Charge transfer in DNA: Hole charge is confined to a single base pair due to solvation effects

Alexander A. Voityuk[a)]
Institució Catalana de Recerca i Estudis Avançats (ICREA), and Institute of Computational Chemistry, Universitat de Girona, 17071 Girona, Spain

We include solvation effects in tight-binding Hamiltonians for hole states in DNA. The corresponding linear-response parameters are derived from accurate estimates of solvation energy calculated for several hole charge distributions in DNA stacks. Two models are considered: (A) the correction to a diagonal Hamiltonian matrix element depends only on the charge localized on the corresponding site and (B) in addition to this term, the reaction field due to adjacent base pairs is accounted for. We show that both schemes give very similar results. The effects of the polar medium on the hole distribution in DNA are studied. We conclude that the effects of polar surroundings essentially suppress charge delocalization in DNA, and hole states in $(GC)_n$ sequences are localized on individual guanines. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1924551]

INTRODUCTION

DNA-mediated charge transfer currently attracts considerable interest because of its relevance for the oxidative damage and mutations of DNA and its potential importance for molecular electronics. The status of experimental and theoretical investigations on charge transport through DNA has recently been the subject of several reviews in the book.[1] The physical framework for a quantitative treatment of charge transfer in DNA has also been considered in papers by Beratan and co-workers[2,3] and Ratner and co-workers.[4,5]

Solvation effects play an important role in the charge transfer mediated by DNA. In particular, solvent reorganization is a key parameter, which influences the dynamics of electron-hole migration through DNA. Estimates of the interaction energy between an excess charge in the interior of the double helix and its environment can essentially depend on a model employed for the calculation. For instance, to estimate the reorganization energy for hole transfer in DNA, one applied the Poisson equation solver[6] to heterogeneous dielectric models consisting of several different dielectric zones surrounding the hole donor and acceptor sites.[2,5,7,8] The calculation results are quite different because of uncertainties concerning the construction of the dielectric model.[8]

The surrounding polar medium affects also the delocalization of an electron hole in DNA over adjacent base pairs. There has been an interesting discussion as to whether the hole charge in DNA is confined to a single base pair or delocalized over several adjacent base pairs (see, for instance, Ref. 1 and references therein). Recent computational studies gave conflicting results. (1) Based on a simple cylindrical cavity model, where the charge was concentrated on the axis of the cylinder, Basko and Conwell accounted for the solvation effects in the tight-binding Hamiltonian and concluded that the hole charge is spread over five or more

adjacent (GC) base pairs.[9] The hole wave function found within this model is similar to that calculated, without taking into account the polar medium. (2) By contrast, Beratan and co-workers showed that the interaction with surroundings considerably affects the charge distribution in DNA.[2] They employed a heterogeneous dielectric model comprised of a DNA zone and a solvent zone. Although solvation terms were shown to essentially favor localization of the hole delocalization over two or three guanines is found to be energetically feasible. Beratan and co-workers estimated the total energies of the system with localized and uniformly delocalized hole, while they did not consider the corresponding correction of the Hamiltonian for the effects of polar surroundings.[2]

Many interesting theoretical results concerning charge transfer in DNA have been obtained by using tight-binding Hamiltonians. Olofsson and Larsson studied the effect of structural reorganization of nucleobases on the delocalization of an excessive charge in DNA.[10] Effects of static and dynamic structural fluctuations on the hole mobility in DNA were considered recently by Grozema *et al.*[11] Roche studied the dependence of the DNA-mediated conduction on the sequence base pairs.[12] Hole transfer through π stacks containing chemically modified nucleobases was also considered.[13] However, in these studies, the effects of the environment have not been accounted for. As already noted the solvation term was included in the Hamiltonian within a simple cylindrical model.[9] In this paper we consider a scheme to derive the polar-medium correction terms for tight-binding Hamiltonians, using solvation energies computed within more elaborated schemes based on heterogeneous (multizone) dielectric approach or molecular-dynamics simulations. Then we will study the effects of a polar environment on the hole distribution in $(GC)_n$ sequences.

