Base Sequence Effects on Transport in DNA

Esther M. Conwell* and Steven M. Bloch

Departments of Chemistry and Physics, University of Rochester, Rochester, New York 14627

Received: September 22, 2005; In Final Form: January 20, 2006

Given the success of the polaron model based on solvation in accounting for the width of a hole polaron on an all-adenine (A) sequence on DNA, we extend the calculations to other sequences. We find excellent agreement with the free energy differences measured by Lewis et al. (*J. Am. Chem. Soc.* **2000**, *122*, 12037−12038) between a guanine (G) cation and a pair of bases, GG, or a triple of bases, GGG, in all cases surrounded by As, by treating AGGA and AGGGA as solvated polarons. There is additional support for hole polaron formation in DNA from experiments in which oxidative damage due to injected holes is investigated in sequences involving Gs and As. Theory and comparison with transport measurements on repeated sequences involving multiple thymines (Ts) or combinations such as ATs or GCs, where C is cytosine, led to the suggestion that the basic sequences in these cases must be polarons whose wave functions have substantial amplitudes on both chains in a duplex. The size of an electron polaron in DNA is predicted to be similar to that of a hole polaron, approximately 4 or 5 bases. Although experiments have shown that polaron hopping is the dominant mode of charge transport in DNA with repeated sequences such as AGGA, further investigations, particularly of temperature dependence of site energies and transfer integrals, are needed to determine to what extent hole transport takes place by polaron hopping for arbitrary DNA sequences.

Introduction

In a polarizable medium a charge and the polarization to which it gives rise must move together, the resulting composite being called a polaron. In a one-dimensional case the polaron is spread over a number of sites,[1] thus delocalized. There has been considerable debate as to whether the wave function of a hole in DNA is localized or delocalized over a number of bases, with the majority opinion holding for the former. Recently, however, it has been demonstrated experimentally by Barton and colleagues that the hole wave function is delocalized.[2] In their experiments the group used a modified cytosine, N4-cyclopropylcytosine (^{CP}C), which is a fast hole trap, to at least partially overcome rapid back charge transfer. The ^{CP}C, set into a DNA duplex a number of bases away from a photooxidant (e.g., an anthraquinone), showed appreciable oxidative damage on excitation of the photooxidant. When the trapping rates of cytosine and guanine were made comparable by cyclopropyl substitution of both and the two modified bases were neighbors on the same strand, the oxidative damage of the two due to the photooxidant was found to be comparable in magnitude, clear evidence of delocalization of the hole wave function. This suggests that, given an appropriate sequence, the hole would form a polaron. It is expected that an excess electron would similarly form a polaron because the situation is essentially the same.

There are two ways in which a hole or an excess electron on the DNA base stack may become a polaron. The charge may give rise to polarization by distorting the arrangement of a number of adjacent bases on which its wave function is centered, decreasing their relative spacing by greater amounts toward the center of the polaron. This is the type of polaron formed in conducting polymers, and it was first suggested that this is the type of polaron formed in DNA.[3,4] Its properties depend on the wave function overlap of adjacent bases, represented by the transfer integral t, the rate of change of t with interbase spacing, and the elastic constant K that relates to change in spacing. On the basis of what seemed to be reasonable values for these parameters, the calculated size of the polaron ranged from 4 to 6 bases, depending somewhat on which bases were included.[4] Its binding energy was found to vary from $\sim$ 0.3 to 0.03 eV as the transfer integral, assumed to be the same for all pairs of bases, varied from 0.3 to 0.2 eV.[4] Because more recent, improved calculations of t give values less than 0.2 eV, the binding energy of this type of polaron is quite small.

The second type of polaron is due to the polarization by the excess charge of the medium surrounding the DNA, water and ions. To distinguish it from the first type of polaron, which we will call the distortion polaron, we will call it the solvated polaron. The properties of this type of polaron were calculated by Kurnikov et al.[5] and by Basko and Conwell.[6] For a duplex with one strand consisting of three guanines (Gs) surrounded by adenines (As) Kurnikov et al. found that the polaron was 1−3 base pairs in length.[5] Basko and Conwell calculated the properties of this type of polaron for all bases the same, e.g., As, with a model to be described further below. The presence of the water and ions, which represent a highly polarizable environment, gave a binding energy of 0.5 eV and a width for an all-A polaron of 4 or 5 sites.[6]

Significantly, O'Neill and Barton found that transfer of a hole introduced into an all-A sequence by photoexcitation was more rapid through a length of 4−5 As, or 8 As, than through other numbers of As in the sequence.[7] They interpreted this finding as evidence that 4 or 5 bases is the characteristic size of a region they called a "domain", possessing, as a result of thermal fluctuations, at least instantaneously a specific, well-coupled conformation of DNA bases.[7] Charge transfer through DNA was attributed by them to hopping among such well-stacked

* Author to whom correspondence should be addressed. Phone: (585) 275-5841. Fax: (585) 276-0205. E-mail: conwell@chem.rochester.edu.

