

# Mapping of Charge Delocalization Domain in Fluctuating DNA

Zhongwei Li and Keli 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

How to deal with the charge delocalized domain in theory has become a restriction on the understanding of the charge transport mechanism of DNA, especially for those charge transport events that exhibit nonmonotonic distance dependence. This work established a method to quantitatively characterize the hole delocalization domain of a given  $\pi$ -stacked conformation of nucleobases through a quantity called *orbital delocalization index* which uses the frontier occupied molecular orbitals of the neutral system as an approximation to predict the hole distribution in the cationic system. Wherein the concept of the delocalized domain can be expressed in an intuitive manner. For the large number of the snapshots generated by molecular dynamics simulations on the fluctuating DNA, the  $k$ -nearest neighbor algorithm in Machine Learning is used to classify the conformations into various delocalization structures. Through statistical analysis of these structures, we reproduced the nonmonotonic distance dependence exhibited by those DNA charge transport events and revealed that the formation of the delocalized domain is the origin of the nonmonotonicity. This work provides a different

---

\*To whom correspondence should be addressed

insight into the mechanism of DNA-mediated charge transport and is of generalizability to other  $\pi$ -stack systems.

## INTRODUCTION

Charge delocalization on a  $\pi$ -stack system is an appealing topic. It is due to the weak interaction of inter-molecules, which is different from the covalent interactions of intra-molecular.<sup>1</sup> So the general definitions such as bond order and electronic delocalization index in the atoms-in-molecules (AIM) theory<sup>2</sup> are no longer suitable to be directly used. A famous example is hole delocalization on DNA.<sup>3-7</sup> The absence of an effective method to describe the weak interaction between stacking bases associated with hole delocalization has led to the debates about the mechanisms of DNA charge transport (CT) lasting for three decades.<sup>8-11</sup> In this work, we aim at to build a method to quantitatively characterize the hole delocalization on a  $\pi$ -stack system and use it to interpret several confusing experiments relevant to the mechanism of DNA CT. If it proves to be effective, it could be used to predict the delocalization structure of a given DNA sequence and other soft  $\pi$ -stacking systems.

When one says DNA CT, it usually means DNA hole transport (HT), what attracts scientists for its relevance to irradiation-induced DNA lesions.<sup>12-14</sup> A hole is generally captured by a guanine (G) for its lowest ionization potential (IP) among the four nucleobases.<sup>15,16</sup> Interestingly, despite adenine (A) has a higher IP, the hole is possibly transmitted between two G sites intervened by the inter- or intra-strand A-sequence.<sup>3,17-24</sup> According to the experimentally measured distance dependence of HT rates,<sup>19,21,22</sup> proposed mechanisms in the early stage include single-step superexchange and multistep incoherent hopping.<sup>11,25-32</sup> Wherein superexchange is thought acts in short-distance and the donor and acceptor are energetically separated from the bridge. So it is a tunneling process causes exponential decay of HT rates with distance.<sup>19,20,22,33,34</sup> As the bridge grows, the tunneling probability decreases rapidly and HT occurs by relying on the multistep hopping,<sup>18,35-39</sup> which is thought a random inco-

herent process with weak distance dependence. The hopping barrier for a hole from G to A is thought overcome by the environmental polarization and structural fluctuations.<sup>40</sup> Hole is thought localized on a single base in these models.<sup>31,32</sup>

At the same time, there are many experiments that are difficult to interpret with superexchange and incoherent hopping,<sup>39,41–46</sup> thereby raising the needs for delocalization models. Renger et al. developed a variable-range hopping model to deal with the delocalization/localization switch due to dynamic disorder.<sup>47</sup> Conwell et al. have established a model in which a polaron is formed around the charge due to the polarization of environment and the structural distortion of base stack. A polaron is considered delocalized over 4~6 base pairs and DNA HT is thought actually the migration of the polaron.<sup>46,48–55</sup> Coincidentally, Barton and co-workers observed a periodic distance dependence of HT yields and the hole is likely to transfer via the CT-active conformations or between two delocalized domains.<sup>41,42,44,45,56,57</sup> Conwell thought that the concept of CT-active domain is actually the polaron,<sup>50</sup> but Barton thought they are different.<sup>56–58</sup> A CT-active domain is a well-stacked conformation consisted of 3~5 base pairs where  $\pi$  orbitals of bases are transiently delocalized. An essential distinguishment of it from the polaron is that a charge is not necessary.

Recently, Zhang and Beratan et al. proposed a flickering resonance (FR) model to replace the superexchange as the mechanism of short-distance DNA CT.<sup>8,59</sup> Wherein the donor-bridge-acceptor (D-B-A) system is transiently energetically degenerate due to structural fluctuations so that the assumption of tunneling is no longer needed. Using it Landi et al. simulated and reproduced Giese's experiments, in which a strong-to-weak change of distance dependence of HT rates was observed and thought an evidence of the crossover between superexchange and incoherent hopping.<sup>22</sup> Besides, Liu et al.<sup>60</sup> and Xiang et al.<sup>61</sup> used the concepts of energy resonance and the delocalized domain to explain the even-odd oscillations of the resistance in G-blocks.

More and more clues have indicated that the exploration of the model of hole delocalized domain is the key to gain deeper insight into the mechanism of DNA HT, especially

for the experiments in which the distance dependence of HT events shows a nonmonotonicity.<sup>17,41,42,44,45,60-62</sup> In this work, we at first establish a method within the fragment orbital (FO) framework<sup>63-66</sup> to quantitatively characterize the  $\pi$  delocalization/localization structure of a given stack. It uses the frontier occupied molecular orbitals (MOs) of a neutral system as the approximation of hole density, since that a converged wavefunction in the desired cation state of a  $\pi$  stack is not always available. Previous study shows that such an approach is effective.<sup>33,63-65,67-71</sup> By a new defined quantity, called *orbital delocalization index* herein, the potential hole delocalized domains can be intuitively expressed. Then, we use the *kNN* (*k*-nearest neighbor) algorithm<sup>72</sup> of Machine Learning to classify the obtained molecular dynamics (MD) snapshots of the fluctuating DNA systems into various delocalization structures. Wherein a typical D-B-A system  $GA_nG_2$  ( $n = 1 \sim 4$ , herein the complementary strand is omitted on writing) is selected. By statistics analysis of the delocalization structures, we reproduced the nonmonotonicity of distance dependence observed in the experiments and revealed its origin.

Incidentally, some concepts in DNA exciton study, such as the length of exciton delocalization,<sup>73</sup> exciton delocalization index,<sup>74</sup> etc., should be distinguished from the terms used in this article. Although our scheme is based on a neutral system, the aim is to simulate hole delocalization of a radical cation state. Yet exciton delocalization is a relaxed excited state of neutral system, in which charge separation is more cared.<sup>75-77</sup> In addition, the polarization effect of the environment, which is thought to suppress charge delocalization,<sup>78-81</sup> or provide driving force of CT,<sup>82,83</sup> is not considered in the present work. As mentioned, DNA CT is of high complexity. Our goal is not to advocate a certain mechanism by denying other ones, but is hoping to provide a new perspective on the mechanism of DNA CT.

# THEORY AND METHODS

## Extract the $\pi$ -Stack Hole Transfer Space

Suppose we now have a stack structure of base pairs, which is obtained from MD simulations or crystal diffraction experiments. If considers one base pair as a fragment, the MOs of the whole system can be expanded in the FO basis  $\Phi_{\text{loc}}$ :

$$\Psi_{\text{sys}} = C_{\text{loc}}^{\text{sys}} \Phi_{\text{loc}}. \quad (1)$$

$C_{\text{loc}}^{\text{sys}}$  is the FO coefficient matrix. The details about how to obtain  $C_{\text{loc}}^{\text{sys}}$  are attached in the Supporting Information (SI). In common HT process (deep-hole transfer<sup>84</sup> is beyond the scope of this work), only the frontier occupied MOs of fragments are involved. This makes it possible to extract an independent and self-consistent  $\pi$ -stack space for hole transfer from the whole system. If select the localized HOMOs of individuals as the basis  $\Phi_{\pi\text{loc}}$ , the  $\pi$ -stack HT space  $\Psi_{\pi\text{stk}}$  in the basis  $\Phi_{\pi\text{loc}}$  is

$$\Psi_{\pi\text{stk}} = C_{\pi\text{loc}}^{\pi\text{stk}} \Phi_{\pi\text{loc}}. \quad (2)$$

$\Psi_{\pi\text{stk}}$ ,  $C_{\pi\text{loc}}^{\pi\text{stk}}$  and  $\Phi_{\pi\text{loc}}$  are respectively the subspaces of  $\Psi_{\text{sys}}$ ,  $\Phi_{\text{loc}}$ , and  $C_{\text{loc}}^{\text{sys}}$ . For a  $m$ -stack, eq 2 is merely a  $m \times m$  matrix equation which contains the main interactions involved in the HT process and is much simpler in comparison with eq 1.

$\Psi_{\pi\text{stk}}$  is not necessarily a set of the highest  $m$  occupied orbitals of the whole system. Therefore the orthonormality of its eigenvectors  $\{\Psi_{\pi\text{stk}}^{\alpha}\}$  must be verified to ensure that it is not mixed in by other FOs (see the SI for detailed discussion). In addition, the basis  $\{\Phi_{\pi\text{loc}}^i\}$  are absolutely localized on individual fragments but usually are not strictly orthogonal. It is more convenient to use a set of orthogonal basis  $\Phi'_{\pi\text{loc}}$ . This can be achieved by any orthogonalization procedure.<sup>85</sup> Orthogonalized basis of  $\{\Phi''_{\pi\text{loc}}\}$  are still centralized on the individual fragments with the associated coefficient matrix being transformed to  $C_{\pi\text{loc}}^{\pi\text{stk}'}$ . In

the following text, we defaultly use the orthogonalized basis of  $\Phi'_{\pi\text{loc}}$  and  $C_{\pi\text{loc}}^{\pi\text{stk}'}$  but omit the apostrophe. And for simplicity,  $\Psi_{\pi\text{stk}}^\alpha$ ,  $\Psi_{\pi\text{stk}}$ ,  $C_{\pi\text{loc}}^{\pi\text{stk}}$  and  $\Phi_{\pi\text{loc}}$  are respectively called the HT eigenstate, HT space, HT coefficient and HT basis.

