Solvent Effects on Charge Spatial Extent in DNA and Implications for Transfer

Yves A. Mantz,[*] Francesco Luigi Gervasio, Teodoro Laino,[†] and Michele Parrinello[‡]

*Computational Science, Department of Chemistry and Applied Biosciences, ETH Zurich, US I Campus, Via Giuseppe Buffi 13,
CH-6900 Lugano, Switzerland*

To clarify the role played by water in facilitating long-range DNA charge transport, carefully designed, state-of-the-art, self-interaction corrected density-functional quantum mechanical and molecular mechanical (SIC-QM/MM) simulations are performed for the first time on two ionized adenine:thymine bridge models in explicit water solvent at finite temperature. For random solvent configurations, the charge is partially delocalized. However, a charge localization on different, well-separated adenines can be induced and is correlated with a restructuring of their first solvation shells. Thus, the importance of water in the mechanism of long-range charge transport is explicitly demonstrated, and the microscopic conditions for a charge localization are revealed.

In the last decade, charge transfer processes in DNA have been the subject of intense theoretical and experimental investigations [1]. This renewed interest is due to the fact that DNA-based components are envisaged for use in molecular electronic devices [2,3] and to the role played by DNA's conductivity in oxidatively generated damage, and, possibly, repair mechanisms [4,5]. However, contradictory experimental results, with conductivities that span the entire range from metallic to insulator, have alimented a still ongoing controversy [1,6]. Single DNA molecules in dry conditions are currently thought to be wide-band-gap semiconductors [1,6], while solvated oligomers can carry charge over at least 200 Å [7,8].

Two mechanistic regimes of charge transfer have been demonstrated in solvated DNA. Over short distances, the transport mechanism is coherent [9] and can be successfully viewed as a tunneling process [10,11] or interpreted classically as radical cation migration [12]. Over long distances, the mechanism is rather different [9]. In some sequences, the rate of charge transfer is observed to be essentially independent of the adenine:thymine (A:T) bridge length [9] and can be modeled as a localized charge's thermal activation onto the A:T bridge and incoherent hopping from 1 A to another [11]. In other cases [7,8], however, an exponential decrease in the rate of charge transfer is observed over long distances. Accordingly, charge transfer can also be viewed as a delocalized charge's phonon-assisted polaronlike hopping [7], conformationally gated hopping [13–15], ion-gated transport [16], and/or response to a rearrangement of the solvation shell [17]. Importantly, the reactivity of the base radical cations formed in solvated DNA is different from that of partially hydrated dry DNA. Thus, the transport mechanisms of DNA under dry conditions may be different.

According to Marcus theory [6,18], a necessary step toward understanding charge transfer processes in DNA is describing how DNA couples to its environment. The application of this idea to complex systems, however, is quite challenging. A few years ago, the effect of a counterion rearrangement in facilitating charge transfer was shown [16]. However, recent work [13] suggests that ion-gated charge transport is not a predominant mechanism but instead that conformational changes and solvent rearrangement play a prominent role in polaron transport [19]. Our objective is to verify the latter hypothesis by observing the rearrangement of water in response to the positive charge at the molecular level.

In previous studies, a role played by water in long-range charge transport involving defect-free poly(A:T) was indirectly suggested. Specifically, the thermally activated doping of oligomers at their edges by water states was recently demonstrated, suggesting a means of transport involving damage sites in DNA [20]. Also, it was previously shown that the neutral, aperiodic poly(A:T)'s highest occupied molecular orbital (HOMO) is localized, generally attributed to dynamical variations in DNA intramolecular interactions and coupling of DNA with its environment [21]. Nonetheless, the role played by water is presently unclear. Perhaps due to this, the effect of water on the electron hole (positive charge) wave function at the molecular level is controversial and currently the subject of debate [17,19,22].

