Published in final edited form as:

J Phys Chem B. 2009 July 16; 113(28): 9402–9415. doi:10.1021/jp904295q.

First principles effective electronic couplings for hole transfer in natural and size-expanded DNA

Agostino Migliore^{*,†}, Stefano Corni[‡], Daniele Varsano[‡], Michael L. Klein[†], and Rosa Di Felice^{*,‡}

[†] Center for Molecular Modeling and Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104-6323, USA

[‡] National Center on nanoStructures and bioSystems at Surfaces (S3) of INFN-CNR, c/o Dipartimento di Fisica, Università di Modena e Reggio Emilia, Via Campi 213/A, 41125 Modena, Italy

Abstract

Hole transfer processes between base pairs in natural DNA and size-expanded DNA (xDNA) are studied and compared, by means of an accurate first principles evaluation of the effective electronic couplings (also known as transfer integrals), in order to assess the effect of the base augmentation on the efficiency of charge transport through double-stranded DNA. According to our results, the size expansion increases the average electronic coupling, and thus the CT rate, with potential implications in molecular biology and in the implementation of molecular nanoelectronics. Our analysis shows that the effect of the nucleobase expansion on the charge-transfer (CT) rate is sensitive to the sequence of base pairs. Furthermore, we find that conformational variability is an important factor for the modulation of the CT rate. From a theoretical point of view, this work offers a contribution to the CT chemistry in π -stacked arrays. Indeed, we compare our methodology against other standard computational frameworks that have been adopted to tackle the problem of CT in DNA, and unravel basic principles that should be accounted for in selecting an appropriate theoretical level.

I. Introduction

DNA-mediated charge transport has been the subject of intensive studies in recent years. The driving factors towards research in this field are its biological relevance, on the one hand, e.g., in controlling the mechanisms of the oxidative damage to DNA and repair strategies,^{1,2,3} and its possible implications for nanoelectronics, on the other hand.^{4,5,6} Indeed, these two topics are strictly connected to each other,^{2,7,8} as a consequence of two main features of charge transport along DNA molecules: (i) Long-range charge transfer (CT), in particular hole transfer, with a shallow distance dependence can occur via a suitable combination of the super-exchange and hopping mechanisms;^{9,10} (ii) Efficiency and rates of CT through DNA are very sensitive to the sequence, the presence of mismatches (even concerning a single base)¹¹ and the electronic structure changes in situations with DNA-protein binding.¹² Such inherent

*CORRESPONDING AUTHORS: Agostino Migliore. Center for Molecular Modeling, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, PA 19104-6323, USA. Phone: +1-215-898-7058. Fax: +1-215-573-6233. E-mail: migliore@seas.upenn.edu. Rosa Di Felice. National Center for nanoStructures and bioSystems at Surfaces (S3) of INFN-CNR, Via Campi 213/A, 41100 Modena, Italy. Phone: +39-059-2055320. Fax: +39-059-2055651. E-mail: rosa.difelice@unimore.it.

Supporting information available. Cartesian coordinates of base-pair dimers; derivation of Eq. 3 and discussion about CDFT diabatic states; transfer integral and problematic convergence in the A-A system; HOMO and LUMO isosurface plots of various A-A systems. This material is available free of charge via the Internet at <http://pubs.acs.org>.

features can in principle be exploited for manufacturing molecular devices with diverse functionalities.^{4,5,6}

As a matter of fact, despite the recent achievements in the construction of DNA nanostructures,¹³ DNA-based nanoelectronics continues to be a challenging issue, because most of the chemical and physical properties related to charge mobility are still unclear. A variety of experimental reports of an insulating behavior for long (≥ 40 nm) DNA molecules deposited on substrates¹⁴ were recently counterbalanced by the evidence of significant currents for short (≤ 10 nm) DNA molecules suspended between electrodes with controlled molecule-electrode contacts and without non-specific molecule-substrate interactions.¹⁵ This controversial background calls for more theoretical investigations, able to understand the dependence of the electric conduction on the base sequence and on conformational changes.