## Characterize the Delocalization Structure of a Stack

According to eq 2, a  $m$ -stack HT space has  $m$  eigenvectors  $\{\Psi_{\pi\text{stk}}^\alpha, \alpha = 1 \sim m\}$  in the HT basis  $\{\Phi_{\pi\text{loc}}^i, i = 1 \sim m\}$ . Now, we define a quantity

$$D_{ij} = \sum_{\alpha}^m W_{\alpha i} W_{\alpha j}, \quad (3)$$

as the *orbital delocalization index* (ODI) between fragments  $i$  and  $j$ . Wherein

$$W_{\alpha i} = [C_{\pi\text{loc}}^{\pi\text{stk}}]_{\alpha i}^2 \quad (4)$$

is the weight of the  $i$ -th HT base  $\Phi_{\pi\text{loc}}^i$  in the  $\alpha$ -th eigenvector of the HT space  $\Psi_{\pi\text{stk}}^\alpha$ . Note that ODI is different from the delocalization indices based on the electron density,<sup>86</sup> because it is the weight matrix here instead of the so-called density matrix. The diagonal term of ODI, which in fact characterizes the degree of orbital localization, is also not the charge density defined within the FO framework.<sup>64,65,67,69,71</sup> The matrix of ODI is simply called **D** matrix in this article.

Before practical application, we think it is necessary to use an ideal model, i.e. Hückel model to illustrate the physical significance of the ODI because it is a new concept. If use  $p$  orbital to represent  $\pi$  orbital, Hückel basis of  $\pi$ -stack is schematically shown in Figure 1a. If there are interactions between them (no matter what interactions they are), the corresponding bond order matrix is actually the same as  $\pi$  conjugation system of hydrocarbon chains. Details of Hückel solutions are attached in the SI. Here we show the results of the  $\pi$  delocalized (2, 3, 4)-stacks (Figure 1b) to discuss the meaning of ODI. Their Hückel **D**

matrices are listed as below:

$$\mathbf{D}(B_1=B_2, \text{Hückel}) = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}, \quad (5)$$

$$\mathbf{D}(B_1=B_2=B_3, \text{Hückel}) = \begin{pmatrix} 0.375 & 0.25 & 0.375 \\ 0.25 & 0.5 & 0.25 \\ 0.375 & 0.25 & 0.375 \end{pmatrix}, \quad (6)$$

$$\mathbf{D}(B_1=B_2=B_3=B_4, \text{Hückel}) = \begin{pmatrix} 0.3 & 0.2 & 0.2 & 0.3 \\ 0.2 & 0.3 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.3 & 0.2 \\ 0.3 & 0.2 & 0.2 & 0.3 \end{pmatrix}. \quad (7)$$

As eqs 5-7 show,  $\mathbf{D}$  is symmetric and what it actually reflects is the sum of all the involved interactions between any pair of fragments in the HT space. Wherein, the diagonal terms characterize the ability of the corresponding sites to localize a hole. The larger the value, the stronger the localization ability, and *vice versa*. In contrast, the off-diagonal terms characterize the ability to delocalize a hole between the corresponding sites.

Hückel ODIs indicate two interesting properties of  $\pi$ -stack. One is the odd-even effect. As eqs 5 and 7 show, the diagonal terms are equal for even  $m$ . But that is not true when  $m$  is odd in eq 6, in which the middle site shows stronger ability to localize a hole. Another one is the interaction at a distance. In a  $\pi$ -stack, it is possible for  $D_{ij}$  to have a larger value when the distance between sites  $i$  and  $j$  is further. This is closely related to single-step at-a-distance hole transfer.

Hückel model of  $\pi$ -stack is completely symmetrical, which results in the maximum degree of orbital interactions. In a real DNA system, the symmetry is broken by the helix structure, conformational fluctuations and the diversity of base sequences. These lead to more or less deviations from Hückel ODIs. As an extreme situation, if there is no interaction between

orbitals at all,  $\mathbf{D}$  becomes an unit matrix, which means all-localized orbitals in the HT space. In order to measure the degree of the orbital interactions in a stack (a part or the whole of the system), we use Hückel ODIs as a reference to define

$$R = 1 - \sqrt{\sum_{ij} [D_{ij} - D_{ij}(\text{Hückel})]^2} \quad (8)$$

as the  $\pi$ -stack delocalization index ( $\pi$ DI) of the stack. The second term in eq 8 can be regarded as the distance of the real stack from Hückel model. The larger the  $\pi$ DI of it, the closer it is to Hückel ideal delocalization situation.

But in a real system, not every pair of neighboring fragments has interactions. So we next introduce another concept, namely delocalization map, to characterize the delocalization/localization mixed structure of a given stack. In Figure 1c, still take (2, 3, 4)-stacks as examples, we show all the possible delocalization maps of a stack. A  $m$ -stack has  $2^{m-1}$  possible delocalization maps. And a delocalized domain is the area connected by one or several continuous “=” in a delocalization map. The corresponding Hückel  $\mathbf{D}$  matrices of the maps in Figure 1c can be constructed as below. The data for pure (2, 3, 4)-delocalization maps ( $B_1=B_2$ ,  $B_1=B_2=B_3$ ,  $B_1=B_2=B_3=B_4$ ) have been shown in eqs 5-7, while those for all-localization maps ( $B_1-B_2$ ,  $B_1-B_2-B_3$ ,  $B_1-B_2-B_3-B_4$ ) are the (2, 3, 4)-order unit matrices respectively. For a mixed delocalization map, take  $B_1=B_2-B_3$  for example, the corresponding Hückel  $\mathbf{D}$  matrix can be constructed as

$$\mathbf{D}(B_1=B_2-B_3, \text{Hückel}) = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (9)$$

Therefore, as have been emphasized, what  $\mathbf{D}$  matrix reflects is not the hole distribution but the delocalization structure of a given stack. For a given stack, its closest delocalization structure can be determined by comparing the  $\pi$ DI values of all possible delocalization maps,



i.e. by the so-called *kNN* algorithm<sup>72</sup> in Machine Learning. It is worth noting that the structure with multiple delocalized domains, e.g.  $B_1=B_2-B_3=B_4$ , implies an inter-domain hopping model which is important to the conformationally gated CT mechanism.<sup>44</sup>

One may note that as  $m$  increases, setting of Hückel  $\mathbf{D}$  matrix becomes enoumerous. Fortunately, when  $i$  and  $j$  are far apart, the ODI of them goes to near zero. So the  $\mathbf{D}$  matrix of a high order can always be divided into several non-zero block matrices of lower-orders. Thus, one needs to do is just determining the delocalization structures of these block matrices and then join them into an entire one. It is a technical tip.

So far, the scheme about how to identify the delocalization structure of a given stack has been presented. Since it is “wet” DNA that mostly used in DNA CT experiments, we mimic the fluctuating DNA in solution by MD simulations with AMBER<sup>87</sup> package (setting details are in the SI). And then for each system of interested, 500 stacking conformations are randomly picked out from 2 ns MD trajectories for delocalization analysis. *Ab initio* HF/3-21G method embedded in GAMESS<sup>88</sup> package is used for electronic structure calculations and the orthonormality check of  $\Psi_{\pi\text{stk}}$  is performed (Figure S3).

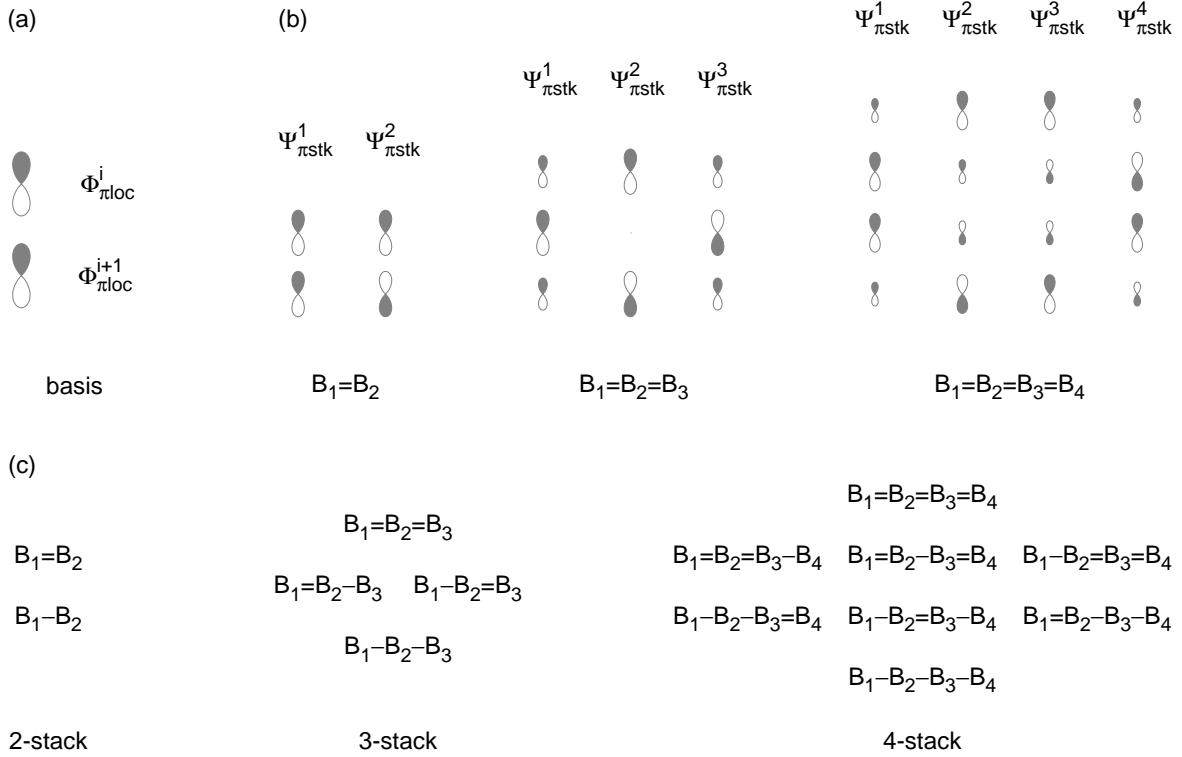


Figure 1: Schematic representations of (a) Hückel basis of  $\pi$ -stack. (b) Hückel solutions of delocalized (2, 3, 4)-stacks. (c) Possible delocalization maps of (2, 3, 4)-stacks. “=” means orbital delocalization and “-” means not.

