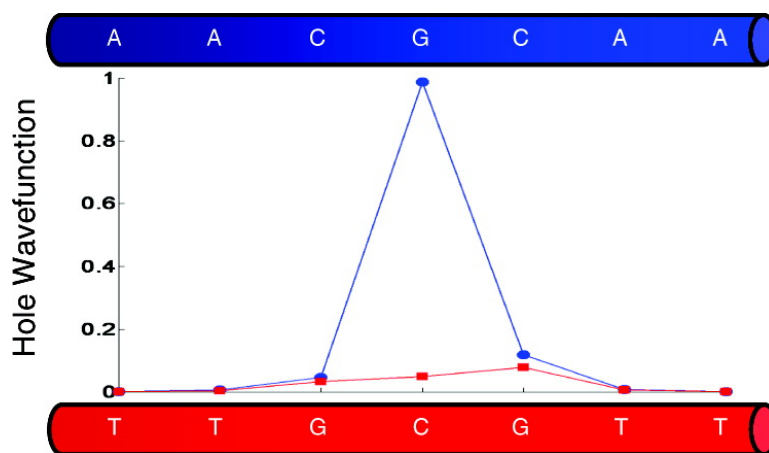


Duplex Polarons in DNA

Esther M. Conwell, Steven M. Bloch, Patrick M. McLaughlin, and Denis M. Basko

J. Am. Chem. Soc., **2007**, 129 (29), 9175-9181 • DOI: 10.1021/ja0691472 • Publication Date (Web): 22 June 2007

Downloaded from <http://pubs.acs.org> on February 16, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)

Duplex Polarons in DNA

Esther M. Conwell,* Steven M. Bloch, Patrick M. McLaughlin, and Denis M. Basko

Contribution from the Departments of Chemistry and Physics, University of Rochester, Rochester, New York 14627, and Department of Physics, Columbia University, New York, New York

Received December 20, 2006; E-mail: conwell@chem.rochester.edu

Abstract: In earlier work we calculated the wavefunction and energy of the solvated polaron in DNA with a simple model in which the charge was taken to be on a single chain of bases at the center of the double helix. To better approximate the actual situation, we have now extended the calculations to the case in which the charge is distributed on two chains of bases, complementary to each other, one on each side of the center. The binding energy of the resulting polaron is somewhat larger than that obtained for the single-chain polaron, the result of each chain of the polaron being closer to some of the polarization charge it induces. Carrying out the calculations for a number of different sequences, we find that the polaron wavefunction is predominantly on one of the two chains, this usually being the one on which the charge was originally placed, despite the availability of lower energy sites on the other chain. This finding is in agreement with recent experiments of Schuster's group (Joy, A.; Ghosh, A. K.; Schuster, G. B. *J. Am. Chem. Soc.* **2006**, *128*, 5346–5347). Thus, in contradiction to the ideas of many researchers, there is no transport in which a hole zigzags from one chain to the other, as has been suggested for a sequence of guanines and cytosines (GCGCGC....), for example.

Introduction

In the earliest discussions of the behavior of an excess electron or hole on a DNA chain it was assumed that it became a polaron by inducing nonuniform spacing of the bases on which it resided,² as is the case for conducting polymers. When the effect on the excess electron or hole of water and ions surrounding the DNA was considered, it was realized that the polarization of its surroundings by the charge made a much larger contribution to the binding energy of the polaron than possible changes in base spacing.³ Although displacements of the bases in the polaron might still occur, their effect should be negligible.

For some years the prevailing picture for an excess electron or hole on DNA had been (and still is in some of the current literature) that it is localized on a single base and moves by tunneling or hopping between bases. This picture is based on experiments of Giese and collaborators, in which holes were introduced onto a guanine (G) followed by a series of adenines (As).⁴ The holes were found to decrease rapidly in number as they traversed the first three adenines, but the survivors were

able to go through many more adenines with little further decrease.⁵ Giese's explanation of these results was that the holes tunneled through the first three As, then became charge carriers owing to endothermic oxidation of As by adjacent G radical cations, and moved further by hopping between adjacent As.⁵ This picture was first seriously questioned by Bixon and Jortner, who found that it was not possible to explain both the steep decrease in the number of holes as they traversed the initial three As and the very slow decrease as they traversed further As.⁶ The rapid decrease found by Giese et al. of the number of holes penetrating successive As up to four should be attributed to back-recombination. We have suggested that the much slower rate of decrease in the number of holes as they traverse further As beyond three is due to the formation of a polaron on the first four As; this lowers the hole energy sufficiently to greatly diminish back-recombination as the polaron travels to further sites.^{7,8}

The possibility of a polaron being created was also suggested by Kendrick and Giese to explain later experiments in which they injected a hole directly onto an A flanked by two As on one side and eight As on the other side, with Cs surrounding the As.⁹ They found that detectors (GGGs) placed at the two ends of the chain complementary to that containing the As received approximately equal numbers of holes and suggested

(1) Joy, A.; Ghosh, A. K.; Schuster, G. B. *J. Am. Chem. Soc.* **2006**, *128*, 5346–5347.

(2) (a) Ly, D.; Sani, L.; Schuster, G. B. *J. Am. Chem. Soc.* **1999**, *121*, 9400–9410. (b) Henderson, P. T.; Jones, D.; Hampikian, G.; Kan, Y.; Schuster, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8353–8358. (c) Conwell, E. M.; Rakhmanova, S. V. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 4556–4560.

(3) Basko, D. M.; Conwell, E. M. *Phys. Rev. Lett.* **2002**, *88*, 098102.

(4) (a) Meggers, E.; Michel-Beyerle, M. E.; Giese, B. *J. Am. Chem. Soc.* **1998**, *120*, 12950–12955. (b) Giese, B.; Wessely, S.; Spormann, M.; Lindemann, U.; Meggers, E.; Michel-Beyerle, M. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 996–998. Giese, B. *Acc. Chem. Res.* **2000**, *33*, 631–636. (c) Bixon, M.; Giese, B.; Wessely, S.; Langenbacher, T.; Michel-Beyerle, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11713–11716. (d) Jortner, J.; Bixon, M.; Langenbacher, T.; Michel-Beyerle, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12759–12765.