The relevance of our work to the field of molecular electronics is provided by the connection between the electrical conductance of single (DNA) molecules and the electron transfer (ET) rate,¹⁶ which, in turn, depends on the effective electronic coupling between the donor and the acceptor.^{17,18} In this paper, similarly to many previous works, the framework of Marcus' theory^{17,19} is essentially preserved. At any rate, the theoretical analysis of the effective electronic couplings between adjacent units is a crucial step for understanding the charge transfer through DNA sequences. This theoretical step still deserves careful investigation in order to resolve unclear issues, to explain discrepancies present in the literature and to address sequence manipulation. Moreover, it is useful for practical purposes, e.g., in connection with the possibility of detecting single-nucleobase mismatches and with the use of synthetic DNA chains. Since guanine (G) is the most easily oxidized base, followed by adenine (A),²⁰ in natural and synthetic DNA double helices G and A are the stepping stones for hole transfer. Notwithstanding the different orders of magnitude of conductance in short synthetic nucleic acids and in long biological DNA duplexes,^{14,15} the nature of CT essentially rests on a common mechanism, which involves the aromatic stacks between adjacent base pairs, here studied in terms of their electronic couplings.

Within the above controversy about the efficiency and mechanisms of CT in DNA one finds a new size-expanded derivative called xDNA.^{21,22,23,24} Double strands of xDNA are obtained by inserting a benzene ring into one nucleobase per each Watson-Crick pair.^{23,24} This modification yields higher thermodynamic stability for sequences in which all the base pairs are size-expanded.²³ Kool and coworkers suggested that the higher stability may be imputed to a higher degree of stacking due to the aromatic insertions. Higher stacking could also be revealed in larger transfer integrals and better suitability to mediate charge motion.^{25,26} Therefore, it is desirable to study the CT behavior of xDNA, starting from the transfer integral between adjacent base pairs.

The above mentioned theoretical and experimental arguments represent the motivations for this work, where we perform a detailed analysis of the electronic coupling matrix elements for hole transfer between dimers of nucleobase pairs in natural DNA and xDNA. The objective is to determine whether xDNA may sustain faster ET than natural DNA. The comparison between the respective transfer integrals is established. The effective electronic coupling is calculated both in idealized stacks obtained with nucleic acid builders²⁷ and in base pair sequences from real DNA and xDNA structures. The connection between transfer integrals and conformational changes is explored, giving insight into the limits of the Condon approximation. The comparison between the electronic couplings in the natural and size-expanded DNAs is read in relation to the effects of the local flexibility and the conformational dynamics. The emerging picture describes features of CT in xDNA sequences potentially useful to future nanoelectronics. Specifically, we find that the transfer integrals for selected xDNA stacked geometries are larger than those characteristic of DNA stacks. This superior CT capability is

rather robust against structural changes, though a more systematic account for geometrical changes could resolve quantitative aspects in a realistic dynamical condition.

The article is organized as follows: we first describe the computational methodology, we then report the results on DNA and xDNA base-pair dimers and show that xDNA is potentially capable of hosting a faster charge transfer than natural DNA, and finally discuss theoretical implications of our methodology and its accuracy relative to other theoretical frameworks.

