Hole Traps in DNA Calculated with Exponential Electron−Lattice Coupling

Ji-Hoon Park,[†] Han-Yong Choi,*[,†,‡] and Esther M. Conwell[§]

Department of Physics, BK21 Physics Research Division, and Institute for Basic Science Research, Sung Kyun Kwan University, Suwon 440-746, Korea, Asia Pacific Center for Theoretical Physics, Pohang 790-784, Korea, and Department of Chemistry, University of Rochester, Rochester, New York 14627

Received: July 8, 2004; In Final Form: September 10, 2004

Pairs or triples of guanine, G, on DNA are readily oxidized, forming hole traps. Lewis et al. (*J. Am. Chem. Soc.* **2000**, ***122***, 12037) have measured the free energy liberated when a hole goes from the radical cation G^+ to GG or GGG. We calculated these free energies for the sequences used experimentally using a simple tight-binding model in which the transfer integral was assumed to vary linearly with the spacing between adjacent bases (*J. Am. Chem. Soc.* **2001**, ***123***, 11441). Later calculations of the transfer integral indicated that it is smaller than the value we used and varies exponentially with the spacing. We have recalculated using this information, again taking into account polaron formation. Good agreement is again obtained with Lewis et al. Because the measured free energies are the differences of two energies, the effects of the environment must essentially cancel; the agreement we have found with the present calculations is thus good evidence for the hole wave function being delocalized over a number of bases.

There has been considerable discussion as to whether the wave function of an extra electron or hole on a DNA molecule is confined to a single base or spread over a number of bases as a polaron. In early calculations, where the effect of the surrounding medium was not included, the Su−Schrieffer−Heeger (SSH) Hamiltonian[1] was used to find the properties of the polaron.[2,3] It was found that, with the parameters chosen, the polaron is spread out over ∼5 sites and is stable at room temperature, although its binding energy is small. The delocalization is the result of the kinetic energy of the polaron tending to make it spread, while the distortion of the DNA (decrease in base spacing within the polaron) tends to confine it.

Subsequently, more realistic calculations were done, again with the SSH Hamiltonian, but with the DNA surrounded by water and ions. Due to the polarizability of the surrounding medium the binding energy of the polaron increased to ∼0.6 eV although the extent of the polaron was still ∼5 sites.[4] In a rather different calculation, Kurnikov et al., noting that water would counteract the tendency of the kinetic energy to make the hole wave function spread, found that an excess hole on DNA would be confined to 1 to 2.5 sites.[5] Their calculations were based on the B DNA geometry being maintained (i.e., no chain distortion) and the charge being spread uniformly over the guanines.

A good test for resolving the question of the extent of the excess hole wave function is a comparison of the theoretical prediction with experimental results such as those of Lewis et al. on the dynamics of an extra hole on DNA.[6] Lewis et al. made measurements on a hairpin incorporating 1, 2, or 3 guanines (G's) surrounded by adenines (A's). They found the magnitude of the free energy change due to transfer of a hole from G to GG, Δ_{GG}, to be 0.052 ± 0.006 eV, while that for transfer from G to GGG, Δ_{GGG}, is 0.077 ± 0.005 eV. Thus multiple G's are shallow traps.

Calculations of Δ_{GG} and Δ_{GGG} were carried out neglecting the surrounding medium. Although the effect of the environment is not small, as was seen in the large increase in binding energy of the polaron mentioned above, it is difficult to believe that its neglect can lead to significant error in Δ_{GG} or Δ_{GGG}. What is being calculated in those cases is the difference of two energies, both including the effect of the environment, but differing only in the replacement of one or two A/T's (where T is thymine) with one or two G/C's, respectively. Calculations carried out for the hole assumed confined to a single base led to 0.3 to 0.13 eV for either Δ_{GG} or Δ_{GGG}, depending on the nature of the surrounding bases.[7] Much better agreement with experiment was obtained with our calculations based on the polaron effect.[3] Values of the transfer integral t_0 for the bases separated by their normal distance in B DNA, 3.4 Å, of 0.2 and 0.3 eV were used and excellent agreement with the results of Lewis et al. was obtained with the difference in ionization potentials of A and G taken as 0.17 eV.[3] This difference is about half the difference measured for isolated A and G in solution.[8,9] But, as has been noted many times, the ionization potential of a base can be much affected by its neighbors. In any case, an upper limit of ∼0.2 eV on the difference in ionization potentials of A and G is set by the finding that a hole can surmount the barrier between G and A with thermal energy at room temperature.[10]