(5) (a) Williams, T. T.; Odom, D. T.; Barton, J. K. *J. Am. Chem. Soc.* **2000**, *122*, 9048–9049. (b) Giese, B.; Amaudrut, J.; Kohler, A.-K.; Spormann, M.; Wessely, S. *Nature* **2001**, *412*, 318–320. (c) Giese, B.; Spichty, M. *ChemPhysChem* **2000**, *1*, 195–198.

(6) Bixon, M.; Jortner, J. *Chem. Phys.* **2002**, *281*, 393–408.

(7) Conwell, E. M.; Park, J.-H.; Choi, H.-Y. *J. Phys. Chem. B* **2005**, *109*, 9760–9763.

(8) Conwell, E. M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8795–8799.

(9) Kendrick, T.; Giese, B. *Chem. Commun.* **2002**, *2*, 2016–2017.

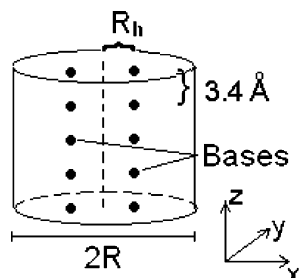


Figure 1. Structure for DNA used in the calculations.

that this result could indicate that the injected holes were delocalized along the As.¹⁰

The importance of polarons has been given a boost by some recent experiments. In one of these it was demonstrated that the hole wavefunction is delocalized over a number of bases, including cytosines and presumably thymines (Ts), as well as Gs and As.¹⁰ In another set of experiments O'Neill and Barton excited an isomer of adenine, 2-aminopurine, (Ap), to create a hole, which traveled from Ap* to a guanine separated from Ap* by a number n As, for different values of n . They found the resulting fluorescence quenching to be greater for $n = 4$ than for $n = 3, 5, 6$, etc. From this they inferred that the rate of hole transport was greater for $n = 4$ than for the other values of n . This they interpreted as evidence for formation of a coherent entity four sites in length on the adenines for $n \geq 4$.¹¹ We identify this coherent entity as a polaron.⁸ Their result is in agreement with our earlier calculations, where we found the width of a polaron on a series of As to be ~ 4 As.^{3,12}

In addition to the overlap of wavefunctions on adjacent bases on the same strand, there is overlap between the hole wavefunctions on a base and its complement. This is characterized by a transverse transfer integral whose value, it will be noted, is comparable to the values of transfer integrals for adjacent bases along the strand.

In our previous work we modeled the DNA as placed inside a cylindrical cavity having the diameter of the double helix. Water and ions were outside the cavity but none inside owing to the hydrophobicity of the bases. The calculation of the energy of a hole on the base stack was simplified by allowing for only a single chain of bases placed on the axis of the cylinder. In the calculations that follow we take the bases to be on two straight chains at distances $+R_h$ and $-R_h$ from the axis of the cylinder, as shown in Figure 1. It will be seen that neglecting the helical structure of the base stack does not lead to significant error.

In the present calculation we assume the hole to be delocalized over the bases of the two chains in the duplex. Following the lines of our previous calculation,³ we first solve the electrostatic problem to obtain the interaction of the delocalized hole with the polarization it induces in the water. (Interaction with the ions was shown in ref 3 to be negligible for the usual ion concentrations.) The solution results in formal expressions for a set of interaction constants $g^{jj'}_{n-n'}$, where j and j' designate the two chains of bases, and n and n' the two bases interacting with each other in the polarization field. After a discussion of the choices for the parameters involved, the interaction constants

are evaluated and used in the Hamiltonian for the problem to solve for the hole wavefunctions on the two chains. We carry out this program for a number of different sequences. Finally the results of the calculations are discussed and some conclusions drawn.

Calculations

Generalizing our earlier calculations³ to the case of two chains of bases, we may write the Hamiltonian

$$\mathbf{H}(\psi_{n,j}, \psi_{n,j}^*) = \mathbf{H}_0(\psi_{n,j}, \psi_{n,j}^*) + (1/2) \sum_{n,n'} \sum_{j,j'} g^{jj'}_{n-n'} |\psi_{n,j}|^2 |\psi_{n',j'}|^2 \quad (1)$$

where

$$\mathbf{H}_0(\psi_{n,j}, \psi_{n,j}^*) = - \sum_{n,j} t_{n,j;n+1,j} \psi_{n,j} \psi_{n+1,j}^* + \sum_{n,j} \Delta_{n,j} \psi_{n,j} \psi_{n,j}^* - t_{\perp} \sum_n \psi_{n,j} \psi_{n,j'}^* + \text{c.c.} \quad (2)$$

Here $\psi_{n,j}$ is the amplitude of the wavefunction at the n th site (base) on the j th chain. The summations over n, n' are to be taken over all the sites on a chain, and those over $j, j' = 1, 2$ cover the two chains of the duplex. The first term in \mathbf{H}_0 is the usual tight-binding term, with $-t_{n,j;n+1,j}$ the transfer integral. The second term in \mathbf{H}_0 allows for different energy of the hole on different bases, with $\Delta_{n,j}$ the hole energy on the n th base of the j th chain relative to that of the HOMO of adenine, which we have chosen to be our zero of energy. The last term allows interchain transfer, with the quantity $-t_{\perp}$ representing the interchain coupling. The c.c. refers to the complex conjugates of the terms in $t_{n,j;n+1,j}$ and t_{\perp} . Finally, the last term on the right of eq 1 represents the effect on the hole energy of the polarization charge it induces in the water.

The assumption of straight chains for the bases rather than helical ones should lead to only small errors if, for the significant terms in the summation, the product of $n-n'$ and a , the distance between bases, is small compared to the period of the helix, 3.4 nm, which is equal to $10a$ for B-DNA. This condition is certainly satisfied for the largest term in the summations, for which $n = n'$. In our experience with the single chain calculations,³ terms with $(n - n')a > 3a$ or $4a$ made little contribution. Thus, with the only sizable contributions made by terms with $(n - n')a \leq 2a$, modeling the bases as straight chains rather than helical is not a source of significant error.

