

**Embedding method for conductance of DNA**

O. R. Davies and J. E. Inglesfield

*School of Physics and Astronomy, University of Wales Cardiff, P.O. Box 913, Cardiff, CF24 3YB, United Kingdom*

(Received 20 August 2003; revised manuscript received 29 January 2004; published 26 May 2004)

Using a technique based on embedding in a local-orbital formalism, the electronic structure and electron transmission properties of long biological molecules may be calculated. The electronic structure is found by adding one structural unit at a time to the molecule, and calculating an embedding potential for adding the next structural unit. At present an extended Hückel scheme is used to evaluate the matrix elements. The transmission is also calculated within the embedding scheme, taking the molecule-metal contacts into account. Results for the density of states and transmission are presented for several structures of DNA. The transmission is highly energy dependent, and is also greatly influenced by the orbitals to which contact is made. The implications of these calculations for conductance are discussed.

DOI: 10.1103/PhysRevB.69.195110

PACS number(s): 71.15.-m, 72.80.Le, 87.14.Gg, 73.23.-b

**I. INTRODUCTION**

There is much interest at the moment in the possibility of using biological molecules as electrical conductors, for example, DNA as a molecular wire.<sup>1</sup> The hope is to take advantage of the self-assembling properties of these molecules, which may allow for the production of electronic devices on the nanometer scale. Electron transport is related to electron transfer<sup>2</sup>—this is very important in many biological processes, such as photosynthesis,<sup>3</sup> respiration<sup>4</sup> and oxidative damage.<sup>5</sup> Further work is needed in this field to fully understand the processes involved.

We have chosen to begin our study of large molecules with a treatment of DNA. Eley and Spivey<sup>6</sup> were the first to suggest that DNA could be a conductor more than 40 years ago. However, the debate about the conductive properties of DNA still rages on. Some experiments have shown that DNA is an insulator,<sup>7</sup> while others suggest that it can be a good linear conductor,<sup>8</sup> or a wide-band-gap semiconductor.<sup>1</sup> Some published data even claim that DNA can superconduct.<sup>9</sup> It is clear that further work is required in this area to resolve many differences in both the theoretical and experimental results and, in particular, to study the relationship between structure and conductance.

The DNA double helix is formed of two twisting, hydrophilic, sugar-phosphate backbones. Attached to each sugar unit along the backbone is a hydrophobic base unit (adenine, guanine, cytosine, thymine), which is roughly perpendicular to the axis of the helix. The two strands are bound together by hydrogen bonds between the bases. These bases form complimentary pairs—an adenine base will always bond with a thymine and a cytosine base will always bond with a guanine.

It is generally believed that the pathway for electron transport runs through the bases in the center of the double-helix molecule.<sup>6</sup> It has been suggested that delocalized  $\pi$  orbitals in consecutive bases overlap to form a channel for the movement of electrons through the center of the molecule.<sup>6</sup> Several models have been proposed for the transport of carriers through the base-pair sequence—the dominant mechanisms in the case of DNA are coherent transport via extended molecular orbitals, and thermal hopping. In the

case of coherent transport, the carriers may tunnel through the potential barrier formed by certain base pairs, leading to an exponential decay of transmission with distance.<sup>10</sup> Thermal hopping is an incoherent mechanism in which the electron moves in a series of thermally activated steps between localized orbitals along the molecule. Over large distances the thermal hopping mechanism dominates over coherent tunneling.<sup>11</sup> In this work we have not considered the effect of phonon interactions, therefore we only consider coherent electron transfer processes.

Extensive work has been done in this area, using a variety of approaches. Much of the work involves using model Hamiltonians to describe the molecule, including the work of Cuniberti *et al.*<sup>12</sup> who study the origin of an energy gap in the  $I$ - $V$  characteristics of poly(G)-poly(C) DNA. Roche<sup>13</sup> also uses a model Hamiltonian to study the effect of base sequence, temperature, and length on transmission through the molecule. Hartree-Fock or density-functional theory provides a framework for treating the electronic structure of the molecule self-consistently. Amongst such studies is the work of Adessi *et al.*<sup>14</sup> who consider the effect of environment and structure on DNA conduction, while Hjort and Stafström<sup>15</sup> use a self-consistent approach to study poly(G)-poly(C) DNA, including temperature-induced disorder, and find temperature-dependent semiconducting behavior. A general framework for studying nonequilibrium processes self-consistently has been given by Damle *et al.*<sup>16</sup> and Xue *et al.*<sup>17</sup>

Fundamental to all the models of electron transport is the nature of the one-electron states and the role of transmission through these states. In this paper we present techniques, based on the embedding method,<sup>18</sup> for calculating the one-electron properties of large molecules. The embedding method is used to partition the molecule into convenient subunits, each of which is treated separately, but with their mutual interaction fully included via the embedding potential. The interaction of the molecule with the contacts is also treated in this way, which turns out to be identical to the self-energy methods widely used in electron transfer calculations.<sup>19–21</sup> Embedding is described in Sec. II. Our calculations are done within a nonorthogonal, localized orbital formalism, given in Sec. III, and at this stage we use the

extended Hückel method for determining the matrix elements. The extended Hückel method has been extensively used to describe the electronic structure of large molecules,<sup>12,14,22</sup> but it is relatively straightforward to apply our formalism in a self-consistent method. The methods we use for calculating transmission and conductance are described in Sec. V, and results are given in Sec. VI for several structures of DNA. The efficiency of our method has enabled us to study the effect of many different contact geometries on transmission, and we shall describe the relationship between transmission and electronic structure. Our method is both flexible and efficient, scaling linearly with the size of the molecule, and with only minor changes it can be used to perform electronic structure calculations for many different types of biological molecules.

In this paper we use Hartree atomic units (a.u.), in which 1 a.u. of energy is 27.21 eV and 1 a.u. of length is 0.529 Å.

## II. GROWING MOLECULES BY EMBEDDING

Embedding was developed as a way of solving the Schrödinger equation in large systems that can be subdivided into smaller units.<sup>18</sup> 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 wave functions correctly matched onto the surroundings. The method was originally used to calculate electronic structure of surfaces and interfaces, within a plane-wave basis set. In this paper, we develop a method of embedding using a tight-binding formalism that can be applied to large molecules. Tight-binding embedding uses Dyson's equation and Green's 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 localized orbitals method,<sup>23</sup> and density-matrix methods,<sup>24</sup> 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.*<sup>25</sup> to "grow" large interface systems, adding atomic layer by atomic layer.

A schematic diagram of embedding one Sec. of the molecule onto another is shown in Fig. 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 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 Fig. 1 represent the regions of each section in which orbital overlap occurs, within some cutoff (this is taken to be 8 a.u. in our extended Hückel calculation). Let us assume that we have already calculated the GF for section 1: we can now find an embedding potential for section 2 which contains all the interactions with 1. From the GF for section

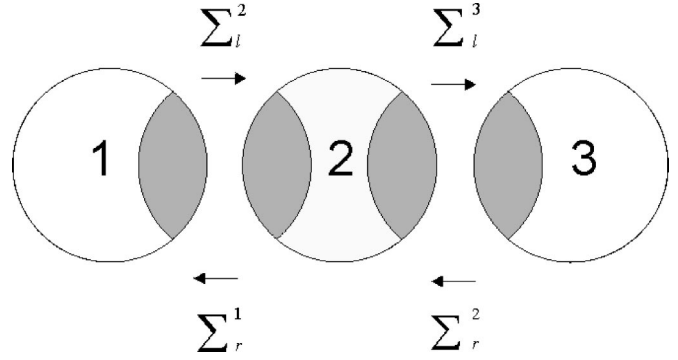


FIG. 1. Schematic representation showing three sections of the DNA molecule, with embedding potentials  $\Sigma_l$  and  $\Sigma_r$  which embed to the left and right.