## RESULTS AND DISCUSSION

### Distance Dependence of D-B-A Delocalization

As the length of the D-B-A system increases, the strong-to-weak distance dependence of HT rates has been a widely accepted feature of DNA HT.<sup>3,19,22,35,36</sup> It is thought due to the mechanism switch from superexchange to incoherent hopping.<sup>25,27,28,34,89</sup> Differently, the FR model interpret the changing distance dependence by the drop of the probability of forming the D-B-A energy resonance state as the length increases.<sup>8,59,90</sup> A similar relationship could also be found in terms of the delocalization map.

In the fluctuating  $GA_nG_2$  systems, of 500 randomly picked out snapshots of each system, the D-B-A delocalization maps detected by our procedures are shown in Figure 2a. The

ordinate indicates the  $\pi$ DI value of a map. It can be seen that as  $n$  increases, the proportion of D-B-A delocalization maps decreases sharply. Of the fluctuating  $\text{GAG}_2$  system, conformations of D-B-A delocalization structures ( $\text{G}=\text{A}=\text{G}-\text{G}$  and  $\text{G}=\text{A}=\text{G}=\text{G}$ ) accounts for 7.2%. And the proportion drops to 3.2% in  $\text{GA}_2\text{G}_2$ . In  $\text{GA}_3\text{G}_2$  and  $\text{GA}_4\text{G}_2$  systems, there was only one D-B-A delocalization map detected respectively. If fit the equation  $P_{\text{DBA}} = \exp(-\beta\Delta r)$  via the D-B-A delocalization formation probability  $P_{\text{DBA}}$  and the D-A distance  $\Delta r$ , the value of  $\beta$  is about  $0.3\sim 0.4 \text{ \AA}^{-1}$  for  $n = 1 \sim 3$ . For  $n > 3$ , the probability of forming the D-B-A delocalization nears zero. Such a change in the distance dependence from strong to weak is consistent with many experiments,<sup>3,18,19,22,35</sup> e.g. Giese's experiments of  $\text{GT}_n\text{G}_3$  systems,<sup>22</sup> where the HT efficiency is measured by the yield ratio of  $P_{\text{GGG}}/P_{\text{G}}$  (Figure 2b) and the reported  $\beta$  is  $0.6 \text{ \AA}^{-1}$ . This indicates that the presented method of identifying the hole delocalization structure of a given stack is effective. Importantly, we did not make any priori assumption about HT mechanism, e.g. the separated or degenerated energy relationship of D-B-A, tunneling or incoherent hopping, etc. We just identified the D-B-A delocalized structures from a large number of MD snapshots and the results show a clear from-strong-to-weak distance dependence. It could be considered a supportance for the FR model. In fact, the D-B-A delocalization map is naturally related to the D-B-A energy resonance state, what can be seen clearly in the next section.

One may note that some maps in Figure 2a have low  $\pi$ DI, e.g.  $R(\text{G}=\text{A}=\text{A}=\text{A}=\text{A}=\text{G}=\text{G}, \text{No.222}, \text{GA}_4\text{G}_2)$  is merely 0.116. But it is still the closest point to corresponding Hückel  $\mathbf{D}$  matrix in all possible delocalization maps of  $\text{GA}_4\text{G}_2$  at the No.222 snapshot. Generally, the shorter the delocalized domain, the easier it is to have a high  $\pi$ DI, and *vice versa*. So the longer the bridge, the more difficult to form the D-B-A delocalization structure, and long distance HT has to rely on multistep mechanism.

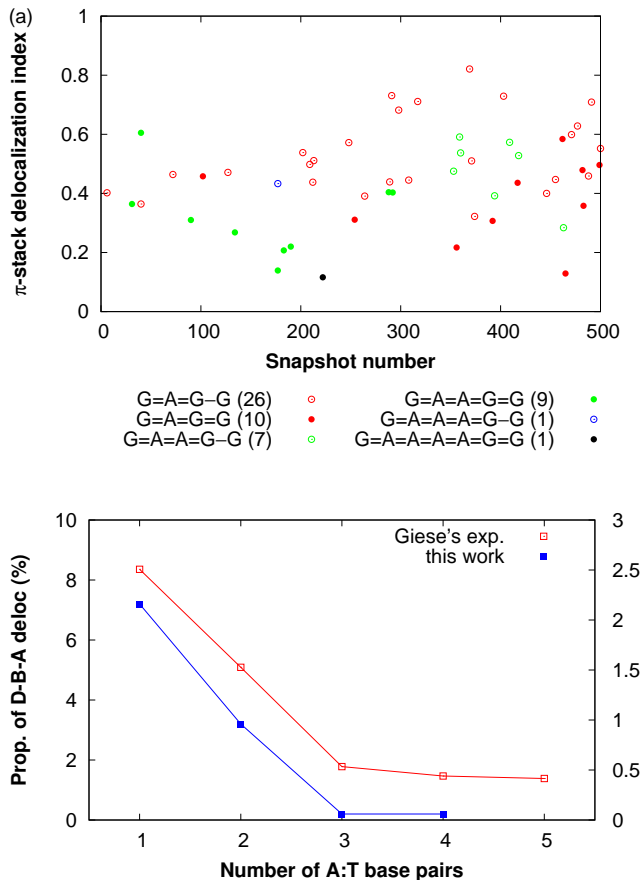


Figure 2: (a)  $\pi$ -stack delocalization indices of the D-B-A delocalization maps detected in the fluctuating  $GA_nG_2$  systems. Map counts are listed in parentheses. (b) Distance dependence respectively of the predicted proportion of D-B-A delocalized structure in  $GA_nG_2$  systems (blue) and the experimentally measured HT product ratio of  $GA_nG_3$  systems<sup>22</sup> (red).

## Electronic Structure of Delocalized Domain

In addition to the D-B-A delocalization map, the multiple domain map indicates another important structure, which is the basis of the conformationally gated CT mechanism.<sup>41,42,44,58</sup> Figure S6 shows a typical example of the multiple domain map, namely  $G-A=A-G=G$ , in terms of the associated  $\pi$ DIs. To understand the delocalization map representation better, the corresponding orbital representations of three typical delocalization maps, namely the D-B-A delocalization, multiple domain delocalization, and all-localization are shown in Figure 3. The associated FO coefficients,  $\mathbf{D}$  matrices, and  $\pi$ DIs are attached in Table S2 and eqs S9-S11. Onsite energies (eq S4) and electronic couplings (eq S5) are also shown at the

bottom of Figure 3.

All-localization structure is the easiest to understand (Figure 3c). Five eigenstates of the HT space of the stack are respectively dominated by five individual localized HOMOs. In No.29 snapshot,  $G_{(5)}$  has a higher onsite energy than A-sites have. It is a transient result due to the dynamic disorder.

No.209 snapshot is the multiple domain map  $G-A=A-G=G$  (Figure 3b). The HT eigenstates  $\Psi_{\pi\text{stk}}^1$  and  $\Psi_{\pi\text{stk}}^2$  are from the strong interactions between the localized HOMOs on the individual  $G_{(4)}$  and  $G_{(5)}$  sites. While  $\Psi_{\pi\text{stk}}^4$  and  $\Psi_{\pi\text{stk}}^5$  are from that of  $A_{(2)}$  and  $A_{(3)}$  sites. And  $\Psi_{\pi\text{stk}}^5$  is mainly from  $G_{(1)}$ . The meaning of the map is as follows. If a hole is transmitted to the  $G_{(1)}$  site at the current structure, it is considered localized. But if it is transmitted to  $A_{(2)}$  or  $A_{(3)}$  site, it has a probability to be delocalized on them, i.e. forms a  $-A_{(2)}=A_{(3)}$  delocalized domain.  $-G_{(4)}=G_{(5)}$  is the same. The delocalization degree is characterized by the  $\pi$ DI. Both the two two-site delocalized domains have high  $\pi$ DI (see eqs S13-S14), which are closed to the ideal Hückel situation. So a multiple domain map does not mean that a hole is simultaneously delocalized on these domains. It is a representation that allows different delocalization/localization states in the HT space to be visible in an intuitive formula.

To deal with the inter-domain HT model is not an easy task that requires more rigorous theoretical derivation.<sup>91,92</sup> We here simply consider a domain as a site to give a preliminary discussion about it. The electronic coupling and the onsite energy calculations were performed for the 91 all-localization maps  $G-A-A-G-G$  and 41  $G-A=A-G=G$  maps (Table 1) and then took the average over the dynamic disorder, i.e.  $\langle |V_{ij}| \rangle$  and  $\langle \epsilon_i \rangle$ . Wherein the absolute values are used for the couplings. As a result, the inter-domain coupling for HT between  $-A_{(2)}=A_{(3)}$  and  $-G_{(4)}=G_{(5)}$  is 0.028 eV that is smaller than the inter-base coupling 0.082 eV between  $A_{(3)}$  and  $G_{(4)}$  sites in the all-localization map. At the same time, the Gibbs free energy ( $\Delta G_{ij} = \epsilon_i - \epsilon_j$ ) for inter-domain HT is 0.45 eV that is greatly reduced in comparison with that for inter-base HT 0.72 eV. According to the Marcus electron transfer

theory,<sup>47,93</sup> an inter-domain HT rate should not be inferior to the inter-base HT.