10.1021/jp0553986 CCC: $33.50
Published on Web 02/22/2006

domains, called conformationally gated charge transfer.[7] We are in agreement with their interpretation, identifying the "domain" as a polaron.[8] It is known from studies of conducting polymers such as poly(phenylene vinylene), PPV, that an electron or hole enters the polymer, in that case usually from a metal contact, as a polaron when the appropriate chain configuration is prepared by fluctuations.[9] That is also an example of conformationally gated charge transport, as is the continued hopping of polarons after they have entered the PPV.

The value of the binding energy of the solvated polaron was found to be insensitive to the size of the transfer integral, changing less than 10% when the transfer integral was decreased from 0.2 to 0.1 eV.[6] Because the contribution to the binding energy of the distortion is quite small for such a small t_0 we concluded that it would be sufficiently accurate to calculate polaron properties taking into account the environment and neglecting the distortion.[10]

It is our purpose in the present paper to extend our earlier calculations of the properties of hole polarons, using the Hamiltonian of ref 6, to other base sequences, showing that they lead to results in good agreement with experiment. We will also discuss briefly the properties of electron polarons in DNA. Finally, we will discuss the implications of our results for the mechanism of transport of excess electrons and holes.

Theory for the Hole Polaron in Solution

In ref 6 a term was added to the Hamiltonian to represent the hole self-localization energy arising from the polarization of the environment. A practical way of describing the environment is to consider the DNA molecule to be placed inside a cavity, in this case arising from the sugar−phosphate backbone and the hydrophobicity of the DNA bases. We take the cavity to be cylindrical, with radius R determined by the size of the helix; thus $R = 1.02$ nm. The water and ions are assumed to be outside the cavity.[11,12] We neglect the polarizability of the backbone and the bases. The hole wave function is represented as a linear combination of molecular orbitals

$$\psi(r) = \Sigma_n \psi_n \phi_n(r - r_n) \tag{1}$$

where ϕ_n is the orbital of the nth base and ψ_n is the probability amplitude for the hole on this base. With eq 1 the hole energy may be expressed as[6]

$$H(\psi_n,\psi_n^*) = H_0(\psi_n,\psi_n^*) + (^1/_2)\Sigma_{n,n}{}'g_{n-n}{}'|\psi_n|^2|\psi_n{}'|^2 \tag{2}$$

where H_0 corresponds to free hole motion, i.e., nearest neighbor hopping with the transfer integral t_0

$$H_0(\psi_n,\psi_n^*) = \Sigma_n \Delta_n \psi_n \psi_n^* - t_0 \Sigma_n (\psi_n \psi_{n+1}^* + \psi_n^* \psi_{n+1}) \tag{3}$$

Here Δ_n is the energy of the hole on the nth base. The hole spin is not included because it is not essential to the problem. The second term in eq 2 represents the effect on the hole energy of the polarization of the environment, water and ions outside the cavity.

In ref 6 we showed that, due to the high dielectric constant of water, the interaction with the ions, at the concentrations usually used, can be neglected; the interaction with water is dominant. To simplify the geometry we assume the charge, which we take to be located on one strand of the DNA, is concentrated on the axis of the cylinder, where the coordinate $r_\perp = 0$, at evenly spaced points $z_n = na$ corresponding to the bases. The problem of determining the energy due to polarization can then be reduced to solving Laplace's equation for the region $0 < r_\perp < R$ subject to the appropriate boundary conditions.[6] Taking advantage of the cylindrical symmetry we may express the charge density as a Fourier transform

$$\rho(r) = \delta(x)\delta(y) \int \mathrm{d}k\, (2\pi)\rho_k \mathrm{e}^{ikz} \tag{4}$$

where the integration is from $k = -\infty$ to $+\infty$. The Fourier component of the potential φ_k may then be expressed in terms of modified Bessel functions K_0 and I_0, and finally the coefficients in eq 2 are found to be[6]

$$g_n = -(\mathrm{e}^2/R)\nu(na/R) \tag{5}$$

where

$$\nu(\xi) \equiv (2/\pi) \int_0^\infty [K_0(\mathrm{q})/I_0(\mathrm{q})] \cos q\xi \, \mathrm{d}q \tag{6}$$

and a is the interbase spacing, taken as the value characteristic of B-DNA, 0.34 nm. To summarize, what is used in the determination of the key parameter g_n are the experimental fact that the hole wave function is delocalized over a number of bases, the high dielectric constant of water, and the approximation of taking the charged bases to be located along the axis of the cylinder, resulting in cylindrical symmetry. The values of $\nu(\xi)$ vs ξ (na/R) for n ranging from 0 to 8 are given in Table 1 of ref 6. It is seen that $\nu(\xi)$ falls off gradually with increasing ξ, decreasing by only a factor 2 as ξ increases from 0 to 7.

