

Long-range Stacking Effects of Nucleobases in Charge Transfer

Zhong-wei Li and Ke-li Han*

State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Science, Dalian 116023 (P. R. China)

E-mail: klhan@dicp.ac.cn

Abstract

Base-stacked structure is an important feature of DNA molecules. Previous studies on the stacking effect concerning DNA-mediated hole transfer have revealed the influence of neighboring bases on onsite energies. But the neighboring base effect acts only in a short-distance. Besides it, a long-range (longer than three base pairs) stacking effect called squeezing effect in this paper has not yet been reported. Such a squeezing effect causes the bases near the middle of a sequence consisting of same type base pairs have lower onsite energies than the bases near the terminals. We predict it by Hückel analysis in an unconventional way and confirmed it by semiempirical calculations combined with molecular dynamics simulations. The results suggest that in order to obtain a reasonable onsite energy map when study charge transfer on DNA, the stacking effects should be considered in a long-distance as possible. The consideration of squeezing effect also provides a new suggestion on the driving force of fluctuation-assisted DNA charge transfer. The method used to calculate the onsite energies in a base stack can be generalized to other π -stacked systems.

*To whom correspondence should be addressed

INTRODUCTION

In cell, DNA-mediated charge transfer (DNA CT) is found to transmit biosignals acting as live wire and the function has been found to play an amazing role in gene repair, replication, and transcription.¹⁻³ DNA damage caused by ultraviolet is also found to be associated with radiation-induced hole transfer (HT).⁴ Besides the biosystems, it is also thought that DNA conductivity has potential application in nanotechnology.⁵ These behaviors attract a lot of attention on the mechanism of DNA CT. Due to the difficulty of direct observation in cell, scientists have synthesized various DNA conjugates to measure their abilities of transmitting a charge in solutions. These systems include Sa-Sd (stilbenedicarboxamide-stilbenediether) series,⁶⁻¹³ Ap (2-aminopurine)¹⁴⁻¹⁷ and E (ethidium) series,^{18,19} NDI (phenothiazine)^{20,21} and PTZ (phenothiazine) series,²² etc. Either way, it is crucial to obtain a reasonable energy relationship of the donor-bridge-acceptor system because it is relevant to what mechanism CT obeys, e.g. superexchange or multi-step hopping.²³

However, due to structural fluctuations and base stacking structures of DNA, the energy relationship involved in charge transfer is complicated because that onsite energy of a charge in a stack is different from that when the site molecules are isolated. A typical example is the fact that GG doublets and GGG triplets have lower ionization potential (IP) values than the isolated G bases.^{24,25} And subsequent studies were reported concerning the change of the onsite energy of hole on a nucleobase (B) when it is in different triplets, denoted as XB^+Y , where X and Y represent the nearest neighboring bases of B.²⁵⁻²⁸ In addition, the onsite energy difference generated by polarization effect²⁹ between two A bases is thought a driving force of hole hopping along a consecutive A-contained sequence as it fluctuates.³⁰ This is essentially also a stacking effect.

Besides the inherent properties of base stacks, consideration of the environment also affects the energy relationship between nucleobases. For example, Grozema et al.³⁰ used a gradually rising onsite energy map by considering the Coulomb interaction between donor and bridge to explain the mechanism conversion from superexchange to hopping. Voityuk³¹

and Kubar et al.^{32,33} took the solvent effect into account to explain the driving force of hole hopping from G base to A base, and the latter has a higher IP.

On the other hand, report on the stacking effect in a longer distance is relatively rare since the interactions between non-adjacent bases are thought to be ignorable. But meanwhile it is also noticed that in some theoretical studies a strong or weak correlation exists between non-adjacent bases.³¹⁻³³ An intuitive performance is that the nucleobases have different predicted onsite energies due to their different locations in a stack even it consists of same type base pairs, e.g., denoted as $B \cdots BB^+B \cdots B$. Since it is just a numerical result without direct experimental evidence and no more theoretical analysis, little discussion was carried out on it.

At the beginning of this work, we employ Hückel theory to analyze π -stacked system for a qualitative understanding on its bonding nature. Hückel theory had been used on the stacked nucleobases by Jakobsson et al.³⁴ but we here present a different discussion. As a result, a general relationship of onsite energies in a stack consisting of same type bases is derived. Concretely, it is a curved onsite energy map that, take HT for example, the bases in the middle region of the stack should have lower onsite energies than that at the terminals.

To verify this qualitative prediction, classical molecular dynamics (MD) simulations combining with quantum mechanism (QM) calculations are performed on fluctuating consecutive A:T sequences and the results support Hückel analysis. In order to distinguish this long-range stacking effect from those act in a short distance, we call it “squeezing” effect in this work. Of course, they are essentially caused by the base-stacking structure. Although the squeezing effect can be exhibited by high level *ab initio* and DFT (density functional theory) methods by our calculations on tested short sequences containing three to six base pairs, we here only show semiempirical (MNDO³⁵) results for its feasibility to longer DNA strands.

The influences of the neighboring base effect and the solvent effect on the squeezing effect are also discussed. The solvent can weaken the squeezing effect but not eliminate it. Our work shows that long-range stacking effect affects the onsite energy map significantly so it

is unignorable in DNA CT studies. Besides, the potential well generated by the squeezing effect provides a new suggestion for the driving force of DNA CT. Incidentally, some technical skills are also developed at calculating the onsite energy in a stack and can be generalized to other complex molecular system having charge transfer behaviors.