Our first step is to determine the coefficients $g^{jj'}_{n-n'}$ in the Hamiltonian (1). We set up a coordinate system with the chains at $x = \pm R_h$, $y = 0$ and extending along the z -axis. Consider first the effect of the charge on the chain at $+R_h$. The charge density $\rho(r)$ on this strand is given by

$$\rho(r) = e \sum_n |\psi_n|^2 |\phi_n(r - r_n)|^2 \quad (3)$$

where $\phi_n(r - r_n)$ is the occupied orbital on the n th base. Proceeding as we did in ref 3, we introduce the Fourier transform for the charge density on the chain at $+R_h$:

$$\rho(r) = \delta(x - R_h) \delta(y) \int_{-\infty}^{\infty} (dk/2\pi) \rho_k e^{ikz} \quad (4)$$

We can use Laplace's equation to determine $\phi_k(r_{\perp})$, the Fourier component of the potential corresponding to ρ_k . In cylindrical coordinates Laplace's equation is

$$-(\partial^2 \phi_k / \partial r_{\perp}^2) - (1/r_{\perp})(\partial \phi_k / \partial r_{\perp}) + k^2 \phi_k = 0 \quad (5)$$

The solution may be sought in the form, analogous to that in eq 11 of ref 3:

$$\phi_k = 2\rho_k [K_0(|k|\{(x - R_h)^2 + y^2\}^{1/2}) - \Phi_k(r_{\perp}, \varphi)] \quad (6)$$

(10) Shao, F.; O'Neill, M. A.; Barton, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17914–17919.

(11) O'Neill, M. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 11471–11483.

(12) Conwell, E. M.; Bloch, S. M. *J. Phys. Chem. B* **2006**, *110*, 5801–5806.

where K_0 represents the modified Bessel function, shifted from the corresponding term in eq 11 to represent the potential of a chain of bases at $+R_h$ in the x direction rather than at $r = 0$. This term represents the potential of this chain of bases in free space. The term $-2\rho_k\Phi_k$ is the potential due to the induced charges on the walls of the cavity; it must satisfy Laplace's equation inside the cavity with the boundary condition, discussed in ref 3, that the potential vanishes at $r_\perp = R$.

To find Φ_k we use the general solution of the modified Bessel eq 5

$$\Phi_k(r_\perp, \varphi) = \sum_{m=-\infty}^{\infty} \Phi_{k,m} e^{im\varphi} [c_1 K_m(|k|r_\perp) + c_2 I_m(|k|r_\perp)] \quad (7)$$

where c_1 and c_2 are arbitrary constants and $I_m(|k|r_\perp)$ is the other modified Bessel function. Requiring the solution to be finite inside the cavity, we set $c_1 = 0$. To obtain an expression for the coefficients $\Phi_{k,m}$ we use the condition that the total potential $\phi_k(r_\perp)$ must vanish at the cylindrical surface of the cavity, $r_\perp = R$. With this condition, absorbing c_2 into $\Phi_{k,m}$ and using eqs 6 and 7 we obtain

$$K_0[(kR \cos \varphi - kR_h)^2 + k^2 R^2 \sin^2 \varphi]^{1/2} = \sum_{m=-\infty}^{\infty} \Phi_{k,m} e^{im\varphi} I_m(|k|R) \quad (8)$$

Multiplying eq 8 by $e^{-im\varphi}$ and integrating over φ leads to the desired expression for the coefficients $\Phi_{k,m}$:

$$\Phi_{k,m} = (2\pi I_m(kR))^{-1} \int_{-\pi}^{\pi} d\varphi e^{-im\varphi} K_0[(k^2 R^2 + k^2 R_h^2 - 2k^2 R R_h \cos \varphi)^{1/2}] \quad (9)$$

Integration over φ in eq 9 allows for the angular variation over the cavity of the induced charge. In eq 7 there is no integration over φ ; φ refers to the angular position of the chain in space. In polar coordinates chain 1 is at $\varphi = 0$ and $r_\perp = +R_h$, while chain 2 is at $\varphi = \pi$ and $r_\perp = -R_h$. Thus, since \sum_m in eq 7 is over all integers, for $g^{11}_{n-n'}$ the factor $\sum_m e^{im\varphi}$ in eq 7 will yield, finally, $(+1)^m$, while $\sum_m e^{-im\varphi}$ will yield $(-1)^m$.

Calculation of the energy shift of the charge on chain 1 due to the polarization charge it induces requires the definition of $g^{ij}_{n-n'}$ in eq 1, and eqs 4, 6, 7, and 9. With ρ_k in the expression for the energy shift appearing in both the charge density and the potential, $-2\rho_k\Phi_k$, integration over z requires that $k' = -k$. We use also the inverse Fourier transform, which for chain 1 leads to

$$\rho_k = e \sum_n |\psi_n|^2 e^{-ikna} \quad (10)$$

Proceeding as in ref 3, we obtain finally

$$g^{11}_{n-n'} = g^{22}_{n-n'} = -e^2 \int_{-\infty}^{\infty} (dk/2\pi) 2\Phi_k(R_h, 0) e^{ika(n-n')} \quad (11a)$$

$$g^{12}_{n-n'} = g^{21}_{n-n'} = -e^2 \int_{-\infty}^{\infty} (dk/2\pi) 2\Phi_k(R_h, \pi) e^{ika(n-n')} \quad (11b)$$

where

$$\left. \begin{aligned} \Phi_k(R_h, 0) \\ \Phi_k(R_h, \pi) \end{aligned} \right\} = \sum_{m=-\infty}^{\infty} (\pm 1)^m \frac{I_m(kR_h)}{I_m(kR)} \int_{-\pi}^{\pi} \frac{d\varphi}{2\pi} e^{-im\varphi} K_0 \left(\sqrt{k^2 R^2 + k^2 R_h^2 - 2k^2 R R_h \cos \varphi} \right) \quad (12)$$

It is seen that $g^{12} < g^{11}$, as expected because the polarization charge induced by chain 2 on the surface of the cavity close to chain 1 will always be less than that induced by chain 1 itself on that part of the surface. Similarly $g^{21} < g^{22}$.