The SSH Hamiltonian incorporates the coupling of nearest neighbors through a transfer integral *t*, which varies linearly with their separation. This form of coupling is reasonable for close neighbors, as in polyacetylene with a separation between nearest neighbors of 1.2 Å, where the overlap does not vary strongly with separation. For the bases in DNA with a separation of 3.4 Å, where the overlap involves the tails of the wave functions, an exponential dependence on base separation is more

* Corresponding author. Phone: 82-31-290-7048. Fax: 82-31-290-7055. E-mail: hychoi@skku.ac.kr.
† Sung Kyun Kwan University.
‡ Asia Pacific Center for Theoretical Physics.
§ University of Rochester.

10.1021/jp046968p CCC: $27.50
Published on Web 11/20/2004

reasonable. Voityuk et al. have found that for DNA the transfer integral varies exponentially with the separation. Specifically, for the duplex sequence AT/AT they found the variation of t with base separation R is well approximated by 0.027 exp[$-a(R - 3.38)$], where $a = 2.0$ Å^{-1}.[11] This corresponds to a considerably smaller transfer integral for the normal B DNA spacing than the 0.2 or 0.3 eV used in our earlier calculations. Other recent calculations of the transfer integral have also obtained smaller values, e.g., 0.08 eV.[12,13] On the basis of these results, we have carried out new calculations of the difference in free energies for a hole on G vs GG or GGG, in all cases surrounded by A's as in the experiments of Lewis et al., with the transfer integral given by

$$t = t_0 \exp[-2.0(R - 3.38)] \tag{1}$$

with t_0 taken as a parameter. As will be shown, for t_0 values of 0.075 eV we again get good agreement with the experimental free energy results of Lewis et al.

The SSH Hamiltonian for a 1-dimensional (1-D) electron−lattice coupled system with a free boundary condition is given by

$$H = \sum_{l=1}^{N}\left[-t_l(c_l^{+}c_{l+1} + c_{l+1}^{+}c_l) + V_l c_l^{+}c_l + \frac{1}{2}M\dot{u}_l^{2}\right] + \sum_{l=1}^{N-1}\frac{1}{2}Ky_l^{2} \tag{2}$$

where c_l^{+} (c_l) is the electron creation (annihilation) operator at the site l, V_l the energy of the lth site, t_l the transfer integral between l and $l + 1$ sites, $y_l = u_{l+1} - u_l$, u_l the displacement of the l-th site from its equilibrium position, N the number of sites, K the elastic constant, and M the mass of the entity at each site. The spin index is suppressed because the physical properties, like the polaron binding energy, we consider in this paper are independent of spin. The transfer integral is usually taken to vary linearly as

$$t_l = t_0 - \alpha y_l \tag{3}$$

where α is the electron−lattice coupling constant. But, for the reasons given above, we take the transfer integral to be exponential in form

$$t_l = t_0 \exp\left(-\frac{\alpha}{t_0}y_l\right) \tag{4}$$

which reduces to the linear form of eq 3 for small y_l.

The Schrödinger equation for the nth eigenstate of the Hamiltonian of eq 2 with the ionic kinetic part neglected can be written as

$$\epsilon_n\psi_n(l) = V_l\psi_n(l) - t_l\psi_n(l+1) - t_{l-1}\psi_n(l-1) \tag{5}$$

where ϵ_n is the energy of the nth one-particle state and $\psi_n(l)$ is the amplitude of the nth state wave function at the site l. The total energy is the sum of the electron and lattice parts given as

$$E_{\mathrm{tot}} = 2\sum_{n}{}'\epsilon_n + \frac{1}{2}K\sum_{l=1}^{N-1}y_l^{2} \tag{6}$$

where the factor of 2 is from the spin degeneracy and the prime indicates that the summation is over the occupied states. A minimization of the total energy leads to the following self-consistency equation:

$$y_l = -2\frac{\alpha}{K}\exp\left(-\frac{\alpha}{t_0}y_l\right)\sum_{n}{}'[\psi_n^{*}(l)\psi_n(l+1) + \psi_n^{*}(l+1)\psi_n(l)] \tag{7}$$

