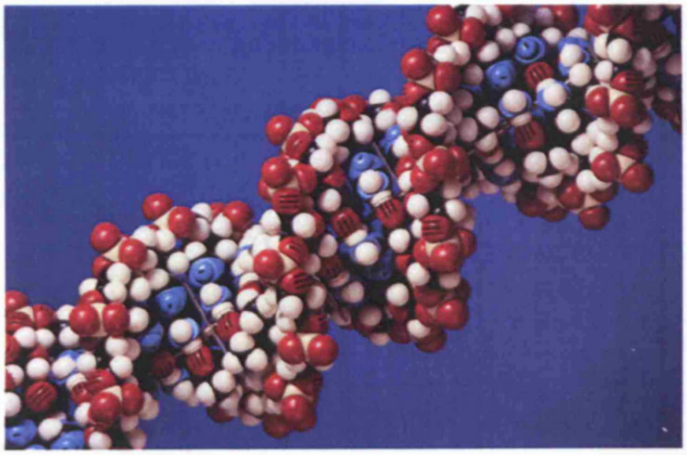
Electronic Structure and Conductance of DNA

Owen R. Davies



A thesis submitted to

Cardiff University

for the degree of

Doctor of Philosophy

March 2005

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

UMI Number: U584783

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U584783
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

During my studies I have been fortunate enough to have been supervised by John Inglesfield. John must be commended for his patience, enthusiasm and eternal optimism, and has become as much a friend as a supervisor.

I would also like to thank Ian Merrick and Hannah Loebl for the generous gift of much of their time while helping me to overcome various computing problems. Part of the joy of studying in Cardiff has been the friendly environment that is so conducive to co-operative work. I would therefore like to thank all of my peers within the department, and especially Dean Trolley. The departmental sporting activities such as the weekly football matches, regular squash and the bi-annual cricket matches have all added to the fun of research in Cardiff. I would also like to acknowledge a studentship from EPSRC.

Special mention must be given to my parents for their continual emotional and financial support throughout my studies, without which life would have been much harder. Finally, I would like to thank Anna Mitchell; over the past three years she has been a constant source of encouragement, motivation and moreover, comic relief. I will forever love her.

ABSTRACT

In this thesis an efficient method has been developed for calculating the electronic structure and conductance of large biological molecules. The embedding method has been adapted to allow the splitting of large molecules such as DNA into smaller component blocks, completely preserving the information of the macro-system. The computational time required for this method increases as $O(N)$ with the size of the system, instead of the traditional $O(N^3)$. The semi-empirical extended Hückel theory is used to describe the electron wavefunctions within a tight-binding scheme, taking the effect of the metal-molecule contacts into account.

Presented in this thesis are the results for several different DNA molecules and structures. It has been determined that the transmission through DNA depends sensitively on the energy at which it is evaluated, and the atoms to which the metallic leads are connected. It is also found that poly(G)-poly(C) DNA conducts charge better than DNA with mixed bases, and that energy-minimised DNA with less structural disorder conducts better than DNA obtained from x-ray diffraction experiments. The electrical conduction of DNA that has undergone stretching has been investigated, and the distorted structure gives very small currents.

The embedding method has also been applied to the small aromatic molecule OPE, to determine its electronic properties. Metallic conductivity is found for this molecule, and it is able to carry currents 1000 times greater than DNA, giving possible applications in molecular electronics.

PUBLICATIONS

O. R. Davies and J. E. Inglesfield. Electronic structure and conductance of large molecules and DNA. *Progress in Surface Science* **74** (1-8) 161-176, December 2003.

O. R. Davies and J. E. Inglesfield. Embedding method for conductance of DNA. *Phys. Rev. B* **69** 195110, May 2004.

Reprinted in *Virtual Journal of Nanoscale Science & Technology* **9** 22 (2004).

O. R. Davies and J. E. Inglesfield. Embedding methods for conductance studies of large molecules. Invited book chapter for *Modern methods for theoretical physical chemistry of biomolecules*, eds. E. B. Starikov, S. Tanaka and J. P. Lewis. (Elsevier 2005), in preparation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The field of molecular electronics is motivated by the desire to engineer ever smaller electronic devices. A smaller electronic device allows for the production of denser integrated circuits, which is the key to producing faster and more powerful computers. It has been predicted that by 2014 we will reach the limitations of conventional lithographic techniques [1]. Present transistor junctions are 180 nm across, close to the limit of about 25 nm, beyond which a tunneling current will short circuit conventional transistors [2]. The inherent size, and the ability of organic molecules to self-assemble, may allow for the production of switches, transistors and memory circuits on the nanometer scale. This relatively new field of mesoscopics requires expertise from the fields of biology, chemistry and physics.

The first break-through in molecular electronics is credited to H. Shirakawa, when in the mid 1970's he and a co-worker at the Tokyo Institute of Technology accidentally discovered the first conducting polymer. They had inadvertently added too much catalyst when trying to form polyacetylene from a solution of the monomer (ethyne), and found a thin silvery film had formed at the surface, which as its

appearance suggested was later found to be conducting [3]. In 1987 the group marketed the first plastic battery which boasted a higher capacity, higher voltage and longer shelf-life than traditional batteries. In 2000 Shirakawa was jointly awarded the Nobel prize for chemistry for his discovery and subsequent work in the field.

However, McGinness *et al.* claim to have made the first organic semiconductor [4]. They used melanin, a derivative of polyacetylene that gives skin its colour, and found that when subject to an electric field the molecule behaves as a bistable switch. The results were published 3 years before Shirakawa.

Another step forward for molecular electronics was made in 1991 when Iijima made the first carbon nanotube [5]. Carbon nanotubes are hollow cylinders of graphite with a diameter of around 1 nm. These fascinating structures have a tensile strength up to 63 GPa [6], and the inter-carbon bonds are stronger than in diamond. It is possible to make long chains of carbon nanotubes by linking individual tubes under high pressure. This has led to proposed applications for carbon nanotubes as strong wires with no limit on their length [7]. Carbon nanotubes are also good conductors of heat and may have applications as heat dissipators in nano-electronics. The electrical properties of these molecules is also very interesting, as they can have metallic or semiconducting properties depending on the radius of the tube and how the tube is wrapped [8]. It is in fact possible to make a diode by joining two tubes of different diameters. Nanoscale transistors have also been sucessfully made using carbon nanotubes [9]. There are however fabrication difficultites with carbon nanotubes, as it is difficult to control whether the tube will be metallic or semiconducting. However, with their wide range of desirable properties it seems certain that carbon nanotubes will have a role to play in the future of molecular electronics.

One of the main applications for organic molecular electronics is in display screens. With low power and low cost, organic light emitting diodes (OLEDs) are likely to replace existing LCD displays. The first OLEDs were produced by Tang and Slyke in 1987 [10]. Nowadays, a variety of electroluminescent polymers are synthesized from monomers such as oxetane, aromatic dibromides and diboronic esters. These devices are constructed by vapour deposition or spin coating. Companies such as Philips, Pioneer and Cambridge Display Technology are producing OLEDs for mobile phones, stereos and digital camera screens.

Many molecules have been investigated to determine their suitability for electronic applications, including small organic polymers, large biological molecules, nanotubes and fullerenes. However, it is DNA which still provokes much debate concerning the future of molecular electronics. With its stable double helix structure and self-assembly properties DNA has been proposed for numerous practical applications, from DNA computers [11, 12], to bio-sensors for the pharmaceutical industry [13] and as the building blocks of a nanoscale construction kit [14, 15]. In this thesis I concentrate on a theoretical study of the electrical conductance of DNA, and also apply my methods to other molecules.

The idea of using DNA as a conductor is not new. In 1962, Eley and Spivey [16] suggested that DNA could transfer electrical charge via π-bonding between consecutive bases in the chain. Later, low temperature experiments showed that radiation-induced conductivity can only be due to charge migration through frozen water molecules rather than through the DNA itself [17]. Interest continued in the conduction of DNA with the investigation of the electron transfer involved in the oxidizing damage done to DNA, which has been linked to cancer [18]. In the early nineties interest in the field was renewed when Murphy *et al.* [19] suggested that

DNA conductivity was responsible for observed fluorescence quenching experiments. This restarted the debate about the charge transport properties of DNA. The debate still continues, with a wide range of experimental results. Some experiments have shown that DNA is an insulator [20], while others suggest that it can be a good linear conductor [21], or a wide-band-gap semiconductor [22]. Some published data even claim that DNA can superconduct [23].

In order to understand these conflicting results theorists have developed many ways of calculating transport properties based on a variety of electronic structure methods including Hartree-Fock, density functional theory and semi-empirical approaches. The effects of structure, base-pair order, temperature and environment are all important effects which are treated in different ways to try and understand the electronic properties of the DNA molecule. It is clear that further work is required in this area to resolve the many differences in both theoretical and experimental results, and in particular, to study the relationship between structure and conductance.

In my work I have developed a method of tight-binding embedding to perform efficient electronic structure and conductance calculations on large molecules. The method is flexible and can be used for many different molecules, although I have concentrated my study on DNA. I have investigated several different molecules of DNA and discovered differences in the conduction of these molecules related to the structure and composition of the molecules themselves.

1.1 DNA structure

DNA (Deoxyribose Nucleic Acid) is the molecule contained in every living cell, holding all the genetic information required to assemble a living being. This amazing double helix molecule is the basis of research in many varied scientific fields, ranging from the human genome project [24] and genetic engineering [25], to gene therapy [26] and DNA conductivity [21].

The building blocks of DNA are sugars, phosphates and bases, and are arranged as shown in fig 1.1. Sugars are composed of a ring of five carbon atoms, and each sugar is connected to both a phosphate and a base, this combination being known as a nucleotide. The sugar-phosphate bonds form the two twisting strands of the backbone of the helical molecule, with the bases holding the two interwoven backbones together via hydrogen bonds, as shown in fig 1.1. There are four types of bases: adenine (A), thymine (T), guanine (G), and cytosine(C). Bases G and A are pyrimidine molecules consisting of two rings of carbon atoms, and bases C and T are purine molecules with only one carbon ring.

Before the 1950's it was believed that proteins rather than DNA were the carriers of genetic information. DNA was thought to have a regular sequence of bases, for example ATCGATCG... and therefore unable to carry genetic information. Today we know that DNA is a long chain of deliberately sequenced nucleotides, represented by the letters A, T, G and C, which is a code that contains all the information that our bodies require to develop and function. In 1953 Watson and Crick published their now famous paper, proposing a structure for DNA [29], based on x-ray crystallography data. Their radical new theory suggested the double helix structure shown in fig 1.2. The reason for the helical structure is explained by the molecule's interaction

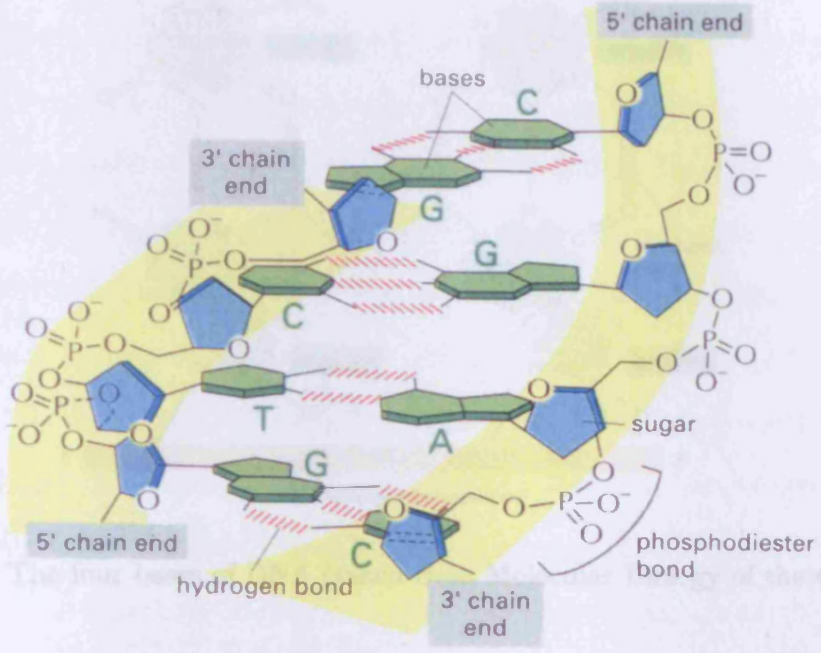


Figure 1.1: Structure of DNA (taken from Molecular Biology of The Cell [27]).

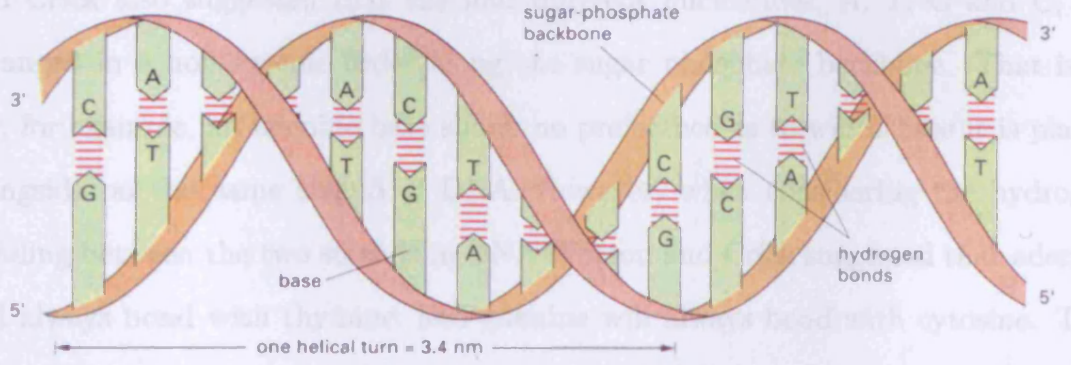


Figure 1.2: Double helix structure of DNA molecule (taken from Molecular Biology of the Cell [28]).

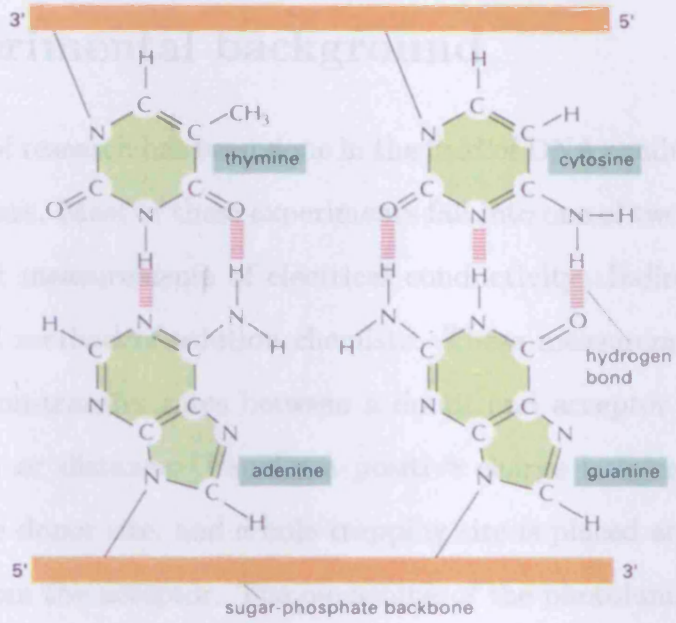


Figure 1.3: The four bases of DNA (taken from Molecular Biology of the Cell [28]).

with water – the backbone of DNA is hydrophilic while the bases are hydrophobic, hence by twisting, the DNA molecule reduces the space between the bases and therefore reduces the amount of water present in the centre of the molecule. Watson and Crick also suggested that the four different nucleotides, A, T, G and C, are arranged in a non-specific order along the sugar phosphate backbone. That is to say, for example, an adenine base shows no preference as to which base it is placed alongside on the same strand of DNA. However, when considering the hydrogen bonding between the two strands in DNA Watson and Crick suggested that adenine will always bond with thymine, and guanine will always bond with cytosine. This revolutionary theory meant that not only could DNA carry genetic information, but there is a mechanism for it to be reproduced.

1.2 Experimental background

A great deal of research has been done in the field of DNA conductivity, especially in the last ten years. Most of these experiments fall into one of two categories, either indirect or direct measurements of electrical conductivity. Indirect measurements are the preferred method of solution chemists. These measurements are aimed at measuring electron-transfer rates between a donor and acceptor site as a function of base sequence or distance. Usually a positive charge is injected into the base stack to form the donor site, and a hole trapping site is placed at a pre-determined distance away from the acceptor. The quenching of the photoluminescent acceptors is used to find electron transfer rates, giving an average result over a large number of molecules. These methods eliminate the effect of any metal contacts, thought to play a major role in conductivity [30], allowing for the measurement of the molecules by themselves.

Solid state physicists favour the use of direct measurements of electrical conductivity. These electrical transport measurements aim to find the current passed by the molecule while under an externally applied electric field [22]. In some experiments molecules are positioned between two metal electrodes [31]. Another technique is to use an atomic force microscope (AFM) as a second contact allowing for distance-dependent measurements [32]. Direct contact to the molecule allows the charge state on the helix to remain unaltered, since charge is supplied from the attached metal reservoirs. Whether the charge carriers are electrons or holes depends on the availability of electron states and the position of the Fermi energy.

1.3 Indirect measurements

Interest in DNA conduction was kick-started by Murphy *et al.* [19] in the early 1990's. They performed experiments on a 15 base-pair double stranded DNA molecule in which metal complexes acting as donor and acceptor sites were non-covalently bound to the base stack (intercalated). Murphy *et al.* found that when the donor complexes were photo-excited in solution, without the presence of the acceptor complexes, the complexes fluoresced. However, when the acceptors were introduced to the system the fluorescence was quenched. Murphy *et al.* suggested that the DNA was acting as a conductor, moving charge carriers from the donor to the acceptor sites.

Lincoln *et al.* [33] later performed a similar experiment to Murphy *et al.* and after modeling the data they suggested that the metal complexes bind adjacent to each other along the base stack, thus not giving the random distribution that Murphy *et al.* expected, and allowing for the fast rate of charge transfer with a small distance dependence. Subsequent work using indirect measurements only seemed to produce contradictory results [34–38].

In 2000 the Barton group at Caltech [39] investigated the effect of base sequence on electron transfer rates. They used rhodium intercalators as the donors and acceptors and concluded that DNA can transfer charge over long distances up to hundreds of angstroms, with the rate of transfer highly dependent on the base sequence. They increased the number of TA base pairs between the donor and acceptor sites – TA base pairs have a higher ionisation energy than a GC pair and therefore present a potential barrier to electron transfer. They indeed find that the charge transfer rate dropped dramatically as the number of TA steps were increased from 1 to 3,

with further increases yielding only a small reduction in charge transfer. The Barton group suggests that there is a change in the transfer mechanism from guanine hopping, which has a sharp distance dependence, to hole hopping through all base types, which has a much shallower distance dependence.

Despite the encouraging results from Barton's group, further contactless experiments indicated that both λ-DNA and poly(G)-poly(C) DNA do not conduct electricity [40–43]. These contradictary results only serve as fuel to the debate on DNA conductivity.

Tran *et al.* [44] used a spectroscopic method for their experiment. They reported insulating behaviour for a 17 μm long molecule of λ-DNA, a multi-base sequence of DNA found in the bacteria E-coli. Tran *et al.* measured the change in quality factor of resonant cavities as the DNA molecule is introduced into the cavity. The results show a strong temperature dependence near room temperature, and a weak temperature dependence at lower temperatures. They also found that the presence of a buffer solution increases the conductivity. Tran *et al.* suggest that this may be due to increased disorder in dry DNA which leads to low conductivity.

1.4 Direct measurements

The first direct measurement of electrical transport on 16 μm long λ-DNA was published by Braun *et al.* [20] in 1998. In this experiment two gold electrodes were spaced by 12-16 μm, as shown in figure 1.4, each electrode being primed with a short sequence of 12 nucleotides. The λ-DNA then has the complementary oligonucleotide sequence attached to each end, and hence the connection is made between the λ-DNA and the electrode. The results show insulating behaviour for voltages up to

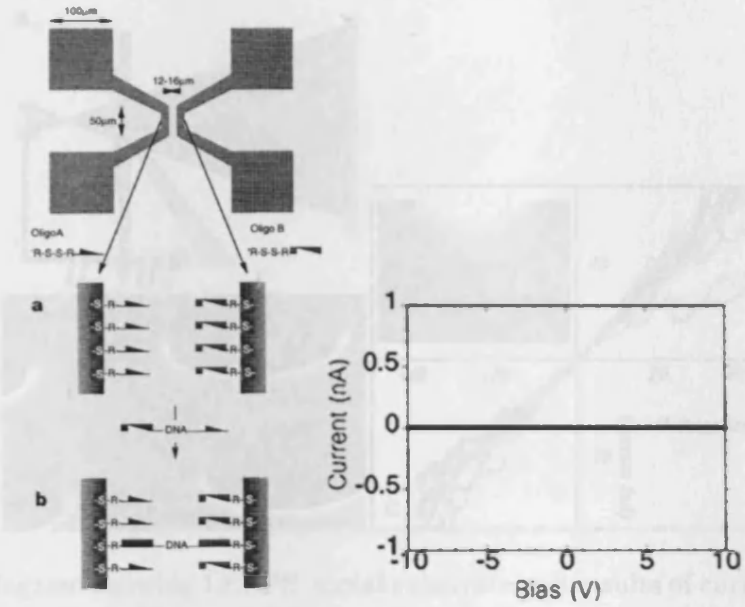


Figure 1.4: Schematic of experiment used to determine conductivity of DNA, and Results (taken from Braun *et al.*[20]).

±10 V (figure 1.4).

Contrary to the work of Braun *et al.*, Fink and Schonenberger [21] report ohmic behaviour for a similar molecule of λ-DNA. In their experiment Fink and Schonenberger use DNA molecules of a few hundred nanometres in length to form ropes which are stretched across 2 μm wide gaps in a metal covered transmission electron microscope grid, as shown in figure 1.5. The two terminals in this experiment were the metal grid and a tungsten tip, which was aligned using a holographic image created with a low energy electron point source (LEEPS). The results, given in figure 1.5 showed ohmic behaviour for a bias voltage up to ± 20 mV, beyond which the results fluctuate. From the linear part of the *I-V* graph the resistance was calculated as 25 MΩ. Since this experiment was done in vacuum, ionic conduction via a buffer solution cannot explain the conductivity. However, it has been

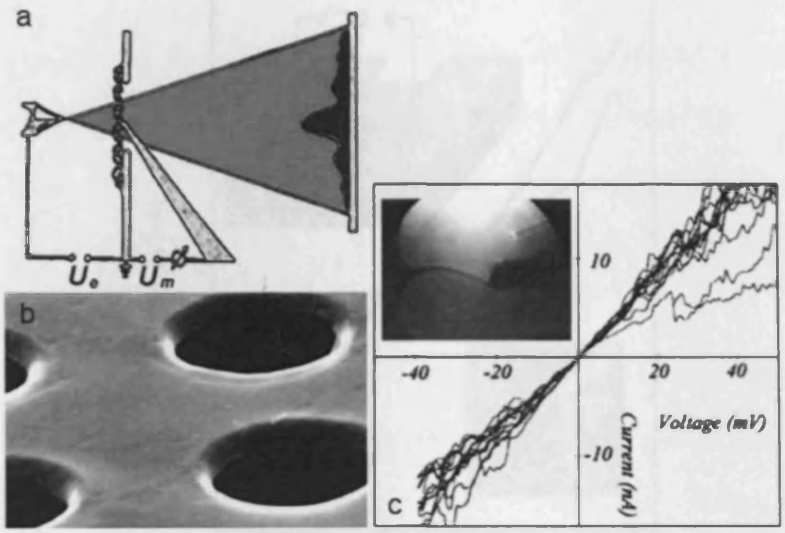


Figure 1.5: Diagram showing LEEPS, metal substrate and results of current versus voltage for λ-DNA, (taken from Fink and Schonenberger [21]).

suggested that LEEPS imaging can damage the DNA [45], and this may have led to the conductivity observed in this experiment.

In 2000 Porath *et al.* [22] published their results for the conductivity of a 10 nm long molecule of poly(G)-poly(C) DNA. This molecule has only guanine bases on one strand with the complementary cytosine bases on the other strand. It is thought that this sequence will conduct charge more efficiently than DNA with a random base order. In their experiment Porath *et al.* electrostatically trapped the DNA molecule between two platinum electrodes separated by 8 nm. The molecule was then dried with nitrogen and current-voltage measurements were taken at a range of temperatures. Their results given in figure 1.6 show that no current is observed at low voltages, then beyond a threshold voltage the current begins to increase – this is semiconducting behaviour for this short poly(G)-poly(C) DNA chain. They also found that the voltage required to produce a current flow through the molecule

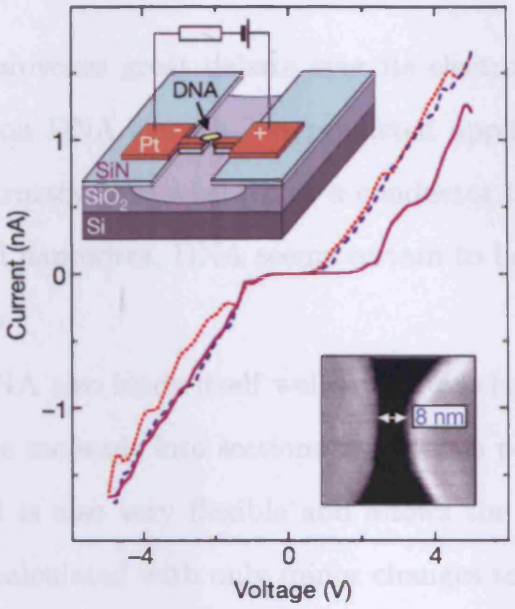


Figure 1.6: Schematic of experiment and results of current versus voltage for poly(G)-poly(C) DNA, taken from [22].

increases as the temperature is raised, suggesting an energy-dependent bandgap.

A number of other experiments have followed this work, unfortunately still producing contradictory results; some showed insulating behaviour [45–47], some found semiconducting properties [48–50] and others linear characteristics [51].

1.5 Motivation for this work

The motivation for this thesis derives from the exciting possibilities of molecular electronics. There are unlimited applications for this technology in a wide variety of fields, from producing superfast supercomputers to helping to solve the mysteries of oxidative damage and cancer.

There are many conducting molecules in use in industry at the moment, however,

it is DNA which still provokes great debate over its electronic properties. I have chosen to concentrate on DNA since it has predicted applications in many fields from computing to pharmacy, and whether as a conductor itself or as an assembly molecule that can build nanowires, DNA seems certain to have a role in the future of molecular electronics.

The structure of DNA also lends itself well to the method of embedding that I use, which separates the molecule into sections in order to perform the calculations efficiently. My method is also very flexible and allows the electronic structure of many molecules to be calculated with only minor changes to the code.

With this work I hope to add to the field of knowledge in this extremely interesting area of computational physics, and we will move closer to a unified theory of electrical conduction through DNA.

1.6 Outline of Thesis

The structure of this thesis is as follows: we begin with a discussion of the various methods of calculating electronic structure, including Hartree-Fock, density functional theory, and tight-binding. In chapter 3 the background of conductance theory is presented, and a summary of recent electronic structure calculations is given. In chapter 4 I derive the density of states and transmission within our embedding scheme. The results for the density of states, transmission and current-voltage behaviour for DNA are presented in chapter 5. Chapter 6 shows the electronic properties of DNA that has undergone stretching, a topic suggested by recent molecular dynamics studies. An investigation of the electrical properties of the small aromatic molecule OPE is given in chapter 7. Finally, a summary and proposed future work

is discussed in chapter 8.

Atomic units are generally used in this thesis, where $e = \hbar = m = 1$. The atomic units for distance and energy are 0.529 Å and 27.2 eV respectively.

Chapter 2

Electronic Structure Methods

All materials including metals, semiconductors, biomolecules and ultimately the living world are held together by electrons, and an understanding of the behaviour of these electrons is fundamental to understanding the properties of these materials. Moreover, modern calculations of electronic structure can be used not only to explain experimental results but also to predict properties such as geometrical structure. In this chapter I summarise the background of electronic structure to which I refer in my work.

The starting point for any quantum mechanical calculation is the Schrödinger equation. The stationary states of the electrons are found by solving the stationary Schrödinger equation,

$$H\Psi = E\Psi, \tag{2.1}$$

where H is the Hamiltonian operator of the system, E gives the possible energy eigenvalues and Ψ is the many-electron wavefunction, from which all the physical

properties of the system can be determined. The electron Hamiltonian is given by

$$H = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i V_{nuc}(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.2}$$

where the first term corresponds to the kinetic energy, the second term is the interaction between the atomic nuclei and the electrons and the third term describes the electron-electron interaction. This equation already invokes the Born-Oppenheimer approximation to simplify the problem [52]. The Born-Oppenheimer approximation states that the motion of the electrons can be considered independently of the motion of the much heavier nuclei; so that when considering electron motion the nuclei are at rest, and when considering nuclear motion the electrons react essentially instantaneously. This greatly simplifies the problem. It has been suggested in the literature [53] that polaron effects may be important for transport in DNA, which means that the separation of electron and nuclear motion is no longer valid. However, in our model we assume that the Born-Oppenheimer approximation is valid.

Even the one-electron Schrödinger equation is impossible to solve exactly except for particular cases like the hydrogen atom and in general the solutions have to be computed. The many-body problem is of course vastly more complicated, but it becomes tractable under different approximations if we apply the variational principle. This says that the ground-state energy E satisfies the relation [54],

$$E \leq \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{2.3}$$

where Ψ is a trial function, and by varying Ψ to minimise E we can obtain a good

estimate of both the ground-state energy and wavefunction. We can make progress by expanding Ψ in terms of basis functions Φ (which at this stage we assume to be orthonormal), and treating the coefficients, C_n as variational parameters,

$$\Psi = \sum_n C_n \Phi_n. \tag{2.4}$$

The Schrödinger equation can then be written in matrix form, giving the matrix eigenvalue problem,

$$\sum_m (H_{nm} - E\delta_{nm}) C_m = 0. \tag{2.5}$$

The eigenvalues and eigenvectors contain all the necessary information for describing the electronic structure. This approach is used in the one-electron problem. However, in many-electron systems a straightforward expansion like (2.4) is impossible, because of all the different possible many-electron excitations. We shall see in section 2.1 how a variational approach can simplify the many-electron problem. As Dirac said soon after the Schrödinger equation was formalised,

> "The general theory of quantum mechanics is now almost complete. The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact applications of these laws leads to equations much too complicated to be soluble."
>
> P.A.M. Dirac 1929

In the remainder of this chapter I will discuss the main methods that are used to perform electronic structure calculations.

2.1 Hartree-Fock

Hartree-Fock theory is the method of choice among many chemists. In this method, which is a mean field theory, each electron feels the average Coulomb repulsion of all the other electrons (the Hartree potential), and also the non-local exchange potential due to the antisymmetric nature of the wavefunctions.

Hartree-Fock, which is based on the variational principle, equation (2.3), uses a many-electron wavefunction in which Ψ is represented by a Slater determinant of occupied one-electron wavefunctions $\psi_i(\mathbf{x}_i)$, where $\mathbf{x}_i$ includes both spin and spatial co-ordinates,

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \psi_1(\mathbf{x}_3) & \cdots & \psi_1(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \psi_2(\mathbf{x}_3) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & & & & \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \psi_N(\mathbf{x}_3) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.6}$$

The total wavefunction Ψ is antisymmetric since electrons are fermions and therefore must obey the antisymmetry principle, which states that a wavefunction describing fermions should be antisymmetric with respect to the interchange of any set of space-spin co-ordinates. A Slater determinant is the simplest function which satisfies the antisymmetry requirement. In this form all the electrons are indistinguishable and each electron is associated with every orbital.

