Hole Traps in DNA

E. M. Conwell* and D. M. Basko

Contribution from the Department of Chemistry, University of Rochester, Rochester, New York 14627

Received April 4, 2001. Revised Manuscript Received July 13, 2001

Abstract: Sequences of guanines, GG and GGG, are known to be readily oxidized, forming radical cations, i.e., hole traps, on DNA. The trapping probability of GG is less than that of GGG. Lewis et al. (*J. Am. Chem. Soc.* **2000**, *122*, 12037) have used measurements on synthetic hairpins to determine the free energy liberated when a hole goes from the radical cation G^+ to GG or to GGG. They find these free energies to be of the order of thermal energy at room temperature, in contradiction to the expectation by many of much greater trap depths. We have calculated the wave function of a hole on G, on GG, and on GGG surrounded by adenines, as in the Lewis et al. experiments, using a simple tight-binding model. We find that to account for the shallow traps found by them it is necessary that the difference in ionization potentials of contiguous guanine and adenine be smaller by about 0.2 eV than the 0.4 eV found for isolated bases. Using this value and taking into account polaron formation, we find the wave functions of holes trapped on G, GG, or GGG to extend over ∼6 sites (bases) and with energy level differences in good agreement with the values found by Lewis et al.

I. Introduction

The motion of a hole (radical cation) on DNA has been studied by observing its trapping at a series of sites with low ionization potential incorporated in the DNA. Hole traps commonly used for this purpose are the sequences GG and GGG. To detect the trapping use is made of the fact that the radical cation may react irreversibly with water or oxygen, resulting, with perhaps further chemical treatment, in cleavage of the DNA at the site of the trapped hole. With this technique it has been shown that, on a DNA strand with a series of GG units separated by a number of other bases, injected holes can travel distances of ∼100 Å with only a small percentage trapped at each individual GG.[1,2] It is considered that the relative reactivity of a hole trap can be determined by densitometric assay of the cleavage bands seen in high-resolution polyacrylamide gel electrophoresis. Measurements of relative reactivity made with this technique show that the GGG traps are more reactive, although not greatly so, than GG. For a duplex DNA containing a G, a GG, and a GGG, Hickerson et al. report a cleavage ratio of 1:3.7:5.3, respectively.[3] These numbers could vary somewhat, depending on the exact sequence of surrounding bases, but many other determinations also found similar low cleavage ratios.[4−6]

Two different explanations have been advanced for the different trapping rates of GG and GGG. Berlin, Burin, and Ratner base their explanation on the assumptions that (i) both traps are deep, i.e., ∼0.5 eV for GG and ∼0.7 eV for GGG, the latter being so effective as to stop further hopping of the hole, and (ii) the ionization potential (IP) of guanine is lower than that of adenine and the other nucleobases by at least 0.4 eV.[7] The different reactivities of the two traps were ascribed by them to different relaxation times after the hole encountered the trap. In their model the GG units were taken to have a long relaxation time, so that a hole is likely to make a further hop before the trap closes on it, while the relaxation time of the GGG units is relatively short, faster than the hopping time. The assumption that the traps are deep was based on data for one-electron redox potentials of nucleobases in solution,[8,9] experimental values of their ionization potentials in vapors[11,12] and ab initio computational results for single bases.[10,13] These data reflect the oxidation potentials of individual bases. As emphasized by Schuster,[14] among others, there are many types of evidence that the ionization potential of a base can be much affected by its neighbors, as will be discussed below.

Lewis et al.[15] have measured the trap depths with experiments on synthetic hairpins that included on a strand GG or GGG units among A's, where A stands for adenine. They found that the free energy liberated in a hole transfer from G^+ to a GG is 0.052 eV, while hole transfer from G^+ to a GGG resulted in a free energy of 0.077 eV. Thus the traps are fairly shallow. It

* Address correspondence to this author: (e-mail) conwell@chem.rochester.edu.