For the calculations the parameter α/t_0 was taken as 2.0 Å^{-1} and K as 0.85 eV/Å^2, as in ref 2. The on-site energy for adenine was taken as $V_A = 0$, while that for guanine V_G was taken as a parameter. When there were 2 or 3 adjacent G's they were taken as having the same V_G. There is a great deal of experimental data showing that the 5′G in a GG sequence is preferentially oxidized, consistent with the finding of many theoretical calculations that the 5′G has lower energy. However, as is well-known, the relative energies of two adjacent G's depend on the neighboring bases, particularly the nearest neighbors. The finding that the 5′G is preferentially oxidized holds for many sets of nearest neighbors but not all. For the sequence AGGA O'Neill et al. find about equal damage induced on each of the Gs by 193 nm light followed by Fpg protein treatment.[14] An earlier measurement of the ratio of 5′ to 3′ reactivity for the sequence CAGGAT under piperidine cleavage gave the result that the 3′G is slightly more reactive than the 5′G.[15] There is a recent theoretical calculation that finds the 5′G much more reactive than the 3′G in the sequence AGGA.[16] Given the uncertainties that still remain in theoretical calculations of such quantities, we have chosen to weight the experimental results more heavily and take V_G the same for adjacent G's.

TABLE 1: The "Polaron Binding Energies" of G, GG, and GGG (in eV) with $t_0 = 0.05$ and 0.075 eV and $V_G = 0.17$ and 0.20 eV Relative to V_A[a]

t_0 (eV)	V_G (eV)	G	GG	GGG
0.075	0.20	0.1107	0.1665 (0.0558)	0.1833 (0.0726)
0.075	0.17	0.0888	0.1384 (0.0496)	0.1542 (0.0654)
0.050	0.20	0.1261	0.1677 (0.0416)	0.1825 (0.0564)
0.050	0.17	0.1004	0.1390 (0.0386)	0.1531 (0.0527)

[a] The binding energies of GG and GGG with respect to G are given in parentheses for GG and GGG.

To compare our calculations with the results of Lewis et al., we performed calculations for sequences similar to those used in their experiments. We took 100 sites along a 1-D lattice with a free boundary condition where a G, GG, or GGG is surrounded by A's with 199 electrons (or 1 hole). For a given sequence of A's and G's, we solved eqs 5 and 7 self-consistently via numerical iterations. The solutions converge after a few tens of iterations and produce the eigen energies, wave functions, and equilibrium lattice configurations.

We present our results in terms of the "polaron binding energy" and shape. The "polaron binding energy" in Table 1 is the difference between the total energy for a given sequence and the reference energy of $-2t_0$. The reference energy corresponds to the total energy of a 1-D lattice of the same number of sites with $\alpha = V_l = 0$. The "polaron binding energy" represents contributions from the lattice relaxation and the difference in on-site potentials of the G and A. It does not include the polarization of the surrounding medium by the charge on the polaron, which increases the binding energy by ~0.5 eV.[4] As discussed earlier, the latter contribution cancels out when taking the binding energy differences between G and GG or GGG to compare with experiment.

In Table 1, we list the "polaron binding energies" for G, GG, and GGG surrounded by A's. We considered $t_0 = 0.05$ and

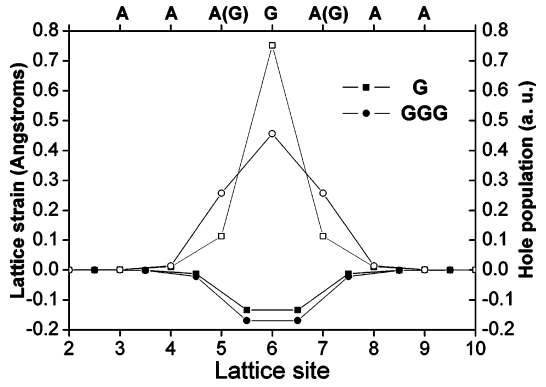

Figure 1. The hole populations $|\psi_n(l)|^2$ (open symbol) and lattice strain $y_l = u_{l+1} - u_l$ (filled symbols), in Å, plotted at $l + {}^1/_2$ for G (squares) and GGG (circles) surrounded by A's with $t_0 = 0.075$ eV, $V_A = 0$, $V_G = 0.2$ eV, $\alpha = 0.15$ eV/Å, and $K = 0.85$ eV/Å^2.