METHODOLOGY

Hückel analysis is discussed in the next section. In this section we briefly introduce MD simulations and the calculation of onsite energy in a stack.

Classical MD simulations are performed to give DNA fluctuating trajectories in solution. Concretely, BSC1 DNA force field and TIP3P water force field in Amber16 package are used.³⁶ Sodium cations are added to neutralize negative charge of phosphate in skeleton. 2 ns MD (300 K and 1 atm) with time step of 2 fs is performed for each DNA sequence of interest. The time interval of two outputted adjacent MD points is set 0.2 ps referring to previous studies.^{30,33} Thus 10000 consecutive MD points are obtained for each sequence. To reduce the computational cost without losing the description of dynamic disorder of structural fluctuations, 500 snapshots are randomly picked up for subsequent QM calculations. Wherein the sugar-phosphate backbones are removed since CT is assumed occurs only via the stacked nucleobases.^{23,37}

Based on the fixed nuclei configurations, QM calculations are performed to evaluate the onsite energy. According to the tight-binding approximation, suppose that the wave function of the charge can be expressed as a linear combination of the frontier molecular orbitals on the individual base pairs,^{28,30,38,39} the onsite energy of a charge localized at site i can be evaluated by⁴⁰

$$\epsilon_i^{\text{HT}} = -\langle \psi_i^{\text{HOMO}} | \hat{F} | \psi_i^{\text{HOMO}} \rangle \quad (1)$$

for HT and

$$\epsilon_i^{\text{EET}} = \langle \psi_i^{\text{LUMO}} | \hat{F} | \psi_i^{\text{LUMO}} \rangle \quad (2)$$

for EET (excess electron transfer). Wherein \hat{F} is the Fock operator of the charge in the stacked system and ψ_i the frontier molecular orbital localized on i -th base pair. HOMO (the highest occupied molecular orbital) is used for HT while LUMO (the lowest unoccupied molecular orbital) for EET. Charge localization on a single base in DNA has been an accepted model for many years.^{24,41} Renaud et al. recently introduced the concept of partial localization (i.e., the charge is partial localized on a single base while others delocalized over several base pairs) to improve theoretical CT rates in comparison with experiments.⁴² Here we use traditional localized model with a trick that one site is set as a base pair rather a single base.

For practical calculations of ϵ , Senthilkumar et al.^{27,28,30,43} employed the charge transfer integrals module included in the Amsterdam Density Functional (ADF) theory program.⁴⁴ We here refer to the projective method proposed by Kirkpatrick.⁴⁵ First, for a n -stacking system, construct the coefficient matrix of C_{loc} which is consisted of the eigenvectors of the localized MOs,

$$C_{\text{loc}} = \begin{bmatrix} c_{m_1, \psi_1}^{\chi_1} & c_{m_1, \psi_1}^{\chi_2} & \cdots & & \\ c_{m_1, \psi_2}^{\chi_1} & \cdots & \cdots & & \\ \cdots & \cdots & \cdots & & \\ & & & \cdots & \\ & & & & c_{m_n, \psi_1}^{\chi_1} & c_{m_n, \psi_1}^{\chi_2} & \cdots \\ & & & & c_{m_n, \psi_2}^{\chi_1} & \cdots & \cdots \\ & & & & \cdots & \cdots & \cdots \end{bmatrix}, \quad (3)$$

wherein $c_{m_i, \psi_j}^{\chi_k}$ means the MO coefficient of the j -th MO on the k -th atomic orbital (AO) basis of the site i molecule. The blank is filled with zeros. Then, the MO coefficients of the whole stack in the basis of the localized MOs can be projected onto C_{loc} by

$$C_{\text{sys}}^{\text{loc}} = C_{\text{loc}} S_{\text{sys}}^{\text{AO}} C_{\text{sys}}^{\text{T}}, \quad (4)$$

wherein C_{sys} are the MO coefficients of the whole system in the same AO basis set as C_{loc} use. And $S_{\text{sys}}^{\text{AO}}$ is the overlap matrix. Note that Eq. (4) is different from the equation used by Kirkpatrick et al. in which orthogonal AO basis is assumed.^{45,46} Finally, the Fock matrix of the stack in the basis of localized MOs can be formed by

$$F_{\text{sys}}^{\text{loc}} = C_{\text{sys}}^{\text{loc}} \epsilon_{\text{sys}} C_{\text{sys}}^{\text{loc}^T}, \quad (5)$$

wherein ϵ_{sys} is the diagonal matrix of MO eigenvalues of the stack. ϵ defined in Eqs. (1) and (2) can be directly picked up from $F_{\text{sys}}^{\text{loc}}$. Electronic structure calculations are performed in GAMESS package.⁴⁷

RESULTS AND DISCUSSIONS

Hückel Analysis on π -Stacked Systems

Since that all the HOMOs as well as the LUMOs of the four nucleobases are π orbitals, it is easy to build a model of the π channel for a n -stacked nucleobases of same type by Hückel theory with the tight-binding approximation. Hückel bases, bond matrix, solutions and MO shapes are attached in the SI. As a result, the variation of energy levels of the Hückel MOs for n -stack ($n=2\sim 6$) is plotted as in Figure 1. From the MO phases (see Figure S2) it can be known that what we should pay attention to is just ϕ_1 . Because that of the n MOs only ϕ_1 actually reflects the nature of the stacking structure what is a linear combination of individual π orbitals via the weak interaction, while others are actually covalent bonds what could be hard to form in a stack. This is different from the discussion of Jakobsson et al. in which all the Hückel MOs were used for CT dynamics simulations.³⁴ It can also be seen from Figure 1 that as n increases, the energy level of ϕ_1 becomes higher and higher due to a stronger antibonding interaction. This is qualitatively consistent with the fact that GG doublets and GGG triplets have lower ionization potential than the isolated G bases.^{24,25}