The parameters used in the calculations and the resulting g values will be discussed in the next section. Once the g values are obtained, the polaron wavefunctions and energies can be determined by diagonalizing the matrix obtained from the Hamiltonian (1) and (2). The

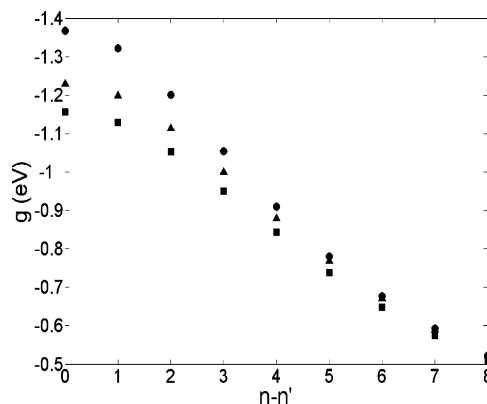


Figure 2. Values in eV vs the number of sites between n and n' of (\blacktriangle) $g^{11}_{n-n'}$ (from ref 3), (\bullet) $g^{12}_{n-n'}$, (\blacksquare) $g^{21}_{n-n'}$.

calculations were simplified by taking the diagonal elements of the matrix as $\Delta_{nj} + (1/2)\sum_{j'} g^{ij}_{n-n'} |\psi_{n',j'}|^2$. Sites 1 through 25 were assigned to one chain, sites 26 through 50 to the other. Periodic boundary conditions were used. The calculations were carried out for a number of base sequences.

Choice of Parameters. The parameter R_h was chosen as 0.3 nm, thus $0.3R$, this being fairly closely the distance between the center of each of the four bases and the axis of the helix. The resulting $g^{11}_{n-n'}$ and $g^{12}_{n-n'}$ are shown as functions of $n-n'$ in Figure 2. It is seen that only for small $n-n'$ do these parameters differ much from $g_{n-n'}$ obtained for the case of the single chain at the center of the cavity in ref 3. The differences between $g^{11}_{n-n'}$ and $g^{12}_{n-n'}$ are not great even though the distance from chain 1 to the far side of the cavity is almost twice its distance to the near side. The differences would be considerably larger for larger R_h ; $g^{11}_{n-n'}$ is almost a factor of 3 larger than $g^{12}_{n-n'}$ for $R_h = 0.7$ nm, for example. It is seen also that for large $n-n'$ the g values are much smaller and there is little difference between $g^{11}_{n-n'}$, $g^{12}_{n-n'}$ and $g_{n-n'}$. This is expected because in the limit of large $n-n'$ the distance $(n-n')a \gg R_h$; in that limit $g^{ij}_{n-n'} \approx e^2/(n-n')a$.

The energy of a hole on each of the bases is an important quantity in determining the characteristics of the polaron. We have chosen to use the values measured for the individual bases in solution.^{13,14} Relative to the value for adenine, these are cytosine, 0.5 eV; thymine, 0.7 eV; guanine, -0.4 eV. For the transfer integrals $t_{n,j;n+1,j}$ we used the calculated values of Voityuk et al.^{15,16} As will be discussed below, with some exceptions the transfer integral values are found to have only a weak influence on the wavefunction of the polaron.

The hydrogen-bond coupling of a base pair was first considered in connection with the phonon modes arising from the stretching of the hydrogen bonds.¹⁷ The coupling was represented by a Morse potential consisting of D , the depth of the potential well, and a multiplying factor involving the relative motion of the bases. Because we are considering only stationary states of DNA, well below the denaturing temperature, our only concern is the well depth. It has been emphasized that D represents not only the hydrogen bonds connecting the two bases in the pair, but also the repulsive interaction of the phosphate backbone and the surrounding solvent effects. It has been used to simplify the problem by assuming that D is an average potential representing the two or three bonds that connect the bases in a complementary pair. An

(13) Orlov, V. M.; Smirnov, A. N.; Varshavsky, Y. M. *Tetrahedron Lett.* **1976**, *48*, 4377–4378.

(14) Steenken, S.; Jovanovic, S. C. *J. Am. Chem. Soc.* **1997**, *119*, 617–618.

(15) Voityuk, A. A.; Röscher, N.; Bixon, M.; Jortner, J. *J. Phys. Chem. B* **2000**, *104*, 9740–9745.

(16) A set of quite different values of $t_{n,j;n+1,j}$ was published subsequently by Troisi, A.; Orlandi, G. *Chem. Phys. Lett.* **2001**, *344*, 509–518.

(17) Gao, Y.; Devi-Prasad, K. V.; Prohofsky, E. W. *J. Chem. Phys.* **1984**, *80*, 6291–6298.

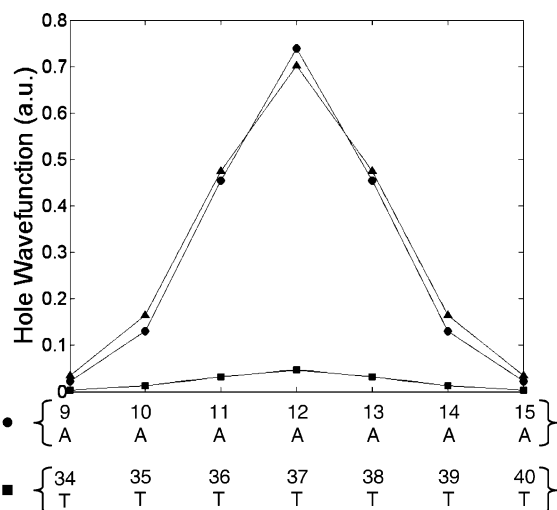


Figure 3. Calculated wavefunction of polaron (▲) on a single all-A chain at center of helix, (●) on A chain of A/T duplex, (■) on T chain of A/T duplex.

estimate of the value of D , with all these effects taken into account, yielded $D = 0.04$ eV,¹⁸ which we have taken as the value of t_{\perp} .

Results

The values of $g_{n-n'}^{ij}$ obtained in the calculations are displayed in Figure 2.