Varying the one-electron wavefunctions ψ_i to minimise the total energy in equation (2.3), subject to the normalisation constraint, a single-particle equation then

results,

$$\left[-\frac{1}{2}\nabla^2 + V_{nuc}(\mathbf{r})\right]\psi_i(\mathbf{x}) + V_H(\mathbf{r})\psi_i(\mathbf{x}) + \int dx' V_x(\mathbf{x},\mathbf{x}')\psi_i(\mathbf{x}') = \epsilon_i\psi_i(\mathbf{r}), \qquad (2.7)$$

where the Hartree potential, V_H and the exchange potential, V_x can be expressed as

$$V_H(\mathbf{r}) = \sum_j \int dx' \psi_j(\mathbf{x}')\psi_j^*(\mathbf{x}')\frac{1}{|\mathbf{r}-\mathbf{r}'|} \qquad (2.8)$$

$$\int dx' V_x(\mathbf{x},\mathbf{x}')\psi_i(\mathbf{x}') = -\sum_j \int dx' \psi_i(\mathbf{x}')\psi_j^*(\mathbf{x}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{x}). \qquad (2.9)$$

Equation (2.7) can then be solved usually with a basis set expansion. This is much easier since we have reduced the many-electron problem to a one-electron problem. However, the Hartree and exchange potentials must be found self-consistently; that means starting with an initial guess for the Hartree and exchange potentials, the eigenfunctions and eigenvalues are then found, from this the Hartree and exchange potentials are recalculated, and this process is then repeated until the input and output potentials are acceptably close together.

This method is only an approximation but it works well for calculating total energies and equilibrium geometries of molecules. Hartree-Fock methods have been found to break down in metals [55] and semiconductors [56], where it is important to have a more accurate description of the electron-electron interaction. In metals it is known that Hartree-Fock gives a very poor description of the density of states, for example with a singularity at the Fermi energy.

2.2 Density functional theory

The drawback of Hartree-Fock theory is that it completely neglects the correlated motion of the electrons, though it includes the effects of exchange. Both exchange and correlation are important effects that are taken into account in density functional theory (DFT). The exchange interaction ensures that no two electrons with parallel spin can occupy the same space at the same time, thus giving rise to an effective repulsion. Correlation arises from the dynamical effects of the Coulomb repulsion between electrons, and unlike exchange it keeps electrons with anti-parallel spin as well as parallel spin apart. Both exchange and correlation have the effect of lowering the energy of the system.

DFT [57] is a method in which the quantities we are interested in, such as the ground state energy, are written in terms of the electron density rather than the traditional methods, which use the vastly more complicated many-electron wavefunctions. For an N-electron system the many-body wavefunction is a function of $3N$ spatial variables, whereas DFT depends only on 3 spatial variables (x, y, z), and as such it is in principle much easier to implement.

Chemists have traditionally gone beyond Hartree-Fock by using the technique of configuration interaction in which a trial function consisting of a linear combination of Slater determinants is used to approximate the many electron wavefunction. This method is impossible for solid state physicists to use due to the large number of electrons involved, and this was largely the motivation for the development of DFT. However, it is now also widely used by chemists. DFT returns quite accurate results for ground state properties at a relatively low computational cost, and DFT is presently the most commonly used method for electronic structure calculations.

The first density functional theory was developed by Thomas [58] and Fermi [59] in the 1920's. They calculated the energy of an atom by representing its kinetic energy as a functional of the electron density, and combining this with the classical expression for the nuclear-electron and electron-electron interactions. The Thomas-Fermi equation however lacked accuracy. One reason for this was that no attempt was made to represent the exchange energy of an atom, predicted by the Hartree-Fock theory. Therefore, Dirac added an exchange energy functional in 1930 [60]. However, the theory was still not accurate enough, since it proved difficult to represent kinetic energy accurately with a density functional, and the effect of electron correlation was entirely neglected.

A big step forward for DFT came when in 1964 Hohenberg and Kohn [57] proved that it was possible to calculate the ground state electronic structure exactly, with only a knowledge of the electron density. They wrote the energy as a functional of the electron density in the form

$$E[\rho(\mathbf{r})] = \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + F[\rho(\mathbf{r})], \tag{2.10}$$

where V_{ext} is the external potential arising from the Coulomb interaction of the electrons and the nuclei. $F[\rho(\mathbf{r})]$ is a non-local functional which contains the sum of the kinetic energy of the electrons and the contribution from the electron-electron interactions. Equation (2.10) can then be minimised using a variational approach to find the ground state energy. However, the problem with equation (2.10) is that we do not know the form of the functional $F[\rho(\mathbf{r})]$, the difficulty being the treatment of the electron kinetic energy as well as exchange and correlation in the electron-electron interaction.

The year after Hohenberg and Kohn published their paper, Kohn and Sham [61] presented an approximation to the kinetic energy functional which greatly improved the accuracy of the method. They suggested that $F[\rho(\mathbf{r})]$ should be represented by the sum of three terms,

$$F[\rho(\mathbf{r})] = E_{KE}[\rho(\mathbf{r})] + E_H[\rho(\mathbf{r})] + E_{XC}[\rho(\mathbf{r})], \tag{2.11}$$

where $E_{KE}[\rho(\mathbf{r})]$ is the kinetic energy of a system of non-interacting electrons with the same density $\rho(\mathbf{r})$ as the real system. E_H, the Hartree energy, and E_{XC}, the exchange-correlation energy, describe the electron-electron Coulomb interaction. E_H is the electrostatic energy of the smeared out charge density and E_{XC} contains the corrections due to exchange and correlation.

Writing the terms in equation (2.11) explicitly, where the density $\rho(\mathbf{r})$ of the system is taken as the sum of the square moduli of a set of one-electron orthonormal orbitals we obtain the one-electron Kohn-Sham equation,

$$\left[-\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M}\frac{z_A}{r_{1A}} + \int\frac{\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_2 + V_{XC}(\mathbf{r}_1)\right]\psi_i(\mathbf{r}_1) = \epsilon_i\psi_i(\mathbf{r}_1), \tag{2.12}$$

where the first term represents the kinetic energy, the second term is the external potential interaction with M nuclei, the third term is the electron-electron Coulombic energy, the fourth term is the exchange-correlation potential, and ϵ_i are the orbital energies. The exchange-correlation potential is given by the functional derivative of the exchange-correlation energy,

$$V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}}{\delta\rho(\mathbf{r})}. \tag{2.13}$$

The ground state energy is then given by

$$E_0 = \sum_{i=1}^{N} \epsilon_i - \frac{1}{2}\int d\mathbf{r} V_H(\mathbf{r})\rho(\mathbf{r}) - \int d\mathbf{r} V_{XC}(\mathbf{r})\rho(\mathbf{r}) + E_{XC}[\rho(\mathbf{r})]. \tag{2.14}$$

Equation (2.12) is solved self-consistently, with an initial guess for the density yielding a set of orbitals from which an improved value for the density can be obtained; this is then used in the next iteration, and so on until the method converges. It is worth noting that the ϵ_i's in equation (2.12) are not really to be interpreted as one-electron energies, nor the ψ_i's as one-electron wavefunctions. It is only the combination given in equation (2.14) and the total charge density which are meaningful.

The key to success with DFT is having a good approximation for E_{XC} and V_{XC}. One of the simplest methods for calculating E_{XC} is the local density approximation (LDA). In this approximation it is assumed that the density varies slowly with distance, and the energy density at a point $\mathbf{r}$ in the real electron gas is equal to the energy density in a homogeneous electron gas that has the same electron density at point $\mathbf{r}$,

$$E_{XC} \approx \int dr \rho(\mathbf{r}) \epsilon_{XC}(\rho(\mathbf{r})), \tag{2.15}$$

where ϵ_{XC} is the exchange-correlation energy per electron of an infinite homogeneous electron gas with density equal to the local density $\rho(\mathbf{r})$.

The most common parameterisation for the exchange-correlation energy is that of Perdew and Zunger [62], which is based on the quantum Monte Carlo calculations of Ceperley and Alder of electron gases at various densities [63]. The LDA accurately describes many ground-state properties of solids, giving lattice constants and bulk moduli to within 1–2 % and 5–10 % respectively, of measured quantites [64]. However, excited-state properties like the bandgap of semiconductors are underes-

timated, based on the assumption that the ϵ_i's are one-electron excitation energies.

The most commonly used improvement to the LDA is the generalised gradient approximation (GGA). In GGA non-local functionals are used to describe the gradient of the density at each point in space, not just its value. Some examples of these functionals are given by the acronyms PW91 [65], B3YLP [66] and Becke88 [67].

Now that DFT has reduced the many-electron problem to a one-electron form we have to solve the Schrödinger equation (2.12). The wide range of methods and computer programs available use a variety of basis set expansions for the wavefunctions in equation (2.4). These include plane wave expansions combined with replacing the deep nuclear potential by the weaker pseudopotential, and the use of localised basis functions, as in SIESTA which is a very popular code for surfaces [68], carbon nanotubes [69, 70] and biomolecules [71–73]. DFT is now widely used due to its efficiency and accuracy, with successful applications in many fields including periodic systems such as crystals [74] and also in molecular systems including biomolecules [75]. We shall refer frequently to DFT calculations of the systems we study.

2.3 Tight-binding

The electronic structure method we choose to use in this work is tight-binding (TB) in the form of extended Hückel theory (EHT). Although the method is not self-consistent, extended Hückel is well established in its application to a wide range of organic molecules [76–78]. The overwhelming reason for this choice is the simplicity and speed of the method, which is of paramount importance when dealing with large molecules such as DNA.

Slater and Koster [79] developed the tight-binding method, following on from Bloch's [80] earlier work on the Linear Combination of Atomic Orbitals method (LCAO). In this method the wavefunctions are expanded in terms of a linear combination of localised orbitals, taken to be atomic-like functions χ, centred on each atom in the system,

$$\psi_n(\mathbf{r}) = \sum_i c_{n,i}\chi_i(\mathbf{r}). \tag{2.16}$$

To find the coefficients $c_{n,i}$ we use the variational principle to minimise the energy. However, atomic orbitals on one site are not orthogonal to those on other sites, giving a non-unit overlap matrix, S. Hence the secular equation giving the eigenvalues E_n has the form

$$\mid H_{ij} - E_n S_{ij} \mid = 0, \tag{2.17}$$

where the Hamiltonian matrix elements are given by

$$H_{ij} = \int dr \chi_i^*(\mathbf{r}) H \chi_j(\mathbf{r}), \tag{2.18}$$

and H is given by the expression in square brackets in equation 2.12. The overlap matrix is given by

$$S_{ij} = \int dr \chi_i^*(\mathbf{r}) \chi_j(\mathbf{r}). \tag{2.19}$$

Such a localised orbital basis set can be used in a first principles approach, but the method we choose to use is the semi-empirical extended Hückel theory, in which H_{ij} is proportional to S_{ij} [81].

An important simplification in this method is the two-centre approximation. Two-centre integrals occur when the potential and the orbital of atom i are in the same location, and the orbital of atom j is at another site. Three-centre integrals

occur when the interacting orbitals and the potentials are all in different locations. The three-centre integrals are smaller than the two-centre integrals and are thus ignored. If we consider the potential energy part of the Hamiltonian as the sum of spherical potentials located on the atoms of the system, and disregard the three-centre integrals, then the only part of the potential energy that remains is due to the sum of the spherical potentials located on the two atoms where the atomic orbitals lie.

We consider a vector $(\mathbf{R}_J - \mathbf{R}_I)$ between two atoms I and J to be an axis. We can express the atomic orbitals χ_i on atoms I and J, as a combination of functions which are space quantised with respect to that axis. For example a p-orbital may be expressed as a combination of $p\sigma$ and $p\pi_\pm$ functions with respect to the axis, where σ and π are the components of angular momentum around the axis. A σ bond lies along the axis, while a π bond lies perpendicular to the axis. This approach is useful since only orbital components of the same type, that is both σ or both π, give non-zero matrix elements, thus reducing the number of integrals.

Following Slater and Koster's formalism we symbolise p_x, p_y, and p_z functions by x, y, and z respectively. The direction cosines of the vector $(\mathbf{R}_J - \mathbf{R}_I)$ are given by l, m, n. For example, one integral may be symbolised as $E_{x,y}(l, m, n)$, meaning an integral in which the function χ_i is a p_x function and χ_j is a p_y type function, where both χ_i and χ_j can be represented by a combination of $p\sigma$ and $p\pi$.

Table 2.1 shows the energy integrals in terms of two-centre integrals for s, p and d orbitals. The indices are cycled, and the direction cosines changed to produce the entire range of formulae that describe all combinations of these orbitals.

Table 2.1: Inter-atomic matrix elements as functions of direction cosines (taken from Slater and Koster [79]).

$E_{s,s}$	$(ss\sigma)$
$E_{s,x}$	$l(sp\sigma)$
$E_{x,x}$	$l^2(pp\sigma) + (1-l^2)(pp\pi)$
$E_{x,y}$	$lm(pp\sigma) - lm(pp\pi)$
$E_{x,z}$	$ln(pp\sigma) - ln(pp\pi)$
$E_{s,xy}$	$\sqrt{3}lm(sd\sigma)$
E_{s,x^2-y^2}	$\frac{1}{2}\sqrt{3}(l^2-m^2)(sd\sigma)$
$E_{s,3z^3-r^2}$	$[n^2 - \frac{1}{2}(l^2+m^2)](sd\sigma)$
$E_{x,xy}$	$\sqrt{3}l^2m(pd\sigma) + m(1-2l^2)(pd\pi)$
$E_{x,yz}$	$\sqrt{3}lmn(pd\sigma) - 2lmn(pd\pi)$
$E_{x,zx}$	$\sqrt{3}l^2n(pd\sigma) + n(1-2l^2)(pd\pi)$
E_{x,x^2-y^2}	$\frac{1}{2}\sqrt{3}l(l^2-m^2)(pd\sigma) + l(1-l^2+m^2)(pd\pi)$
E_{y,x^2-y^2}	$\frac{1}{2}\sqrt{3}m(l^2-m^2)(pd\sigma) - m(1+l^2-m^2)(pd\pi)$
E_{z,x^2-y^2}	$\frac{1}{2}\sqrt{3}n(l^2-m^2)(pd\sigma) - n(l^2-m^2)(pd\pi)$

2.3.1 Extended Hückel Theory

There are various methods for implementing tight-binding electronic structure calculations. One empirical approach is Harrison's tight-binding method [82], in which the overlap matrix is simplified further by assuming that the basis set is orthonormal, so that

$$S_{ij} = \delta_{ij}. \tag{2.20}$$

The diagonal elements of the Hamiltonian matrix are given by the atomic ionization energies given in table 2.2, and the off-diagonal elements are given by

$$H_{ij} = V_{ij}, \tag{2.21}$$

where V_{ij} are the integrals treated as parameters – these integrals are obtained by considering the interactions between nearest neighbours only. However, this method does have some limitations, and it is unable to describe charge transfer for example, which is very important when considering the electrical properties of materials.

The method we use to calculate electronic structure is the semi-empirical extended Hückel theory, mainly chosen because of its simplicity. Originally applied to organic molecules, EHT has had many applications in calculating solid state band structures [83–87], and has also been used to calculate the Hamiltonian and overlap matrix elements for calculations of the electrical conductance through molecular wires, with success in describing experimental results [78, 88–92].

In the EHT scheme an orthonormal atomic orbital (AO) basis set is chosen for each atom I. The diagonal Hamiltonian matrix elements are given by the empirical valence shell ionization potentials. The non-diagonal elements between two atomic orbitals α, β centred on atoms I, J (where $I \neq J$) are approximated by setting them equal to the corresponding overlap, $S_{\alpha I,\beta J}$, of the two AO's, where

$$H_{\alpha I,\beta J} = K_{\alpha I,\beta J} S_{\alpha I,\beta J}. \tag{2.22}$$

$K_{\alpha I,\beta J}$ only depends on the on-site energies, $E_{\alpha I}$ and $E_{\beta J}$. There are different prescriptions for representing $K_{\alpha I,\beta J}$. In this work we have used

$$K_{\alpha I,\beta J} = K_{EHT} \frac{E_{\alpha I} + E_{\beta J}}{2}, \tag{2.23}$$

where K_{EHT} is a dimensionless constant, taken to be 1.75, following the work of Cerda [93].

In calculating the overlap integrals, each atomic orbital is approximated by a single Slater-type orbital, given by a product of a radial wave function and an appropriate spherical harmonic [94],

$$\chi(r,\theta,\phi) = R_{N,\eta}(r)Y_{l,m}(\theta,\phi), \tag{2.24}$$

where

$$R_{N,\eta}(r) = \frac{1}{\sqrt{(2N)!}}(2\eta)^{N+\frac{1}{2}}e^{-\eta r}r^{N-1}, \tag{2.25}$$

which includes a normalising factor. N is the principle quantum number, r is the distance from the centre of the atom, and η is the Slater exponent, describing the decay of the atomic orbital with distance from the centre. For d-orbitals the radial part of the wavefunction is expressed in terms of two Slater functions with coefficients η_1 and η_2. A list of the Slater exponents used is given in Table 2.2.

Table 2.2: Parameters for extended Hückel calculation. Valence ionisation potentials (H_{ii}) are in eV, Slater exponents η in au [95].

	shell	H_{ii}			η		η_1	η_2
atom	N	s	p	d	s	p	d	
H	1	−13.6	—	—	1.3	—	—	—
C	2	−21.4	−11.4	—	1.625	1.625	—	—
N	2	−26.0	−13.4	—	1.950	1.950	—	—
O	2	−32.3	−14.8	—	2.275	2.275	—	—
P	3	−27.3	−13.6	—	2.123	2.123	—	—
S	3	−20.0	−13.3	—	1.817	1.817	—	—
Cu	4	−11.4	−6.06	—	2.2	2.2	—	—
	3	—	—	−13.0	—	—	2.3(0.537)	5.95(0.593)
Au	6	−10.9	−5.6	—	2.60	2.58	—	—
	5	—	—	−15.0	—	—	6.16(0.685)	2.79(0.597)

A rather neat method for evaluating the multicentre integrals was devised by Sharma [96]. In his method a Slater type orbital may be expanded in the form $r^{k-l}e^{\pm\eta r}$, where k is an integer l is the order of the coefficient and η is the Slater exponent. The coefficients are given by a relatively simple algebraic expression. To begin with, I used Sharma's method, as implemented by a previous Cardiff student, to calculate the matrix S given in equation 2.19. However, I soon discovered spurious large eigenvalues, when studying ethylene, which were traced to the old computer code. Unfortunately we could not reconcile the problem, and therefore we wrote our own program to calculate the overlap matrix directly from the formulas given above.

EHT has proven to be an efficient and accurate method in my work, giving results for density of states and conductance that compare well with literature values.

Chapter 3

Conductance Theory

To understand the electronic properties of a molecule we must know its conductance, which gives us a quantitative measure of the suitability of molecules for molecular electronics. In this chapter I will describe how conductance may be expressed in terms of transmission, and I will also discuss possible mechanisms of transport through molecular wires.

Early descriptions of electrical conduction in metals and semiconductors were semiclassical, based on the Boltzmann equation [97] with the force on the electrons due to the applied field balanced by back scattering due to phonons and lattice defects. It was not until the 1950's when full quantum mechanical descriptions began to be used. Kubo's formalism became a popular method for finding conductance [98]. Kubo's formalism is a linear response theory, which describes how a system responds to a perturbation, for example the polarisation of a system in an electric field.

Nowadays the dimensions of electrical samples can be made so small that the quantum mechanical nature of electrons is essential to an understanding of the

conductance. First we consider a system in which a perfect conductor joins two infinite reservoirs. The conductor is taken as narrow enough only to contain a single transverse eigenstate, thus making this a one-dimensional problem. We apply a potential difference between the two reservoirs, taking the chemical potential on the left μ_1 to be higher than the chemical potential on the right μ_2, and so in the ground state of the system there is a current, j, flowing from left to right given by

$$j = -(\mu_1 - \mu_2)ev\frac{dn}{d\mu}, \tag{3.1}$$

where v is the velocity component along the tube, and $dn/d\mu$ is the density of states allowing for spin degeneracy, and only considering electrons travelling from left to right.

We can re-write equation 3.1, since in a one-dimensional system $dn/d\mu$ may be written as $(1/\pi\hbar v)$ and $(\mu_1 - \mu_2) = -e(V_1 - V_2)$, including spin degeneracy and considering only electrons travelling from left to right, giving,

$$j = \frac{e^2}{\pi\hbar}(V_1 - V_2). \tag{3.2}$$

Conductance Γ is defined as inverse resistance, in other words,

$$\Gamma = \frac{j}{V_1 - V_2}. \tag{3.3}$$

Then substituting for j in equation 3.3 the conductance can be expressed as,

$$\Gamma = \frac{2e^2}{h}. \tag{3.4}$$

This is the conductance for a perfect one-dimensional conductor. If we now introduce a potential barrier into the tube we then obtain the Landauer-Büttiker formula for the conductance

$$\Gamma = \frac{2e^2}{h}T, \tag{3.5}$$

where T is the transmission probability of an electron moving through the barrier.

The above expression for conductance can be generalised to include more than one eigenstate in a channel. To begin we consider a wave, $\phi(\mathbf{r})$ in the left-hand contact incident on the boundary S_1 in figure 3.1. From this figure it can be seen that we have split the system into sections. This sectioning is necessary for the implementation our embedding method. The embedding potential Σ_l is a tool for including the effects of the substrate in the Hamiltonian, a more detailed discussion of which is given in chapter 4.

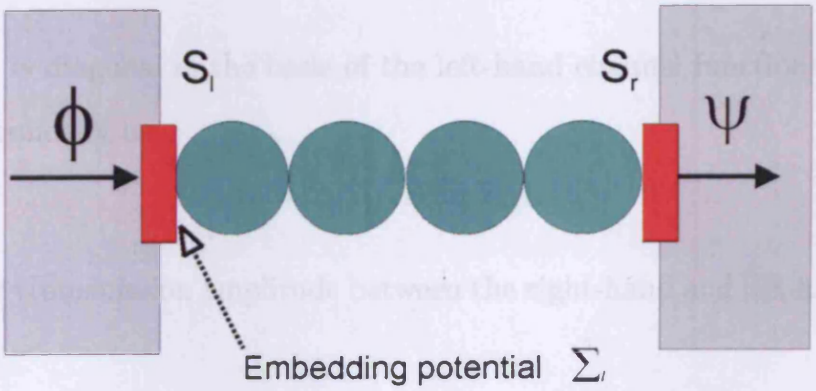


Figure 3.1: Diagram showing how the molecule has been split into sections, and embedded onto metal contacts.

If we consider the example shown in figure 3.1 with a wave $\phi(\mathbf{r})$ incident on the left contact from the bulk, then the full wavefunction everywhere to the right is

given by

$$\psi(\mathbf{r}) = 2i \int_{S_l} d\mathbf{r}_s \int_{S_l} d\mathbf{r}'_s G(\mathbf{r}, \mathbf{r}_s) \Im\mathrm{m}\Sigma_l(\mathbf{r}_s, \mathbf{r}'_s) \phi(\mathbf{r}'_s), \tag{3.6}$$

where G is the Green function for the whole system, that is, the conducting system embedded onto the contacts on either side [99].

We now take the incident amplitude $\phi(\mathbf{r}_s)$ to be the left-hand channel function $\phi_i(\mathbf{r}_s)$ – we shall define the channel functions to be the normalized eigenfunctions of the operator $\Im\mathrm{m}\Sigma_l(\mathbf{r}_s, \mathbf{r}'_s)$ (this is a real, symmetric operator). We also project $\psi(\mathbf{r})$ onto one of the right-hand channel functions $\phi_j(\mathbf{r}_s)$, an eigenfunction of the right-hand embedding potential. We then have a transmission amplitude between these channel functions given by

$$t_{ji} = 2i \int_{S_r} d\mathbf{r}_s \int_{S_l} d\mathbf{r}'_s \int_{S_l} d\mathbf{r}''_s \phi_j(\mathbf{r}_s) G(\mathbf{r}_s, \mathbf{r}'_s) \Im\mathrm{m}\Sigma_l(\mathbf{r}'_s, \mathbf{r}''_s) \phi_i(\mathbf{r}''_s). \tag{3.7}$$

As $\Im\mathrm{m}\Sigma_l$ is diagonal in the basis of the left-hand channel functions, this can be written symbolically as

$$t_{ji} = 2iG(j, i) \Im\mathrm{m}\Sigma_l(i, i). \tag{3.8}$$

Similarly, the transmission amplitude between the right-hand and left-hand channels is given by

$$t_{ij} = 2iG(i, j) \Im\mathrm{m}\Sigma_r(j, j). \tag{3.9}$$

Taking into account the flux in each channel, the transmission probability from channel i on the left to channel j on the right, and vice versa, may be written as

$$T_{ji} = t_{ji} t^*_{ij} = 4G(j, i) \Im\mathrm{m}\Sigma_l(i, i) G^*(i, j) \Im\mathrm{m}\Sigma_r(j, j). \tag{3.10}$$

Summing over all channels, $T = \sum_{i,j} T_{i,j}$ [100], we obtain the final expression for the total transmission (in atomic units)

$$T = 4\mathrm{Tr}[G_{lr}\Im m\Sigma_r G_{rl}^{*}\Im m\Sigma_l]. \tag{3.11}$$

This leads to the conductance formula for small voltages,

$$\Gamma = \frac{4}{\pi}\mathrm{Tr}[G_{lr}\Im m\Sigma_r G_{rl}^{*}\Im m\Sigma_l]. \tag{3.12}$$

This is the generalisation of equation (3.4), and has been given previously by Levi Yeyati *et al.* [101].

In realistic cases the transmission varies rapidly with energy and we must generalise equation 3.1 to find the current at finite voltage,

$$I = \frac{e}{h}\int dE[f_l(E-\mu_l) - f_r(E-\mu_r)]T(E), \tag{3.13}$$

where f_l and f_r are the Fermi functions in the left and right contacts, and $(\mu_l - \mu_r)$ is the finite voltage, V, taking $E = 0$ at the molecular Fermi energy, E_F. Equation (3.13) may then be conveniently written as,

$$I = \frac{2e}{h}\int_{E_F-eV/2}^{E_F+eV/2} T(E)dE. \tag{3.14}$$

An interesting question was raised by Landauer in 1981 [102], who pointed out that for a perfect conductor, equation 3.4 gives a resistance of h/Ne^2, where N is the number of channels. However, we would expect a perfect conductor to have zero resistance. The correct interpretation of this example was given by Imry in 1986 [103]

when he realised that the resistance was due to a Sharvin point contact resistance [104] – the carrier distribution in both reservoirs is described by an equilibrium Fermi function. However, in the conductor there is a net movement of charge from the high to low potential. The transition from an equilibrium distribution to a current carrying distribution at both ends of the conductor is a dissipative process, giving rise to the resistance.

The effect of contact resistance becomes more important as we consider ever smaller conductors. For a perfect conductor the contact resistance is given by

$$\Gamma_c^{-1} = \frac{h}{2e^2 N} \approx \frac{12.9k\Omega}{N}, \tag{3.15}$$

hence, a two terminal measurement of a narrow conductor that supports only one transverse mode would yield a resistance of $\sim 12.9k\Omega$ which is appreciable. This resistance decreases for larger conductors as the number of available modes increases.

We can calculate the number of modes available for transport by assuming periodic boundary conditions, giving allowed values of the transverse modes separated by $2\pi/W$, where W is the width of the conductor [105]. Hence, the number of modes can be written as

$$N = Int\left[\frac{W}{\lambda_F/2}\right], \tag{3.16}$$

where $Int(x)$ is a function that returns the integer that is just smaller than x, and λ_F is the Fermi wavelength. DNA has a width of around 15 nm, so assuming DNA is a semiconductor with a typical Fermi wavelength of 30 nm this gives only one allowed mode for transport within the molecule. Let us now consider the number of modes available at the contact between the gold reservoir and a single sulphur atom on the end of the DNA. Making use of equation 3.16 we can substitute the Fermi

wavelength of gold (~ 0.1 nm), and the atomic radius of sulphur (~ 0.1 nm) to find that there are only 4 allowed modes, and hence an appreciable contact resistance.

If we now introduce a scattering site into our perfect conductor, then we can ask the question, where is the energy of the electrons dissipated after tunneling through a potential barrier? If we consider a conductor with N modes containing just one scatterer with a transmission probability of T, then we may split the total observed resistance into that due to the contact plus that of the scatterer in series [105]. Hence, we may write equation (3.5) as

$$\Gamma^{-1} = \frac{h}{2e^2NT} = \frac{h}{2e^2N} + \frac{h}{2e^2N}\frac{(1-T)}{T}. \tag{3.17}$$

The first term on the right hand side is the contact resistance and the second term is due to the scatterer. The potential drop across the scatterer, V_s, is then given by

$$V_s = \frac{1}{e}(1-T)(\mu_1 - \mu_2), \tag{3.18}$$

but where is the heat associated with the resistance, j^2/Γ_s, lost? If the scatterer is rigid with no internal degrees of freedom then it cannot dissipate the heat itself. Instead the heat is dissipated in the contact reservoirs as the relatively high energy electrons suffer phonon scattering in the reservoirs that brings their energy down to that of the Fermi level in the reservoir. This is similar to the contact resistance discussed earlier. Experiments have confirmed this theory; in 1998 Frank *et al.* [106] recorded currents flowing through carbon nanotubes that should have produced temperatures of 20,000 K if the energy had been dissipated in the tube.

The discussion above concerning a conducting tube connected at either end by

a contact may be extended to a three or four terminal measurement, where it is desirable to eliminate the contact resistance by measuring the current and voltage from different contacts [100]. Here we are only concerned with two terminal measurements.

3.1 Electron transport and transfer

It is important to make a distinction between the seemingly similar processes of electron transport and electron transfer. As described in section 1.3 chemists usually use indirect measurement techniques instead of the direct measurements preferred by physicists. It turns out that the chemists are measuring electron transfer and the physicists electron transport [107].

Charge transfer is the question of how quickly and effectively an electron or hole can be transferred from one part of the system to another. Charge carriers are usually injected by another molecule attached to a specific site, then removed by another molecule at a site a known distance away [108]. These processes are important for the correct functioning of a molecule. However, it does not directly tell us about the conductivity of the molecule. To determine the conductivity we must consider electron transport. For electron transport there must always be charge carriers moving through the molecule, even with no externally applied force the charge carriers move but with a net charge of zero [109]. This is the process that physicists are currently trying to understand.

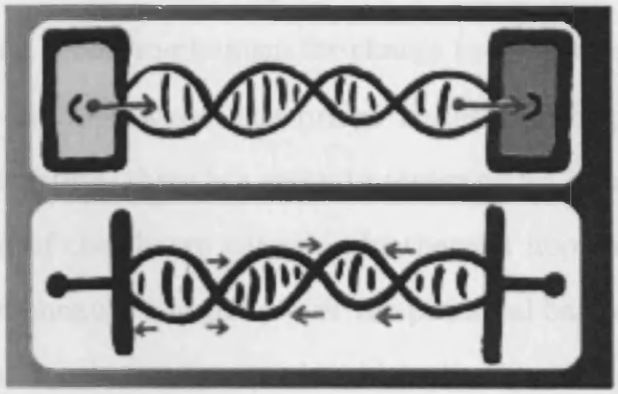


Figure 3.2: Diagram of charge transfer (top) and charge transport (bottom), taken from [109].

3.2 Transport mechanisms

In this short section I will describe the two main processes by which charge carriers can move through a molecule, and in particular DNA.

The first process is coherent transfer. In this model an electron must tunnel through a potential barrier to pass from a donor to an acceptor site. For example, an electron on a G base, which has a relatively low energy, must tunnel through the barrier posed by a neighbouring T or A base, until it reaches another G base. This rate of transfer decays exponentially with distance. The process is termed coherent since the electron does not exchange any energy with the molecule during transfer, and the electron is never localised. This process is often referred to as superexchange in chemistry literature [110].

The coherent transfer mechanism described above can be applied to the movement of electrons at low temperatures, where the thermal energy of the carrier is much less than the potential barriers that oppose conduction. However, at higher temperatures coherent transfer breaks down since the thermal disorder further lo-

calises the charge, and a new mechanism for charge transfer must be applied.

For higher temperatures, where the bridge states are comparable or lower in energy than the initial state, there is a second process called thermal hopping which describes the motion of the charge carriers. In thermal hopping the electrons are thermally excited via phonons and jump over the potential barrier rather than going through it. This is an incoherent process in which the electrons are localised on the molecule and exchange energy with it. The rate of thermal hopping decays more slowly with distance than coherent transfer.

In the model that we use for conductance calculations we assume $T = 0$ K, hence we have no thermally assisted hopping and therefore only consider the mechanism of coherent transfer.

3.3 Theoretical studies of DNA

In order to try to understand the conflicting experimental evidence of conduction experiments on DNA, many theorists have been concentrating their work on solving this transport problem. In this chapter I will discuss some of the main models that have been used to calculate electronic structure and transport properties of DNA.