(1) Henderson, P. T.; Jones, D.; Hampikian, G.; Kan, Y.; Schuster, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8353−8358.
(2) Ly, D.; Sanii, L.; Schuster, G. B. *J. Am. Chem. Soc.* **1999**, *121*, 9400−9410.
(3) Hickerson, R. P.; Prat, F.; Muller, J. G.; Foote, C. S.; Burrows, C. J. *J. Am. Chem. Soc.* **1999**, *121*, 9423−9428.
(4) Muller, J. G.; Hickerson, R. P.; Perez, R. J.; Burrows, C. J. *J. Am. Chem. Soc.* **1997**, *119*, 1501−1506.
(5) Saito, I.; Takayama, M.; Sugiyama, H.; Nakatani, K. *J. Am. Chem. Soc.* **1995**, *117*, 6406−6407.
(6) Nakatani, K.; Dohno, C.; Saito, I. *J. Am. Chem. Soc.* **1999**, *121*, 10854−10855.
(7) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. *J. Am. Chem. Soc.* **2001**, *123*, 260−268.
(8) Seidel, C. A. M.; Schultz, 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) Sugiyama, H.; Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 7063−7068.
(11) Hush, N. S.; Cheung, A. S. *Chem. Phys. Lett.* **1975**, *34*, 11−13.
(12) Lifschitz, C.; Bergman, E.; Pullman, B. *Tetrahedron Lett.* **1967**, *46*, 4583−4586.
(13) Prat, F.; Houk, K. N.; Foote, C. S. *J. Am. Chem. Soc.* **1998**, *120*, 845−846.
(14) Schuster, G. B. *Acc. Chem. Res.* **2000**, *33*, 253−260.
(15) Lewis, F. D.; Liu, X.; Liu, J.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2000**, *122*, 12037−12038.

10.1021/ja015947v CCC: $20.00
Published on Web 10/26/2001

was suggested by Lewis et al. that this would account for their relatively small trapping.

Among the evidence that the ionization potential of a base is affected by its neighbors is the fact that ab initio calculations of the ionization potentials of pairs or triples of stacked bases show considerable differences from the single base values.[10,16] Thus, for example, the calculated IP of GG was found to be smaller than that of G by 0.5 eV.[10] The calculated IP of AG was 0.5 eV smaller than that of A and larger than that of GG by 0.23 eV.[10] Cleavage ratios, although the amount of cleavage undoubtedly depends on factors other than IP, are roughly consistent with the calculated IP values for stacked pairs of bases.[10] Another type of evidence comes from hole motion observed in different sequences. From single-base values the barrier between G and A would be ~0.4 eV. Yet there is strong evidence that the barrier can be breached with thermal excitation.[17] Although experiments have shown that a hole starting from a G has increasing difficulty getting through three successive A's,[18] beyond four it moves easily and rapidly through a further succession of A's once it has gotten past the G to the first A.[19,14]

In general a radical cation or anion increases its stability by self-trapping in a structural distortion of the medium in which it is immersed, i.e., forming a polaron, and this is to be expected in DNA also.[14,20] In a one-dimensional (1D) situation the polaron must be a large polaron, spread out over a number of sites.[21] The effect of neighboring bases on the IPs mentioned above is indirect evidence that such spreading occurs in DNA. Calculations of polaron properties have been carried out for various base sequences in DNA.[20] It was found that the extent of the polaron wave function and the distortion are 5 to 7 sites (bases) for reasonable values of the parameters. The distortion consists of a decrease in the interbase spacing, with the maximum decrease being $\lesssim$0.4 Å, depending on the value of the hole transfer integral.

In what follows we use the Su−Schrieffer−Heeger (SSH) Hamiltonian to calculate the properties of a trapped hole created by removing an electron from a stack of bases. This is done for a stack containing one G, a stack with a GG, and one with a GGG, in all cases surrounded by adenines. We find that the resulting polarons are similar to each other and to the polarons obtained earlier for random sequences. The results are not sensitive to the value of the transfer or resonance integral but quite sensitive to the difference in the ionization potentials of G and A. We find that, with the usual approximation of neglecting the contribution of environmental and structural effects to the trapping energies,[10,16] we can match quite well the results of Lewis et al. for the differences between the energies of G and GG or of G and GGG traps. The best fit was obtained for the IP difference of 0.17 eV; the significance of this figure will be discussed below. The calculations are done with long sequences of adenines, but we show the results are little changed when the adenine sequences are short, as in the experiments. As will be discussed, our results are consistent also with the results of cleavage experiments on the sequences for which we carried out the calculations.