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

$$(H - ES)G_0 = I, \quad (1)$$

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

$$(H + \delta - ES)G = I. \quad (2)$$

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 Fig. 1, and we treat these as the perturbation  $\delta$ .

If we multiply Eq. (1) with  $G$  and multiply Eq. (2) with  $G_0$ , and then subtract these two equations, we obtain Dyson's equation,

$$G = G_0 - G_0 \delta G. \quad (3)$$

Expanding Eq. (3) gives

$$G = G_0 - G_0 \delta G_0 + G_0 \delta G_0 \delta G. \quad (4)$$

Applying Eq. (4) directly to our problem we obtain

$$\tilde{G}_{22} = G_{22}^0 + G_{22}^0 \delta_{21} G_{11}^0 \delta_{12} \tilde{G}_{22}, \quad (5)$$

where  $\tilde{G}$  is the GF of a section embedded only on the left and  $G_{11}^0$  is the unperturbed GF in section 1. The second term in Eq. (4) goes to zero since  $G_{12}^0$  is zero, because there are no links between regions 1 and 2 in the unperturbed system. Comparing Eq. (5) with Eq. (3) we can see that the series is the same as if we take the perturbation  $\delta$  in Eq. (3) to be  $-\delta_{21} G_{11}^0 \delta_{12}$  acting entirely within the space of region 2. Hence, we take the perturbation to be given by

$$\Sigma_l^2 = -(H_{21} - ES_{21})G_{11}^0(H_{12} - ES_{12}), \quad (6)$$

which is the embedding potential, embedding section 2 onto 1. The overlap matrix  $S$  in Eq. (6) has to be taken into account due to the nonorthogonality of the basis set. This perturbation is then added to the unperturbed Hamiltonian matrix of section 2 to give an effective Hamiltonian,  $H_{\text{eff}}$ , and then the GF for this section, embedded to the left, can be calculated from

$$\tilde{G}_{22} = (H_{\text{eff}} - ES)^{-1}. \quad (7)$$

We note that our result (6) for the embedding potential has been given previously,<sup>14</sup> 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

$$\Sigma_l^3 = -(H_{32} - ES_{32})\tilde{G}_{22}(H_{23} - ES_{23}). \quad (8)$$

Using  $\tilde{G}_{22}$  in Eq. (8) includes the effects of all the previous sections. Once we have finished adding units to the right, we have obtained the GF for each subunit 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

$$\Sigma_r^2 = -(H_{23} - ES_{23})G_{33}^0(H_{32} - ES_{32}) \quad (9)$$

and

$$\Sigma_r^1 = -(H_{12} - ES_{12})\tilde{G}_{22}(H_{21} - ES_{21}), \quad (10)$$

where  $\tilde{G}$  is the GF of a section embedded on the right. Once we have finished adding subunits 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.

This gives us a method of obtaining the GF, which scales linearly with the size of the system. Our method can be applied directly to all molecules that have a linear sequence, with the assumption that only neighboring sections have direct orbital overlap. This method of “growing” molecules can be related to methods of diagonalizing tridiagonal block matrices.<sup>26</sup> In Sec. V 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. We intend to work on including the extra interactions which this folding causes.

### III. THEORETICAL MODEL

The molecules of DNA we investigate are 12 base-pair long, consisting of 760–780 atoms, and we include all of the base and backbone atoms in our model. These fairly large molecules have been studied previously using first-principles methods, such as density-functional theory<sup>27,28</sup> and quantum chemistry techniques.<sup>29</sup> However, for simplicity we use a tight-binding formalism to represent the electronic wave

functions, which 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.

In the tight-binding method the wave functions are expanded in terms of a linear combination of localized orbitals, taken to be atomiclike functions  $\chi_i$ , centered on each atom in the system,

$$\psi_n(\mathbf{r}) = \sum_i c_{n,i} \chi_i(\mathbf{r}). \quad (11)$$

As the atomic orbitals on one site are not orthogonal to the orbitals on other sites, we have a nonunit overlap matrix  $S$  giving the secular equation for the eigenvalues  $E_n$  of the form

$$|H_{ij} - E_n S_{ij}| = 0, \quad (12)$$

where the Hamiltonian matrix elements are given by

$$H_{ij} = \int d\mathbf{r} \chi_i^*(\mathbf{r}) H \chi_j(\mathbf{r}) \quad (13)$$

and the overlap matrix elements are given by

$$S_{ij} = \int d\mathbf{r} \chi_i^*(\mathbf{r}) \chi_j(\mathbf{r}). \quad (14)$$

Such a localized orbital basis set can be used in a first-principles approach, but the method we choose to use is the semiempirical extended Hückel theory, in which  $H_{ij}$  is proportional to  $S_{ij}$ .<sup>30</sup> The basis functions in Eq. (14) are taken as Slater orbitals with parameters taken from the literature,<sup>31</sup> and the Hückel parameters relating  $H$  to  $S$  are taken from the work of Cerda.<sup>32</sup> Extended Hückel was mainly chosen due to its simplicity, which is advantageous when considering large molecules, but it is well established in its application to a wide range of organic molecules.<sup>33–35</sup>

In this paper we present calculations performed on three DNA molecules. The first is a 12 base-pair *B*-DNA molecule, d(CGTAGATCTACG). The spatial coordinates of the atoms were obtained from single-crystal x-ray-diffraction experiments, performed at 15 °C with 2.25 Å resolution.<sup>36</sup> Hydrogen atoms were then added using the computer program VIEWERLITE,<sup>37</sup> as H atoms are not detected by x-ray-diffraction experiments.

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 computer package HYPERCHEM,<sup>38</sup> in which the Amber force field is used to produce the structure with a minimum energy. In the minimization, the Coulomb interactions between the atoms were modeled to fall off as  $1/r^2$  rather than  $1/r$  to simulate the screening effect of a solvent.

The last molecule we investigate is poly(G)-poly(C) DNA, since both experiment<sup>1</sup> and theory<sup>15</sup> 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 minimize its energy with the Amber force field.

#### IV. DENSITY OF STATES

##### A. 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. We now show how it can be calculated in our tight-binding embedding formalism.

The DOS is given by

$$N(E) = \sum_n \delta(E - E_n), \quad (15)$$

where  $n$  runs over all the states in the system. We can relate this to the full GF  $G(\mathbf{r}, \mathbf{r}'; E)$  using

$$N(E) = \frac{1}{\pi} \int d\mathbf{r} \text{Im} G(\mathbf{r}, \mathbf{r}; E + i\epsilon). \quad (16)$$

Expanding  $G$  in terms of the basis functions we have

$$G(\mathbf{r}, \mathbf{r}; E) = \sum_{i,j} G_{ij}(E) \chi_i(\mathbf{r}) \chi_j(\mathbf{r}), \quad (17)$$

and substituting Eq. (17) into Eq. (16) gives the DOS as

$$N(E) = \frac{1}{\pi} \text{Im} \text{Tr}(GS). \quad (18)$$

The trace of this product matrix can be written as

$$\text{Tr}(GS) = \sum_{(n,i),(m,j)} G_{(n,i),(m,j)} S_{(m,j),(n,i)}, \quad (19)$$

where  $n$  and  $m$  label neighboring sections of DNA (as in Fig. 1), with  $i$  and  $j$  labeling the orbitals in  $n$  and  $m$ , respectively. The sum in Eq. (19) 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

$$\begin{aligned} \text{Tr}(GS) = & \sum_n \left( \sum_{i,j;m=n} G_{(n,i),(m,j)} S_{(m,j),(n,i)} \right. \\ & + \sum_{i,j;m=n-1} G_{(n,i),(m,j)} S_{(m,j),(n,i)} \\ & \left. + \sum_{i,j;m=n+1} G_{(n,i),(m,j)} S_{(m,j),(n,i)} \right), \quad (20) \end{aligned}$$