a)Electronic mail: alexander.voityuk@icrea.es

122, 204904-1

METHOD

Effective Hamiltonian

Distribution of the hole charge in a π stack can be described using a tight-binding Hamiltonian $\widetilde{H}$. Only one state per site (base pair b_i) is taken into account. Thus a *system* $b_1 b_2 \ldots b_n$ has n states, where the hole is localized on one of the site. The corresponding state functions are assumed to be orthonormalized (overlap between states is neglected). Matrix elements of $\widetilde{H}$ are determined as

$$\widetilde{H}_{ii} = \varepsilon_i^0 + s_i,$$

$$\widetilde{H}_{ij} = V_{ij}.$$

The diagonal elements $\widetilde{H}_{ii}$ can be estimated as a sum of the oxidation potential ε_i^0 and the solvation correction s_i. Off-diagonal elements between adjacent base pairs ($j = i \pm 1$) can be approximated by the corresponding electronic coupling V_{ij} and neglected in other cases. For the canonical structure of B-DNA, the electronic coupling V of adjacent GC pairs calculated using different schemes is about 0.08 eV.[14,15]

In turn, the vertical oxidation potential ε_i^0 may be approximated by the relative ionization energy of a base pair b_i in the sequence. These energies depend in an essential fashion on the nature of adjacent base pairs j ($j = i \pm 1$).[16,17] However, in sequences $(^*)(GC)(GC)_n(GC)(^*)$ all base pairs within the $(GC)_n$ fragment have very similar oxidation potentials, $\varepsilon_i^0 = \varepsilon^0$ for all i.

Let us consider now how to estimate the solvation term s_i. According to the reaction field theory, a charge immersed in a dielectric medium will induce an electric field in the solvent. In turn, this field will stabilize the solute. Within the linear-response model s_i can be written as

$$s_i = s_i^0 + \zeta_i^0 (q_i - 1) + \sum_{j \neq 0} \zeta_i^j (q_{i-j} + q_{i+j}). \tag{1}$$

The quantity s_i^0 corresponds to a reference state in which the positive charge is completely localized on a single base pair ($q_i = 1, q_{i-j} = q_{i+j} = 0$); the second term $\zeta_i^0 (q_i - 1)$ is due to medium polarization by the charge q_i, $\zeta_i^0 < 0$; the last term $\sum_{j \neq 0} \zeta_i^j (q_{i-j} + q_{i+j})$ is due to charges q_{i-j} and q_{i+j} on other base pairs (e.g., for nearest neighboring pairs $j = 1$); one can expect that $|\zeta_i^j| < |\zeta_i^0|$. When all base pairs in the systems are identical the quantities s_i^0, ζ_i^0, and ζ_i^j do not depend on a base pair b_i and the subscript i can be dropped. We will consider two models. In the model A, the correction s_i, Eq. (1), depends only on the charge localized on the corresponding site, i.e., for all j, $\zeta_i^j = 0$. In the model B, the reaction field due to nearest base pairs $j = \pm 1$ is also accounted for, $\zeta_i^1 \neq 0$. Then the solvation term corresponding to a state with charge distribution $\{q_i\}$ can be expressed as

$$\Delta E^{\text{solv}}(q_1, q_2, \ldots, q_n)$$

$$= \sum_{i=1}^{n} q_i s_i = \sum_{i=1}^{n} q_i (s^0 + \zeta^0 (q_i - 1) + \zeta^1 q_{i-1} + \zeta^1 q_{i+1}). \tag{2}$$

When a hole is delocalized over n base pairs the corresponding solvation term ΔE_n^{solv} is

TABLE I. Hole charge delocalization in $(GC)_n$ sequences. The effects of polar environment are neglected.

Sequence	$(GC)_3$	$(GC)_5$	$(GC)_7$	$(GC)_9$
q_0	0.5	0.333	0.25	0.2
$q_{\pm 1}$	0.25	0.25	0.213	0.181
$q_{\pm 2}$		0.083	0.125	0.131
$q_{\pm 3}$			0.037	0.069
$q_{\pm 4}$				0.019
Sequence	$(GC)_2$	$(GC)_4$	$(GC)_6$	$(GC)_8$
$q_{\pm 1}$	0.5	0.362	0.272	0.216
$q_{\pm 2}$		0.138	0.175	0.167
$q_{\pm 3}$			0.054	0.092
$q_{\pm 4}$				0.026

$$\Delta E_n^{\text{solv}} = s^0 + \left(\frac{1}{n} - 1 \right) \zeta^0 + \frac{2n - 2}{n^2} \zeta^1. \tag{3}$$