To obtain the eigenfunctions and eigenvalues we used eqs 2 and 3 with the g values calculated from eqs 5 and 6. The hole was assumed to be on a DNA oligomer 100 bases in length, with periodic boundary conditions. The zero of energy was taken at the bottom of the free hole band on a chain with $\Delta_n = 0$, i.e., all bases the same.

Application to Sequences with Guanines and Adenines

We consider first the parameters to use in applying the formalism to hole polarons. For an all-A polaron the only parameter required is t_0. Values in the literature for t_0 of a pair of As in B-DNA vary from 0.03[13] to 0.156 eV.[14] A value can also be obtained from the calculated valence bandwidth of an adenine stack; this gives 0.14 eV.[15] In the calculations of ref 6 the values of 0.2 and 0.1 eV were used, leading to polaron binding energies of 0.52 and 0.56 eV, respectively, and a width of $|\psi\psi^*|$ of ~5 sites. Repeating the calculations for $t_0 = 0.03$ eV, we obtained a polaron binding energy of 0.59 eV and a $|\psi\psi^*|$ width of 3 or 4 sites. The width of ψ itself is 4 or 5 sites, in agreement with the results of ref 7. The higher binding energy and smaller width for the smaller t_0 are expected because the solvation forces tend to localize the hole. As noted earlier, the calculated polaron width for an all-A polaron is in good agreement with measurements of O'Neill and Barton.[7,8]

For calculations involving bases other than As we need also the relative energies of the highest occupied molecular orbitals (HOMOs) of the different bases. One possibility is to use the measured values of the adiabatic ionization energies of the isolated bases in solution. Another possibility is to use for a given base the value calculated for a triplet of bases with the given base flanked by its actual neighbors in the sequence.[16,17] The justification stated for using the values obtained for triplets was that sequences with repeated G bases show higher reactivity toward oxidation than isolated G bases, and this has been attributed to the formation of delocalized hole traps at GG and GGG sequences.[18] Since our treatment is based on the wave functions being delocalized, it makes little sense to start with

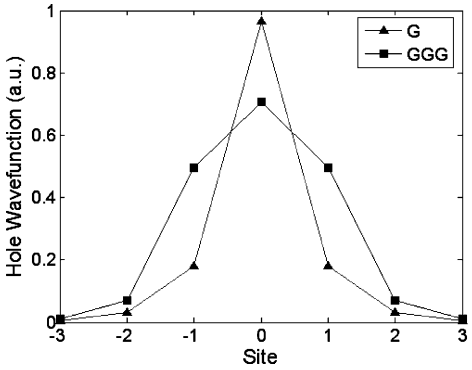

Figure 1. Hole wave function vs lattice site number for one guanine at site 0 (triangles) and three guanines at sites −1, 0, 1 (squares), in both cases surrounded by adenines.

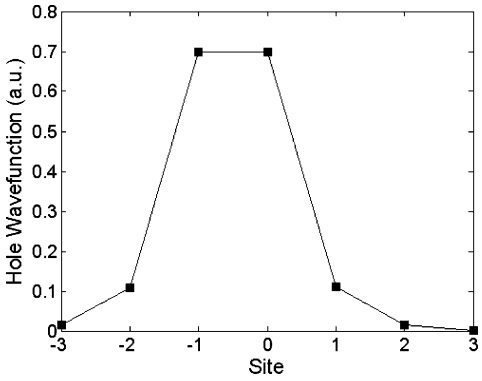

Figure 2. Hole wave function vs lattice site number for two guanines, at sites −1 and 0, surrounded by adenines.

values obtained from the wave function being partly delocalized over triplets of bases. Further, as will be seen, the triplet values lead to poor agreement with experiment. We therefore used for site energies the literature values of the adiabatic ionization energies of the isolated bases in solution. With the zero of energy chosen at the HOMO of A, the relative energies of the HOMOs were taken: −0.40 for guanine (G),[19] +0.70 for thymine (T),[20] +0.51 for cytosine (C).[20] Note, however, that the only really quantitative calculations of this paper are done for sequences with only Gs and As. Of the additional t_0 values needed, for a pair of adjacent Gs with normal B-DNA spacing and relative twist angle, as pointed out by Grozema et al.,[21] the calculated values cluster closely around 0.083 eV.[15,13,21] A value close to this, 0.089 eV, was calculated for t_0 of G−A neighbors.[13]