II. Theoretical and Computational methods

Quantum mechanical first principles calculations in the framework of density functional theory (DFT) are carried out on the DNA dimer sequences shown in Figure 1: GC-GC, xGC-xGC, AT-AT, xAT-xAT. The sequences are illustrated in Figure 1 for ideal structures, but we also consider real structures with the same sequences: structural details are specified later and the atomic coordinates are given in the Supporting Information. In all of them, the natural or expanded purines are in one strand and the natural pyrimidines are in the opposite strand. These oligomers were chosen in accordance with their theoretical and practical relevance, as sketched above. In addition to the structures of Figure 1, we also consider single-stranded G-G and A-A stacks that we use to carry out extensive analysis of the methodological performance. The sugar-phosphate DNA backbone is excluded from our first principles calculations without loss of significance. In fact, hole transfer and electron transfer processes proceed through the stacked nucleobases,^{28,29} while the effect of the backbone on the transfer integrals is negligible.^{30,31} The role of the backbone in the charge transport essentially arises from the duplex motion, which causes conformation changes in the π -stacking with a consequent significant effect on the CT between the nucleobases. Throughout the text we adopt the symbol Q to denote the configurational coordinate. A subscript i (r) to the right indicates ideal (real) configurations. The presence (absence) of a subscript x to the right indicates xDNA (natural DNA) configurations. A superscript G (A) to the left indicates stacks of guanines (adenines) without complementary bases. A superscript GC (AT) to the left indicates stacks of GC-GC (AT-AT) hydrogen-bonded pairs.

The charge transfer is expected to fall within the nonadiabatic regime,³² so that the CT rate constant is proportional to the square of the transfer integral between the donor and acceptor pairs.^{17,33} The first principles calculation of the transfer integrals is performed here by means of the formula³⁴

$$V_{\text{IF}} = \left| \frac{\Delta E_{\text{IF}} ab}{a^2 - b^2} \right|. \quad (1)$$

$\Delta E_{\text{IF}} \equiv E_{\text{I}} - E_{\text{F}}$ is the energy difference, at the given nuclear configuration, between the ET initial (I) and final (F) diabatic states. a and b are the coefficients of I and F, respectively, in the expression of the adiabatic ground state, which in the two-state model is given by $|\psi\rangle = a|\psi_{\text{I}}\rangle + b|\psi_{\text{F}}\rangle$. Note, however, that Eq. 1 can also be used when the two-state approximation does not exactly hold.³⁴ The reactant state vector is defined as $|\psi_{\text{I}}\rangle = |D^+\rangle |A\rangle$ and the product state vector as $|\psi_{\text{F}}\rangle = |D\rangle |A^+\rangle$. $|D^+\rangle$ ($|D\rangle$) represents the oxidized (reduced) ground state of the isolated donor group, which is the base pair where the hole is localized before the CT process. $|A\rangle$ ($|A^+\rangle$) is the reduced (oxidized) ground state of the isolated acceptor group, where the hole remains localized after the CT process. Eq. 1 does not require the knowledge of the exact transition state coordinate. Moreover, it provides directly the value of the effective electronic coupling matrix element, which is the quantity entering the expression for the ET rate constant. The quantity V_{IF} that appears in Eq. 1 is related to the electronic coupling matrix element, i.e.,

$H_{\text{IF}} = \langle \psi_{\text{I}} | H | \psi_{\text{F}} \rangle$ (where H is the electronic Hamiltonian of the system), by the relation $V_{\text{IF}} \cong H_{\text{IF}} - S_{\text{IF}}(E_{\text{I}} + E_{\text{F}})/2$, under the assumption $|S_{\text{IF}}|^2 \equiv |\langle \psi_{\text{I}} | \psi_{\text{F}} \rangle|^2 \gg 1$ that is widely satisfied by all the systems studied in the present work. The quantity ΔE_{IF} in Eq. 1 is given by

$$\Delta E_{\text{IF}} = (E_{\text{D}^+} + E_{\text{A}}) - (E_{\text{D}} + E_{\text{A}^+}) + W_{\text{D}^+ - \text{A}} - W_{\text{D} - \text{A}^+}. \quad (2)$$

Here E_{D^+} and E_{A^+} (E_{D} and E_{A}) are the ground-state energies of the isolated hole donor and acceptor groups, respectively, in their oxidized (reduced) state of charge. $W_{\text{D}^+ - \text{A}}$ ($W_{\text{D} - \text{A}^+}$) is the energy of essentially electrostatic interaction between donor and acceptor in the initial (final) diabatic state. The computational details on the evaluation of the interaction terms are reported in the Supporting Information, where it is also shown that the vertical excitation energy (i.e., the energy difference between the adiabatic ground state and the first excited state, at the given nuclear configuration) can be calculated by means of the formula

$$\Delta E_{\text{v}} = \left| \frac{a^2 + b^2}{a^2 - b^2} \Delta E_{\text{IF}} \right|. \quad (3)$$