Thus, for $n = 1$, $\Delta E_1^{\text{solv}} = s^0$; for $n = 2$, $\Delta E_2^{\text{solv}} = s^0 - 1/2\zeta^0 + 1/2\zeta^1$; and so on. Parameters ζ^0 and ζ^1 can be found by fitting of ΔE_n^{solv} ($n \geq 2$) given by Eq. (3) to corresponding values computed for the "real" models.

As already noted quite complicated schemes must be used to estimate solvation energies ΔE_n^{solv} of hole states in DNA. Kurnikov *et al.* calculated the stabilization energy of different hole states within DNA duplexes.[2] In particular, they considered several systems with the charge uniformly delocalized over $(GC)_n$ clusters embedded in AT run. ΔE_n^{solv} is found to be -1.857, -1.451, -1.190, and -1.013 eV, for $n = 1$, 2, 3, and 4, respectively. Based on these results we can estimate the quantities s^0, ζ^0, and ζ^1. The fitted values we obtain are $\zeta_A^0 = -1.020$ eV for the model A, and $\zeta_B^0 = -1.433$ eV and $\zeta_B^1 = -0.629$ eV for the model B. Using these parameters, one can well reproduce reference values of ΔE_n^{solv} [the standard deviation of solvation energies estimated within the models A and B (Eq. (2)) are 0.066 and 0.006 eV, respectively]. Note that the parameter $s^0 = \Delta E_1^{\text{solv}}$ does not influence the charge distribution in the system, but equally shifts all energy levels. As will be shown, the models A and B provide very similar results.

The energies E_i and the coefficients matrix C are determined by the eigenvalue problem for the effective Hamiltonian $\widetilde{H}$. Because diagonal matrix elements of $\widetilde{H}$ depend on the charge distribution, an iterative procedure is used. An initial density matrix is assumed to be diagonal, with matrix elements corresponding to a delocalized hole.

RESULTS AND DISCUSSION

In this section, we study the effects of a polar environment on the charge distribution in $(GC)_n$ sequences. First, we consider charge distributions in the sequences when the solvation effects are neglected (Table I). The hole charge is well delocalized over π stacks. For instance, in systems $(GC)_{2k+1}$ the largest charge, which is localized on the central base pair q_0, decreases as $1/(k+1)$ (the charge is equal to 0.5, ~ 0.33, 0.25, and 0.20 when k increases from 1 to 4). If the surrounding medium is not taken into account, all diagonal matrix elements are of the same value and the charge distribution

TABLE II. Hole charge distribution in $(GC)_n$ sequences. The effects of a polar environment are included within the models A and B.

Sequence Model	$(GC)_3$		$(GC)_5$		$(GC)_7$		$(GC)_9$	
	A	B	A	B	A	B	A	B
q_0	0.988	0.981	0.987	0.980	0.987	0.980	0.987	0.980
$q_{\pm 1}$	0.006	0.010	0.006	0.010	0.006	0.010	0.006	0.010
$q_{\pm 2}, q_{\pm 3}, q_{\pm 4}$			0.000	0.000	0.000	0.000	0.000	0.000

Sequence Model	$(GC)_2$		$(GC)_4$		$(GC)_6$		$(GC)_8$	
	A	B	A	B	A	B	A	B
q_0	0.994	0.990	0.988	0.980	0.987	0.980	0.987	0.980
$q_{\pm 1}$	0.006	0.010	0.006	0.010	0.006	0.010	0.006	0.010
$q_{\pm 2}, q_{\pm 3}, q_{\pm 4}$			0.000	0.000	0.000	0.000	0.000	0.000

does not depend on the magnitude of electronic coupling V. Thus the results are very similar to those obtained by Basko and Conwell.[9]