II. Model

The essential ingredients of our model are the following:

(i) The hole is assumed to be confined to a single strand of the DNA duplex. The strand is represented as a 1D lattice, each site n of which corresponds to a DNA base, either a guanine or an adenine. The corresponding ionization potentials Δ_G and Δ_A for isolated guanine and adenine determine the on-site hole energies V_n (the diagonal elements of the hole Hamiltonian). Since we are going to study hole trapping on guanines placed in a chain of adenines, it is convenient to set $V_n = 0$ if the nth site is an adenine, and $V_n = \Delta_G - \Delta_A < 0$ if it is a guanine. Thus a guanine (or a sequence of guanines) plays the role of an effective potential well for the hole.

(ii) As in the SSH Hamiltonian, we assume that the electronic wave functions on adjacent bases overlap, which leads to nonzero values of the hole transfer integrals $t_{n+1,n}$ between the sites $n + 1$ and n (the off-diagonal elements of the hole Hamiltonian).

(iii) The displacements of the bases u_n are also included in the model. The elastic restoring force acting on the base n from the neighboring bases is assumed to be given by

$$F_n = K(u_{n+1} - u_n) - K(u_n - u_{n-1}) \tag{1}$$

where K is the elastic constant. The lattice displacements may be treated classically due to the large mass of a base.

(iv) The hole motion is coupled to the lattice displacements via the dependence of the transfer integral $t_{n+1,n}$ on the distance between the bases, i.e., on $u_{n+1} - u_n$. The simplest assumption, adopted in the SSH model, is that the displacements are small enough so that the dependence may be described by a linear term

$$t_{n+1,n} = t_0 - \alpha(u_{n+1} - u_n) \tag{2}$$

α being the derivative of $t_{n+1,n}$ with respect to the net displacement and playing the role of the coupling constant. In principle, the dependence of $t_{n,n+1}$ on $u_{n+1} - u_n$ is more complicated than just a linear one. For large base separations it should be exponential, as prescribed by the asymptotical behavior of π-electron wave functions and obtained in the calculations of Sugiyama and Saito.[10] However, the character of this dependence for $u_{n+1} - u_n < 0$ (i.e. for two bases squeezed together), which is the relevant case for polarons, is not known. The results of ref 10 do not suggest any better law for the base separation less than 3.4 Å than a linear one. Therefore we use the linear approximation.

Another vibrational degree of freedom important for DNA is the relative twisting angle $\theta_{n+1} - \theta_n$ between the neighboring bases. It also couples to the transfer integral via the dependence of $t_{n,n+1}$ on $\theta_{n+1} - \theta_n$, and its effect on the charge transport was studied in ref 22. At small angles this dependence may be approximated by a quadratic one, as seen in Figure 4 of ref 10. This means that the equilibrium value of θ_n is not affected by the presence of the hole and thus this degree of freedom does not contribute to the energies of stationary states (the only ones to be studied in the present work). This interaction, however, may lead to renormalization of the transfer integral t_0, and in particular, introduce a temperature dependence. Still, since the precise values of t_0 and α are not well known (see the discussion below), we do not take this interaction into account.

(16) Voityuk, A. A.; Jortner, J.; Bixon, M.; Rosch, N. *J. Chem. Phys.* **2001**, *114*, 5614−5620.
(17) Giese, B.; Spichty, M. *CHEMPHYSCHEM* **2000**, *1*, 195−198.
(18) Meggers, E.; Michel-Beyerle, M. E.; Giese, B. *J. Am. Chem. Soc.* **1998**, *120*, 12950−12955.
(19) Williams, T. T.; Odom, D. T.; Barton, J. K. *J. Am. Chem. Soc.* **2000**, *122*, 9048−9049.
(20) Conwell, E. M.; Rakhmanova, S. V. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 4556−4560.
(21) Emin, D. *Phys. Rev. B* **1986**, *33*, 3973−3975.
(22) Bruinsma, R.; Grüner, G.; D'Orsogna, M. R.; Rudnick, J. *Phys. Rev. Lett.* **2000**, *85*, 4393−4396.

