Strong localization of positive charge in DNA

D. B. Uskov and A. L. Burin
Department of Chemistry, Tulane University, New Orleans LA, 70118
(Dated: March 15, 2008)

Microscopic mechanisms of positive charge transfer in DNA remain unclear. A quantum state of electron hole in DNA is determined by the competition of a pi-stacking interaction $b$ smearing a charge between different base pairs and interaction $\lambda$ with the local environment which attempts to trap charge. To determine which interaction dominates we investigated charge quantum states in various $(GC)_n$ sequences choosing DNA parameters satisfying experimental data for the balance of charge transfer rates $G^+ \leftrightarrow G^+_{n^*}$, $n = 2.3$ [1]. We show that experimental data can be consistent with theory only under an assumption $b \ll \lambda$ meaning that charge is typically localized within a single $G$ site. Consequently any DNA sequence including the one consisting of identical base pairs behaves more like an insulating material than a molecular conductor.

PACS numbers: 7080.0e, 72.20.Ee, 72.25.-b, 87.14.Gg

1. Positive charge transfer in DNA is being extensively investigated since its experimental discovery [2], while to our knowledge there is no experimental demonstration of the negative charge transfer. Electron transfer in DNA can be responsible for the oxidative DNA damage [2, 3, 4, 5, 6] and is possibly important for DNA repairing [2, 8]. Also an ability of DNA to promote long distant charge transfer can be used in molecular electronics [3]. Therefore this problem is extensively investigated during the past decade but the microscopic mechanism of charge transfer is not understood yet. We hope that this work may shed some light on this challenging problem.

DNA contains two different sorts of base pairs AT and GC forming quasi-random sequences. The lowest ionization potential is attributed to a GC pair (essentially G-base [10]). Since the electron transfer integral $b$ between adjacent bases does not exceed AT - GC ionization potential difference $b \leq \Delta \sim 0.5$eV, the quantum state of charge in a frozen environment will be localized near some G base and its localization length is comparable to the interbase distance [11]. Interaction with the environment breaks down this localization, inducing charge hopping between adjacent quantum states localized at G sites. Indeed, according to experimental studies [2, 3] and theoretical model [12, 13] the sequence dependent charge transfer in DNA can be represented as the series of charge hops between adjacent G bases serving as centers of localized states. An addition of AT pair between adjacent GC pairs dramatically reduces the charge hopping rate [2, 12, 13] and therefore optimum base sequence for the most efficient charge transfer consists of identical base pairs. Therefore we are going to study a charge (hole) quantum state in sequences of identical GC base pairs.

The thermal energy at room temperature is very small compared to other characteristic energies of the problem so we can use the ground state of the hole coupled to environment as a representative state. Then the spatial size of the hole state is determined by the competition of the delocalization of charge due to the pi-stacking interaction of heterocyclic groups belonging to adjacent bases and the localization caused by the environment polarization around the charge. The charge delocalization energy can be characterized by the effective electron transfer integral $b$ and the localization energy is given by the medium reorganization energy $\lambda$ [14, 15, 16, 17, 18]. Delocalization of charge over $k$ base pairs leads to the gain in the energy $E_{del} \sim -2b + b/k^2$, while the reorganization energy scales with the size of charge wavefunction as $E_{loc} \sim -\lambda/k$ (see e. g. [14]). Optimization of the total energy with respect to the number of sites $k$ estimates the size of the ground state charge wavefunction $k \sim 2b/\lambda$. At zero temperature the hole is localized, while at finite temperature it can hop to different states because of its interaction with the fluctuating environment. In the translationally invariant system ($(GC)_n$ or $(AT)_n$) the potential barrier separating two configurations can be estimated as the energy price for increasing the size of the wavefunction by one more site $k \rightarrow k+1$ compared to its optimum state. If $b \ll \lambda$ then this energy is given by the half of the reorganization energy $\Delta \approx \lambda/2 \gg k_B T$, while in the opposite limit we get a very small value $\Delta \sim \lambda^2/(16b^2)$ which quickly becomes negligible at moderately large $b/\lambda$.

