, Jan. 1926, doi: 10.1039/TF9262100536.

[17] C. A. M. Seidel, A. Schulz, and M. H. M. Sauer, “Nucleobase-Specific Quenching of Fluorescent Dyes. 1. Nucleobase One-Electron Redox Potentials and Their Correlation with Static and Dynamic Quenching Efficiencies,” pp. 5541–5553, 1996, doi: 10.1021/jp951507c.

[18] P. Slavíček, B. Winter, M. Faubel, S. E. Bradforth, and P. Jungwirth, “Ionization energies of aqueous nucleic acids: Photoelectron spectroscopy of pyrimidine nucleosides and ab initio calculations,” *J Am Chem Soc*, vol. 131, no. 18, pp. 6460–6467, May 2009, doi: 10.1021/ja8091246.

[19] M. Xu, R. G. Endres, and Y. Arakawa, “The Electronic Properties of DNA Bases,” *Small*, vol. 3, no. 9, pp. 1539–1543, Sep. 2007, doi: 10.1002/smll.200600732.

[20] S. Kilina *et al.*, “Electronic properties of DNA base molecules adsorbed on a metallic surface,” *Journal of Physical Chemistry C*, vol. 111, no. 39, pp. 14541–14551, 2007, doi: 10.1021/jp070805u.

[21] D. M. Close, “Calculation of the ionization potentials of the DNA bases in aqueous medium,” *Journal of Physical Chemistry A*, vol. 108, no. 46, pp. 10376–10379, 2004, doi: 10.1021/jp046660y.

[22] D. D. Eley and D. I. Spivey, “Semiconductivity of organic substances. Part 9. - Nucleic acid in the dry state,” *Transactions of the Faraday Society*, vol. 58, no. 0, pp. 411–415, Jan. 1962, doi: 10.1039/TF9625800411.

[23] P. T. Henderson, D. Jones, G. Hampikian, Y. Kan, and G. B. Schuster, “Long-distance charge transport in duplex DNA: The phonon-assisted polaron-like hopping mechanism,” *Proceedings of the National Academy of Sciences*, vol. 96, no. 15, pp. 8353–8358, 1999, doi: 10.1073/pnas.96.15.8353.

[24] B. Giese, “Long-distance charge transport in DNA: The hopping mechanism,” *Accounts of Chemical Research*, vol. 33, no. 9, pp. 631–636, 2000, doi: 10.1021/ar990040b.

[25] F. D. Lewis, X. Liu, J. Liu, S. E. Miller, R. T. Hayes, and M. R. Wasielewski, “Direct measurement of hole transport dynamics in DNA,” *Nature*, vol. 406, no. 6791, pp. 51–53, Jul. 2000, doi: 10.1038/35017524.

[26] B. Giese, J. Amaudrut, A.-K. Kohler, M. Spormann, and S. Wessely, "Direct observation of hole transfer through DNA by hopping between adenine bases and by tunnelling," vol. 412, no. JULY, pp. 3–5, 2001.

[27] S. Delaney and J. K. Barton, "Long-Range DNA Charge Transport," pp. 6475–6483, 2003, doi: 10.1021/jo030095y.

[28] B. Xu and N. J. Tao, "Measurement of single-molecule resistance by repeated formation of molecular junctions," *Science (1979)*, vol. 301, no. 5637, pp. 1221–1223, Aug. 2003, doi: 10.1126/science.1087481.

[29] H. Cohen, C. Nogues, R. Naaman, and D. Porath, "Direct measurement of electrical transport through single DNA molecules of complex sequence," 2005.

[30] N. Kang, A. Erbe, and E. Scheer, "Electrical characterization of DNA in mechanically controlled break-junctions," *New Journal of Physics*, vol. 10, 2008, doi: 10.1088/1367-2630/10/2/023030.

[31] D. Xiang, H. Jeong, T. Lee, and D. Mayer, "Mechanically controllable break junctions for molecular electronics," *Advanced Materials*, vol. 25, no. 35, pp. 4845–4867, 2013, doi: 10.1002/adma.201301589.

[32] K. Wang, "DNA-based single-molecule electronics: From concept to function," *Journal of Functional Biomaterials*, vol. 9, no. 1, 2018, doi: 10.3390/jfb9010008.

[33] G. R. Abel, L. E. Korshoj, P. B. Otoupal, S. Khan, A. Chatterjee, and P. Nagpal, "Nucleotide and structural label identification in single RNA molecules with quantum tunneling spectroscopy," *Chemical Science*, vol. 10, no. 4, pp. 1052–1063, 2019, doi: 10.1039/c8sc03354d.

[34] R. Venkatramani, E. Wierzbinski, D. H. Waldeck, and D. N. Beratan, "Breaking the simple proportionality between molecular conductances and charge transfer rates," *Faraday Discussions*, vol. 174, pp. 57–78, 2014, doi: 10.1039/c4fd00106k.

[35] G. Khandelwal and J. Bhyravabhotla, "A phenomenological model for predicting melting temperatures of dna sequences," *PLoS ONE*, vol. 5, no. 8, pp. 1–9, 2010, doi: 10.1371/journal.pone.0012433.

[36] J. D. Slinker, N. B. Muren, S. E. Renfrew, and J. K. Barton, "DNA charge transport over 34 nm," *Nature Chemistry*, vol. 3, no. 3, pp. 228–233, Mar. 2011, doi: 10.1038/nchem.982.

[37] F. Chen, X. Li, J. Hihath, Z. Huang, and N. Tao, "Effect of Anchoring Groups on Single-Molecule Conductance: Comparative Study of Thiol-, Amine-, and Carboxylic-Acid-Terminated Molecules," 2006, doi: 10.1021/ja065864k.

[38] J. M. Beebe, V. B. Engelkes, L. L. Miller, and C. Daniel Frisbie, "Contact Resistance in Metal-Molecule-Metal Junctions Based on Aliphatic SAMs: Effects of Surface Linker and Metal Work Function," *J. AM. CHEM. SOC*, vol. 124, pp. 11268–11269, 2002, doi: 10.1021/ja0268332.

[39] A. Bellec, J. Lagoute, and V. Repain, "Molecular electronics: Scanning tunneling microscopy and single-molecule devices," *Comptes Rendus Chimie*, vol. 21, no. 12, pp. 1287–1299, Dec. 2018, doi: 10.1016/j.crci.2018.06.001.

[40] Y. Tsuji, T. Stuyver, S. Gunasekaran, and L. Venkataraman, "The Influence of Linkers on Quantum Interference: A Linker Theorem," *Journal of Physical Chemistry C*, vol. 121, no. 27, pp. 14451–14462, 2017, doi: 10.1021/acs.jpcc.7b03493.

[41] B. Xu, P. Zhang, X. Li, and N. Tao, "Direct conductance measurement of single DNA molecules in aqueous solution," *Nano Letters*, vol. 4, no. 6, pp. 1105–1108, 2004, doi: 10.1021/nl0494295.

[42] Y. Li, L. Xiang, J. L. Palma, Y. Asai, and N. Tao, "Thermoelectric effect and its dependence on molecular length and sequence in single DNA molecules," *Nature Communications*, vol. 7, no. 1, pp. 1–8, Apr. 2016, doi: 10.1038/ncomms11294.

[43] H. Zhang, W. Xu, X. Liu, F. Stellacci, and J. T. L. Thong, "Capturing a DNA duplex under near-physiological conditions," *Applied Physics Letters*, vol. 97, no. 16, p. 163702, Oct. 2010, doi: 10.1063/1.3505152.

[44] B. Tan, M. Hodak, W. Lu, and J. Bernholc, "Charge transport in DNA nanowires connected to carbon nanotubes," *Physical Review B - Condensed Matter and Materials Physics*, vol. 92, no. 7, pp. 1–7, 2015, doi: 10.1103/PhysRevB.92.075429.

[45] A. Y. Kasumov, "Proximity-Induced Superconductivity in DNA," *Science (1979)*, vol. 291, no. 5502, pp. 280–282, Jan. 2001, doi: 10.1126/science.291.5502.280.

[46] C. Nogues, S. R. Cohen, S. Daube, N. Apter, and R. Naaman, "Sequence dependence of charge transport properties of DNA," *Journal of Physical Chemistry B*, vol. 110, no. 18, pp. 8910–8913, May 2006, doi: 10.1021/jp060870o.

[47] A. Laikhter and K. Linse, "The Chemical Synthesis of Oligonucleotides." https://www.biosyn.com/tew/the-chemical-synthesis-of-oligonucleotides.aspx (accessed Apr. 18, 2020).

[48] P. W. K. Rothemund, "Folding DNA to create nanoscale shapes and patterns," *Nature*, vol. 440, no. 7082, pp. 297–302, Mar. 2006, doi: 10.1038/nature04586.

[49] Y. Ke *et al.*, "Multilayer DNA Origami Packed on a Square Lattice", doi: 10.1021/ja906381y.

[50] Y. Ke, L. L. Ong, W. M. Shih, and P. Yin, "Three-Dimensional Structures Self-Assembled from DNA Bricks," *Science (1979)*, vol. 338, no. 6111, pp. 1177–1183, Nov. 2012, doi: 10.1126/science.1227268.

[51] Y. Li *et al.*, "Detection and identification of genetic material via single-molecule conductance," *Nature Nanotechnology*, vol. 13, no. 12, pp. 1167–1173, Dec. 2018, doi: 10.1038/s41565-018-0285-x.

[52] E. Wierzbinski *et al.*, "Effect of backbone flexibility on charge transfer rates in peptide nucleic acid duplexes," *J Am Chem Soc*, vol. 134, no. 22, pp. 9335–9342, 2012, doi: 10.1021/ja301677z.

[53] M. Xu, R. G. Endres, and Y. Arakawa, “The Electronic Properties of DNA Bases,” *Small*, vol. 3, no. 9, pp. 1539–1543, Sep. 2007, doi: 10.1002/smll.200600732.

[54] J. Qi, N. Govind, and M. P. Anantram, “The role of cytosine methylation on charge transport through a DNA strand,” *The Journal of Chemical Physics*, vol. 143, no. 9, p. 094306, Sep. 2015, doi: 10.1063/1.4929909.

[55] Y. Zhang *et al.*, “Conformationally Gated Charge Transfer in DNA Three-Way Junctions,” 2015, doi: 10.1021/acs.jpclett.5b00863.

[56] C. Adessi, S. Walch, and M. P. Anantram, “Environment and structure influence on DNA conduction,” *Physical Review B - Condensed Matter and Materials Physics*, vol. 67, no. 8, pp. 1–4, 2003, doi: 10.1103/PhysRevB.67.081405.

[57] Z. Xie, T. Z. Markus, S. R. Cohen, Z. Vager, R. Gutierrez, and R. Naaman, “Spin specific electron conduction through DNA oligomers,” *Nano Letters*, vol. 11, no. 11, pp. 4652–4655, Nov. 2011, doi: 10.1021/nl2021637.

[58] Y. Zhang *et al.*, “Spin-resolved transport properties of DNA base multi-functional electronic devices,” *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 383, no. 17, pp. 2069–2075, 2019, doi: 10.1016/j.physleta.2019.03.034.

[59] B. Göhler *et al.*, “Spin selectivity in electron transmission through self-assembled monolayers of double-stranded DNA,” *Science (1979)*, vol. 331, no. 6019, pp. 894–897, Feb. 2011, doi: 10.1126/science.1199339.

[60] N. Kang, A. Erbe, and E. Scheer, “Observation of negative differential resistance in DNA molecular junctions,” *Applied Physics Letters*, vol. 96, no. 2, pp. 94–97, 2010, doi: 10.1063/1.3291113.

[61] C. Guo *et al.*, “Molecular rectifier composed of DNA with high rectification ratio enabled by intercalation,” *Nature Chemistry*, vol. 8, no. 5, pp. 484–490, May 2016, doi: 10.1038/nchem.2480.

[62] T. Harashima, C. Kojima, S. Fujii, M. Kiguchi, and T. Nishino, “Single-molecule conductance of DNA gated and ungated by DNA-binding molecules,” *Chemical Communications*, vol. 53, no. 75, pp. 10378–10381, 2017, doi: 10.1039/c7cc02911j.

[63] L. Xiang, J. L. Palma, Y. Li, V. Mujica, M. A. Ratner, and N. Tao, “Gate-controlled conductance switching in DNA,” *Nature Communications*, vol. 8, pp. 1–10, 2017, doi: 10.1038/ncomms14471.

[64] M. Yoon *et al.*, “Charge Transport in 2D DNA Tunnel Junction Diodes,” *Small*, vol. 13, no. 48, pp. 1–7, 2017, doi: 10.1002/smll.201703006.

[65] Y. Cho, J. Lee, J. Y. Lim, S. Yu, Y. Yi, and S. Im, “DNA-based small molecules for hole charge injection and channel passivation in organic heptazole field effect transistors,” *Journal of Physics D: Applied Physics*, vol. 50, no. 6, 2017, doi: 10.1088/1361-6463/50/6/065107.

[66] J. Lee *et al.*, “DNA-base guanine as hydrogen getter and charge-trapping layer embedded in oxide dielectrics for inorganic and organic field-effect transistors,” *ACS Applied Materials and Interfaces*, vol. 6, no. 7, pp. 4965–4973, 2014, doi: 10.1021/am405998d.

[67] T. Bayrak *et al.*, “DNA-Mold Templated Assembly of Conductive Gold Nanowires,” *Nano Letters*, vol. 18, no. 3, pp. 2116–2123, 2018, doi: 10.1021/acs.nanolett.8b00344.

[68] K. D. Robertson, “DNA methylation and human disease,” *Nature Reviews Genetics*, vol. 6, no. 8. pp. 597–610, Aug. 2005. doi: 10.1038/nrg1655.

[69] M. Tsutsui, K. Matsubara, T. Ohshiro, M. Furuhashi, M. Taniguchi, and T. Kawai, “Electrical Detection of Single Methylcytosines in a DNA Oligomer,” *J Am Chem Soc*, vol. 133, no. 23, pp. 9124–9128, Jun. 2011, doi: 10.1021/ja203839e.

[70] P. T. Bui, T. Nishino, H. Shiigi, and T. Nagaoka, “One-by-one single-molecule detection of mutated nucleobases by monitoring tunneling current using a DNA tip,” *Chemical Communications*, vol. 51, no. 9, pp. 1666–1669, Jan. 2015, doi: 10.1039/c4cc08227c.