With the numbers just given we calculated the properties of the polarons arising from a hole on a single G, a pair of Gs, and a triplet of Gs, in all cases surrounded by As. We had done this calculation earlier using the distortion model of the polaron[22,23] and, of course, different parameters. The resulting solvated polarons are plotted in Figures 1 and 2. Our results can be compared with those of Kurnikov et al.,[5] who, as noted earlier, found the polaron for a sequence of 3Gs surrounded by As to be localized to 1 to 3 sites. For the distortion type of polaron we obtained a length of ∼6 sites for this sequence, while for the solvated polaron, as seen in Figure 1, we find most of the polaron on the central 3 sites, ψ having decreased by a factor of 10 on the fourth site. Thus our results are somewhat in disagreement with those of Kurnikov et al. The fact that this polaron is shorter than the all-A polaron is due to the deeper well provided by the Gs.

The results of calculations of hole wave functions for the alternating sequence $(GA)_{11}G$ surrounded by As are shown in Figure 3. The calculations were done for 100 sites with a G on every odd site from 39 to 62. Site 0 in the plot corresponds to site 50, at which there is an A. Initially the hole was placed at site 50. Using $t_0 = 0.1$ eV, as given in ref 13, we find that the wave function in this case extends over about 5 sites, peaking strongly at the central Gs as expected.

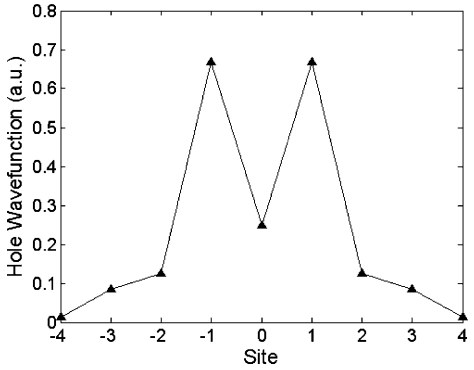

Figure 3. Hole wave function vs lattice site number for the sequence $(GA)_{11}G$ surrounded by As.

Comparison with Experiment

The above calculations make possible a comparison with experimental results of Lewis et al.[24] Lewis et al. have measured the free energy liberated when a hole goes from the radical cation G^+ to GG or GGG, in all cases surrounded by As. For G^+ to GG they obtained a free energy change of 0.052 eV, while for G^+ to GGG the energy difference is 0.077 eV. Using the distortion model of the polaron we were able to obtain good agreement with these results, provided the difference in on-site energies of A and G was taken as 0.2 eV[22,23] rather than the measured energy difference of ∼0.4 eV for isolated A and G in solution. At the time this discrepancy was considered acceptable, even to be expected, because Giese's interpretation of his experimental results for the sequence $G(A)_nGGG$ as showing thermal activation of holes from G to A for $n \geq 4$ [25] required an energy difference between G and A no greater than 0.2 eV. In detail, Giese's idea was that after tunneling through n As ($n \leq 4$) a hole would jump onto the bridge of As and go further by hopping between neighboring As. Bixon and Jortner found that Giese's model could not account for the observed very weak dependence on bridge size of the relative chemical yields and the ratio of the rates for tunneling and hopping.[26] Since that time an alternative explanation of Giese's data in terms of polarons has been advanced.[27,8] Using the solvation model of the polaron, with the on-site energy difference between G and A taken as 0.4 eV and t_0 values of 0.08 eV for both G−G and G−A, in line with the discussion above, we found the energy difference between the case of the hole on G and the hole on GG to be 0.055 eV and that between the hole on G and on GGG to be 0.075 eV, in excellent agreement with the experimental results of Lewis et al.[24] Calculations with the same value of t_0 and the on-site energy difference taken as −0.2 eV led to differences of 0.042 and 0.57 eV, while for the on-site energy difference taken as −0.3 eV they were 0.050 and 0.068 eV.

A different type of experiment with which we can compare our results is that in which the relative reactivity for oxidation of guanines with different neighbors is measured. In these experiments it is usual to expose the DNA to UV laser pulses and then treat it with Fpg (formamidopyrimidine) protein or piperidine, which creates strand breaks at the sites with strong oxidation. Such photocleavage experiments show that the reactivity of a given guanine depends very much on the

surrounding sequence. The strongly oxidized sites are expected to be those with the largest hole charge. For a GG pair flanked by Ts, i.e., the sequence TGGT, much stronger damage is found on the G at the 5′ site than on the G at the 3′ site,[28,29] indicating that the hole charge is more concentrated on the 5′ G. It has been stated many times in the literature that the charge is always more concentrated on the 5′ G, whatever the flanking bases. However, for the sequence AGGA (actually CAGGAT) Spassky and Angelov found that there was almost equal damage at the two Gs, that on the 3′ G being 5% larger, for either the Fpg treatment or the piperidine treatment.[28] Melvin et al. also found the two Gs about equally oxidized in the sequence CAGGAA.[30] (Saito et al.[29] do not give experimental data for this sequence.) Theoretical calculations of Senthilkumar et al.[17] for the AGGA sequence gave the fraction of the charge on the 5′ G 5 times as large as that on the 3′ G. We suggest that this discrepancy is mainly due to the use of the triplet calculations[17] to obtain the site energies, which makes the energy of the 5′ G, obtained from the triplet AGG, considerably lower than that of the 3′ G, obtained from GGA.