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

We find the first term in Eq. (20) directly from our embedding procedure (section 2). In the second term,  $G_{(n,i),(m,j)}$ , which is the GF between orbitals in one section with those in the preceding section, can be derived using Dyson's equation, and is given by

$$G_{(n,i),(m,j)} = - \sum_k \tilde{G}_{(n,i),(n,k)} \delta_{(n,k)(m,l)} G_{(m,l),(m,j)}, \quad (21)$$

where  $\delta_{(n,k)(m,l)}$  in Eq. (21) is given by

$$\delta_{(n,k)(m,l)} = H_{(n,k)(m,l)} - ES_{(n,k)(m,l)}. \quad (22)$$

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

The third term in Eq. (20) 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 Eq. (20), multiplying the second term by a factor of 2.

In this way, the total DOS of the molecule can be calculated by inverting and diagonalizing the Hamiltonian and overlap matrices for each section, without having to invert the Hamiltonian matrix of the entire molecule.

##### B. Density of states of DNA

We now apply this method to calculate the DOS of the different structures of DNA. In our calculation the overlap between orbitals is cut off beyond 8 a.u., and there is only significant overlap between neighboring sections, due to the localization of the Slater-type orbitals. A small imaginary part of 0.005 a.u. is added to the energy, broadening the  $\delta$  functions that represent the discrete electronic states of the molecule [Eq. (15)]. The results for DNA taken from the x-ray structure are shown in Fig. 2; the DOS of the energy-minimized structure of this molecule is essentially the same, and therefore not given. The DOS of poly(G)-poly(C) DNA is plotted in Fig. 3.

For the x-ray-diffraction DNA, we calculate the energy of the highest occupied molecular orbital (HOMO) to be  $-0.43$  a.u., and the lowest unoccupied molecular orbital (LUMO) to be at  $-0.32$  a.u., giving a band gap of 0.11 a.u. When considering the transport properties of the molecule, the most important states are those on either side of the band gap, as these will dominate conduction through the molecule in the limit of small applied voltages. We find that states near the HOMO and LUMO are all located on atoms in the bases, in agreement with the generally accepted theory that conduction through DNA occurs via the  $\pi$  orbitals in the bases.<sup>14</sup>

Comparing the DOS for the poly(G)-poly(C) DNA molecule shown in Fig. 3 with the results for the x-ray-diffraction DNA with different base pairs, Fig. 2, it can be seen that the results are surprisingly similar, with only minor changes in the fine detail. The band gap remains 0.11 a.u. (3 eV).

Most literature values for the band gap of poly(G)-poly(C) DNA vary between 1.12 eV and 3.2 eV for a variety of methods.<sup>13,15,39,40</sup> We conclude therefore that our value of 3 eV for the band gap agrees very well with published data.

The band gap we calculate for a single GC pair is actually the same as for the whole molecule. This contrasts with the density-functional theory (DFT) calculations of Lewis *et al.*<sup>40</sup> who find a band gap of 3.37 eV for a single GC section, and a narrower band gap of 1.40 eV for a ten-base

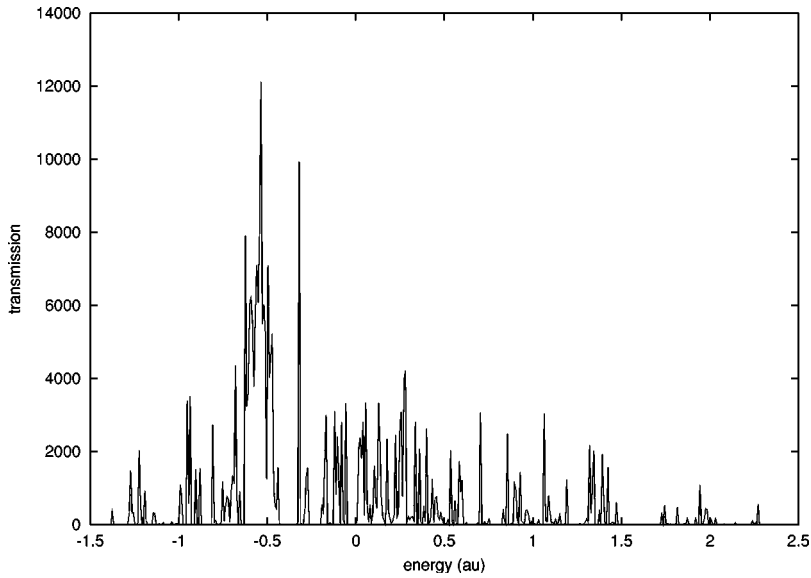


FIG. 2. DOS of mixed base DNA dodecamer obtained from x-ray-diffraction structure. HOMO is at  $-0.43$  a.u. and LUMO at  $-0.32$  a.u.

system. So, although our results compare very well for the single-base case, we do not find the same band gap narrowing when the whole molecule is considered. Lewis *et al.* also report a valence-band width of 1.1 a.u. in good agreement with our result of 0.9 a.u. Their conduction-band width is smaller by a factor of 2, 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.

## V. CONDUCTANCE AND TRANSMISSION

### A. Green's functions and transmission

There has been a great deal of theoretical work performed on calculating the conductance of microscopic systems such

as molecules. This shows that the conductance 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.<sup>41</sup> These channels usually correspond to the incident and transmitted Bloch states at a particular energy. If we consider electrons incident on the molecule from each side, the current in each channel contains velocity and DOS factors which cancel out, giving the net current as

$$I = \frac{e}{h} \int_{E_{F,l}}^{E_{F,r}} T(E) dE, \quad (23)$$

where the integral is between the Fermi energies in the left and right contacts, and  $T$  is the transmission summed over all left and right channels,  $T = \sum_{i,j} T_{ij}$ .<sup>41</sup> In the limit of small

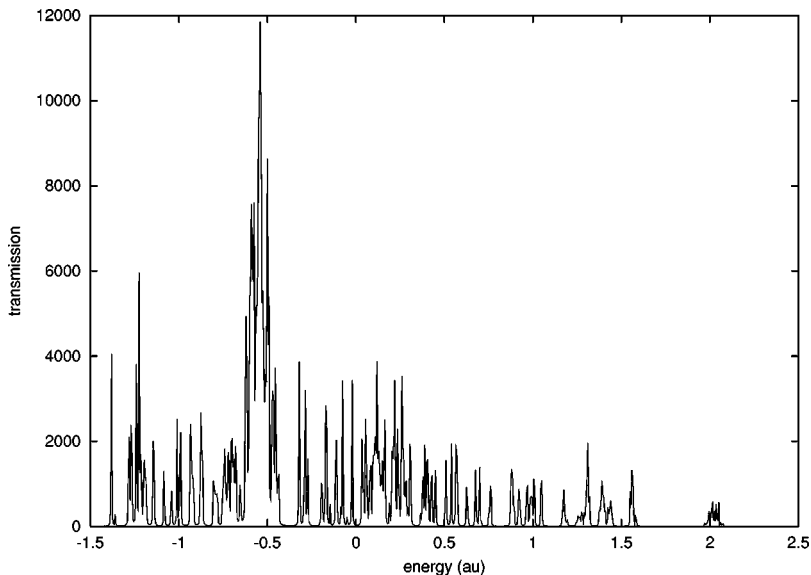


FIG. 3. DOS for poly(G)-poly(C) DNA dodecamer. HOMO is at  $-0.43$  a.u. and LUMO at  $-0.32$  a.u.