[71] J. Shendure *et al.*, “DNA sequencing at 40: past, present and future,” *Nature*, vol. 550, no. 7676, pp. 345–353, Oct. 2017, doi: 10.1038/nature24286.

[72] S. J. Heerema and C. Dekker, “Graphene nanodevices for DNA sequencing,” *Nature Nanotechnology*, vol. 11, no. 2. Nature Publishing Group, pp. 127–136, Feb. 01, 2016. doi: 10.1038/nnano.2015.307.

[73] B. Merriman, I. T. R&D Team, and J. M. Rothberg, “Progress in Ion Torrent semiconductor chip based sequencing,” *ELECTROPHORESIS*, vol. 33, no. 23, pp. 3397–3417, Dec. 2012, doi: 10.1002/elps.201200424.

[74] J. Karow, “Roswell Biotechnologies Harnesses Molecular Electronics for Chip-Based DNA Sequencing,” 2018.

[75] J. K. Sowa, J. A. Mol, G. Andrew, D. Briggs, and E. M. Gauger, “Beyond Marcus theory and the Landauer-Büttiker approach in molecular junctions: A unified framework ARTICLES YOU MAY BE INTERESTED IN,” *J. Chem. Phys*, vol. 149, p. 154112, 2018, doi: 10.1063/1.5049537.

[76] T. Kubař, R. Gutiérrez, U. Kleinekathöfer, G. Cuniberti, and M. Elstner, “Modeling charge transport in DNA using multi-scale methods,” *Physica Status Solidi (B) Basic Research*, vol. 250, no. 11, pp. 2277–2287, 2013, doi: 10.1002/pssb.201349148.

[77] Y. Li *et al.*, “Comparing Charge Transport in Oligonucleotides: RNA:DNA Hybrids and DNA Duplexes,” *Journal of Physical Chemistry Letters*, vol. 7, no. 10, pp. 1888–1894, 2016, doi: 10.1021/acs.jpclett.6b00749.

[78] H. Mehrez and M. P. Anantram, “Interbase electronic coupling for transport through DNA,” *Physical Review B - Condensed Matter and Materials Physics*, vol. 71, no. 11, 2005, doi: 10.1103/PhysRevB.71.115405.

[79] M. P. Anantram and J. Qi, “Modeling of electron transport in biomolecules: Application to DNA,” in *Technical Digest - International Electron Devices Meeting, IEDM*, 2013, pp. 798–801. doi: 10.1109/IEDM.2013.6724736.

[80] T. Kubař and M. Elstner, "What governs the charge transfer in DNA? The role of DNA conformation and environment," *Journal of Physical Chemistry B*, vol. 112, no. 29, pp. 8788–8798, 2008, doi: 10.1021/jp803661f.

[81] J. Qi, N. Edirisinghe, M. G. Rabbani, and M. P. Anantram, "Unified model for conductance through DNA with the Landauer-Büttiker formalism," *Physical Review B - Condensed Matter and Materials Physics*, vol. 87, no. 8, pp. 1–10, 2013, doi: 10.1103/PhysRevB.87.085404.

[82] M. Kilgour and D. Segal, "Charge transport in molecular junctions: From tunneling to hopping with the probe technique," vol. 024111, no. 2015, 2015, doi: 10.1063/1.4926395.

[83] R. Gutiérrez, S. Mandai, and G. Cuniberti, "Quantum transport through a DMA wire in a dissipative environment," *Nano Letters*, vol. 5, no. 6, pp. 1093–1097, 2005, doi: 10.1007/s12094-006-0148-z.

[84] R. Gutiérrez, S. Mandal, and G. Cuniberti, "Dissipative effects in the electronic transport through DNA molecular wires," *Physical Review B - Condensed Matter and Materials Physics*, vol. 71, no. 23, pp. 1–9, 2005, doi: 10.1103/PhysRevB.71.235116.

[85] P. Karasch, D. A. Ryndyk, and T. Frauenheim, "Vibronic Dephasing Model for Coherent-to-Incoherent Crossover in DNA," Feb. 2018, doi: 10.1103/PhysRevB.97.195401.

[86] G. Penazzi, A. Pecchia, V. Gupta, and T. Frauenheim, "A Self Energy Model of Dephasing in Molecular Junctions," 2016, doi: 10.1021/acs.jpcc.6b04185.

[87] A. J. White, M. A. Ochoa, and M. Galperin, "Nonequilibrium atomic limit for transport and optical response of molecular junctions," *Journal of Physical Chemistry C*, vol. 118, no. 21, pp. 11159–11173, 2014, doi: 10.1021/jp500880j.

[88] H. Mohammad *et al.*, "Role of intercalation in the electrical properties of nucleic acids for use in molecular electronics," *Nanoscale Horizons*, vol. 6, no. 8, pp. 651–660, 2021, doi: 10.1039/D1NH00211B.

[89] X.-Q. Li and Y. Yan, "Electrical transport through individual DNA molecules," *Appl. Phys. Lett*, vol. 79, p. 2190, 2001, doi: 10.1063/1.1407860.

[90] H. Kim, M. Kilgour, and D. Segal, "Intermediate Coherent-Incoherent Charge Transport: DNA as a Case Study," *Journal of Physical Chemistry C*, vol. 120, no. 42, pp. 23951–23962, 2016, doi: 10.1021/acs.jpcc.6b07602.

[91] R. Venkatramani, S. Keinan, A. Balaeff, and D. N. Beratan, "Nucleic acid charge transfer: Black, white and gray," *Coordination Chemistry Reviews*, vol. 255, no. 7–8, pp. 635–648, 2011, doi: 10.1016/j.ccr.2010.12.010.

[92] E. Wierzbinski *et al.*, "Charge transfer through modified peptide nucleic acids," *Langmuir*, vol. 28, no. 4, pp. 1971–1981, 2012, doi: 10.1021/la204445u.

[93] R. Gutiérrez, R. Caetano, P. B. Woiczikowski, T. Kubar, M. Elstner, and G. Cuniberti, "Structural fluctuations and quantum transport through DNA molecular wires: A combined molecular dynamics and model hamiltonian approach," *New Journal of Physics*, vol. 12, 2010, doi: 10.1088/1367-2630/12/2/023022.

[94] E. M. Conwell, "Charge transport in DNA in solution: The role of polarons," *Proceedings of the National Academy of Sciences*, vol. 102, no. 25, pp. 8795–8799, 2005, doi: 10.1073/pnas.0501406102.

[95] E. M. Conwell and S. V Rakhmanova, "Polarons in DNA," *Proceedings of the …*, vol. 97, no. 9, pp. 4556–60, 2000, doi: 10.1073/pnas.050074497.

[96] A. A. Voityuk, "Charge transfer in DNA: Hole charge is confined to a single base pair due to solvation effects," *Journal of Chemical Physics*, vol. 122, no. 20, p. 204904, May 2005, doi: 10.1063/1.1924551.

[97] J. C. Genereux and J. K. Barton, "Mechanisms for DNA Charge Transport", doi: 10.1021/cr900228f.

[98] D. J. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratch, "Gaussian 16, Revision C.01." Gaussian, Inc., Wallingford CT, 2016.

[99] M. A. González, "Force fields and molecular dynamics simulations," *Collection SFN*, vol. 12, pp. 169–200, 2011, doi: 10.1051/sfn/201112009.

[100] D. A. Case *et al.*, "The Amber biomolecular simulation programs," *Journal of Computational Chemistry*, vol. 26, no. 16. pp. 1668–1688, Dec. 2005. doi: 10.1002/jcc.20290.

[101] F. Jensen, *Introduction to computational chemistry*. John Wiley & Sons, 2007.

[102] J. M. Soler *et al.*, "The SIESTA method for ab initio order-N materials simulation Related content O(N) methods in electronic structure calculations D R Bowler and T Miyazaki-Topical Review Mark E Tuckerman-B-splines in atomic and molecular physics H Bachau, E Cormier, P Decleva et al.-Recent citations The SIESTA method for ab initio order-N materials simulation," 2002.

[103] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," *Physical Review B - Condensed Matter and Materials Physics*, vol. 54, no. 16, pp. 11169–11186, Oct. 1996, doi: 10.1103/PhysRevB.54.11169.

[104] P. Giannozzi *et al.*, "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials Related content QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials," *J. Phys.: Condens. Matter*, vol. 21, p. 395502, 2009, doi: 10.1088/0953-8984/21/39/395502.

[105] G. te Velde *et al.*, "Chemistry with ADF," *Journal of Computational Chemistry*, vol. 22, no. 9, pp. 931–967, Jul. 2001, doi: 10.1002/jcc.1056.

[106] E. Cauët, M. Valiev, and J. H. Weare, "Vertical ionization potentials of nucleobases in a fully solvated DNA environment," *Journal of Physical Chemistry B*, vol. 114, no. 17, pp. 5886–5894, 2010, doi: 10.1021/jp9120723.

[107] N. Russo, M. Toscano, and A. Grand, "Theoretical Determination of Electron Affinity and Ionization Potential of DNA and RNA Bases," *Journal of Computational Chemistry*, vol. 21, no. 14, pp. 1243–1250, 2000, doi: 10.1002/1096-987X(20001115)21:14<1243::AID-JCC3>3.0.CO;2-M.

[108] E. Cauët, D. Dehareng, and J. Liévin, "Ab initio study of the lonization of the DNA bases: Ionization potentials and excited states of the cations," *Journal of Physical Chemistry A*, vol. 110, no. 29, pp. 9200–9211, 2006, doi: 10.1021/jp0617625.

[109] M. Félix and A. A. Voityuk, "DFT performance for the hole transfer parameters in DNA π stacks," *International Journal of Quantum Chemistry*, vol. 111, no. 1, pp. 191–201, Jan. 2011, doi: 10.1002/QUA.22419.

[110] T. Yamamoto, T. Uda, T. Yamasaki, and T. Ohno, "Hydration effect on the optical property of a DNA fiber: First-principles and molecular dynamics studies," *Physical Chemistry Chemical Physics*, vol. 12, no. 32, pp. 9300–9311, 2010, doi: 10.1039/b924678a.

[111] E. Pluhařová, P. Slavíček, and P. Jungwirth, "Modeling photoionization of aqueous DNA and its components," *Accounts of Chemical Research*, vol. 48, no. 5, pp. 1209–1217, 2015, doi: 10.1021/ar500366z.

[112] E. D. Raczyńska, P. Michalec, M. Zalewski, and M. Sapuła, "Effects of ionization on stability of 1-methylcytosine — DFT and PCM studies," *Journal of Molecular Modeling*, vol. 22, no. 7, Jul. 2016, doi: 10.1007/S00894-016-3020-2.

[113] "SCRF, Gaussian." https://gaussian.com/scrf/ (accessed Apr. 18, 2020).

[114] S. Datta, "Nanoscale device modeling: the Green's function method," *Superlattices and Microstructures*, vol. 28, no. 4, pp. 253–278, 2000, doi: 10.1006/spmi.2000.0920.

[115] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. S. S. Lundstrom, "A simple quantum mechanical treatment of scattering in nanoscale transistors," *Journal of Applied Physics*, vol. 93, no. 9, pp. 5613–5625, 2003, doi: 10.1063/1.1563298.

[116] M. P. Anantram, M. S. Lundstrom, D. E. D. E. Nikonov, and S. M. Ieee, "Modeling of Nanoscale Devices," *Proceedings of the IEEE*, vol. 96, no. 9, pp. 1511–1550, Sep. 2008, doi: 10.1109/JPROC.2008.927355.

[117] A. Erpenbeck, R. Härtle, and M. Thoss, "Effect of nonadiabatic electronic-vibrational interactions on the transport properties of single-molecule junctions," *Physical Review B - Condensed Matter and Materials Physics*, vol. 91, no. 19, pp. 1–14, 2015, doi: 10.1103/PhysRevB.91.195418.

[118] M. Buttiker, "Coherent and sequential tunneling in series barriers," *IBM Journal of Research and Development*, vol. 32, no. 1, pp. 63–75, 1988, doi: 10.1147/rd.321.0063.

[119] H. Rahman, P. Karasch, D. A. Ryndyk, T. Frauenheim, and U. Kleinekathöfer, "Dephasing in a Molecular Junction Viewed from a Time-Dependent and a Time-Independent Perspective," *The Journal of Physical Chemistry C*, 2019, doi: 10.1021/acs.jpcc.9b00955.

[120] R. Korol and D. Segal, “From Exhaustive Simulations to Key Principles in DNA Nanoelectronics,” *Journal of Physical Chemistry C*, vol. 122, no. 8, pp. 4206–4216, 2018, doi: 10.1021/acs.jpcc.7b12744.

[121] M. Kilgour and D. Segal, “Inelastic effects in molecular transport junctions: The probe technique at high bias,” *Journal of Chemical Physics*, vol. 144, no. 12, 2016, doi: 10.1063/1.4944470.

[122] K. Senthilkumar *et al.*, “Absolute rates of hole transfer in DNA,” *J Am Chem Soc*, vol. 127, no. 42, pp. 14894–14903, 2005, doi: 10.1021/ja054257e.

[123] S. R. Patil *et al.*, “Quantum Transport in DNA Heterostructures: Implications for Nanoelectronics,” *ACS Applied Nano Materials*, vol. 4, no. 10, pp. 10029–10037, Oct. 2021, doi: 10.1021/acsanm.1c01087.

[124] L. D. Landau and E. M. Lifshits, *Electrodynamics of Continuous Media 2nd Edition*, Second Edition., vol. 8. Butterworth-Heinemann, 1984.

[125] C. J. O. Verzijl, J. S. Seldenthuis, and J. M. Thijssen, “Applicability of the wide-band limit in DFT-based molecular transport calculations,” *Journal of Chemical Physics*, vol. 138, no. 9, 2013, doi: 10.1063/1.4793259.

[126] J. L. de Amato and H. M. Pastawski, “Conductance of a disordered linear chain including inelastic scattering events,” *Physical Review B*, vol. 41, no. 11, pp. 7411–7420, 1990, doi: 10.1103/PhysRevB.41.7411.

[127] J. Hihath, S. Guo, P. Zhang, and N. Tao, “Effects of cytosine methylation on DNA charge transport,” *Journal of Physics Condensed Matter*, vol. 24, no. 16, 2012, doi: 10.1088/0953-8984/24/16/164204.