Note that the diabatic states can also be derived by the constrained DFT (CDFT) method.³⁵ In this event, a partition of ΔE_{IF} , as in Eq. 2, is not required. CDFT is, however, not feasible in the presence of basis sets with many diffuse functions. In our work we do not use CDFT. Rather, we use tensor product states, $|\psi_{\text{I}}\rangle$ and $|\psi_{\text{F}}\rangle$, that satisfy the two-state approximation and give an appropriate description of CT. A brief discussion of the performance of the CDFT diabatic states relative to our tensor product states is given in the Supporting Information and some CDFT values are also reported in Table 6. Here we wish to point out that the tensor product and CDFT diabatic electronic states yield very similar values of ΔE_{IF} (e.g., within 2% for the cases reported in Table 6), therefore corroborating our implementation (see the Supporting Information) of Eq. 2.

All the quantities that appear in Eqs. 1–3 are obtained from full-electron spin-unrestricted DFT calculations carried out with the NWChem package.³⁶ The Becke half-and-half (here denoted BHH) hybrid exchange-correlation functional³⁷ is adopted: it is made of 1/2 Hartree-Fock (HF) exchange (which is essentially equal to the exact Kohn-Sham exchange), 1/2 Slater exchange and 1/2 PW91-LDA correlation. It derives from the rigorous first principles formula for the exchange-correlation energy of Kohn-Sham DFT, known as the *adiabatic connection* formula,³⁸ after linear interpolation of the pertinent inter-electronic coupling-strength parameter. Therefore, it rests on a clear theoretical basis, as compared with other hybrid functionals using different amounts of exact exchange. The accuracy of the chosen computational framework is analyzed in the Discussion and compared against other theoretical/computational setups.

The adopted hybrid DFT scheme offers the best compromise between accuracy and feasibility within a theoretical framework that includes correlation effects.^{39,40} It allows us to explore the whole size range of the considered base stacks, which is not manageable with post-HF methods. Previous works suggest that half-and-half exchange-correlation functionals are the best available functionals to describe long-range CT states.⁴¹ Indeed, some recent works claim that BHH is the only hybrid DFT functional able to give reliable excitation energies for the π - π CT complexes.⁴² Moreover, the Becke half-and-half choice has been successfully applied to the study of various properties of many π -stacked aromatic complexes, yielding results in good agreement with high-level post-HF calculations and experimental data.^{43,44,45,46} Finally, its good performance on the individual Watson-Crick nucleobase pairs is also supported by very recent works^{47,48} on polypeptides, which indicate the half-and half DFT scheme as a viable

alternative to highly correlated *ab initio* methods in calculating the interaction energies of weakly polar interactions,⁴⁷ and recommend its use in the study of other systems with intramolecular hydrogen bonds.⁴⁸ In our work, the BHH functional is validated against other commonly adopted hybrid functionals on the test case of the G-G stack, and against previous post-HF results for G-G and A-A stacks. Our results, along with those of other authors,^{43,44, 45,46,49} indicate that the hybrid DFT using BHH can be a valid alternative to post-HF methods in the treatment of relatively large π -stacked systems.

III. Results

Effective electronic couplings in DNA and xDNA base-pair dimers

In this section our main target is the detection of a possible enhancement of the transfer integral in xDNA base-pair dimers relative to those of natural DNA. To this purpose, we investigate GC-GC, xGC-xGC, AT-AT and xAT-xAT dimers (Figure 1). We consider “ideal” stacks constructed with standard rise and twist parameters for B-DNA and “real” stacks extracted from PDB files. In addition to this physico-chemical goal, our results also allow us to comment on critical issues in the literature from a theoretical perspective; but this we postpone to the Discussion.