Figure 3a is the mentioned D-B-A delocalization map. Five HOMOs of the individuals are strongly interacted that result in the five HT eigenstates delocalized over the entire stack space. So the map itself can be considered as a delocalized domain. One may note that in No.40 and No.209 snapshots there are large couplings and highly degenerate onsite energies inside the delocalized domains. If plot CT parameters as functions of  $\pi$ DI, as Figure 4 shows, it can be seen that the greater the  $\pi$ DI of a domain, the larger the couplings and the higher the degeneracy of onsite energies inside. It indicates that the energetically degenerate state in the FR model<sup>59</sup> and the CT-active delocalized domain in the conformationally gated CT mechanism<sup>44</sup> mean the same thing. The  $\pi$ DI is in function similar to the energy matching probability in the FR model.<sup>59</sup> In addition, the correlations between large coupling and low standard deviation of onsite energy have also been noticed by Meng et al. when they recently used the FR model to study DNA CT in hydrated ionic liquids.<sup>94</sup> Voityuk et al. used the correlations of onsite energies as well as the correlations of electronic couplings between neighboring base pairs to characterize hole size.<sup>60,61,81,82,95-98</sup> The rationality of these work is supported by the analysis of the delocalized domain. Finally, the localized domains ( $\pi$ DI=0 in Figure 4) with large couplings and small  $\sigma_\epsilon$  values correspond the model of incoherent hopping.

By this section discussion, we show that the **D** matrix,  $\pi$ DI, and the delocalization map are effective to characterize the delocalization structure of a stack. This avoids the ways that by subjective observation on the graphical representations of MO (or Kohn-Sham orbital) or charge density<sup>79,99-103</sup> or by artificially setting of threshold parameters<sup>33,89</sup> to characterize the delocalization.

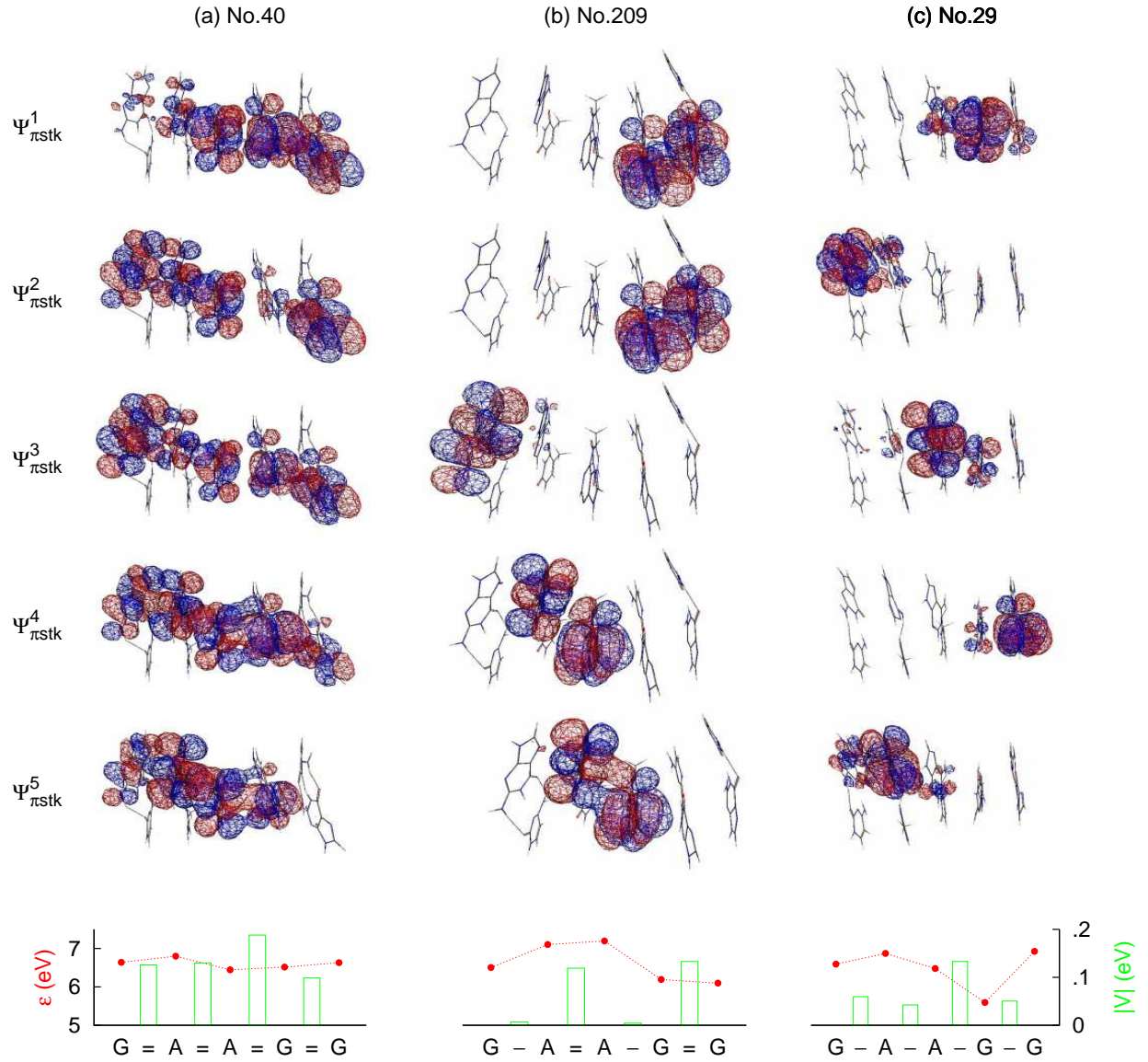


Figure 3: Snapshots of three typical delocalization maps in orbital representation. (a) D-B-A delocalization. (b) Multiple domains. (c) All-localization. Corresponding onsite energies  $\epsilon$  (red dot) and absolute values of the electronic coupling  $|V|$  (green rectangle) are also shown at the bottom.

Table 1: Distribution of the delocalization maps of fluctuating  $GA_2G_2$  system.

Delocalization map	Count	Percentage
G-A=A-G-G	200	40.0%
G-A-A-G-G	95	19.0%
G=A=A-G-G	68	13.6%
G-A=A-G=G	41	8.2%
G=A-A-G-G	25	5.0%
G-A-A-G=G	12	2.4%
G-A-A=G=G	10	2.0%
G=A=A=G=G	9	1.8%
G=A=A-G=G	9	1.8%
G-A=A=G=G	8	1.6%
G=A=A=G-G	7	1.4%
G-A-A=G-G	6	1.2%
G=A-A=G=G	5	1.0%
G-A=A=G-G	3	0.6%
G=A-A=G-G	1	0.2%
G=A-A-G=G	1	0.2%
Total	500	100%



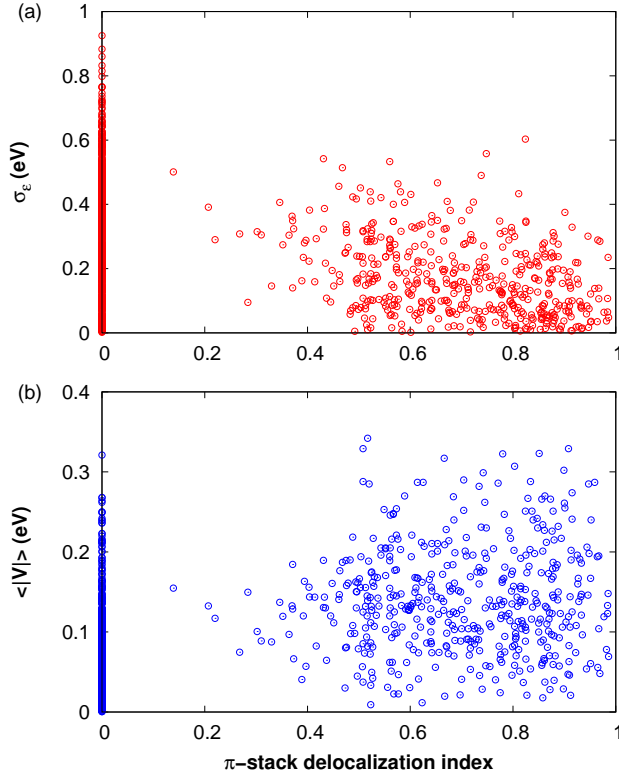


Figure 4: The dependences of (a) unbiased standard deviation of onsite energy  $\sigma_\epsilon$  (small  $\sigma_\epsilon$  means highly degenerate onsite energies, see eq S15), and (b) mean of absolute values of electronic couplings  $\langle |V| \rangle$  on the  $\pi$ -stack delocalization index of a domain. Refer to Table 1 for the selection of the delocalized domains. The areas connected by “—” are regarded as the localized domains whose  $\pi$ DIs are set to zero.

## Nonmonotonicity of Distance Dependence of CT Events

In addition to the most frequently discussed strong-to-weak distance dependence of HT rates,<sup>3,19,22,25,62,89</sup> there are other important phenomena being helpful to reveal the mechanism of DNA HT. Such as the odd-even oscillations of resistance in  $G_n$  systems,<sup>60,61</sup> and periodic distance dependence of the HT yield in  $A_nG$  systems.<sup>41,42,44</sup> A feature of these phenomena is that the nonmonotonicity of HT rates or HT yields versus distance. It is thought due to the formation of the delocalized domains with different lengths. We performed statistical analysis on the delocalized domains of  $GA_nG_2$  systems so as to seek the root cause of the nonmonotonicity.