[128] R. M. Stoltenberg and A. T. Woolley, “DNA-templated nanowire fabrication,” *Biomedical Microdevices*, vol. 6, no. 2, pp. 105–111, 2004, doi: 10.1023/B:BMMD.0000031746.46801.7d.

[129] A. v. Malyshev, “DNA double helices for single molecule electronics,” *Physical Review Letters*, vol. 98, no. 9, pp. 1–4, 2007, doi: 10.1103/PhysRevLett.98.096801.

[130] B. Göhler *et al.*, “Spin selectivity in electron transmission through self-assembled monolayers of double-stranded DNA,” *Science (1979)*, vol. 331, no. 6019, pp. 894–897, Feb. 2011, doi: 10.1126/science.1199339.

[131] C. H. Wohlgamuth, M. A. Mcwilliams, and J. D. Slinker, “DNA as a Molecular Wire: Distance and Sequence Dependence,” 2013, doi: 10.1021/ac401229q.

[132] C. Liu, L. Guo, B. Zhang, and L. Lu, “Graphene quantum dots mediated electron transfer in DNA base pairs,” *RSC Advances*, vol. 9, no. 54, pp. 31636–31644, 2019, doi: 10.1039/c9ra05481b.

[133] A. K. Mahapatro, K. J. Jeong, G. U. Lee, and D. B. Janes, “Sequence specific electronic conduction through polyion-stabilized double-stranded DNA in nanoscale break junctions,” *Nanotechnology*, vol. 18, no. 19, p. 195202, 2007, doi: 10.1088/0957-4484/18/19/195202.

[134] A. Svizhenko and M. P. Anantram, “Effect of scattering and contacts on current and electrostatics in carbon nanotubes,” *Physical Review B - Condensed Matter and Materials Physics*, vol. 72, no. 8, 2005, doi: 10.1103/PhysRevB.72.085430.

[135] S. Hood, N. Zarrabi, P. Meredith, I. Kassal, and A. Armin, “Measuring Energetic Disorder in Organic Semiconductors Using the Photogenerated Charge-Separation Efficiency,” 2019, doi: 10.1021/acs.jpclett.9b01304.

[136] B. C. Hoffman, S. Pazoki, A. Apperson, and D. B. Dougherty, “Spatially Uniform Shallow Trap Distribution in an Ultrathin Organic Transistor,” *Physica Status Solidi - Rapid Research Letters*, vol. 13, no. 5, pp. 4–7, 2019, doi: 10.1002/pssr.201800486.

[137] G. Horowitz, “Validity of the concept of band edge in organic semiconductors,” *Journal of Applied Physics*, vol. 118, no. 11, p. 115502, Sep. 2015, doi: 10.1063/1.4931061.

[138] R. J. Meier, “On art and science in curve-fitting vibrational spectra,” *Vib Spectrosc*, vol. 39, no. 2, pp. 266–269, Oct. 2005, doi: 10.1016/j.vibspec.2005.03.003.

[139] K. Akaike, “Advanced understanding on electronic structure of molecular semiconductors and their interfaces,” in *Japanese Journal of Applied Physics*, 2018, vol. 57, no. 3. doi: 10.7567/JJAP.57.03EA03.

[140] W. W. Parson, “Temperature Dependence of the Rate of Intramolecular Electron Transfer,” *Journal of Physical Chemistry B*, vol. 122, no. 38, pp. 8824–8833, 2018, doi: 10.1021/acs.jpcb.8b06497.

[141] W. W. Parson, “Vibrational Relaxations and Dephasing in Electron-Transfer Reactions,” *The Journal of Physical Chemistry B*, vol. 120, no. 44, pp. 11412–11418, Nov. 2016, doi: 10.1021/acs.jpcb.6b08803.

[142] D. A. Case *et al.*, “Amber 2016,” *University of California, San Francisco*, 2016.

[143] F. Zahid, M. Paulsson, and S. Datta, “Electrical Conduction through Molecules,” in *Advanced Semiconductor and Organic Nano-Techniques*, Elsevier Inc., 2003, pp. 1–41. doi: 10.1016/B978-012507060-7/50022-2.

[144] M. Baldo, “6.701 Introduction to Nanoelectronics,” *Massachusetts Institute of Technology: MIT OpenCourseWare*, no. Spring. 2010. Accessed: Feb. 09, 2022. [Online]. Available: https://ocw.mit.edu

[145] M. P. Anantram and A. Svizhenko, “Multidimensional modeling of nanotransistors,” *IEEE Transactions on Electron Devices*, vol. 54, no. 9, pp. 2100–2115, 2007, doi: 10.1109/TED.2007.902857.

[146] A. Troisi, “Charge dynamics through pi-stacked arrays of conjugated molecules: Effect of dynamic disorder in different transport/transfer regimes,” *Molecular Simulation*, vol. 32, no. 9, pp. 707–716, 2006, doi: 10.1080/08927020600857305.

[147] A. Troisi and G. Orlandi, “Dynamics of the intermolecular transfer integral in crystalline organic semiconductors,” *Journal of Physical Chemistry A*, vol. 110, no. 11, pp. 4065–4070, 2006, doi: 10.1021/jp055432g.

[148] S. Stafström, "Electron localization and the transition from adiabatic to nonadiabatic charge transport in organic conductors," *Chemical Society Reviews*, vol. 39, no. 7, pp. 2484–2499, 2010, doi: 10.1039/b909058b.

[149] B. Kim, J. M. Beebe, Y. Jun, X. Y. Zhu, and G. D. Frisbie, "Correlation between HOMO alignment and contact resistance in molecular junctions: Aromatic thiols versus aromatic isocyanides," *J Am Chem Soc*, vol. 128, no. 15, pp. 4970–4971, Apr. 2006, doi: 10.1021/ja0607990.

[150] S. Datta, *Quantum transport: Atom to transistor*, vol. 9780521631457. Cambridge University Press, 2005. doi: 10.1017/CBO9781139164313.

[151] E. Wierzbinski *et al.*, "Effect of backbone flexibility on charge transfer rates in peptide nucleic acid duplexes," *J Am Chem Soc*, vol. 134, no. 22, pp. 9335–9342, 2012, doi: 10.1021/ja301677z.

[152] M. Xu, R. G. Endres, and Y. Arakawa, "The electronic properties of DNA bases," *Small*, vol. 3, no. 9. pp. 1539–1543, 2007. doi: 10.1002/smll.200600732.

[153] J. Qi, N. Govind, and M. P. Anantram, "The role of cytosine methylation on charge transport through a DNA strand," *Journal of Chemical Physics*, vol. 143, no. 9, 2015, doi: 10.1063/1.4929909.

[154] Y. Zhang *et al.*, "Conformationally Gated Charge Transfer in DNA Three-Way Junctions," *The Journal of Physical Chemistry Letters*, vol. 6, no. 13, pp. 2434–2438, Jul. 2015, doi: 10.1021/acs.jpclett.5b00863.

[155] J. M. Artés, Y. Li, J. Qi, M. P. Anantram, and J. Hihath, "Conformational gating of DNA conductance," *Nature Communications*, vol. 6, pp. 1–8, 2015, doi: 10.1038/ncomms9870.

[156] C. Adessi, S. Walch, and M. P. Anantram, "Environment and structure influence on DNA conduction," *Physical Review B - Condensed Matter and Materials Physics*, vol. 67, no. 8, pp. 1–4, 2003, doi: 10.1103/PhysRevB.67.081405.

[157] C. Bruot, L. Xiang, J. L. Palma, and N. Tao, "Effect of Mechanical Stretching on DNA Conductance," *ACS Nano*, vol. 9, no. 1, pp. 88–94, Jan. 2015, doi: 10.1021/nn506280t.

[158] Y. Li *et al.*, "Detection and identification of genetic material via single-molecule conductance," *Nature Nanotechnology*, vol. 13, no. 12, pp. 1167–1173, 2018, doi: 10.1038/s41565-018-0285-x.

[159] T. Bayrak *et al.*, "DNA-Mold Templated Assembly of Conductive Gold Nanowires," *Nano Letters*, vol. 18, no. 3, pp. 2116–2123, 2018, doi: 10.1021/acs.nanolett.8b00344.

[160] Y. Zhang *et al.*, "Spin-resolved transport properties of DNA base multi-functional electronic devices," *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 383, no. 17, pp. 2069–2075, 2019, doi: 10.1016/j.physleta.2019.03.034.

[161] H. M. J. Al-Ta'ii, V. Periasamy, and Y. M. Amin, "Humidity-dependent characteristics of DNA thin film-based Al/DNA/Al surface-type cell," *Sensors and Actuators, B: Chemical*, vol. 232, pp. 195–202, 2016, doi: 10.1016/j.snb.2016.03.093.

[162] M. Yoon *et al.*, “Charge Transport in 2D DNA Tunnel Junction Diodes,” *Small*, vol. 13, no. 48, p. 1703006, Dec. 2017, doi: 10.1002/smll.201703006.

[163] J. Lee *et al.*, “DNA-base guanine as hydrogen getter and charge-trapping layer embedded in oxide dielectrics for inorganic and organic field-effect transistors,” *ACS Applied Materials and Interfaces*, vol. 6, no. 7, pp. 4965–4973, 2014, doi: 10.1021/am405998d.

[164] Y. Cho, J. Lee, J. Y. Lim, S. Yu, Y. Yi, and S. Im, “DNA-based small molecules for hole charge injection and channel passivation in organic heptazole field effect transistors,” *Journal of Physics D: Applied Physics*, vol. 50, no. 6, p. 065107, Feb. 2017, doi: 10.1088/1361-6463/50/6/065107.

[165] L. Kékedy-Nagy and E. E. Ferapontova, “Directional Preference of DNA-Mediated Electron Transfer in Gold-Tethered DNA Duplexes: Is DNA a Molecular Rectifier?,” *Angewandte Chemie - International Edition*, vol. 58, no. 10, pp. 3048–3052, 2019, doi: 10.1002/anie.201809559.

[166] K. L. Jiménez-Monroy *et al.*, “High Electronic Conductance through Double-Helix DNA Molecules with Fullerene Anchoring Groups,” *The Journal of Physical Chemistry A*, vol. 121, no. 6, pp. 1182–1188, Feb. 2017, doi: 10.1021/acs.jpca.7b00348.

[167] M. DeLuca, Z. Shi, C. E. Castro, and G. Arya, “Dynamic DNA nanotechnology: toward functional nanoscale devices,” *Nanoscale Horizons*, vol. 5, no. 2, pp. 182–201, Feb. 2020, doi: 10.1039/C9NH00529C.

[168] G. R. Abel, L. E. Korshoj, P. B. Otoupal, S. Khan, A. Chatterjee, and P. Nagpal, “Nucleotide and structural label identification in single RNA molecules with quantum tunneling spectroscopy,” *Chemical Science*, vol. 10, no. 4, pp. 1052–1063, 2019, doi: 10.1039/c8sc03354d.

[169] T. Furuhata, T. Ohshiro, G. Akimoto, R. Ueki, M. Taniguchi, and S. Sando, “Highly Conductive Nucleotide Analogue Facilitates Base-Calling in Quantum-Tunneling-Based DNA Sequencing,” *ACS Nano*, vol. 13, no. 5, pp. 5028–5035, May 2019, doi: 10.1021/acsnano.9b01250.

[170] S. Afsari, L. E. Korshoj, G. R. Abel, S. Khan, A. Chatterjee, and P. Nagpal, “Quantum Point Contact Single-Nucleotide Conductance for DNA and RNA Sequence Identification,” *ACS Nano*, vol. 11, no. 11, pp. 11169–11181, Nov. 2017, doi: 10.1021/acsnano.7b05500.

[171] J. B. Chaires, “Drug—DNA interactions,” *Current Opinion in Structural Biology*, vol. 8, no. 3, pp. 314–320, Jun. 1998, doi: 10.1016/S0959-440X(98)80064-X.

[172] S. Bhaduri, N. Ranjan, and D. P. Arya, “An overview of recent advances in duplex DNA recognition by small molecules,” *Beilstein Journal of Organic Chemistry*, vol. 14, pp. 1051–1086, May 2018, doi: 10.3762/bjoc.14.93.

[173] H. S. Rye *et al.*, “Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: properties and applications,” *Nucleic Acids Research*, vol. 20, no. 11, pp. 2803–2812, 1992, doi: 10.1093/nar/20.11.2803.

[174] H. M. Berman and P. R. Young, “The Interaction of Intercalating Drugs with Nucleic Acids,” *Annual Review of Biophysics and Bioengineering*, vol. 10, no. 1, pp. 87–114, Jun. 1981, doi: 10.1146/annurev.bb.10.060181.000511.

[175] J. B. Chaires, “A thermodynamic signature for drug–DNA binding mode,” *Archives of Biochemistry and Biophysics*, vol. 453, no. 1, pp. 26–31, Sep. 2006, doi: 10.1016/j.abb.2006.03.027.

[176] H. Lei, X. Wang, and C. Wu, “Early stage intercalation of doxorubicin to DNA fragments observed in molecular dynamics binding simulations,” *Journal of Molecular Graphics and Modelling*, vol. 38, pp. 279–289, Sep. 2012, doi: 10.1016/j.jmgm.2012.05.006.

[177] A. A. Almaqwashi, T. Paramanathan, I. Rouzina, and M. C. Williams, “Mechanisms of small molecule–DNA interactions probed by single-molecule force spectroscopy,” *Nucleic Acids Research*, vol. 44, no. 9, pp. 3971–3988, May 2016, doi: 10.1093/nar/gkw237.

[178] L. Maganti and D. Bhattacharyya, “Sequence specificity in DNA–drug intercalation: MD simulation and density functional theory approaches,” *Journal of Computer-Aided Molecular Design*, vol. 34, no. 1, pp. 83–95, Jan. 2020, doi: 10.1007/s10822-019-00268-y.

[179] W. D. Sasikala and A. Mukherjee, “Intercalation and de-intercalation pathway of proflavine through the minor and major grooves of DNA: roles of water and entropy,” *Physical Chemistry Chemical Physics*, vol. 15, no. 17, p. 6446, 2013, doi: 10.1039/c3cp50501d.