However, this situation changes dramatically when the solvation effects are included. Table II shows the corresponding charges obtained using the models A and B. First, the hole charge is almost completely (more than 98%) confined to a single pair. Less than 1% of the charge is found on each nearest site. All the more distant base pairs do not carry any charge. This result is independent of the length of a π stack. Both models give a very similar picture for the charge distribution. Note that in systems with an even number of base pairs, the charge distribution is not symmetric. For instance, in the dimer $(GC)_2$ the charge can be distributed as (0.99, 0.01) or (0.01, 0.99). Obviously, both states are of the same energy. In $(GC)_{2k+1}$ the ground state of the radical cation corresponds to a state where the hole is localized on the middle base pair. Starting with a different density matrix, one can also obtain states where the charge is confined to other sites. Let us consider the stack $(GC)_5$. In the ground state, the hole is on the third (middle) site. The calculated energies of states, where the charge is localized on the first and the second base pairs, are very close to that of the ground state (the difference is about 0.02 kT). This estimation suggests that a hole can be localized on any base pair within the π stack. The same result is also obtained for other $(GC)_n$ systems. Thus we can conclude that the solvation effects lead to very localized hole states, which can reside on any (GC) pair.

This conclusion is at variance with the result obtained by Basko and Conwell.[9] They conclude that the shape of the hole wave function is not essentially different from the one obtained without taking into account solvation effects, and

the hole is spread over three to five sites. The main reason for this disagreement can be explained as follows. A solvation correction to $\tilde{H}_{ii}$ due to the charge q_i calculated within the cylindrical model is comparable to that caused by charges on neighboring pairs [the parameters ζ_i^j, $j = 0, 1, \ldots, k$, in Eq. (1) varies slowly with j].[9] In turn, the weak dependence of ζ_i^j on j is due to the fact that the distances between charges on the axis of the cylinder and the polar medium are essentially longer than the distance between adjacent pairs, 3.4 Å. This model appears to be rather crude, and it is not supported by the results of molecular-dynamics simulations of DNA[18,19] or even by the more extended electrostatic models using a heterogeneous dielectric medium.[2,6–8] As expected, the calculation carried out using $\zeta^0 = \zeta^1 = \zeta^2 = 1.0$ leads to quite delocalized hole states in $(GC)_n$. For instance, we obtain that in $(GC)_5$ $q_0 = 0.476$, $q_{\pm 1} = 0.251$, $q_{\pm 2} = 0.011$.

The electronic coupling V between base pairs is known to be a parameter, which essentially determines the delocalization of the hole wave function. As already discussed in the literature[14,20,21] electronic couplings between base pairs are very sensitive to conformational changes of the DNA. Therefore, it is worthwhile to consider how the variation of the coupling V will affect the hole distribution in a stack. While $V = 0.08$ eV seems to be very reasonable, we carry out calculations of $(GC)_5$ with larger values of this parameter. Table III compares charges calculated with $V = 0.08$, 0.16, 0.24, and 0.32 eV. Note that the results obtained within both solvation models are in good agreement. As expected, the hole delocalization increases with the electronic coupling; however, the charge remains essentially confined to a single site with

TABLE III. Hole charge distribution in $(GC)_5$ calculated using different values of the electronic coupling V (in eV). The effects of a polar environment are included within the models A and B.

V (in eV) Solvation model	0.08		0.16		0.24		0.32	
	A	B	A	B	A	B	A	B
q_0	0.987	0.980	0.947	0.922	0.868	0.830	0.727	0.724
$q_{\pm 1}$	0.006	0.010	0.026	0.039	0.063	0.083	0.124	0.131
$q_{\pm 2}$	0.000	0.000	0.001	0.000	0.004	0.002	0.013	0.007

one exception obtained at $V=0.32$ eV. In this instance, the hole charge is found to be remarkably ($\sim 30\%$) delocalized. However, a strongly overestimated value of the electronic coupling was employed for the calculation [$V=0.32$ eV is by a factor of four larger than the value of 0.078 eV obtained by averaging coupling matrix elements, calculated for different mutual position of (GC) pairs[15]].

Besides the solvation term there is another effect resulting in the hole confinement to one base pair. This is an internal (structural) reorganization of nucleobases caused by an excess charge. This effect was recently considered in detail by Olofsson and Larsson.[10] They found that spatially well-localized hole states are energetically stabilized due to the internal reorganization of nucleobases. Obviously, the internal reorganization term will reinforce the hole localization to a single base pair.