In this Letter, water-gated charge transport is elucidated at the molecular level by performing self-interaction corrected [23] density-functional quantum mechanical and molecular mechanical [24,25] (SIC-QM/MM) computations using two models of the A:T bridge in explicit water solvent. The first, termed "ideal," is the experimentally averaged x-ray structure of Arnott B-DNA. The second, labeled "disordered," is a distorted conformation taken from an AMBER [26] molecular dynamics simulation performed similarly as in reference [27]. Both $d(5'\text{-AA}_8\text{A-}3')$ models consist of 10 A:T base pairs, 18 sodium counter-
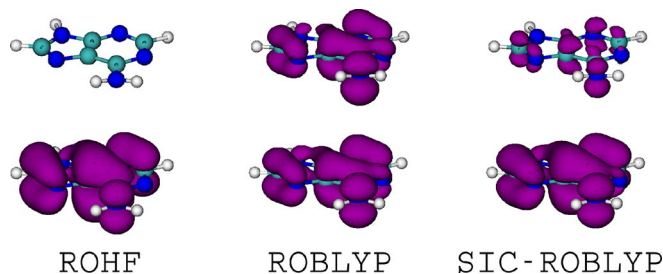
FIG. 1 (color online). In the stacked radical cation adenine dimer (6 Å separation gap), the positive charge, defined by the dark (purple online) isosurface of $m(\mathbf{r}) = 0.002$, is correctly localized at the restricted open-shell Hartree-Fock (ROHF) level of theory (left). In contrast to uncorrected ROBLYP (middle), the charge is 85% localized at the SIC-ROBLYP level (right).
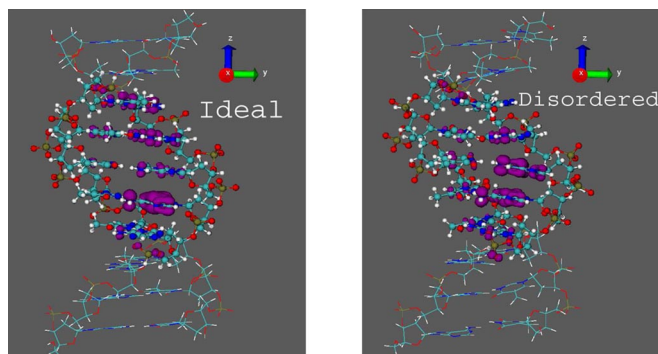

FIG. 2 (color online). The quantum subsystem (as spherical atoms) within the ideal (left) and disordered (right) models, excluding solvent molecules and counterions for clarity, and the spatial extent of the positive charge as dark (purple online) lobes, $m(\mathbf{r}) = 0.001$.

ions, and 2000 water molecules (6656 atoms total). Based on experimental [13–15] and certain theoretical [17,19] studies indicating 4–5 bases are needed to support a delocalized charge in certain sequences, 5 A:T base pairs and their sugar plus backbone atoms are initially included in the quantum subsystem (318 atoms), terminated by four capping atoms between backbone atoms.

Using both ideal and disordered models single-point SIC-QM/MM calculations are performed, followed by carefully designed Born-Oppenheimer SIC-QM/MM simulations (in which the wave function is optimized at every MD step) to allow the solvent molecules to rearrange, and then repeating the static calculations to see how the spatial extent of the positive charge is affected, quantified using the experimentally observable spin density, $m(\mathbf{r})$. All calculations are performed using CP2K [28]. As shown by ourselves [23], the self-interaction error (SIE) [29] of the unpaired electron can lead to spurious results when describing the electronic structure of stacked radical cation DNA base pairs. Accordingly, the Kohn-Sham energy functional expression [28,30] describing the quantum subsystem at the restricted open-shell DFT-Becke-Lee-Yang-Parr [31,32] level, ROBLYP, is corrected by applying a new empirical self-interaction correction (SIC) [23] with the choice of parameters $a = 0.8$ and $b = 0.5$, extensively validated (Fig. 1 and reference [23]). The remaining energy terms describing the molecular mechanical system and the coupling between QM and MM regions are standard [24,25]. Other computational details are given elsewhere [33–35].

In both ideal and disordered structures, the spin density, $m(\mathbf{r})$, is initially spread out along several bases, supporting experimental [13–15] and certain theoretical [17,19,36] studies. As shown by a selected isosurface of $m(\mathbf{r})$ (Fig. 2), for the ideal model, the charge is favoring adenine no. 5, the second base in the QM region, which is more easily ionized than thymine [23]. In the disordered model, it is shared primarily between adenines no. 5 and no. 6 (Fig. 2). Nonetheless, thymine oxidation products may still be formed due to a higher reactivity of the thymine (versus adenine) radical cation [37]. Qualitatively similar results

are obtained upon enlarging the QM region to include a 6–8 Å-thick shell of water molecules and counterions. These results are physically reasonable: In both models, the charge is favoring adenines whose electronegative nitrogen atoms, N7 and N3, are forming longer or bent hydrogen bonds with water. Most likely, the energetic destabilization of the system is minimized by depleting the electron density from these bases' nitrogen atoms, making them less effective hydrogen bond acceptors.