It is seen that $g_{n-n'}^{11}$, which is due to the interaction of chain 1 with the polarization to which it gives rise, is greater than $g_{n-n'}^{12}$, which is due to the interaction of chain 1 with the polarization to which chain 2 gives rise, as expected. Also, as expected, their values bracket those of $g_{n-n'}$, the value obtained when the chain is at the center of the helix, equidistant from all parts of the cavity wall. The difference between g^{11} and g^{12} , or g^{22} and g^{21} for given $n-n'$ would be larger if R_h were larger.

The wavefunctions for a duplex with one chain being all As, the other being all Ts, A/T, are shown in Figure 3. Also shown are the results obtained earlier for a single chain with all As.¹² In obtaining these results $t_{n,n+1}$ was taken as 0.03 eV for adjacent adenines, from the calculations of refs 15, 19, and 20, and 0.16 eV for adjacent thymines, from the calculations of ref 15. As expected because of the much higher energy of the hole on thymine, almost all of the hole wavefunction is on the adenines. At the peak the wavefunction on the T chain is only a few % of that on the A chain of the duplex. It is also to be expected then that the wavefunction on the A chain of the duplex is only a few % different from that on the single chain of As.

The calculated energy of the A/T polaron was -0.7103 eV. In the absence of a hole the lowest electronic energy level calculated for the A/T duplex was -0.0636 eV, only 0.0036 below the lowest energy level, $-2t_{n,n+1}$, for the single A chain without a hole. The small size of the increment is also the result of the high energy of the hole on T. In any case, the numbers just cited give the binding energy of the polaron on A/T as 0.65 eV. The binding energy of the hole polaron on a single A chain running along the axis of the helix with the same value of $t_{n,n+1}$ was calculated to be 0.59 eV.¹² The larger binding energy

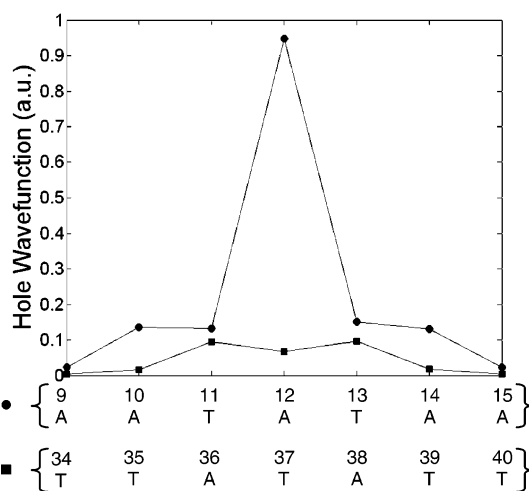


Figure 4. Calculated wavefunction of polaron for duplex in which ● denotes the chain with the sequence ATATA surrounded by As, and ■ denotes the chain with the sequence TATAT surrounded by Ts. Hole was inserted at site 12. Below the figure are the site numbers and the letters specifying the base at each site.

in the duplex is due to the increased polarization charge induced by the hole when it is moved closer to the surface of the cavity.

Calculations were carried out for duplexes with two or three of the central As replaced by Gs and the complementary Ts replaced by Cs. The results on the mainly A chain are quite similar to those obtained in ref 12 for a single A chain with the same number of Gs. The difference in energy between the duplex with 3Gs and the duplex with 2Gs is 0.02 eV, which is the same result as we obtained for the single chain case in ref 12. This value is in good agreement with the measured value of this difference, 0.027 eV, obtained by Lewis et al.²¹

For the duplexes in which different bases are mixed on the same chain, as mentioned earlier, we used the values of $t_{n,j;n+1,j}$ calculated by Voityuk et al., listed in Table 2 of ref 15. For a given pair of bases XY the value of the transfer integral given there is for n on the 5' side, $n+1$ on the 3' side. Transfer integrals are not easy to calculate, and there is no reason to believe that the calculated values are accurate. We felt, however, that, rather than assuming the transfer integrals are all the same, as is most frequently done, it would be worthwhile to incorporate a set of calculated values to give an idea of the effects that result from different values for different n,n' .

We found that the wavefunctions could be changed greatly by changing the site at which the hole is inserted at the beginning of the calculation. This is illustrated above for the sequence obtained from A/T by inserting Ts on the A chain (consisting of sites 1 through 25) at sites 11 and 13, and the complementary As at sites 36 and 38 on the T chain (consisting of sites 26 through 50). In Figure 4 we show the results of inserting the hole at site 12, an A flanked by Ts. The wavefunction is much larger on the mainly A chain than on the mainly T chain.

It is almost symmetric about site 12 on the former and site 37 on the latter, the small asymmetries being due to the difference between the calculated values¹⁵ of $t_{n,j;n+1,j}$ for TA (0.086 eV) and AT (0.185 eV). The total energy with the polaron on the chain is -0.709 eV, only a little larger than the energy

(18) Dauxois, T.; Peyrard, M.; Bishop, A. R. *Phys. Rev.* **1973**, *47*, 684–698.

(19) Grozema, F. C.; Siebbeles, L. D. A.; Berlin, Y. A.; Ratner, M. A. *ChemPhysChem* **2002**, *2*, 536–539.

(20) Senthilkumar, K.; Grozema, F. C.; Fonseca Guerra, C.; Bickelhaupt, F. M.; Lewis, F. D.; Berlin, Yu. A.; Ratner, M. A.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2005**, *127*, 14894–14903.

(21) Lewis, F. D.; Liu, X.; Liu, J.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2000**, *122*, 12037–12038.

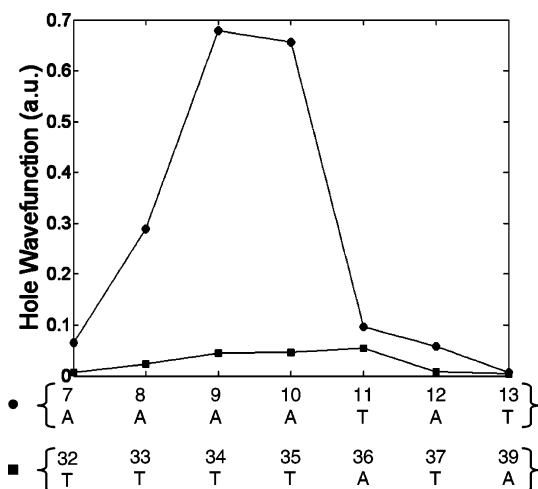


Figure 5. Calculated wavefunction of polaron for same duplex as in Figure 4 with hole inserted at site 10.