Together with the shape variation of ϕ_1 (as shown in Figure 2) a more general conclusion can be inferred. Namely, as the basis in the middle position of the stack contributes the most to the energy increase, it is in turn implicating that the energy level of the localized π orbital on the middle site should raise higher than that at the ends. So far, since no assumption that the charge is positive or negative is made, it is generalized for HT and EET. But for the following discussion concerned, it means HT if there is no special explanation. For HT along the nucleobases of same type, it means that the base in the middle of the stack has lower onsite energy than the bases at the terminals, what can be seen more straightly from QM calculations in the next section.

Squeezing Effect

The averaged values during 2 ns of the fluctuating onsite energies for a hole in the $[A:T]_n$ systems predicted by MNDO are shown in Figure 3. For the sake of clarity, relative values of three stacks i.e. $[A:T]_3$, $[A:T]_6$, and $[A:T]_{10}$ are exhibited. The curved pattern of onsite energies appears clearly in all the examples. And the “well” becomes deeper as the stack grows. For $n=3$, the energy gap between terminals and the middle is around 0.1 eV. As n increases to 10, the gap enlarges to 0.3 eV. At the current computational level, it is even comparable to the IP gaps between the four nucleobases (MNDO predicted relative $\langle\epsilon_i^{\text{HT}}\rangle$ for the four isolated bases are respectively A 0.38 eV, C 0.76 eV and T 1.16 eV with respect to G, which are reasonable in comparison with the values predicted by SAOP/TZ2P method²⁷ and experimental results⁴⁸). In addition, 5'-A exhibits a lower onsite energy than 3'-A. It is not surprising if we note the experiments that 5'-G was found to be more easily ionized than 3'-G.⁴⁹ But we are not to discuss this more here.

The squeezing effect implicates a driving force neglected previously. In the suggested mechanisms of fluctuation-assisted DNA CT, the driving force was found to act via the following ways: (i) Strong electronic couplings induced by interbase motions.⁵⁰ (ii) Fluctuating onsite energy gaps generated by the polarization effect between neighboring bases.^{30,31}

(iii) Electrostatic potentials from the environment.^{32,33,51} By considering one or more factors of them, the proposed models include ion-gated CT,⁵¹ solvent fluctuations driven CT,³³ phonon-assisted polaron-like hopping CT,^{41,52} and CT-active conformations gated CT,¹⁴ et al. In the last two models, both have a well-like domain consists of approximately 3~5 base pairs but have different meanings from the well in Figure 3. In the phonon-assisted polaron-like hopping mechanism, an energy minimum is thought to be created due to a structural distortion induced by charge delocalization. And the CT-active conformation means a well-coupled domain with a period of 4~5 base pairs over which the charge can be delocalized on transiently extended π -orbitals. A charge is required in the former while not in the latter. Either way, the concept of charge delocalization over several base pairs is supposed, what actually remains a debate. Consideration of squeezing effect avoids the assumption of charge delocalization. In fact, during a period in which CT occurs, it is difficult to maintain a well-stacked conformation in a long range due to the flexibility of DNA (see Figure 4). More likely, a potential well of onsite energies is formed in a short range and provides a wave-like driving force to transmit the charge to a distance as the strand fluctuates.

We found no obvious influence of squeezing effect on charge transfer integrals so are not to discuss much about it.

Neighboring Base Effect

As above, the squeezing effect has been clearly exhibited on $[A:T]_n$ systems. However, the diversity of DNA sequences makes the squeezing effect not appear as explicit as it is in consecutive A:T sequences. Another influence on onsite energy map is the effect of neighboring bases. Knowing that $X-A_n-Y$ (X, Y are G or modified G bases) systems are usually used as models to study DNA HT experimentally,^{10,12,53-60} we take $G:C[A:T]_n[G:C]_2$ and $C:G[A:T]_n[C:G]_2$ systems as examples to illustrate the variety of onsite energy maps due to stacking effects (Figure 5). As Figure 5a shows, since G bases are in the different strand from A bases, it is in fact C bases that form a π stacked structure with A bases. This makes

the onsite energies of neighboring A bases rise.^{26,28} Thus the part of the onsite energy map of consecutive A:T bases remains the pattern as Figure 3 shows. On the other hand, Figure 5b shows a different pattern. G bases are in the same strand with A bases and cause the onsite energies of neighboring A bases down. This makes the local energy minimum at the middle of a consecutive A:T sequence transform to an energy maximum. Another difference is the energy order of G' and G'' . In Figure 5a, neighboring T base causes the onsite energy of G' rise and G'' is the hole acceptor. While in Figure 5b, G' becomes the hole acceptor due to the squeezing effect of A- G' - G'' .