It is important to determine the true relationship of $b$ and $\lambda$ in DNA. Indeed, in the regime $b \gg \lambda$ a DNA molecule made of identical base pairs would behave as a one-dimensional conductor, while in the opposite limit DNA always acts like an insulator. Existing estimates in literature [10, 14, 15, 16, 17, 18, 19] do not help much in resolving this issue, because there is a large controversy between different approximations. In particular, various estimates for the electron transfer integral $b$ range from 0.05eV [17] to 0.5eV [10], and all calculations of $b$ ignore vibrational rearrangements [21]. The estimates for the reorganization energy $\lambda$ using the continuous medium approach range from 0.25eV [18] to more than 1eV [14] due to uncertainty in the water dielectric constant value near the DNA molecule. Therefore based on different ap-
proaches different conclusions are made about the charge state including a prediction of a propagating intermediate size polaron \(^{10, 18, 19}\) or a small radius polaron essentially localized within a single base\(^{17}\).

To resolve this problem we suggest an alternative method to study the charge quantum state within the DNA molecule, using experimental data sensitive to the relationship of two key parameters of the theory, \(b\) and \(\lambda\). Namely, we exploit the rate constants for the balance of charge transfer rates between different \((GC)\_n\) complexes, measured by Lewis and coworkers \(^1\)

\[
\begin{align*}
G^+ + GG & \xrightleftharpoons{k_{GG}^{\text{GC}}}{k_{GG}^{\text{GC}}}^{-1} G + (GG)^+, \quad k_{GG}^{\text{GC}} = 7.7 \pm 1 \\
G^+ + GGG & \xrightleftharpoons{k_{GGG}^{\text{GC}}}{k_{GGG}^{\text{GC}}}^{-1} G + (GGG)^+, \quad k_{GGG}^{\text{GC}} = 20 \pm 1.
\end{align*}
\]

(1)

In the thermal equilibrium these ratios are determined by the base pair partition functions

\[
\begin{align*}
r_2 &= \frac{k_{GG}^{\text{GC}}}{k_{GG}^{\text{GC}}} = \frac{Z_{2+}Z_1}{Z_2Z_1}, \\
r_3 &= \frac{k_{GGG}^{\text{GC}}}{k_{GGG}^{\text{GC}}} = \frac{Z_{3+}Z_1}{Z_2Z_1}.
\end{align*}
\]

(2)

where \(Z_{n+}\) stands for the partition function for \(G_n\) sequence containing a single hole in it, while \(Z_n\) is the partition function for the same base sequence, but without the hole. Since \(G\) and \(G_n\) complexes are separated by some AT bridge we can ignore their interactions.

Both ratios in Eq. (1) depend on two parameters \(b\) and \(\lambda\) and the thermal energy at room temperature \(k_B T \sim 0.026\text{eV}\). Below we calculate both ratios using tight binding model for \(G_n\) complexes and standard linear response theory for charge interaction with the environment \(^{15, 20}\). Theory determines the domain of parameters \(b\) and \(\lambda\) satisfying experimental data Eq. (1).

We demonstrate that any choice of \(\lambda\) and \(b\), satisfying Eq. (7), corresponds to the regime \(b \ll \lambda\), where the hole in its ground state is localized essentially in a single \(G\) base (see Fig. 7).