Statistics on the delocalization maps of  $GA_2G_2$  system have been shown in Table 1 and

that of other systems are shown in Table S3-S5 in the SI. The probability of finding a particular delocalization map can be known from these tables. Take  $GA_2G_2$  for example (Table 1), the  $-A=A-$  containing maps account for nearly half (48.2%) that exceeds the proportion of all-localization maps (19%).  $-G=G$  containing maps account for 12.6%, indicating that it is also a significant delocalized domain. From this point of view, a hole should be easier to be delocalized on A-sequence than on G-sequence, what is in line with current consensus.<sup>50,104</sup> Recently, Jin et al. adopted the method of localized orbital scaling correction (LOSC) DFT to give a contrary result,<sup>95</sup> making it somewhat controversial in theory. But the electron spin resonance (ESR) experiments showed that the signal of hole delocalization on A-stack is more obvious than on G-stack.<sup>6,105</sup>

In addition to  $-A=A-$  and  $-G=G$ , there are other notable delocalized domains in  $GA_nG_2$  systems, e.g.  $G=A-$ ,  $-A=G-$  and  $-A=A=A-$  (if any). In Figure 5a, we plot the change in number of the significant domains as  $n$  increases. We use the number rather than the percentage as a function of  $n$  because there is a situation that more than one identical delocalized domains may coexist in one map, e.g., the  $-A=A-$  in  $G-A=A-A=A-G-G$ . So the denominator is uncertain. But for the domains which are unique in a map, percentage is still usable (denominator is 500). Figure 5a contains a lot of information. Some is universal to all DNA sequences while some is related only to the current case. By the combination with Figure 5b and Figure 5c, we summarized them as below:

(a) In A-sequence,  $-A=A-$  is the easiest formed domain. In fact, both ESR spectra<sup>6</sup> and pulse radiolysis<sup>5</sup> experiments have found that the most obvious signals of a hole on A-sequence come from A-dimer cation radical.

(b)  $-G=G$  proportion is around 10% and is independent of the A-bridge length.

(c)  $G=A-$  proportion is about 1/3 in  $GAG_2$  system. As the A-bridge grows, the proportion drops to below 10% due to the appearance of  $-A=A-$  and no longer decreases markedly.

(d)  $-A=G-$  is the most difficult formed domain in  $GA_nG_2$  systems and the proportion

is around 1%. The reason is that its competitors include  $-G=G$ ,  $-A=A-$  or  $G=A-$  which are more easily formed.

(e) All-localization structures account for more than 50% in  $GAG_2$  system. In  $GA_2G_2$ , the proportion drops to 20%. It is worth noting that a plateau appears from  $n = 2$  to  $n = 3$  and drops again by  $n = 4$ . Such a  $A_2$ -to- $A_3$  plateau was also observed in Harris's experiments,<sup>3</sup> which include the charge injection dynamics of donor- $A_nG$  systems (Figure 5a). The causes can be understood better from Figure 5b, which shows the delocalization indices of the individual sites (i.e.  $1 - D_{ii}$ ) of  $GA_nG_2$ . Compared with  $GA_2G_2$ , the added  $A_{(4)}$  site in  $GA_3G_2$  almost has no effect on the DIs of  $A_{(2)}$  and  $A_{(3)}$  sites, what indicates that the predominant delocalization structure of  $A_3$ -sequence should be the  $-A=A-A-$  (Table S4). That is why  $A_2G$  and  $A_3G$  have similar hole injecting dynamics in Harris's experiments. The existence of  $-A=A-A-$  can also be used to analyze the comparative dynamics of  $A_2G_3$  and  $A_3G_2$  reported by Olshansky et al.,<sup>62</sup> in which the added  $A_{(3)}$  site greatly increases the charge recombination barrier due to its localized role.

(f) Almost all the displayed domains in Figure 5a exhibit odd-even oscillations. Wherein  $-A=A-$ ,  $-G=G-$  and  $-A=G-$  are synchronous, while that  $-A=A=A-$  and  $G=A-$  are asynchronous. So it is reasonable to attribute the experimentally observed resistance oscillations of  $G_n$  blocks (red line in Figure 5c) to the formation of the delocalized domains, e.g.  $-G=G-$ , what has been suggested by Xiang et al.<sup>61</sup> Hückel model shows that the difference between odd and even stacks is the inherent property of  $\pi$ -stack interaction (see Figure S5 and eqs 5-7).

(g) Figure 5b shows that the site DIs of  $GA_nG_2$  increase obviously when  $n = 2$  and  $n = 4$ . It means that the highly delocalized domain is formed only at a certain length. This is consistent with the CT-active domain model,<sup>41,42,44</sup> which was proposed to interpret the CT events of periodic distance dependence (blue curve in Figure 5c). Due to the limitation of the current computational costs, we did not perform the analysis as above in a longer range. Replacing *ab initio* method with semiempirical method in the electronic structure

calculation step makes it possible to investigate longer systems, what is the research subject in future.

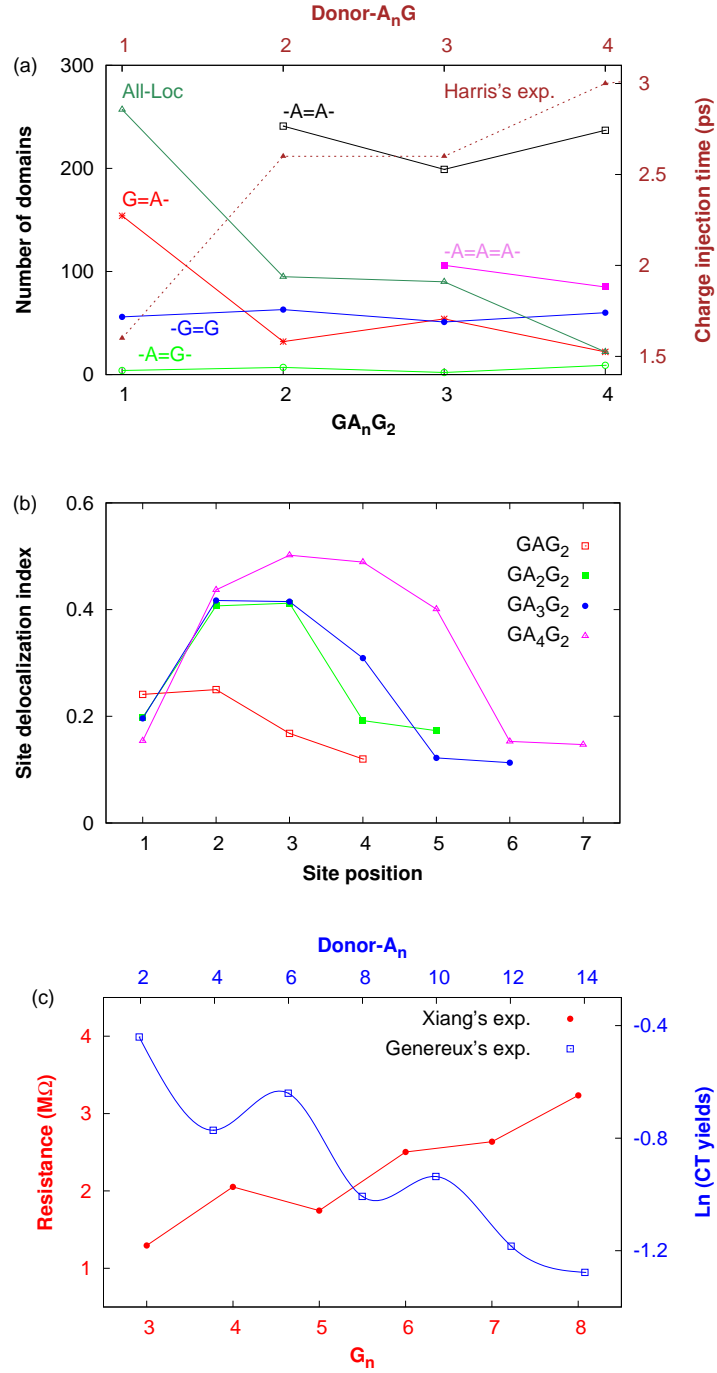


Figure 5: (a) Count of several important delocalized domains in  $GA_nG_2$  systems. Harris's experiments on the charge injection dynamics of donor- $A_nG$  systems<sup>3</sup> (brown dashed line) are used to compare with the all-localization domain. Both show a plateau from  $n=2$  to  $n=3$ . (b) The delocalization indices of individual sites of  $GA_nG_2$  systems. Note the slight change from  $n=2$  to  $n=3$  and the great change from  $n=3$  to  $n=4$ . (c) Xiang's experiments on the resistances of  $G_n$  blocks (red).<sup>61</sup> Note its odd-even oscillations which is reproduced in panel (a). Genereux's experiments on the HT yields of donor- $A_n$  systems (blue).<sup>42</sup> Note the situation that longer strands have higher HT yields indicating that highly delocalized domains are formed only at special lengths, what is reflected in panel (b).

## Special Role of the First Few Bridge Sites

As mentioned, in many DNA-mediated HT experiments, HT events through the short bridge (less than four base pairs) usually exhibits a changeable distance dependence.<sup>3,19,22,35</sup> As the bridge grows, the change tends to be predictable. So most of the models focus on the explanations of short distance HT,<sup>25,59,89,90</sup> while that of long distance is attributed to the incoherent hopping between neighboring base pairs.<sup>18,19,38,106</sup> But recently, it has been realized that not only it is insufficient to rely merely one mechanism during the short distance, but also it is inappropriate to simply attribute the long distance HT to the incoherent hopping of the localized hole.<sup>8,41,42,44,46,50,56,57</sup> Herein we give a discussion on the diversity of DNA HT mechanisms from the perspective of the delocalized domains.

In Figure 6, the delocalized domains of the fluctuating  $GA_nG_2$  systems are counted with different lengths. By combinations with Figure 2 and Figure 5a, it can be known that in short DNA strands, the D-B-A delocalization, the all-localization and other ordinary delocalization structures can coexist. This may result in a simultaneous action of the coherent, incoherent and partial coherent HT. As the strand becomes long, the D-B-A delocalization as well as the all-localization disappear and only ordinary delocalization structures left. The mechanism of long distance HT becomes monotonous. But it not necessarily is the incoherent hopping of the localized charge, it is more likely the coexistence of hole delocalization/localization and to switch as the structure fluctuates. The above differences cause the complexity of HT mechanisms in short systems while the regularity in long ones.