Different onsite energy maps implicate different charge dynamics. Suppose a hole is initially generated at the first G base and then transported to a distant G' or G'' acceptor via the $[A]_n$ bridge, $C:G[A:T]_n[C:G]_2$ and $G:C[A:T]_n[G:C]_2$ systems should have different hole dynamics. In Figure 5b, if the hole is transmitted by hopping mechanism, it has to overcome an energy barrier like climbing a mountain. At each step before reaching the peak in the middle of the sequence, the hole have a high probability of returning and result a low rate of arrival at the acceptor. In Figure 5a the situation is improved as in the middle is a saddle point which is helpful to decrease the difficulty of hole arrival at the acceptor. This might be an explanation of that why most of the experiments on long-distance hole transfer are carried out on $C:G[A:T]_n[C:G]_2$ systems rather than on $G:C[A:T]_n[G:C]_2$ systems,⁵⁹ and why a hole has a faster rate of arrival at G when G is at the different strand from A.^{10,54,56,58}

Influence of Solvent on the Stacking Effects

The above calculations are in vacuo model. It has been theoretically found that consideration of electrostatic interactions from the environment (water, ions etc.) could reduce the IP difference of G-A.³¹⁻³³ Advanced simulations of the environment require QM/MM methods. For simplicity, we use polarizable continuum model (PCM)⁶¹ to consider the solvent effect. It is not a stacking effect but could have influence on the stacking effects. Concretely, we calculate C_{loc} and C_{sys} (meanwhile S_{sys}^{AO} and ϵ_{sys}) in Eqs. (4) and (5) with PCM and then

use them to obtain the onsite energy.

By such a treatment, the onsite energy maps in Figure 5 transform to that as Figure 6 shows. Corresponding values in the figures are collected in table 1. It can be seen from the two Figures that the IP difference between the G:C (or C:G) pair and its neighboring A:T pair indeed reduced from ~ 0.8 eV to ~ 0.55 eV at the current computational level, what is helpful to G-A hopping. However, the depth of the well and the height of the arch formed by $[A:T]_n$ regions are reduced from ~ 0.3 eV to ~ 0.03 eV, which nearly makes the squeezing effect vanished. Such an influence is actually overestimated in our model. Because when construct C_{loc} of Eq. (3), PCM is added to each individual base pair and the unique effects of environment on individuals are eliminated. In spite of that the squeezing effect is still exhibited. A more sophisticated model of the environment may remain the squeezing effect more while reduce the IP difference of G-A. So we retain previous discussions about the influence of stacking effects on hole dynamics.

CONCLUSIONS

In this work we proposed a long-range stacking effect of nucleobases called squeezing effect. Previous studies on the stacking effects usually focus on the effect from neighboring bases so the squeezing effect was neglected. We used Hückel analysis to predict the existence of it and verified by performing MD/MNDO calculations on consecutive A:T sequences. One consequence of it is that in DNA-mediated hole transfer, the bases near the middle of the sequence have lower onsite energies than the bases near the terminals and the depth of the potential well can reach $0.1\sim 0.3$ eV in vacuo. Although considering the solvent effect can reduce the depth but not eliminate it. In addition, the combination of the squeezing effect with neighboring base effect (e.g. embedding a G:C pair into the A:T sequence) makes the relationship of onsite energies in a sequence more complicated. Given that small changes in onsite energy can significantly affect charge dynamics, we recommend taking stacking effects

into account in a long-distance as possible as one can when study DNA CT.

The squeezing effect also provides a new suggestion for DNA CT mechanism. A charge carried by the well can be transported to a distant as DNA fluctuates, just like on a wave. From this perspective, the so-called incoherent hopping is actually somewhat coherent. Further studies on long-range stacking effects on DNA CT require creditable charge dynamics, what is under developing in our group.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (Grant No. 21703234).

Supporting Information Available

This material is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Service, R. F. *Science* **2014**, *346*, 1284–1287.
- (2) Grodick, M. A.; Muren, N. B.; Barton, J. K. *Biochemistry* **2015**, *54*, 962–973.
- (3) O’Brien, E.; Holt, M. E.; Thompson, M. K.; Salay, L. E.; Ehlinger, A. C.; Chazin, W. J.; Barton, J. K. *Science* **2017**, *355*, eaag1789.
- (4) Hall, D. B.; Holmlin, R. E.; Barton, J. K. *Nature* **1996**, *382*, 731–735.
- (5) Endres, R. G.; Cox, D. L.; Singh, R. R. P. *Rev. Mod. Phys.* **2004**, *76*, 195–214.
- (6) Conron, S. M. M.; Thazhathveetil, A. K.; Wasielewski, M. R.; Burin, A. L.; Lewis, F. D. *J. Am. Chem. Soc.* **2010**, *132*, 14388–14390.