2. The chain of \(n\) \(GC\) base pairs can be described by the tight binding Hamiltonian coupled to the classical environment represented by coordinates \(X_i, i = 1, \ldots, n\) coordinates for each DNA site

\[
\hat{H} = \hat{H}_{\text{hole}} + \hat{H}_{\text{media}} + \hat{V}_{\text{int}},
\]

\[
\hat{H}_{\text{hole}} = -b \sum_{i=1}^{n-1} (c_i^+ c_{i+1} + c_{i+1}^+ c_i),
\]

\[
\hat{H}_{\text{media}} = \frac{1}{2\lambda} \sum_{i=1}^{n} X_i^2, \quad \hat{V}_{\text{int}} = -\sum_{i=1}^{n} X_i (c_i^+ c_i).
\]

(3)

Here \(c_i, c_i^+\) are operators of creation and annihilation of electron hole in a site \(i\). Classical coordinates \(X_i\) describing the polar environment are directly coupled to the local charge density \(n_i = c_i^+ c_i\). The solvent energy is expressed as a bilinear form with respect to solvent coordinates, which is justified by a standard assumption that polarization fields are small compared to atomic fields \(^{20}\) so we can ignore \(X^3\) terms. We assume that only classical degrees of freedom with excitation energy comparable or less than the thermal energy are left in Eq. (3), while high energy modes are integrated out. This may lead to the renormalization of parameters in the system Hamiltonian Eq. (3) (see e.g. \(^{21, 22}\)) and we assume that this renormalization is made. We do not include off-diagonal terms \(X_iX_j, i \neq j\) into the Hamiltonian. This is justified because they are smaller than the diagonal ones \(^{16}\). It can be shown that for \(G_2\) sequence the problem including off-diagonal terms can be reduced to the diagonal model Eq. (3) with the replacement of the single site reorganization energy \(\lambda\) with the reorganization energy for charge transfer between adjacent sites. For \(GGG\) sequence the similar replacement with removal off-diagonal terms remains a good approximation\(^{23}\). Note that the addition of \(A\) or \(T\) bases surrounding \(G_n\) sequences leads to small changes in our results because the electron transfer integral is smaller in average between \(A\) and \(G\) base than between two \(G\) bases and also because of the large ionization potential difference of \(A\) and \(G\) bases \(^{17}\). Also we ignore the second \(C\) strand because of the weak coupling between strands \(^{17}\) and large difference of their ionization potentials \(^{10}\).

We study the ratios of charge transfer rates Eq. (2). Each partition function is given by \(Z_n = \int dX_1 \ldots dX_n Tr e^{-\beta H_n}\), where \(H_n\) is \(G_n\) Hamiltonian Eq. (3), trace is taken only over the states with the single hole \((Z_{n+})\) or no holes \((Z_n)\) and \(\beta = 1/(k_B T)\). If there is no hole calculations are reduced to multiple evaluation of a Gaussian integral leading to

\[
Z_n = c^n, \quad c = \sqrt{\frac{2\pi \lambda}{\beta}}.
\]

(4)

For the sequences containing a hole an analytical expression can be obtained only for \(n = 1\)

\[
Z_{1+} = c e^{\beta \lambda/2}.
\]

(5)

For \(n = 2, 3\) one can perform analytical integration over a “center of mass” coordinate \(X_1 + \ldots X_n\) which is coupled to the conserving operator of the total number of particles \(c_1^+ c_1 + \ldots + c_n^+ c_n = 1\). This reduces the calculations to single and double integrals, respectively. Below we give the expression for \((GG)^+\) partition function

\[
Z_{2+} = (\sqrt{2}c) e^{\beta \lambda/4} \int_{-\infty}^{+\infty} du \frac{e^{-u^2/4}}{\cosh \left(\beta \sqrt{u^2/4 + b^2}\right)},
\]

(6)

while the expression for \(Z_{3+}\) is more complicated and will be published elsewhere\(^{23}\).
We have performed numerical evaluations of ratios in Eq. 2 to find domains of parameters $b$ and $\lambda$ satisfying Eq. 1 and show these domains in Fig. 1. The upper (lower) border of each domain is defined by the maximum (minimum) value of ratios $r_2$ and $r_3$ Eq. 2 within the experimental error (8.7 (6.7) for $GG$ and 21 (19) for $GGG$). The acceptable domain of parameters for the $GGG$ sequence fully belongs to the corresponding domain for the $GG$ sequence. Thus the domains for $GG$ and $GGG$ base sequences are completely consistent with each other. Therefore we cannot determine parameters $\lambda$ and $b$ better then using the “dark” domain for $GGG$. This information is still sufficient to consider the localization of the hole wavefunction in $G_n$ aggregates.