CONCLUSIONS

(1) A simple scheme has been proposed for estimating the solvation effects within the tight-binding Hamiltonian and related methods employed for modeling of charge transfer in DNA. We showed how the corresponding linear-response parameters can be derived from solvation energies calculated for several charge distributions in DNA using accurate computational approaches. Two models were considered: (A) where the correction to $\tilde{H}_{ii}$ depends only on the charge localized on the same site, and (B) where, in addition, the effect of the reaction field due to adjacent base pairs is accounted for. We found that both schemes give similar results.

(2) The effects of polar surroundings on the hole distribution in DNA were studied. At variance with previous results[9] we conclude that hole states in (GC)$_n$ sequences are localized on individual base pairs. The solvation effects suppress essentially the charge delocalization in DNA, leading to radical cation states confined to a single (GC) site. The result remains unchanged when considerable deviations of the electronic coupling from its average value are accounted for.

Thus, our statement on the hole confinement to one site appears to be quite robust and independent of other terms included in the Hamiltonian.

It is not very uncommon that a hole delocalization over several base pairs is assumed, at least implicitly, when interpreting experimental data on one-electron oxidation of DNA or/and considering possible mechanisms of charge migration in DNA.[1] Our results may be helpful by justifying such explanations.

[1] Long-Range Charge Transfer in DNA, in Topics in Current Chemistry Vols. 236–237, edited by G. B. Shuster (Springer, Berlin, 2004).

[2] I. V. Kurnikov, G. S. M. Tong, M. Madrid, and D. N. Beratan, J. Phys. Chem. B **106**, 7 (2002).

[3] G. S. M. Tong, I. V. Kurnikov, and D. N. Beratan, J. Phys. Chem. B **106**, 2381 (2002).

[4] D. Beljonne, G. Pourtois, M. A. Ratner, and J. L. Bredas, J. Am. Chem. Soc. **125**, 14510 (2003).

[5] D. N. LeBard, M. Lilichenko, D. V. Matyushov, Yu. A. Berlin, and M. A. Ratner, J. Phys. Chem. B **107**, 14509 (2003).

[6] W. Rocchia, E. Alexov, and B. J. Honig, J. Phys. Chem. B **105**, 6507 (2001).

[7] H. L. Tavernier and M. D. Fayer, J. Phys. Chem. B **104**, 11541 (2000).

[8] K. Siriwong, A. A. Voityuk, M. D. Newton, and N. Rösch, J. Phys. Chem. B **107**, 2595 (2003).

[9] D. M. Basko and E. M. Conwell, Phys. Rev. Lett. **88**, 098102 (2002).

[10] J. Olofsson and S. Larsson, J. Phys. Chem. B **105**, 10398 (2001).

[11] F. C. Grozema, L. D. A. Siebbeles, Yu. A. Berlin, and M. A. Ratner, ChemPhysChem **3**, 536 (2002).

[12] S. Roche, Phys. Rev. Lett. **91**, 108101 (2003).

[13] A. A. Voityuk and N. Rösch, J. Phys. Chem. B **106**, 3013 (2002).

[14] A. A. Voityuk, J. Jortner, M. Bixon, and N. Rösch, J. Chem. Phys. **114**, 5614 (2001).

[15] A. A. Voityuk and N. Rösch, J. Chem. Phys. **117**, 5607 (2002).

[16] A. A. Voityuk, J. Jortner, M. Bixon, and N. Rösch, Chem. Phys. Lett. **324**, 430 (2000).

[17] K. Senthilkumar, F. C. Grozema, C. F. Guerra, F. M. Bickelhaupt, and L. D. A. Siebbeles, J. Am. Chem. Soc. **125**, 13658 (2003).

[18] V. Makarov, B. M. Pettitt, and M. Feig, Acc. Chem. Res. **35**, 376 (2002).

[19] M. A. Young, G. Ravishanker, and D. L. Beveridge, Biophys. J. **73**, 2313 (1997).

[20] A. A. Voityuk, K. Siriwong, and N. Rösch, Phys. Chem. Chem. Phys. **3**, 5431 (2001).

[21] A. Troisi and G. Orlandi, J. Phys. Chem. B **106**, 2093 (2002).