The formal treatment of the problem is analogous to that of ref 20, the only difference being that in the present work we do not calculate the energy levels for all the electrons in the stack, but focus on the behavior of the hole (the missing electron). We do not include the hole spin since it is not essential in this problem. Let ψ_n be the hole wave function in a stationary state, which may be chosen real. It obeys the stationary Schrödinger equation

$$(E_h - 2t_0)\psi_n = V_n\psi_n + [t_0 - \alpha(u_{n+1} - u_n)]\psi_{n+1} + [t_0 - \alpha(u_n - u_{n-1})]\psi_{n-1} \quad (3)$$

where E_h is the hole energy. The presence of $2t_0$ on the left-hand side indicates our choice of the zero energy, which is taken to be the bottom of the free hole band in a chain with $V_n = 0$, $u_n = 0$. The lattice displacements are found from the equilibrium condition (the requirement of zero net force acting on a base):

$$K(u_{n+1} - u_n) - K(u_n - u_{n-1}) = 2\alpha(\psi_{n+1}\psi_n - \psi_n\psi_{n-1}) \quad (4)$$

where the left-hand side is the elastic force, and the right-hand side is the electronic force. From eq 4 one obtains simply

$$u_{n+1} - u_n = \frac{2\alpha}{K}\psi_{n+1}\psi_n \quad (5)$$

Note that only the strain $u_{n+1} - u_n$ enters the equations, not the displacements themselves. We take the total energy of the system in the stationary state to be the sum of the hole energy and the lattice deformation energy:

$$E_{tot} = E_h + E_{lat} = E_h + \frac{K}{2}\sum_n (u_{n+1} - u_n)^2 \quad (6)$$

As noted above, this expression neglects the contribution of environmental and structural effects. With our choice of the zero hole energy (eq 3) the sign of E_{tot} indicates whether the hole is free or trapped.

As discussed earlier,[20] values of t_0 and α can be obtained from the calculations of Sugiyama and Saito[10] for the interaction of two G's as a function of the distance between them. This led to the choice $t_0 = 0.3$ eV and $\alpha = 0.6$ eV/Å. Application of superexchange theory to the observed tunneling from G/C through A/T led to $t_0 = 0.2$ eV.[23] Our calculations were done for both these values of t_0 and also for $t_0 = 0.1$ eV. Because α is the derivative of t with respect to displacement, when we used $t_0 = 0.2$ or 0.1 eV we scaled α, taking it as 0.4 or 0.2 eV/Å, respectively. The value of the elastic constant K was taken as 0.85 eV/Å, derived[20] from the measured value of the sound velocity in DNA.

III. Properties of the Hole Traps

To compare with the results of Lewis et al.[15] we carried out the calculations for sequences similar to those used by them. Thus we calculated the hole trapping energies from eq 6, the hole wave functions, and the lattice distortion for a single G, GG, and GGG, in all cases surrounded by A's. In Figure 1 are shown the results for a single guanine at site $n = 0$, surrounded by adenines, for two different values of the transfer integral. The lattice strain (always negative) is represented in the figures by solid symbols placed at half-integer abscissas $\nu = ..., -1/2, 1/2, 3/2, ...$, which correspond to the differences $u_{\nu+1/2} - u_{\nu-1/2}$. The hole population $|\psi_n|^2$ (the probability of finding the hole at the site n) is shown by open symbols. The well depth $\Delta_A -$