3.3.1 Density Functional Methods

The basis of DFT has been given in Section 2.2. This method has become a very popular way for determining the electronic properties of molecules, and in 1997 Lewis *et al.* [111] calculated the first electronic structure of 10 base-pair poly(G)-poly(C) B-DNA, which is a right-handed helix with 10 base-pairs per turn that is the most common conformation *in vivo*. In their model they ignored solvent

effects and the presence of counter-ions, and considered the two strands of DNA separately, splitting the Hamiltonian into two parts. Firstly they treated the strong intramolecular interactions within each strand, with DFT using the LDA and the pseudopotential method. Secondly they used a simplified theory accounting for the intermolecular H-bonds between the two strands of DNA, where the Hamiltonian of the weak interaction is a sum of the electrostatic and exchange contributions based on many-body interactions, plus an overlap contribution coming from a Löwdin orthogonalisation transformation between strands. Lewis *et al.* then used a linear scaling method to solve the resulting sparse Hamiltonian and overlap matrices. They presented their results for the density of states of a single GC base-pair and also for the 10 base-pair molecule. They find that there is a bandgap separating the HOMO (highest occupied molecular orbital) from the LUMO (lowest unoccupied molecular orbital) of 3.4 eV in the single base case, which narrows to 1.4 eV for the full molecule. They explain that this is due to the addition of the backbone states and broadening due to coupling with the other base-pairs.

In 2001 Hjort and Stafström [112] calculated the current-voltage characteristics of a 20 base-pair poly(G)-poly(C) DNA molecule using DFT and the Landauer formula. The structure of the molecule was obtained by taking experimental coordinates of one GC base-pair then translating and rotating the pair to build up the full regular structure of the 20 base-pair molecule. Hjort and Stafström also included the effects of metallic contacts via a tight-binding scheme, and introduced temperature effects by rotating the base stack around the backbone by a random amount, then calculating the conductance. To reduce the computational cost of their calculations they ignored the backbone, replacing it with a hydrogen atom on each base to satisfy the dangling bonds; they also treated the two base stacks individually, ignoring

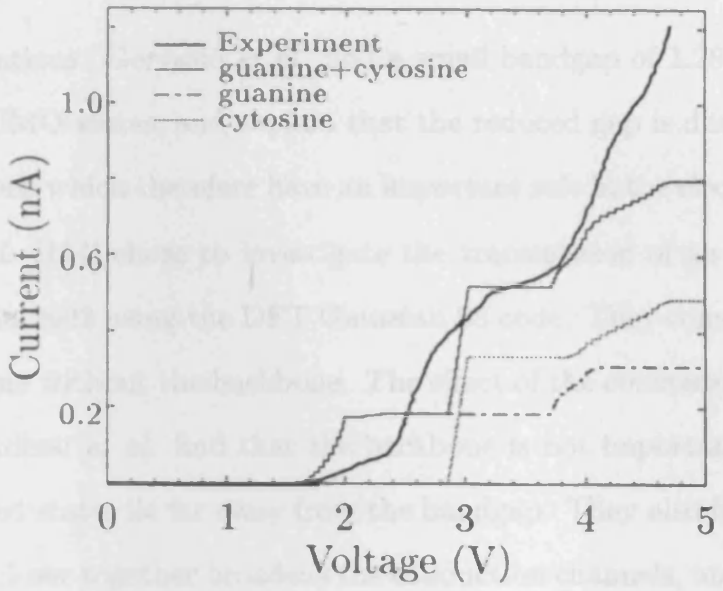


Figure 3.3: Results of Hjort and Stafström calculations for poly(G)-poly(C) DNA (taken from [112]).

interactions across the H-bond. The results show semiconducting behaviour for poly(G)-poly(C) DNA, in agreement with the experimental results of Porath *et al.* [22]. They find a stepped *I-V* curve due to the contribution to the conductance from individual bands of transmission. Hjort and Stafström also find an increase in threshold voltage with increasing temperature, due to increased localisation of the charge carriers arising from structural disorder within the molecule.

In 2002 Gervasio *et al.* [113] used DFT to calculate the electronic structure and charge density of a 12 base-pair Z-DNA molecule; this form of DNA is a left-handed helix with 12 base-pairs per turn, which is narrower and more elongated than B-DNA. They included the effects of solvents and counterions by explicitly including 138 H_2O molecules and 24 Na atoms in the calculation. The structure was optimised using the Car-Parrinello molecular dynamics code simultaneously with the electronic

stucture calculations. Gervasio *et al.* find a small bandgap of 1.28 eV between the HOMO and LUMO states, and explain that the reduced gap is due to the presence of the counterions which therefore have an important role in the electronic structure.

Adessi *et al.* [114] chose to investigate the transmission of an infinite chain of poly(G) DNA in 2003 using the DFT Gaussian 98 code. They consider two models, one with and one without the backbone. The effect of the counterions is included in their model. Adessi *et al.* find that the backbone is not important for conduction as the associated states lie far away from the bandgap. They also find that bringing the base-pairs closer together broadens the conduction channels, and that a periodic arrangement of the Na counterions improves conduction while aperiodically spaced counterions have the opposite effect.

The inclusion of counterions varies between models. There is usually one counterion per nucleotide, included to counteract the negative charge on the phosphate groups along the backbone of the DNA molecule. In my model I do not include counterions, however, the screening effect of water molecules is included by taking experimental geometries or having a $1/r^2$ distance dependence when minimising computer generated structures.

3.3.2 Tight-Binding Methods

As discussed in Section 2.3, tight-binding methods are often used since they tend to be computationally less expensive than more sophisticated descriptions of the electron Hamiltonian. This allows the study of much larger systems, which would otherwise be prohibitively expensive. In 2003 Tada *et al.*[77] used the extended Hückel molecular orbital method and the Landauer formalism to calculate

the current-voltage characteristics of a DNA molecule consisting of 12 AT base-pairs. The structure was taken from the protein data bank; however, the AT base-pairs were arranged so that the A bases do not all lie on the same strand and there is also no regular pattern in the base sequence. Tada *et al.* use a cluster of 80 gold atoms, with a 4x4 atom surface, to probe different parts of the DNA to see how the conductance varied. Since they use a finite cluster of atoms rather than a semi-infinite reservoir the Green function describing the leads does not have the correct boundary conditions and will give a spiky density of states. Presumably, they broaden this out to obtain the density of states and the Green function corresponding to gold joined to a reservoir. In their model of the contacts each atom on the surface of the cluster interacts with all the nearby atoms on the DNA molecule via weak coupling, and there are no covalent bonds between the metallic leads and the DNA. Their results suggest that DNA behaves like a semiconductor when the leads are attached to the bases; however, when contact is made to the backbone insulating behaviour is observed. They also find larger currents when the gold leads contact the A bases rather than the T bases, with values of conductance ranging between $10^{-9}\Omega^{-1} \rightarrow 10^{-11}\Omega^{-1}$.

In 2002 Zwolak and Di Ventra [115] used tight-binding methods to calculate spin-dependent transport in DNA between ferromagnetic Fe contacts. They use the Landauer-Büttiker formula to calculate the current, taking the effect of the metallic leads into account via self-energies. Zwolak and Di Ventra find that spin-dependent transport can be found in short DNA molecules that are between ferromagnetic contacts, and they suggest that this might have applications in molecular electronics.

3.3.3 Model Hamiltonian Methods

In model Hamiltonian studies the system is simplified even further than in the tight-binding models described above. Here the DNA is represented by a strictly one-dimensional chain with nearest neighbour hopping. This approach has been mainly used to study temperature dependence and incoherent effects in transport.

In 2001 Yu and Song [116] performed model tight-binding calculations on λ-DNA. They treat the molecule as a one-dimensional disordered system, and include the effect of temperature by twisting the bases by modifying the hopping term correspondingly. They find a strong temperature dependence of the conductivity at high temperatures and a low temperature dependence at lower temperatures. They explain this result by a variable range hopping mechanism, where the probability of hopping is maximised by the interplay between distance, temperature and the energy difference between energy levels.

Roche [117] used the same model Hamiltonian as Yu and Song to calculate the transmission probabilities of coherent transport through both λ-DNA and poly(G)-poly(C) DNA molecules. Roche includes the effect of two semi-infinite electrodes, describing them by using a tight-binding Hamiltonian. He includes temperature effects by twisting the bases. This is done by modifying the hopping term correspondingly, and he finds that increased temperature leads to less transmission. Roche also increases the length of the λ-DNA chain and finds that the tranmission is reduced with increasing length.

Zhang and Ulloa [118] used a one-dimensional tight-binding model to calculate *I-V* curves for a model system of random DNA sequences containing 562 base-pairs. They perform their calculations at 300 K, and the temperature effect is included via

the Fermi function that describes the electron distribution in the reservoirs, and in the first part of the work they do not include any molecular vibrations. Zhang and Ulloa use the Landauer-Büttiker formula to calculate the current and find behaviour ranging from insulating to metallic as they improve the level of order in the system. Subsequently they also investigate torsional motion of DNA bases and find that this motion suppresses electron movement.

It is a consequence of the variety of experimental results for DNA that has given rise to such a wide range of theoretical models. Most models find semiconducting behaviour for short DNA molecules, with a range of bandgaps, although some results show insulating properties or metallic conduction. Effects such as temperature, base sequence, counter-ions and the role of contacts are treated differently in different models, with the hope of explaining some of the contradictary experimental results.

Chapter 4

Growing Molecules by Embedding

Embedding was developed as a way of solving the Schrödinger equation in large systems that can be sub-divided into smaller units [119]. An embedding potential, added onto the Hamiltonian for part of the system, allows the Schrödinger equation to be solved for just this part, with the wavefunctions correctly matched onto the surroundings. The method was originally used to calculate electronic structure of surfaces and interfaces, within a plane-wave basis set. We now use the concepts of embedding and apply it to our tight-binding problem to develop a method that can be applied to large molecules. Tight-binding embedding uses Dyson's equation and Green functions (GF) to find the embedding potential. We can then treat the molecule as a series of sections, adding a section at a time to build up the entire chain, embedding as we go along. This method has the advantage that the computational time for solving the Schrödinger equation scales linearly with the size of the system, unlike traditional methods that scale as $O(N^3)$. There are other order-N methods for solving these sorts of problems, such as the localised orbitals method [120], and density matrix methods [121]; however, our method directly yields the

GF, and seems well suited for conductance problems. The ability to perform order-N calculations allows us to treat larger systems than would otherwise be feasible. An analogous method to ours was applied by Crampin *et al.* [122] to "grow" large interface systems, adding atomic layer by atomic layer.

A schematic diagram of embedding one section of the molecule onto another is shown in figure 4.1. What we aim to do is to find a term to be added to the Hamiltonian of each section, which replaces the effect of the rest of the system. Figure 4.1 shows three sections – in the case of DNA we split the 12 base-pair molecule into 12 such sections, each containing a base-pair plus the associated sugar-phosphate backbone. The shaded areas in figure 4.1 represent the regions of each section in which orbital overlap occurs, within some cut-off (this is taken to be 8 au in our extended Hückel calculation). Let us assume that we have already calculated the GF for the isolated section 1: we can now find an embedding potential for section 2 which contains all the interactions with 1. From the GF for section 2 embedded onto 1 we can find an embedding potential for 3 onto 2, allowing us to find the GF for 3 embedded onto the entire system to the left, and so on.

To find the embedding potential in a tight-binding system, we use Dyson's equation. The GF of an unperturbed system is given by

$$(H - ES)G_0 = I, \tag{4.1}$$

where E is the energy, and H and S are the Hamiltonian and overlap matrices respectively. When we include a perturbation δ to the system the equation becomes

$$(H + \delta - ES)G = I. \tag{4.2}$$

Let us first consider embedding section 2 onto section 1 on the left, in which case the unperturbed Hamiltonian matrix of the two sections consists of H_{11} and H_{22}, which do not interact with each other. The two sections are coupled together by H_{12}, H_{21} from the interacting region represented by the shaded region in figure 4.1, giving the following Hamiltonian,

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \tag{4.3}$$

and we treat H_{12} and H_{21} as the perturbation δ in (4.2).

If we now multiply (4.1) with G and multiply (4.2) with G_0, then subtracting these two equations we obtain Dyson's equation,

$$G = G_0 - G_0 \delta G. \tag{4.4}$$

Expanding (4.4) gives

$$G = G_0 - G_0 \delta G_0 + G_0 \delta G_0 \delta G. \tag{4.5}$$

Applying equation (4.5) directly to our problem we obtain

$$\tilde{G}_{22} = G^0_{22} + G^0_{22} \delta_{21} G^0_{11} \delta_{12} \tilde{G}_{22}, \tag{4.6}$$

where the tilde on $\tilde{G}$ indicates that this is the GF of a section embedded only on the left, and G^0_{11} is the unperturbed GF in section 1. The second term in (4.5) goes to zero since G^0_{12} is zero, because there are no links between regions 1 and 2 in the unperturbed system. Comparing (4.6) with (4.4) we can see that the series is the same as if we take the perturbation δ in (4.4) to be $-\delta_{21} G^0_{11} \delta_{12}$ acting entirely within

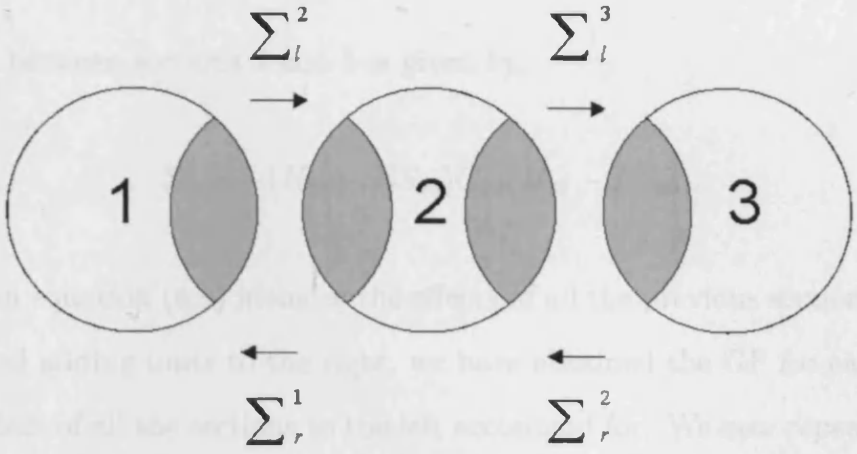


Figure 4.1: Schematic representation showing three sections of the DNA molecule, with embedding potentials Σ_l and Σ_r which embed to the left and right.

the space of region 2. Hence, we take the perturbation to be given by

$$\Sigma_l^2 = -(H_{21} - ES_{21})G_{11}^0(H_{12} - ES_{12}), \tag{4.7}$$

which is the embedding potential, embedding 2 onto 1. The overlap matrix S in (4.7) has to be taken into account due to the non-orthogonality of the basis set. This perturbation is then added to the unperturbed Hamiltonian matrix of section 2 to give an effective Hamiltonian H_{eff}, and then the GF for this section embedded to the left, can be calculated from

$$\tilde{G}_{22} = (H_{\text{eff}} - ES)^{-1}. \tag{4.8}$$

We note that our result (4.7) for the embedding potential has been given previously [114], described as a self-energy.

In this way we proceed to build up the chain from left to right so that the

interaction between sections 2 and 3 is given by,

$$\Sigma_l^3 = -(H_{32} - ES_{32})\tilde{G}_{22}(H_{23} - ES_{23}). \tag{4.9}$$

Using $\tilde{G}_{22}$ in equation (4.9) includes the effects of all the previous sections. Once we have finished adding units to the right, we have obtained the GF for each sub-unit with the effect of all the sections to the left accounted for. We now repeat the above process, but this time building up the molecule from right to left, so that

$$\Sigma_r^2 = -(H_{23} - ES_{23})G_{33}^0(H_{32} - ES_{32}), \tag{4.10}$$

and

$$\Sigma_r^1 = -(H_{12} - ES_{12})\bar{G}_{22}(H_{21} - ES_{21}), \tag{4.11}$$

where $\bar{G}$ is the GF of a section embedded on the right. Once we have finished adding sub-units to the left, we now have the left and right embedding potentials for all sections. We then simply add these embedding potentials to the unperturbed Hamiltonian for each section, and in this way we are able to calculate the GF for each section of DNA, no matter where it lies in the chain, with the effect of all the other sections in the molecule taken into account. So in the example above,

$$G_{22} = (H_{22} + \Sigma_l^2 + \Sigma_r^2 - ES_{22})^{-1}. \tag{4.12}$$

This gives us a method of obtaining the GF, which scales linearly with the size of the system, and can be applied to any Hamiltonian with localised orbitals, allowing for self-consistent calculations. Our method can be applied directly to all molecules

that have a linear sequence, with the assumption that only neighbouring sections have direct orbital overlap. This method of "growing" molecules can be related to methods of diagonalising tri-diagonal block matrices [123]. In section 4.2 we describe how we embed metal contacts onto the molecule in the same way. Of course many biological molecules such as proteins, though of underlying linear structure, are folded back on themselves. This adds extra interactions between sections that are not nearest neighbours, and must be included. Since we use a linear strand of DNA these interactions are not included in our model.

4.1 Density of States, an Embedding Approach

The density of states (DOS) is the most basic quantity of electronic structure, and is fundamental for determining the properties of a material. The DOS is a measure of the number of electronic states within a narrow interval of energy,

$$n(E) = \sum_i \delta(E - E_i), \tag{4.13}$$

where i runs over all the states in the system. From the density of states we can find the Fermi energy of the system and hence the energy of the highest occupied molecular orbital (HOMO), and the lowest unoccupied molecular orbital (LUMO). This information can tell us about the electrical conduction properties of a molecule – in the case of a semiconductor the Fermi energy will lie within a bandgap between the HOMO and the LUMO, whereas for a good conductor there will be no gap between the valence and conduction bands. The density of states also provides information about the stability of the system. If the Fermi energy lies on a peak in

the density of states then the material will tend to be unstable. The material will then undergo a structural transistion, in which the atoms are rearranged so that the energy of the occupied states is lowered, opening up an energy gap in which the Fermi energy will now lie [124].

I now show how the DOS can be calculated in our tight-binding embedding formalism. We can relate equation (4.13) to the full GF $G(\mathbf{r}, \mathbf{r}'; E)$ using,

$$n(E) = \frac{1}{\pi} \int d\mathbf{r} \Im m G(\mathbf{r}, \mathbf{r}; E + i\epsilon). \tag{4.14}$$

Expanding G in terms of the basis functions we have

$$G(\mathbf{r}, \mathbf{r}; E) = \sum_{i,j} G_{ij}(E) \chi_i(\mathbf{r}) \chi_j(\mathbf{r}), \tag{4.15}$$

and substituting (4.15) into (4.14) gives the DOS as

$$n(E) = \frac{1}{\pi} \Im m \mathrm{Tr}(GS). \tag{4.16}$$

The trace of this product matrix can be written as

$$\mathrm{Tr}(GS) = \sum_{(n,i),(m,j)} G_{(n,i),(m,j)} S_{(m,j),(n,i)}, \tag{4.17}$$

where n and m label neighbouring, or identical sections of DNA (as in figure 4.1), with i and j labeling the orbitals in n and m, respectively. The sum in (4.17) runs over all sections. However, there are only contributions to the sum when $m = n, (n-1)$, or $(n+1)$, due to the short-range of the overlap. Therefore, we can

rewrite the trace as

$$\mathrm{Tr}(GS) = \sum_n \left(\sum_{i,j;n} G_{(n,i),(n,j)} S_{(n,j),(n,i)} + \sum_{i,j;n-1} G_{(n,i),(n-1,j)} S_{(n-1,j),(n,i)} + \sum_{i,j;n+1} G_{(n,i),(n+1,j)} S_{(n+1,j),(n,i)} \right), \tag{4.18}$$

where i is an orbital in section n, and j is an orbital in section m.

We find the first term in (4.18) directly from our embedding procedure – as we have seen we obtain the GF between all orbitals within the same section directly from equation (4.12). The second term involves $G_{(n,i),(n-1,j)}$, which is the GF between orbitals in one section with those in the previous section, and can be derived using Dyson's equation (4.4), given by

$$G_{(n,i),(n-1,j)} = G^0_{(n,i),(n-1,j)} - \sum \tilde{G}_{(n,i),(n,k)} \delta_{(n,k)(n-1,l)} G_{(n-1,l),(n-1,j)}. \tag{4.19}$$

The first term on the RHS in equation (4.19) is zero since there are no links between sections n and m in the unperturbed system. Hence,

$$G_{(n,i),(n-1,j)} = -\sum \tilde{G}_{(n,i),(n,k)} \delta_{(n,k)(n-1,l)} G_{(n-1,l),(n-1,j)}, \tag{4.20}$$

where $\delta_{(n,k)(n-1,l)}$ in (4.20) is given by

$$\delta_{(n,k)(n-1,l)} = H_{(n,k)(n-1,l)} - E S_{(n,k)(n-1,l)}. \tag{4.21}$$

As the tilde implies, $\tilde{G}_{(n,i),(n,k)}$ is the GF of the nth section embedded only on the left.

The third term in (4.18) counts the contributions between the current section and the next. However, these same contributions have already been calculated when $m = n - 1$, therefore we need not calculate them again and can simply drop the third term in (4.18), multiplying the second term by a factor of 2.

In this way we were able to write programs that calculate the total DOS of the molecule. We did this by inverting and diagonalising the Hamiltonian and overlap matrices for each section. This allows us to work with a number of small matrices, rather than having to invert the Hamiltonian matrix of the entire molecule, saving an enormous amount of time.

4.2 Embedding Approach to Transmission

The quantity from which we can obtain a quantitative measure of the electrical properties of DNA is the conductance, and it can be formulated in terms of the transmission coefficients, T_{ij}, between electron channels i, j in the contacts at each end of the molecule [100]. These channels usually correspond to the incident and transmitted Bloch states at a particular energy.

As explained in chapter 3, the total transmission in the embedding scheme can be written concisely in terms of the GF between the contacts as [125, 126]

$$T = 4\mathrm{Tr}(\hat{G}_{lr}\Im\mathrm{m}\Sigma_r\hat{G}^*_{rl}\Im\mathrm{m}\Sigma_l). \tag{4.22}$$

In this formula, which does not include spin degeneracy, $\hat{G}_{lr}$ is the GF for the molecule connected to the contacts, between the left-hand and right-hand contacting orbitals, and Σ_l and Σ_r are the embedding potentials which couple these orbitals

to the corresponding contacts. The trace contains the sums over channels, and as this is independent of representation, we need not worry about the explicit form of the channels. This result has been known in a local orbital representation for several years [127, 128], though, as mentioned earlier in this chapter, in this context Σ is usually called the self-energy [114]. This is exactly the same as our embedding potential.

Recently, the same result has been derived and used in the framework of embedding theory, in which the embedding potential is defined over an embedding plane separating the embedded region from the substrate [125]. It can be used to find, for example, the conductance of an interface between metals, when $\hat{G}_{lr}$ is the GF for the whole system between the left and right embedding planes, and $\Sigma_{l/r}$ are the embedding potentials on those planes. This was applied by Wortmann *et al.* [125] to study spin-polarised transmission through a ferromagnetic Co monolayer sandwiched between Cu.

In our application we couple a single atom on the molecule to a single atom on the metallic contact. This is an approximation to the usual experimental arrangement where the metal-DNA contact extends over several atoms [21], but the extension to multi-atom contact is straightforward. In this case, the matrices given in equation (4.25) are simply extended to include all of the contacting atoms. However, we note that single-atom contact experiments are indeed possible, as shown in the work of Agraït *et al.* [129]. This makes our assumption of a single contact atom more plausible, and in any case it is revealing about the effect of changing the contact atom.

If we are to calculate the transmission between two atoms using (4.22), we must first find the GF $\hat{G}_{lr}$ linking these atoms, which may be located anywhere along the

molecule. The GF coupling the two atoms embedded onto the leads, $\hat{G}_{lr}$, can be derived from Dyson's equation,

$$\hat{G}_{lr} = G_{lr} - G_{ll}\Sigma_l\hat{G}_{lr} - G_{lr}\Sigma_r\hat{G}_{rr}. \tag{4.23}$$

Here G_{lr} is the unperturbed GF linking atom l to atom r, without the metal contacts. Σ_l and Σ_r are the embedding potentials linking the left and right metal reservoirs to the molecule on these atoms. However, we do not know the quantity $\hat{G}_{rr}$, which is the GF of the right hand atom connected to the metal contact. We can, however, write $\hat{G}_{rr}$ in terms of other quantities that we do know,

$$\hat{G}_{rr} = G_{rr} - G_{rl}\Sigma_l\hat{G}_{lr} - G_{rr}\Sigma_r\hat{G}_{rr}. \tag{4.24}$$

Re-arranging (4.24) for $\hat{G}_{rr}$ and substituting into (4.23) we obtain the GF linking the two atoms as

$$\hat{G}_{lr} = [1+G_{ll}\Sigma_l - G_{lr}\Sigma_r(1+G_{rr}\Sigma_r)^{-1}G_{rl}\Sigma_l]^{-1}[G_{lr} - G_{lr}\Sigma_r(1+G_{rr}\Sigma_r)^{-1}G_{rr}]. \tag{4.25}$$

This GF formula takes into account the effect of the metal contacts on both ends of the molecule. This effect has also been included in the work of Cuniberti *et al.* [31] and Damle *et al.* [130], through the use of self-energies. In our model we choose to use two metal reservoirs of either Au or Cu to make contact with the molecule. The embedding potential Σ_l (4.7), embedding the molecule onto the left reservoir is given by

$$\Sigma_l = -(H_{l\mathrm{Cu}} - ES_{l\mathrm{Cu}})G^0_{\mathrm{CuCu}}(H_{\mathrm{Cu}l} - ES_{\mathrm{Cu}l}), \tag{4.26}$$

where G^0_{CuCu} is the unperturbed GF in the left metal lead, with a similar expression for Σ_r.

In the case of atom-to-atom contact between the molecule and the metal lead the GFs and embedding potentials are given by matrices, because of the multiple orbitals involved. However, when beginning our investigation of G_{lr} we considered only a single metallic s-orbital making contact with a single orbital on the molecule. This single orbital contact between the metal and molecule allows for only one conduction channel in and out of the molecule. The coupling terms in (4.25) are evaluated using the extended Hückel method, assuming realistic metal–molecule distances [131], but with an arbitrary angle between the molecule and the metal surface. To describe the metal contact we use a full electronic structure calculation for the surface of semi-infinite Cu (001) or Au (001). This uses the embedded linearised augmented plane wave method [132], which gives very accurate results for the density of states on the surface atoms.

Since we only have a single atom/orbital molecule-atom contact in this model, we project the total surface density of states of the metal onto the atom/orbital making contact. From this, the imaginary part of the GF is given by,

$$\Im m G^0_{CuCu}(E) = \pi n_{Cu}(E), \tag{4.27}$$

where n_{Cu} is the total surface density of states on the metal contact atom/orbital. The real part of the GF can then be found from the Kramers-Kronig relation [98], and is given by

$$\Re e G^0_{CuCu}(E) = \int_{-\infty}^{\infty} dE' \frac{n_{Cu}(E')}{E' - E}. \tag{4.28}$$

In the multiple-orbital case the s, p, and d components of the density of states of

a Au surface, are projected onto a single Au atom. The Green function can therefore be written in the spherically symmetric form,

$$G^0_{\mathrm{CuCu}} = \sum_L G_L \phi_L(r)\phi_L(r'), \tag{4.29}$$

where L is angular momentum. This ignores cross terms caused by the lack of spherical symmetry.

When calculating the transmission there are two different ways in which we treat the density of states of the metal contact. In one method we find the GF, G^0_{CuCu}, evaluated at the Fermi energy of Cu, and use this fixed value to determine Σ_l, Σ_r, $\hat{G}_{lr}$, and hence the transmission of the molecule over the energy range of the molecular DOS. The method represents doping the molecule to calculate transmission for small voltages. The second, and more realistic method is to shift the Fermi energy and hence the DOS of the left contact up by a fixed energy and shift the Fermi energy of the right contact down by a fixed energy (see figure 4.2). This second method gives a more realistic representation of the finite potential difference between the two contacts. For a complete description of the metal-molecule contact we would need to shift in energy the density of states of the metal by the energy at which the transmission is calculated. However, since our current-voltage calculations involve an integration we would need to calculate the whole range of transmission for each small energy step. This would be prohibitively slow and therefore we choose to fix the energy shift of the metal density of states to half of the full voltage range.

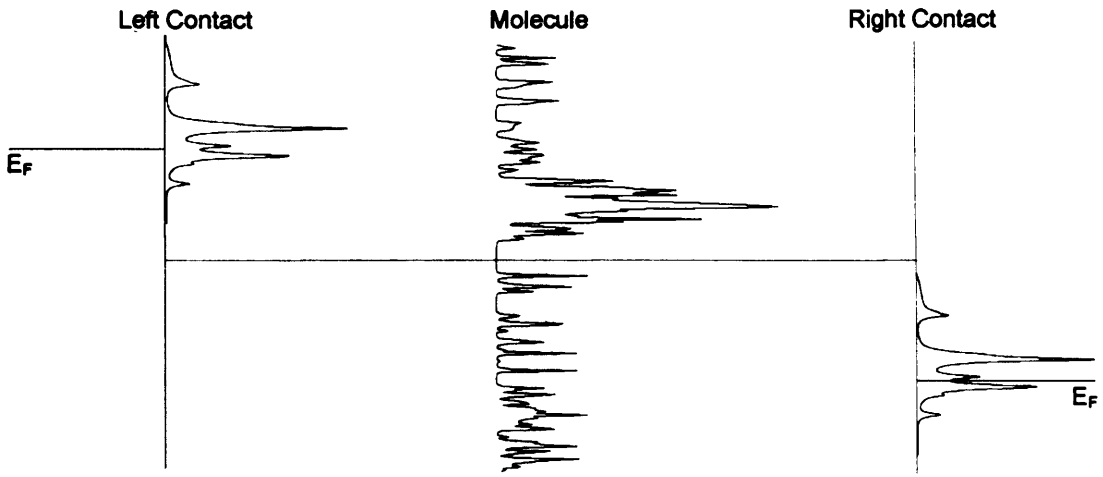


Figure 4.2: Schematic representation showing the left and right metallic contact with the DOS shifted to represent a voltage difference between the two sides.

4.3 Transmission Results for Carbon Chain

Before applying our embedding method to DNA, we test it on a model system of a linear chain of 12 C atoms aligned along the x direction, spaced by 1.53 Å. A Cu reservoir is attached to the p_x orbitals on either end of the chain, and the total end-to-end transmission of the carbon chain is calculated as a function of energy. The DOS and transmission of this small system are calculated both directly, without any partitioning of the system, and with embedding. Both cases give exactly the same results. We choose this simple system to test our method because we know that the orbitals will overlap well to give good transmission. The results for transmission are shown in figure 4.3, along with the DOS of the C-chain uncoupled to the Cu contacts. The DOS is calculated with an imaginary part of 0.005 au added to the energy to broaden the discrete states. It can be seen that the transmission is very peaky, with only a few states contributing to the conductance between the end atoms. The transmission states are calculated with zero imaginary part, and are not broadened

at all except by the interaction with the Cu reservoirs. The maximum possible total transmission in this case is 1, as there is only one channel at input and output, and it can be seen from figure 4.3 that the transmission through the carbon chain at the peak energies is very close to 1. This is as we would expect, since we have identical p_x orbitals aligned along the chain, providing a good pathway for conduction, with no reflection within the chain.

It can be seen from figure 4.3 that there are fewer transmission peaks than DOS peaks – this happens since transmission is only appreciable at energies that correspond to those wavefunctions extending from one end of the chain to the other, with appreciable weight at each end.

When we compare the transmission peaks with the DOS peaks in figure 4.3, we see that they do not align. However, in figure 4.4 the same transmission peaks are shown to line up exactly with the total density of states when we include the effects of the Cu reservoirs. This energy shift of the states is a result of the coupling of the Cu contacts with the C-chain, via equation (4.25). Different states are shifted by varying amounts, and it can be seen that the shift is largest in the middle of the conduction band, near 0.5 au. Also, we note that the more a state is shifted, the wider the transmission peak, due to the interaction with the Cu contact. This example shows that the effects of the metal contacts must be taken into account when describing the relationship between the DOS and the transmission of relatively short molecules. On the other hand, while investigating the DOS of DNA we find that states are not noticeably shifted when coupled to the Cu contacts. This is due to the length of the DNA compared with the C-chain – the states are much more extended, and the contact provides a relatively smaller perturbation.