Next, SIC-QM/MM simulations are performed to search efficiently for solvent configurations that can induce a charge localization at different points along the A:T bridge, which would implicate water as a key player in facilitating charge transfer. After the positive charge is constrained to reside on a particular base by redefining the quantum subsystem in the ideal model to include only either adenine no. 5 or no. 8 (and a capping atom on nitrogen atom N9), runs no. 1 and no. 2 are commenced, respectively. Runs no. 3 and no. 4 are similarly carried out using the disordered model. Each run within the microcanonical (NVE) ensemble at $T = 350$ K and employing a 0.5 fs time step is performed for ≈ 5 ps, permitting several hydrogen bond
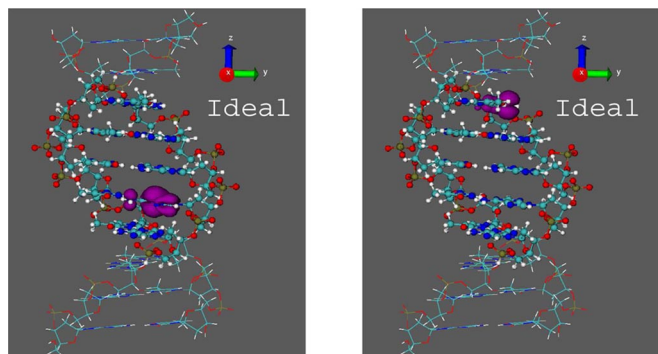

FIG. 3 (color online). Snapshots depicting the localization of the positive charge in the ideal model after run no. 1 (left) or run no. 2 (right). (Details are given in the legend of Fig. 2.) Results for the disordered model are similar.
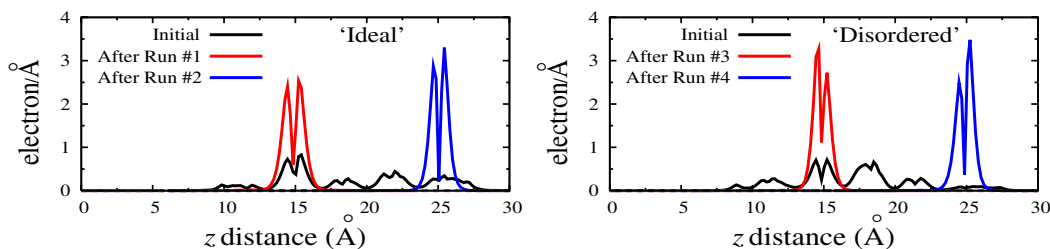
FIG. 4 (color online).   Projections of $m(\mathbf{r})$ along the z-axis for both ideal (left) and disordered (right) models of the A:T bridge.

breaking and formation events, while (both QM and MM) atoms in the original quantum subsystem of 318 atoms are held fixed. Random configurations from the last 3–4 ps of the simulations are then analyzed by recalculating $m(\mathbf{r})$ using the original definition of the quantum subsystem, e.g., Fig. 3. As shown earlier (Fig. 2), the positive charge is initially spread out in both models [Fig. 4, delocalized (black) curves]. However, in all structures chosen from the runs, it is located on a single adenine base and not shared with any others [Figure 4, localized (red or gray and blue or dark gray) curves]. These results are essentially unchanged upon expanding the quantum subsystem.

This charge redistribution (Figs. 3 and 4) is due exclusively to a local solvent rearrangement. Because the geometry of the 318-atom subsystem is frozen during the simulations, the localization of $m(\mathbf{r})$ cannot be due to bridge movement. Furthermore, during the four short runs, the counterions are always located near their initial positions around the negatively charged phosphate groups. Thus, a solvent reorganization must be responsible for this effect. Note, the global solvent arrangements are entropically driven and quite different, as measured by their net dipole moments.

During the simulations, a significant restructuring of the solvent around the electronegative nitrogen atoms of the positively charged adenines is seen. To illustrate this fact, radial distribution functions, $g_{\mathrm{N7\text{-}HW}}(r)$'s, between N7 of adenine in the ideal and disordered models and the water hydrogen atoms are examined (Fig. 5). The first peak of the $g_{\mathrm{N7\text{-}HW}}(r)$'s describing (randomly chosen) neutral adenines, shown as dashed lines, is quite pronounced and yields a coordination number of one, indicating a "strong" nitrogen-water hydrogen bond. In marked contrast, the functions describing the charged adenines, shown as solid

lines in Fig. 5, are lacking this feature, indicating a deficiency of donors. Similar conclusions apply to a second atom, N3. In conjunction with a breakdown of radial order, the distributions of the hydrogen bond angle, $\theta = \angle \mathrm{N\text{-}H\text{-}O}$, are significantly broadened and shifted away from 180° (Fig. 6). Very recently, the reorientation of water molecules around guanine was also seen to localize the charge on that base [36,38].