- (7) Lewis, F.; Daublain, P.; Cohen, B.; VuraWeis, J.; Wasielewski, M. *Angew. Chem. Int. Ed.* **2008**, *47*, 3798–3800.
- (8) Lewis, F. D.; Zhu, H.; Daublain, P.; Fiebig, T.; Raytchev, M.; Wang, Q.; Shafirovich, V. *J. Am. Chem. Soc.* **2006**, *128*, 791–800.
- (9) Lewis, F. D.; Zhu, H.; Daublain, P.; Cohen, B.; Wasielewski, M. R. *Angew. Chem. Int. Ed.* **2006**, *45*, 7982–7985.
- (10) Lewis, F. D.; Liu, J.; Zuo, X.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2003**, *125*, 4850–4861.
- (11) Lewis, F. D.; Liu, J.; Weigel, W.; Rettig, W.; Kurnikov, I. V.; Beratan, D. N. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12536–12541.
- (12) Lewis, F. D.; Zuo, X.; Liu, J.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2002**, *124*, 4568–4569.
- (13) Lewis, F. D.; Liu, X.; Miller, S. E.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2002**, *124*, 11280–11281.
- (14) O'Neill, M. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 11471–11483.
- (15) O'Neill, M. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 13234–13235.
- (16) O'Neill, M. A.; Becker, H.; Wan, C.; Barton, J. K.; Zewail, A. H. *Angew. Chem. Int. Ed.* **2003**, *42*, 5896–5900.
- (17) Wan, C.; Fiebig, T.; Schiemann, O.; Barton, J. K.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 14052–14055.
- (18) Valis, L.; Wang, Q.; Raytchev, M.; Buchvarov, I.; Wagenknecht, H.-A.; Fiebig, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10192–10195.

- (19) Wan, C.; Fiebig, T.; Kelley, S. O.; Treadway, C. R.; Barton, J. K.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6014–6019.
- (20) Renaud, N.; Harris, M. A.; Singh, A. P. N.; Berlin, Y. A.; Ratner, M. A.; Wasielewski, M. R.; Lewis, F. D.; Grozema, F. C. *Nat. Chem.* **2016**, *8*, 1015–1021.
- (21) Takada, T.; Kawai, K.; Cai, X.; Sugimoto, A.; Fujitsuka, M.; Majima, T. *J. Am. Chem. Soc.* **2004**, *126*, 1125–1129.
- (22) Kawai, K.; Takada, T.; Tojo, S.; Majima, T. *J. Am. Chem. Soc.* **2003**, *125*, 6842–6843.
- (23) Genereux, J. C.; Barton, J. K. *Chem. Rev.* **2010**, *110*, 1642–1662.
- (24) Sugiyama, H.; Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 7063–7068.
- (25) Saito, I.; Nakamura, T.; Nakatani, K.; Yoshioka, Y.; Yamaguchi, K.; Sugiyama, H. *J. Am. Chem. Soc.* **1998**, *120*, 12686–12687.
- (26) Voityuk, A. A.; Jortner, J.; Bixon, M.; Rsch, N. *Chem. Phys. Lett.* **2000**, *324*, 430–434.
- (27) Senthilkumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2003**, *125*, 13658–13659.
- (28) Senthilkumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Lewis, F. D.; Berlin, Y. A.; Ratner, M. A.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2005**, *127*, 14894–14903.
- (29) Valeev, E. F.; Coropceanu, V.; da Silva Filho, D. A.; Salman, S.; Brdas, J.-L. *J. Am. Chem. Soc.* **2006**, *128*, 9882–9886.
- (30) Grozema, F. C.; Tonzani, S.; Berlin, Y. A.; Schatz, G. C.; Siebbeles, L. D. A.; Ratner, M. A. *J. Am. Chem. Soc.* **2008**, *130*, 5157–5166.
- (31) Voityuk, A. A.; Siriwong, K.; Rsch, N. *Angew. Chem. Int. Ed.* **2004**, *43*, 624–627.

- (32) Kubar, T.; Elstner, M. *J. Phys. Chem. B* **2008**, *112*, 8788–8798.
- (33) Kubar, T.; Kleinekathofer, U.; Elstner, M. *J. Phys. Chem. B* **2009**, *113*, 13107–13117.
- (34) Jakobsson, M.; Stafström, S. *J. Chem. Phys.* **2008**, *129*, 125102.
- (35) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comput. Chem.* **2006**, *27*, 1101–1111.
- (36) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2309–2309.
- (37) Gorodetsky, A. A.; Green, O.; Yavin, E.; Barton, J. K. *Bioconjugate Chem.* **2007**, *18*, 1434–1441.
- (38) Siebbeles, L. D.; Berlin, Y. A. *Chem. Phys.* **1998**, *238*, 97–107.
- (39) Grozema, F. C.; Berlin, Y. A.; Siebbeles, L. D. A. *Int. J. Quantum Chem.* **1999**, *75*, 1009–1016.
- (40) Newton, M. D. *Chem. Rev.* **1991**, *91*, 767–792.
- (41) Henderson, P. T.; Jones, D.; Hampikian, G.; Kan, Y.; Schuster, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8353–8358.
- (42) Renaud, N.; Berlin, Y. A.; Lewis, F. D.; Ratner, M. A. *J. Am. Chem. Soc.* **2013**, *135*, 3953–3963.
- (43) Senthilkumar, K.; Grozema, F. C.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Chem. Phys.* **2003**, *119*, 9809–9817.
- (44) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931–967.