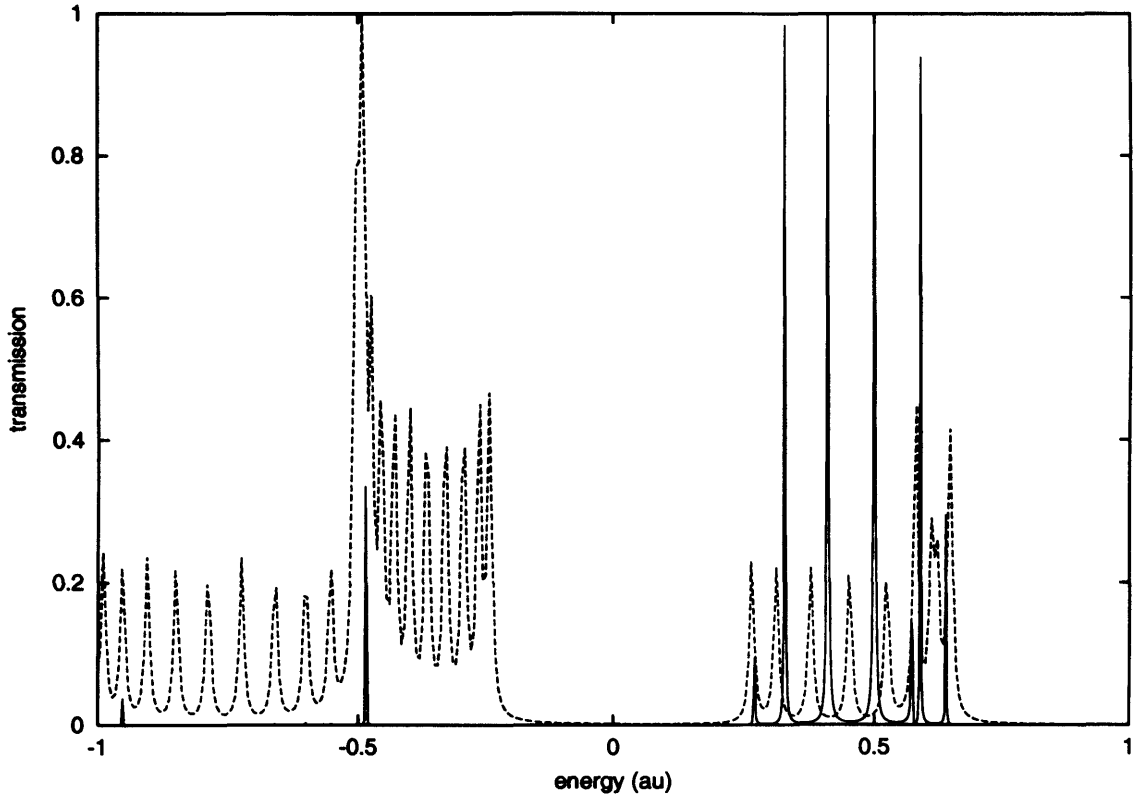


Figure 4.3: End-to-end transmission (solid line), and scaled DOS (dashed line) of a chain of 12 carbon atoms.

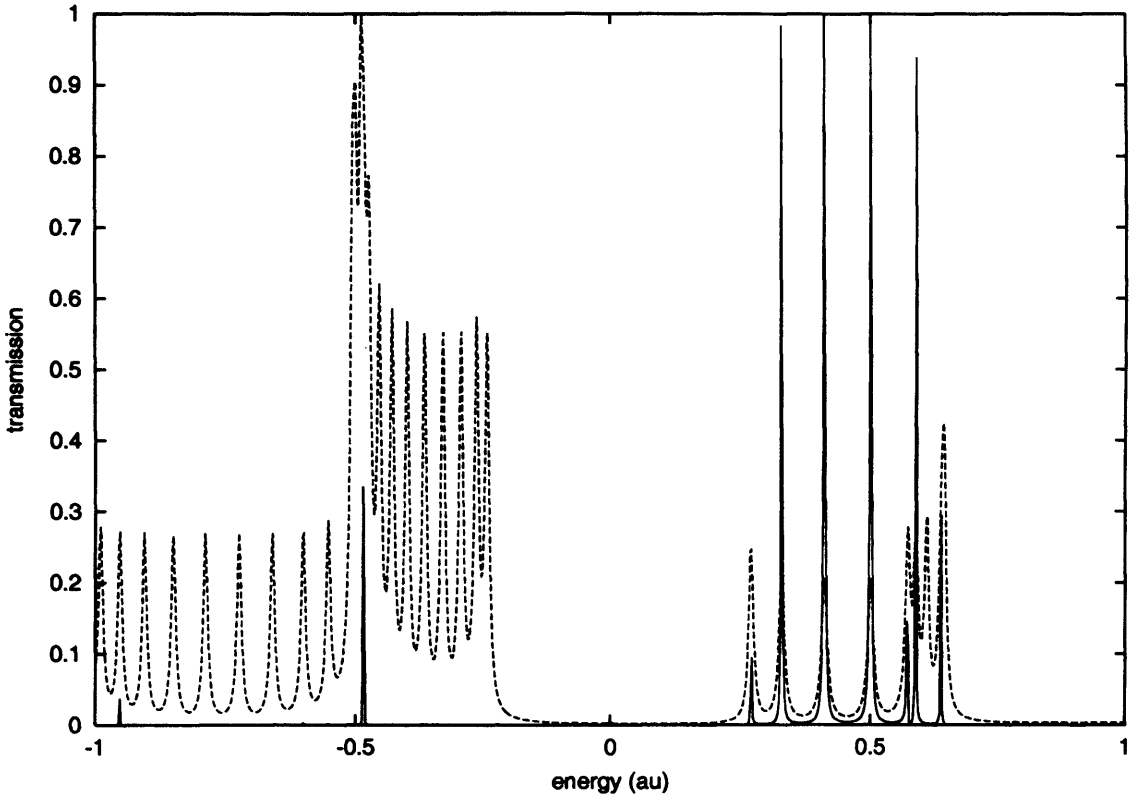


Figure 4.4: End-to-end transmission (solid line), and scaled total DOS (dashed line) of a chain of 12 carbon atoms attached to Cu reservoir.

Chapter 5

Results for DNA

The DNA molecules we investigate are 12 base-pairs long, these dodecamers consist of between 760 and 780 atoms. This includes all of the base and backbone atoms. These fairly large molecules have been studied previously using first-principles methods, such as density functional theory [113, 133], and quantum chemistry techniques [134]. However, for simplicity we use a tight-binding formalism to represent the electronic wave-functions. As explained in chapter 2.3 this is a more approximate, but quicker method of solving the Schrödinger equation. This will allow us greater flexibility for testing our embedding method, and readily exploring the effect of structure and molecule-metal contact on transmission and conductance.

Initial calculations were performed on three DNA molecules. The first is a 12 base-pair *B*-DNA molecule, (CGTAGATCTACG). The spatial co-ordinates of the atoms were obtained from single crystal x-ray diffraction experiments, performed at 15° C with 2.25 Å resolution [135]. Since x-ray diffraction experiments do not detect hydrogen atoms, we used the computer program ViewerLite [136] to add H atoms to the dangling bonds of the structure.

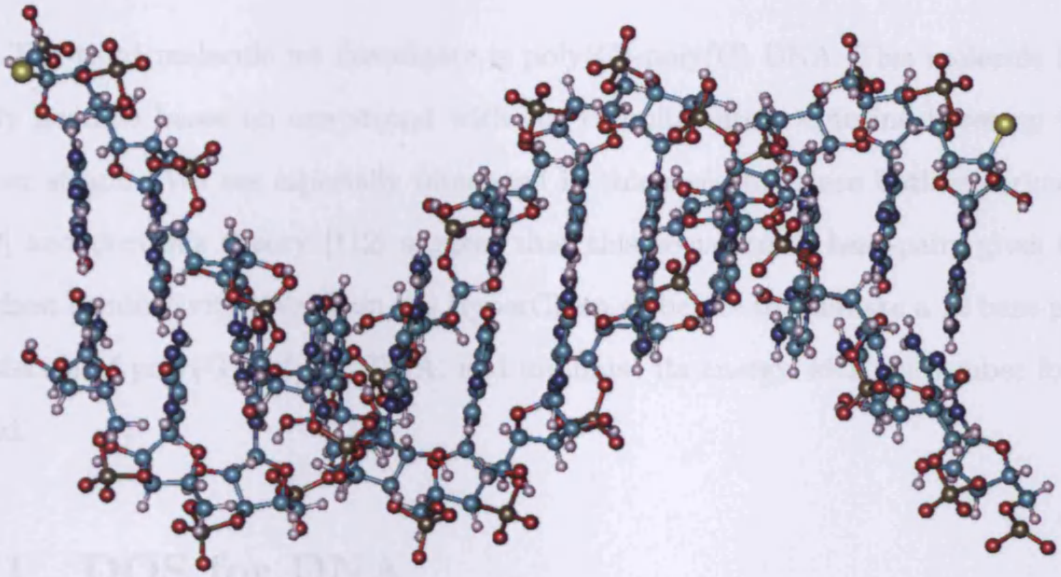


Figure 5.1: 12 base-pair poly(G)-poly(C) DNA molecule. The double ring guanine bases lie consecutively along the same backbone strand, with the complementary single ring cytosine bases along the opposite strand.

To investigate the effects of a more ordered structure of DNA, the second molecule we choose to study has the same 12 base-pair sequence as the first molecule. However, instead of using the x-ray diffraction structure, the new molecule is constructed using the molecular dynamics computer package HyperChem [137], in which the AMBER forcefield is used to produce the structure with a minimum energy. The AMBER (Assisted Model Building and Energy Refinement) forcefield is widely used for protein and DNA molecular dynamics simulations [138–140]. Originally developed by Peter Kollman in the University of California San Francisco, it uses a classical forcefield with parameter sets for proteins, nucleic acids and organic molecules to obtain energy minimised structures. In the minimisation, the Coulomb interactions between the atoms were modelled to fall off as $1/r^2$ rather than $1/r$, to simulate the screening effect of a solvent.

The third molecule we investigate is poly(G)-poly(C) DNA. This molecule has only guanine bases on one strand with the complimentary cytosine bases on the other strand. We are especially interested in this molecule since both experiment [22] and previous theory [112] suggest that this sequence of base-pairs gives the highest conductivity. We again use HyperChem as before to simulate a 12 base-pair molecule of poly(G)-poly(C) DNA, and minimise its energy with the Amber force field.

5.1 DOS for DNA

We now apply our embedding method, as described in section 4.1, to calculate the DOS of the three different structures of DNA described above. In our calculations the overlap between orbitals is cut off beyond 8 au, and there is only significant overlap between neighbouring sections, due to the localisation of the Slater-type orbitals. A small imaginary part of 0.005 au is added to the energy, broadening the δ-functions that represent the discrete electronic states of the molecule (equation (4.13)). The results for DNA taken from the mixed base structure are shown in figure 5.2; the DOS of the energy-minimised structure of this molecule is very similar, and therefore not given. The DOS of poly(G)-poly(C) DNA is plotted in figure 5.3.

One important property that can be determined from a DOS plot is the position of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). These values tell us where the valence band ends and the conduction band begins. We find the position of the HOMO by integrating the graph in figure 5.2 with respect to energy, using the trapezium rule. Allowing for spin we fill up the states, starting with the lowest energy, with the known number of

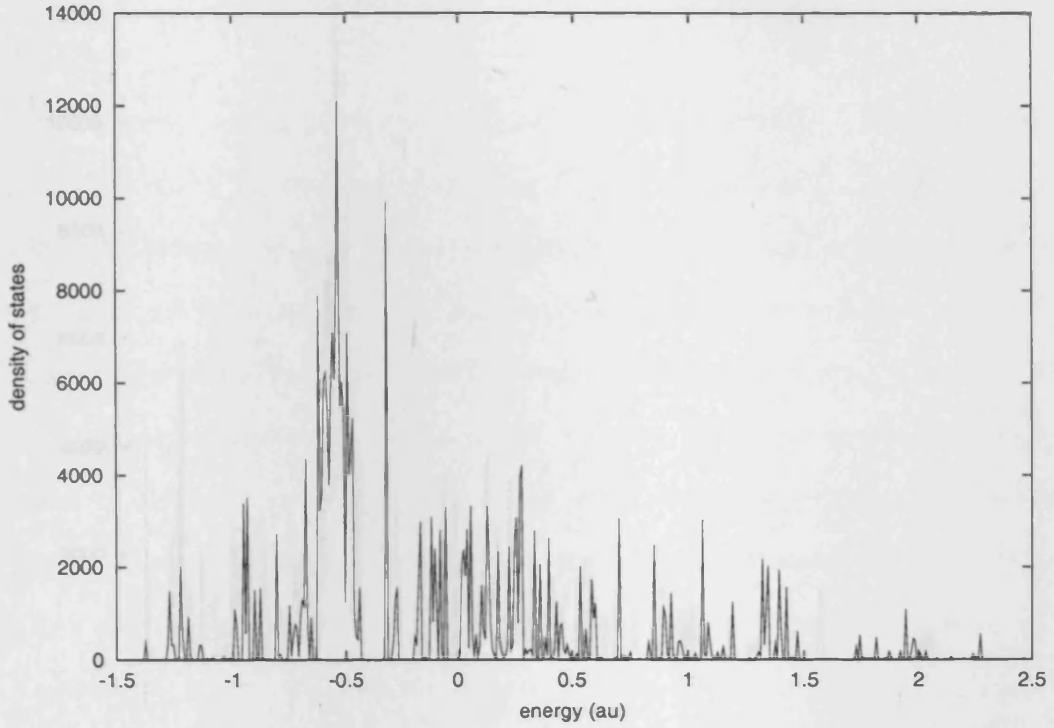


Figure 5.2: DOS of mixed base DNA dodecamer obtained from x-ray diffraction structure. HOMO is at -0.434 au and LUMO at -0.327 au.

electrons in our system. The highest occupied state is then labelled HOMO and the neighbouring unoccupied state is labelled LUMO. A full discussion of the position of the Fermi energy is given later in section 5.3.

For the mixed base DNA, we calculate the energy of the HOMO to be -0.434 au, and the LUMO to be at -0.327 au, giving a bandgap of 0.107 au. When considering the transport properties of the molecule, the most important states are those on either side of the bandgap, as these will dominate conduction through the molecule in the limit of small applied voltages. When we investigate the charge density of the mixed base DNA molecule, we find that states near the HOMO and LUMO are all located on atoms in the bases. This is in agreement with the generally

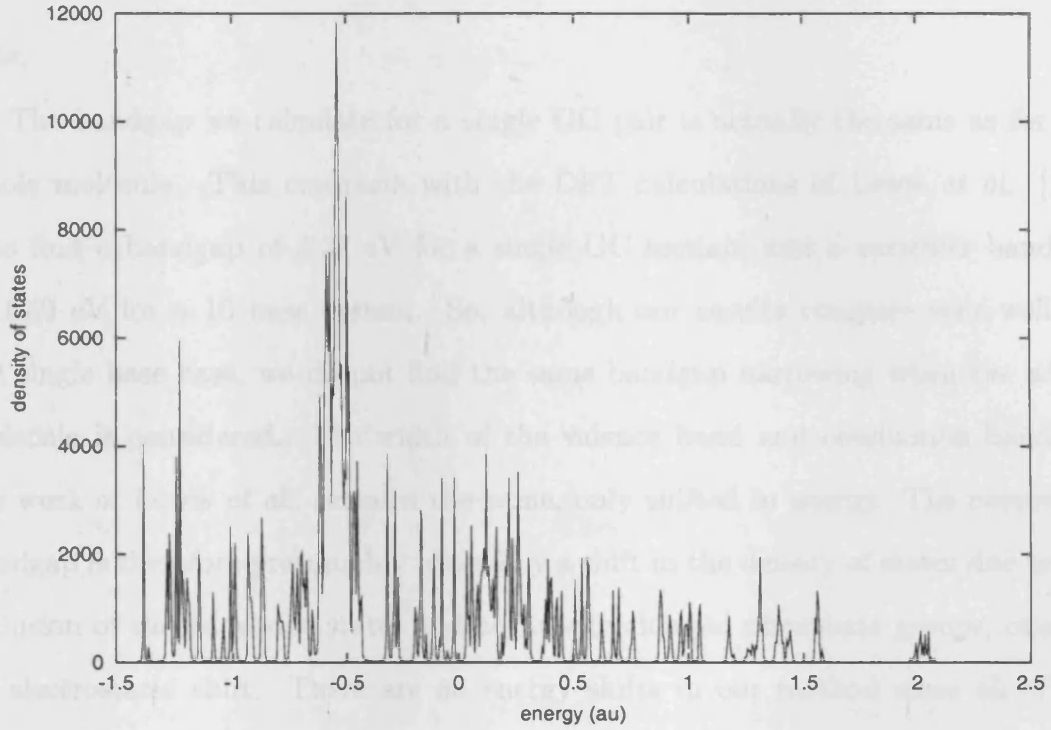


Figure 5.3: DOS for poly(G)-poly(C) DNA dodecamer. HOMO is at -0.435 au and LUMO at -0.327 au.

accepted theory that conduction through DNA occurs via the π orbitals in the bases [114].

Comparing the DOS for the poly(G)-poly(C) DNA molecule shown in figure 5.3 with the results for the mixed base DNA, figure 5.2, it can be seen that the results are surprisingly similar, with only minor changes in the fine detail. For both types of DNA the HOMO and LUMO values are -0.44 au and -0.33 au respectively, the bandgap remaining 0.11 au (3 eV).

Most literature values for the bandgap of poly(G)-poly(C) DNA vary between 1.12 eV and 3.2 eV for a variety of methods [111, 112, 115, 117]. We conclude therefore that our value of 3 eV for the bandgap agrees very well with published

data.

The bandgap we calculate for a single GC pair is actually the same as for the whole molecule. This contrasts with the DFT calculations of Lewis *et al.* [111] who find a bandgap of 3.37 eV for a single GC section, and a narrower bandgap of 1.40 eV for a 10 base system. So, although our results compare very well for the single base case, we do not find the same bandgap narrowing when the whole molecule is considered. The width of the valence band and conduction bands in the work of Lewis *et al.* remains the same, only shifted in energy. The narrowing bandgap is therefore presumably caused by a shift in the density of states due to the inclusion of the backbone states and negatively charged phosphate groups, causing an electrostatic shift. There are no energy shifts in our method since all of the atomic potentials are preset, and the calculation is not self-consistent. Lewis *et al.* also report a valence band width of 1.1 au in good agreement with our result of 0.9 au. Their conduction band width is smaller by a factor of two, an effect in extended Hückel theory that we have also encountered when dealing with small molecules. Apart from this discrepancy, the overall DOS in the extended Hückel scheme is in relatively good agreement.

5.2 Transmission

As with the C-chain in section 4.3, we are interested in the end-to-end transmission. This will tell us about the ability of the DNA molecule to transport charge from one end to the other. At this stage we use a single Cu orbital to contact a base orbital on one end of the DNA molecule, and then calculate the transmission through the molecule to another Cu orbital contacting a base orbital at the other

end of the DNA molecule. There is no covalent bonding between the molecule and the metal reservoirs. We vary the energy over the energy range of the DOS, and calculate the corresponding transmission. Direct contact to the bases is expected to give greater transmission since the electrons do not have to tunnel through the backbone states to reach the electron pathway of the bases.

5.2.1 Mixed-base DNA

The results for the transmission of the mixed-base DNA molecule obtained from x-ray diffraction are shown in figure 5.4. It can immediately be seen that the transmission is very peaky. This is due to the finite size of the system we are considering which gives rise to individual states, not bands. It has also been reported by Pendry *et al.* that disorder in one-dimensional systems, such as DNA, leads to extremely peaky transmission as a function of energy [141], a finding which is consistent with the results of figure 5.4. The width of the individual transmission peaks varies, but is generally very narrow, of the order of 10^{-6} au – the peak width comes from the interaction of the DNA with the continuum of states in the electrodes. However, since the states are much more extended in DNA compared with the carbon chain (figure 4.3), a much narrower peak width is observed. A recent extended Hückel study of $(AT)_{12}$ DNA by Tada *et al.* [77] gives very small transmission, again peaky, but with broader peaks than we find. This is a consequence, we believe, of their multi-orbital contacts between the metal and the DNA molecule.

From figure 5.4 we can see that only a few states give appreciable transmission, and the maximum transmission is 0.44. The rest of the energy range gives transmission too low to register on figure 5.4. The large transmission peaks around

−0.5 au in figure 5.4 lie in the large peak in the DOS below the HOMO. There is a small peak near the LUMO at −0.32 au and another slightly larger one high in the conduction band at 1.33 au. Experimental work and DFT calculations performed by de Pablo *et al.* [45] show poor conductivity for DNA with a random base-pair sequence, consistent with this work.

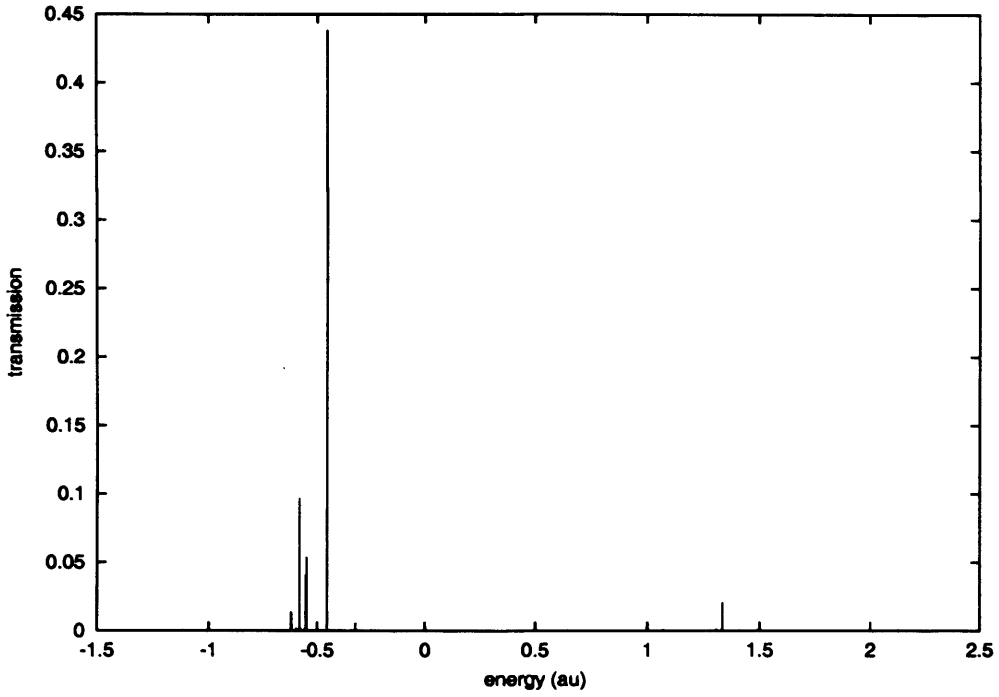


Figure 5.4: End-to-end transmission of mixed base DNA, for varying energy.

While investigating the transmission through the DNA, we find that the values of transmission, and the energies at which it is significant, greatly depend on which orbitals are attached to the Cu contacts – this effect is also seen both experimentally by Kushmerick *et al.* [76], and theoretically by Damle *et al.* [130]. In order to investigate the contact orbital dependence, we fix the energy at that of a particular state, and calculate the transmission between all 32,000 or so combinations of orbitals in the two end sections of the molecule, including the backbone.

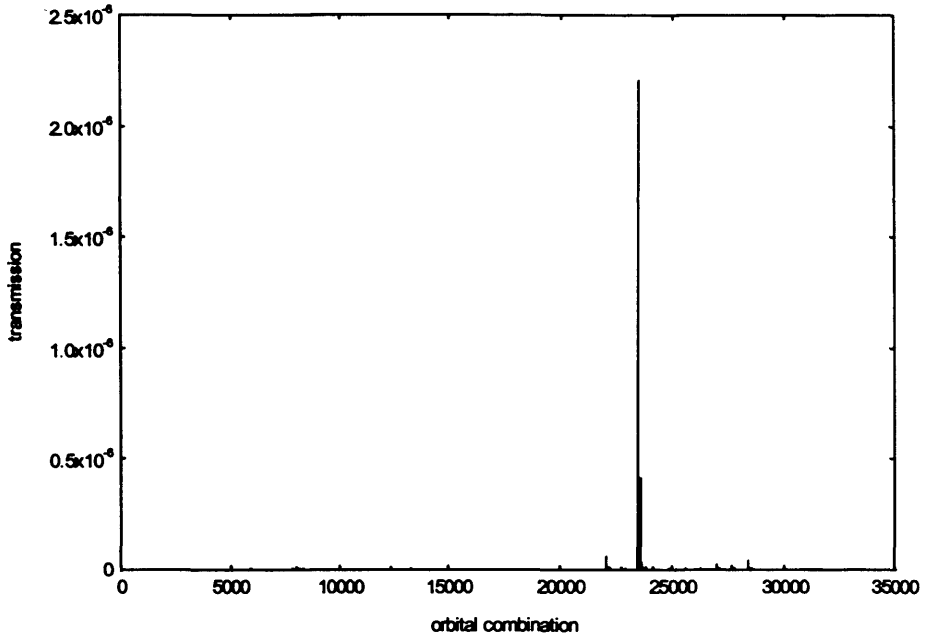


Figure 5.5: End-to-end transmission of mixed base DNA for all combinations of orbitals in sections 1 and 12, for HOMO.

We apply this method to states near the bandgap, which are the most important for conduction [114], since they provide the transmission channels at small voltages. In figure 5.5 the transmission of the HOMO state is investigated, and we see that the transmission is very small for all orbital combinations, with a maximum value of 2.2×10^{-6}. The other states near the HOMO all give extremely low transmissions, of the order of 10^{-14}. However, we find a series of relatively large transmission peaks for a state just above the LUMO in the conduction band at energy -0.320, figure 5.6. We see that for this state there is a very large number of combinations of contact orbitals giving an appreciable transmission, with a maximum value of 0.04, which is reasonably large for transmission in this molecule. This implies that the corresponding wavefunction is well distributed over these orbitals. This contrasts with

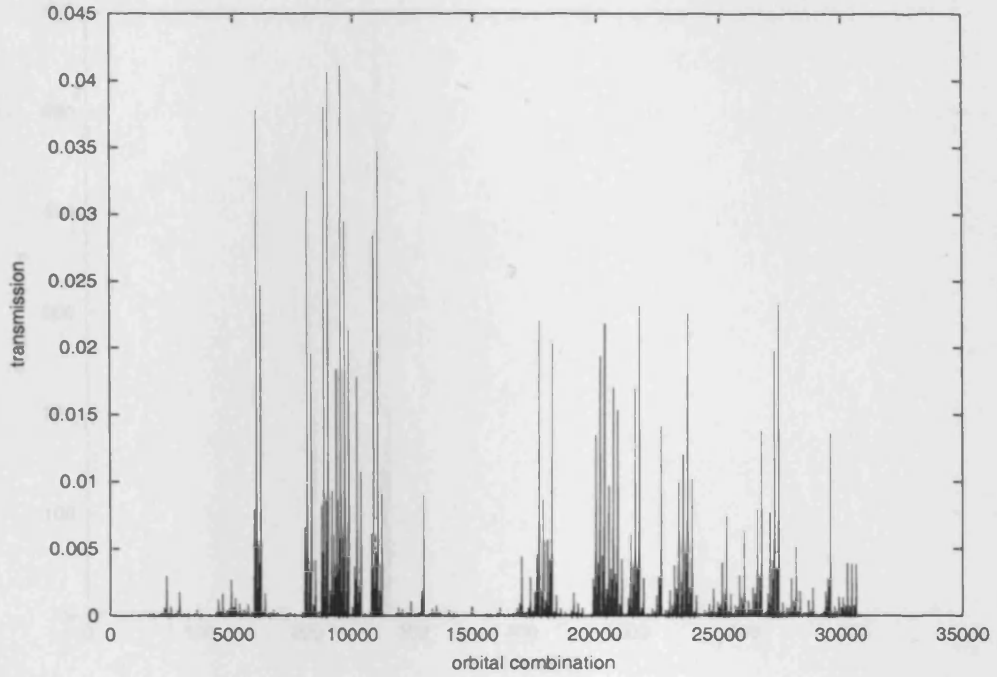


Figure 5.6: End-to-end transmission of mixed base DNA for all combinations of orbitals in sections 1 and 12, for conduction band state at -0.320 au.

the transmission shown in figure 5.5 for the HOMO state, where the transmission is very poor. The maximum transmission for this state is very small $\sim 10^{-6}$, and only a few orbital combinations have transmission large enough to register on the graph. Analysing the results from figure 5.6 we find that all of the major peaks correspond to orbital combinations between base orbitals at either end of the molecule. This supports the idea that DNA conducts charge via the base stack. To investigate this further we look at the distribution of charge density along the molecule. There are 760 atoms in the mixed-base DNA molecule, which are labelled 1 to 380 from base-pairs 1 to 12 along a single strand of DNA, then 381 to 760 from base-pairs 12 to 1 along the complementary single DNA strand. Figures 5.7 and 5.8 show the distribution of charge along the molecule for the HOMO and the conduction band

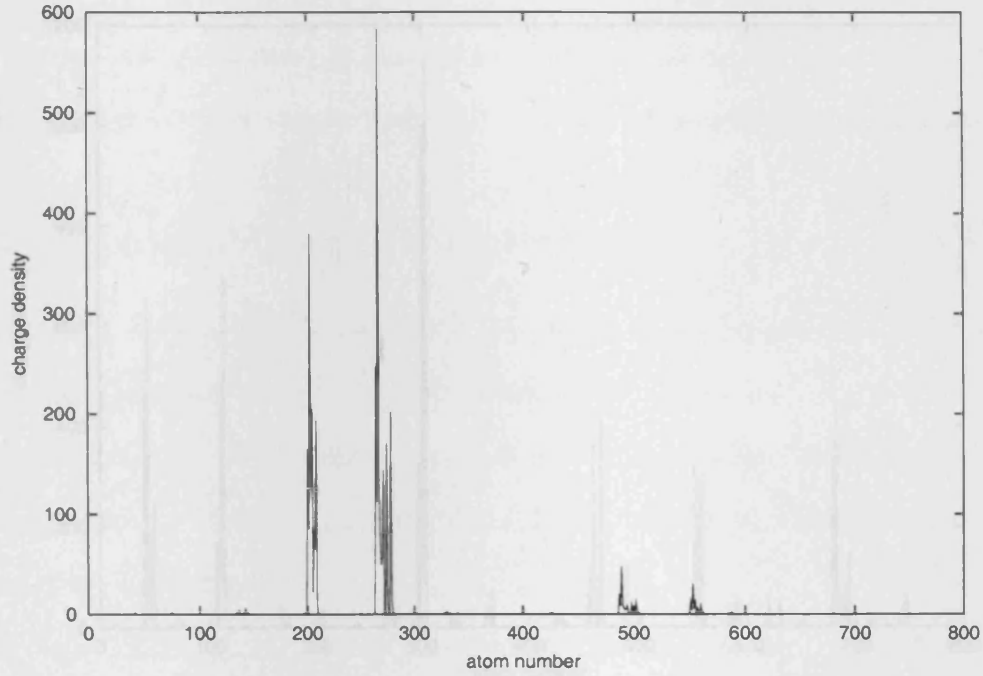


Figure 5.7: Distribution of charge density along mixed base DNA molecule, for HOMO, the atom number describes the position along the chain.

state at -0.320 au respectively. Comparing the two graphs it can be seen that for the conduction band state the charge is distributed more evenly along the molecule, allowing for electron transfer along the chain. However, for the HOMO state there are large gaps between regions of high charge density, corresponding to a more localised state, hence reducing the probability of electron transport from one end of the molecule to the other. This is clearly why there is such an enormous difference in the transmission of the two states (figures 5.5 and 5.6). A discussion of the charge density on contact atoms and their role in transmssion is given in section 5.4.

Using figure 5.8 we can determine the atoms, and hence the orbitals that have a large charge density for the transmission state at -0.320 au. As we expect most of the peaks are on p_z-orbitals in the bases, which lie along the axis of the molecule.