Table 1 reports the values of transfer integrals for two different GC-GC dimers at different theoretical levels. The transfer integral in the ideal GC-GC stack from B-DNA, $V_{\text{IF}}(^{\text{GC}}Q_i)$, is slightly larger than in the absence of the cytosines (cf. Table 1 and Table 6). This confirms – at the full-electron level of our calculations – that CT essentially occurs through the guanines also in the presence of the complementary cytosines. The increase in V_{IF} upon guanine-cytosine pairing is in general agreement with previous results,^{29,32,50} which is neither trivial nor valid for any sequence.^{29,50} In this work, this trend gains additional support from the fact that it is consistently reproduced by using different basis sets. Our best value of 0.075 eV for $V_{\text{IF}}(^{\text{GC}}Q_i)$ is rather close to the SCC-DFTB50 and root mean square (rms) INDO/S¹⁸ values and midway between the two HF estimates from Refs. 32 and ⁵¹. The discrepancies with other data are discussed later. Our best value of 0.058 eV for $V_{\text{IF}}(^{\text{GC}}Q_r)$ is in excellent agreement with the result of the only other available calculation on a real GC-GC geometry.⁵¹

Table 2 reports various quantities from transfer integral calculations on an ideal xGC-xGC dimer at two different reaction coordinates $^{\text{GC}}Q_{x1i}$ and $^{\text{GC}}Q_{x2i}$, both corresponding to the base step parameters for the ideal xDNA structure: $^{\text{GC}}Q_{x1i}$ is close to the transition state coordinate $^{\text{GC}}Q_{xt}$, while $^{\text{GC}}Q_{x2i}$ deviates from $^{\text{GC}}Q_{xt}$. No structural experimental data are available for xDNA containing a fragment with this sequence; hence we cannot investigate a real xGC-xGC dimer. No reference data exist, which endows theoretical calculations with a predictive role.

Even considering the maximum range of the coupling for the B-DNA GC-GC stack reported in Ref. 51 (0.099 eV), our V_{IF} evaluations for the ideal xGC-xGC stack indicate a considerable increase of the transfer integral, hence of the related hole transfer rate, after guanine augmentation (Table 2). This agrees with previous expectations based on the easier oxidability of xG relative to G and the strengthened π - π interactions characterizing xDNA.^{23,26,55} In particular, according to our results at the optimal BHH/6-311++g(3df,3pd) level, the effective electronic coupling in the ideal xGC-xGC stack is almost three times the coupling in the ideal GC-GC system. The relative effects of the basis set on V_{IF} are reduced in the presence of the expanded bases (cf. Tables 1 and 2). In particular, the V_{IF} values from the four best basis sets differ by less than 10% at $^{\text{GC}}Q_{x1i}$ and less than 5% at $^{\text{GC}}Q_{x2i}$.

The values of ΔE_{IF} , which can be chosen as the reaction coordinate,⁵⁸ and V_{IF} for AT-AT base stacks at different theoretical/computational levels and configurational coordinates are

reported in Table 3. Due to the large scattering found among published values of the transfer integrals for AT-AT and A-A stacks, which is instead not the case for guanine stacks, we inspect more structures in order to glean the origin of the controversial data. The two ideal structures $^{AT}Q_{2i}$ and $^{AT}Q_{3i}$ are superposed in Figure 2a. The real structure $^{AT}Q_{1r}$ is illustrated in Figure 2b, while Figure 2c shows a fragment from a nicked B-DNA duplex that incorporates both $^{AT}Q_{2r}$ and $^{AT}Q_{3r}$, in red and blue, respectively. While we noted above that the transfer integral does not change much when adding the complementary cytosines to a G-G stack, indicating the negligible role of the pyrimidines in hole transfer processes within G-rich portions, the situation is at first sight more complicated for A-rich stacks. Table 3 denotes a strong structural dependence of V_{IF} in the ideal conformation of AT-AT that prevents an immediate interpretation of whether or not the thymines are active in hole transfer from direct comparison to the values of the ideal A-A stack (Discussion section). However, we recognize that the values of V_{IF} do not appreciably change by removing the thymine from $^{AT}Q_{1i}$, $^{AT}Q_{2i}$ and $^{AT}Q_{3i}$ (to yield the A-A stacks denoted by $^AQ_1'$, $^AQ_2'$ and $^AQ_3'$ in Table 7), indicating that the pyrimidines do not significantly interfere with charge transfer also in A-rich segments. In fact, the values of V_{IF} for $^{AT}Q_{1i}$, $^{AT}Q_{2i}$ and $^{AT}Q_{3i}$ in Table 3 should be compared to those for $^AQ_1'$, $^AQ_2'$ and $^AQ_3'$ in Table 7. We note that the transfer integrals for the real structures $^{AT}Q_{1r}$, $^{AT}Q_{2r}$ and $^{AT}Q_{3r}$ significantly differ among them and with respect to the ideal structures, indicating a strong dependence on the conformation⁵⁹ and on the details of the structure within a given conformation.