- (45) Kirkpatrick, J. *Int. J. Quantum Chem.* **2008**, *108*, 51–56.
- (46) Reha, D.; Barford, W.; Harris, S. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5436–5444.
- (47) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
- (48) Hush, N.; Cheung, A. S. *Chem. Phys. Lett.* **1975**, *34*, 11–13.
- (49) Saito, I.; Takayama, M.; Sugiyama, H.; Nakatani, K.; Tsuchida, A.; Yamamoto, M. *J. Am. Chem. Soc.* **1995**, *117*, 6406–6407.
- (50) Voityuk, A. A. *J. Chem. Phys.* **2008**, *128*, 045104.
- (51) Barnett, R. N.; Cleveland, C. L.; Joy, A.; Landman, U.; Schuster, G. B. *Science* **2001**, *294*, 567–571.
- (52) Ly, D.; Sanii, L.; Schuster, G. B. *J. Am. Chem. Soc.* **1999**, *121*, 9400–9410.
- (53) Meggers, E.; Michel-Beyerle, M. E.; Giese, B. *J. Am. Chem. Soc.* **1998**, *120*, 12950–12955.
- (54) Giese, B.; Wessely, S.; Spormann, M.; Lindemann, U.; Meggers, E.; Michel-Beyerle, M. E. *Angew. Chem. Int. Ed.* **1999**, *38*, 996–998.
- (55) Nakatani, K.; Dohno, C.; Saito, I. *J. Am. Chem. Soc.* **1999**, *121*, 10854–10855.
- (56) Williams, T. T.; Odom, D. T.; Barton, J. K. *J. Am. Chem. Soc.* **2000**, *122*, 9048–9049.
- (57) Nakatani, K.; Dohno, C.; Saito, I. *J. Am. Chem. Soc.* **2000**, *122*, 5893–5894.
- (58) Lewis, F. D.; Letsinger, R. L.; Wasielewski, M. R. *Acc. Chem. Res.* **2001**, *34*, 159–170.
- (59) Giese, B.; Amaudrut, J.; Khler, A.; Spormann, M.; Wessely, S. *Nature* **2001**, *412*, 318–320.

- (60) Lewis, F. D.; Liu, J.; Liu, X.; Zuo, X.; Hayes, R. T.; Wasielewski, M. R. *Angew. Chem. Int. Ed.* **2002**, *41*, 1026–1028.
- (61) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3094.

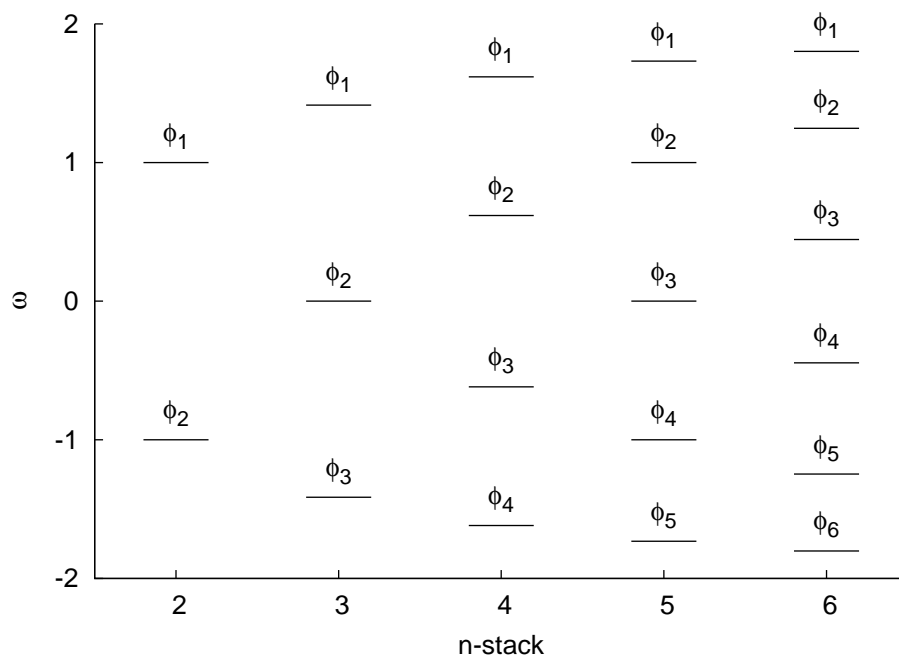


Figure 1: MO energy levels (indicated by ω) of n -mer π -stacked systems predicted by Hückel theory. MO shapes see Figure S2.

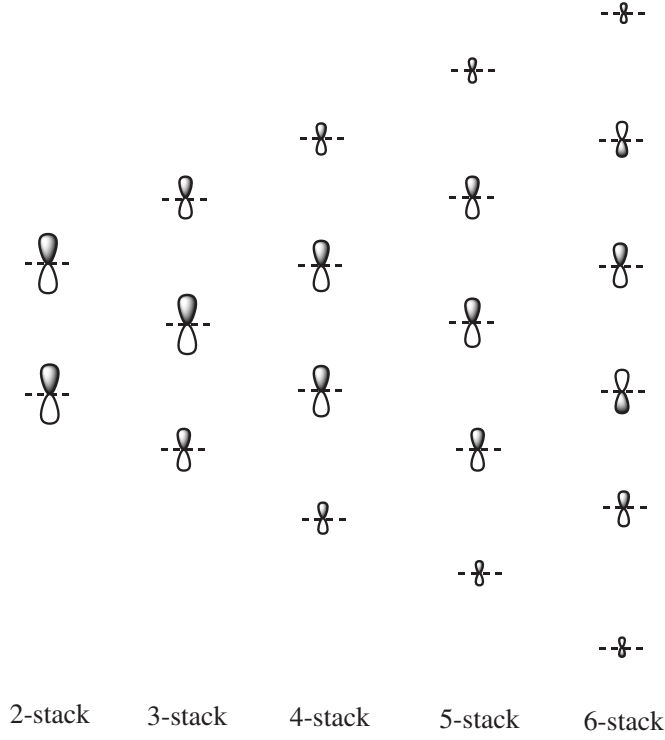


Figure 2: Diagrams of the highest orbitals (ϕ_1) in a n -mer π -stacked systems predicted by Hückel theory.