To verify these results, two DFT-SIC-ROBLYP calculations [28,33] are performed using a couple of adenine molecules in the Arnott B-DNA geometry and two water molecules. First, one water molecule is manually placed at a hydrogen bond distance of 1.7 Å from one adenine's N3, and a second is placed near that adenine's N7, before computing $m(\mathbf{r})$. Next, the water molecules are moved around the other adenine, and $m(\mathbf{r})$ is recomputed. In both cases, the (orbital of the) positive charge is localized on the base without water molecules.

Thus, the role played by water in gating charge transport is explicitly demonstrated for the first time and is envisioned to involve the breaking and formation of water-purine nitrogen hydrogen bonds in the first solvation shell occurring on the order of the simulation time scale (ps). This process is expected to compete with reaction (e.g., deprotonation, hydration) of the base radical cations and is open to experimental verification. Also, the microscopic conditions for a charge localization on bases with longer or bent hydrogen bonds are revealed, helping to reconcile seemingly contradictory theoretical predictions that the positive charge can be both delocalized [17] and localized [22]. These accomplishments are made by applying a self-interaction correction to DFT, representing an unprecedented level of accuracy in treating large (biological) systems.
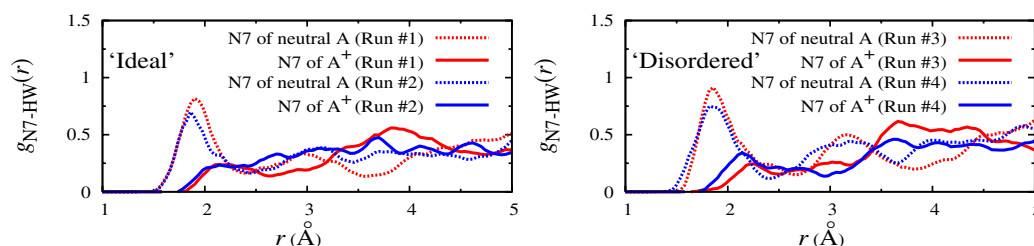

FIG. 5 (color online).   Radial distribution functions calculated for ideal (left) and disordered (right) models.
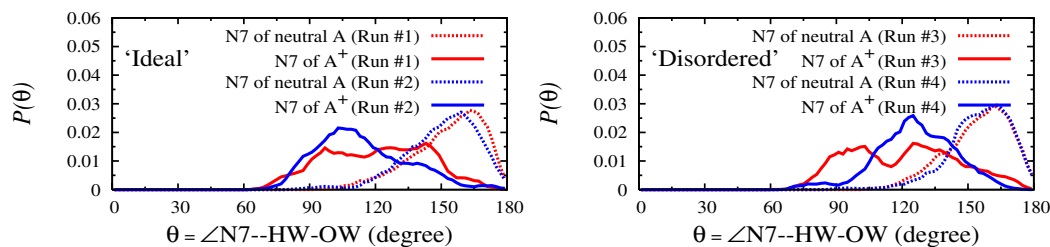
FIG. 6 (color online).   Angular distribution functions calculated for ideal (left) and disordered (right) models.

*ymantz@phys.chem.ethz.ch

†Present address: Physical Chemistry Institute, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

‡parrinello@phys.chem.ethz.ch

[1] *Long-Range Charge Transfer in DNA I-II: Topics in Current Chemistry*, edited by G. B. Schuster (Springer, Heidelberg, 2004), Vol. 236–237.

[2] N. C. Seeman, Nano Lett. **1**, 22 (2001).

[3] M. Di Ventra and M. Zwolak, in *Encyclopedia of Nanoscience and Nanotechnology*, edited by H. Singh-Nalwa (American Scientific Publishers, New York, 2004).

[4] E. M. Boon, A. L. Livingston, N. H. Chmiel, S. S. David, and J. K. Barton, Proc. Natl. Acad. Sci. U.S.A. **100**, 12543 (2003).