Our results for ΔE_{IF} and V_{IF} of xAT-xAT dimers at different theoretical levels and different configurational coordinates are summarized in Table 4. No reference data exist. The transfer integrals of the ideal xAT-xAT dimers are sensibly larger than those of the ideal AT-AT dimers reported in Table 3. In the case of the real systems, this is not always true: $V_{IF}(^{AT}Q_{1r})$ is definitely larger than $V_{IF}(^{AT}Q_{2r})$ but quite similar to $V_{IF}(^{AT}Q_{3r})$ and smaller than $V_{IF}(^{AT}Q_{2r})$. Namely, the comparison suffers from the structure/conformation variability that we have just pointed out for the AT-AT stacks. These arguments are elaborated in some detail in the Discussion section.

Summarizing this section (see Table 5), we find that the aromatic base expansion yields a strong increase of the electronic coupling in some systems, which is however affected by conformational fluctuations. In fact, the increase is noticeable in xGC-xGC, for which we could only inspect an ideal configuration, relative to GC-GC. It is also true for ideal xAT-xAT relative to ideal AT-AT, but less clear for the real xAT-xAT fragments extracted from pdb files. This suggests that the efficiency of the charge transfer in a realistic situation will eventually be affected by the dynamical flexibility.^{10,60} Anyway, an overall increase of the effective electronic couplings upon size expansion emerges from the summary of Table 5.

Nucleobase expansion and CT rate: implications for nanoelectronics

Various experimental data on the hole transfer through DNA systems have been interpreted in terms of a multi-hopping mechanism, which involves both the G and A bases as charge carriers.^{9,61,62} The G-hopping through short $(AT)_n$ bridges, up to $n = 3-4$, has been described to occur by coherent superexchange (unistep superexchange hopping), which becomes an elementary ET process in the case of adjacent guanines. In the presence of a long bridge, the charge is carried by the intervening adenine-containing base pairs, after thermally induced endothermic hole excitation to an intermediate A. The overall CT rate clearly depends on the effective electronic couplings between nearest neighbor bases. The unistep superexchange rate constant has an approximate exponential dependence on the donor-acceptor distance R , given by the relation^{62,63}

$$k = \frac{2\pi}{\hbar} \langle |V(D, B_1)|^2 \exp(-\beta R) \rangle \rho_{FC} \quad (4)$$

where

$$\beta = \frac{2}{R_0 n} \left[\ln \frac{V(B_n, A)}{\Delta E(D, B_1)} + \sum_{j=1}^{n-1} \ln \frac{V(B_j, B_{j+1})}{\Delta E(D, B_{j+1})} \right]. \quad (5)$$

B_j are the base pairs in the bridge between donor and acceptor, ρ_{FC} is the thermally averaged density of states, weighted by the Franck-Condon factor, $V(D, B_1)$, $V(B_j, B_{j+1})$, and $V(B_n, A)$ are the transfer integrals between adjacent base pairs, $\Delta E(D, B_j)$ are the respective (diabatic) energy gaps, and R_0 is the nearest neighbor spacing. In Eq. 4 the average on the electronic factor accounts for the effects of duplex conformational changes on the CT rate in the limit of slow modulation,⁶⁴ whereas the dynamical coupling between nuclear and electronic coordinates is neglected.