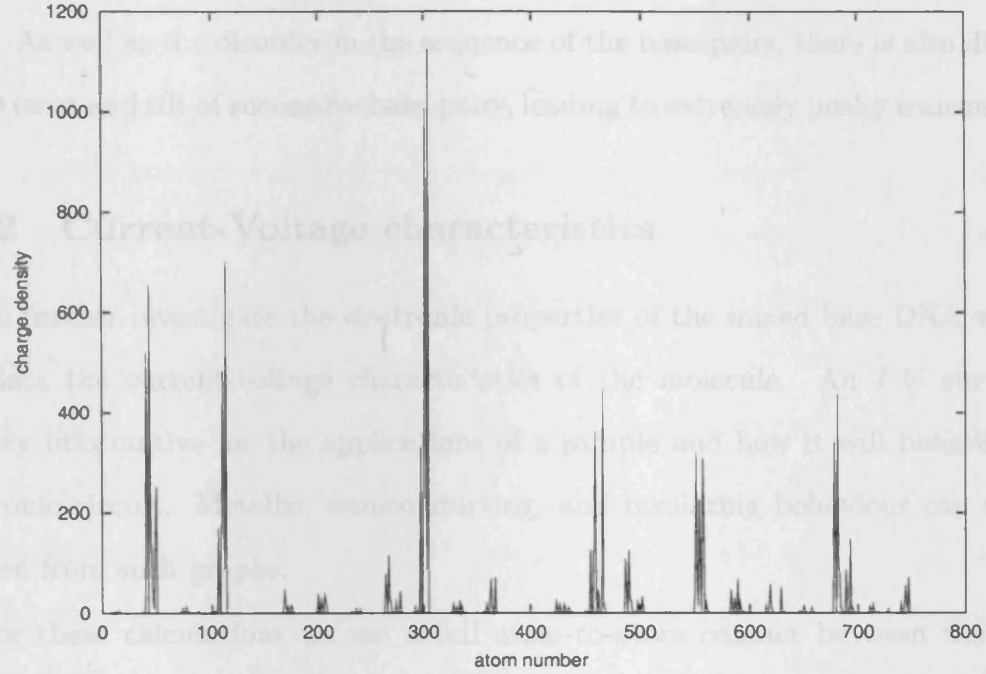


Figure 5.8: Distribution of charge density along mixed base DNA molecule, for conduction band state at -0.320 au, the atom number describes the position along the chain.

This is consistant with transport along the base stack. However, there appears to be no preference for the charge to lie on a specific base, with a similar number of charge density peaks on each of the four bases. This behaviour is repeated for many of the transmissions peaks in the mixed-base DNA molecule. We would expect more charge to lie on the G bases, since they have the lowest ionisation energy. This suggests that for this molecule we are observing hopping between all base types as seen by the Barton group [39].

The low transmission in this DNA molecule for nearly all energies is a result of a combination of the composition of the DNA molecule and the degree of disorder. The molecule consists of a mixture of AT and GC base-pairs. The AT base-pairs present a potential barrier to charge carriers on a GC base-pair, thus reducing conduction

[142]. As well as the disorder in the sequence of the base-pairs, there is also disorder in the twist and tilt of successive base-pairs, leading to extremely peaky transmission.

5.2.2 Current-Voltage characteristics

To further investigate the electronic properties of the mixed base DNA we now calculate the current-voltage characteristics of the molecule. An I-V curve can be very informative for the applications of a sample and how it will behave in an electronic circuit. Metallic, semiconducting, and insulating behaviour can readily be seen from such graphs.

For these calculations we use a full atom-to-atom contact between the metal reservoir and the molecule. We now use gold as the metal contact, and project the surface density of states of the s, p and d electrons onto the contact orbitals of a single Au atom. For a more realistic experimental representation we substitute a H atom from a sugar on each end of the molecule and replace it with a sulphur atom, as the divalent S atom covalently bonds the DNA molecule to the Au leads. Sulphur atoms are often used to improve the bonding between a molecule and metal surface, since sulphur chemisorbs well to metal surfaces [47, 143].

We use equation (3.14) to calculate the I-V characteristics over a voltage range of 0-10 V, typical of the voltages expected of an electrical component, and those used in experiments. The transmission in this range was calculated very precisely, and the integration performed numerically. The Fermi energy of the molecule was placed half way between the HOMO and LUMO, in the middle of the large bandgap above -0.43 au. The integration was calculated taking the middle of the bandgap as the starting point and then extending the limits equally in energy into the valence

band and conduction band. When calculating the current at a voltage, V, we should shift the Fermi energy of the Au reservoirs by $\pm eV/2$ either side of the molecular Fermi energy (see figure 4.2). However, since the calculation of the current involves an integral over energy, we would need to repeat the transmission calculations for each energy point, shifting the Fermi energy each time. We therefore fix the Fermi energy of the left and right metal contact at +2.0 V and −2.0 V respectively, which is in the middle of the typical voltage range. This assumption reduces the computing time by a factor of 10^5.

The results for the mixed base DNA are shown in figure 5.9. It can be seen that for very small voltages there is no current. Then at just above 3 V the current rises in a series of steps. This threshold voltage, up to which there is no current flow, is typical of semiconductors. Each step represents the inclusion of extra transmission

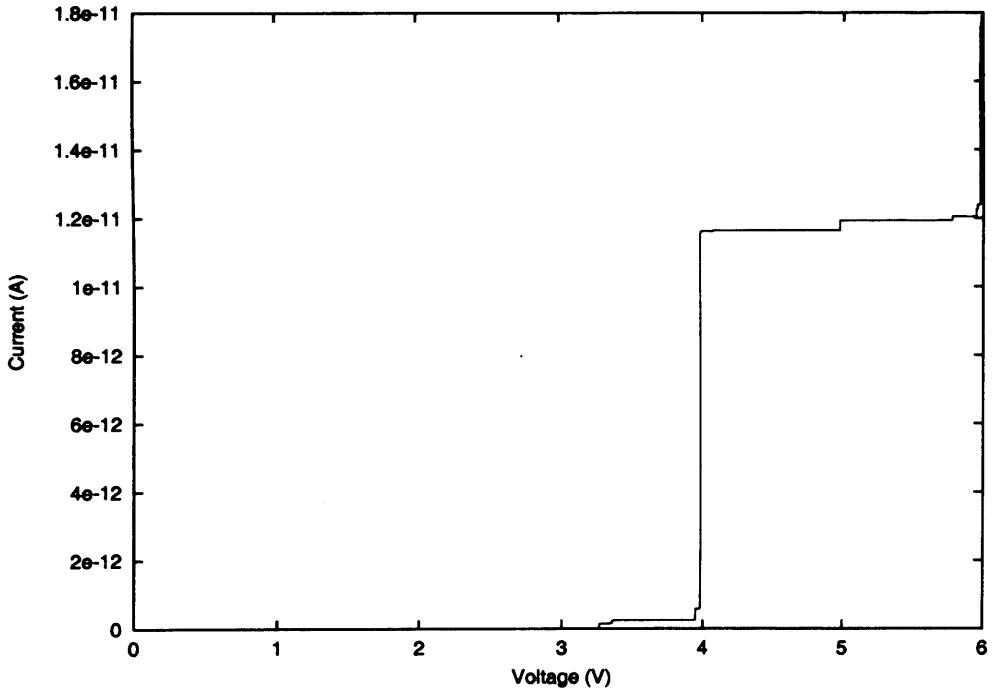


Figure 5.9: Current-Voltage characteristics for mixed base DNA.

states as the voltage increases. The current then continues to increase in a step-like fashion as the voltage is increased further. This step-like behaviour is due to the fact that the transmission is so peaky. Typical current values of $\sim 10^{-12}$ A correspond to quite poor conduction. These results agree well with the extended Hückel calculations of Tada *et al.* [77], who find currents of a few picoamps for their $(AT)_{12}$ molecule, the random arrangement of AT base-pairs being similar to our mixed base case. Tada *et al.* also observe step-like behaviour, and find a threshold voltage of 4 V, very similar to our own. As explained previously, the poor conduction of the mixed base DNA is due to the potential barrier posed by the AT base-pairs, and the structural disorder of the system.

5.2.3 Energy-minimised mixed-base DNA

We now consider the results from our energy-minimised molecule of DNA. This molecule has the same mixed-base sequence as the one previously considered. However, the conformation has been manipulated using HyperChem, to yield the minimum potential energy. We consider this model, in order to investigate the effect of changes in structure on transmission.

To find the transmission of this energy-minimised molecule, we once again couple a single Cu s-orbital to a single base orbital, and calculate the transmission over the energy range of the density of states. The results are shown in figure 5.10. We can again see that the transmission is very peaky. However, compared with figure 5.4 we can see that there are many more transmission peaks for the energy-minimised molecule, and that the maximum transmission is now close to 1. There is also good transmission for some of the states below the HOMO at ~ -0.5 au. It is these states

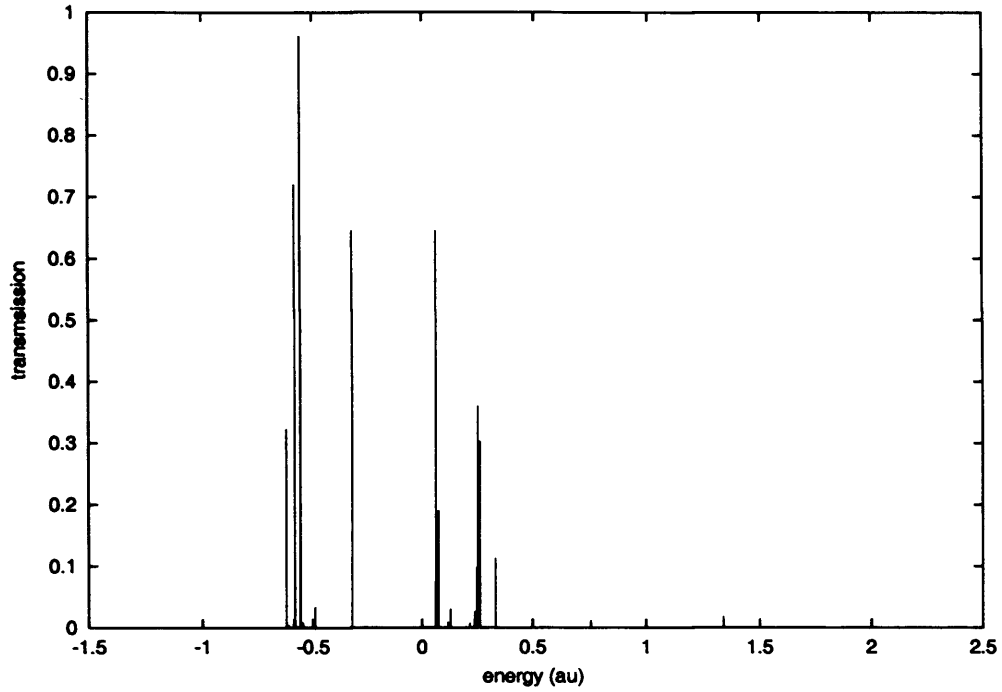


Figure 5.10: End-to-end transmission of energy-minimised mixed-base DNA for varying energy.

near the HOMO and LUMO that are the most important for conduction.

There are several more states with large transmission in the conduction band, just above the LUMO. Figure 5.11 shows the transmission of one such state at -0.320 au, as a function of orbital combination, for which the maximum transmission is 0.64 – this is comparable to that of the carbon chain (figure 4.4). The transmission still has huge fluctuations when varying the contact orbitals. These great changes in transmission in going from the x-ray diffraction structure to the energy-minimised structure show how small changes in structure can dramatically affect transmission. These initial results suggest that the minimised structure will conduct electric current more easily than the structure obtained by x-ray diffraction.

The current-voltage characteristics of the energy-minimised molecule are shown

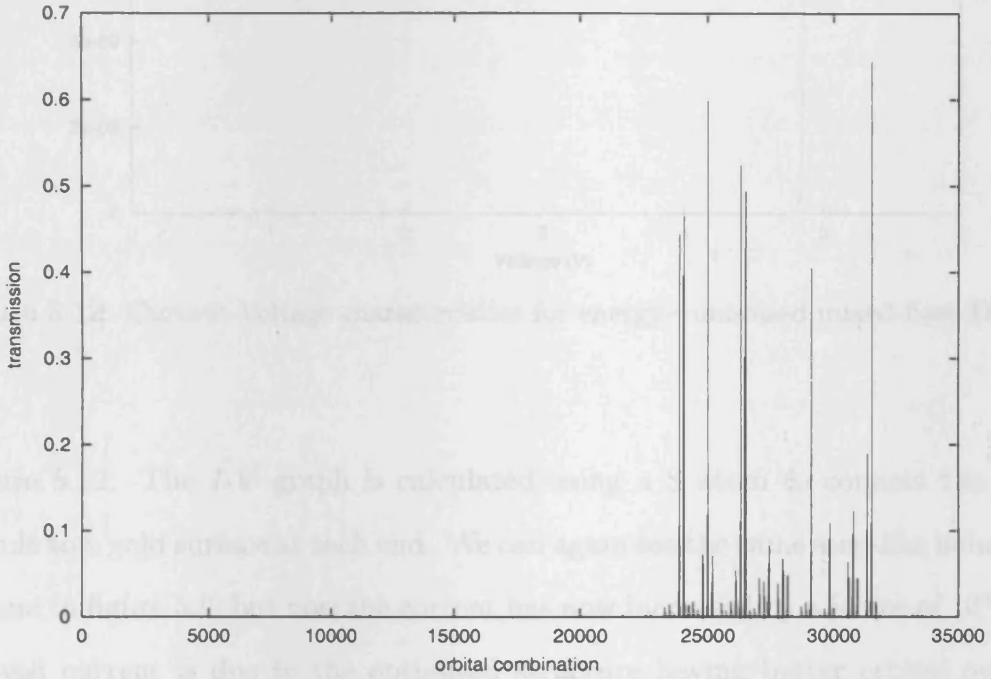


Figure 5.11: End-to-end transmission of energy-minimised mixed-base DNA for all combinations of orbitals in sections 1 and 12, for a conduction band state at -0.320 au.

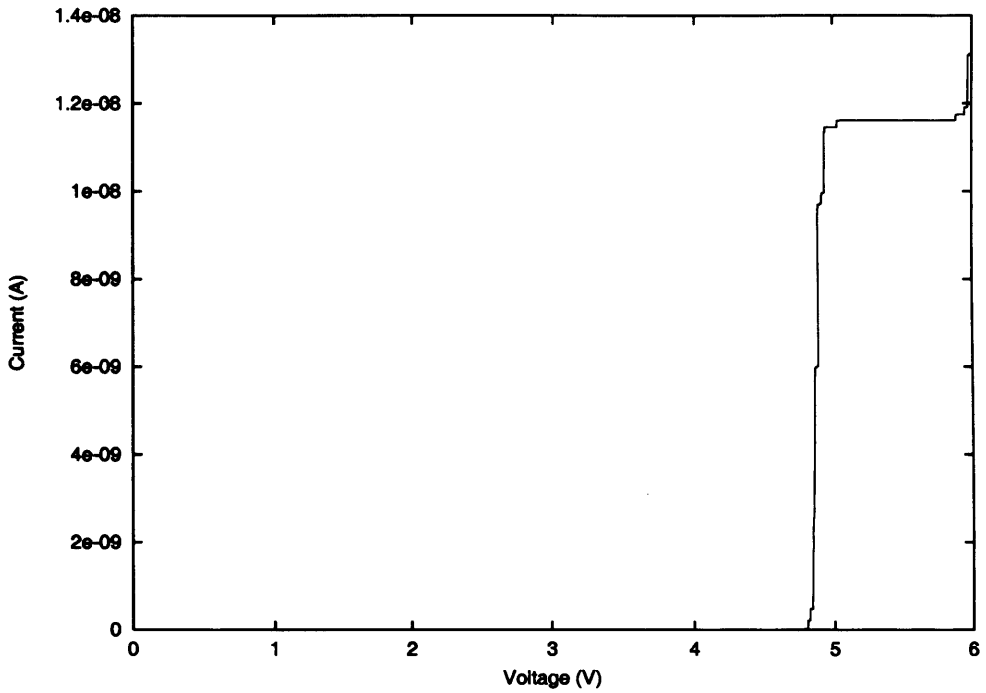


Figure 5.12: Current-Voltage characteristics for energy-minimised mixed-base DNA.

in figure 5.12. The *I-V* graph is calculated using a S atom to connect the DNA molecule to a gold surface at each end. We can again see the same step-like behaviour we found in figure 5.9, but now the current has now increased by a factor of 10^3. The improved current is due to the optimised structure having better orbital overlap, and being less disordered than the structure obtained from the x-ray diffraction experiments. Comparing figures 5.12 and 5.9 it appears that the threshold voltage for the energy-minimised structure has increased to 4.75 V compared with 3.25 V for the x-ray structure. However, when we examine the *I-V* graph for the energy-minimised structure more closely we find that there is a small current of $\sim 10^{-12}$ A, which begins to increase in a series of steps above 3.25 V. This behaviour is however, too small to register on figure 5.12. We conclude therefore that the current is sensitive to both changes in contact orbitals and changes in structure.

5.2.4 Energy-minimised poly(G)-poly(C) DNA

The third molecule we investigate is poly(G)-poly(C) DNA. We have minimised its energy using HyperChem. This ordered system, consisting of the same repeated bases, is thought to have greatly improved transmission over mixed base-pair DNA [45].

First we investigate the end-to-end transmission through the molecule by fixing a single Cu orbital to a single s-orbital on the DNA bases at either end of the molecule, and calculating the transmission as a function of energy. The results are given in the peaky graph shown in figure 5.13. It can be seen that there are a number of peaks with remarkably good transmission, many more than for the mixed-base DNA (figure 5.4) and the energy-minimised mixed-base DNA (figure 5.10). However, for

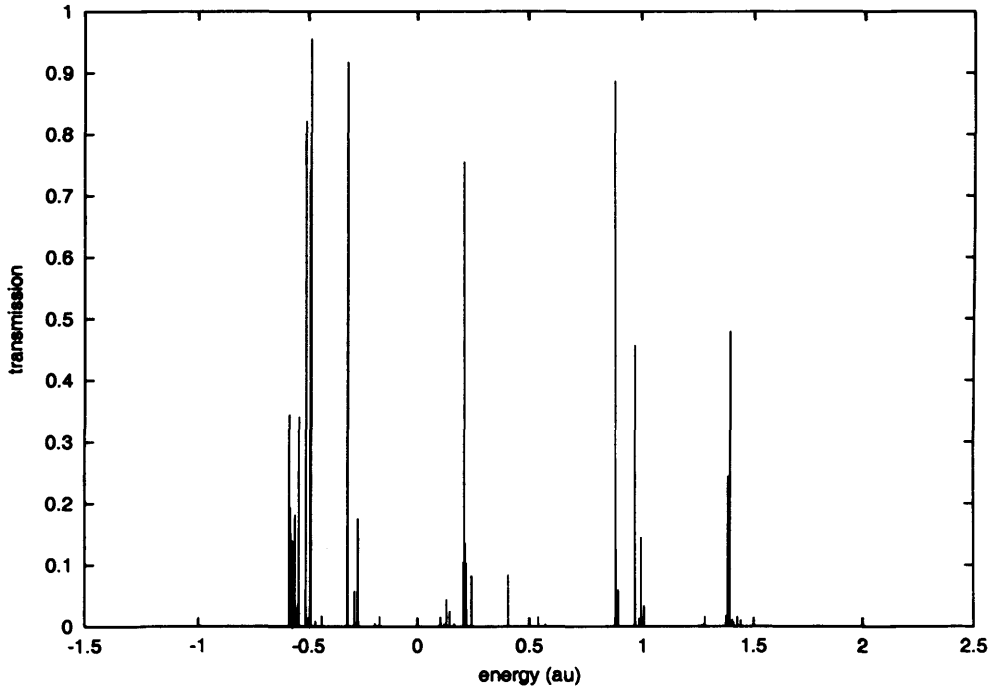


Figure 5.13: End-to-end transmission of poly(G)-poly(C) DNA, for varying energy.

nearly all energies the transmission still remains effectively zero, though there are groups of states around the HOMO (−0.43 au) and LUMO (−0.32 au) with large transmission between the metal contacts at the ends of the molecule.

The results in figure 5.13 can be compared with previous studies of transmission of poly(G)-poly(C) DNA, all of which show peaky transmission, though there are significant differences with our work. Adessi *et al.* [114], in their DFT studies of poly(G) DNA without a backbone, find discrete narrow blocks of complete transmission, presumably a result of the formation of narrow bands due to the infinite structure which they use. The large fluctuations in transmission as a function of energy result from the finite structure and disorder of our system – although the structure has been determined by energy minimisation it is not completely ordered, leading to the fluctuations, characteristic of one-dimensional disordered systems [141]. Similar large fluctuations are seen in the work of Roche [117] using a model Hamiltonian to show the effect of temperature on transmission through poly(G)-poly(C) DNA.

Once again we study the effect of contact orbitals on the transmission of our molecule. For several states near the HOMO, many orbital combinations give almost complete transmission. Figure 5.14 shows the transmission for a state just below the HOMO at −0.4368 au, with a maximum transmission of 0.96, compared with 10^{-6} for the peak transmission shown in figure 5.5 for the mixed-base DNA. However, we see for figure 5.14 that again, there is an extreme dependence on orbital combination, and the same holds for states near the LUMO.

It is clear from our results that poly(G)-poly(C) DNA has markedly different behaviour to the mixed-base DNA, and with its much higher transmission for states near both the HOMO and LUMO, it will be a much better electrical conductor. This is in agreement with previous theoretical work [112] and experiment [48].

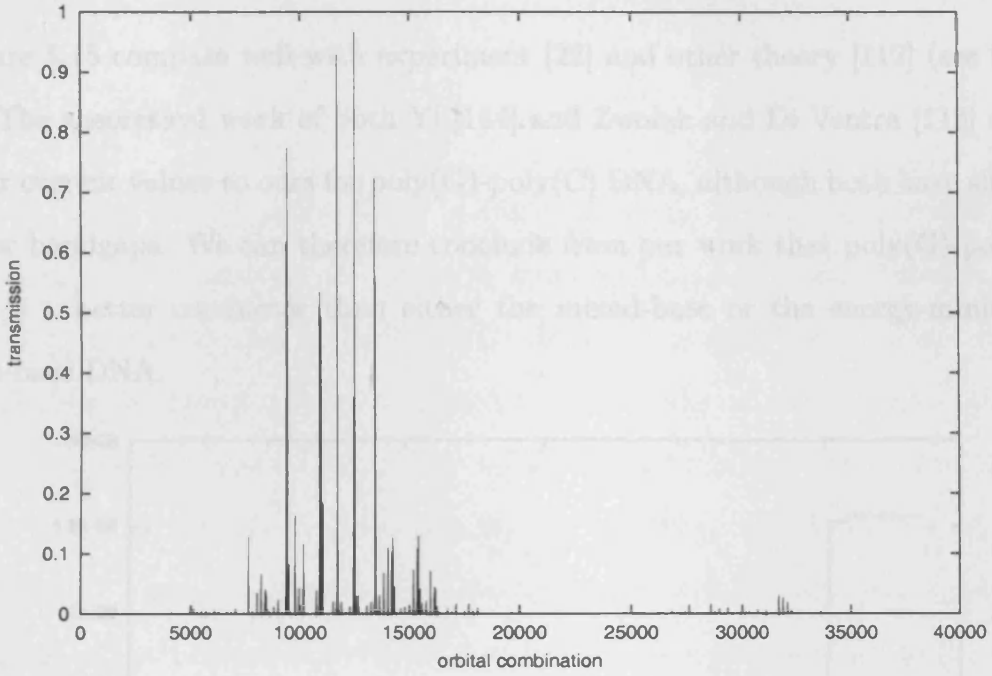


Figure 5.14: End-to-end transmission of poly(G)-poly(C) DNA for all combinations of orbitals in sections 1 and 12, for a valence band state at -0.4368 au.

We now calculate the current-voltage characteristics of poly(G)-poly(C) DNA, again using a S atom to connect the molecule and metal. We can see from figure 5.15 that this DNA molecule displays the same step-like behaviour as seen in both previous DNA molecules (figures 5.9 and 5.12). The threshold voltage of poly(G)-poly(C) DNA is just above 3 V, similar to that of the mixed-base DNA, but smaller than the energy-minimised mixed-base DNA, making poly(G)-poly(C) DNA a wide bandgap semiconductor. We can see that compared to the mixed-base DNA the current is 10^3 times larger, with typical values of 10^{-9} A, similar to the energy-minimised mixed-base DNA. Moreover, the small threshold voltage of the poly(G)-poly(C) DNA means that a much smaller voltage is required to produce an appreciable current, making poly(G)-poly(C) DNA a better conductor. The results

in figure 5.15 compare well with experiment [22] and other theory [112] (see figure 3.3). The theoretical work of both Yi [144] and Zwolak and Di Ventra [115] shows similar current values to ours for poly(G)-poly(C) DNA, although both have slightly smaller bandgaps. We can therefore conclude from our work that poly(G)-poly(C) DNA is a better conductor than either the mixed-base or the energy-minimised mixed-base DNA.

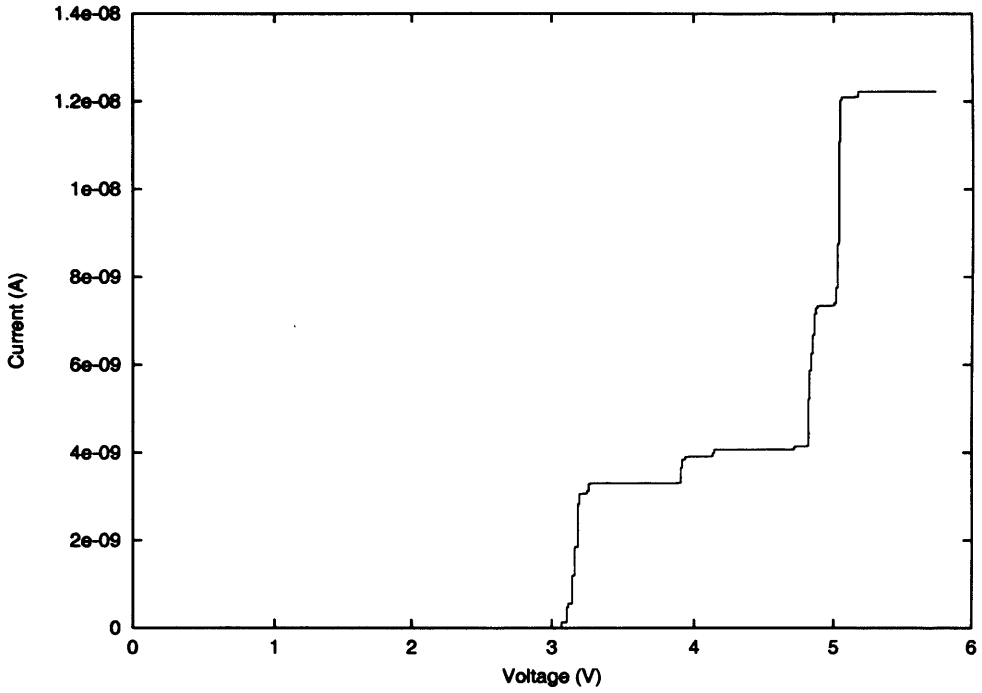


Figure 5.15: Current-Voltage characteristics for poly(G)-poly(C) DNA.

5.3 Position of Fermi energy

As we now show, the position of the Fermi energy plays a crucial role for determining the electrical properties of a molecule predicted by our theoretical model. We treat DNA as an intrinsic semiconductor, and as such the position of the Fermi

energy is given by [145],

$$E_F = \frac{E_g}{2} + \frac{3}{4}kT\ln\frac{m_h^*}{m_e^*}, \tag{5.1}$$

where E_g is the energy width of the bandgap, T is the absolute temperature and m_h^* and m_e^* are respectively the effective masses of the holes and electrons, with the zero of energy located at the HOMO. Since we assume $T = 0$ in our model, we find that the Fermi energy is located in the middle of the bandgap half-way between the HOMO and LUMO. Thus accurately determining the position of the HOMO is vitally important. To determine the energy of the HOMO we integrate the density of states, multiplying by a factor of 2 to account for spin degeneracy. This gives us the number of electrons that can be accommodated up to the upper limit of the integral. In the case of poly(G)-poly(C) DNA, our 12 base-pair molecule consists of the following atoms – 262 H, 228 C, 168 O, 96 N, 24 P, and 2 S. Therefore considering only the valence shell of each atom we have a total of 2794 electrons available for bonding. DNA however is an acid, and has one extra electron per phosphate group (figure 1.1). Therefore, we have an extra 24 electrons to include. For completeness we also need to include an extra electron for each sulphur atom to satisfy the dangling bonds at the contacts. This brings the total number of electrons up to 2820. Previously, the integration was performed using the trapezium rule, and we found that the HOMO lies at the bottom of the relatively large bandgap just above -0.5 au (see figure 5.3). This puts the Fermi energy in the middle of the gap, giving us the I-V graph shown in figure 5.9. However, to find the position of the HOMO exactly, we reduce the imaginary part of the GF to reach the resolution of individual states (figure 5.17). We find now however, that the density of states becomes so peaky that we are unable to get an accurate answer using the trapezium

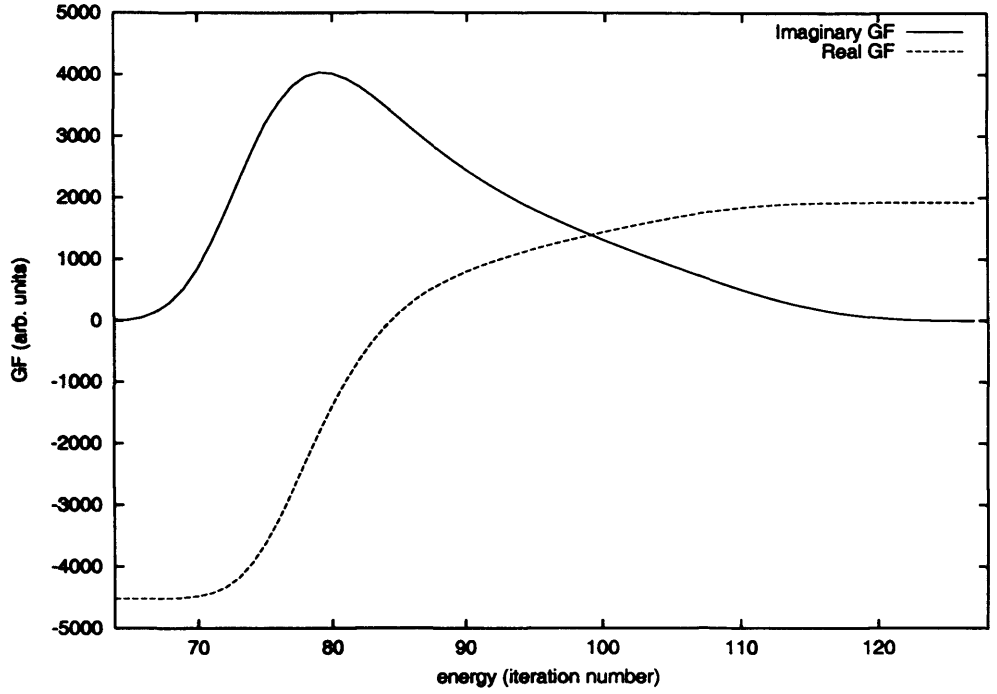


Figure 5.16: Real and imaginary parts of the Green function over the energy range of the valence band of poly(G)-poly(C) DNA, between $E = -1.5$ and -0.38 (see appendix A).

method of integration. Instead we use a highly accurate method based on contour integration of the analytic Green function in the upper-half complex plane (see appendix A). It can be seen from figure 5.16 that the density of states is smoothed out, with a large imaginary part for the GF in the middle of the range that tends to zero at both ends. This method allows for a very accurate calculation of the integral.