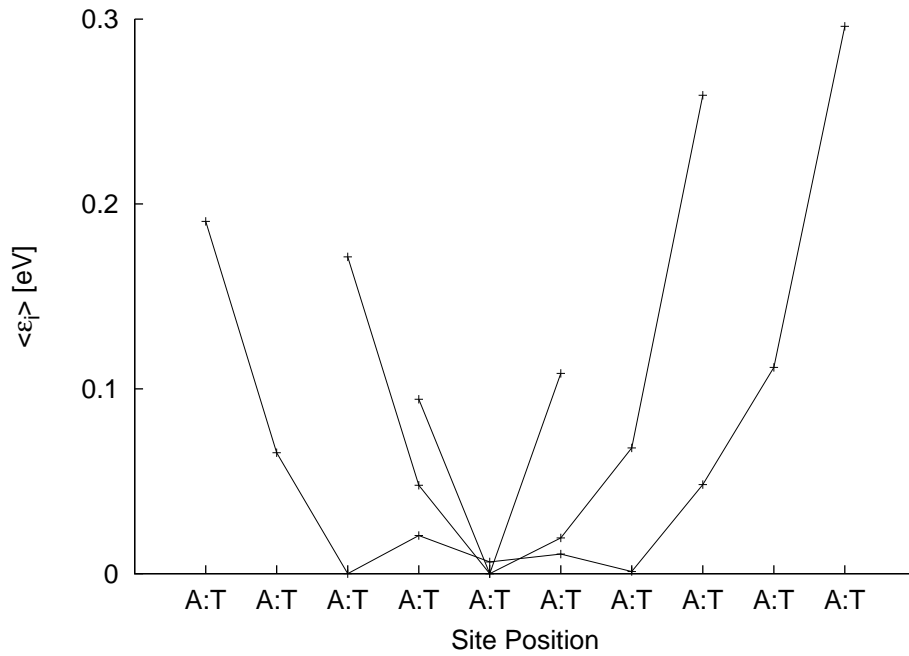


Figure 3: MNDO predicted relative onsite energies for a hole on n -stacked systems. The examples are from fluctuating $[A:T]_3$, $[A:T]_6$ and $[A:T]_{10}$.

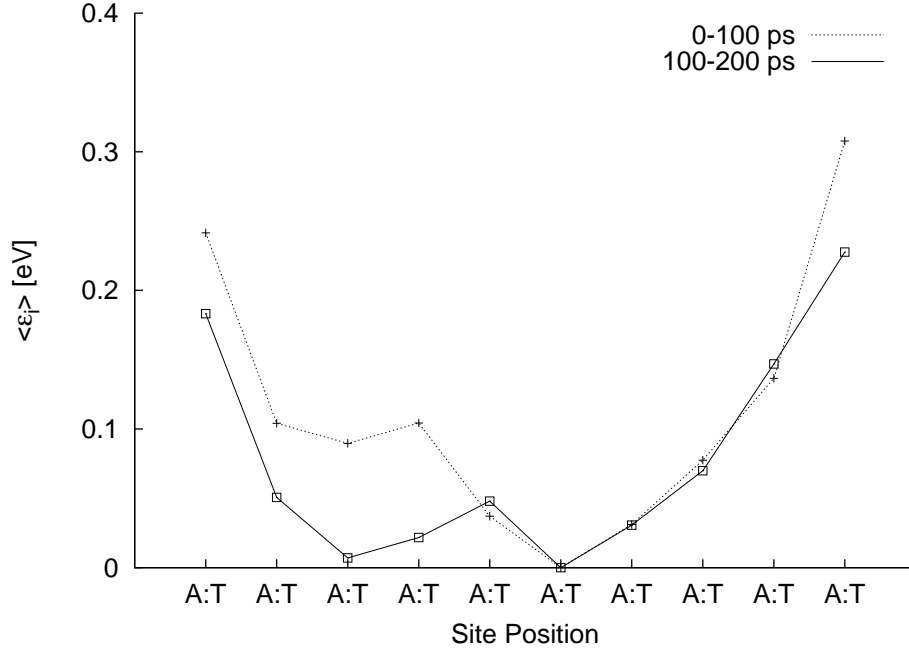


Figure 4: MNDO predicted relative onsite energies for a hole on $[A:T]_{10}$ system. The values are averaged on 500 samples from the first and the second 100 ps MD simulations, respectively.