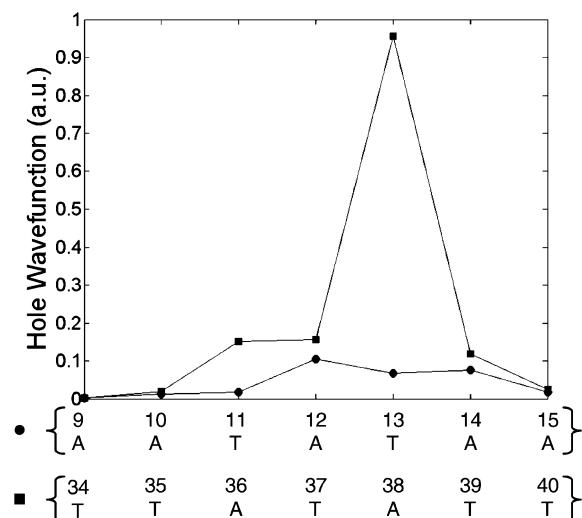


Figure 6. Calculated wavefunction of polaron for same duplex as in Figure 4 with hole inserted at site 37.

found earlier for the polaron on A/T, although the wavefunctions are rather different. In Figure 5 insertion of the hole at site 10 in the same sequence results in the wavefunction having approximately the same value at site 9 as at site 10, both As. The slightly larger value at site 9 is due to there being another A at 8 flanking the A at 9 while a T is flanking the A at 10. The larger amplitude on the left side of site 10 is due to the sites on the left being all As. It is interesting to note that the large value of $t_{n,j,n+1,j}$ for the neighboring pair AT has considerable influence on the wavefunction. If it is reduced from 0.105 to 0.03 eV, for example, the peak amplitude of the wavefunction is shifted from site 9 to site 8, the amplitudes on the A at site 10 and on T at site 11 are decreased, the former by more than a factor 3, the latter by an order of magnitude. Further decrease in the AT coupling, even to 0, does not change the wavefunction further.

Still another behavior is shown for this duplex in Figure 6, with the hole introduced on site 37. The wavefunction is now principally on the mainly T chain rather than the mainly A chain, although it peaks on site 38 rather than the site on which it was introduced. The wavefunction is almost symmetric about site 38 on chain 2. It is interesting that if the wavefunction is calculated on the assumption that $t_{n,j,n+1,j}$ is the same for TA as

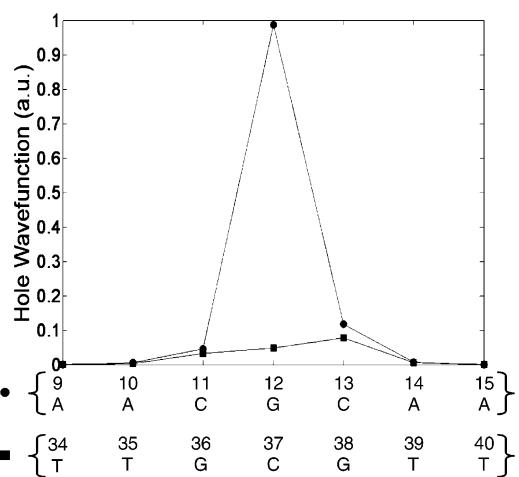


Figure 7. Calculated wavefunction of polaron for duplex in which ● denotes the chain with the sequence CGC surrounded by As, and ■ denotes the chain with the sequence GCG surrounded by Ts. Hole was inserted at site 12.

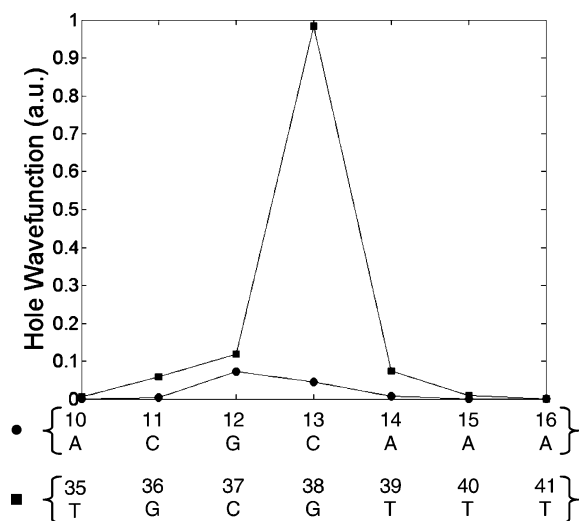


Figure 8. Calculated wavefunction of polaron for same duplex as in Figure 7 with hole inserted at site 38.

for AT, the wavefunction has two peaks of the same size at the As flanking the T at site 37, where the hole was introduced. Thus the larger value of $t_{n,n+1}$ for AT than for TA means that a hole on site 37 sees a bigger coupling to site 38 than to 36, which results in the peak at 36 being suppressed. (Note that 3' and 5' are opposite on the predominantly T chain from the predominantly A chain.)

Calculations for the sequence in which the As on sites 11 to 13 are replaced by CGC and the Ts from 36 to 38 replaced by GCG are shown in Figures 7 and 8. In Figure 7, where the hole is inserted at site 12 the wavefunction is strongly peaked at that site, with very little amplitude on the other chain. It is only slightly asymmetric, owing to the difference between the transfer integrals for CG, 0.042 eV, and GC, 0.110 eV.¹⁵ Insertion of the hole at site 13 resulted in the same wavefunctions as shown in Figure 7. As is seen in Figure 8, insertion of the hole at site 38 resulted in the wavefunction being strongly peaked at that site, and very little of the wavefunction is seen on the other chain.

Comparing Figures 4–8 with Figure 3 we see that the peak of the polaron is much narrower in the former figures and the

polaron is more confined. The increased localization is the result of a purine being flanked on both sides by pyrimidines, on which the hole energy is much greater.