Another important topic is the hole size. For example, a polaron is thought varies from four to six base pairs,<sup>48-51,54,107</sup> and a CT-active domain is thought extends over three to five base pairs.<sup>41,44,58</sup> And there are other debates on it.<sup>43,78,81,95,97,98,108</sup> In our model the hole size is fluctuated and the fluctuating range varies with the lengths of the delocalized domains (Figure 6). It is the switch between delocalization/localization of a domain delivers a hole to the distant place.

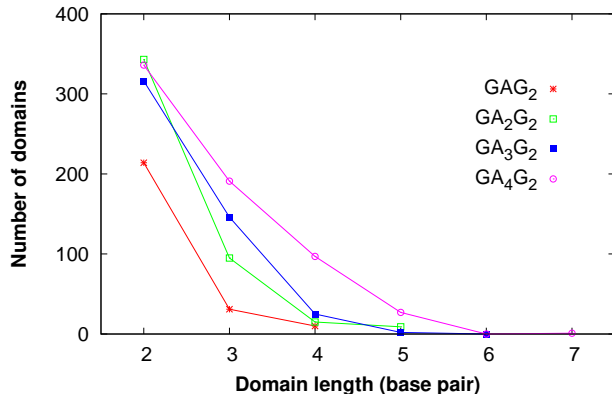


Figure 6: The number of delocalized domains of different lengths in  $GA_n G_2$ .

## CONCLUSIONS

In this article, we established a method to study the instantaneous potential hole delocalized structure in the fluctuating DNA system. The hole size and the position in different eigenstates of the HT space can be simultaneously expressed in a delocalization map. By virtue of it we interpreted several significant phenomena that are crucial to understand the mechanism of DNA HT, e.g., the even-odd oscillations of the resistance of  $G_n$  blocks, the periodic distance dependence of HT yields, and the special role of the first few bridge sites, etc. These phenomena are all related to the formation of the delocalized domains, what has been proposed by different experimental scientists. Our model illustrates the delocalized domain in an intuitive manner.

From the methodology point view, the  $kNN$  algorithm in Machine Learning is used to classify a large number of MD snapshots into various delocalization structures. In the electronic structure calculation step, more economic methods (e.g. semiempirical INDO/S<sup>40</sup>) can be employed if one is interested in larger systems. Conceptually, the quantities and terms introduced in this work, e.g. orbital delocalization index,  $\pi$ -stack delocalization index, and delocalization map etc. help us to quantitatively characterize the noncovalent  $\pi$ -stacking interactions associated with charge delocalization. These should also be suitable to other

$\pi$ -stacking systems.

How to deal with the intermediate CT mechanism between coherence and incoherence or partial coherence has been a nodus in DNA charge dynamics for a long time. This work should assist with its further development.

## Acknowledgement

This work is supported by the National Natural Science Foundation of China (Grant No. 21703234). Authors thank to Dr. Lu, Tian (Beijin KeYin Company) for valuable discussions about the interactions of inter-molecules. Authors also thank to Dr. Yang, Songqiu (Dalian Institute of Chemical Physics, CAS) for valuable discussions on experimental literatures.

## Supporting Information Available

Normality checking, Hückel solutions, MD details, and distributions of delocalization maps are attached in the supporting information. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## References

- (1) Johnson, E. R.; Keinan, S.; Mori-Snchez, P.; Contreras-Garca, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.
- (2) Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **1991**, *91*, 893–928.
- (3) Harris, M. A.; Mishra, A. K.; Young, R. M.; Brown, K. E.; Wasielewski, M. R.; Lewis, F. D. Direct Observation of the Hole Carriers in DNA Photoinduced Charge Transport. *J. Am. Chem. Soc.* **2016**, *138*, 5491–5494.



- (4) Capobianco, A.; Caruso, T.; D'Ursi, A. M.; Fusco, S.; Masi, A.; Scrima, M.; Chatgililoglu, C.; Peluso, A. Delocalized Hole Domains in Guanine-Rich DNA Oligonucleotides. *J. Phys. Chem. B* **2015**, *119*, 5462–5466.
- (5) Kobayashi, K. Evidence of Formation of Adenine Dimer Cation Radical in DNA: The Importance of Adenine Base Stacking. *J. Phys. Chem. B* **2010**, *114*, 5600–5604.
- (6) Adhikary, A.; Kumar, A.; Khanduri, D.; Sevilla, M. D. Effect of Base Stacking on the Acid-Base Properties of the Adenine Cation Radical [A<sup>+</sup>] in Solution: ESR and DFT Studies. *J. Am. Chem. Soc.* **2008**, *130*, 10282–10292.
- (7) Sugiyama, H.; Saito, I. Theoretical Studies of GG-Specific Photocleavage of DNA via Electron Transfer: Significant Lowering of Ionization Potential and 5-Localization of HOMO of Stacked GG Bases in B-Form DNA. *J. Am. Chem. Soc.* **1996**, *118*, 7063–7068.
- (8) Beratan, D. N. Why Are DNA and Protein Electron Transfer So Different? *Annu. Rev. Phys. Chem.* **2019**, *70*, 71–97.
- (9) Ferapontova, E. E. Electron Transfer in DNA at Electrified Interfaces. *Chem. Asian J.* **2019**, *14*, 3773–3781.
- (10) Genereux, J. C.; Barton, J. K. Mechanisms for DNA Charge Transport. *Chem. Rev.* **2010**, *110*, 1642–1662.
- (11) Jortner, J.; Bixon, M.; Langenbacher, T.; Michel-Beyerle, M. E. Charge transfer and transport in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 12759–12765.
- (12) Grodick, M. A.; Muren, N. B.; Barton, J. K. DNA Charge Transport within the Cell. *Biochemistry* **2015**, *54*, 962–973.
- (13) Hall, D. B.; Holmlin, R. E.; Barton, J. K. Oxidative DNA damage through long-range electron transfer. *Nature* **1996**, *382*, 731–735.

- (14) Murphy, C. J.; Arkin, M. R.; Jenkins, Y.; Ghatlia, N. D.; Bossmann, S. H.; Turro, N. J.; Barton, J. K. Long-Range Photoinduced Electron Transfer Through a DNA Helix. *Science* **1993**, *262*, 1025–1029.
- (15) Saito, I.; Nakamura, T.; Nakatani, K. Mapping of Highest Occupied Molecular Orbitals of Duplex DNA by Cobalt-Mediated Guanine Oxidation. *J. Am. Chem. Soc.* **2000**, *122*, 3001–3006.
- (16) Saito, I.; Takayama, M.; Sugiyama, H.; Nakatani, K.; Tsuchida, A.; Yamamoto, M. Photoinduced DNA Cleavage via Electron Transfer: Demonstration That Guanine Residues Located 5' to Guanine Are the Most Electron-Donating Sites. *J. Am. Chem. Soc.* **1995**, *117*, 6406–6407.
- (17) Kawai, K.; Majima, T. Hole Transfer Kinetics of DNA. *Acc. Chem. Res.* **2013**, *46*, 2616–2625.
- (18) Conron, S. M. M.; Thazhathveetil, A. K.; Wasielewski, M. R.; Burin, A. L.; Lewis, F. D. Direct Measurement of the Dynamics of Hole Hopping in Extended DNA G-Tracts. An Unbiased Random Walk. *J. Am. Chem. Soc.* **2010**, *132*, 14388–14390.
- (19) Lewis, F. D.; Zhu, H.; Daublain, P.; Fiebig, T.; Raytchev, M.; Wang, Q.; Shafirovich, V. Crossover from Superexchange to Hopping as the Mechanism for Photoinduced Charge Transfer in DNA Hairpin Conjugates. *J. Am. Chem. Soc.* **2006**, *128*, 791–800.
- (20) Lewis, F. D.; Liu, J.; Zuo, X.; Hayes, R. T.; Wasielewski, M. R. Dynamics and Energetics of Single-Step Hole Transport in DNA Hairpins. *J. Am. Chem. Soc.* **2003**, *125*, 4850–4861.
- (21) Lewis, F. D.; Zuo, X.; Liu, J.; Hayes, R. T.; Wasielewski, M. R. Dynamics of Inter- and Intrastrand Hole Transport in DNA Hairpins. *J. Am. Chem. Soc.* **2002**, *124*, 4568–4569.

- (22) Giese, B.; Amaudrut, J.; Khler, A.-K.; Spormann, M.; Wessely, S. Direct observation of hole transfer through DNA by hopping between adenine bases and by tunnelling. *Nature* **2001**, *412*, 318–320.
- (23) Giese, B.; Wessely, S.; Spormann, M.; Lindemann, U.; Meggers, E.; Michel-Beyerle, M. E. On the Mechanism of Long-Range Electron Transfer through DNA. *Angew. Chem., Int. Ed.* **1999**, *38*, 996–998.
- (24) Meggers, E.; Michel-Beyerle, M. E.; Giese, B. Sequence Dependent Long Range Hole Transport in DNA. *J. Am. Chem. Soc.* **1998**, *120*, 12950–12955.
- (25) Grozema, F. C.; Tonzani, S.; Berlin, Y. A.; Schatz, G. C.; Siebbeles, L. D. A.; Ratner, M. A. Effect of Structural Dynamics on Charge Transfer in DNA Hairpins. *J. Am. Chem. Soc.* **2008**, *130*, 5157–5166.
- (26) Shimazaki, T.; Asai, Y.; Yamashita, K. Theoretical Rate Constants of Super-Exchange Hole Transfer and Thermally Induced Hopping in DNA. *J. Phys. Chem. B* **2005**, *109*, 1295–1303.
- (27) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. Elementary steps for charge transport in DNA: thermal activation vs. tunneling. *Chem. Phys.* **2002**, *275*, 61–74.
- (28) Bixon, M.; Jortner, J. Long-range and very long-range charge transport in DNA. *Chem. Phys.* **2002**, *281*, 393–408.
- (29) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. Charge Hopping in DNA. *J. Am. Chem. Soc.* **2001**, *123*, 260–268.
- (30) Grozema, F. C.; Berlin, Y. A.; Siebbeles, L. D. A. Mechanism of Charge Migration through DNA: Molecular Wire Behavior, Single-Step Tunneling or Hopping? *J. Am. Chem. Soc.* **2000**, *122*, 10903–10909.