voltages  $V$  across the molecule, taking the difference in Fermi functions, and dividing by  $V$ , gives the Landauer-Büttiker formula for the conductance,

$$\Gamma = \frac{e^2}{h} T(E_F). \quad (24)$$

The total transmission is evaluated at the Fermi energy and can be written concisely in terms of the GF between the contacts as<sup>42,43</sup>

$$T = 4 \text{Tr}(\hat{G}_{lr} \text{Im} \Sigma_r \hat{G}_{rl}^* \text{Im} \Sigma_l). \quad (25)$$

Here,  $\hat{G}_{lr}$  is the GF for the molecule connected to the contacts, between the left-hand and right-hand contacting orbitals;  $\Sigma_l$  and  $\Sigma_r$  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,<sup>44,45</sup> though, as we remarked in Sec. II, in this context  $\Sigma$  is usually called the self-energy.<sup>14</sup> 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.<sup>42</sup> It can be used to find, for example, the conductance of an interface between metals. Then  $\hat{G}_{lr}$  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.*<sup>42</sup> to study spin-polarized transmission through a ferromagnetic Co monolayer sandwiched between Cu.

### B. Embedding approach to transmission

In this preliminary application we couple a single orbital on the DNA to a single orbital on the metallic contact. This means that there is only one channel available for transmission at each end of the molecule. This is an approximation to the usual experimental arrangement where the metal-DNA contact extends over several orbitals,<sup>8</sup> but the extension to multi-orbital contact is straightforward. However, we note that single-atom contact experiments are indeed possible, as shown in the work of Agrait *et al.*<sup>46</sup> This makes our assumption of a single contact orbital more plausible.

If we are to calculate the transmission between two orbitals using Eq. (25), we must first find the GF  $\hat{G}_{lr}$  linking these orbitals, which may be located anywhere on the DNA chain. The GF coupling the two orbitals embedded onto the leads,  $\hat{G}_{lr}$ , can be derived from Dyson's equation,

$$\hat{G}_{lr} = G_{lr} - G_{ll} \Sigma_l \hat{G}_{lr} - G_{lr} \Sigma_r \hat{G}_{rr}. \quad (26)$$

Here  $G_{lr}$  is the unperturbed GF linking orbital  $l$  to orbital  $r$ , without the metal contacts.  $\Sigma_l$  and  $\Sigma_r$  are the embedding potentials linking the left and right metal reservoirs to the molecule on these orbitals. However, we do not know the quantity  $\hat{G}_{rr}$ , which is the GF of the right-hand orbital con-

nected to the metal contact. We can, however, write  $\hat{G}_{rr}$  in terms of other quantities that we do know,

$$\hat{G}_{rr} = G_{rr} - G_{rl} \Sigma_l \hat{G}_{lr} - G_{rr} \Sigma_r \hat{G}_{rr}. \quad (27)$$

Rearranging Eq. (27) for  $\hat{G}_{rr}$  and substituting into Eq. (26) we obtain the GF linking the two orbitals as

$$\hat{G}_{lr} = [1 + G_{ll} \Sigma_l - G_{lr} \Sigma_r (1 + G_{rr} \Sigma_r)^{-1} G_{rl} \Sigma_l]^{-1} \times [G_{lr} - G_{lr} \Sigma_r (1 + G_{rr} \Sigma_r)^{-1} G_{rr}]. \quad (28)$$

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.*<sup>12</sup> and Damle *et al.*,<sup>16</sup> though in a somewhat different form. In our model we choose to use two Cu reservoirs to make contact with the DNA molecule. The embedding potential  $\Sigma_l$ , Eq. (6), embedding the molecule onto the left reservoir is given by

$$\Sigma_l = -(H_{l\text{Cu}} - ES_{l\text{Cu}}) G_{\text{CuCu}}^0 (H_{\text{Cu}l} - ES_{\text{Cu}l}), \quad (29)$$

where  $G_{\text{CuCu}}^0$  is the unperturbed GF in the left Cu lead, with a similar expression for  $\Sigma_r$ .

Equation (28) may be used immediately to describe a multiorbital contact between the metal and the molecule, in which case the GF's and embedding potentials are given by matrices. However, for simplicity we consider a single Cu  $s$  orbital, which makes contact with a single orbital on the DNA molecule. This single contact between the metal and molecule allows for only one conduction channel in and out of the molecule. The coupling terms in Eq. (29) are evaluated using the extended Hückel method, assuming realistic Cu-molecule distances,<sup>47</sup> but with an arbitrary angle between the molecule and the Cu surface. To describe the Cu contact we use a full electronic structure calculation for the surface of semi-infinite Cu (001). This uses the embedded linearized augmented plane-wave method,<sup>48</sup> which gives very accurate results for the density of states on the surface atoms.

As we are using only a single Cu  $s$  orbital at this stage of our treatment of the metal-molecule contact, we project the total surface density of states of the Cu onto the orbital. From this, the imaginary part of the GF is given by

$$\text{Im} G_{\text{CuCu}}^0(E) = \pi n_{\text{Cu}}(E), \quad (30)$$

where  $n_{\text{Cu}}$  is the total surface density of states on the Cu  $s$  orbital. The real part of the GF can then be found from the Kramers-Kronig relation, and is given by

$$\text{Re} G_{\text{CuCu}}^0(E) = \int_{-\infty}^{\infty} dE' \frac{n_{\text{Cu}}(E')}{E' - E}. \quad (31)$$

The GF  $G_{\text{CuCu}}^0$  is evaluated at the Fermi energy of Cu, and we use this value to determine  $\Sigma_l$ ,  $\Sigma_r$ ,  $\hat{G}_{lr}$ , and hence the transmission of the molecule over the energy range of the molecular DOS.

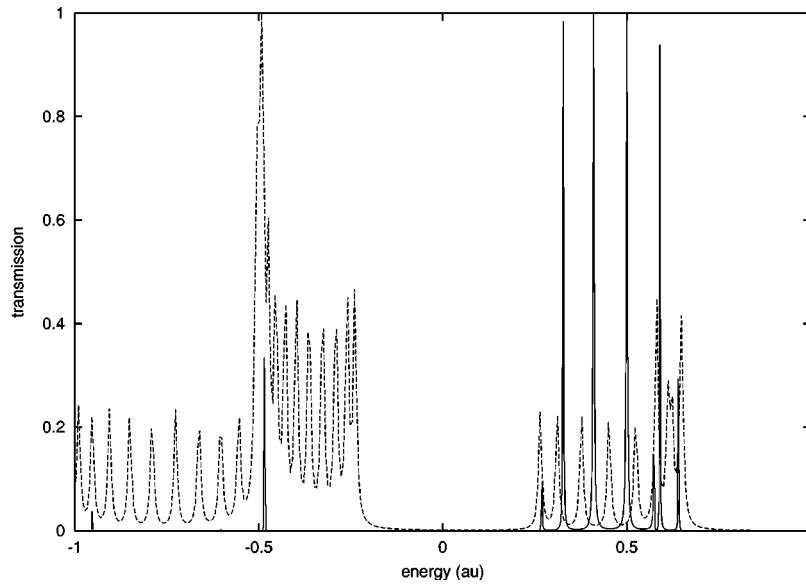


FIG. 4. End-to-end transmission (solid line) and scaled DOS (dashed line) of a chain of 12 carbon atoms.