Since the thermal energy $k_B T \approx 0.026$eV is smaller than other characteristic energies of the system (remember that the minimum estimate for the reorganization energy is $\lambda \sim 0.25$eV [18]) we can characterize the wavefunction using the system ground state at coordinates $X_i$ minimizing the ground state energy. In the relevant domain of parameters in Fig. 1 ($\lambda > 0.25$) the ground state wavefunction is centered at one of $G$ bases (left or right ones for the $GG$ sequence and the central one for a $GGG$ sequence). We characterize this state by the probability $P_0$ for the particle in the ground state to be in this central site. One can show that $P_0 = X_i/\lambda$, where $i$ is the central site [22].

For instance, for the $GG$ sequence the expression for the hole ground state energy at arbitrary coordinates $X_1$, $X_2$ reads $E_2 = \frac{X_1^2 + X_2^2}{2\lambda} - \frac{X_1 + X_2}{\lambda} \sqrt{(X_1^2 + X_2^2) + b^2}$. In the regime of interest $2b < \lambda$ (see Fig. 1) the minimum of energy is given by

$$E_{2\text{min}} = -\lambda/2 - b^2/\lambda,$$

and it is realized at $X_1 = \lambda/2 \pm \sqrt{(\lambda/2)^2 - b^2} = \lambda - X_2$. Accordingly $P_0 = \frac{X_1}{\lambda} = \frac{\lambda + \sqrt{\lambda^2 - 4b^2}}{2\lambda}$. Note that if $2\lambda < b$ the ground state wavefunction is symmetric in the minimum $X_1 = X_2 = \lambda/2$ and the energy of this state is given by

$$E_{2\text{symm}} = -\lambda/4 - b.$$

In the case of $2b < \lambda$ this symmetric state is the transition state (saddle point in the energy function $E_2(X_1, X_2)$) between the energy minima centered in the first and the second $G$’s.

For the $GGG$ complex the probability to reside in the central site was evaluated numerically. Both probabilities $P_0$ obtained for the $\lambda - b$ line corresponding to the ratio $r_2 = 7.7$ Eq. 2 are shown in the inset in Fig. 1. It is clear from this graph that for both $GG$ and $GGG$ sequences the wavefunction of the hole is essentially localized in the single $G$ site. For instance at the minimum value of $\lambda \sim 0.25$eV we have 85% and 78% of the probability to find the particle in that site for $GG$ and $GGG$ sequences, respectively. These probabilities increase with increasing of the reorganization energy to 1eV up to 96% and 94%, respectively.

Thus we come to the important conclusion that wavefunctions of hole are essentially localized in the single $G$ site for $G_n$ sequences. This conclusion differs from the predictions of previous work [8, 18, 19] where the polaron of an intermediate range was used to describe the quantum state of the hole. Our conclusion is justified by the agreement with the experiment [1].

An impressive consistency between $GG$ and $GGG$ in Fig. 1 is not accidental and can be explained by the strong localization of charge wavefunctions. In the regime of strong localization the partition function $Z_{n+}$ for $n \geq 2$ consists of $n$ contributions of energy minima corresponding to wavefunctions centered in all $n$ $G$ sites with coordinates $X$ realizing the corresponding energy minimum $E(X) \approx \lambda \gg X_k, k \neq i$ for the state centered at site $i$. Since in the zero order approximation in $b/\lambda$ each quantum state is localized at one site one can neglect the difference in preexponential factors for the case of $b = 0$ and we can approximate the partial $i$th contribution to the partition function as $Z_{n+} = e^{E_i - \beta E_i}$ (see Eq. 1), where $E_i$ is the energy of the ground state for coordinates $X$ realizing the local minimum. First order correction to the energy $E_i^{(0)} = -\lambda/2$ is important because it is in the exponent and multiplied by the large factor $\beta$. For the two states at the edges this correction can be expressed as $E_1 = E_n \approx -\lambda/2 - b^2/\lambda$, which coincides