- (31) Giese, B. Long-Distance Charge Transport in DNA: The Hopping Mechanism. *Acc. Chem. Res.* **2000**, *33*, 631–636.
- (32) Bixon, M.; Giese, B.; Wessely, S.; Langenbacher, T.; Michel-Beyerle, M. E.; Jortner, J. Long-range charge hopping in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 11713–11716.
- (33) Hatcher, E.; Balaeff, A.; Keinan, S.; Venkatramani, R.; Beratan, D. N. PNA versus DNA: Effects of Structural Fluctuations on Electronic Structure and Hole-Transport Mechanisms. *J. Am. Chem. Soc.* **2008**, *130*, 11752–11761.
- (34) Senthilkumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Lewis, F. D.; Berlin, Y. A.; Ratner, M. A.; Siebbeles, L. D. A. Absolute Rates of Hole Transfer in DNA. *J. Am. Chem. Soc.* **2005**, *127*, 14894–14903.
- (35) Vura-Weis, J.; Wasielewski, M. R.; Thazhathveetil, A. K.; Lewis, F. D. Efficient Charge Transport in DNA Diblock Oligomers. *J. Am. Chem. Soc.* **2009**, *131*, 9722–9727.
- (36) Lewis, F. D.; Zhu, H.; Daublain, P.; Cohen, B.; Wasielewski, M. R. Hole Mobility in DNA A Tracts. *Angew. Chem., Int. Ed.* **2006**, *45*, 7982–7985.
- (37) Bixon, M.; Jortner, J. Incoherent charge hopping and conduction in DNA and long molecular chains. *Chem. Phys.* **2005**, *319*, 273–282.
- (38) Takada, T.; Kawai, K.; Cai, X.; Sugimoto, A.; Fujitsuka, M.; Majima, T. Charge Separation in DNA via Consecutive Adenine Hopping. *J. Am. Chem. Soc.* **2004**, *126*, 1125–1129.
- (39) Kawai, K.; Takada, T.; Tojo, S.; Majima, T. Kinetics of Weak Distance-Dependent Hole Transfer in DNA by Adenine-Hopping Mechanism. *J. Am. Chem. Soc.* **2003**, *125*, 6842–6843.

- (40) Voityuk, A. A.; Siriwong, K.; Rosch, N. Environmental Fluctuations Facilitate Electron-Hole Transfer from Guanine to Adenine in DNA Stacks. *Angew. Chem., Int. Ed.* **2004**, *43*, 624–627.
- (41) Genereux, J. C.; Wuerth, S. M.; Barton, J. K. Single-Step Charge Transport through DNA over Long Distances. *J. Am. Chem. Soc.* **2011**, *133*, 3863–3868.
- (42) Genereux, J. C.; Augustyn, K. E.; Davis, M. L.; Shao, F.; Barton, J. K. Back-Electron Transfer Suppresses the Periodic Length Dependence of DNA-Mediated Charge Transport across Adenine Tracts. *J. Am. Chem. Soc.* **2008**, *130*, 15150–15156.
- (43) Zeidan, T. A.; Carmieli, R.; Kelley, R. F.; Wilson, T. M.; Lewis, F. D.; Wasielewski, M. R. Charge-Transfer and Spin Dynamics in DNA Hairpin Conjugates with Perylenediimide as a Base-Pair Surrogate. *J. Am. Chem. Soc.* **2008**, *130*, 13945–13955.
- (44) O’Neil, M. A.; Barton, J. K. DNA Charge Transport: Conformationally Gated Hopping through Stacked Domains. *J. Am. Chem. Soc.* **2004**, *126*, 11471–11483.
- (45) Shao, F.; O’Neill, M. A.; Barton, J. K. Long-range oxidative damage to cytosines in duplex DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17914–17919.
- (46) Liu, C.-S.; Schuster, G. B. Base Sequence Effects in Radical Cation Migration in Duplex DNA: Support for the Polaron-Like Hopping Model. *J. Am. Chem. Soc.* **2003**, *125*, 6098–6102.
- (47) Renger, T.; Marcus, R. A. Variable-Range Hopping Electron Transfer through Disordered Bridge States: Application to DNA. *J. Phys. Chem. A* **2003**, *107*, 8404–8419.
- (48) Conwell, E. M.; McLaughlin, P. M.; Bloch, S. M. Charge-Transfer Excitons in DNA. *J. Phys. Chem. B* **2008**, *112*, 2268–2272.

- (49) Conwell, E. M.; Bloch, S. M.; McLaughlin, P. M.; Basko, D. M. Duplex Polarons in DNA. *J. Am. Chem. Soc.* **2007**, *129*, 9175–9181.
- (50) Conwell, E. M.; Bloch, S. M. Base Sequence Effects on Transport in DNA. *J. Phys. Chem. B* **2006**, *110*, 5801–5806.
- (51) Conwell, E. M. Charge transport in DNA in solution: The role of polarons. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 8795–8799.
- (52) Park, J.-H.; Choi, H.-Y.; Conwell, E. M. Hole Traps in DNA Calculated with Exponential ElectronLattice Coupling. *J. Phys. Chem. B* **2004**, *108*, 19483–19486.
- (53) Basko, D. M.; Conwell, E. M. Effect of Solvation on Hole Motion in DNA. *Phys. Rev. Lett.* **2002**, *88*, 098102.
- (54) Conwell, E. M.; Rakhmanova, S. V. Polarons in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 4556–4560.
- (55) Schuster, G. B. Long-Range Charge Transfer in DNA: Transient Structural Distortions Control the Distance Dependence. *Acc. Chem. Res.* **2000**, *33*, 253–260.
- (56) Arnold, A.; Grodick, M.; Barton, J. DNA Charge Transport: from Chemical Principles to the Cell. *Cell Chem. Biol.* **2016**, *23*, 183–197.
- (57) Muren, N. B.; Olmon, E. D.; Barton, J. K. Solution, surface, and single molecule platforms for the study of DNA-mediated charge transport. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13754–13771.
- (58) Shao, F.; Augustyn, K.; Barton, J. K. Sequence Dependence of Charge Transport through DNA Domains. *J. Am. Chem. Soc.* **2005**, *127*, 17445–17452.
- (59) Zhang, Y.; Liu, C.; Balaeff, A.; Skourtis, S. S.; Beratan, D. N. Biological charge transfer via flickering resonance. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 10049–10054.

- (60) Liu, C.; Xiang, L.; Zhang, Y.; Zhang, P.; Beratan, D. N.; Li, Y.; Tao, N. Engineering nanometre-scale coherence in soft matter. *Nat. Chem.* **2016**, *8*, 941–945.
- (61) Xiang, L.; Palma, J. L.; Bruot, C.; Mujica, V.; Ratner, M. A.; Tao, N. Intermediate tunnellinghopping regime in DNA charge transport. *Nat. Chem.* **2015**, *7*, 221–226.
- (62) Olshansky, J. H.; Young, R. M.; Wasielewski, M. R. Charge Separation and Recombination Pathways in Diblock DNA Hairpins. *J. Phys. Chem. B* **2019**, *123*, 1545–1553.
- (63) Castet, F.; Ducasse, L.; Fritsch, A. From organic superconductors to DNA: Fragment orbital-based model. *Int. J. Quant. Chem.* **2006**, *106*, 734–746.
- (64) Fritsch, A.; Ducasse, L.; Castet, F.; Ramasesha, S. Charge transport in DNA strands using fragment orbital theory. *Synthetic Met.* **2005**, *155*, 418–421.
- (65) Senthilkumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Siebbeles, L. D. A. Mapping the Sites for Selective Oxidation of Guanines in DNA. *J. Am. Chem. Soc.* **2003**, *125*, 13658–13659.
- (66) Dapprich, S.; Frenking, G. Investigation of Donor-Acceptor Interactions: A Charge Decomposition Analysis Using Fragment Molecular Orbitals. *J. Phys. Chem.* **1995**, *99*, 9352–9362.
- (67) Felix, M.; Voityuk, A. A. DFT performance for the hole transfer parameters in DNA stacks. *Int. J. Quant. Chem.* **2011**, *111*, 191–201.
- (68) Blancafort, L.; Duran, M.; Poater, J.; Salvador, P.; Simon, S.; Sola, M.; Voityuk, A. A. Excess charge delocalization in organic and biological molecules: some theoretical notions. *Theo. Chem. Acc.* **2009**, *123*, 29–40.
- (69) Blancafort, L.; Voityuk, A. A. CASSCF/CAS-PT2 Study of Hole Transfer in Stacked DNA Nucleobases. *J. Phys. Chem. A* **2006**, *110*, 6426–6432.