## VI. TRANSMISSION RESULTS

### A. 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  $1.53 \text{ \AA}$ . 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. We choose this simple system to test our method because we know that the orbitals will overlap well to give good transmission. The results are shown in Fig. 4, along with the DOS of the C chain uncoupled to the Cu contacts. The DOS was calculated with an imaginary part of  $0.005 \text{ a.u.}$  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 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 Fig. 4 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 Fig. 4 that there are fewer transmission peaks than DOS peaks—this happens since transmission is only appreciable at energies that correspond to those wave functions 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 Fig. 4, we see that they do not align. However, in Fig. 5 the same transmission peaks are shown to line up exactly with the total density of states when we include the

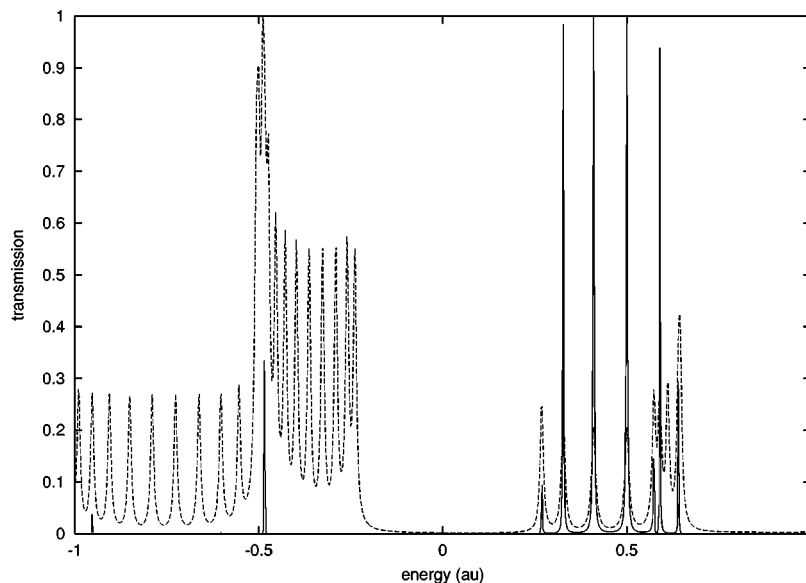


FIG. 5. End-to-end transmission (solid line), and scaled total DOS (dashed line) of a chain of 12 carbon atoms attached to Cu reservoir.

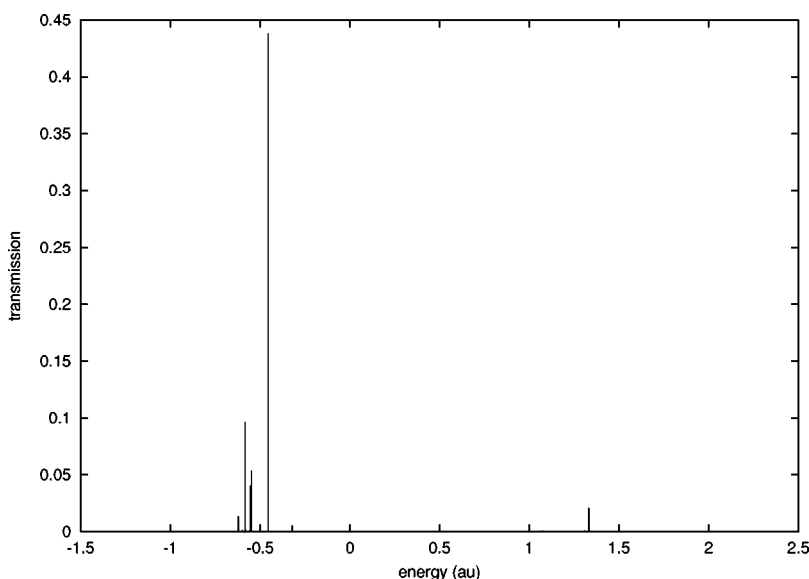


FIG. 6. End-to-end transmission of x-ray-diffraction DNA for varying energy.

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 Eq. (28). Different states are shifted by varying amounts, and it can be seen that the shift is the largest in the middle of the conduction band, near 0.5 a.u. 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.

### B. X-Ray-Diffraction DNA

We now turn from the short model molecule to the more complex DNA double helix. As with the C chain, we are interested in the end-to-end transmission. First, we calculate the transmission between two fixed contacts in the bases, at either end of the molecule, and vary the energy over the energy range of the DOS. The results are shown in Fig. 6, and it can immediately be seen that the transmission is very peaky. 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 Fig. 6. The large transmission peaks around  $-0.5$  a.u. in Fig. 6 lie in the large peak in the DOS below the HOMO. There is a small peak near the LUMO at  $-0.32$  a.u. and another slightly larger one high in the conduction band at  $1.33$  a.u. Theoretical and experimental work by de Pablo *et al.*<sup>49</sup> shows poor conductivity for DNA with a random base-pair sequence, consistent with this work.

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.*<sup>33</sup> and theoretically by Damle *et al.*<sup>16</sup> 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 combinations of orbitals in the two end sections of the molecule.

We apply this to states near the band gap, which are the most important for conduction.<sup>14</sup> In Fig. 7 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  $2.2 \times 10^{-6}$ . The other states near the HOMO all gave extremely low transmissions, in 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 (Fig. 8) (LUMO+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 wave function is well distributed over these orbitals. This contrasts with the transmission shown in Fig. 7 for the HOMO state, where significant transmission only occurs for a few orbital combinations.

To investigate this further we look at the distribution of charge density along the molecule. Figures 9 and 10 show the distribution of charge along the molecule for the HOMO and the LUMO+6, respectively. Comparing the two graphs it can be seen that for the LUMO+6 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 localized state, hence, reducing the probability of electron transport from one end of the molecule to the other. This is clear why there is such an enormous difference in the transmission of the two states (Figs. 7 and 8).

The low transmission in this DNA molecule for nearly all energies is partly a consequence of the potential barrier



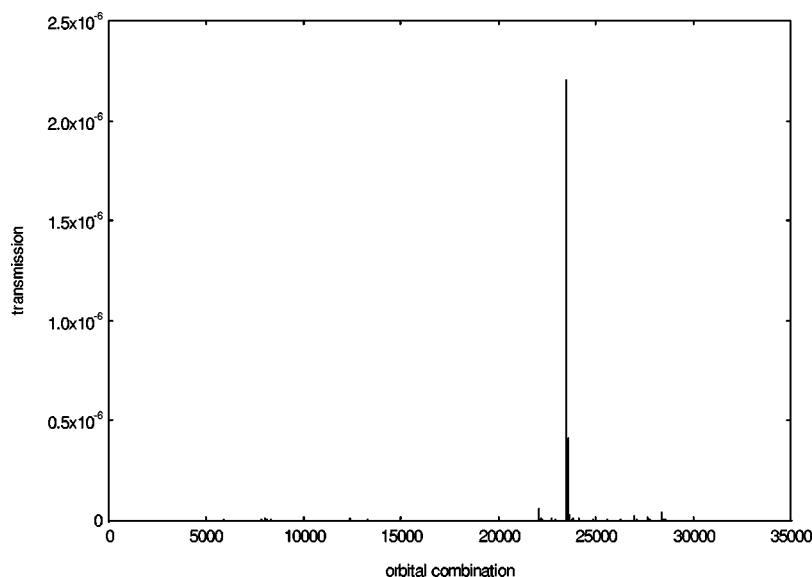


FIG. 7. End-to-end transmission of x-ray-diffraction DNA for all combinations of orbitals in sections 1 and 12 for HOMO.

formed by the AT base pairs,<sup>11</sup> but also of disorder in the molecule. 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. Disorder in one-dimensional systems leads to extremely peaky transmission as a function of energy,<sup>50</sup> consistent with the results of Fig. 6.