[180] C. Guo *et al.*, “Molecular rectifier composed of DNA with high rectification ratio enabled by intercalation,” *Nature Chemistry*, vol. 8, no. 5, pp. 484–490, 2016, doi: 10.1038/nchem.2480.

[181] X. Wang, L. Gao, B. Liang, X. Li, and X. Guo, “Revealing the direct effect of individual intercalations on DNA conductance toward single-molecule electrical biodetection,” *Journal of Materials Chemistry B*, vol. 3, no. 26, pp. 5150–5154, 2015, doi: 10.1039/c5tb00666j.

[182] L. Xiang, J. L. Palma, Y. Li, V. Mujica, M. A. Ratner, and N. Tao, “Gate-controlled conductance switching in DNA,” *Nature Communications*, vol. 8, pp. 1–10, 2017, doi: 10.1038/ncomms14471.

[183] A. Aggarwal, A. K. Sahoo, S. Bag, V. Kaliginedi, M. Jain, and P. K. Maiti, “Fine-tuning the DNA conductance by intercalation of drug molecules,” *Physical Review E*, vol. 103, no. 3, p. 032411, Mar. 2021, doi: 10.1103/PhysRevE.103.032411.

[184] H. Xi, D. Gray, S. Kumar, and D. P. Arya, “Molecular recognition of single-stranded RNA: Neomycin binding to poly(A),” *FEBS Letters*, vol. 583, no. 13, pp. 2269–2275, Jul. 2009, doi: 10.1016/j.febslet.2009.06.007.

[185] N. N. Shaw, H. Xi, and D. P. Arya, “Molecular recognition of a DNA:RNA hybrid: Sub-nanomolar binding by a neomycin–methidium conjugate,” *Bioorganic & Medicinal Chemistry Letters*, vol. 18, no. 14, pp. 4142–4145, Jul. 2008, doi: 10.1016/j.bmcl.2008.05.090.

[186] N. N. Degtyareva *et al.*, “Antimicrobial Activity, AME Resistance, and A-Site Binding Studies of Anthraquinone-Neomycin Conjugates,” *ACS Infectious Diseases*, 2017, doi: 10.1021/acsinfecdis.6b00176.

[187] T. Harashima, C. Kojima, S. Fujii, M. Kiguchi, and T. Nishino, “Single-molecule conductance of DNA gated and ungated by DNA-binding molecules,” *Chemical Communications*, vol. 53, no. 75, pp. 10378–10381, 2017, doi: 10.1039/c7cc02911j.

[188] J. S. Al-Otaibi, P. Teesdale Spittle, and T. M. El Gogary, “Interaction of anthraquinone anti-cancer drugs with DNA:Experimental and computational quantum chemical study,” *Journal of Molecular Structure*, vol. 1127, pp. 751–760, Jan. 2017, doi: 10.1016/j.molstruc.2016.08.007.

[189] B. Gatto, G. Zagotto, C. Sissi, and M. Palumbo, “Preferred interaction of d-peptidyl-anthraquinones with double-stranded B-DNA,” *International Journal of Biological Macromolecules*, vol. 21, no. 4, pp. 319–326, Dec. 1997, doi: 10.1016/S0141-8130(97)00076-7.

[190] D. J. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratch, “Gaussian 16, Revision C.01.” Gaussian, Inc., Wallingford CT, 2016.

[191] H. Kim, M. Kilgour, and D. Segal, “Intermediate Coherent-Incoherent Charge Transport: DNA as a Case Study,” *Journal of Physical Chemistry C*, vol. 120, no. 42, pp. 23951–23962, 2016, doi: 10.1021/acs.jpcc.6b07602.

[192] J. Qi, N. Edirisinghe, M. G. Rabbani, and M. P. Anantram, “Unified model for conductance through DNA with the Landauer-Büttiker formalism,” *Physical Review B - Condensed Matter and Materials Physics*, 2013, doi: 10.1103/PhysRevB.87.085404.

[193] A. Aggarwal, S. Bag, R. Venkatramani, M. Jain, and P. K. Maiti, “Multiscale modelling reveals higher charge transport efficiencies of DNA relative to RNA independent of mechanism,” *Nanoscale*, vol. 12, no. 36, pp. 18750–18760, 2020, doi: 10.1039/d0nr02382e.

[194] X. Wang, L. Gao, B. Liang, X. Li, and X. Guo, “Revealing the direct effect of individual intercalations on DNA conductance toward single-molecule electrical biodetection,” *Journal of Materials Chemistry B*, vol. 3, no. 26, pp. 5150–5154, 2015, doi: 10.1039/c5tb00666j.

[195] A. Patra and C. Richert, “High Fidelity Base Pairing at the 3′-Terminus,” *J Am Chem Soc*, vol. 131, no. 35, pp. 12671–12681, Sep. 2009, doi: 10.1021/ja9033654.

[196] J. N. Lisgarten, M. Coll, J. Portugal, C. W. Wright, and J. Aymami, “The antimalarial and cytotoxic drug cryptolepine intercalates into dna at cytosine-cytosine sites,” *Nature Structural Biology*, vol. 9, no. 1, pp. 57–60, Dec. 2002, doi: 10.1038/nsb729.

[197] Kim, J. M. Beebe, Y. Jun, X.-Y. Zhu, and C. D. Frisbie, “Correlation between HOMO Alignment and Contact Resistance in Molecular Junctions: Aromatic Thiols versus Aromatic Isocyanides,” *J Am Chem Soc*, vol. 128, no. 15, pp. 4970–4971, Apr. 2006, doi: 10.1021/ja0607990.

[198] Y. Tsuji, T. Stuyver, S. Gunasekaran, and L. Venkataraman, “The Influence of Linkers on Quantum Interference: A Linker Theorem,” *The Journal of Physical Chemistry C*, vol. 121, no. 27, pp. 14451–14462, Jul. 2017, doi: 10.1021/acs.jpcc.7b03493.

[199] B. Xu and N. J. Tao, “Measurement of single-molecule resistance by repeated formation of molecular junctions,” *Science (1979)*, vol. 301, no. 5637, pp. 1221–1223, Aug. 2003, doi: 10.1126/science.1087481.

[200] L. E. Ratcliff, S. Mohr, G. Huhs, T. Deutsch, M. Masella, and L. Genovese, “Challenges in large scale quantum mechanical calculations,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 7, no. 1, p. e1290, Jan. 2017, doi: 10.1002/wcms.1290.

[201] G. Peng, M. Strange, K. S. Thygesen, and M. Mavrikakis, “Conductance of conjugated molecular wires: Length dependence, anchoring groups, and band alignment,” *Journal of Physical Chemistry C*, vol. 113, no. 49, pp. 20967–20973, 2009, doi: 10.1021/jp9084603.

[202] X. J. Lu and W. K. Olson, “3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures,” *Nucleic Acids Research*, vol. 31, no. 17, pp. 5108–5121, Sep. 2003, doi: 10.1093/nar/gkg680.

[203] D. Porath, N. Lapidot, and J. Gomez-herrero, “Charge transport in DNA-based devices,” *Lecture Notes in Physics*, vol. 680. Springer, Berlin, Heidelberg , pp. 411–444, 2006. doi: 10.1007/3-540-31514-4_16.

[204] J. Qi, M. G. Rabbani, S. Edirisinghe, and M. P. Anantram, “Transport of charge in DNA heterostructures,” *Proceedings of the IEEE Conference on Nanotechnology*, no. 1, pp. 487–491, 2011, doi: 10.1109/NANO.2011.6144620.

[205] C. M. Niemeyer, “Nanoparticles, Proteins, and Nucleic Acids: Biotechnology Meets Materials Science,” *Angewandte Chemie International Edition*, vol. 40, no. 22, pp. 4128–4158, Nov. 2001, doi: 10.1002/1521-3773(20011119)40:22<4128::AID-ANIE4128>3.0.CO;2-S.

[206] C. M. Niemeyer, “DNA as a Material for Nanotechnology,” *Angewandte Chemie International Edition in English*, vol. 36, no. 6, pp. 585–587, Apr. 1997, doi: 10.1002/anie.199705851.

[207] E. di Mauro and C. P. Hollenberg, “DNA technology in chip construction,” *Advanced Materials*, vol. 5, no. 5, pp. 384–386, May 1993, doi: 10.1002/adma.19930050512.

[208] T. Liedl, T. L. Sobey, and F. C. Simmel, “DNA-based nanodevices,” *Nano Today*, vol. 2, no. 2, pp. 36–41, Apr. 2007, doi: 10.1016/S1748-0132(07)70057-9.

[209] A. Landi, A. Capobianco, and A. Peluso, “Coherent Effects in Charge Transport in Molecular Wires: Toward a Unifying Picture of Long-Range Hole Transfer in DNA,” *The Journal of Physical Chemistry Letters*, vol. 11, no. 18, pp. 7769–7775, Sep. 2020, doi: 10.1021/acs.jpclett.0c01996.

[210] L. Xiang, J. L. Palma, C. Bruot, V. Mujica, M. A. Ratner, and N. Tao, “Intermediate tunnelling-hopping regime in DNA charge transport,” *Nature Chemistry*, vol. 7, no. 3, pp. 221–226, 2015, doi: 10.1038/nchem.2183.

[211] T. Liu and J. K. Barton, “DNA Electrochemistry through the Base Pairs Not the Sugar−Phosphate Backbone,” *J Am Chem Soc*, vol. 127, no. 29, pp. 10160–10161, Jul. 2005, doi: 10.1021/ja053025c.

[212] Y. Li, J. M. Artés, and J. Hihath, “Long-Range Charge Transport in Adenine-Stacked RNA:DNA Hybrids,” *Small*, vol. 12, no. 4, pp. 432–437, Jan. 2016, doi: 10.1002/smll.201502399.

[213] P. Xie, K. Liu, F. Gu, and Y. Aoki, “Counter-ion effects of A- and B-type poly(dG)·Poly(dC) and poly(dA)·Poly(dT) DNA by elongation method,” *International Journal of Quantum Chemistry*, vol. 112, no. 1, pp. 230–239, Jan. 2012, doi: 10.1002/qua.23230.

[214] R. Zhuravel *et al.*, “Backbone charge transport in double-stranded DNA,” *Nature Nanotechnology*, vol. 15, no. 10, pp. 836–840, Oct. 2020, doi: 10.1038/s41565-020-0741-2.

[215] X. H. Li, Y. Y. Zhang, T. Liu, and K. L. Wang, “Further Discussion on Polaron Existence in Dry DNA,” *Chinese Physics Letters*, vol. 26, no. 12, p. 128701, Dec. 2009, doi: 10.1088/0256-307X/26/12/128701.

[216] M. A. Abolghassemi Fakhree, D. R. Delgado, F. Martínez, and A. Jouyban, “The Importance of Dielectric Constant for Drug Solubility Prediction in Binary Solvent Mixtures: Electrolytes and Zwitterions in Water + Ethanol,” *AAPS PharmSciTech*, vol. 11, no. 4, pp. 1726–1729, Dec. 2010, doi: 10.1208/s12249-010-9552-3.

[217] W. W. Parson, “Effects of Free Energy and Solvent on Rates of Intramolecular Electron Transfer in Organic Radical Anions,” 2017, doi: 10.1021/acs.jpca.7b08579.

[218] G. Cuniberti, L. Craco, D. Porath, and C. Dekker, “Backbone-induced semiconducting behavior in short DNA wires,” *Physical Review B*, vol. 65, no. 24, p. 241314, Jun. 2002, doi: 10.1103/PhysRevB.65.241314.

[219] H. Wang, T. E. Cheatham, P. M. Gannett, and J. P. Lewis, “Differential electronic states observed during A–B DNA duplex conformational transitions,” *Soft Matter*, vol. 5, no. 3, pp. 685–690, Jan. 2009, doi: 10.1039/B800462E.

[220] M. Taniguchi and T. Kawai, “Electronic structures of A- and B-type DNA crystals,” *Physical Review E*, vol. 70, no. 1, p. 011913, Jul. 2004, doi: 10.1103/PhysRevE.70.011913.

[221] A. Capobianco *et al.*, “Delocalized Hole Domains in Guanine-Rich DNA Oligonucleotides,” *The Journal of Physical Chemistry B*, vol. 119, no. 17, pp. 5462–5466, Apr. 2015, doi: 10.1021/acs.jpcb.5b02940.

[222] S. Afsari, L. E. Korshoj, G. R. Abel, S. Khan, A. Chatterjee, and P. Nagpal, “Quantum Point Contact Single-Nucleotide Conductance for DNA and RNA Sequence Identification,” *ACS Nano*, vol. 11, no. 11, pp. 11169–11181, 2017, doi: 10.1021/acsnano.7b05500.

[223] K. M. Koo, A. A. I. Sina, L. G. Carrascosa, M. J. A. Shiddiky, and M. Trau, “DNA-bare gold affinity interactions: Mechanism and applications in biosensing,” *Analytical Methods*, vol. 7, no. 17. Royal Society of Chemistry, pp. 7042–7054, Sep. 07, 2015. doi: 10.1039/c5ay01479d.

[224] Z. E. Hughes, G. Wei, K. L. M. Drew, L. Colombi Ciacchi, and T. R. Walsh, “Adsorption of DNA Fragments at Aqueous Graphite and Au(111) via Integration of Experiment and Simulation,” *Langmuir*, vol. 33, no. 39, pp. 10193–10204, Oct. 2017, doi: 10.1021/acs.langmuir.7b02480.

APPENDIX: CODE DEVELOPMENT

The MATLAB codes developed in this work.