An important feature, seen in all the figures, is that the hole wavefunction is predominantly on one chain, the one on which the hole was originally inserted, despite there being lower energy sites available on the other chain. Apparently the interchain transfer integral is not large enough to cause switching between chains for the sequences examined here. It can be anticipated, however, that for some sequences there would be a switch from the hole being predominantly on one chain in one part of the duplex to being predominantly on the other chain in another part of the same duplex. As a specific example consider a chain, I, with 25 As followed by 20 Ts, hybridized with chain II having 25 Ts followed by 20 As. If a hole were inserted around the middle of the A region in chain I it would form a polaron similar to the one shown in Figure 3, with its wavefunction predominantly on chain I. If there were an electric field pushing the polaron toward the T region of chain I, after crossing into the T region it would switch so that its wavefunction would be predominantly on the As in chain II. Because of the delocalization of the polaron, switching requires that there be a number of unfavorable, that is, high energy, sites in the path of the polaron. Reasonably that number should be of the order of the length of a polaron or greater. We have shown that this is indeed the case in calculations for appropriate sequences that are included in Supporting Information.

We have also carried out calculations for a duplex where one strand has three Ts inserted in a series of As. This was intended for comparison with the experimental results of Joy et al., one of whose sequences consisted of $[(A)_2(T)_3]_n$.¹ The resulting wavefunctions are shown in the Supporting Information. We found that for TTT inserted into sites 10 to 12 on the chain with all the other sites As, and the hole inserted on any site between 7 and 13, the wavefunction spread over approximately six sites around that site and was large on the As, and less than $1/7$ as large on the Ts. On the other strand the wavefunction was spread over approximately six sites opposite to those it occupied on the first strand but was smaller everywhere than on the first strand. We conclude that the introduction of three high-energy sites is not enough for a switch of the majority of the wavefunction to the other chain. This is in agreement with the results of ref 1, as will be discussed further in the next section.

Implications for Transport. On the view that a hole is localized on a single site, transport would occur by hopping even in a sequence made up of a single base pair repeated. A much studied case is an all A/T sequence. Hole motion on such a sequence has been dubbed A-hopping.⁵ Experiments of Kawai et al.²² and Takada et al.²³ have shown that hole motion in an all-A/T sequence is quite rapid. In the latter paper the assumption that the motion was hopping led to a hopping rate of $2 \times 10^{10} \text{ s}^{-1}$.²³ Later investigation of an all A/T sequence by Lewis et al. found that, if the motion were hopping, it required a hopping rate smaller by a factor of 40.²⁴ It should be noted, however,

that the rates obtained in these studies are expected to be field-dependent²⁵ and therefore cannot be compared.

It has been suggested that the motion on an all A/T sequence is not hopping of a hole localized to a single site but drift of a hole polaron in the electric field.^{3,26} The dependence of the rate of drift on the electric field and other parameters has been calculated.²⁶ To establish that hole motion on an all A/T sequence is polaron drift rather than hopping would, however, require detailed knowledge of the spatial variation of the electric field. The field is expected to be quite nonuniform in experiments such as those of refs 22–24.²⁶

In recent experiments Osakada et al. observed the transport of holes injected into a GCGCGC... sequence and found it to be comparably rapid to hole transport in an all-A sequence.²⁷ On the assumption that a hole is localized on one site, and taking into account the fact that the hole energy is lowest on a guanine, there are two possibilities for the hole transport: hopping from one G to the next one on the same strand or hopping from a G on one strand to the closest G, which is on the other strand. They concluded that the interstrand hopping is more likely.²⁷ Although early experiments of Kelley and Barton were interpreted by them as evidence for intrastrand hopping,²⁸ experimentalists have generally interpreted the results of their experiments in terms of interstrand hopping.^{5b,c,27} A number of theorists have also favored interstrand hopping.^{4c,d,16,30}

Our calculations, as discussed above, favor the intrastrand hopping option. Although the hole wavefunction is split between the two complementary bases, unless, as discussed earlier, there are too many high-energy bases in the sequence, the large majority of the hole wavefunction stays on one strand of the duplex, the strand on which it was originally placed. An experimental proof of this was provided recently by experiments of Joy et al.¹ In the experiments an anthraquinone (AQ) group is excited by light to create a hole on a DNA duplex to which it is connected. One of the duplexes they studied consisted of $(AATT)_2(A)_5(AATT)_2$ on one strand and its complementary sequence on the other strand, with the TT next to the AQ on the 5' side and AA next to the AQ on the 3' side. Because the oxidation potential of A is much less than that of T it is expected that the hole will be created on the 3' A next to AQ. The hole, that is, the radical cation, is expected to migrate down the strand and ultimately be trapped in a reaction of the cation with H_2O or O_2 . Evidence of the trapping is provided by strand cleavage on subsequent treatment with piperidine. In the usual experiments of this type the hole is trapped at a guanine, which has the lowest oxidation potential of the bases. With no Gs present, it was expected that the hole would be trapped at an A, which has the next lowest ionization potential, most likely in the $(A)_5$ segment.¹ Instead the hole visited all four TT segments and gave rise to strand cleavage at a T in each one of the four TTs. In similar experiments with the sequence $[(A)_2(T)_3]_n$ Joy et al. found piperidine caused strand cleavage predominantly at the central T of the TTT sites. Thus, as noted earlier, the presence

- (22) Kawai, K.; Takada, T.; Tojo, S.; Majima, T. *J. Am. Chem. Soc.* **2003**, *125*, 6842–6843.
(23) Takada, T.; Kawai, K.; Cai, X.; Sugimoto, A.; Fujitsuka, M.; Majima, T. *J. Am. Chem. Soc.* **2004**, *126*, 1125–1129.
(24) Lewis, F. D.; Zhu, H.; Daublain, P.; Cohen, B.; Wasielewski, M. R.; *Angew. Chem., Int. Ed.* **2006**, *45*, 1–4.

- (25) The electric field referred to here includes the field due to the charge on the hole donor and the force per unit charge due to the difference in oxidation potentials of the hole donor and hole acceptor.
(26) Conwell, E. M.; Basko, D. M. *J. Phys. Chem. B* **2006**, *110*, 23603–23606.
(27) Osakada, Y.; Kawai, K.; Fujitsuka, M.; Majima, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18072–18076.
(28) Kelley, S. O.; Barton, J. K. *Science* **1999**, *283*, 375–381.
(29) Troisi, A.; Orlandi, G. *J. Phys. Chem.* **2002**, *106*, 2093–2101.
(30) Voityuk, A. A.; Siri Wong, K.; Rösch, N. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5421–5425.

of three high-energy sites did not cause significant switching of the wavefunction to the other strand.