We now consider the results from our energy-minimized molecule of DNA. This molecule has the same 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. The transmission of the valence states near the HOMO is still very small, of the order of  $10^{-12}$  for all the end-to-end orbital combinations, much the same as for the x-ray-diffraction DNA. However, there are several more states with large transmission in the conduction band, just above the

LUMO. Figure 11 shows the transmission of the LUMO + 1 state, for which the maximum transmission is 0.64—this is comparable to that of the carbon chain (Fig. 5). This is a consequence of the optimized structure having better orbital overlap than the x-ray-diffraction structure, and being less disordered. However, the transmission still has huge fluctuations when varying the contact orbitals. These great changes in transmission in going from the x-ray-diffraction DNA to the energy-minimized structure show how small changes in structure can dramatically affect transmission.

### C. Energy-minimized poly(G)-poly(C) DNA

The last molecule we investigate is poly(G)-poly(C) DNA. We have minimized its energy using the Amber force field within HYPERCHEM.<sup>38</sup> This ordered system, consisting

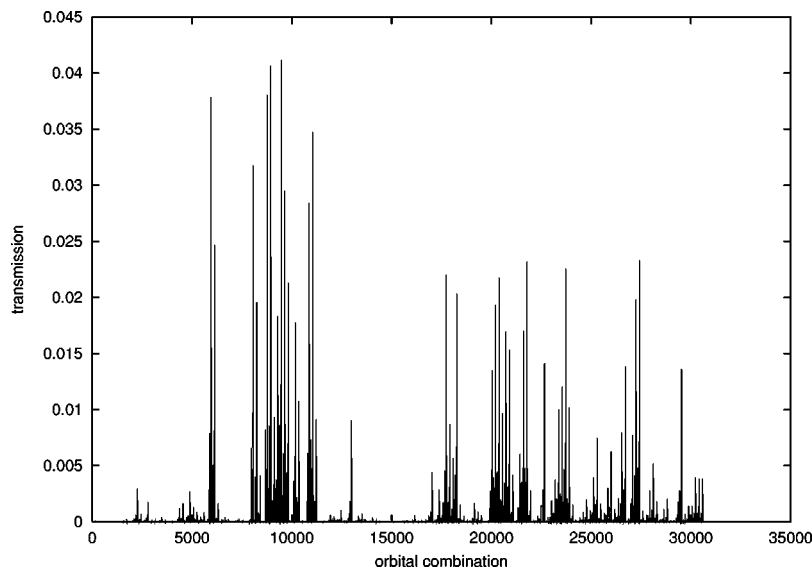


FIG. 8. End-to-end transmission of x-ray-diffraction DNA for all combinations of orbitals in sections 1 and 12 for LUMO+6.

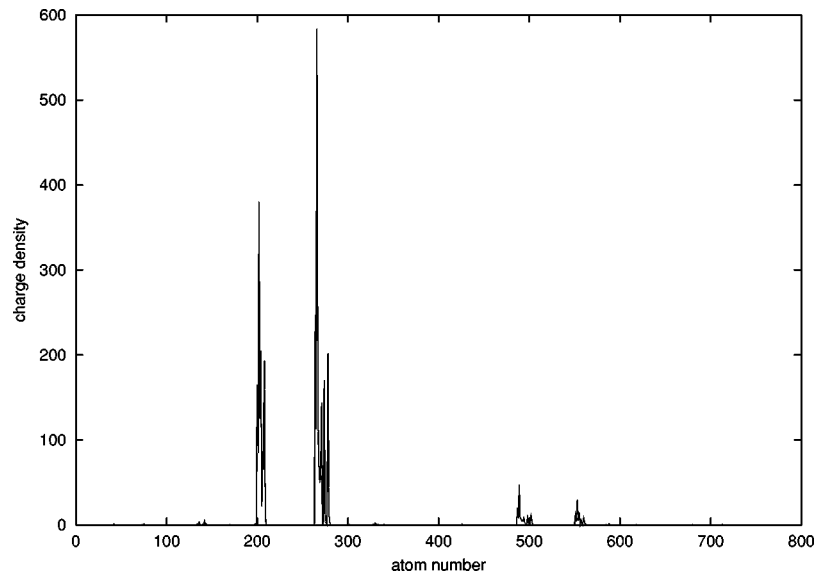


FIG. 9. Distribution of charge density along x-ray-diffraction DNA molecule for HOMO; the atom number describes the position along the chain.

of the same repeated bases, is known to have greatly improved transmission over mixed base-pair DNA.<sup>49</sup>

Fixing the metal-DNA contact orbitals on bases at either end of the molecule, and varying the energy at which the transmission is calculated, gives us the peaky graph shown in Fig. 12. It can be seen that there are a number of peaks with remarkably good transmission, many more than for the x-ray-diffraction DNA (Fig. 6). However, for nearly all energies the transmission still remains effectively zero, though there are groups of states around the HOMO and LUMO with large transmission between the metal contacts at the ends of the molecule.

The results in Fig. 12 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.*,<sup>14</sup> 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. The large fluctuations in transmission as a function of energy result from the structure of our system—although the structure has been determined by energy minimization it is not completely ordered, leading to the fluctuations, characteristic of a one-dimensional disordered systems.<sup>50</sup> Similar large fluctuations are seen in the work of Roche<sup>13</sup> using a model Hamiltonian to show the effect of temperature on transmission through poly(G)-poly(C) DNA. The transmission peaks in our results are very narrow, of the order of  $10^{-6}$  a.u. in width—the peak width comes from the interaction of the DNA with the continuum of states in the electrodes. A recent

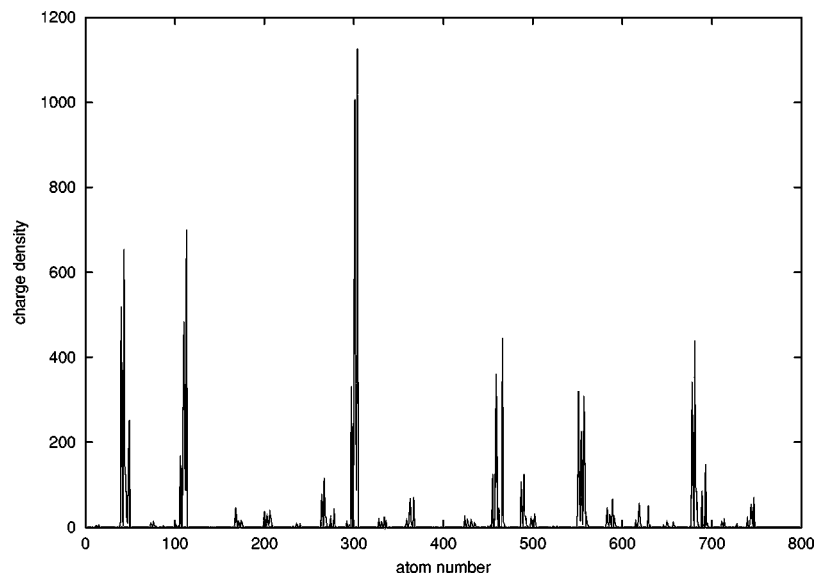


FIG. 10. Distribution of charge density along x-ray-diffraction DNA molecule for LUMO+6; the atom number describes the position along the chain.