1. Code for generating the input file

```
%%% Reads gjf file, extracts the structure atoms and
%%% generates the number of orbitals and number of electrons per atom based
%%% on the (b3lyp/6-31g(d,p)) basis set. Then finds the sum_Orbitals &
%%% sum_electrons. This code assumes no Na (outputs # of Phosphates to get
%%% correct HOMO eigenvalue) --> set num_P = 0 if Na = # Phosphates
%%% Different Buttiker Probes values per group of
%%% atoms can be set (see bProbe variable)
%%%
%%% Parameters GammaL, GammaR, and Lambda can each be given multiple values.
%%% Each combination of parameters generates a unique parameters .txt file.
clear
clc

%%%% User Input: Set Run Parameters  %%%%
%%%% IMPORTANT: If BB+Base partitioning, make sure to have Inject and
%%%% Extract Sites to correspond to the full BB+Base.

structure_name = 'BTCT1';
energy_range = [-6:0.001:0];
InjectSite =[1:33] ;
ExtractSite =[192:225];
lambda = [50]/1000;             %for Exp-decay: Gamma_BP * exp[-(E-
eig)/lambda]
GammaL = [0.1 0.5 1];            % eV
GammaR = [0.1 0.5 1];            % eV
Broadening = 0;          % eV (for DOS)
mix_contact_coupling = false;

%%% Manual Input
%%%%% Last element is always kept 0
%%%%% Last cell block is bprobe value for the remaining atoms (specified by
0)
bProbe = {100e-3 0}; % manual input { eV atoms; eV atoms;..; eV 0 }
partitioning = 0;   %if 1: create BB+Base partitioning file, otherwise no
file and transport code assumes atomic blocking

manual_input = 1;  % set to 1 for bProbe defined above, 0: automated (uses
pdb file for BB and Base atoms)
%%% If the manual_input variable is 0, then
%%% Automated BP: Base & BB are extracted from pdb file
%%% Enter Decoherence per Nucleotide in this order: [G C A T(U)]
BP_base = [0.1 0.1 0.1 0.1];
BP_bb = [0.1 0.1 0.1 0.1];
%% Import data from gjf
Name1=[structure_name '.gjf'];
```

```
Name2=[structure_name '.com'];
if exist(Name1, 'file') == 2
filename=Name1;
elseif exist(Name2, 'file') == 2
    filename=Name2;
else
    disp('Gaussian File Not Found')
end


startRow = 2;formatSpec = '%2s%[^\n\r]';fileID = fopen(filename,'r');
textscan(fileID, '%[^\n\r]', startRow-1, 'WhiteSpace', '', 'ReturnOnError',
false, 'EndOfLine', '\r\n');
dataArray = textscan(fileID, formatSpec, 'Delimiter', '', 'WhiteSpace', '',
'ReturnOnError', false);
dataArray{1} = strtrim(dataArray{1});fclose(fileID);
read_loc = contains(dataArray{2},'(PDBName=');
%%%%%%%% If gjf file does not contain (PDBName=) format:
if ~read_loc
      for LL=1:10
        Lngth(LL) = length(dataArray{2}{LL});
    end
      start_loc = find(diff((diff(Lngth)==0))==1,1,'last')+1;

      for LL=length(dataArray{2})-5:length(dataArray{2})
        Lngth(LL) = length(dataArray{2}{LL});
      end
    end_loc = find((diff(Lngth)<0), 1, 'last' );    %assumes last line
contains text that is different from coordinates of last atom

    read_loc = zeros(length(dataArray{2}),1);
    read_loc(start_loc:end_loc)=1;
    read_loc=logical(read_loc);
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Atoms = dataArray{:, 1};
Atoms = Atoms(read_loc);
% clearvars filename startRow formatSpec fileID dataArray ans;
%% Generating Orbital Numbers per Atom (H,O,C,N,P,X=Na)
Atoms = cell2mat(Atoms);
num_P=0;    %number of Phosphates, to find net charge (assuming no Na)
Orbitals=nan(length(Atoms),1);    %-1 because last value of Atoms is removed
('eps')
electrons=Orbitals;
%first and last values of Atoms variable are removed
for i=1:length(Atoms)
        switch Atoms(i)
            case 'H'
               Orbitals(i)=5;
               electrons(i)=1;
            case 'O'
               Orbitals(i)=15;
               electrons(i)=8;
            case 'C'
               Orbitals(i)=15;
```

```
               electrons(i)=6;
            case 'N'
               Orbitals(i)=15;
               electrons(i)=7;
            case 'P'
               Orbitals(i)=19;
               electrons(i)=15;
               num_P = num_P+1; %number of Phosphates, to find net charge
(assuming no Na)
            case 'X'        %X=Na, rename in gjf file
               Orbitals(i)=19;
               electrons(i)=11;
        end
end

%% Setting up BProbe Values and their Corresponding Atoms %%
if manual_input == 0
%%% Automated Buttiker Probes
 [BB,Base,Nucleotide]=bb_base_extract(structure_name); %extracts info from
PDB file

 Nucleotide = strrep(Nucleotide,'ANT','1'); %Assuming Anthraquinone (ANT in
pdb) has high decoherence
 Nucleotide = strrep(Nucleotide,'G','1');
 Nucleotide = strrep(Nucleotide,'C','2');
 Nucleotide = strrep(Nucleotide,'A','3');
 Nucleotide = strrep(Nucleotide,'T','4');
 Nucleotide = strrep(Nucleotide,'U','4');

 Nucleotide = str2double(Nucleotide);


ProbeSite=setdiff(1:length(Atoms),[InjectSite ExtractSite]);
bProbe=100*ones(1,length(ProbeSite));   %initialization, after the for loop
no element is 100

if length(Base)==length(BB)
    for ii=1:length(Base)
        [~,b]=intersect(ProbeSite,Base{ii}); %get intersect location of Base
in ProbeSite
        bProbe(b)=BP_base(Nucleotide(ii));
        [~,b]=intersect(ProbeSite,BB{ii}); %get intersect location of BB in
ProbeSite
        bProbe(b)=BP_bb(Nucleotide(ii));
    end
else %Sometimes (if Aq) Base>BB
        for ii=1:length(Base)
        [~,b]=intersect(ProbeSite,Base{ii}); %get intersect location of Base
in ProbeSite
        bProbe(b)=BP_base(Nucleotide(ii));

        end
        BB{length(Base)} = [];  %for partitioning (Base>BB)

    for ii=1:length(BB)
```

```
        [~,b]=intersect(ProbeSite,BB{ii}); %get intersect location of BB in
ProbeSite
        bProbe(b)=BP_bb(Nucleotide(ii));
    end
end

 %%% BB+Base partitioning
 for k = 1:length(Base)
    Partition{k} = sort([BB{k} Base{k}]);
 end

else
    a=bProbe;
    clearvars bProbe
    ProbeSite=setdiff(1:length(Atoms),[InjectSite ExtractSite]);
    bProbe=a{end,1}*ones(1,length(ProbeSite));
    for i=1:size(a,1)-1
    [~,b]=intersect(ProbeSite,a{i,2}); %get intersect location in ProbeSite
    if isempty(b)
    disp('incorrect bprobes input format, please make sure only last element
is 0')
    end
    bProbe(b)=a{i,1};
    end
%%%%%%%%%%%%%%%%%%%%%


end

%% Generating the Parameters file %%
mkdir Parameters_Folder
if mix_contact_coupling
for u = 1:1:length(lambda)
    for i = 1:1:length(GammaL)
        for j = 1:1:length(GammaR)
            param_file_index = j + (i-1)*length(GammaL) + (u-
1)*length(GammaL)*length(GammaR)
            param_filename = sprintf("Parameters%d.txt",param_file_index)

            fileID = fopen(strcat('Parameters_Folder/',param_filename),'w');
            fprintf(fileID, [structure_name '\r\n']);
            fprintf(fileID,'%f\n',[num_P sum(electrons)]);
            fprintf(fileID, 'Orbitals set\r\n');
            fprintf(fileID,'%f\n',Orbitals);
            fprintf(fileID, '\r\nEnergy Range\r\n');
            fprintf(fileID,'%f\n',energy_range);
            fprintf(fileID, '\r\nInject Site (atoms number)\r\n');
            fprintf(fileID,'%f\n',InjectSite);
            fprintf(fileID, '\r\nExtract Site (atoms number)\r\n');
            fprintf(fileID,'%f\n',ExtractSite);
            fprintf(fileID, '\r\nGammaL\r\n');
            fprintf(fileID,'%f\n',GammaL(i));
            fprintf(fileID, '\r\nGammaR\r\n');
            fprintf(fileID,'%f\n',GammaR(j));
            fprintf(fileID, '\r\nProbes Site (atoms number)\r\n');
```

```
            fprintf(fileID,'%f\n',ProbeSite);
            fprintf(fileID, '\r\nBroadening (for DOS)\r\n');
            fprintf(fileID,'%f\n',Broadening);
            fprintf(fileID, '\r\nLambda (exp-decay)\r\n');
            fprintf(fileID,'%f\n',lambda(u));
            fprintf(fileID, '\r\nProbe (for Decoh)\r\n');
            fprintf(fileID,'%f\n',bProbe);
            fclose(fileID);
            disp('Parameters file is ready!')

            if partitioning ==1
            %%%%% Create Partitioning File (e.g. BB+Base)
            fileID = fopen([structure_name '_partitioning.txt'],'w');
            for k=1:length(Partition)
            fprintf(fileID,'%3G,',Partition{k});
            fprintf( fileID, '\r\n\r\n' ) ;
            end

            fclose(fileID);
            disp('Partitioning file is ready!')
            end
        end
    end
end

else
counter = 0;
    for u = 1:length(lambda)
    for i = 1:length(GammaL)
            param_file_index = counter + 1;
            counter = counter + 1;
            param_filename = sprintf("Parameters%d.txt",param_file_index);

            fileID = fopen(strcat('Parameters_Folder/',param_filename),'w');
            fprintf(fileID, [structure_name '\r\n']);
            fprintf(fileID,'%f\n',[num_P sum(electrons)]);
            fprintf(fileID, 'Orbitals set\r\n');
            fprintf(fileID,'%f\n',Orbitals);
            fprintf(fileID, '\r\nEnergy Range\r\n');
            fprintf(fileID,'%f\n',energy_range);
            fprintf(fileID, '\r\nInject Site (atoms number)\r\n');
            fprintf(fileID,'%f\n',InjectSite);
            fprintf(fileID, '\r\nExtract Site (atoms number)\r\n');
            fprintf(fileID,'%f\n',ExtractSite);
            fprintf(fileID, '\r\nGammaL\r\n');
            fprintf(fileID,'%f\n',GammaL(i));
            fprintf(fileID, '\r\nGammaR\r\n');
            fprintf(fileID,'%f\n',GammaR(i));
            fprintf(fileID, '\r\nProbes Site (atoms number)\r\n');
            fprintf(fileID,'%f\n',ProbeSite);
            fprintf(fileID, '\r\nBroadening (for DOS)\r\n');
            fprintf(fileID,'%f\n',Broadening);
            fprintf(fileID, '\r\nLambda (exp-decay)\r\n');
            fprintf(fileID,'%f\n',lambda(u));
            fprintf(fileID, '\r\nProbe (for Decoh)\r\n');
```

```
            fprintf(fileID,'%f\n',bProbe);
            fclose(fileID);
            disp('Parameters file is ready!')

            if partitioning ==1
            %%%%% Create Partitioning File (e.g. BB+Base)
            fileID = fopen([structure_name '_partitioning.txt'],'w');
            for k=1:length(Partition)
            fprintf(fileID,'%3G,',Partition{k});
            fprintf( fileID, '\r\n\r\n' ) ;
            end

            fclose(fileID);
            disp('Partitioning file is ready!')
            end
        end
        end
end
```

2. *Code for extracting backbone and base atoms*

```
function [BB,Base,Nucleotide]=bb_base_extract(structure_name)
%%% This function generates bases from PDB file
%%% Note: Bases should be numbered sequentially (edit in notepad++ if not
sequential)
Name=[structure_name '.pdb'];
filename=Name;
disp('Loading from PDB')
    %% Import data from PDB file.
% Read columns of data as text:
formatSpec = '%4s%7s%4s%1s%3s%1s%2s%4s%12s%8s%8s%6s%6s%s%[^\n\r]';
fileID = fopen(filename,'r');
dataArray = textscan(fileID, formatSpec, 'Delimiter', '', 'WhiteSpace', '',
'TextType', 'string',  'ReturnOnError', false);
fclose(fileID);
% Replace non-numeric text with NaN.
raw = repmat({''},length(dataArray{1}),length(dataArray)-1);
for col=1:length(dataArray)-1
    raw(1:length(dataArray{col}),col) = mat2cell(dataArray{col},
ones(length(dataArray{col}), 1));
end

AtomType = join([raw{1:end,3}; raw{1:end,4}],1)';   %Extracts AtomType (e.g.
H5' or C4 ..)
AtomType = erase(AtomType,' '); %Removes spaces

nucleotide = join([raw{1:end,5}; raw{1:end,6}],1)';   %Extracts Nucleotide
name (GCAT) [string variabale]
nucleotide = erase(nucleotide,' '); %Removes spaces


% a=cellfun(@(x) ismember(x,' '),nucleotide);
```

```
% if length(nucleotide(a))/length(nucleotide) > 0.5
%     nucleotide = raw(:,5);          %Sometimes extra space exists between
columns, need to shift to correct column
% end
numericData = NaN(size(dataArray{1},1),size(dataArray,2));
%%
for col=[2,4,6,8,9,10,11,12,13]
    % Converts text in the input cell array to numbers. Replaced non-numeric
    % text with NaN.
    rawData = dataArray{col};
    for row=1:size(rawData, 1)
        % Create a regular expression to detect and remove non-numeric
prefixes and
        % suffixes.
        regexstr = '(?<prefix>.*?)(?<numbers>([-
]*(\d+[\,]*)+[\.]{0,1}\d*[eEdD]{0,1}[-+]*\d*[i]{0,1})|([-
]*(\d+[\,]*)*[\.]{1,1}\d+[eEdD]{0,1}[-+]*\d*[i]{0,1}))(?<suffix>.*)';
        try
            result = regexp(rawData(row), regexstr, 'names');
            numbers = result.numbers;

            % Detected commas in non-thousand locations.
            invalidThousandsSeparator = false;
            if numbers.contains(',')
                thousandsRegExp = '^\d+?(\,\d{3})*\.{0,1}\d*$';
                if isempty(regexp(numbers, thousandsRegExp, 'once'))
                    numbers = NaN;
                    invalidThousandsSeparator = true;
                end
            end
            % Convert numeric text to numbers.
            if ~invalidThousandsSeparator
                numbers = textscan(char(strrep(numbers, ',', '')), '%f');
                numericData(row, col) = numbers{1};
                raw{row, col} = numbers{1};
            end
        catch
            raw{row, col} = rawData{row};
        end
    end
end
rawNumericColumns = raw(:, [2,4,6,8,9,10,11,12,13]);
rawStringColumns = string(raw(:, [1,3,5,7,14]));

R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),rawNumericColumns); % Find
non-numeric cells
rawNumericColumns(R) = {NaN}; % Replace non-numeric cells

for catIdx = [1,2,3,4,5]
    idx = (rawStringColumns(:, catIdx) == "<undefined>");
    rawStringColumns(idx, catIdx) = "";
end

%%% Create output variable
TABLE = table;
```

```
TABLE.NAME = rawStringColumns(:, 1);
TABLE.AtomNum = cell2mat(rawNumericColumns(:, 1));
% Cleaning file from TER and END rows (keeping only ATOM)
A = ismember(TABLE{:,1}, 'ATOM');
A = A(:,end);
AtomType = AtomType(A,1);
nucleotide = nucleotide(A,1);

%%% Locating Base Number (ResNum)
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),raw); % Find non-numeric
cells
raw(R) = {NaN}; % Replace non-numeric cells
PDB = cell2mat(raw);
clearvars formatSpec fileID dataArray ans raw col numericData rawData row
regexstr result numbers invalidThousandsSeparator thousandsRegExp me R;
Group_no=PDB(1:end,8);
Group_no(A==0)=[];
loc=find(diff(Group_no)~=0);
BB_Base = [1 loc(1)];
for i=2:length(loc)
BB_Base=[BB_Base loc(i-1)+1 loc(i)];
end
BB_Base = [BB_Base loc(end)+1 length(Group_no)];
%%% Extracting BB & Base atoms
for ii=1:2:length(BB_Base)
    range = BB_Base(ii):BB_Base(ii+1);
    bb=range(contains(AtomType(range),''''));
    bb = [bb range(contains(AtomType(range),'P'))];
    bb = [bb range(contains(AtomType(range),'H5T'))];    %Terminal Nucleotide
in CHARMM
    bb = [bb range(contains(AtomType(range),'H3T'))];    %Terminal Nucleotide
in CHARMM

    BB{ii} = sort(bb);
    Base{ii} = setxor(range,BB{ii});
    Nucleotide(ii)=nucleotide(range(2));    %any value within range is ok
end
%%
 BB = BB(~cellfun('isempty',BB));
 Base = Base(~cellfun('isempty',Base));
 Nucleotide = rmmissing(Nucleotide);

 %%%%% If 'CHARMM', Nucleotide Names are different
 Nucleotide = strrep(Nucleotide,'CYT','C') ;   %CYT = C
 Nucleotide = strrep(Nucleotide,'THY','T') ;   %THY = T
 Nucleotide = strrep(Nucleotide,'GUA','G') ;   %GUA = G
 Nucleotide = strrep(Nucleotide,'ADE','A') ;   %ADE = A
 Nucleotide = strrep(Nucleotide,'URA','U') ;   %URA = U

%%%%%% Cleaning
 Nucleotide = strrep(Nucleotide,'D','') ;
 Nucleotide = strrep(Nucleotide,' ','') ;
 Nucleotide = regexprep(Nucleotide,'\d','') ;
end
```