Using this method we find that the HOMO for poly(G)-poly(C) DNA in fact lies at -0.4356 au. Figure 5.17 shows a detailed density of states just below the large bandgap, with an imaginary energy of only 0.0001 au added, so that we can identify individual states. The accurate calculation puts the HOMO as the third state from the right in figure 5.17. If we take this to be the HOMO then the Fermi

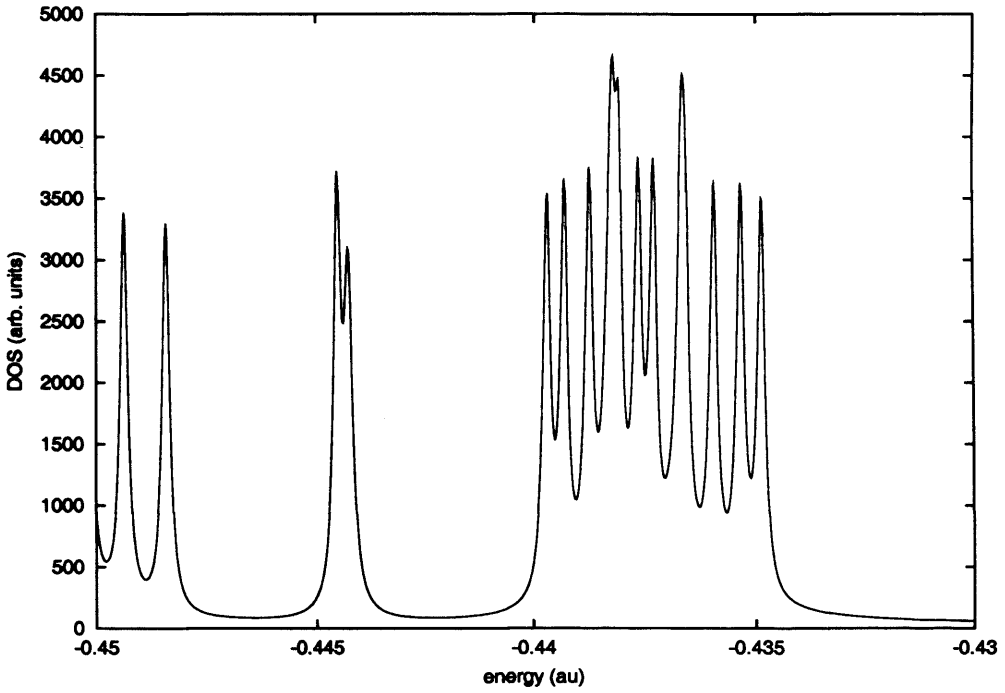


Figure 5.17: DOS for poly(G)-poly(C) DNA, for states just below the bandgap.

energy would lie between the second and third states from the right, and the *I-V* characteristics for poly(G)-poly(C) DNA would be completely changed (see figure 5.18).

Since the Fermi energy is now much closer to the conducting states, the size of the bandgap has been cut effectively to zero. The step-like behaviour and current values are similar to figure 5.15, but the characteristics of the molecule are changed completely. This gives *I-V* results which are now more characteristic of a finite metallic chain than a semiconductor. On physical grounds we would in fact expect the HOMO to be the state at the bottom of the bandgap, where we assumed it to be in the calculations of *I-V* in sections 5.2.2, 5.2.3 and 5.2.4. A bandgap above the HOMO stabilises the structure, and in any case the more accurate density functional calculations of Lewis *et al.* [111] place the HOMO at the bottom of this gap. The

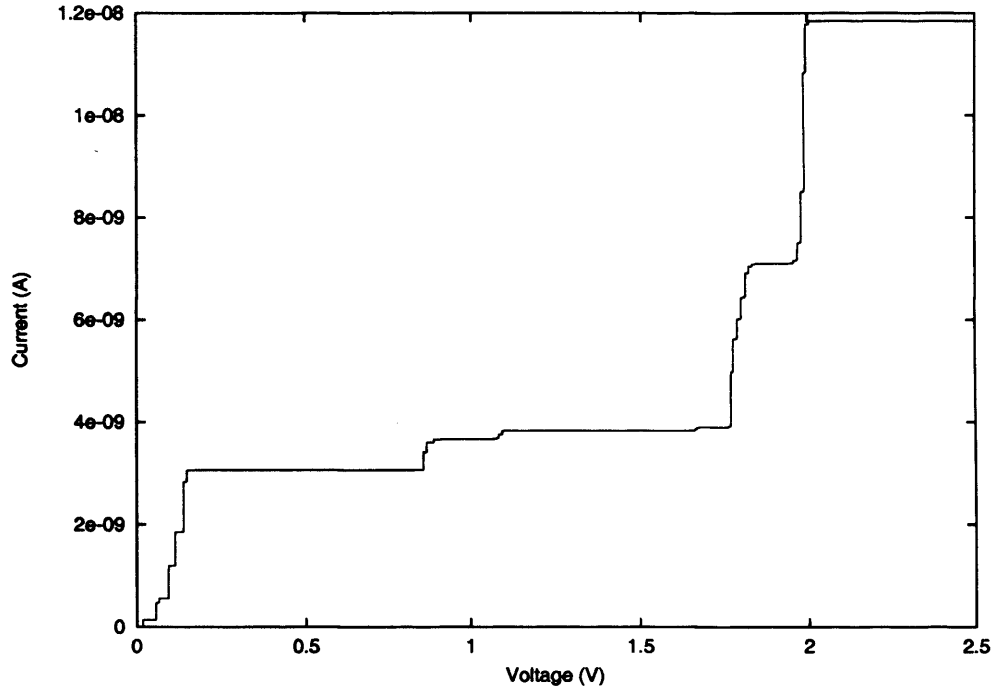


Figure 5.18: Current-Voltage characteristics for poly(G)-poly(C) DNA, with E_F=−0.4356, just below the bandgap.

discrepancy between the position of the HOMO and the bottom of the bandgap may be due simply to the inadequacies of extended Hückel. Extended Hückel is not a self-consistent method, and we believe that this may allow two states to drop from the conduction band into the valence band.

As a check on the accuracy of our method, we have integrated the DOS of poly(G)-poly(C) DNA over the whole energy range of the DOS, using the complex contour method of integrating described in appendix A. For the poly(G)-poly(C) molecule there should be a total of 2334 states in the conduction and valence bands, corresponding to the number of basis functions in our minimum basis set calculation. However, the numerical integration gives 2333.4 states, this means that we have lost 0.6 of a state. As the density of states calculation involves the numerical integration

of the overlap matrices, the error is presumably due to limiting the distance of the overlap calculations to 8 au and rounding errors. This margin of error corresponds to a percentage error of 0.03 %, with which we are well satisfied, and means that the HOMO calculations above will be accurate to within 0.8 of an electron, or less than half a state.

To further demonstrate the effect of shifting the Fermi energy, we now consider the I-V characteristics of the neutral molecule, that is, without the extra 24 electrons on the phosphate groups. The HOMO is now placed at -0.4444 au, which is between the two peaks located just above -0.445 au in figure 5.17.

The results are shown in figure 5.19. We can see that the bandgap is 0.25 V which is much smaller than our original calculation, but larger than the metallic

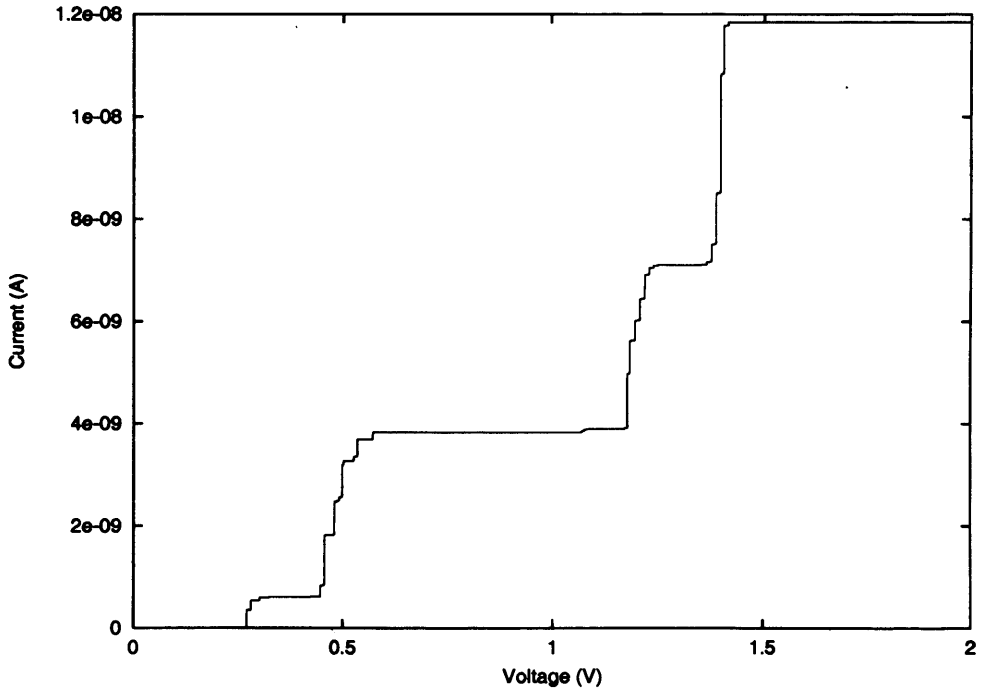


Figure 5.19: Current-Voltage characteristics for poly(G)-poly(C) DNA, with E_F=-0.4444 au.

I-V graph (figure 5.18), and we can say in this case that the DNA molecule is a small bandgap semiconductor. This shows the importance of including all of the appropriate electrons in order to get an accurate description of DNA's electrical properties. We have now seen that the current-voltage characteristics depend not only on which orbitals the metal reservoirs contact, but also the structure, base-pair composition, and now the chosen position of the Fermi energy. With so many variables it is perhaps not suprising how many different results have been quoted for DNA conductivity.

5.4 Charge density on contact atoms

It seems intuitive that in order to have good transmission from one end of the DNA molecule to the other, the corresponding electron state must have a large charge density situated on the atoms that are connected to the metal leads, the S atoms in the case of poly(G)-poly(C) DNA. However, we find that it is indeed necessary to have charge on these connecting atoms, but the maximum charge density is often located elsewhere, with only a relatively small amount of charge on the S atoms. For example, in figure 5.20 there is a good distribution of charge along the molecule resulting in a transmission of 0.95. However, the charge on the contact atoms (9 and 403) is only 35 and 2 (arb. units) respectively, compared with a maximum charge density of 350 (arb. units).

We find that the transmission depends crucially on the full Green function between the connecting atoms. From equation 4.22 we can see that the transmission depends on both the real and imaginary parts of the Green function G_{lr}, whereas the charge density on the end atoms depends only on the imaginary part of G_{ll} or G_{rr}.

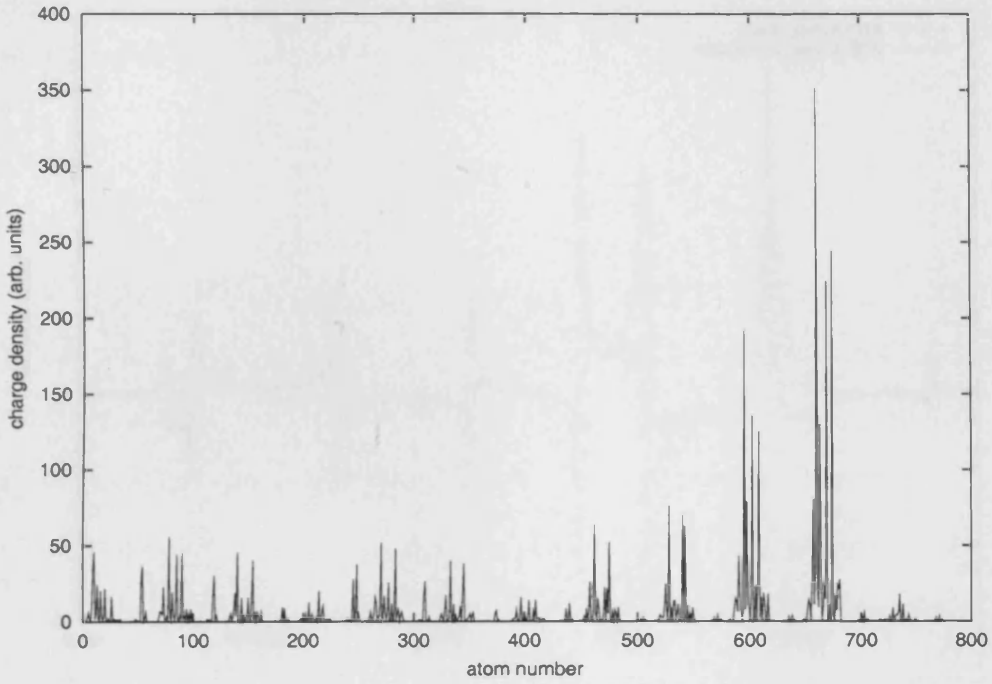


Figure 5.20: Charge density distribution for a state at $E = -0.4886$ au in poly(G)-poly(C) DNA.

Figure 5.21 shows an example of the Green function G_{lr} connecting atoms 9 and 403, the S contact atoms at the end of the molecule. We see that the structure is very complicated, but we note that both $\Re eG_{lr}$ and $\Im mG_{lr}$ are large at $E = -0.4886$, the energy of the state under discussion. We conclude that it is difficult to associate the transmission with simple concepts like the charge density on the connecting atoms, and we must consider the full Green function G_{lr} rather than just the imaginary part on the contact atom.

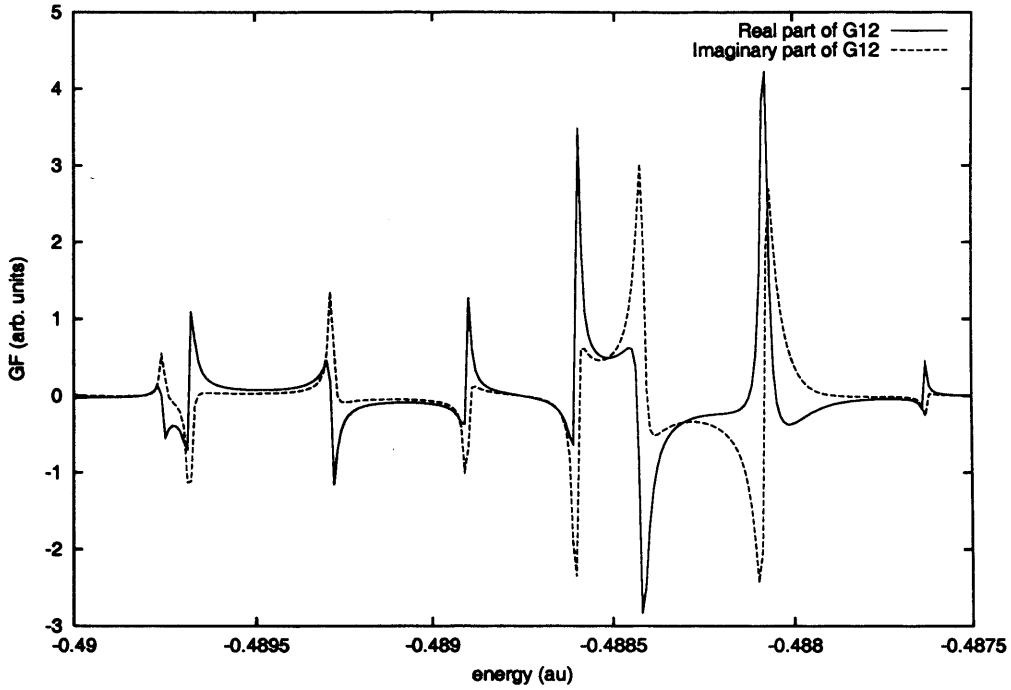


Figure 5.21: Real and imaginary part of GF for poly(G)-poly(C) DNA.

5.5 Period of vibrations

Our model so far has assumed that DNA is a static molecule. Of course this is not true. DNA has many vibrational modes, with twisting and tilting of the base-pairs as well as bending modes of the whole molecule. So does this mean that our model is inappropriate, since these vibrations will surely disrupt the pathway for electrical conduction. The important question to ask is, over what timescales do these vibrations occur? and how does this compare with the time for a charge carrier to move from one end of the molecule to the other? If the time taken for a charge carrier to move through the molecule is much less than the period of vibration then the molecule can be considered to be static.

To estimate the traversal time Δt of the carriers through the DNA, we can

employ the uncertainty principle,

$$\Delta E \Delta t \approx \hbar, \tag{5.2}$$

where ΔE is the width of the corresponding transmission peak. This approach is used in tunneling through quantum devices [146], but there is much discussion in the literature about the true definition of traversal time [147]. The widths of the transmission peaks vary depending on the interaction of the states with the metal contacts. Taking the root mean square width of the narrowest peaks, $\Delta E \sim 2 \times 10^{-6}$ au, we find that $\Delta t = 12$ ps, which is 1 ps per base-pair. Experimentally, Wan *et al.* [148] give the tunneling time through 5 base-pairs as 5 ps, the same as our estimate. Literature values for the frequency of twist fluctuations of DNA are given between 10^{11}–10^{12} s^{-1} by Kats and Lebedev [149]. We can see therefore that the time for tunneling through a base-pair in our model is comparable with the period of vibration of the bases. Therefore, we may not be completely justified in assuming a static model of DNA within the Born-Oppenheimer approximation.

Even within a static model, we can only represent one of the many structures of vibrating DNA frozen in time. As the model calculations of Roche [117] indicate, the twisting of the base-pairs will affect the conduction of DNA, giving different transmission results for the same molecule. However, our assumptions enable us to make preliminary conclusions about the condcutance of DNA.

CARDIFF UNIVERSITY

Chapter 6

Stretched DNA

One way in which DNA can be attached to a substrate is to use molecular combing [150]. In this method DNA is stretched across a surface substrate and can be positioned with nanoscale precision. An array of DNA molecules made in this way may have possible uses in high-resolution genomic studies, or as a template for nanoscale electrical devices [151]. It is therefore important to have an understanding of the properties of stretched DNA. In this chapter I will present the results of electronic structure calculations performed on stretched DNA using our embedding method.

DNA is a surprisingly elastic molecule when subjected to longitudinal forces. As far back as 1951, before Watson and Crick had published the structure of DNA, Wilkins *et al.* [152] were performing stretching experiments on DNA that suggested that when stretched far enough, DNA undergoes a structural transition allowing for elongation up to twice the length of the relaxed DNA molecule.

In the last 10 years many technological developments have allowed further study of this stretching transition, with techniques such as optical and magnetic tweezers

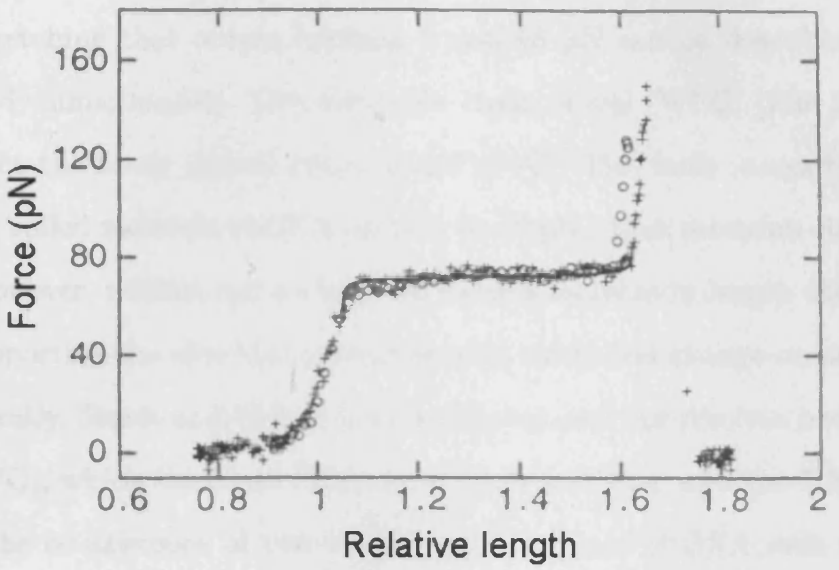


Figure 6.1: Graph of force against extension for 15.1 μm λ-DNA, with a pulling velocity of 1 μm/s ($\circ$) and 10 μm/s (+), taken from [153].

providing useful results. In 1996 both Cluzel *et al.* [153] and Smith *et al.*[154] found that when a force of ~70 pN is applied to a single molecule of double stranded DNA a structural transition occurs, after which a very small increase in force produces a large increase in the length of the DNA. This transition occurs over a very small range of only ~2 pN (see figure 6.1). For applied forces between 10 pN and 65 pN Smith *et al.* find that DNA stretches elastically, in accordance with Hooke's law, and when the force is removed the DNA molecule will return to its original shape. From this force region a value for the Young's modulus of DNA can be determined and is given by Smith *et al.* as 3.46×10^8 Pa, which is similar to that of wood, but much less than steel (10^{12} Pa) [155]. Both groups report that the DNA stretches to 1.7 times its original length, but beyond this any further stretching results in snapping the molecule.

The stretching that occurs between 0 and 65 pN can be described by various molecular dynamic models. The worm-like chain model (WLC) [156–158] and to a lesser extent the freely jointed chain model (FJC) [159] both successfully describe how a long coiled molecule of DNA unfolds (0-10 pN), then stretches elastically (10-65 pN). However, neither can explain the sudden increase in length observed above 65 pN, supporting the idea that a fundamental structural change occurs above this force. Recently, Storm and Nelson [160] have proposed the discrete persistent chain model (DPC), which combines elements of both the WLC and the FJC. They also allow for the co-existence of two conformational states of DNA each with its own parameter set. Their results correspond extremely well with overstretched DNA over the whole range of forces.

When the DNA is in its stretched state, after the structural transition, the molecule is given the name S-DNA. However, the exact structure of the molecule is unknown. Numerical simulations by Lebrun and Lavery [161] suggest two types of conformation depending on the type of stretching. Lebrun and Lavery find that if both 3′-3′ ends are pulled (see figure 1.2) then the DNA unwinds with the final structure resembling a ladder. However, if both 5′-5′ ends are pulled then the molecule remains helical, with the bases inclined. Both cases result in a rupture of the molecule once the original length has been doubled.

An alternative structure has been suggested by Harris [162]. Harris has performed molecular dynamic simulations on a DNA molecule containing 12 mixed base-pairs, (CGCAAAAAAGCG). Harris finds that when DNA is subjected to a longitudinal force, the width of the molecule decreases and the bases become tilted (see figure 6.2). As the DNA is stretched further a conformational change occurs, and either one or two "dislocations" can form which have the effect of reducing the

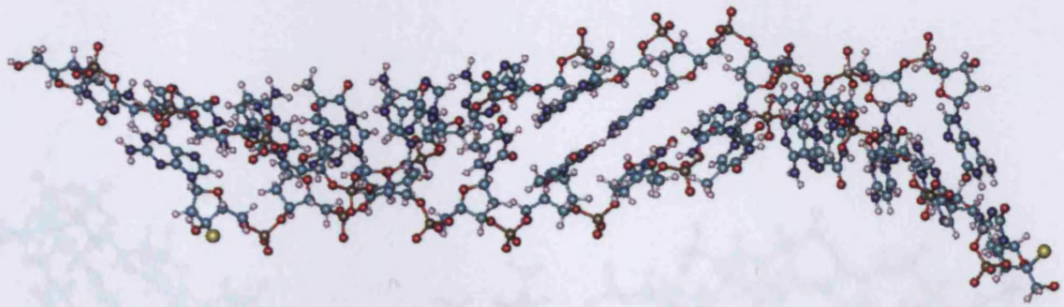


Figure 6.2: Stretched DNA.

tension on the molecule (see figures 6.3 and 6.4).

One possible way to discover the true conformation of S-DNA may be to perform electrical conduction experiments. The hope is that the different conformations of S-DNA will have different electrical characteristics, since the π-orbital pathway for conduction will be disrupted in different ways.

To this end I have applied my computational methods to this example in order to determine the merit of performing such experiments on this system. All structures were taken from Harris [162]. Three structures have been investigated, the first being DNA that has been stretched but without any holes (figure 6.2). The other two structures have undergone further stretching and have developed one and two holes respectively (figures 6.3 and 6.4).

For each of the three structures we substituted a S atom for a H atom on a sugar ring on the backbone at each end of the molecule to make contact with a gold reservoir, taking account of the increased bond length between the S and the DNA, taken from [131]. The density of states of the Au surface was projected onto a single gold atom and this was used to make contact with the backbone atom. The same contact atoms were used for each conformation of the stretched DNA molecule. The

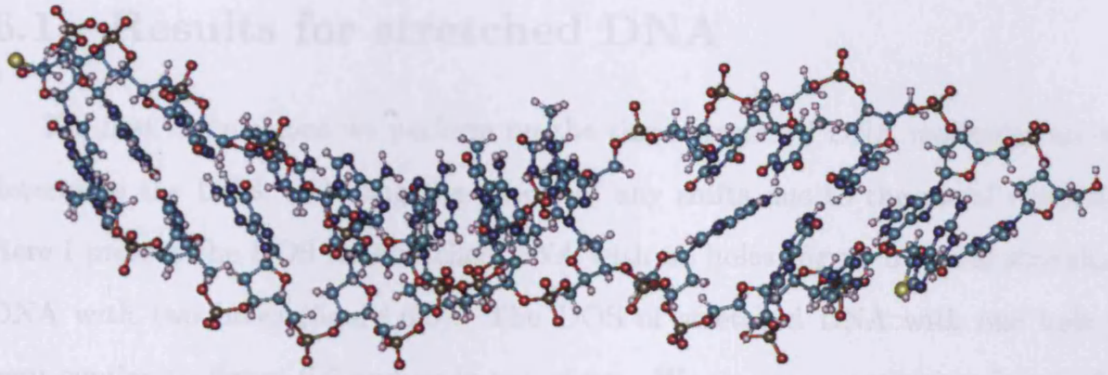


Figure 6.3: DNA that has undergone further stretching, showing one hole.

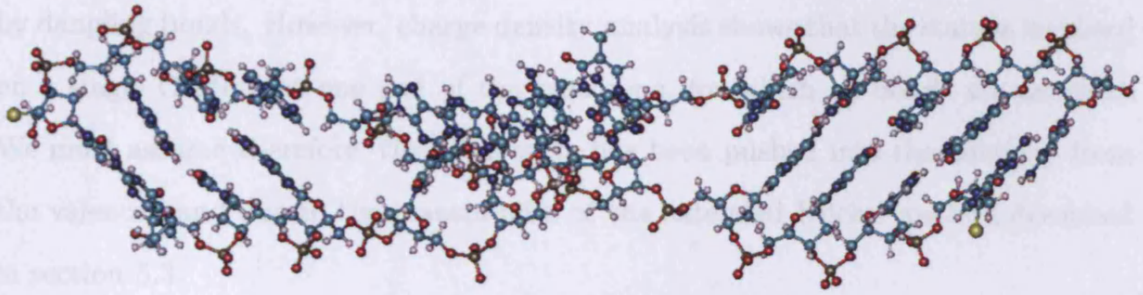


Figure 6.4: DNA that has undergone further stretching, showing two holes.

results of these calculations are presented in the following section.

6.1 Results for stretched DNA

The first calculations we perform on the three stretched DNA molecules are to determine the DOS, including the effects of any shifts due to the metal contacts. Here I present the DOS for stretched DNA with no holes (figure 6.5) and stretched DNA with two holes (figure 6.6). The DOS of stretched DNA with one hole is very similar to figure 6.6 and so is not given. We can see, comparing figures 6.5 and 6.6 for the stretched DNA with figure 5.3 for poly(G)-poly(C) DNA, that they look suprisingly similar. They all have essentially the same width of valence and conduction bands, with a large peak at ~ -0.5 au. There are however minor changes in the fine detail of the graphs, and this, as we shall see, will produce different conducting characteristics. If we examine more closely figure 6.5 for the stretched DNA with no holes, we can see that there is a single state in the middle of the bandgap at $E = -0.381$ au. At first we assumed that this was a defect state, caused by dangling bonds. However, charge density analysis shows that the state is localised on a single O atom at one end of the backbone, for which all bonds are satisfied. We must assume therefore, that this state has been pushed into the bandgap from the valence band due to the inaccuracies of the extended Hückel method discussed in section 5.3.

The position of the Fermi energy is taken as half way between the HOMO and LUMO, in the middle of the large bandgap at -0.381 au, -0.372 au, and -0.382 au for stretched DNA, DNA with one hole and DNA with two holes respectively.

The results for the three different conformations of stretched DNA are shown

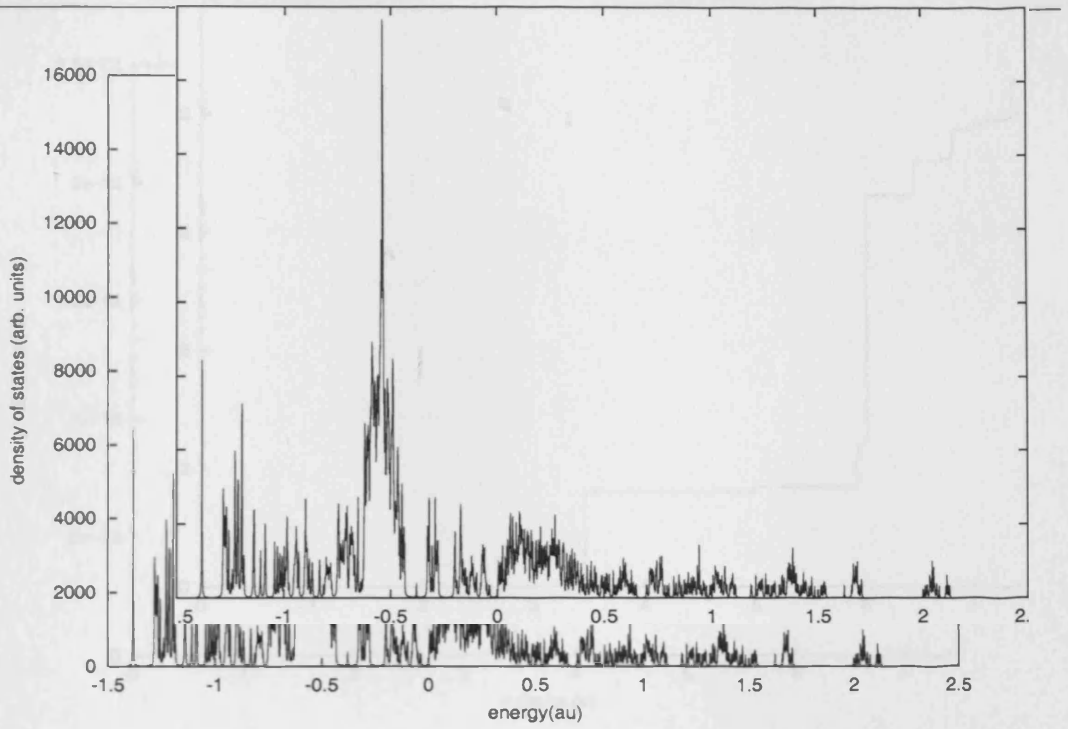


Figure 6.5: Density of states for stretched DNA with no holes.

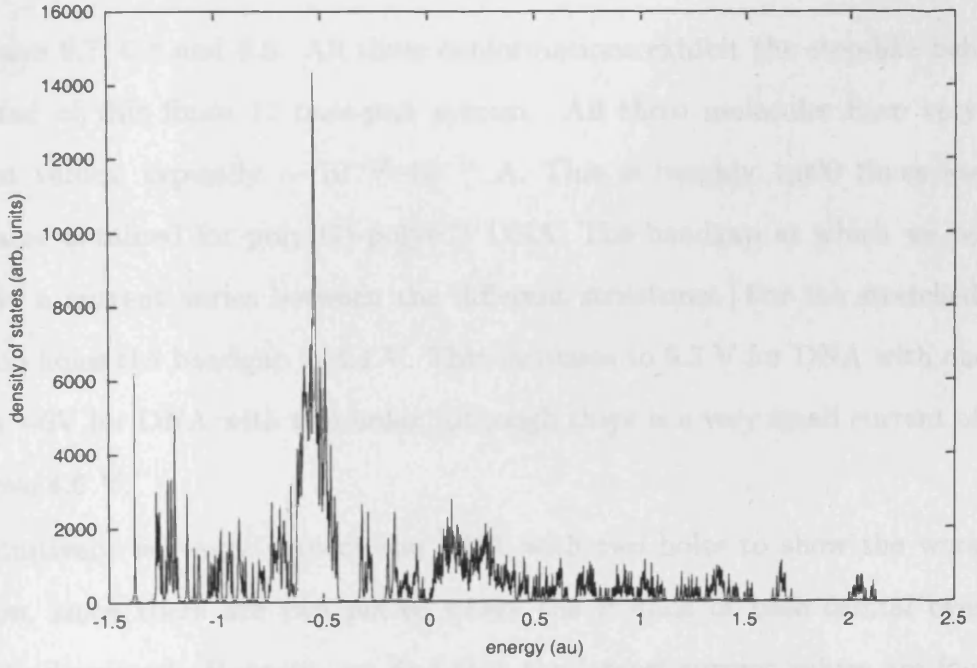


Figure 6.6: Density of states for stretched DNA with two holes.