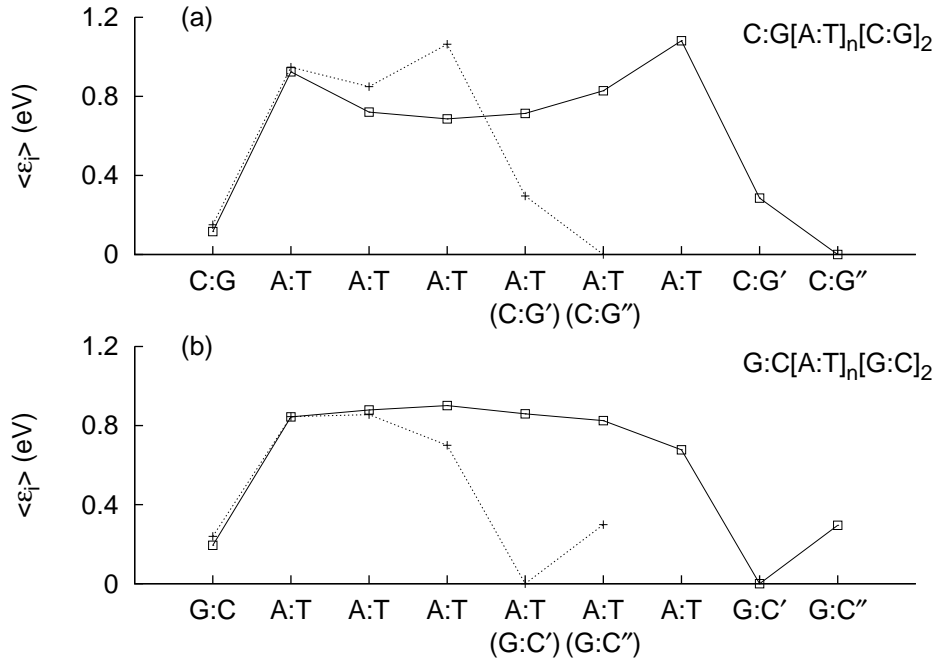


Figure 5: MNDO predicted relative onsite energies for a hole on $C:G[A:T]_n[C:G]_2$ and $G:C[A:T]_n[G:C]_2$ systems, $n=3$ and 6 . The last two G-contained base pairs are marked as G' and G'' respectively.

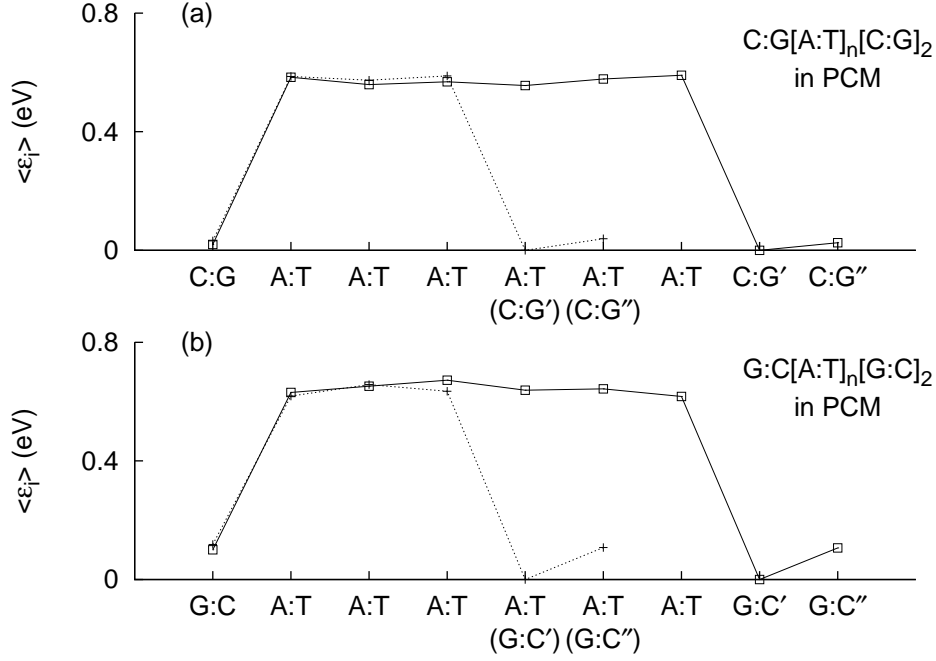


Figure 6: MNDO predicted relative onsite energies for a hole on $C:G[A:T]_n[C:G]_2$ and $G:C[A:T]_n[G:C]_2$ systems in PCM models, $n=3$ and 6. Compare with Figure 5.

Table 1: Onsite energies (in eV) in Figures 3-6. Corresponding sequences are denoted by one strand.

	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	A ₉	A ₁₀
Figure 3										
A ₃	0.09	0	0.11							
A ₆	0.17	0.05	0	0.02	0.07	0.26				
A ₁₀	0.19	0.07	0	0.02	0.01	0.01	0	0.05	0.11	0.3
Figure 4										
100 ps	0.24	0.1	0.09	0.1	0.04	0	0.03	0.08	0.14	0.31
200 ps	0.18	0.05	0.01	0.02	0.05	0	0.03	0.07	0.15	0.23
	G	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	G'	G''	
Figure 5										
CA ₃ C ₂	0.15	0.95	0.85	1.06				0.3	0	
CA ₆ C ₂	0.12	0.92	0.72	0.69	0.71	0.83	1.08	0.29	0	
GA ₃ G ₂	0.24	0.85	0.86	0.67				0	0.3	
GA ₆ G ₂	0.19	0.84	0.88	0.9	0.86	0.83	0.68	0	0.3	
Figure 6										
CA ₃ C ₂	0.03	0.59	0.57	0.59				0	0.04	
CA ₆ C ₂	0.02	0.58	0.56	0.57	0.56	0.58	0.59	0	0.03	
GA ₃ G ₂	0.12	0.62	0.66	0.64				0	0.11	
GA ₆ G ₂	0.1	0.63	0.65	0.67	0.64	0.64	0.62	0	0.11	