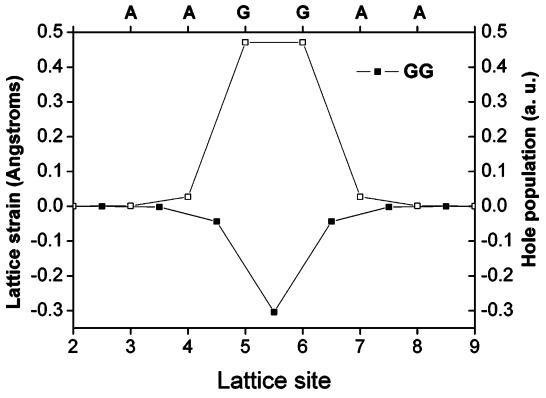

Figure 2. The hole populations $|\psi_n(l)|^2$ (open) and lattice strain (filled) for GG surrounded by A's with the same parameters as Figure. 1.

0.075 eV and $V_G = 0.17$ and 0.20 eV. The numbers inside the parentheses for GG and GGG are the binding energies (in eV) relative to the single G. As noted above, Lewis et al. reported that the binding energy difference between GG and G is 0.052 ± 0.006 eV, and that between GGG and G is 0.077 ± 0.005 eV. Our calculation with $t_0 = 0.075$ and $V_G = 0.2$ eV gives 0.0558 and 0.0726 eV, respectively, in good agreement with the Lewis results. Clearly, still better agreement could have been obtained with small variations of the parameters.

We now discuss the case of $t_0 = 0.075$ and $V_G = 0.2$ eV in more detail. In Figure 1, we show the lattice configuration for G and GGG surrounded by A's by plotting $y_l = u_{l+1} - u_l$ at $l + {}^1/_2$ and the hole populations $|\psi_n(l)|^2$ as a function of l. The polaron is approximately 4 to 5 sites wide, about as found in the calculations of ref 3. The GGG has bigger strains than the single G. In Figure 2, we show the GG case with the same parameters as used for Figure 1.

It is seen that the polaron model, with the dependence on base spacing of the transfer integral, changed from linear to exponential form to agree with the calculations of ref 11, results in very good agreement with the measured differences in free energy of G, GG, and GGG traps. The good agreement is obtained with a value of t_0 much closer to recently calculated values, although somewhat larger. On the basis of the fact that the various degrees of freedom of DNA are not independent of each other, and the finding of Voityuk et al. that the transfer integral is quite sensitive to the other possible motions of DNA, it is reasonable that the t_0 value be larger than they found for "rise" only. It is also seen that the maximum strains are smaller than were found for the linear dependence of t on base spacing: 0.17 Å vs 0.3 Å for GGG, 0.3 Å vs 0.55 Å for GG. Smaller strains appear more reasonable physically.

As another check on our results we can compare the hole populations $|\psi_n(l)|^2$ on the three G sites in GGG. According to Figure 1 there are equal populations on the two outer G's, at sites 5 and 7, while the population at the central G, at site 6, is ~1.7 times as large. This is in very good agreement with the ratios seen in piperidine cleavage of the sequence AGGGA in ref 15. There also the two outer G's have equal populations, while that of the central G is 1.65 times as large. In contrast the calculations of ref 16 give the relative populations as 38: 59:5 for the three G's.

Calculations to determine whether the wave function of an extra hole is localized on a single base or delocalized have also been carried out by Olofsson and Larsson[17] and Cramer, Krapf, and Koslowski.[18] In the former case the investigators use as the criterion for localization that the reorganization energy λ be ~4 or more times the coupling between a pair of nearest neighbor bases. On this basis they conclude that the wave function of a hole in a series of cytosines or of thymines or of alternating adenines and thymines would be delocalized, while a series of guanines or adenines would have weakly trapped holes. Their calculations include the effects of changes in bond length, which ours do not. However, their calculations refer to zero temperature; they neglect chain vibrations, which of course we include, and disorder resulting from them. They also neglect the important outer-sphere contribution of the surrounding water.[4] We can do this in the calculations of this paper because we are calculating the differences of two energies where the effect of the water should essentially cancel.