3. *Code for main function to calculate transmission*

```
%%% fpath is the path to the functions directory, e.g.:
%%% '/gscratch/stf/hashemm/EdepBP/functions/'
%%% mpath is the path to the matrices dir, e.g.:
%%% '/gscratch/stf/hashemm/EdepBP/matrices/'

function main_code(run_num,Edep,DOS,fpath,mpath,Vbias,linVdrop,Vprcnt)
if nargin < 8
Vprcnt = 0.4;
end

addpath(fpath);

if DOS ==1
        dosBlock_calc(run_num,mpath,Edep,Vbias,linVdrop,Vprcnt)
else
    if Edep ==1
        DNATransmission_Decoherence_EBP(run_num,mpath,Vbias,linVdrop,Vprcnt)
    else

DNATransmission_Decoherence_Eindep(run_num,mpath,Vbias,linVdrop,Vprcnt)
    end

end
```

4. *Code for DOS calculation*

```
%DOS Calculation for the Molecule
%Note1: if GammaL=GammaR = 0, code assigns bprobe values instead (assumes
%bprobe values for all locations).
%Note2: this code uses fine energy mesh (6000 points) from Energy(1) to
%Energy (end), see line 90.

function  dosBlock_calc(run_num,mpath,Edep,Vbias,linVdrop,Vprcnt)
%% Initialization (Loads Parameters from subdir run_num)
format long

apath = strcat(pwd,'/');
addpath(apath)

d=['run' num2str(run_num)];
dir_path=  strcat(apath,d,'/');    %loads subdirectory
addpath(dir_path)
workdir = dir_path;

FNAME = dir([workdir '/' 'Parameters*.txt']);
x=fopen(FNAME.name);
dataArray=textscan(x,'%s','WhiteSpace','\r\n');
loc1=find(~cellfun(@isempty,strfind(dataArray{1,1},'Energy')));
strand=char(dataArray{1,1}(1));
num_P = dataArray{1,1}(2);
num_P = str2double(cell2mat(num_P));
```

```
sum_e = dataArray{1,1}(3);
sum_e = str2double(cell2mat(sum_e));

Orbitals=dataArray{1,1}(5:loc1-1);
Orbitals=cellfun(@str2num,Orbitals)';

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Inject')));
Energy=dataArray{1,1}(loc1+1:loc2-1);
Energy=cellfun(@str2num,Energy)';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Extract')));
Lsite=dataArray{1,1}(loc1+1:loc2-1);
Lsite=cellfun(@str2num,Lsite)';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'GammaL')));
Rsite=dataArray{1,1}(loc1+1:loc2-1);
Rsite=cellfun(@str2num,Rsite)';

loc1=loc2;
loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'GammaR')));
gammaL=dataArray{1,1}(loc1+1:loc2-1);
gammaL=str2double(cell2mat(gammaL))';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Probes')));
gammaR=dataArray{1,1}(loc1+1:loc2-1);
gammaR=str2double(cell2mat(gammaR));
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Broadening')));
Dsites=dataArray{1,1}(loc1+1:loc2-1);
Dsites=cellfun(@str2num,Dsites)';

lambda=dataArray{1,1}(loc2+3);
lambda=str2double(cell2mat(lambda));

bprobe=dataArray{1,1}(loc2+5:end);
bprobe=cellfun(@str2num,bprobe)';

fclose(x);
clearvars loc1 loc2 dataArray x
%%
%%%%%%%%%%%%%Load Matrices%%%%%%%%%%%%%%%%%%%%%%%%%%%
Fm = ['load ',mpath, strand, '.mat'];
eval(Fm);
H0 = eval(strand);
    if linVdrop == 1
        H0 = bias_drop(workdir,strand,H0,Orbitals,Vbias);
    else
        H0 = bias_drop_nonlin(workdir,strand,H0,Orbitals,Vbias,Vprcnt);
    end
```

```
[H,Lsite,Rsite,Dsites,Orbitals,HOMO] =
partitionH(H0,Lsite,Rsite,Dsites,Orbitals,num_P,sum_e,workdir);    %U'H0U, H-
blocking [diag(H)=EV(block)]
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
sizeH = size(H, 1);  % size of Hamiltonian
disp(['Size of the Hamiltonian = ' num2str(sizeH)])
%%%%%%%%%%Using Fine Energy Grid%%%%%%%%%%%%%%%%%%
Energy = Energy(1) :(Energy(end)-Energy(1))/6000: Energy(end);
eta=0;  %defining eta as zero, broadening from BP
Nsite = length([Lsite Rsite Dsites]);

switch Edep
    case 1  %Edep
%% E-dep Decoherence Loop Initialization %%
NE = length(Energy);
%%%%%%%% Check available files from Checkpoint %%%%%%%%%%%
Dname=strcat(workdir,strand,'_Edep_DOS_Vbias',num2str(Vbias),'.mat');
mat=dir(Dname);

if length(mat)~=0
load(mat.name);
qq=find(DOS~=-1, 1, 'last' )+1;
else
    qq=1;
    DOS = -1*ones(1, NE);
    DOSBlock = -1*ones(Nsite, NE);
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Entering the loop to calculate E-dep DOS %%
EV = diag(H);
for nE = qq : NE
    nE
    E = Energy(nE)
%%%%%%%%%set sumSig %%%%%%%%%%%%%%%%%%%%%
GAMMA = zeros(sizeH, sizeH);
%%%%%%%%%%% Contacts Coupling %%%%%%%%%%%%%%
sites = [Lsite Rsite];
if any(gammaL)
%If contact coupling is included for DOS calculation
gamma = [gammaL*ones(1,length(Lsite)) gammaR*ones(1,length(Rsite))];

for ii = 1 : length(sites)
    isite = sites(ii);
    %%%% locating cooridnates in H corresponding to isite
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    GAMMA(range,range) = gamma(ii) * eye(Len);
end

%%%%%%%%%%% Buttiker Probes %%%%%%%%%%%%%%
sites = Dsites;
gamma = bprobe;
else
```

```
%If no contact coupling, apply bprobe value for all atoms
    sites = [Lsite Dsites Rsite];
    gammaL = bprobe(1);
    gammaR = gammaL;
    gamma = [gammaL*ones(1,length(Lsite)) bprobe gammaR*ones(1,length(Rsite))
];
end

for ii = 1 : length(sites)
    isite = sites(ii);
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    %exp decay per Atomic/Molecular Orbital (EV of H{block(ii)})
    f_bp= exp(-abs(E-EV(range))/lambda);
    %exp decay per Atomic/Molecular Orbital (HOMO cutoff)
    range2 = EV(range) > HOMO(1) & EV(range) <= mean(HOMO);
    f_bp(range2) = exp(-abs(E-HOMO(1))/lambda);
   %exp decay per Atomic/Molecular Orbital (LUMO cutoff)
    range2 = EV(range) < HOMO(2) & EV(range) > mean(HOMO);
    f_bp(range2) = exp(-abs(E-HOMO(2))/lambda);

    f_bp(f_bp<10^-20)=0; %cutoff for BP=0 is <10^-20

    GAMMA(range,range) = gamma(ii) * f_bp .* eye(Len);
end



sumSig = -1i * GAMMA / 2;

    Gr = ((E + 1i * eta) * eye(sizeH) - H - sumSig) \ eye(sizeH);

%%%%%%%%%%%%%%%%%%% DOS %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
     tempM = imag(diag(Gr));
     DOS(nE)=-sum(tempM) / pi;

     t1 = 1;
     t2 = Orbitals(1);

     for nOrbitals = 1 : Nsite
         DOSBlock(nOrbitals, nE) = -sum(tempM(t1 : t2)) / pi;
         if nOrbitals < Nsite
            t1 = t1 + Orbitals(nOrbitals);
            t2 = t2+Orbitals(nOrbitals + 1);
         end
     end
%%%%%%% Save after every iteration %%%%
save(Dname,'Energy','DOS','DOSBlock');
end

    otherwise % Eindep
        %% E-indep Decoherence Loop Initialization %%
NE = length(Energy);
```

```
%%%%%%%% Check available files from Checkpoint %%%%%%%%%%%
Dname=strcat(workdir,strand,'_Eindep_DOS_Vbias',num2str(Vbias),'.mat');

mat=dir(Dname);

if length(mat)~=0
load(mat.name);
qq=find(DOS~=-1, 1, 'last' )+1;
else
    qq=1;
    DOS = -1*ones(1, NE);
    DOSBlock = -1*ones(Nsite, NE);
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Entering the loop to calculate E-indep DOS %%
EV = diag(H);
for nE = qq : NE
    nE
    E = Energy(nE)
%%%%%%%%%set sumSig %%%%%%%%%%%%%%%%%%%%%
GAMMA = zeros(sizeH, sizeH);
%%%%%%%%%%% Contacts Coupling %%%%%%%%%%%%%%
sites = [Lsite Rsite];
if any(gammaL)
gamma = [gammaL*ones(1,length(Lsite)) gammaR*ones(1,length(Rsite))];

for ii = 1 : length(sites)
    isite = sites(ii);
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    GAMMA(range,range) = gamma(ii) * eye(Len);
end

%%%%%%%%%%% Buttiker Probes %%%%%%%%%%%%%%
sites = Dsites;
gamma = bprobe;
else
    sites = [Lsite Dsites Rsite];
    gammaL = bprobe(1);
    gammaR = gammaL;    %assuming the same bprobe value for all atoms
    gamma = [gammaL*ones(1,length(Lsite)) bprobe gammaR*ones(1,length(Rsite))
];
end

for ii = 1 : length(sites)
    isite = sites(ii);
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    f_bp = 1;

    GAMMA(range,range) = gamma(ii) * f_bp .* eye(Len);
```

```
end



sumSig = -1i * GAMMA / 2;

    Gr = ((E + 1i * eta) * eye(sizeH) - H - sumSig) \ eye(sizeH);

%%%%%%%%%%%%%%%%%%% DOS %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
     tempM = imag(diag(Gr));
     DOS(nE)=-sum(tempM) / pi;

     t1 = 1;
     t2 = Orbitals(1);

     for nOrbitals = 1 : Nsite
         DOSBlock(nOrbitals, nE) = -sum(tempM(t1 : t2)) / pi;
         if nOrbitals < Nsite
            t1 = t1 + Orbitals(nOrbitals);
            t2 = t2+Orbitals(nOrbitals + 1);
         end
     end
%%%%%%% Save after every iteration %%%%
save(Dname,'Energy','DOS','DOSBlock');
end
end
%% Fin

disp('Finished DOS!')
```