FIG. 1: The domains consistent with the experimental ratios of reaction rates Eq. 1, dark grey for $GGG$ and light grey for $GG$. Inset shows the fraction of the quantum charge state belonging to the central site ($P_0$) vs. the reorganization energy $\lambda$. 

with the ground state energy for $GG$ Eq. [7]. This is not surprising because the contributions of non-neighboring sites is negligible due to the strong localization of charge. For $n - 2$ remaining states the correction to the energy should be doubled because of the addition of contributions of two neighbors so we get $E_i^{(1)} = -\lambda/2 - 2b^2/\lambda$, $1 < i < n$. Consequently, we can approximate the ratio $r_n$ (cf. Eq. [2]) as

$$r_n = Z_{n+1}/Z_n Z_{1+} \approx 2e^{3b^2/\lambda} + (n - 2)e^{2b^2/\lambda}. \quad (9)$$

Particularly, one can show that $(Z_{i+1}/Z_n Z_{1+}) \approx (Z_{i+1}/Z_2 Z_{1+}) = ((Z_{i+1}/Z_2 Z_{1+})^2) + (Z_{i+1}/Z_2 Z_{1+}^2)$ and this relationship is satisfied for the experimental values of ratios within the accuracy of the experiment. This explains the consistency of domains for $GG$ and $GGG$ Fig. 1. Using Eq. [9] one can predict that ratios $r_n$ form arithmetic series. Particularly for the balance between $G$ and $G_4$ sequence we predict the ratio $r_4 = 2r_3 - r_2 = 32.3$. This estimate agrees with our numerical calculations for the $G_4^+$ partition function. One should notice that in case of delocalization $b > \lambda/2$ the difference in energy between $G^+$ and $G_4^+$ is $b \geq 0.15eV$ and one would expect $r_2 \geq e^{b^2/\lambda} \geq 300$ which is not the case in the experiment.

5. Thus we considered the quantum state of the positive charge (hole) in poly-$G$ - poly-$C$ base sequence. It turns out that the agreement with the experimental data for the ratios $r_n$ Eq. [2] for $n = 2$, 3 can be attained only assuming the strong localization of charge within almost a single $G$-base. The charge in DNA then behaves as a small polaron with the size less than the interbase distance. Based on our theory we predict all other ratios $r_n = 7.7 + 12.3 \cdot (n - 2)$. Notwithstanding this prediction we are not able to identify more accurately the electron transfer integral $b$ and the reorganization energy $\lambda$ using experimental data only for the ratios.

Based on our theory we can suggest to find these parameters measuring the temperature dependence of the charge transfer rate through poly-$G$ - poly-$C$ base sequence. We expect that this temperature dependence will be described by the Arrhenius law with the activation energy defined by the difference of charge symmetric transition state energy within $(GG)^+$ state pair Eq. [8] and the charge ground state energy for $(GG)^+$ state Eq. [7] $E_A = \lambda/4 - b + b^2/\lambda$.

This work is supported by the NSF CRC Program, Grant No. 0628092. The authors acknowledge Fred Lewis, Michael Wasielewski, George Schatz, Thorsten Fiebig and Mark Ratner for useful discussions.

[1] F. D. Lewis, X. Liu, J. Liu, S. E. Miller, R. T. Hayes and M. R. Wasielewski, Nature, 406, 51-53 (2000); F. D. Lewis, X. Liu, J. Liu, S. E. Miller, R. T. Hayes and M. R. Wasielewski, J. Am. Chem. Soc., 122, 12037 (2000).