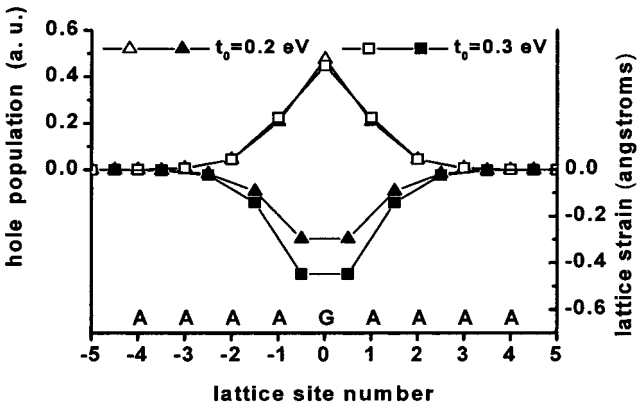

Figure 1. Hole population (open symbols) and lattice strain (filled symbols) for one guanine at $n = 0$ among adenines: $t_0 = 0.2$ eV (triangles) and $t_0 = 0.3$ eV (squares).

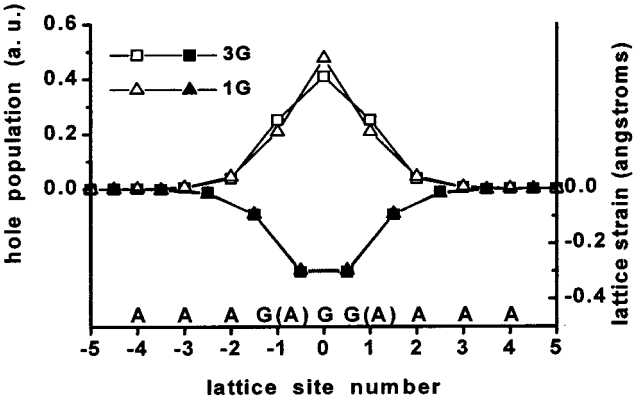

Figure 2. Hole population (open symbols) and lattice strain (filled symbols) for one guanine at $n = 0$ (triangles) and three guanines at $n = -1, 0, 1$ (squares), in both cases surrounded by adenines.

Table 1. Trapping Energies (eV) for a Hole in G, GG, and GGG Traps for $\Delta_A - \Delta_G = 0.17$ eV and Different Values of t_0

	G	GG	GGG
$t_0 = 0.3$ eV	−0.187	−0.240	−0.268
$t_0 = 0.2$ eV	−0.110	−0.161	−0.187
$t_0 = 0.1$ eV	−0.081	−0.130	−0.151

Δ_G is taken as 0.17 eV for all the cases calculated because, as noted earlier, that value was found subsequently to lead to agreement with the results of Lewis et al. The spatial extent of the trapped state is ~6 sites. Calculations with $\Delta_A - \Delta_G$ taken as 0.5 eV instead of 0.17 eV, with the other parameters the same, led to a trap state about 1 site narrower, the larger difference between HOMO levels resulting in greater confinement. The smaller strain and larger energy found for $t_0 = 0.2$ eV are the result of choosing a smaller (scaled according to t_0) value for α. The calculated energies are given in Table 1.

In Figure 2 are shown the results for GGG, at sites $n = -1$, 0, +1, among adenines, contrasted with those for a single G at site 0. The value of $t_0 = 0.2$ eV and the difference between IP's is again 0.17 eV. It is seen that the difference between the wave functions and the strain for G and GGG is quite small, but it does lead to a significant difference in the total energy, as shown in Table 1. For $t_0 = 0.3$ eV the results for GGG are quite similar to those seen for G at this value of t_0 in Figure 1. The results for GG, shown in Figure 3, have different shapes but otherwise are similar to those for G and GGG.

The fact that the hole wave functions are not confined to the guanines raises the possibility that these calculations do not

(23) Bixon, M.; Jortner, J. *J. Phys. Chem. B* **2000**, *104*, 3906−3913.