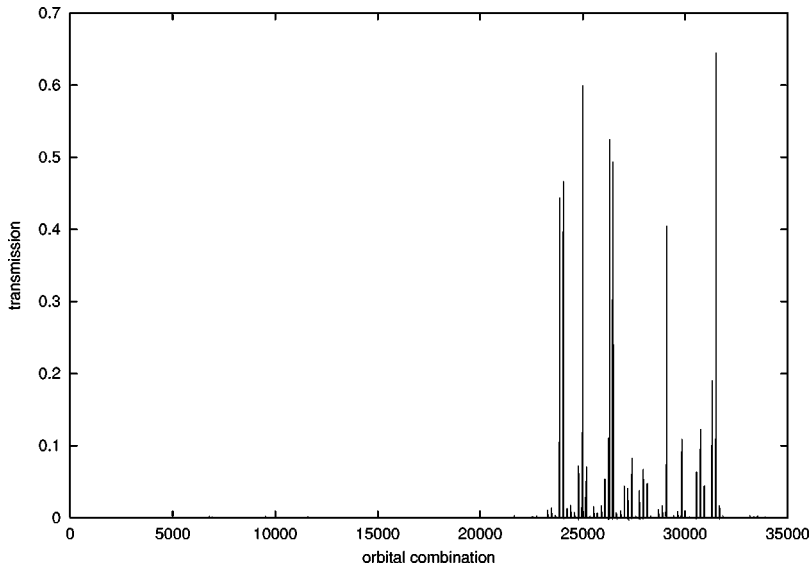


FIG. 11. End-to-end transmission of energy-minimized DNA for all combinations of orbitals in sections 1 and 12 for LUMO+1.

extended Hückel study of  $(AT)_{12}$  DNA (Ref. 34) 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.

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 13 shows the transmission for the HOMO-2 state at energy  $-0.4368$  a.u., with a maximum transmission of 0.96, compared with  $10^{-6}$  for the peak transmission shown in Fig. 7 for the x-ray-diffraction DNA. However, we see for Fig. 13 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 completely different behavior from the mixed base DNA,

and with its much higher transmission for states near both the HOMO and LUMO, will be a much better electrical conductor. This is in agreement with previous theoretical work<sup>15</sup> and experiment.<sup>51</sup>

## VII. CONCLUSIONS

The embedding method we have used in this work efficiently produces the electronic structure of DNA molecules within a tight-binding scheme. We have also shown how transmission and conductance can be calculated within the same scheme. Our results show that the transmission through the DNA molecules has a very peaky energy dependence, and is also extremely sensitive to the choice of contact orbital. This energy dependence means that experiments which measure the conductance of DNA will show a non-uniform

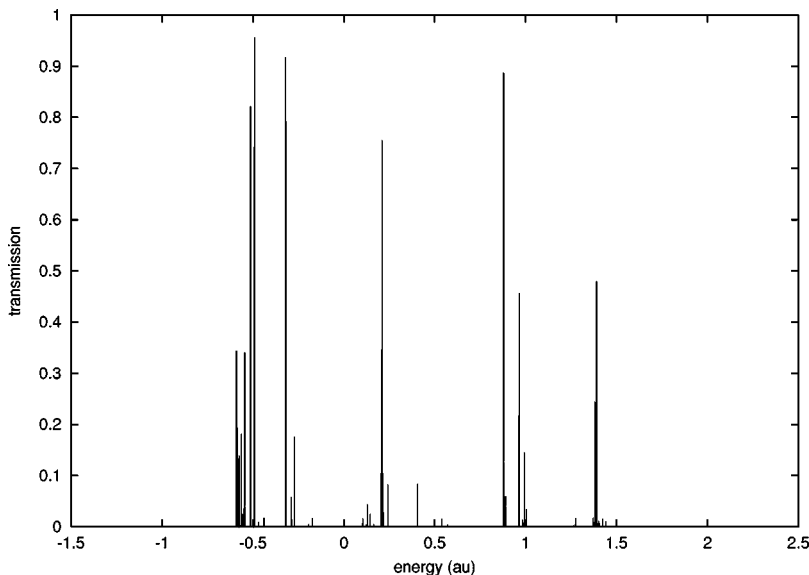


FIG. 12. End-to-end transmission of poly(G)-poly(C) DNA for varying energy.

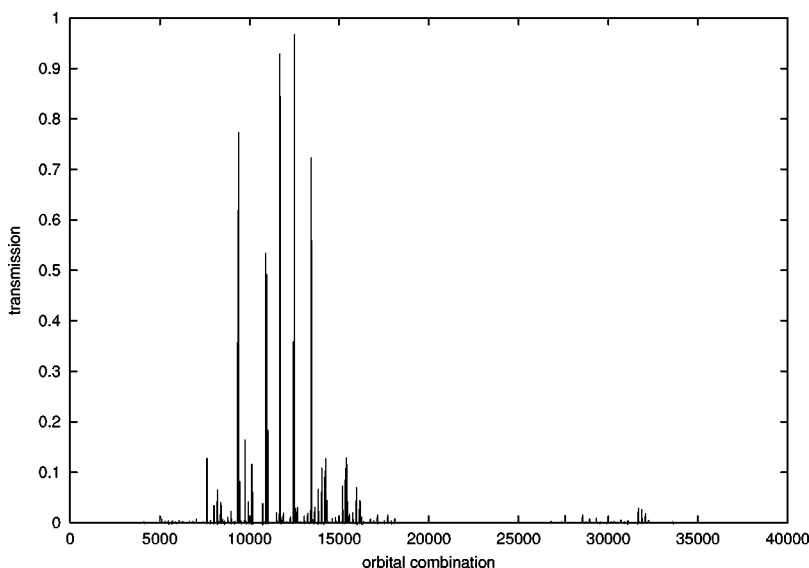


FIG. 13. End-to-end transmission of poly(G)-poly(C) DNA for all combinations of orbitals in sections 1 and 12 for HOMO-2.

increase in current, as the potential across the molecule is increased, and more transmission states are included in the conduction process [Eq. (23)]. This indeed seems to be the case.<sup>15</sup>

The transmission increases when we energy-minimize the system by increasing the overlap between orbitals and reducing disorder, allowing electrons to move more freely through the molecule. We also find poly(G)-poly(C) to be a much better electrical conductor than mixed base DNA, with maximum transmission similar to that of a short carbon chain. The importance of base sequence on conduction has been reported by Hjort and Stafström,<sup>15</sup> and an explanation for the improved conduction of poly(G)-poly(C) has been given by Dekker and Ratner<sup>11</sup>—an AT base pair acts as a potential barrier, reducing conduction.

These calculations correspond to coherent transport at  $T = 0$  K, with no dynamic fluctuations of the molecule and neglecting inelastic and vibronic effects in the transmission.<sup>45</sup> The validity of this approach will depend on the dwell time of the charge carriers in the molecule, com-

pared with the typical vibrational frequencies. The width of our transmission peaks, coming from the coupling of the molecule to the continuum of states in the electrodes, corresponds via the uncertainty principle to  $\Delta t = 6 \times 10^{-12}$  s; whether this is relevant to the dwell time for comparing with the period of the important twist mode of  $10^{-11} - 10^{-12}$  s (Ref. 52) is the subject of debate.<sup>53</sup>

In the future we will be applying this method to investigate the effect of order and sequence more thoroughly on the conductance of DNA. We intend to replace the extended Hückel theory with the more rigorous treatment of self-consistent DFT in a local basis set. After further developments in methodology we will also tackle different biological molecules, such as cytochrome and oligo(phenylene ethynylene).