5. *Code to calculate transmission with energy-independent decoherence*

```
function
DNATransmission_Decoherence_Eindep(run_num,mpath,Vbias,linVdrop,Vprcnt)
%% Initialization (Loads Parameters from subdir run_num)
format long

apath = strcat(pwd,'/');
addpath(apath)

d=['run' num2str(run_num)];
dir_path=  strcat(apath,d,'/');    %loads subdirectory
addpath(dir_path)
workdir = dir_path;

FNAME = dir([workdir '/' 'Parameters*.txt']);
x=fopen(FNAME.name);
dataArray=textscan(x,'%s','WhiteSpace','\r\n');
loc1=find(~cellfun(@isempty,strfind(dataArray{1,1},'Energy')));
strand=char(dataArray{1,1}(1));
num_P = dataArray{1,1}(2);
num_P = str2double(cell2mat(num_P));
```

```
sum_e = dataArray{1,1}(3);
sum_e = str2double(cell2mat(sum_e));

Orbitals=dataArray{1,1}(5:loc1-1);
Orbitals=cellfun(@str2num,Orbitals)';

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Inject')));
Energy=dataArray{1,1}(loc1+1:loc2-1);
Energy=cellfun(@str2num,Energy)';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Extract')));
Lsite=dataArray{1,1}(loc1+1:loc2-1);
Lsite=cellfun(@str2num,Lsite)';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'GammaL')));
Rsite=dataArray{1,1}(loc1+1:loc2-1);
Rsite=cellfun(@str2num,Rsite)';

loc1=loc2;
loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'GammaR')));
gammaL=dataArray{1,1}(loc1+1:loc2-1);
gammaL=str2double(cell2mat(gammaL))';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Probes')));
gammaR=dataArray{1,1}(loc1+1:loc2-1);
gammaR=str2double(cell2mat(gammaR));
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Broadening')));
Dsites=dataArray{1,1}(loc1+1:loc2-1);
Dsites=cellfun(@str2num,Dsites)';

eta=dataArray{1,1}(loc2+1);
eta=str2double(cell2mat(eta));

lambda=dataArray{1,1}(loc2+3);
lambda=str2double(cell2mat(lambda));

bprobe=dataArray{1,1}(loc2+5:end);
bprobe=cellfun(@str2num,bprobe)';

fclose(x);
clearvars loc1 loc2 dataArray x

%%%%%%%%%%%%Prepare Hamiltonian%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Load the matrix file as a struct in case filename is different from
%%% variable name.
eval(['Fm = load(''',mpath, strand, '.mat'')']);
temp = fieldnames(Fm);
H0 = eval(['Fm.' temp{:} ';']);
```

```
    if linVdrop == 1
        H0 = bias_drop(workdir,strand,H0,Orbitals,Vbias);
    else
        H0 = bias_drop_nonlin(workdir,strand,H0,Orbitals,Vbias,Vprcnt);
    end

[H,Lsite,Rsite,Dsites,Orbitals,~] =
partitionH(H0,Lsite,Rsite,Dsites,Orbitals,0,0,workdir);    %U'H0U, H-blocking
[diag(H)=EV(block)]
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
eta = 0;
sizeH = size(H, 1);  % size of Hamiltonian
disp(['Size of the Hamiltonian = ' num2str(sizeH)])

%%
Norb = sum(Orbitals);    %total size of H,total number of orbitals
if(Norb ~= sizeH)
    disp(['sum(Orbitals) = ' num2str(Norb)])
    disp('Total size of H =/= number of Orbitals, please check your input
files')
    return;
end
Nb = length(Dsites);
Nsite = length(Orbitals);   %total number of atoms


%% Loop Initialization %%
NE = length(Energy);
%%%prepare the output name before entering the loop%%%
Tname=strcat(workdir,'Teff_',strand,'_gammaL_',num2str(gammaL),'_gammaR_',num
2str(gammaR),'_bias_',num2str(mean(Vbias)),'_',num2str(run_num),'_Eindep','.m
at');

%%%%%%%% Check available files from Checkpoint %%%%%%%%%%%
mat=dir(Tname);

if length(mat)~=0
load(mat.name);
qq=find(T~=-1, 1, 'last' )+1;
else
    T=-1*ones(1,NE);
    qq=1;
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Entering the loop
EV = diag(H);
for nE = qq : NE
    nE
    E = Energy(nE)
%%%%%%%%%set sumSig %%%%%%%%%%%%%%%%%%%%%
GAMMA = zeros(sizeH, sizeH);
%%%%%%%%%%% Contacts Coupling %%%%%%%%%%%%%%
sites = [Lsite Rsite];
gamma = [gammaL*ones(1,length(Lsite)) gammaR*ones(1,length(Rsite))];
```

```
for ii = 1 : length(sites)
    isite = sites(ii);
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    GAMMA(range,range) = gamma(ii) * eye(Len);
end
%%%%%%%%%%% Buttiker Probes %%%%%%%%%%%%%%
sites = Dsites;
gamma = bprobe;

for ii = 1 : length(sites)
    isite = sites(ii);
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    f_bp= 1;

    GAMMA(range,range) = gamma(ii) * f_bp .* eye(Len);
end

%% Checking BP(AOi/MOi) = 0 to update Orbitals and Dsites and remove from
decoherent transport calculations
a=diag(GAMMA);
ao_nz = find(a~=0); %locates AO(MO) with BP=/=0
%transform orbital into its atom number( or block number, depending on the
%partitioning scheme)
idx = zeros(size(ao_nz));
z= cumsum(Orbitals);
    for ii=1:length(ao_nz)
        k=1;
            while ao_nz(ii)>z(k)
                    k=k+1;
            end
        idx(ii)=k;
    end

trunc_Dsites = setdiff(unique(idx),[Lsite Rsite])';
if ~isempty(trunc_Dsites)
ao_nz = [ao_nz idx]; %add corresponding atom numbers (or block number) per nz
orbital

%% Preparing Gr for Transport Calcuations %%

sumSig = -1i * GAMMA / 2;

    Gr = ((E + 1i * eta) * eye(sizeH) - H - sumSig) \ eye(sizeH);
    Ga = Gr';

T(nE) = decoherent_transport_f(Gr,Ga,GAMMA,Lsite,Rsite,trunc_Dsites,ao_nz);

else %Coherent Transport if all BP = 0
```

```
    sumSig = -1i * GAMMA / 2;
    Gr = ((E + 1i * eta) * eye(sizeH) - H - sumSig) \ eye(sizeH);
    Ga = Gr';

    T(nE) = coherent_transport_f(Gr,Ga,Lsite,Rsite,Orbitals,gammaL,gammaR);
end

%%%%%%% Save after every iteration %%%%
save(Tname,'Energy','T')
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
end


disp('Finished Decoherent Transmission!')
```

6. Code to calculate transmission with energy-dependent decoherence

```
function
DNATransmission_Decoherence_EBP(run_num,mpath,Vbias,linVdrop,Vprcnt)
%% Initialization (Loads Parameters from subdir run_num)
format long

apath = strcat(pwd,'/');
addpath(apath)

d=['run' num2str(run_num)];
dir_path=  strcat(apath,d,'/');    %loads subdirectory
addpath(dir_path)
workdir = dir_path;

FNAME = dir([workdir '/' 'Parameters*.txt']);
x=fopen(FNAME.name);
dataArray=textscan(x,'%s','WhiteSpace','\r\n');
loc1=find(~cellfun(@isempty,strfind(dataArray{1,1},'Energy')));
strand=char(dataArray{1,1}(1));
num_P = dataArray{1,1}(2);
num_P = str2double(cell2mat(num_P));

sum_e = dataArray{1,1}(3);
sum_e = str2double(cell2mat(sum_e));

Orbitals=dataArray{1,1}(5:loc1-1);
Orbitals=cellfun(@str2num,Orbitals)';

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Inject')));
Energy=dataArray{1,1}(loc1+1:loc2-1);
Energy=cellfun(@str2num,Energy)';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Extract')));
Lsite=dataArray{1,1}(loc1+1:loc2-1);
Lsite=cellfun(@str2num,Lsite)';
```

```
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'GammaL')));
Rsite=dataArray{1,1}(loc1+1:loc2-1);
Rsite=cellfun(@str2num,Rsite)';

loc1=loc2;
loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'GammaR')));
gammaL=dataArray{1,1}(loc1+1:loc2-1);
gammaL=str2double(cell2mat(gammaL))';
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Probes')));
gammaR=dataArray{1,1}(loc1+1:loc2-1);
gammaR=str2double(cell2mat(gammaR));
loc1=loc2;

loc2=find(~cellfun(@isempty,strfind(dataArray{1,1},'Broadening')));
Dsites=dataArray{1,1}(loc1+1:loc2-1);
Dsites=cellfun(@str2num,Dsites)';

eta=dataArray{1,1}(loc2+1);
eta=str2double(cell2mat(eta));

lambda=dataArray{1,1}(loc2+3);
lambda=str2double(cell2mat(lambda));

bprobe=dataArray{1,1}(loc2+5:end);
bprobe=cellfun(@str2num,bprobe)';

fclose(x);
clearvars loc1 loc2 dataArray x

%%%%%%%%%%%%%Prepare Hamiltonian%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Load the matrix file as a struct in case filename is different from
%%% variable name.
eval(['Fm = load(''',mpath, strand, '.mat'')']);
temp = fieldnames(Fm);
H0 = eval(['Fm.' temp{:} ';']);

    if linVdrop == 1
        H0 = bias_drop(workdir,strand,H0,Orbitals,Vbias);
    else
        H0 = bias_drop_nonlin(workdir,strand,H0,Orbitals,Vbias,Vprcnt);
    end

[H,Lsite,Rsite,Dsites,Orbitals,HOMO] =
partitionH(H0,Lsite,Rsite,Dsites,Orbitals,num_P,sum_e,workdir);    %U'H0U, H-
blocking [diag(H)=EV(block)]
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
eta = 0;
sizeH = size(H, 1);  % size of Hamiltonian
disp(['Size of the Hamiltonian = ' num2str(sizeH)])

%%
```

```
Norb = sum(Orbitals);    %total size of H,total number of orbitals
if(Norb ~= sizeH)
    disp(['sum(Orbitals) = ' num2str(Norb)])
    disp('Total size of H =/= number of Orbitals, please check your input
files')
    return;
end
Nb = length(Dsites);
Nsite = length(Orbitals);   %total number of atoms


%% Loop Initialization %%
NE = length(Energy);
%%%prepare the output name before entering the loop%%%
Tname=strcat(workdir,'Teff_',strand,'_gammaL_',num2str(gammaL),'_gammaR_',num
2str(gammaR),'_bias_',num2str(mean(Vbias)),'_',num2str(run_num),'_BP(E)','.ma
t');

%%%%%%%% Check available files from Checkpoint %%%%%%%%%%
mat=dir(Tname);

if length(mat)~=0
load(mat.name);
qq=find(T~=-1, 1, 'last' )+1;
else
    T=-1*ones(1,NE);
    qq=1;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Entering the loop
EV = diag(H);
for nE = qq : NE
    nE
    E = Energy(nE)
%%%%%%%%%%set sumSig %%%%%%%%%%%%%%%%%%%%%
GAMMA = zeros(sizeH, sizeH);
%%%%%%%%%%%% Contacts Coupling %%%%%%%%%%%%%%
sites = [Lsite Rsite];
gamma = [gammaL*ones(1,length(Lsite)) gammaR*ones(1,length(Rsite))];

for ii = 1 : length(sites)
    isite = sites(ii);
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    GAMMA(range,range) = gamma(ii) * eye(Len);
end
%%%%%%%%%%%% Buttiker Probes %%%%%%%%%%%%%%
sites = Dsites;
gamma = bprobe;

for ii = 1 : length(sites)
    isite = sites(ii);
    TempLen1 = sum(Orbitals(1 : isite)) - Orbitals(isite);
```

```
    TempLen2 = sum(Orbitals(1 : isite));
    range = TempLen1 + 1 : TempLen2;
    Len = length(range);
    %exp decay per Atomic/Molecular Orbital (EV of H{block(ii)})
    f_bp= exp(-abs(E-EV(range))/lambda);
    %exp decay per Atomic/Molecular Orbital (HOMO cutoff})
    range2 = EV(range) > HOMO(1) & EV(range) <= mean(HOMO);
    f_bp(range2) = exp(-abs(E-HOMO(1))/lambda);
   %exp decay per Atomic/Molecular Orbital (LUMO cutoff})
    range2 = EV(range) < HOMO(2) & EV(range) > mean(HOMO);
    f_bp(range2) = exp(-abs(E-HOMO(2))/lambda);

    f_bp(f_bp<10^-20)=0; %cutoff for BP=0 is <10^-20

    GAMMA(range,range) = gamma(ii) * f_bp .* eye(Len);
end

%% Checking BP(AOi/MOi) = 0 to update Orbitals and Dsites and remove from
decoherent transport calculations
a=diag(GAMMA);
ao_nz = find(a~=0); %locates AO(MO) with BP=/=0
%transform orbital into its atom number( or block number, depending on the
%partitioning scheme)
idx = zeros(size(ao_nz));
z= cumsum(Orbitals);
    for ii=1:length(ao_nz)
        k=1;
            while ao_nz(ii)>z(k)
                    k=k+1;
            end
        idx(ii)=k;
    end

trunc_Dsites = setdiff(unique(idx),[Lsite Rsite])';
if ~isempty(trunc_Dsites)
ao_nz = [ao_nz idx]; %add corresponding atom numbers (or block number) per nz
orbital

%% Preparing Gr for Transport Calcuations %%

sumSig = -1i * GAMMA / 2;

    Gr = ((E + 1i * eta) * eye(sizeH) - H - sumSig) \ eye(sizeH);
    Ga = Gr';

T(nE) = decoherent_transport_f(Gr,Ga,GAMMA,Lsite,Rsite,trunc_Dsites,ao_nz);

else %Coherent Transport if all BP = 0
    sumSig = -1i * GAMMA / 2;
    Gr = ((E + 1i * eta) * eye(sizeH) - H - sumSig) \ eye(sizeH);
    Ga = Gr';

    T(nE) = coherent_transport_f(Gr,Ga,Lsite,Rsite,Orbitals,gammaL,gammaR);
end
```

```
%%%%%%% Save after every iteration %%%%
save(Tname,'Energy','T')
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
end


disp('Finished Decoherent Transmission!')
```