Our calculated results, shown in Figure 2, give an equal amplitude of the wave function on the two Gs, which would result in equal damage on the two sites, in close agreement with the experimental results of Spassky and Angelov. It might be objected, however, that our calculations would also give equal amplitude of the hole wave function on the two Gs for TGGT, there being no source for asymmetry. We suggest that what is left out of our calculations is the asymmetry between 3′ and 5′ Gs. Ab initio molecular orbital calculations of an isolated stacked pair of Gs with one electron missing gave the result that the hole wave function is primarily on the 5′ G of 5′-GG-3′.[31] When Ts or Cs are added on each side of the GG pair, because the energy of a hole on T or C is so much greater than the energy of a hole on G, there will be almost no hole amplitude on the added Ts or Cs. Taking the energy difference between an electron on G and one on T as 0.6 eV[32] and t_0 as the average of the values found for G−G and G−T by Voityuk et al.,[13] we find, using the polaron theory given above, that the amplitude of the hole wave function on the Ts is $\sim 1/100$ of that on the Gs. Thus addition of Ts or Cs should have little effect on the hole wave function, leaving it predominantly on the 5′ G. Flanking the GG pair with As should have a different effect because the energy of a hole on A is sufficiently close to that on G that the wave function spreads over AGGA. This should certainly reduce the asymmetry between the 5′ and 3′ Gs. The photocleavage results of Spassky and Angelov suggest that the asymmetry is wiped out and need not be taken into account in our calculations for AGGA.

For the GGG sequence the calculations of Saito et al., with Cs, Ts, or As flanking the GGG, gave the wave function of the hole, and thus the damage, greatest on the 5′ G, smaller on the middle G, and close to zero on the 3′ G.[29] Although experiments with Cs flanking the GGG gave results in agreement with these calculations, this was not the case for flanking Ts.[29] For the flanking T experiments, many groups found the damage greatest on the middle G and smaller for the two surrounding Gs.[29] (Saito et al. report no experiments with As flanking the GGG.) The calculations of Senthilkumar et al.[17] for GGG with flanking As gave the excess charge the largest on the central G, 65% as large on the 5′ G, and less than 10% as large on the 3′ G. Experiments by Spassky and Angelov[28] on the sequence AGGGA gave the damage on the central G largest and that on the two outside Gs ∼80% as large, with that on the 3′ G ∼5% larger than that on the 5′ G. Our calculations, as shown in Figure 1, gave the largest peak on the central G and equal peaks on the outside Gs, with charge ∼50% as large as that on the central G, thus in good agreement with the main features of the experimental results for this sequence. Apparently, with the hole spread over 3 Gs flanked by A or T, the 5′ G is no longer dominant. The fact that the 5′ G remains dominant with a flanking C can be attributed to the fact that the C is paired with another G, which, according to the calculations of Saito et al.,[29] has some hole amplitude.

Another point of contact with our results is the work of Schuster and colleagues on radical cation transport. With the idea of demonstrating polaron transport, Liu et al. investigated hole transport in a series of DNA oligomers made up of $[(A)_nGG]_m$ or $[(T)_nGG]_m$ segments, where $m = 4$ or 6 and $n = 1-7$ for the $(A)_n$ set, $1-5$ for the $(T)_n$ set.[33] For $n = 2$ the resulting sequence in the former case may, apart from an initial A, be written $(AGGA)_6$. In accounting for the variation of the hopping rate with n, Schuster, in his words, "arbitrarily suggests" that AGGA is the polaron and for $n = 2$ the hole hops from one AGGA to an adjacent one, the hopping rate being determined by the motions of the DNA and its environment.[33] The calculations we carried out above justify his assignment of the polaron to the AGGA sequence, the length of the polaron for the case of two Gs surrounded by As being 4 sites, as shown in Figure 2. For $n = 1$ the sequence of Liu et al. is a series of AGGs. Taking AGGA as the polaron, one can also account for the transport as a series of hops from one AGGA to the next, with the difference that in this case the hop is shorter, the last A of one polaron forming the first A of the next polaron. Liu et al. noted that hole transport is most rapid for the $n = 1$ and $n = 2$ $(A)_n$ sequences and decreases for $n > 2$, where there are extra As between the AGGAs.[33] This can also be seen as an example of conformationally gated hopping.[8]