#### ACKNOWLEDGMENTS

We would like to thank H. C. Loebel, I. Merrick, and J. E. Macdonald for their contribution to this work. O.R.D. acknowledges EPSRC funding.

<sup>1</sup>D. Porath, A. Bezryadin, S. de Vries, and C. Dekker, *Nature* (London) **403**, 635 (2000).

<sup>2</sup>H.W. Fink, *CMLS* **58**, 1 (2001).

<sup>3</sup>G.R. Fleming, J.C. Martin, and J. Breton, *Nature* (London) **333**, 6169 (1988).

<sup>4</sup>G. Schäfer, W. Purschke, and C. Schmidt, *FEMS Microbiol. Rev.* **18**, 173 (1996).

<sup>5</sup>M.E. Nunez, D.B. Hall, and J.K. Barton, *Chem. Biol.* **6**, 85 (1999).

<sup>6</sup>D.D. Eley and D.I. Spivey, *Trans. Faraday Soc.* **58**, 411 (1962).

<sup>7</sup>E. Braun, Y. Elchen, U. Sivan, and G. Ben-Yoseph, *Nature* (London) **391**, 775 (1998).

<sup>8</sup>H.W. Fink and C. Schönenberger, *Nature* (London) **398**, 407 (1999).

<sup>9</sup>A.Y. Kasumov *et al.*, *Science* **291**, 5502 (2001).

<sup>10</sup>F.D. Lewis, J. Liu, W. Weigel, W. Rettig, I.V. Kurnikov, and D.N. Beratan, *PNAS* **99**, 20 (2002).

<sup>11</sup>C. Dekker and M. Ratner, *Phys. World* **14**, 8 (2001).

<sup>12</sup>G. Cuniberti, L. Craco, D. Porath, and C. Dekker, *Phys. Rev. B* **65**, 241314 (2002).

<sup>13</sup>S. Roche, *Phys. Rev. Lett.* **91**, 108101 (2003).

<sup>14</sup>Ch. Adessi, S. Walch, and M.P. Anantram, *Phys. Rev. B* **67**, 081405 (2003).

<sup>15</sup>M. Hjort and S. Stafström, *Phys. Rev. Lett.* **87**, 228101 (2001).

<sup>16</sup>P.S. Damle, A.W. Ghosh, and S. Datta, *Phys. Rev. B* **64**, 201403 (2001).

<sup>17</sup>Y. Xue, S. Datta, and M.A. Ratner, *Chem. Phys.* **281**, 151 (2002).

<sup>18</sup>J.E. Inglesfield, *J. Phys. C* **14**, 26 (1981).

- <sup>19</sup>P.S. Damle, A.W. Ghosh, and S. Datta, Chem. Phys. **281**, 171 (2002).
- <sup>20</sup>G. Cuniberti, F. Großmann, and R. Gutiérrez, Adv. Solid State Phys. **42**, 133 (2002).
- <sup>21</sup>K. Stokbro, J. Taylor, M. Brandbyge, J.L. Mozos, and P. Ordejón, Comput. Mater. Sci. **27**, 151 (2003).
- <sup>22</sup>A. Di Carlo, Semicond. Sci. Technol. **18**, R1 (2003).
- <sup>23</sup>P. Ordejón, D.A. Drabold, M.P. Grumbach, and R.M. Martin, Phys. Rev. B **48**, 14 646 (1993).
- <sup>24</sup>X.-P. Li, R.W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10 891 (1993).
- <sup>25</sup>S. Crampin, J.B.A.N. Van Hoof, and M. Nekovee, J. Phys.: Condens. Matter **4**, 1475 (1992).
- <sup>26</sup>M.P. Anantram and T.R. Govindan, Phys. Rev. B **58**, 8 (1998).
- <sup>27</sup>F.L. Gervasio, P. Carloni, and M. Parrinello, Phys. Rev. Lett. **89**, 108102 (2002).
- <sup>28</sup>Y. Ye and L. Shen, J. Comput. Chem. **21**, 12 (2000).
- <sup>29</sup>Y.J. Ye and Y. Jiang, Int. J. Quantum Chem. **78**, 2 (2000).
- <sup>30</sup>R. Hoffmann, J. Chem. Phys. **39**, 6 (1963).
- <sup>31</sup>R.H. Summerville and R. Hoffmann, J. Am. Chem. Soc. **98**, 23 (1976).
- <sup>32</sup>J. Cerda and F. Soria, Phys. Rev. B **61**, 12 (2000).
- <sup>33</sup>J.G. Kushmerick, D.B. Holt, J.C. Yang, J. Naciri, M.H. Moore, and R. Shashidhar, Phys. Rev. Lett. **89**, 086802 (2002).
- <sup>34</sup>T. Tada, M. Kondo, and K. Yoshiyawa, ChemPhysChem **4**, 11 (2003).
- <sup>35</sup>E.G. Emberly and G. Kirczenow, Phys. Rev. B **58**, 16 (1998).
- <sup>36</sup>G.A. Leonard and W.A. Hunter, J. Mol. Biol. **234**, 198 (1993).
- <sup>37</sup>ViewerLite 4.2, Copyright 2001 by Accelrys Inc.
- <sup>38</sup>HyperChem 7.1, Copyright 2002 Hypercube, Inc.
- <sup>39</sup>M. Zwolak and M. Di Ventra, Appl. Phys. Lett. **81**, 925 (2002).
- <sup>40</sup>J.P. Lewis, P. Ordejón, and O.F. Sankey, Phys. Rev. B **55**, 11 (1997).
- <sup>41</sup>M. Büttiker, *Electronic Properties of Multilayers and Low-Dimensional Semiconductors Structures* (Plenum Press, New York, 1990).
- <sup>42</sup>D. Wortmann, H. Ishida, and S. Blügel, Phys. Rev. B **66**, 075113 (2002).
- <sup>43</sup>D.S. Fisher and P.A. Lee, Phys. Rev. B **23**, 6851 (1981).
- <sup>44</sup>M. Brandbyge, N. Kobayashi, and M. Tsukada, Phys. Rev. B **60**, 17 064 (1999).
- <sup>45</sup>A. Troisi and M.A. Ratner, J. Chem. Phys. **118**, 6072 (2003).
- <sup>46</sup>N. Agrait, A.L. Yeyati, and J.M. van Ruitenbeek, Phys. Rep. **377**, 81 (2003).
- <sup>47</sup>R.C. Evans, *An Introduction to Crystal Chemistry* (Cambridge University Press, Cambridge, 1964).
- <sup>48</sup>I. Merrick (private communication).
- <sup>49</sup>P.J. de Pablo, F. Moreno-Herrero, J. Colchero, J.G. Herrero, A.M. Baró, P. Ordejón, J.M. Soler, and E. Artacho, Phys. Rev. Lett. **85**, 4992 (2000).
- <sup>50</sup>J.B. Pendry, A. Mackinnon, and P.J. Roberts, Proc. R. Soc. London, Ser. A **437**, 1899 (1992).
- <sup>51</sup>L. Cai, H. Tabata, and T. Kawai, Appl. Phys. Lett. **77**, 19 (2000).
- <sup>52</sup>E.I. Kats and V.V. Lebedev, JETP Lett. **75**, 1 (2002).
- <sup>53</sup>H.M. Nussenzveig, Phys. Rev. A **62**, 042107 (2000).