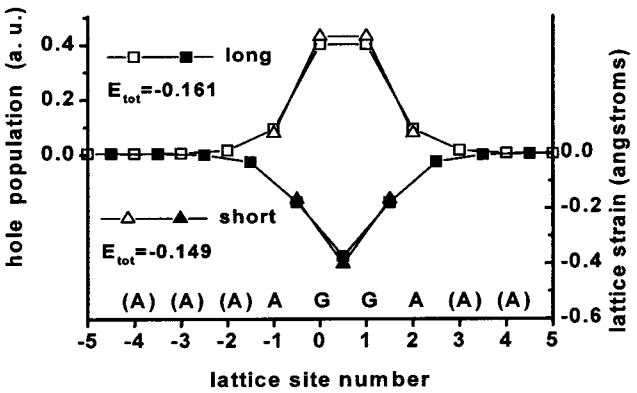

Figure 3. Hole population (open symbols) and lattice strain (filled symbols) for GG surrounded by many adenines (squares) and a short chain consisting of AGGA only (triangles).

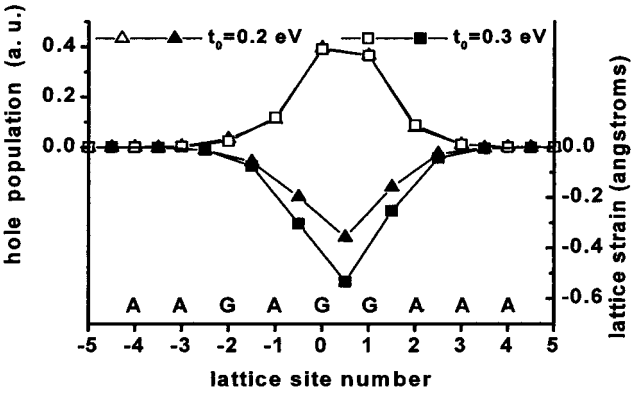

Figure 4. Hole population (open symbols) and lattice strain (filled symbols) for a GG separated by one A from a G: $t_0 = 0.2$ eV (triangles) and $t_0 = 0.3$ eV (squares).

apply to the results of Lewis et al., where usually there was only one A between G and GG or GGG. Indeed it suggests that their results may not be valid for cases where G, GG, and GGG are truly isolated from each other by a long series of A's. We therefore did an additional calculation, the results of which are also shown on Figure 3, for a stack of 4 sites, AGGA. It is seen that the results are little different from those where the guanines are surrounded by a long series of adenines. We also did a calculation, shown in Figure 4, for the situation of G and GG on a stack separated by only one A. It is seen that the wave function is essentially confined to GG, with only a small asymmetry due to the closeness of the G. The energy of the GG is almost unaltered. The good separation seen for the wave functions is largely due to the confinement of the hole by the polaron effect. Thus the results of Lewis et al. may be taken to apply to G, GG, and GGG with any number of A's between them.

IV. Discussion of Results

With the results shown in Table 1 we obtain for the energy released when a hole goes from G to 2G $E_1 - E_2 = 0.051$ eV for $t_0 = 0.2$ eV and $E_1 - E_2 = 0.053$ eV for $t_0 = 0.3$ eV. For a transition from G to GGG the energy released is $E_1 - E_3 = 0.078$ eV for $t_0 = 0.2$ eV and $E_1 - E_3 = 0.081$ eV for $t_0 = 0.3$ eV. It is seen that the agreement with the results of Lewis et al., $E_1 - E_2 = 0.052 \pm 0.006$ eV and $E_1 - E_3 = 0.077 \pm 0.005$ eV, respectively, is excellent. For $t_0 = 0.1$ eV the values for the energy released were within 10% of the values for $t_0 = 0.2$ or 0.3 eV. Thus the agreement is insensitive to the value of t_0. However, it is, as noted earlier, quite sensitive to the difference in energy between the IP values of A and G. A change from 0.17 eV to 0.23 eV in that energy results in a difference of 40% between the calculated and experimental results. As noted earlier, contributions to the energy from environmental effects, such as polarization by the hole of its surroundings and structural changes in the helix due to the motion of the bases, have not been taken into account. It is difficult to believe that these effects are large, because the change from *G* to *GG* or GGG involves only replacement of one or two A/T pairs by one or two G/C pairs, respectively. Nevertheless, we cannot conclude that 0.17 eV is the precise difference between the IP values of G and A. Still, we believe our result is a strong indication that the IP difference is smaller than usually considered. Further, to be consistent with the observation that a hole can make the transition from G to A by thermal escape,[17] this difference cannot be much larger than our result.