Sequences with Thymine and Cytosine as well as Guanine and Adenine

Apart from the propagation being slower, the properties of the sequences $[(T)_nGG]_m$ found by Liu et al. are quite similar to those of the $[(A)_nGG]_m$ sequences.[33] This suggests that polarons are also formed for the $[(T)_nGG]_m$ sequences, where by analogy the polaron would be TGGT. However, as discussed above, the energy difference between an electron on G and one on T is so large that the hole is essentially localized on the Gs, and it is not reasonable to consider TGGT a polaron. In general, for a complementary pair A/T most of the hole wave function should be on A rather than T because the energy on A is much lower. It must be concluded that the polaron involved in the hopping in the case of $[(T)_nGG]_m$ segments occupies both sides of the duplex. Thus, the similarities in the behaviors of $[(T)_nGG]_m$ and $[(A)_nGG]_m$ lead to the conclusion that the polaron in the former case consists of the duplex TGGT/ACCA. The fact that some of the hole wave function is on the As means that the asymmetry of hole occupation of the 3′ and 5′ Gs may be decreased somewhat in the duplex polaron but probably will still have considerable effect. Of course, in the case of the polaron AGGA a small part of the polaron wave function must also reside on the complementary strand, but it should be small enough that calculations neglecting this are reasonably accurate.

There are in the literature some studies in which holes have been found to go through sequences where the closest Gs (taking into account both sides of the duplex) have been separated by 4, 6, 8, or 10 ATs or TTs. Proof of the passage of the holes has been the observation of oxidative damage to proximal GG pairs located just before the beginning of the AT or TT sequence

and to distal GG pairs just after the end. In one set of such studies the distal/proximal oxidation ratios were in general greater than unity,[34] which is rather surprising. It was later suggested that the ratios being greater than unity was the effect of relatively rapid back recombination depleting the proximal pair.[35] The significant point, however, is that the penetration of so many Ts indicates that the transport was not due to tunneling. We suggest that the transport mechanism in these long AT or TT segments is polaron hopping. By the arguments given above, the polarons involved must also represent examples of duplex polarons, consisting of ATAT/TATA or TTTT/AAAA.

Electron Polarons

The theory presented above applies to electrons as well as to holes; it is expected therefore that under appropriate conditions an excess electron will form a polaron. Because thymine and cytosine are the most easily reduced bases, electron polarons should be concentrated on these bases rather than on guanine or adenine. It has been suggested, however, that the tendency of cytosine to gain a proton when it accepts an electron[36] makes it doubtful that it could participate in electron conduction. In what might be called the inverse case, when a guanine is oxidized, it has a tendency to lose a proton. That protons are involved in hole transfer in DNA was verified by the finding of a decrease in charge-transfer efficiency when all acidic protons were replaced by deuterons.[37] This showed, as was implicit in earlier sections of this paper, that deprotonation does not cut off hole conduction in GA sequences properly hybridized.[37] Giese and Wessely attributed the persistence of hole conduction in this case to the deprotonation taking place by the shifting of the proton from nitrogen toward the complementary C, which they suggested stabilizes the hole. If protonation of cytosine on its accepting an electron were to take place by a shift of a proton from the complementary guanine it is possible that cytosine could contribute to electron conduction. Be that as it may, deprotonation is not likely for thymine, so mobile electron polarons are expected in that case.

The properties of electron polarons cannot be predicted as well as those of hole polarons at this time because the t_0 values and the relative energies of an electron on the different bases in water are less well-known. The value of t_0 for an electron on a pair of Ts obtained from the calculated width of the conduction band of a thymine stack is 0.04 eV.[14] Corresponding to this the length of an electron polaron on an all-T sequence would be ~4 or 5 bases. For sequences $(T)_n$ surrounded on both sides by As the properties of the electron polaron would be rather similar to those of the hole polaron in sequences $(G)_n$ surrounded on both sides by As.

Conclusions

Because there is a finite transfer integral the wave function of an electron or hole in DNA in solution will tend to be delocalized, whatever the sequence. Its interaction with water causes the electron or hole to be a polaron, with a wave function such as those described above. The polaron width and binding energy are determined by the transfer integrals of the bases included and the relative energies of the hole or electron on these bases. Typical polaron widths are 4−6 bases. Smaller t_0 leads to increased binding energy and decreased width because the solvation effect, which tends to localize the hole, is relatively greater. The theory for the solvated polaron leads to excellent agreement with the values measured by Lewis et al. for the energy differences between a hole on G or GG or GGG, in all cases surrounded by As. The parameters involved are now sufficiently well-known that this agreement represents a significant endorsement of the theory. The theory is also in substantial agreement with the observed oxidation by injected holes of GG and GGG sequences surrounded by As.