7. Code for the function to calculate decoherent transmission

```
function T = decoherent_transport_f(Gr,Ga,GAMMA,Lsite,Rsite,Dsites,Orbitals)
%%%%%%%%%% Transmission %%%%%%%%%%%%%
%%%%%%%%%% transmission between every 2 probes %%%%%%%%%%%%%
sites = [Lsite Rsite Dsites];
Nsite = length(sites);
Nb = length(Dsites);
Tmat = zeros(Nsite, Nsite);

for ii =  2: Nsite
    rangei = Orbitals(Orbitals(:,2)==sites(ii),1)';
    Gammai = GAMMA(rangei,rangei);

    for jj = 1 : (ii - 1)
        rangej = Orbitals(Orbitals(:,2)==sites(jj),1)';
        Gammaj = GAMMA(rangej,rangej);

        Tmat(ii, jj) = real(trace(Gammai * Gr(rangei, rangej) * Gammaj *
Ga(rangej, rangei)));
        Tmat(jj, ii) = Tmat(ii, jj);
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
LL = length([Lsite Rsite]);

TL = [];
TR = [];
for ii = LL+1 : Nsite         %include only Bprobe sites (no GammaL/GammaR)
    TL = [TL Tmat(1:length(Lsite), ii)];    %Lsite-to-Bprobe coupling
    TR = [TR; Tmat(ii, length(Lsite)+1:length(Lsite)+length(Rsite))]; %Rsite-
to-Bprobe coupling
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Wmat = zeros(Nb);

for ii = 1 : Nb

        Wmat(ii, ii) = sum(Tmat(ii + LL, :));

    for jj = ii + 1 : Nb
        Wmat(ii, jj) = -Tmat(ii + LL, jj + LL);
```

```
    end
end

Wmat = Wmat' + Wmat;
for ii = 1 : Nb
    Wmat(ii, ii) = Wmat(ii, ii) / 2;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Td = Tmat(1:length(Lsite), length(Lsite)+1:length(Lsite)+length(Rsite)) + TL
* (Wmat \ eye(Nb)) * TR;
T = sum(sum(Td));

end
```

8. *Code for the function to calculate coherent transmission*

```
function T= coherent_transport_f(Gr,Ga,Lsite,Rsite,Orbitals,gammaL,gammaR)


%%%%%%%%%% Transmission %%%%%%%%%%%%
%%%%%%%%%% transmission between Left and Right contact atoms %%%%%%%%%%%%%
Tmat = zeros(length(Lsite), length(Rsite));

for ii =  1: length(Lsite)
    isite = Lsite(ii);
    TempLeni1 = sum(Orbitals(1 : isite)) - sum(Orbitals(isite));
    TempLeni2 = sum(Orbitals(1 : isite));
    Leni = TempLeni2 - TempLeni1;
    Gammai = gammaL * eye(Leni);

    for jj = 1 : length(Rsite)
        jsite = Rsite(jj);
        TempLenj1 = sum(Orbitals(1 : jsite)) - sum(Orbitals(jsite));
        TempLenj2 = sum(Orbitals(1 : jsite));
        Lenj = TempLenj2 - TempLenj1;
        Gammaj = gammaR * eye(Lenj);

        Tmat(ii, jj) = real(trace(Gammai * Gr(TempLeni1 + 1 : TempLeni2,
TempLenj1 + 1 : TempLenj2) * Gammaj * Ga(TempLenj1 + 1 : TempLenj2, TempLeni1
+ 1 : TempLeni2)));
    end
end

T = sum(sum(Tmat));   %Ballistic Transmission is the sum of all Tij
end
```

9. *Code to apply linear bias drop across the Hamiltonian*

```
function [H0] = bias_drop(workdir,strand,H0,Orbitals,Vbias)
```

```
%%% Load Partitioning File
%%% This part uses the partitioning (e.g. base-pairs) to apply the voltage
%%% drop accross the molecule
vdrop = true;
Name=[workdir '/' strand '_Vbias_*.txt'];
Name = dir(Name);
try C = csvread(Name.name);
catch
disp('ERROR: Partitioning file not found, assuming no bias drop')
vdrop = false;
end
if vdrop
disp('Bias partitioning file found, applying bias using the defined blocking
scheme...')
b=cell(size(C,1),1);
sum_orbitals = zeros(size(b));
BlockSite = size(C,1);  %number of blocks partitioning the system
base1_range = zeros(BlockSite,2);
compbase_range = base1_range;

%%% Find the two bases making the basepair and save their atom sites. This
%%% helps call the correct indices H(range,range) when modifying the
%%% onsite potential per base-pair.
for ii=1:BlockSite
    b{ii}=nonzeros(C(ii,:))';
    compbase = find(diff(b{ii})>1);  %change in atom number at complementary
base
    base1_range(ii,:)=[b{ii}(1) b{ii}(compbase)];
    compbase_range(ii,:)=[b{ii}(compbase+1) b{ii}(end)];
end


%%%%% Applying Voltage Drop %%%%%%
%%% positive bias:  EfL < EfR
%%% negative bias:  EfL > EfR
%Eqt: En = En + (Vbias *(n-1)/(N-1));   linear drop
% Mu = EfL:Vbias/Nsite-1:EfR

    for ii=1:BlockSite
    %%% Base 1 of the base-pair
    TempLen1 = sum(Orbitals(1 :  base1_range(ii,1))) -
Orbitals(base1_range(ii,1));
    TempLen2 = sum(Orbitals(1 :  base1_range(ii,2)));
    range = TempLen1 + 1 : TempLen2;    %Matrix indices (to check:
length(range) must equal to
sum(Orbitals(base1_range(ii,1):base1_range(ii,2)))

H0(range,range) = H0(range,range) + eye(length(range))*(Vbias *(ii-
1)/(BlockSite-1));

    %%% Complementary Base of the base-pair
    TempLen1 = sum(Orbitals(1 :  compbase_range(ii,1))) -
Orbitals(compbase_range(ii,1));
    TempLen2 = sum(Orbitals(1 :  compbase_range(ii,2)));
```

```
    range = TempLen1 + 1 : TempLen2;    %Matrix indices (to check:
length(range) must equal to
sum(Orbitals(base1_range(ii,1):base1_range(ii,2)))

H0(range,range) = H0(range,range) + eye(length(range))*(Vbias *(ii-
1)/(BlockSite-1));
    end
end


end
```

10. Code to apply nonlinear bias drop across the Hamiltonian

```
function [H0] = bias_drop_nonlin(workdir,strand,H0,Orbitals,Vbias,Vprcnt)

%%% Load Partitioning File
%%% This part uses the partitioning (e.g. base-pairs) to apply the
%%% nonlinear voltage drop (default: 40%:20%:40%) accross the molecule
vdrop = true;
Name=[workdir '/' strand '_Vbias_*.txt'];
Name = dir(Name);
try C = csvread(Name.name);
catch
disp('ERROR: Partitioning file not found, assuming no bias drop')
vdrop = false;
end
if vdrop
disp('Bias partitioning file found, applying nonlinear bias using the defined
blocking scheme...')
b=cell(size(C,1),1);
sum_orbitals = zeros(size(b));
BlockSite = size(C,1);  %number of blocks partitioning the system
base1_range = zeros(BlockSite,2);
compbase_range = base1_range;

%%% Find the two bases making the basepair and save their atom sites. This
%%% helps call the correct indices H(range,range) when modifying the
%%% onsite potential per base-pair.
for ii=1:BlockSite
    b{ii}=nonzeros(C(ii,:))';
    compbase = find(diff(b{ii})>1);  %change in atom number at complementary
base
    base1_range(ii,:)=[b{ii}(1) b{ii}(compbase)];
    compbase_range(ii,:)=[b{ii}(compbase+1) b{ii}(end)];
end


%%%%% Applying Voltage Drop %%%%%
%%% positive bias:  EfL < EfR
%%% negative bias:  EfL > EfR
% Equation for Shifting onsite potential:
% En = { En,n=1
```

```
%       En+qV_bias×(Vprcnt+1-2*Vprcnt*(n-2)/(N-3)),for 2≤n≤N-1
%       En+qV_bias,for n=N }

    for ii=2:BlockSite-1
    %%% Base 1 of the base-pair
    TempLen1 = sum(Orbitals(1 :  base1_range(ii,1))) -
Orbitals(base1_range(ii,1));
    TempLen2 = sum(Orbitals(1 :  base1_range(ii,2)));
    range = TempLen1 + 1 : TempLen2;    %Matrix indices (to check:
length(range) must equal to
sum(Orbitals(base1_range(ii,1):base1_range(ii,2)))

H0(range,range) = H0(range,range) + eye(length(range))*(Vbias * Vprcnt) +
eye(length(range))*(Vbias * (1-Vprcnt*2) * (ii-2)/(BlockSite-3));

    %%% Complementary Base of the base-pair
    TempLen1 = sum(Orbitals(1 :  compbase_range(ii,1))) -
Orbitals(compbase_range(ii,1));
    TempLen2 = sum(Orbitals(1 :  compbase_range(ii,2)));
    range = TempLen1 + 1 : TempLen2;    %Matrix indices (to check:
length(range) must equal to
sum(Orbitals(base1_range(ii,1):base1_range(ii,2)))

H0(range,range) = H0(range,range) + eye(length(range))*(Vbias * Vprcnt) +
eye(length(range))*(Vbias * (1-Vprcnt*2) * (ii-2)/(BlockSite-3));
    end

%%% After the for loop, set potential shift at last site (BlockSite) to
%%% Vdrop
    ii = BlockSite;
    TempLen1 = sum(Orbitals(1 :  base1_range(ii,1))) -
Orbitals(base1_range(ii,1));
    TempLen2 = sum(Orbitals(1 :  base1_range(ii,2)));
    range = TempLen1 + 1 : TempLen2;    %Matrix indices (to check:
length(range) must equal to
sum(Orbitals(base1_range(ii,1):base1_range(ii,2)))

H0(range,range) = H0(range,range) + eye(length(range))* Vbias ;

    %%% Complementary Base of the base-pair of the last block
    TempLen1 = sum(Orbitals(1 :  compbase_range(ii,1))) -
Orbitals(compbase_range(ii,1));
    TempLen2 = sum(Orbitals(1 :  compbase_range(ii,2)));
    range = TempLen1 + 1 : TempLen2;    %Matrix indices (to check:
length(range) must equal to
sum(Orbitals(base1_range(ii,1):base1_range(ii,2)))

H0(range,range) = H0(range,range) + eye(length(range))* Vbias ;
end


end
```

11. Code to partition the Hamiltonian

```
%%% This function Partitions H into Blocks based on the partitioning scheme
%%% defined by the user.
function  [H,Lsite,Rsite,Dsites,Orbitals,HOMO] =
partitionH(H0,Lsite,Rsite,Dsites,Orbitals,num_P,sum_e,workdir)


FNAME = dir([workdir '/*_partitioning.txt']);
try C = csvread(FNAME.name);
catch
    disp('no *partitioning.txt file was found, running atomic-partitioning
calculations')
     block_flag=false;
     C=[];
end

if ~isempty(C)
block_flag=true;   %user-defined partitionning
b=cell(size(C,1),1);


sum_orbitals = zeros(size(b));
for i=1:size(C,1)
    b{i}=nonzeros(C(i,:))';
    sum_orbitals(i)=sum(Orbitals(b{i})); % number of orbitals per block
end
disp('Partitioning file found, using the defined blocking scheme')


H0 = rearrange_H(H0,b,Orbitals); %Re-arrange H0 to have the atoms in serial
order per block definition
end

if block_flag
%%% L defines element range of H per partition (e.g. BB+Base) based on the
%%% number of orbitals per block.
L(1,1) = 1;
L(1,2) = sum_orbitals(1);
for i = 2:length(sum_orbitals)
    L(i,1) = L(i-1,2)+1;
    L(i,2) = sum(sum_orbitals(1:i));
end
%% Re-Define the Orbitals and Sites if block_flag is true (if not atomic
partitioning)
   ii=0;
   idx = 0;
    while ~diag(idx)
        ii=ii+1;
        idx= diag(b{ii}==Lsite') ;
    end

    Lsite = ii;

    ii=0;
   idx = 0;
    while ~diag(idx)
```

```
        ii=ii+1;
        idx= diag(b{ii}==Rsite') ;
    end
    Rsite = ii;

    Dsites = setdiff(1:length(L),[Lsite Rsite]);

    Orbitals = sum_orbitals;


else

%%% L defines H element range per atom
L(1,1) = 1;
L(1,2) = Orbitals(1);
for i = 2:length(Orbitals)
    L(i,1) = L(i-1,2)+1;
    L(i,2) = sum(Orbitals(1:i));
end

end
%% Transform H0 Into Blocked Matrix H
%%% each partition is defined as per '*partitioning.txt' or per atom
disp('Partitioning the Hamiltonian by applying [H=U''H0U]')

%%% Create U matrix per layer(i), then transform H0 to H
U = zeros(size(H0));
Eig_Mat = U;
for i=1:length(L)
    temp = H0(L(i,1):L(i,2), L(i,1):L(i,2));
    [EVec,EVal] = eig(temp);    % eig(block(i)), EV is not sorted by default
    % sorting Evect and Eval
    [~,ind] = sort(diag(EVal));
    EVal = EVal(ind,ind);   %sorted EVal
    EVec = EVec(:,ind);     %sorted EVec
    U(L(i,1):L(i,2), L(i,1):L(i,2)) = EVec;
    Eig_Mat(L(i,1):L(i,2), L(i,1):L(i,2)) = EVal;
    clear temp
end

    H = U' * H0 * U;
%% Extracting HOMO and LUMO
%%% For Edep decoherence cutoff
HOMO = zeros(1,2);
if sum_e > 0
    EV = sort(eig(H));
    homo_loc = ceil((sum_e+num_P)/2);
    HOMO(1) = EV(homo_loc);     %HOMO
    HOMO(2) = EV(homo_loc+1);   %LUMO
end
end
```

12. Code to rearrange the Hamiltonian

```
function [H_new] = rearrange_H(H,blocks,Orbitals)
%%% This function re-arranges H based on the layers order, it exracts
%%% onsite potential as well as the coupling to other bases. This helps
%%% reorganizing added atoms' location to be in between the correct bases
instead of
%%% being at the end as in the Gaussian input file
%%% Note: the variables 'layers' and 'orbitals' do not need to be updated
after
%%% re-arrangement of H
%%% Extract corresponding layers from the Hamiltonian %%%
%%% 1st, get atoms location inside the Hamiltonian %%%
%%% Define range of H to be updated per block number
blockN = [blocks{:}];
%%% Range of Orbitals (H elements) per atom
for ii=1:length(blockN)
    temp_lcn1(ii) = sum(Orbitals(1 : blockN(ii)))-
sum(Orbitals(blockN(ii)))+1;
      temp_lcn2(ii) = sum(Orbitals(1 : blockN(ii)));
end

for ii=1:length(blockN)
    row_loc = temp_lcn1(ii)  : temp_lcn2(ii);
      for jj=1:length(blockN)
            col_loc = temp_lcn1(jj)  : temp_lcn2(jj);
            H_new{ii,jj} = H(row_loc(1):row_loc(end),
col_loc(1):col_loc(end));
        end
end

H_new=cell2mat(H_new);

end
```