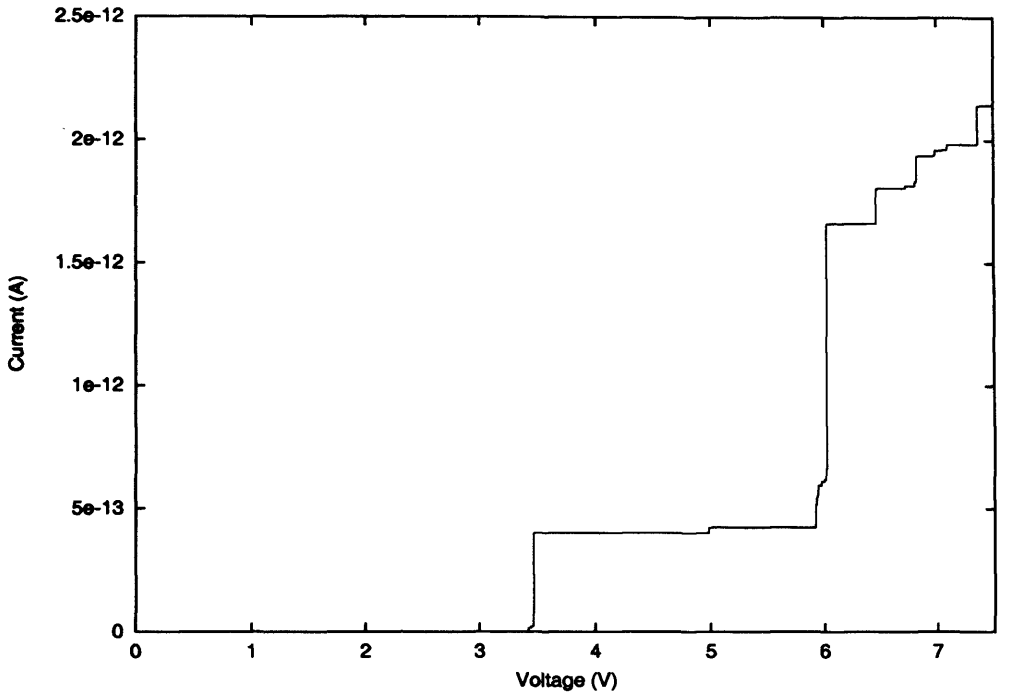


Figure 6.7: Current-Voltage characteristics of stretched DNA.

in figures 6.7, 6.8 and 6.9. All three conformations exhibit the step-like behaviour expected of this finite 12 base-pair system. All three molecules have very small current values, typically $\sim 10^{-12}$–10^{-11} A. This is roughly 1,000 times less than the value obtained for poly(G)-poly(C) DNA. The bandgap at which we begin to observe a current varies between the different structures. For the stretched DNA with no holes the bandgap is 3.4 V. This increases to 6.5 V for DNA with one hole, and is ~6V for DNA with two holes, although there is a very small current of 10^{-13} A above 4.6 V.

Intuitively we would expect the DNA with two holes to show the worst conduction, since there are two places where the π-stack of base orbital overlap is severely disrupted. However, we find that the largest current values are found for this molecule, although it does show a very large bandgap. The currents observed

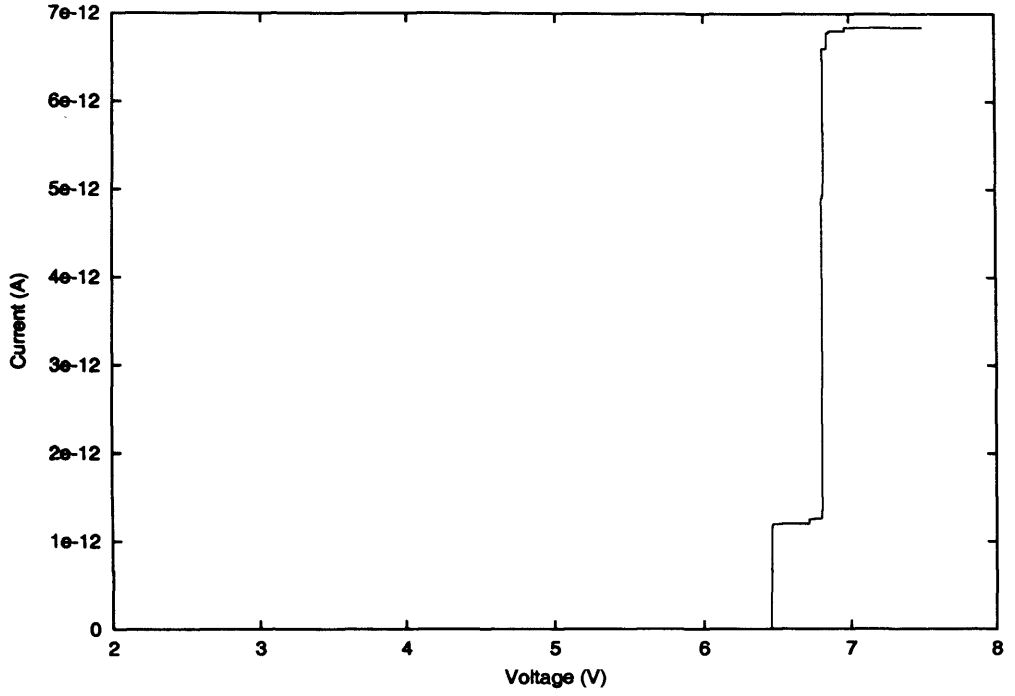


Figure 6.8: Current-Voltage characteristics of stretched DNA with one hole.

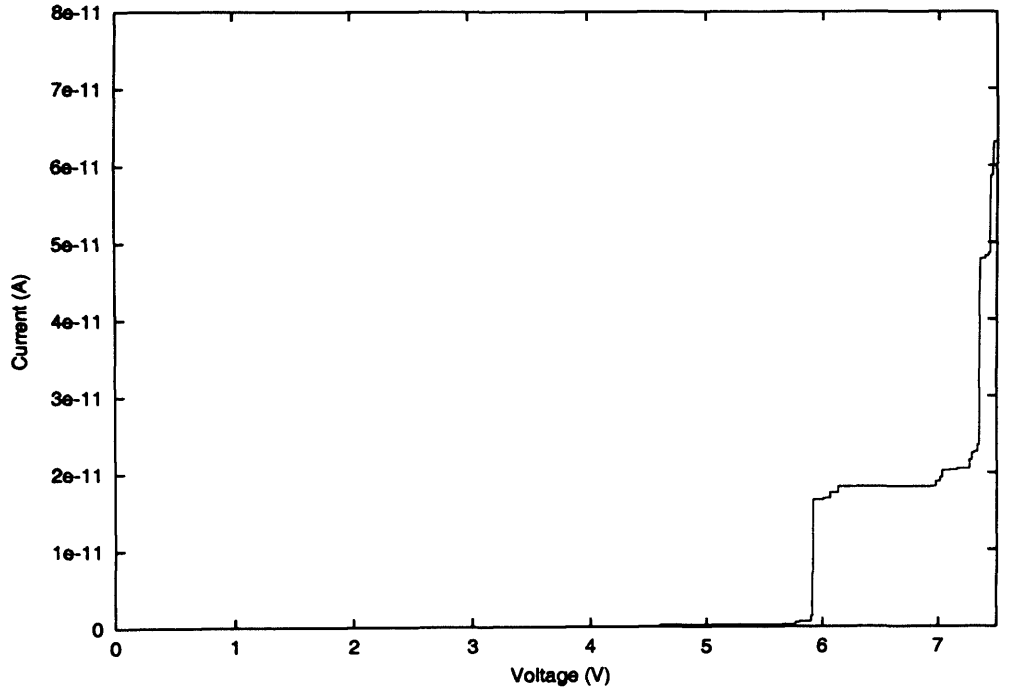


Figure 6.9: Current-Voltage characteristics of stretched DNA with two holes.

in all three structures are very low, and it may be that conduction this small is due to transmission through the backbone states. This may account for the current observed with two holes in the middle of the DNA molecule.

Our work with stretched DNA has shown that the conduction for the mixed-base sequence we use is very poor. Perhaps stretched poly(G)-poly(C) DNA may conduct more efficiently, however with such a disrupted π-stack is seems unlikely. We therefore conclude that it is doubtful that direct contact conduction experiments will provide an answer to the structure of S-DNA.

As for the future of molecular stretching experiments, Strick *et al.* [163] report that the force necessary to pull two strands of DNA apart is correlated with the base sequence, and proposes that this could be a method of determining the sequence of DNA. It has also been proposed that force experiments may help in the understanding of protein folding [164]. So it seems molecular stretching experiments are set to continue to provide an interesting and insightful look at the behaviour of molecules.

Chapter 7

Oligo(phenylene ethynylene)

Large molecules such as DNA are by no means the only types of molecules that physicists are investigating with a view towards molecular electronics applications. Many technological advances in recent years have allowed for the treatment of much smaller molecules. More and more physicists are turning their attention to small aromatic molecules in the hope that they will be the switches, diodes and transistors of the future [165].

In 1995 Tour *et al.* [166] made self-assembled monolayers (SAMs) from a range of oligomers with thiol, dithiol, thioacetyl or dithioacetyl end groups. Building on this work, several years later Reed *et al.* [167] were able to make I-V measurements of a single benzene-1,4-dithiol molecule, using a SAM break junction. They discovered semiconducting behaviour for this small aromatic molecule, with a bandgap of 0.7 V. DFT calculations by Seminario *et al.* [168] and Di Ventra *et al.* [169] confirmed their findings.

Son *et al.* [170] performed conduction experiments on SAMs of dodecanethiol ($S(CH_2)_{12}H$) and hexadecanethiol ($S(CH_2)_{16}H$). They used an interfacial force mi-

croscope, with a tip width of ~300 nm to contact the SAM. Their results show linear current-voltage characteristics at low voltages with a non-linear response at higher bias.

In 2001 Cui *et al.* [171] found linear conduction at low bias with non-linear conduction at higher voltages for the molecule octanedithiol (S-$(CH_2)_8$-S), using an AFM tip to contact single molecules of a SAM. In the same year Holmlin *et al.* [172] formed SAMs on two metal contacts and investigated the *I-V* characteristics for both aliphatic and aromatic molecules. All molecules showed non-linear *I-V* curves.

In another experiment with short conjugated molecules, Read *et al.* [173] attached functional groups to the middle ring of an oligo(phenylene ethynylene) molecule (OPE, figure 7.1), to fabricate a memory device. They used a self-assembled monolayer of the molecule between two gold contacts, and found that as the applied voltage is increased the molecule gives low conduction – this state of the molecule is given the information character 0. However after the initial voltage sweep, subsequent voltage sweeps showed that the molecule was a good conductor, and this state was assigned the information character 1. Read *et al.* find this to be completely reproducable and applying a reverse voltage to state 1 returns the molecule to state 0, the effect of erasing the single bit of information. This memory chip was found to be stable for 15 minutes. Obviously improvements need to be made to compete with silicon microchips, but small aromatic molecules may have a role to play in future computer circuits.

In 2002 Kushmerick *et al.* [76] showed experimentally that when OPE has symmetric contacts, that is S atoms at both ends, then the molecule behaves as a molecular wire. However, when the contacts are asymmetric the molecule behaves

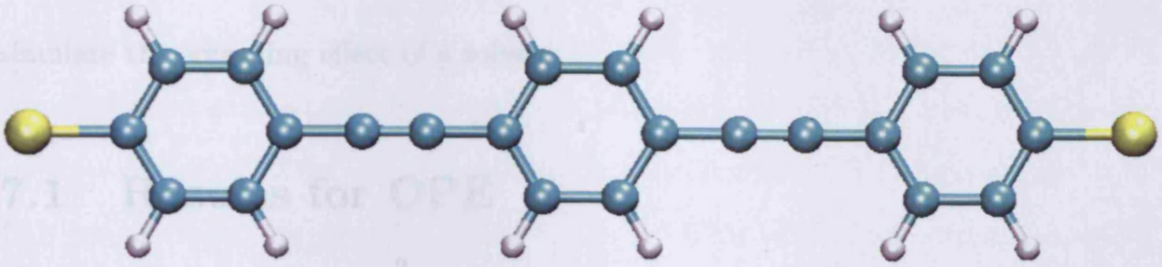


Figure 7.1: Oligo(phenylene ethynylene) molecule, showing three C phenyl rings each separated by two triple bonded carbon atoms, with a S atom at each end of the molecule.

like a molecular diode. Using a Green function approach within the extended-Hückel method Kushmerick *et al.* showed that the rectifying behaviour was a result of the difference between the local density of states on the contact atoms at each end of the OPE molecule. This leads to a non-symmetric division of the voltage at the ends, with the good coupling provided by the Au-S contact pinning the molecular levels of the asymmetric OPE molecule relative to the Fermi energy of that Au electrode.

In order to demonstrate the flexibility of our method, and moreover because these molecules have some interesting properties, we have chosen to investigate the electrical conduction of OPE (see figure 7.1). The OPE molecule consists of 22 C atoms and 12 H. We use a S atom at each end of the chain to bond the molecule to a gold reservoir. The three aromatic rings of the OPE molecule are separated by C triple bonds, giving rise to delocalised π-electrons, making them good candidates for electrical conductors.

We first generate the OPE molecule using the computer package Hyperchem 7, then energy-minimise the molecule with the AMBER force field to obtain the energy minimised spatial co-ordinates. As with the DNA simulations, the Coulomb interactions between the atoms were given a $1/r^2$ dependence rather than $1/r$, to

simulate the screening effect of a solvent.

7.1 Results for OPE

The DOS of the OPE molecule is shown in figure 7.2. This figure includes the energy shift of the states which occurs as a result of the interaction with the gold reservoirs. Figure 7.2 shows all 108 states of the system, each capable of accommodating two electrons. A few isolated states may be seen, for example at 1.25 au and 2.30 au, though, most of the states overlap to form larger features such as those near -0.5 au. This overlap is due not only to the imaginary part introduced artificially to broaden the states, but the states are also broadened by the interaction of the molecule with the gold contact surfaces. We find that when comparing the

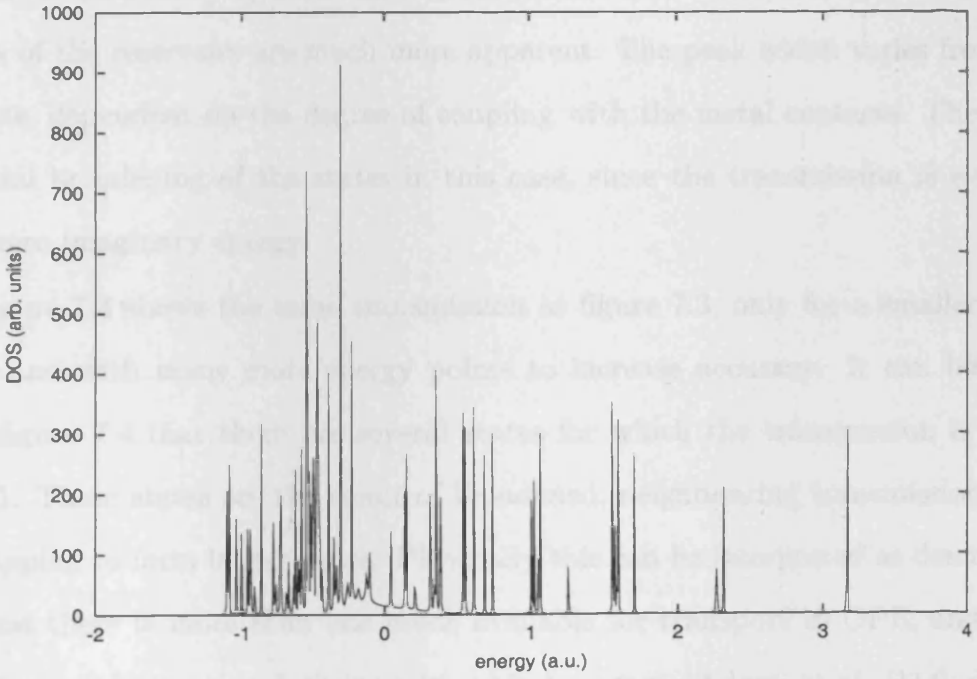


Figure 7.2: DOS of OPE molecule attached to a gold contact at each end.

DOS of the molecule connected to the gold contact, with the unconnected DOS, there is a significant energy shift observed, much more than for DNA. This is presumably because of the size difference between the two molecules. In DNA two atoms out of 760 are contacted compared with two in 36 for OPE. The OPE molecule has 112 electrons, and integrating the DOS using the accurate complex contour integration method, we find that the HOMO lies at -0.452 au.

Making contact with the S atoms at each end, the transmission was calculated as a function of energy, and the results are shown in figure 7.3. It can be seen that there are broad overlapping peaks of transmission giving rise to several transmission bands, with the HOMO lying at the lower edge of one of these bands. This is markedly different to the case of DNA, where we observe much narrower peaks, and more isolated transmission features. The width of the transmission peaks is increased in the case of OPE since the molecule is much smaller than DNA, and so the broadening effects of the reservoirs are much more apparent. The peak width varies from state to state, dependent on the degree of coupling with the metal contacts. There is no artificial broadening of the states in this case, since the transmission is evaluated with zero imaginary energy.

Figure 7.4 shows the same transmission as figure 7.3, only for a smaller energy range and with many more energy points to increase accuracy. It can been seen from figure 7.4 that there are several states for which the transmission is greater than 1. These states are the result of broadened, neighbouring transmission peaks, overlapping to form larger states. Physically this can be interpreted as demonstrating that there is more than one mode available for transport in OPE, and this is indeed possible since we have a multi-orbital contact. Adessi *et al.* [114] also find $T > 1$ for their ordered infinite DNA chain.

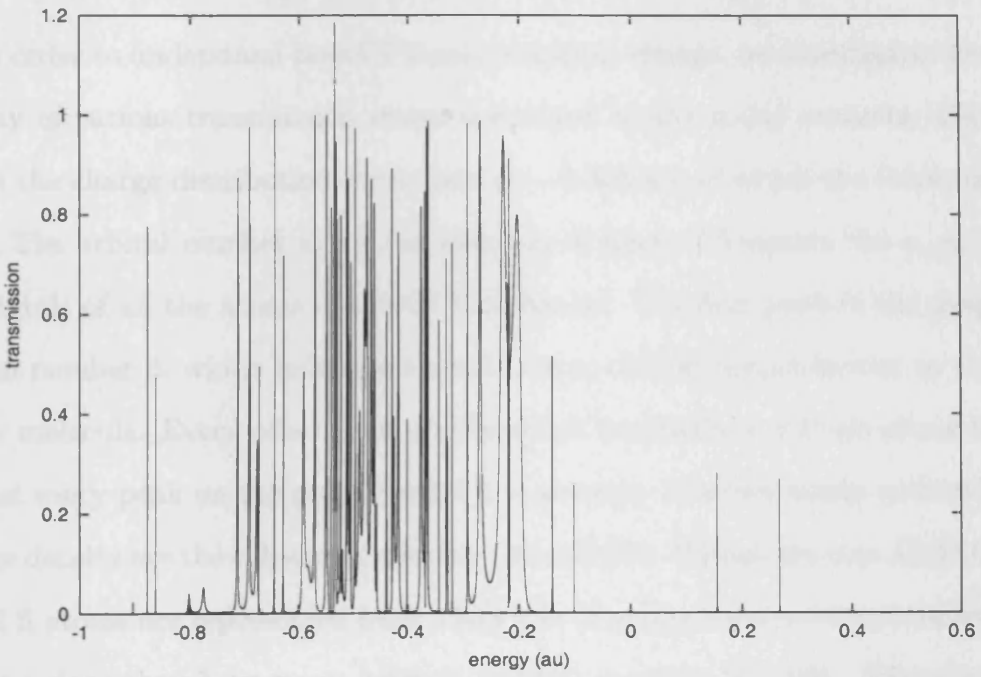


Figure 7.3: Transmission of OPE molecule attached to a gold contact at each end.

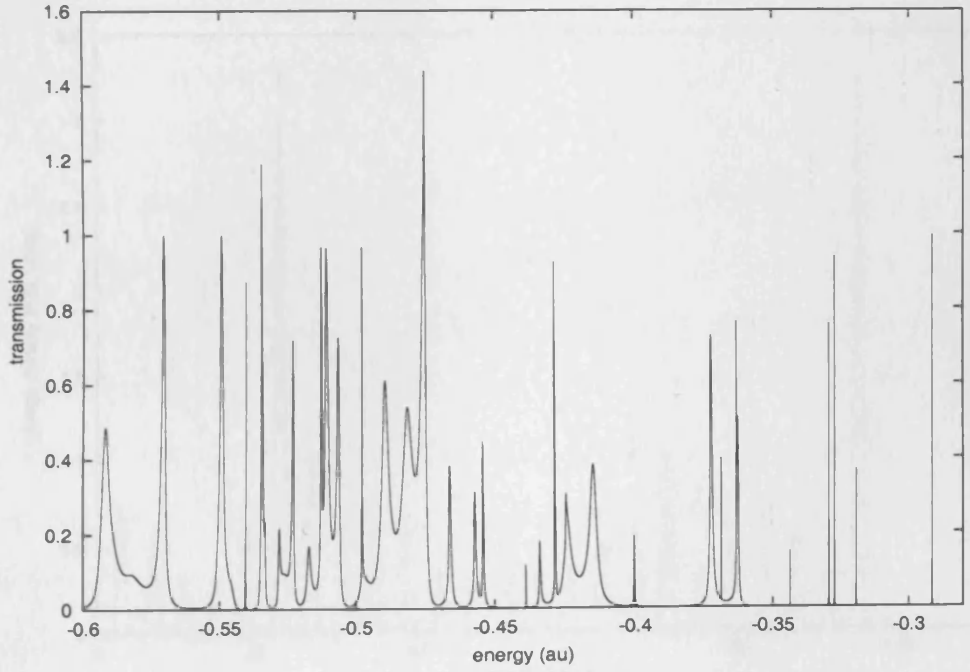


Figure 7.4: Transmission of OPE molecule attached to a gold contact at each end, near the Fermi energy.

In order to understand how OPE can transport charge, we investigate the charge density of various transmission states connected to the metal contacts. Figure 7.5 shows the charge distribution for a state at -0.363 au, of which the transmission is 0.75. The orbital number along the abscissa of figure 7.5 counts the s, p_x, p_z and p_y orbitals of all the atoms in the OPE molecule. The first peak in the graph is at orbital number 3, which indicates that it is a p_z-orbital perpendicular to the plane of the molecule. Every other peak on the graph is exactly 4 orbitals above the last, so that every peak on the graph lies on a p_z orbital. The two peaks with maximum charge density are the sulphur p_z-orbitals (23 and 95). We can see that all 22 C atoms and 2 S atoms are represented here. Only the H atoms show no appreciable charge density, since they have no p_z orbitals (orbital numbers 97–108). Therefore, there will be a plane of delocalised π-electrons above and below the molecule, providing

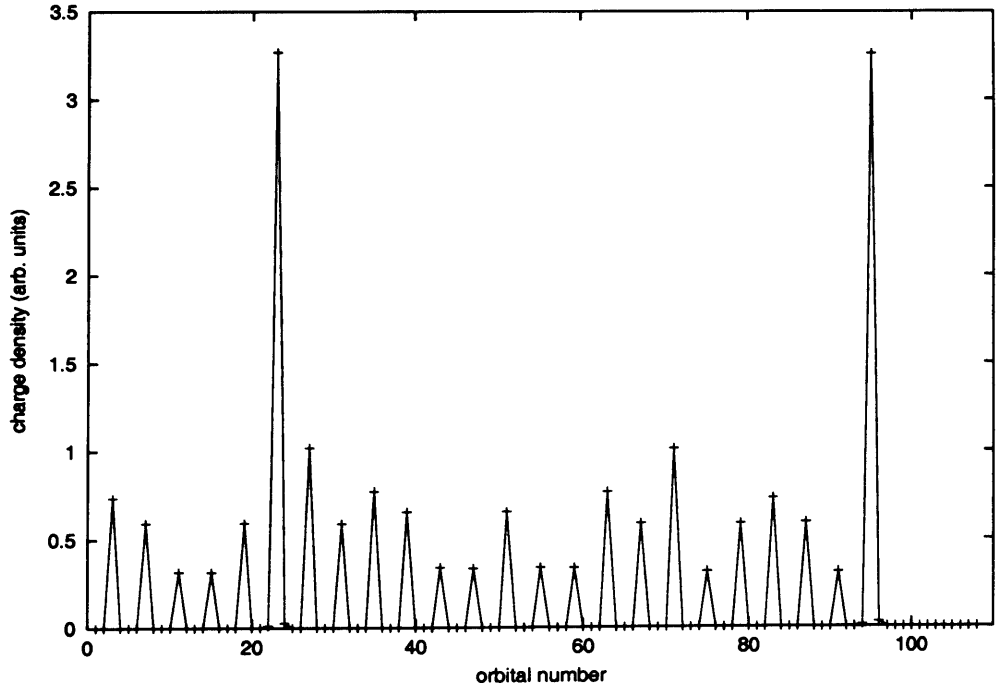


Figure 7.5: Charge density of OPE at energy -0.363 au.

a pathway for charge transport. This explains the large number of states that give good transmission. It can also be shown from the graph that the charge density is exactly symmetric, as we would expect from a symmetric molecule, although this is not obvious due to the way the orbitals are labelled.

Finally we investigate the current-voltage characteristics of OPE. This will tell us what size of current we can expect from this molecule, and also whether it will behave like a semiconductor or a conductor. As with DNA we place the Fermi level halfway between the HOMO and LUMO, and shift the DOS of the left and right Au contacts by +2.0 V and −2.0 V respectively. The *I-V* graph is then obtained by integrating the transmission peaks. The results of the *I-V* calculations are given in figure 7.6. We can again see the step-like behaviour that we have seen with DNA, although the steps are much less defined here, since the broader transmission peaks for OPE give much smoother steps. The current values for OPE are $\sim 10^{-5}$ A, a factor of 10^3 higher than poly(G)-poly(C) DNA, which was the best conductor out of all the DNA molecules we investigated. These values for current compare very well with the experimental and theoretical work of Kushmerick *et al.* [76] who report a current of 2 μA at 1 V compared with 2.2 μA for our model.

It can be seen from figure 7.6 that the *I-V* graph is almost linear, suggesting that it may behave like a metal. However, there appears to be a small bandgap of 0.3 V which suggests that OPE is in fact a small bandgap semiconductor. When we look at this more closely we find that there is no bandgap at all. The current rises steadily from 0 V, although at low voltage the current is of the order of nA and therefore is too small to register on figure 7.6. This occurs since the Fermi energy is sandwiched between two transmission peaks, so even a very small voltage is enough to reach the tails of the transmssion peaks.

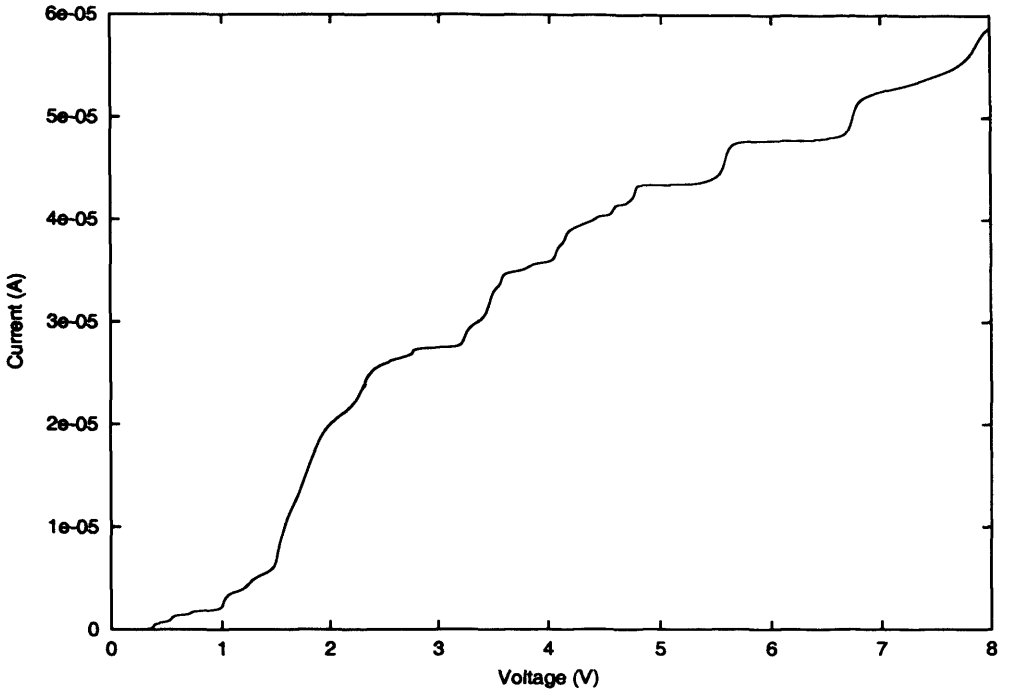


Figure 7.6: Current-Voltage characteristics of OPE.

If we assume that the I-V curve shown in figure 7.6 is linear, then we can calculate the resistance simply using $R = V/I$. This gives a resistance of 133 kΩ, compared with 200 Ω for a Cu wire of about the same dimensions. This shows that although OPE is a good molecular conductor, it still does not conduct as well as traditional conductors. The resistance of OPE is in fact a combination of contact resistance plus the resistance of the molecule itself. As discussed earlier there are at least two modes of transport in OPE, hence from equation (3.15) the contact resistance will be a maximum of 6.45 kΩ, suggesting that molecular resistance dominates in OPE. As discussed in chapter 3, the energy from both the contact and molecular resistances will be dissipated in the metal reservoirs as heat.

Our success with OPE shows that it is extremely simple and effective to apply our method to different molecules. From our results it is clear that OPE is an

excellent molecular conductor, with metallic behaviour, and may prove to be very useful in molecular electronics design.

Chapter 8

Conclusion and Outlook

In this thesis we have presented a new method for performing electronic structure calculations for large molecules. Our embedding method takes advantage of the natural partitioning of nucleotide pairs in DNA to perform efficient order-N calculations of the density of states and conductance of several types of DNA molecules, giving results comparable to both experiment and other theoretical studies.

Our results show that transmission through DNA is highly dependent on the atom to which the metal electrode is connected. Charge transport also varies between DNA molecules containing different bases. We find that poly(G)-poly(C) DNA conducts much better than a DNA molecule with mixed bases, and also the increased order in an energy-minimised DNA molecule allows for better conduction than DNA taken from x-ray diffraction experiments.

The calculations performed on stretched DNA show extremely poor conduction for all three structures studied. We therefore predict that it would be extremely difficult to determine the structure of S-DNA by performing electrical conduction experiments.

We have also demonstrated the flexibility of our method by applying it to OPE. Very minor changes allowed us to calculate the DOS and I-V characteristics for this short conjugated molecule. We find that OPE is a good molecular conductor, due to the delocalised π-electrons above and below the plane of the molecule. We believe that our method is applicable to a wide variety of molecules, although, special care will have to be taken when considering proteins that fold back on themselves, since our method presently only takes into account neighbouring sections when finding the embedding potential.

Although we find our model satisfactory for the investigation of the electrical properties of DNA there are several ways in which it could be improved. One way to improve the accuracy of the model would be to use a larger basis set. In this work we have only used a single Slater type orbital to describe the atomic orbitals, except for the *d*-orbitals on the metal contacts. This can be extended to include a series of Slater type orbitals, making the description more accurate.

Another improvement on the current model would be a better treatment of the electrode-molecule interface. The current method uses a very accurate description of the surface density of states of Au to represent the contact between the electrode and the molecule. However, it is limited to a single atom-to-atom contact. This should be extended to the more realistic case where the contact region extends over a larger area.

A more fundamental improvement on the present model would be to make it self-consistent. This could be done either using the iterative extended Hückel method or using density functional theory. Self-consistent methods usually give more reliable results since they take account of all the electron-electron interactions.

Further areas to study which could immediately be carried out include length

dependent calculations, to see how the charge transport rate varies with increasing distance. Another interesting area of study would be the effect of raising the temperature on conduction. The inclusion of frozen phonons would be relatively straightforward. There are also many more molecules that have interesting properties to investigate, and which would be suitable for our model, such as cytochrome, RNA and proteins.