It is implicit in most of the foregoing that the polaron wave function can with good accuracy be treated as if it were entirely on a single strand of the duplex. This should be true for a hole polaron if the sequence involves only Gs and As. If the sequence includes Ts or Cs, on which the energy of a hole is much greater than on G or A, then the polaron must be described as a duplex. This does not appear to affect the width of the polaron. It does, however, decrease its hopping rate, not surprising because the difficulty of matching the hopped-to wave function with the original wave function, on which the hopping rate depends, should be larger than if the polaron were essentially limited to one strand.

Polaron transport may be by drift or hopping. Polaron drift is characterized by the polaron as it moves taking on a base in the direction of motion and dropping a base at the other end. It can only occur in sequences with all bases the same, e.g., all As. Examples of such motion are seen in ref 7 and, we suggest, in refs 25, 38, and 39, where it is referred to as A hopping. Polaron hopping could also occur in sequences with all bases the same. As demonstrated by Schuster,[33] polaron hopping also occurs in periodically repeated sequences. The period may range from 2 bases, e.g., in GAGAGA..., to a length not much greater than the polaron width.

The question arises as to how extensive is charge transport by polaron hopping on DNA that does not have periodically repeated sequences. Calculations for a hole on an arbitrary DNA sequence, which may be random, in water show that a delocalized wave function can always be formed because its energy is lower than that of a hole localized on a single site. Hopping of a polaron from the set of bases over which it is delocalized to another set is in principle always possible if the hop is energetically allowed. Energetic considerations may, of course, determine the set of possible sites to which the polaron may hop. Additional energy, if needed, may be supplied by the thermal energy of the bases. In any case, if the wave functions on the possible final sites are a poor match to that on the original or the energy differences are considerable, then hopping may be very slow. In that case, another process, such as the tunneling or G hopping envisioned by Giese,[40] may be more rapid.

At first sight it might appear that the difference between polaron wave functions resulting from a couple of bases or more being different would be so large as to make hopping very slow. However, this neglects the effect of thermal motions on the two quantities that determine the polaron wave function: the relative site energies and the coupling between pairs, t_0. Voityuk et al., calculating for a specific sequence, found that thermal fluctuations of the water molecules and counterions have a considerable effect on the relative energy of a G and an A; they can be so large as to make the A energy instanteously lower than that of G.[41] Voityuk et al. have also shown that the thermal motion of the bases can have strong effects on the electronic coupling between the bases.[42] Although the effects are likely to be smaller under the stabilizing effect of polaron formation, they might be considerable. An idea of how strongly they affect room-temperature transport may be obtained from the experiments of O'Neill and Barton. Measuring the fraction of fluorescence quenching of excited aminopurine, Ap*, in the sequence $Ap(A)_nG$, they found a strong increase with increasing temperature from ~30 °C to the denaturation temperature, ~ 60 K.[7] The temperature dependence of oxidative damage induced by

Ap* in the DNA duplex $Ap(A)_3{}^{CP}G$ was found to increase from 0.36% at 10 K to 0.54% at 30 K, indicating an increase in charge transport with temperature even below room temperature. We conclude that it is not possible in the present state of our knowledge about thermal fluctuations to predict how much of the transport in arbitrary sequences of DNA will be due to polaron hopping.

Acknowledgment. S.M.B. is grateful for the support of The Camille and Henry Dreyfus Foundation. We are also grateful to Professor H.-Y. Choi and to J.-H. Park for illuminating discussions.