[5] M. R. Holman, T. Ito, and S. E. Rokita, J. Am. Chem. Soc. **129**, 6 (2007).

[6] R. G. Endres, D. L. Cox, and R. R. P. Singh, Rev. Mod. Phys. **76**, 195 (2004).

[7] P. T. Henderson, D. Jones, G. Hampikian, Y. Kan, and G. B. Schuster, Proc. Natl. Acad. Sci. U.S.A. **96**, 8353 (1999).

[8] M. E. Nùñez, D. B. Hall, and J. K. Barton, Chemistry & Biology **6**, 85 (1999).

[9] B. Giese, J. Amaudrut, A.-K. Kohler, M. Spormann, and S. Wesserly, Nature (London) **412**, 318 (2001).

[10] J. Jortner, M. Bixon, T. Langenbacher, and M. E. Michel-Beyerle, Proc. Natl. Acad. Sci. U.S.A. **95**, 12759 (1998).

[11] M. Bixon and J. Jortner, Chem. Phys. **326**, 252 (2006).

[12] C.-S. Liu, R. Hernandez, and G. B. Schuster, J. Am. Chem. Soc. **126**, 2877 (2004).

[13] M. A. O'Neill and J. K. Barton, J. Am. Chem. Soc. **126**, 11471 (2004).

[14] F. Shao, M. A. O'Neill, and J. K. Barton, Proc. Natl. Acad. Sci. U.S.A. **101**, 17914 (2004).

[15] F. Shao, K. Augustyn, and J. K. Barton, J. Am. Chem. Soc. **127**, 17445 (2005).

[16] R. N. Barnett, C. L. Cleveland, A. Joy, U. Landman, and G. B. Schuster, Science **294**, 567 (2001).

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

[18] R. A. Marcus, J. Chem. Phys. **24**, 966 (1956).

[19] E. M. Conwell, Proc. Natl. Acad. Sci. U.S.A. **102**, 8795 (2005).

[20] A. Hübsch, R. G. Endres, D. L. Cox, and R. R. P. Singh, Phys. Rev. Lett. **94**, 178102 (2005).

[21] J. P. Lewis, T. E. Cheatham, III, E. B. Starikov, H. Wang, and O. F. Sankey, J. Phys. Chem. B **107**, 2581 (2003).

[22] A. A. Voityuk, J. Phys. Chem. B **109**, 10793 (2005).

[23] Y. A. Mantz, F. L. Gervasio, T. Laino, and M. Parrinello, J. Phys. Chem. A **111**, 105 (2007).

[24] T. Laino, F. Mohamed, A. Laio, and M. Parrinello, J. Chem. Theory Comput. **1**, 1176 (2005).

[25] T. Laino, F. Mohamed, A. Laio, and M. Parrinello, J. Chem. Theory Comput. **2**, 1370 (2006).

[26] D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, *et al.*, computer code AMBER 8 (University of California, San Francisco, 2004).

[27] R. Walker, computer code AMBER 8 Tutorial: DNA polyA-polyT Decamer, http://amber.scripps.edu/tutorial/polyApolyTNew/.

[28] J. VandeVondele, M. Krack, F. Mohamad, M. Parrinello, T. Chassaing, and J. Hutter, Comput. Phys. Commun. **167**, 103 (2005).

[29] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[30] J. VandeVondele and M. Sprik, Phys. Chem. Chem. Phys. **7**, 1363 (2005).

[31] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[32] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[33] The orbital transformation algorithm [34] is used; core electrons and nuclei are treated with Goedecker-Teter-Hutter pseudopotentials [35]; and valence electrons are described with a double-ζ valence basis set with polarization functions (DZVP) [30] and an auxiliary plane-wave basis set with a density cutoff of 280 Ry, both checked for convergence.

[34] J. VandeVondele and J. Hutter, J. Chem. Phys. **118**, 4365 (2003).

[35] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).

[36] E. M. Conwell and D. M. Basko, J. Phys. Chem. B **110**, 23603 (2006).

[37] A. Joy, A. K. Ghosh, and G. B. Schuster, J. Am. Chem. Soc. **128**, 5346 (2006).

[38] R. N. Barnett, A. Bongiorno, C. L. Cleveland, A. Joy, U. Landman, and G. B. Schuster, J. Am. Chem. Soc. **128**, 10795 (2006).

[39] W. Humphrey, A. Dalke, and K. Schulten, VMD 1.8.2, J. Mol. Graphics **14**, 33 (1996).