- (70) Lewis, J. P.; Cheatham, T. E.; Starikov, E. B.; Wang, H.; Sankey, O. F. Dynamically Amorphous Character of Electronic States in Poly(dA)Poly(dT) DNA. *J. Phys. Chem. B* **2003**, *107*, 2581–2587.
- (71) Beljonne, D.; Pourtois, G.; Ratner, M. A.; Brdas, J. L. Pathways for Photoinduced Charge Separation in DNA Hairpins. *J. Am. Chem. Soc.* **2003**, *125*, 14510–14517.
- (72) Wu, X.; Kumar, V.; Quinlan, J. R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G. J.; Ng, A.; Liu, B.; Yu, P. S.; Zhou, Z.-H.; Steinbach, M.; Hand, D. J.; Steinberg, D. Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **2008**, *14*, 1–37.
- (73) Nogueira, J. J.; Plasser, F.; Gonzalez, L. Electronic delocalization, charge transfer and hypochromism in the UV absorption spectrum of polyadenine unravelled by multiscale computations and quantitative wavefunction analysis. *Chem. Sci.* **2017**, *8*, 5682–5691.
- (74) Blancafort, L.; Voityuk, A. A. Exciton delocalization, charge transfer, and electronic coupling for singlet excitation energy transfer between stacked nucleobases in DNA: An MS-CASPT2 study. *J. Chem. Phys.* **2014**, *140*, 095102.
- (75) Improta, R.; Santoro, F.; Blancafort, L. Quantum Mechanical Studies on the Photo-physics and the Photochemistry of Nucleic Acids and Nucleobases. *Chem. Rev.* **2016**, *116*, 3540–3593.
- (76) Nielsen, L. M.; Hoffmann, S. V.; Nielsen, S. B. Electronic coupling between photo-excited stacked bases in DNA and RNA strands with emphasis on the bright states initially populated. *Photochem. Photobiol. Sci.* **2013**, *12*, 1273–1285.
- (77) Crespo-Hernandez, C. E.; Cohen, B.; Kohler, B. Base stacking controls excited-state dynamics in AT DNA. *Nature* **2005**, *436*, 1141–1144.
- (78) Kravec, S. M.; Kinz-Thompson, C. D.; Conwell, E. M. Localization of a Hole on an



- AdenineThymine Radical Cation in B-Form DNA in Water. *J. Phys. Chem. B* **2011**, *115*, 6166–6171.
- (79) Mantz, Y. A.; Gervasio, F. L.; Laino, T.; Parrinello, M. Solvent Effects on Charge Spatial Extent in DNA and Implications for Transfer. *Phys. Rev. Lett.* **2007**, *99*, 058104.
- (80) Voityuk, A. A. Charge transfer in DNA: Hole charge is confined to a single base pair due to solvation effects. *J. Chem. Phys.* **2005**, *122*, 204904.
- (81) Kurnikov, I. V.; Tong, G. S. M.; Madrid, M.; Beratan, D. N. Hole Size and Energetics in Double Helical DNA: Competition between Quantum Delocalization and Solvation Localization. *J. Phys. Chem. B* **2002**, *106*, 7–10.
- (82) Kubar, T.; Kleinekathfer, U.; Elstner, M. Solvent Fluctuations Drive the Hole Transfer in DNA: A Mixed QuantumClassical Study. *J. Phys. Chem. B* **2009**, *113*, 13107–13117.
- (83) Barnett, R. N.; Cleveland, C. L.; Joy, A.; Landman, U.; Schuster, G. B. Charge Migration in DNA: Ion-Gated Transport. *Science* **2001**, *294*, 567–571.
- (84) Renaud, N.; Harris, M. A.; Singh, A. P. N.; Berlin, Y. A.; Ratner, M. A.; Wasielewski, M. R.; Lewis, F. D.; Grozema, F. C. Deep-hole transfer leads to ultrafast charge migration in DNA hairpins. *Nat. Chem.* **2016**, *8*, 1015–1021.
- (85) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc. Mineola, New York, 1996.
- (86) Matito, E.; Poater, J.; Sol, M.; Duran, M.; Salvador, P. Comparison of the AIM Delocalization Index and the Mayer and Fuzzy Atom Bond Orders. *J. Phys. Chem. A* **2005**, *109*, 9904–9910.

- (87) 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. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1996**, *118*, 2309–2309.
- (88) Schmidt, M. W.; Baldridge, 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. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
- (89) Renaud, N.; Berlin, Y. A.; Lewis, F. D.; Ratner, M. A. Between Superexchange and Hopping: An Intermediate Charge-Transfer Mechanism in Poly(A)-Poly(T) DNA Hairpins. *J. Am. Chem. Soc.* **2013**, *135*, 3953–3963.
- (90) Landi, A.; Borrelli, R.; Capobianco, A.; Peluso, A. Transient and Enduring Electronic Resonances Drive Coherent Long Distance Charge Transport in Molecular Wires. *J. Phys. Chem. Lett.* **2019**, *10*, 1845–1851.
- (91) Levine, A. D.; Iv, M.; Peskin, U. Formulation of Long-Range Transport Rates through Molecular Bridges: From Unfurling to Hopping. *J. Phys. Chem. Lett.* **2018**, *9*, 4139–4145.
- (92) Renaud, N.; Berlin, Y. A.; Ratner, M. A. Impact of a single base pair substitution on the charge transfer rate along short DNA hairpins. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14867–14871.
- (93) Marcus, R. A. Electron transfer reactions in chemistry. Theory and experiment. *Rev. Mod. Phys.* **1993**, *65*, 599–610.
- (94) Meng, Z.; Kubar, T.; Mu, Y.; Shao, F. A Molecular Dynamics-Quantum Mechanics Theoretical Study of DNA-Mediated Charge Transport in Hydrated Ionic Liquids. *J. Chem. Theory Comput.* **2018**, *14*, 2733–2742.

- (95) Jin, Y.; Ru, X.; Su, N. Q.; Mei, Y.; Beratan, D. N.; Zhang, P.; Yang, W. Revisiting the Hole Size in Double Helical DNA with Localized Orbital Scaling Corrections. *J. Phys. Chem. B* **2020**, *124*, 3428–3435.
- (96) Liu, C.; Beratan, D. N.; Zhang, P. Coarse-Grained Theory of Biological Charge Transfer with Spatially and Temporally Correlated Noise. *J. Phys. Chem. B* **2016**, *120*, 3624–3633.
- (97) Kubar, T.; Elstner, M. What Governs the Charge Transfer in DNA? The Role of DNA Conformation and Environment. *J. Phys. Chem. B* **2008**, *112*, 8788–8798.
- (98) Voityuk, A. A. Are Radical Cation States Delocalized over GG and GGG Hole Traps in DNA? *J. Phys. Chem. B* **2005**, *109*, 10793–10796.
- (99) Kumar, A.; Adhikary, A.; Sevilla, M. D.; Close, D. M. One-electron oxidation of ds(5-GGG-3) and ds(5-G(8OG)G-3) and the nature of hole distribution: a density functional theory (DFT) study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5078–5089.
- (100) Capobianco, A.; Caruso, T.; Peluso, A. Hole delocalization over adenine tracts in single stranded DNA oligonucleotides. *Phys. Chem. Chem. Phys.* **2015**, *17*, 4750–4756.
- (101) Bravaya, K. B.; Epifanovsky, E.; Krylov, A. I. Four Bases Score a Run: Ab Initio Calculations Quantify a Cooperative Effect of H-Bonding and -Stacking on the Ionization Energy of Adenine in the AATT Tetramer. *J. Phys. Chem. Lett.* **2012**, *3*, 2726–2732.
- (102) Kumar, A.; Sevilla, M. D. Density Functional Theory Studies of the Extent of Hole Delocalization in One-Electron Oxidized Adenine and Guanine Base Stacks. *J. Phys. Chem. B* **2011**, *115*, 4990–5000.
- (103) O’Neill, P.; Parker, A. W.; Plumb, M. A.; Siebbeles, L. D. A. Guanine Modifications Following Ionization of DNA Occurs Predominantly via Intra- and Not Interstrand

- Charge Migration: An Experimental and Theoretical Study. *J. Phys. Chem. B* **2001**, *105*, 5283–5290.
- (104) Kubar, T.; Elstner, M. Coarse-Grained Time-Dependent Density Functional Simulation of Charge Transfer in Complex Systems: Application to Hole Transfer in DNA. *J. Phys. Chem. B* **2010**, *114*, 11221–11240.
- (105) Adhikary, A.; Khanduri, D.; Sevilla, M. D. Direct Observation of the Hole Protonation State and Hole Localization Site in DNA-Oligomers. *J. Am. Chem. Soc.* **2009**, *131*, 8614–8619.
- (106) Lewis, F. D.; Daublain, P.; Cohen, B.; Vura-Weis, J.; Shafirovich, V.; Wasielewski, M. R. Dynamics and Efficiency of DNA Hole Transport via Alternating AT versus Poly(A) Sequences. *J. Am. Chem. Soc.* **2007**, *129*, 15130–15131.
- (107) Kucherov, V. M.; Kinz-Thompson, C. D.; Conwell, E. M. Polarons in DNA Oligomers. *J. Phys. Chem. C* **2010**, *114*, 1663–1666.
- (108) Woiczikowski, P. B.; Kubar, T.; Gutierrez, R.; Caetano, R. A.; Cuniberti, G.; Elstner, M. Combined density functional theory and Landauer approach for hole transfer in DNA along classical molecular dynamics trajectories. *J. Chem. Phys.* **2009**, *130*, 215104.

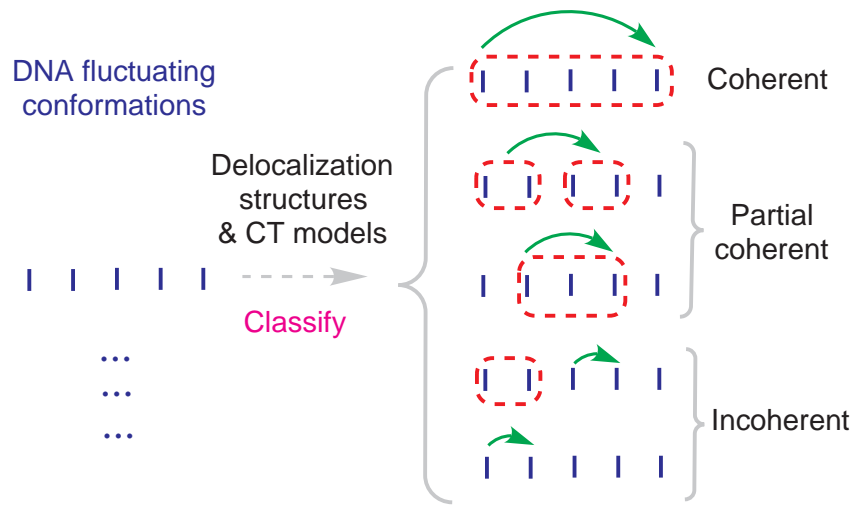


Figure 7: TOC