Evidence that the wave function of a trapped hole must extend over ~5 sites comes from measurements of the relative reactivity of a number of 5′-TXGYT-3′ sequences for different bases X and Y in B form DNA toward photoinduced one-electron oxidation.[24] The reactivity was found to be quite sequence dependent. Within the 6 sequences studied containing GG surrounded by other bases, there was a variation in reactivity of close to a factor 3, the sequences differing by this factor being 5′-TCGGT-3′ and 5′-TGGCT-3′.

From $|\psi_n|^2$ for GGG in Figure 2 we obtain equal populations on sites $n = 1$ and -1, with the population of the middle G, at site $n = 0$, being 1.65 times as large. This matches surprisingly well the ratios seen in piperidine cleavage of the sequence AGGGA,[25] which supports the suggestion that the relative probability of cleavage on a site *n* in a GGG trap is proportional to the probability $|\psi_n|^2$ of finding the hole on it.

In our calculations the populations $|\psi_n|^2$ are necessarily equal for the two guanines in a GG trap, due to the symmetry assumed in the calculation. In line with the suggestion that the cleavage probability is proportional to the hole population, this would correspond to the cleavage ratio 1:1. In the experiments the corresponding cleavage ratio is (strongly or weakly, depending on the specific experimental conditions and the base sequence) different from unity. Typically the guanine on the 5′ side is more reactive. A reason for this has been advanced by Prat et al.[13] Their calculations showed that the electrostatic potential map of a guanine has a significant concentration of negative charge on one of the nitrogens (N7) and one of the oxygens.[13] In B-form DNA N7 of the 3′G is located just below the six-membered ring of the 5′G, resulting in a more attractive potential for the hole on the 5′G and thus more concentration of the hole wave function on the 5′G.

It has been noted that structural factors such as local deviations from perfection of the B-form helix also play a role in the reactivity of a site.[25,14] Whether this role is due to the chemistry underlying cleavage or an effect on the relative population of the bases is not clear. Generally, many determinations of the cleavage ratio for the 5′G to that for the 3′G have led to the statement that the reactivity of the 5′G is 3 to 5 times that of the 3′G.[26] However, these experiments have been performed on sequences other than AGGA.

A measurement of the ratio of 5′ to 3′ reactivity for the sequence CAGGAT under piperidine cleavage gave the result

(24) Saito, I.; Nakamura, T.; Nakatani, K.; Yoshioka, Y.; Yamaguchi, K.; Sugiyama, S. *J. Am. Chem. Soc.* **1998**, *120*, 12686−12687.
(25) Spassky, A.; Angelov, D. *Biochemistry* **1997**, *36*, 6571−6576.
(26) Sistare, M. F.; Codden, S. J.; Heimlich, G.; Thorp, H. H. *J. Am. Chem. Soc.* **2000**, *122*, 4742−4749.

that the 3′G is slightly more reactive than the 5′G.[25] In our modeling this small effect may be reproduced by ascribing slightly different values of V_n to the two guanines, thus shifting down one of them. This makes the hole "prefer" this site to the other one, and can be done without destroying the agreement with the Lewis et al. results for the energy difference $E_1 - E_2$.

In summary, we can account well with a polaron model, having reasonable values of the parameters, for the trap depths of GG and GGG relative to that of G measured by Lewis et al.[15] In this model the wave function is not confined to the G's but is still substantial on the surrounding bases, A's in this case. The polaronic distortion is a decrease in the spacing of adjacent bases, the maximum decrease being 0.4 Å for $t_0 = 0.3$ eV or smaller if t_0 is smaller. The fit is insensitive to the value of the transfer integral, but requires that the difference between IP values of adjacent G and A be ~0.2 eV rather than the 0.4 eV characteristic of the isolated bases. The small trapping found, particularly for GG, is due to the shallowness of the traps rather than relaxation effects. Our results are in agreement with the cleavage or reactivity of a particular G in a sequence being determined by its population relative to that of the other G's in the sequence.

JA015947V