In ref 17 the authors note that their calculated couplings, admittedly difficult to calculate accurately, are comparable to those of Voityuk et al. of ref 7. In that reference, however, the coupling between adjacent bases was not calculated. Rather, what was calculated was an "ionization energy", the difference in energy between a triple of bases centered on the final site and another triple centered on the original site. As we noted earlier, in later papers Voityuk et al., and others calculating for a pair of adjacent bases, found smaller values for the coupling.[11,12] In another early calculation Sugiyama and Saito[19] calculated energy as a function of separation for a pair of guanines assumed exactly overlapped. At a separation of 0.34 nm the two levels the pair gave rise to were separated by 0.72 eV, leading to a coupling of 0.36 eV on a Huckel model. Of course this overestimated the coupling because adjacent Gs are not overlapped in the helical structure of DNA.

Cramer, Krapf, and Koslowski[18] have calculated wave functions of an extra hole using the Su−Schrieffer−Heeger Hamiltonian in an atomic parametrization, extended to include dielectric polarization effects. They considered a B DNA sequence of the length found in a nucleosome (146 base pairs). In this sequence they found, for well-separated G's surrounded by other bases, that an excess hole wave function was centered on a G but extended to other bases with a half-width of approximately one nuclear base. Our earlier calculations, ref 3, are in good agreement with their result, as indeed they point out. As seen in Figures 1 and 2, the extent of the wave functions obtained in our present calculations is quite similar, as were our results also when water was included.[4]

In summary, the fact that the polaron model, with reasonable values of the parameters, is able to account well for the measured free energy differences of the hole traps, while the model based on the hole being localized on a single base gives poor results

for these energies, is good evidence that an extra hole or electron on DNA is not confined to a single site.

Acknowledgment. Financial support was provided by the Korea Science & Engineering Foundation (KOSEF), through grant No. R01-1999-000-00031-0, and by the Ministry of Education, through the BK21 SNU-SKKU Physics program (J.H.P. and H.Y.C.).

References and Notes

(1) Su, W. P.; Schrieffer, J. R.; Heeger, A. J. *Phys. Rev. Lett*. **1979**, *42*, 1698−1701.

(2) Conwell, E. M.; Rakhmanova, S. V. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 4556−4560.

(3) Conwell, E. M.; Basko, D. M. *J. Am. Chem. Soc*. **2001**, *123*, 11441−11445.

(4) Basko, D. M.; Conwell, E. M. *Phys. Rev. Lett*. **2002**, 88, 098102-1-4.

(5) Kurnikov, I. V.; Tong, G. S. M.; Madrid, M.; Beratan, D. N. *J. Phys. Chem. B* **2002**, *106*, 7−10.

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

(7) Voityuk, A. A.; Jortner, J.; Bixon, M.; Rösch, N. *Chem. Phys. Lett*. **2000**, *324*, 430−434.

(8) Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. *J. Phys. Chem*. **1996**, *100*, 5541−5553.

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

(10) Giese, B.; Spichty, M. *ChemPhysChem* **2000**, *1*, 195−198.

(11) Voityuk, A. A.; Siriwong, K.; Rösch, N. *Phys. Chem. Chem. Phys*. **2001**, *3*, 5421−5425.

(12) Grozema, F. C.; Siebbeles, L. D. A.; Berlin, Y. A.; Ratner, M. A. *ChemPhysChem* **2002**, *3*, 536−539.

(13) Zhang, M.-L.; Miao, M. S.; Doren, V. E. V.; Ladik, J. J.; Mintmire, J. W. *J. Chem. Phys.* **1999**, *111*, 8696−8700.

(14) O'Neill, P.; Parker, A. W.; Plumb, M. A.; Siebbeles, L. D. A. *J. Phys. Chem. B* **2001**, *105*, 5283−5290.

(15) Spassky, A.; Angelov, D. *Biochemistry* **1997**, *36*, 6571−6576.

(16) Senthilkumar, K.; Grozema, F. C.; Fonseca Guerra, C.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Am. Chem. Soc*. **2003**, *125*, 13658−13659.

(17) Olofsson, J.; Larsson, S. *J. Phys. Chem. B* **2001**, *105*, 10398−10406.

(18) Cramer, T.; Krapf, S.; Koslowski, T. *Phys. Chem. Chem. Phys*. **2004**, *6*, 3160−3166.

(19) Sugiyama, H.; Saito, I. *J. Am. Chem. Soc*. **1996**, *118*, 7063−7068.