References and Notes

(1) Emin, D.; Holstein, T. *Phys. Rev. Lett.* **1976**, *36*, 323−326.
(2) Shao, F.; O′Neill, M. A.; Barton, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17914−17919.
(3) Henderson, P. T.; Jones, D.; Hampikian, G.; Kan, Y.; Schuster, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8353−8358.
(4) Conwell, E. M.; Rakhmanova, S. V. *Proc. Natl. Acad. Sci. U.S.A.* **2000,** *97*, 4556−4560.
(5) Kurnikov, I. V.; Tong, G. S. M.; Madrid, M.; Beratan, D. N. *J. Phys. Chem. B* **2002,** *106*, 7−10.
(6) Basko, D. M.; Conwell, E. M. *Phys. Rev. Lett.* **2002**, *88*, 098102.
(7) O'Neill, M. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 11471−11483.
(8) Conwell, E. M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8795−8799.
(9) Basko, D. M.; Conwell, E. M. *Phys. Rev. B* **2002**, *66*, 094304.
(10) It should be noted that the strong similarity of the polaron wave function due to distortion and that due to the water and ions that is seen in Figure 1 of ref 6 is an accident of the choice of the parameters for the distortion case. If the transfer integral had been chosen smaller, such as the values found in more recent calculations, then the wave function of the distortion polaron would have been much more extended.
(11) It has been suggested that ions residing in the grooves of the helix have an important effect, acting to gate the motion of the hole.[12] Experiments of O'Neill and Barton, discussed at length in ref 7, establish that they did not see such effects due to the ions, which suggests they are not usual.
(12) Barnett, R. N.; Cleveland, C. L.; Joy, A.; Landman, U.; Schuster, G. B. *Science* **2001**, *294*, 567−571.
(13) Voityuk, A. A.; Rösch, N.; Bixon, M.; Jortner, J. *J. Phys. Chem. B* **2000**, *104*, 9740−9745.
(14) Troisi, A.; Orlandi, G. *Chem. Phys. Lett.* **2001**, *344*, 509−518.
(15) Zhang, M.-L.; Miao, M. S.; Van Doren, V. E.; Ladik, J. J.; Mintmire, J. W. *J. Chem. Phys.* **1999**, *111*, 8696−8700.
(16) Voityuk, A. A.; Jortner, J.; Bixon, M.; Rösch, N. *Chem. Phys. Lett.* **2000,** *324*, 430−434.
(17) Senthilkumar, K.; Grozema, F. C.; Fonseca Guerra, C.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2003**, *125*, 13658−13659.
(18) Saito, I.; Nakamura, T.; Nakatani, K. *J. Am. Chem. Soc.* **2000**, *122*, 3001−3006.
(19) Steenken, S.; Jovanovic, S. C. *J. Am. Chem. Soc.* **1997**, *119*, 617−618.
(20) Orlov, V. M.; Smirnov, A. N.; Varshavsky, Ya. M. *Tetrahedron Lett.* **1976**, *48*, 4377−4378.
(21) Grozema, F. C.; Siebbeles, L. D. A.; Berlin, Yu. A.; Ratner, M. A. *ChemPhysChem* **2002**, *3*, 536−539.
(22) Conwell, E. M.; Basko, D. M. *J. Am. Chem. Soc.* **2001**, *123*, 11441−11445.
(23) Park, J.-H.; Choi, H.-Y.; Conwell, E. M. *J. Phys. Chem. B* **2004**, *108*, 19483−19486.
(24) Lewis, F. D.; Liu, X.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2000**, *122*, 12037−12038.
(25) Giese, B.; Amaudrut, J.; Köhler, A.-K.; Spormann, M.; Wessely, S. *Nature* **2001**, *412*, 318−320.
(26) Bixon, M.; Jortner, *J. Chem. Phys.* **2002**, *281*, 393−408.
(27) Conwell, E. M.; Park, J.-H.; Choi, H.-Y. *J. Phys. Chem. B* **2005**, *109*, 9760−9763,
(28) Spassky, A.; Angelov, D. *Biochemistry* **1997**, *36*, 6571−6576.
(29) Saito, I.; Nakamura, T.; Nakatani, K. *J. Am. Chem. Soc.* **2000**, *122*, 3001−3006.
(30) Melvin, T.; Plumb, M. A.; Botchway, S. W.; O'Neill, P.; Parker, A. W. *Photochem. Photobiol.* **1995**, *61*, 584−591.
(31) Sugiyama, H.; Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 7063−7068.
(32) Seidel, C. A. M.; Schulz, A.; Sauer, H. M. *J. Phys. Chem.* **1996**, *100*, 5541−5553.
(33) Liu, C.-S.; Hernandez, R.; Schuster, G. B. *J. Am. Chem. Soc.* **2004**, *126*, 2877−2884.
(34) Williams, T. T.; Odom, D. T.; Barton, J. K. *J. Am. Chem. Soc.* **2000**, *122*, 9048−9049.
(35) Williams, T. T.; Dohno, C.; Stemp, E. D. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 8148−8158.
(36) Steenken, S. *Free Radical Res. Commun.* **1992**, *16*, 349−379.
(37) Giese, B.; Wessely, S. *Chem. Commun.* **2001**, 2108−2109.
(38) Kawai, K.; Takada, T.; Tojo, S.; Majima, T. *J. Am. Chem. Soc.* **2003**, *125*, 6842−6843.
(39) Takada, T. D.; Kawai, K.; Cai, X.; Sugimoto, A.; Fujitsuka, M.; Majima, T. *J. Am. Chem. Soc.* **2004**, *126*, 1125−1129.
(40) Giese, B. *Acc. Chem. Res.* **2000**, *33*, 631−636.
(41) Voityuk, A. A.; Siriwong, K.; Rösch, N. *Angew. Chem., Int. Ed.* **2004**, *43*, 624−627.
(42) Voityuk, A. A.; Siriwong, K.; Rösch, N. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5421−5425.