In general the future of molecular electronics is looking bright. Many commercial applications already exist for conducting polymers, and with the laboratory fabrication of molecular diodes and transistors, the field is moving apace. However, whether molecules will replace silicon in the semiconductor industry will probably depend more on economic rather than technological reasons.

Appendix A

Contour Method of Integration

As described in section 5.4, in order to find the position of the Fermi energy, we need to integrate a very peaky density of states. Simply integrating the area under the peaks using the trapezium rule is insufficiently accurate. We therefore use a method of contour integration that allows us to quickly and accurately determine the number of states within a given energy range, and this method is described below.

Integrating the DOS we find the total number of states N between the chosen integration limits. Below we show an example of integrating up to the Fermi energy for a DOS similar to that of DNA. The total number of states up to the Fermi energy is given by

$$N = \int_0^{E_F} n(E)dE, \tag{A.1}$$

where $n(E)$ is the density of states, and using equation (4.16) this may be written as,

$$N = \frac{1}{\pi}\int_0^{E_F} \Im m \mathrm{Tr} G(E + i\epsilon)dE, \tag{A.2}$$

where ϵ is infinitesimal. In order to accurately integrate what is normally a very peaky DOS, we vary the imaginary part of the energy over the range of the integration. The imaginary part is large in the middle of the range, tending to zero at the ends (figure A.1), which broadens the peaky states. The energy points used in the method of Gaussian integration, are distributed so that a greater number of them are near the upper and lower limits of the integral, which improves the accuracy of the method. In order to convert the integral from just above the real axis (contour 1) to a contour integral along contour 2 we use the fact that the Green function only has singularities along the real axis, so from Cauchy's theorem

$$\int_1 \mathrm{Tr}G dE = \int_2 \mathrm{Tr}G dE. \tag{A.3}$$

Taking the imaginary part of equation (A.3), and multiplying by $1/\pi$ we find,

$$\frac{1}{\pi}\int_1 \Im m(\mathrm{Tr}G) dE = \frac{1}{\pi}\int_2 \Im m(\mathrm{Tr}G dE). \tag{A.4}$$

From figure A.1 it can be seen that any complex energy lying on contour 2 can be given by

$$E = \frac{E_F}{2}(1 + e^{i\phi}), \tag{A.5}$$

and hence dE is given by

$$dE = i\frac{E_F}{2}e^{i\phi}d\phi. \tag{A.6}$$

So changing variables in equation (A.4) we find,

$$N = \frac{E_F}{2\pi}\int_{\pi}^{0} \Im m(ie^{i\phi}\mathrm{Tr}G)d\phi, \tag{A.7}$$

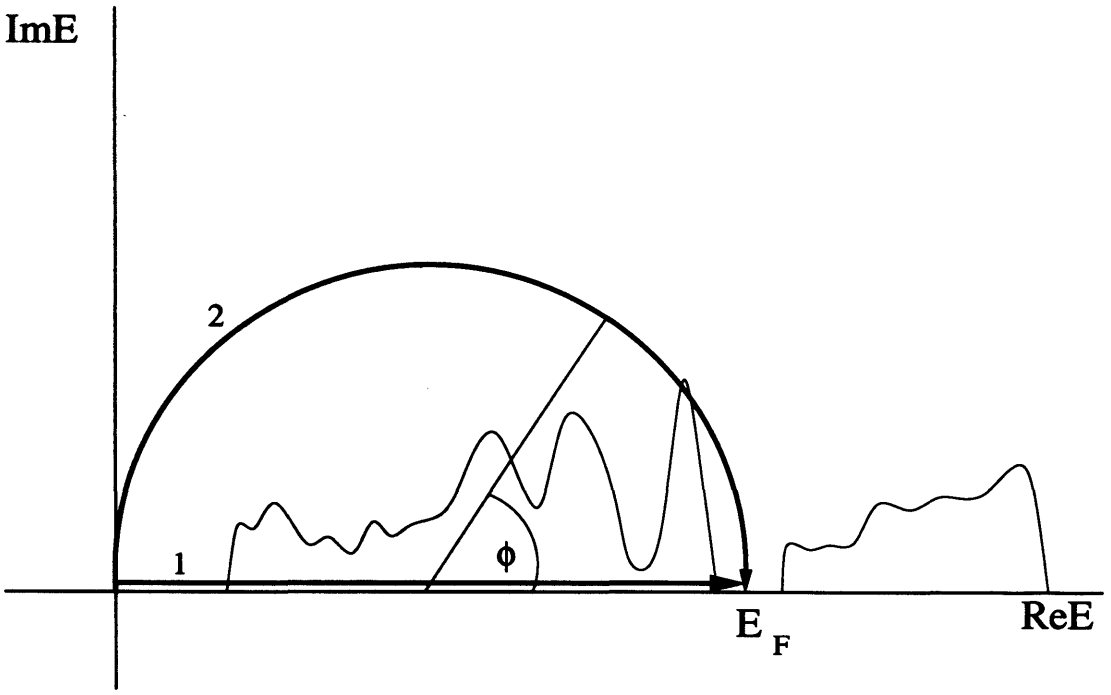


Figure A.1: Diagram showing how the contour integration method varies the imaginary energy over the range of the integral.

and converting to trigonometric functions,

$$\begin{aligned} N &= -\frac{E_F}{2\pi}\int_0^{\pi}(\cos\phi \mathrm{Tr}\Re eG - \sin\phi \mathrm{Tr}\Im mG)d\phi \\ &= \frac{E_F}{2\pi}\int_0^{\pi}(\sin\phi \mathrm{Tr}\Im mG - \cos\phi \mathrm{Tr}\Re eG)d\phi. \end{aligned} \tag{A.8}$$

Equation (A.8) can then be calculated numerically, and due to the broadening of the peaks in the middle of the range we only need a maximum of 128 energy points for an excellent degree of accuracy. The trapezium method of integration used previously needed tens of thousands of energies to achieve the same accuracy. Therefore we can quickly and reliably find the number of states within a given energy range, and the positions of HOMO and LUMO.

Appendix B

Programming

Initially we tried to use some of the computing code of a previous Cardiff student for applying Sharma's method [96]. However, after discovering problems with the code we decided to write our own programs with numerical evaluation of the overlap intgrals. Detailed below is a summary of the logic of the programs and how various quantities are calculated. All of the programming was done in Fortran 90, making use of the inbuilt matrix multiplication function MATMUL which is extremely efficient.

The structure of the program shown in the flow chart demonstrates how easy it is to apply our method to different molecules. After choosing a suitable molecule that may be split into sections, all we require are the spatial co-ordinates and atom types in a standard pdb format. The remainder of the code remains unchanged. If we are to make this method self-consistent then we would need to replace the extended Hückel calculation of the overlap integrals and Hamiltonian matrix elements (steps 2 & 3), with a DFT evaluation including a form of feedback for this iterative process.

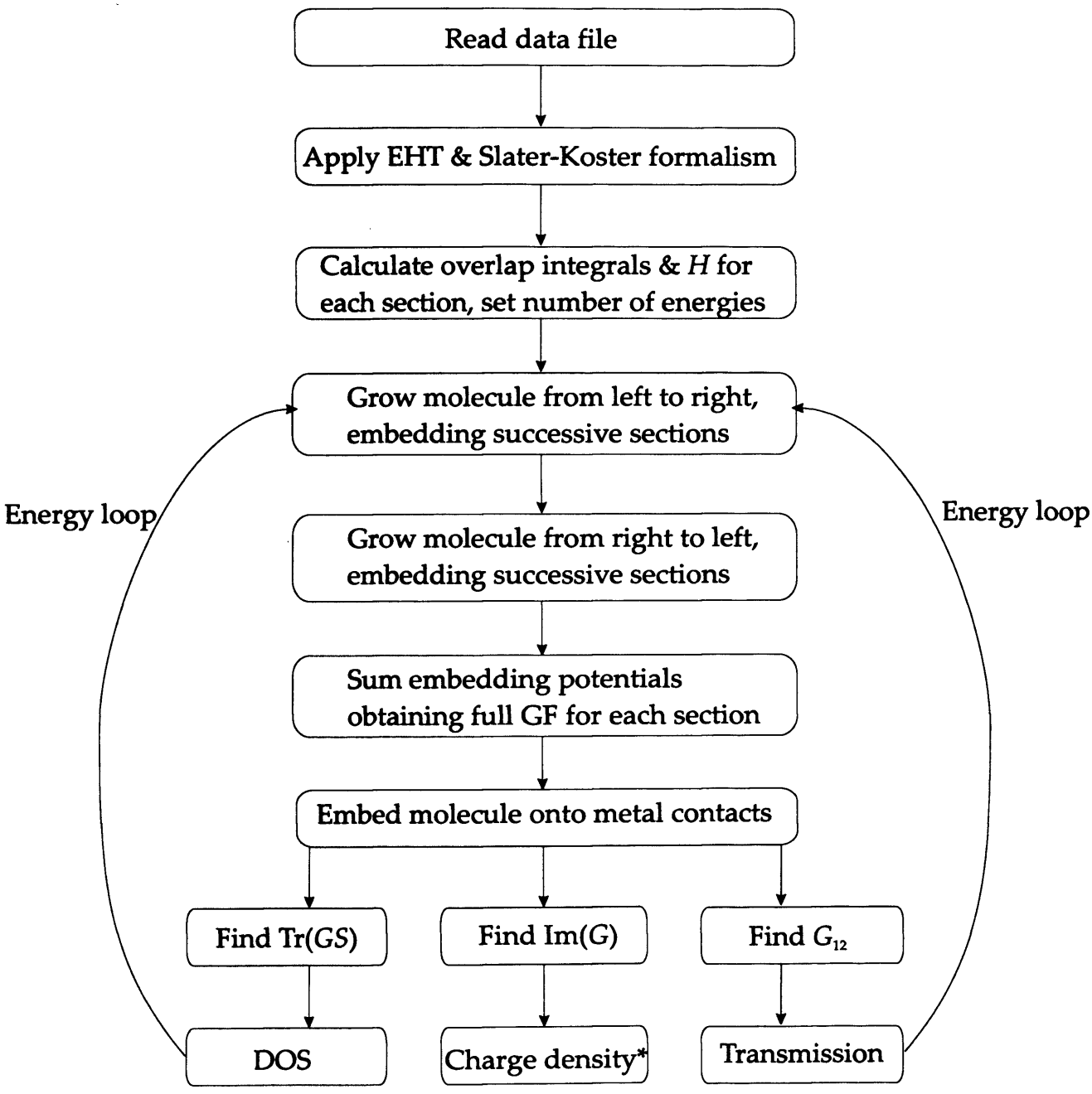


Figure B.1: Flow chart showing the programming logic.*Charge density is evaluated at a single energy.

Bibliography

[1] D. Normile. *Science*, 293(5531):787, 2001.

[2] L. B. Sivitz. *Science News*, 158(22):350, 2000.

[3] H. Shirakawa, E. J. Louis, A. G. Macdiarmid, C. K. Chiang, and A. J. Heeger. *J. C. S. Chem. Comm.*, pages 578–580, 1977.

[4] J. E. McGinness, P. M. Corry, and P. Proctor. *Science*, 183:853–855, 1974.

[5] S. Iijima. *Nature*, 354(56), 1991.

[6] M. Yu, O. Lourie, M. J. Dyer, K. Moloni, T. F. Kelly, and R. S. Ruoff. *Science*, 287:637–640, 2000.

[7] T. Yildirim, O. Gülseren, Ç. Kiliç, and S. Ciraci. *Phys. Rev. B*, 62(19), 2000.

[8] A. Bachtold, M. S. Fuhrer, S. Plyasunov, M. Forero, E. M. Anderson, A. Zettl, and P. L. McEwen. *Phys. Rev. Lett.*, 84(26):6082–6085, 2000.

[9] J. W. G. Wildoer *et al. Nature*, 391:59, 1998.

[10] C. W. Tang and S. A. Slyke. *Appl. Phys. Lett.*, 51(913), 1987.

[11] L. M. Adleman. *Science*, 266(5187):1021–1024, 1994.

[12] Y. Benenson, T. Paz-Elizur, R. Adar, E. Keinan, Z. Livneh, and E. Shapiro. *Nature*, 414(6862):430–434, 2001.

[13] J. Fritz *et al. Science*, 288(5464):316–318, 2000.

[14] J. Chen and N. C. Seeman. *Nature*, 350(6319):631–633, 1991.

[15] N. C. Seeman. *Nature*, 421(6921):427–431, 2003.

[16] D. D. Eley and D. I. Spivey. *Trans. Faraday Soc.*, 58:411–415, 1962.

[17] D. Van Lith, J. M. Warman, M. P. de Haas, A. Hummel, and C. Prinsen. *J. Chem. Soc., Faraday Trans.*, pages 2933–2943, 1986.

[18] M. E. Nunez, D. B. Hall, and J. K. Barton. *Chemistry and Biology*, 6(85), 1999.

[19] C. J. Murphy, M. R. Arkin, Y. Jenkins, N. D. Ghatlia, S. H. Bossman, N. J. Turro, and J. K. Barton. *Science*, 262(1025), 1993.

[20] E. Braun, Y. Elchen, U. Sivan, and G. Ben-Yoseph. *Nature*, 391:775, 1998.

[21] H. W. Fink and C. Schonenberger. *Nature*, 398(6726):407–410, 1999.

[22] D. Porath, A. Bezryadin, S. de Vries, and C. Dekker. *Nature*, 403:635–638, 2000.

[23] A. Y. Kasumov et al. *Science*, 291(5502):280–282, 2001.

[24] L. Helman. *Science*, 293:583–585, 2001.

[25] A. R. Butler et al. *Nature Biotechnology*, 20(7):713–716, 2002.

[26] F. Kreppel, V. Biermann, and S. Kochanek. *Human Gene Therapy*, 13(10):1151–1156, 2002.

[27] Molecular Biology of the Cell. B. Alberts *et al. Garland Publishing.* page 99, 1994.

[28] Molecular Biology of the Cell. B. Alberts *et al. Garland Publishing.* page 101, 1994.

[29] J. Watson and F. Crick. *Nature*, 171(4356):373–374, 1953.

[30] K. W. Hipps. *Science*, 294:536–537, 2001.

[31] G. Cuniberti, L. Craco, D. Porath, and C. Dekker. *Phys. Rev. B*, 65:241314, 2002.

[32] F. Moreno-Herrero, P. Herrero, F. Moreno, J. Colchero, C. Gómez-Navaro, J. Gómez-Herrero, and A. M. Baró. *Nanotechnology*, 14(2):128–133, 2003.

[33] P. Lincoln, E. Tuite, and B. Norden. *J. Am. Chem. Soc.*, 119(1454), 1997.

[34] S. Priyadarshy, S. M. Risser, and D. N. Beratan. *J. Biol. Inorg. Chem.*, 3(196), 1998.

[35] N. J. Turro and J. K. Barton. *J. Biol. Inorg. Chem.*, 3(201), 1998.

[36] T. L. Netzel. *J. Biol. Inorg. Chem.*, 3(210), 1998.

[37] F. D. Lewis and R. L. Letsinger. *J. Biol. Inorg. Chem.*, 3(215), 1998.

[38] E. S. Krider and T. J. Meade. *J. Biol. Inorg. Chem.*, 3(222), 1998.

[39] T. T. Williams, D. T. Odom, and J. K. Barton. *J. Am. Chem. Soc.*, 122:9048–9049, 2000.

[40] C. Gómez-Navarro, M. Álvarez A. Gil, P. J. de Pablo, F. Moreno-Herrero, I. Horcas, R. Fernańdez-Sánchez, J. Colchero, J. Gómez Herrero, and A. M. Baró. *Nanotechnology*, 13(314), 2002.

[41] A. Gil, P. J. de Pablo, J. Colchero, J. Gómez Herrero, and A. M. Baró. *Nanotechnology*, 13(309), 2002.

[42] C. Gómez-Navarro, F. Moreno-Herrero, P. J. de Pablo, J. Colchero, J. Gómez Herrero, and A. M. Baró. *Proc. Nat. Acad. Sci.*, 99(8484), 2002.

[43] M. Bockrath, N. Markovic, A. Shepard, M. Tinkham, L. Gurevich, L. P. Kouwenhoven, M. W. Wu, and L. L. Sohn. *Nano. Lett.*, 2(187), 2002.

[44] P. Tran, B. Alavi, and G. Gruner. *Phys. Rev. Lett.*, 85(7):1564–1567, 1999.

[45] P. J. de Pablo, F. Moreno-Herrero, J. Colchero, J. Gómez Herrero, P. Herrero, A. M. Baró, P. Ordejón, J. M Soler, and E. Artacho. *Phys. Rev. Lett.*, 85(23):4992–4995, 2000.

[46] A. J. Storm, J. van Noort, S. de Vries, and C. Dekker. *Appl. Phys. Lett.*, 79(23):3881–3883, 2001.

[47] Y. Zhang, R. H. Austin, J. Kraeft, E. C. Cox, and N. P. Ong. *Phys. Rev. Lett.*, 89(198102), 2002.

[48] H. Tabata L. Cai and T. Kawai. *Appl. Phys. Lett.*, 77(19), 2000.

[49] H. Watanabe, C. Manabe, T. Shigematsu, K. Shimotani, and M. Shimizu. *Appl. Phys. Lett.*, 79(15):2462–2464, 2001.

[50] K. H. Yoo, D. H. Ha, J. O. Lee, J. W. Park, Jinhee Kim, J. J. Kim, H. Y. Lee, T. Kawai, and Han Yong Choi. *Phys. Rev. Lett.*, 87(19), 2001.

[51] J. S. Hwang, K. J. Kang, D. Ahn, G. S. Lee, D. J. Ahn, and S. W. Hwang. *Appl. Phys. Lett.*, 81(6):1134–1136, 2002.

[52] J. M. Ziman. Principles of the theory of solids. *Camb. Univ. Press*, 1965.

[53] E. M. Conwell and S. V. Rakhmanova. *Proc. Natl. Acad. Sci.*, 97(9):4556–4560, 2000.

[54] F. Mandl. Quantum mechanics. *Wiley*, page 187, 1998.

[55] L. Hedin, S. Lundqvist, H. Ehrenrich, F. Seitz, and D. Turnbull. *Solid State Physics*, 28:225–300, 1973.

[56] M. Bennett and J. C. Inkson. *J. Phys. C*, 10(987), 1977.

[57] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136(B864), 1964.

[58] L. H. Thomas. *Proc. Cambridge Philos. Soc.*, 23:542, 1927.

[59] E. Fermi. *Z. Phys.*, 48(73), 1928.

[60] P. A. M. Dirac. *Proc. Cambridge Philos. Soc.*, 26:376, 1930.

[61] W. Kohn and L. J. Sham. *Phys. Rev.*, 140(A1133), 1965.

[62] J. P. Perdew and A. Zunger. *Phys. Rev. B*, 23(5048), 1981.

[63] P. M. Cepherley and B. J. Alder. *Phys. Rev. Lett*, 45(566), 1980.

[64] L. H. Yang, A. P. Smith, R. Benedek, and D. D. Koelling. *Phys. Rev. B*, 47(24):16101–16106, 1993.

[65] J. P. Perdew and Y. Wang. *Phys. Rev. B*, 33(8800), 1986.

[66] A. D. Becke. *J. Chem. Phys.*, 98(1372), 1993.

[67] A. D. Becke. *Phys. Rev. A*, 38(3098), 1988.

[68] J. Wang, A. Hallmark, D. S. Marshall, W. J. Ooms, P. Ordejón, J Junquera, D. Sánchez-Portal, E. Artacho, and J. M. Soler. *Phys. Rev. B (Rapid Comm.)*, 60(4968), 1999.

[69] D. Sánchez-Portal, E. Artacho, J. M. Soler, A. Rubio, and P. Ordejón. *Phys. Rev. B*, 59(12678), 1999.

[70] M. S. C. Mazzoni and H. Chacham. *Appl. Phys. Lett.*, 76:1561–1563, 2000.

[71] E. Burgos, E. Halac, R. Weht, H. Bonadeo, E. Artacho, and P. Ordejón. *Phys. Rev. Lett.*, 85(2328), 2000.

[72] J. K. Tomfohr and O. F. Sankey. *Phys. Stat. Sol. B*, 233:59–69, 2002.

[73] X. D. Cui, A. Primak, X. Zarate, J. Tomfohr, O. F. Sankey, A. L. Moore, T. A. Moore, D. Gust, L. A. Nagahara, and S. M. Lindsay. *J. Phys. Chem. B*, 106:8609–8614, 2002.

[74] M. D. Towler and M. Causà. *Computer Phys. Comms.*, 98:181–205, 1996.

[75] F. Q. Ban, K. N. Rankin, J. W. Gauld, and R. J. Boyd. *Theor. Chem. Acc.*, 108(1):1–11, 2002.

[76] J. G. Kushmerick, D. B. Holt, J. C. Yang, J. Naciri, M. H. Moore, and R. Shashidhar. *Phys. Rev. Lett.*, 89(8):086802, 2002.

[77] M. Kondo T. Tada and K. Yoshiyawa. *Chem. Phys. Chem*, 4(11), 2003.

[78] E. G. Emberly and G. Kirczenow. *Phys. Rev. B*, 58(16):10911–10920, 1998.

[79] J.C. Slater and G.F. Koster. *Phys. Rev.*, 94(6):1498–1524, 1954.

[80] F. Bloch. *Z. Physik*, 52:555, 1928.

[81] R. Hoffmann. *J. Chem. Phys.*, 39(6):1397–1412, 1963.

[82] W. A. Harrison. Electronic structure and the properties of solids. *Freeman*, San Francisco, 1980.

[83] M. Nishida. *Phys. Rev. B*, 58(11):7103–7112, 1998.

[84] A. Ouammou, M. Mouallembahout, O. Pena, J. F. Halet, J. Y. Saillard, and C. Carel. *J. Solid State Chem.*, 117(1):73–79, 1995.

[85] X. A. Chen, M. Onoda, H. Wada, A. Sato, H. Nozaki, and R. Herbst-Irmer. *J. Solid State Chem.*, 145(1):204–211, 1999.

[86] M. S. Bailey and F. J. DiSalvo. *J. Alloy. Compd.*, 353(1–2):146–152, 2003.

[87] J. F. Britten, O. P. Clements, R. C. Haddon, M. E. Itkis, K. M. Matkovich, and R. T. Oakley. *Chem Mater*, 16(8):1564–1572, 2004.

[88] M. P. Samanta, W. Tian, and S. Datta. *Phys. Rev. B*, 53(12):7626–7629, 1996.

[89] S. Datta, W. Tian, S. Hong, R Reifenberger, J. I. Henderson, and C. P. Kubiak. *Phys. Rev. Lett.*, 79:2530–2533, 1997.

[90] M. Magoga and C Joachim. *Phys. Rev. B*, 56:4722–4729, 1997.

[91] M. Olson, Y. Mao, T. Windus, M. Kemp, M. Ratner, N Léon, and V. Mujica. *J. Phys. Chem. B*, 102:941–947, 1998.

[92] S. N. Yaliraki, M. Kemp, and M. A. Ratner. *J. Am. Chem. Soc.*, 121:3428–3434, 1999.

[93] J. Cerda and F. Soria. *Phys. Rev. B*, 61(12):7965–7971, 2000.

[94] R. H. Dicke and J. P. Wittke. Introduction to quantum mechanics. *Addison-Wesley*, page 148, 1963.

[95] D.M. Medvedev. *J. Am. Chem.*, 122(28):6571–6582, 2000.

[96] R.R. Sharma. *Phys. Rev. A*, 13(2):3682–3693, 1976.

[97] H. Ibach and H. Lüth. Solid-state physics. *Springer-Verlag*, Berlin:168–171, 1993.

[98] J. Callaway. Quantum theory of the solid state. *Academic Press*, (San Diego), 1991.

[99] J. E. Inglesfield, S. Crampin, and H. Ishida. *Phys. Rev. B*, (in press), 2005.

[100] M. Büttiker. Electronic properties of multilayers and low-dimensional semiconductors structures. *Plenum Press, New York*, 1990.

[101] A. Levi Yeyati and M. Büttiker. *Phys. Rev. B*, 52(20):14360–14363, 1995.

[102] R. Landauer. *Phys. Lett.*, 85 A(91), 1981.

[103] Y. Imry in Directions of Condensed Matter Physics. *World Scientific, Singapore*, 1:101, 1986.

[104] Yu. V. Sharvin. *Soviet Phys. JETP*, 21(655), 1965.

[105] S. Datta. Electronic Transport in Mesoscopic Systems. *Camb. Univ. Press*, (Cambridge), 1995.

[106] S. Frank, P. Poncharal, Z. L. Wang, and W. A. de Heer. *Science*, 280(1744), 1998.

[107] D. Porath, G. Cuniberti, and R. Di Felice. *Topics in Current Chemistry*, 237:183–228, 2004.

[108] E. Megan, E. Nunez, and J. K. Barton. *Curr. Opin. Chem. Biol.*, 4:199–206, 2000.

[109] H. W. Fink. *Cellular and Molecular Life Sciences*, 58(1), 1999.

[110] M. Ratner. *Nature*, 397:480–481, 1999.

[111] O. F. Sankey J. P. Lewis, P. Ordejon. *Phys. Rev. B*, 55(11):6880–6887, 1997.

[112] S M. Hjort and Stafström. *Phys. Rev. Lett.*, 87(21):228101, 2001.

[113] F. L. Gervasio, P. Carloni, and M. Parrinello. *Phys. Rev. Lett.*, 89(10):108102, 2002.

[114] Ch. Adessi, S. Walch, and M. P. Anantram. *Phys. Rev. B*, 67(081405), 2003.

[115] M. Zwolak and M. Di Ventra. *Appl. Phys. Lett.*, 81(5):925–927, 2002.

[116] Z. G. Yu and X. Song. *Phys. Rev. Lett.*, 86(26):6018–6021, 2000.

[117] S. Roche. *Phys. Rev. Lett.*, 91(10), 2003.

[118] W. Zhang and S. E. Ulloa. *Microelectronics Journal*, 35:23–26, 2004.

[119] J. E. Inglesfield. *Journal of Physics C-Solid State Physics*, 14(26):3795–3806, 1981.

[120] P. Ordejón, D. A. Drabold, M. P. Grumbach, and R. M. Martin. *Phys. Rev. B*, 48(19):14646–14649, 1993.

[121] X. P. Li, R. W. Nunes, and D. Vanderbilt. *Phys. Rev. B*, 47(16):10891–10894, 1993.

[122] S. Crampin, J. B. A. N. Van Hoof, and M. Nekovee. *J. Phys.: Condens Matter*, 4(1475), 1992.

[123] M. P. Anantram and T. R. Govindan. *Phys. Rev. B*, 58(8), 1998.

[124] J. H. Xu and A. J. Freeman. *Phys. Rev. B*, 40:11927, 1989.

[125] D. Wortmann, H. Ishida, and S. Blügel. *Phys. Rev. B*, 66(075113), 2002.

[126] D. S. Fisher and P. A. Lee. *Phys. Rev. B*, 23(12):6851–6854, 1983.

[127] M. Brandbyge, N. Kobayashi, and M. Tsukada. *Phys. Rev. B*, 60(24):17064–17070, 1999.

[128] A. Troisi and M. A. Ratner. *J. Chem. Phys.*, 118(13):6072–6082, 2003.

[129] N. Agraït, A. L. Yeyati, and J. M. van Ruitenbeek. *Physics Reports*, 377(81), 2003.

[130] P. S. Damle, A. W. Ghosh, and S. Datta. *Phys. Rev. B*, 64:201403, 2001.

[131] R. C. Evans. An introduction to crystal chemistry. *Cambridge university Press*, 1964.

[132] I. Merrick. *Private communication.*

[133] Y. Ye and L. Shen. *J. Comp. Chem.*, 21(12):1109–1117, 2000.

[134] Y. J. Ye and Y. Jiang. *Intl. J. Qu. Chem.*, 78(2):112–130, 2000.

[135] G. A. Leonard and W. A. Hunter. *J Mol Biol*, 234:198, 1993.

[136] ViewerLite 4.2. *Copyright 2001 by Accelrys Inc.*

[137] HyperChem 7.1. *Copyright 2002 Hypercube Inc.*

[138] E. J. Sorin and V. S. Pande. *Biophysical Journal*, 88:2472–2493, 2005.

[139] J. W. Ponder and D. A. Case. *Adv. Prot. Chem.*, 66:27–85, 2003.

[140] T. E. Cheatham and M. A. Young. *Biopolymers*, 56:232–256, 2001.

[141] J. B. Pendry, A. Mackinnon, and P. J. Roberts. *Proc. Roy. Soc. Lond. A*, 437(1899):67–83, 1992.

[142] C. Dekker and M. Ratner. *Physics World*, Aug:29, 2001.

[143] J. S. Hwang, S. W. Hwang, and D. Ahn. *Superlattices and Microstructures*, 34:433–438, 2003.

[144] J. Yi. *Phys. Rev. B*, 68:193103, 2003.

[145] L. Solymar and D. Walsh. Electrical properties of materials. *Oxford University Press*, page 126, 1998.

[146] M. A. Kastner. *Ann. Phys.*, 9:885–894, 2000.

[147] Th. Martin and R. Landauer. *Phys. Rev. A*, 47(3):2023–2030, 1993.

[148] C. Wan, T. Fiebig, S. O. Kelley, C. R. Treadway, J. K. Barton, and A. H. Zewail. *Proc. Natl. Acad. Sci.*, 96:6014–6019, 1999.

[149] E. I. Kats and V. V. Lebedev. *JETP Lett.*, 75(1):37–40, 2002.

[150] A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, and D. Bensimon. *Science*, 265:2096–2098, 1994.

[151] D. C. G. Klein, L. Gurevich, J. W. Janssen, L. P. Kouwenhoven, J. D. Carbeck, and L. L. Sohn. *Appl. Phys. Lett.*, 78(16):2396–2398, 2001.

[152] M. H. F. Wilkins, R. G. Gosling, and W. E. Seeds. *Nature*, 167:759, 1951.

[153] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J-L. Viovy, D. Chatenay, and F. Caron. *Science*, 271:792–794, 1996.

[154] S. B. Smith, Y. Cui, and C. Bustamante. *Science*, 271:795–799, 196.

[155] M. F. Ashby and D. R. H. Jones. Engineering materials 1. *Butterworth Heinemann*, page 34, 2000.

[156] M. Fixman and J. Kovac. *J. Chem. Phys.*, 58:1564–1568, 1973.

[157] C. Bustamante, J. F. Marko, E. D. Sigga, and S. Smith. *Science*, 265:1599–1600, 1994.

[158] C. Bouchiat, M. D. Wang, S. M. Block, J-F Allemand, and V. Croquette. *Biophysical Journal*, 76:409–413, 1999.

[159] S. B. Smith, L. Finzi, and C. Bustamante. *Science*, 258:1122–1126, 1992.

[160] C. Storm and P. C. Nelson. *Phys. Rev. E*, 67(051906), 2003.

[161] A. Lebrun and R. Lavery. *Curr. Opin. Struct. Biol.*, 7:348–354, 1997.

[162] S. Harris. *Private communication.*

[163] T. R. Strick, J. F. Allemand, D. Bensimon, and V. Croquette. *Annu. Rev. Biophys. Biomol. Struct.*, 29:523–543, 2000.

[164] M. Carrion-Vazquez *et al. Proc. Natl. Acad. Sci. USA*, 96(7):3694–3699, 1999.

[165] F. M. Raymo. *Advanced Materials*, 14(6):401, 2002.

[166] J. M. Tour *et al. J. Am. Chem. Soc.*, 117:9529–9534, 1995.

[167] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour. *Science*, 278:252–254, 1997.

[168] J. M. Seminario, A. G. Zacarias, and J. M. Tour. *J. Phys. Chem. A*, 103:7883–7887, 1999.

[169] M. Di Ventra, S. T. Pantelides, and N. D. Lang. *Phys. Rev. Lett.*, 84(5):979–982, 2000.

[170] K. A. Son, H. I. Kim, and J. E. Houston. *Phys. Rev. Lett.*, 86(23):5357–5360, 2001.

[171] X. D. Cui *et al. Science*, 294:571–574, 2001.

[172] Holmlin *et al. J. Am. Chem. Soc.*, 123:5075–5085, 2001.

[173] M. A. Reed, J. Chen, A. M. Rawlett, D. W. Price, and J. M. Tour. *Appl. Phys. Lett.*, 78(23), 2001.