From other experiments Joy et al. learned that the methyl group in the thymine radical cation has an important role in this effect, acting to increase the time the hole spends on a thymine long enough for the thymine to lose a methyl proton and subsequently undergo an irreversible trapping reaction with O₂. The significance for us is that, although there were the lower ionization potential As opposite the Ts, the hole went from the As on which it started to the Ts, staying predominantly on the chain on which it started. Note that it does not appear to be the case that the amplitude of the hole on a thymine, seen in the preceding figures to be small relative to that on an adenine, increased. Finding the trapping on a GG to be much more efficient than trapping on TT, Joy et al. suggest that in the DNA oligomers that lack guanines the hole visits each T many times before reacting.¹ Multiple visits could explain also Giese's finding of the effectiveness of a GGG for trapping holes that were injected on the other strand, presumably resulting in small amplitude of the wavefunction on the GGG.^{4,5} Joy et al. demonstrated trapping of the hole at Ts in other duplexes with different sequences of As and Ts. In all cases the cleavage was found on the strand on which the hole was injected from the AQ,¹ in agreement with our calculations.

We believe that, apart from all-A sequences, polaron hopping is the process that accounts for transport in DNA.^{2a,b} It is well-known that hopping probability is largest when there is a match of wavefunctions between hopped-to and hopped-from regions. To account for hopping between, for example, two different sequences, the thermal vibrations of DNA have been invoked.^{2a,b} Calculations of Voityuk et al. show that the square of the matrix element for hole transfer between adjacent Watson–Crick pairs may vary by a factor of several hundred owing to moderate changes of the duplex conformation.³⁰ They have also shown that thermal fluctuations may result in large changes in energy of the bases, such that, for example, the energy of A⁺ may become lower than the energy of G⁺,³¹ whereas it should be 0.2 to 0.4 eV higher. Troisi and Orlandi, investigating the effective coupling between two Gs separated by four As or four Ts or TATA have shown that the thermal fluctuations enhance their average coupling by about 1 order of magnitude.²⁹ Direct experimental evidence that thermal fluctuations are important for transport is the finding of O'Neill and Barton that below the duplex melting temperature (~60 °C) the yield of charge transfer through DNA increases with increasing temperature.³² In their terminology the role played by the thermal fluctuations is conformational gating.

Conclusions

Our earlier result, obtained with the single-chain model of the polaron, that the width of the polaron wavefunction on an all-A sequence is 4–5 sites holds up on calculations for a duplex with geometry that reasonably approximates the actual duplex in DNA. Thus we retain agreement with experiment on this point. The energy difference between sequences with two Gs

and three Gs, in both cases surrounded by As, was unchanged when the single-chain model of the polaron was replaced by the duplex, thus is still in agreement with the experimental results for this quantity obtained by Lewis et al.²¹

It should be noted that the binding energies we have calculated appear to be somewhat large. Extensive quantum chemical calculations of the energy lowering of a hole on an (AAGG)₃AA sequence in water by Barnett et al. led to a value of 0.4–0.5 eV.³³ We believe that our simple model leads to higher binding energy, 0.65 eV, because we treat water as a continuum. At the interface between water and DNA the situation is more complex. According to the calculations of Barnett et al.³³ the water molecules closest to the DNA are strongly bonded to it and may not be able to completely adjust their dipole moments as we have assumed. This may be viewed as an effective increase of the radius of the cavity in our model, which would have the effect of decreasing the *g* coefficients and the binding energy we obtain.

Because of large differences in energy of a hole on sites belonging to different bases, the hole wavefunction is generally much larger on one chain of duplex DNA than the other, whatever the sequence. It appears that this would still be true if the on-site energies were closer together by a couple of tenths of an eV than was assumed here by taking the values as those for the isolated bases. When a high-energy site, say T, is next to a low-energy site, A or G, if the wavefunction peaks at G it drops strongly at the T, but remains quite small at the A opposite T. The effect of two Ts flanking an A or G is to produce a wavefunction strongly confined to the A or G but still small at all sites on the complementary chain. This is illustrated in Figure 4, for example. The width of the wavefunction on an all A sequence is almost twice that of an A between two Ts, or, more generally, of a purine between two pyrimidines.

As has been pointed out in the discussion above, differences between the transfer integral $t_{n,j;n+1,j}$ for different base pairs may have considerable effect on the wavefunction. The effects are more subtle, however, than those resulting from the differences in on-site energies of the bases.

Finally, our calculations show that, when a hole is introduced on one strand of a duplex DNA, its wavefunction tends to be larger on that strand unless there are considerable changes in sequence seen by the hole as it migrates. We find that a change such as insertion of three thymines in a series of adenines is not sufficient to cause the preponderance of the wavefunction to switch from one strand to the other. This result is in agreement with experiments of Joy et al.¹ Interpretation of experimental data on the basis of the assumption that the hole wavefunction is confined to a single base is liable to lead to incorrect conclusions concerning the path of the hole as it migrates, such as concluding that it zigzags between strands.

Acknowledgment. S.M.B. and P.M.M. are grateful for the support of the Camille and Henry Dreyfus Foundation.

Supporting Information Available: Calculated wavefunctions for DNA with sequence on one strand (A)_l(T)_m(A)_n, 3 ≤ *m* ≤ 5, *l* = 9, *n* ≥ 9, and complementary sequence on the other strand. This material is available free of charge via the Internet at <http://pubs.acs.org>.

(31) Voityuk, A. A.; Siriwong, K.; Rösch, N. *Angew. Chem., Int. Ed.* **2004**, *43*, 624–627.

(32) O'Neill, M. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 11471–11483.

(33) Barnett, R. N.; Bongiorno, A.; Cleveland, C. L.; Joy, A.; Landman, U.; Schuster, G. B. *J. Am. Chem. Soc.* **2006**, *128*, 10795–10800.