[2] C. J. Murphy, M. R. Arkin, Y. Jenkins, N. D. Ghatlia, S. H. Bossmann, N. J. Turro, and J. K. Barton, Science, Vol. 262, 1012 (1993).

[3] M. R. Arkin, E. D. Stemp, R. E. Holmlin, J. K. Barton, A. Hoermann, E. J. C. Olsen, P. F. Barbara, Science, 273, 475 (1996).

[4] D. B. Hall, R. E. Holmlin and J. K. Barton, Nature, 382, 831 (1996).

[5] E. Meggers, M. E. Michel-Beyerle, B. Giese, J. Amer. Chem. Soc. 120, 12950 (1998).

[6] P. T. Henderson, D. Jones, G. Hampikian G, Y. Z. Kan, G. B. Schuster, Proc. Natl. Acad. Sci. 96, 8353 (1999).

[7] S. R. Rajski, B. A. Jackson, and J. K. Barton, Mutat. Res. 447, 49 (2000).

[8] C. T. Shih, S. Roche, R. A. Romer, Phys. Rev. Lett. 100, 018105 (2008).

[9] D. Porath, A. Bezryadin, S. de Vries and C. Dekker, Nature, 403, 635 (2000); T. Takada, K. Kawai, M. Fujitsu, T. Majima, Proc. Natl. Acad. Sci. 101, 14002 (2004).

[10] H. Sugiyama, I. Saito, J. Am. Chem. Soc., 118, 7063 (1996).

[11] All states in one-dimensional random system are localized at scale $l \sim a (\Delta)^b$ where $a$ is the interatomic distance, $b$ is overlap integral and $\Delta$ is disordering in energy, e. g. B. Shapiro, Phys. Rev. Lett. 48, 823 (1982).

[12] Y. A. Berlin, A. L. Burin, M. A. Ratner, J. Phys. Chem. A 104, 443 (2000).

[13] M. Bixon, B. Giese, S. Wessely, T. Langenbacher, M. E. Michel-Beyerle, J. Jortner, Proc. Natl. Acad. Sci. 96, 11713 (1999).

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

[15] Y. A. Berlin, I. V. Kurnikov, D. Beratan, M. A. Ratner, A. L. Burin, Topics in Current Chemistry 237, 1 (2004).

[16] D. LeBard, M. Lilichenko, D. Matyushov, Y. A. Berlin and M. Ratner, J. Phys. Chem. B, 107, 14509 (2003).

[17] A. A. Voityuk, J. Chem. Phys. 122, 204904 (2005); J. Rak, A. A. Voityuk, A. Marquez, N. Rosch, J. Phys. Chem. B 106, 7919 (2002).

[18] J. A. Berashevich, T. Chakraborty, J. Phys. Chem. B 111, 13465 (2007); X. F. Wang and T. Chakraborty, Phys. Rev. Lett., 97, 106602 (2006).

[19] E. M. Conwell, D. M. Basko J. Am. Chem. Soc. 123, 11441 (2001); E. M. Conwell, et al, J. Am. Chem. Soc. 129, 9175 (2007);

[20] R. A. Marcus, Rev. Mod. Phys., 65, 599 (1993).

[21] Existing estimates of charge transfer integral should be treated with caution because this integral is also affected by high energy polar molecular vibrations, which can change their quantum state during charge tunneling. We define the efficient parameter $b$ taking into account all possible renormalizations induced by such modes (see e. g. A. L. Burin, Yu. A. Berlin, M. A. Ratner, J. Phys. Chem. A 105, 2652 (2001)).

[22] A. L. Burin, I. Y. Polishchuk, P. Fulde, Y. Sereda, Phys. Rev. Lett. 96, 025505 (2006).

[23] A. L. Burin, D. B. Uskov, preprint.