According to the present results, the passage from a DNA duplex to the corresponding xDNA duplex with expanded G and A bases can result in a gain of a factor of ~ 8 in the hole transfer rate for any couple of adjacent xG bases, and a significant gain for xAT-xAT stacks related to the average in Eq. 4 (or in a more sophisticated equation when also the relaxed Condon approximation is not satisfied). This enhancement is coherent with the explanation of the higher structural stability of xDNA in terms of enhanced π - π coupling. Future studies on the connection between conformation changes and electronic couplings in DNA⁶⁰ and xDNA stacks are desirable (also in the light of the increased structural stability reported for xDNA^{21,22,23,24}), in order to design sequences with tailored velocities of the hole transfer processes. A quantitative analysis of the CT in xDNA requires further V_{IF} evaluations on different nucleobase sequences, the conformation averaging indicated in Eq. 4, an analysis of the relaxed Condon approximation, and a deeper understanding of the incoherent CT mechanisms.

Despite the need for additional work to unravel the features of CT in xDNA relative to DNA, the present work quantifies the effectiveness of the base expansion on the hole transfer rate, as summarized in Table 5. We also indicate various crucial points that call for further attention for the development of controllable xDNA blocks potentially useful to nanoelectronics.

The nature of the hopping steps and the connection to the decay parameter β for the CT rate are still under debate. A crossover from a relatively high β , in sequences with short adenine-rich bridges between guanine or multiple-guanine traps, to a much smaller β , in bridges including more than 3 adenines, has been observed.⁹ The two β regimes can be interpreted in terms of different CT mechanisms, with the value of β dependent on the relative magnitude of the injection energy⁶⁵ and the transfer integrals between the neighboring nucleobase pairs.¹⁰

If the injection energy is significantly larger than the transfer integral, a high β value is observed for short donor-acceptor distances, with consequent fast decay of the CT rate. The CT occurs through a single-step superexchange mechanism in this limit.

If the magnitudes of the injection energy and of the charge transfer integral are comparable to each other, the injected charge can populate the bridge between the donor and acceptor sites and a low β value is obtained. For long enough donor-acceptor distances, even if β is larger

than the transfer integral, the charge injection into the bridge is favored. In this condition, the CT occurs through an incoherent mechanism: either hopping of localized charges^{9,66,67,68} or polaron diffusive motion.^{69,70,71} According to a recent theoretical work on the solvent reorganization energy for hole transfer in DNA,⁷² although the delocalization of the hole is energetically unfavorable, the transfer of a hole spread over a few bases is characterized by a lower energy barrier than the transfer of a hole localized at a single base.

The CT mechanism is determined by two important parameters:¹⁰ the site energies, which correspond to the localization of the charge on a single base or base pair, and the transfer integral. Both these parameters depend on the fluctuations of the DNA structure. Here the structural dependence of the effective electronic coupling is analyzed only on specific atomic configurations. This allows us to address the methodological approach and to understand some important dynamical effects (see next section).

IV. Discussion of relevant theoretical/computational aspects

In this section we (i) discuss important issues concerning the two-state single-particle model (widely used in the literature); (ii) explore the appropriateness of the Condon approximation and (iii) analyze the performance of our theoretical-computational method by comparing our results with previous electronic coupling estimates in natural DNA and xDNA. Our analysis (i) highlights some shortcomings of the two-state model at the single-particle level of investigation; (ii) gains some useful insights into the limited pertinence of the Condon approximation; (iii) supports the robustness of our approach and (iv) helps clarifying some discrepancies in the existing literature, thus giving useful insights for future transfer integral calculations.

Effective electronic coupling in single-stranded G-G and A-A stacked dimers: paradigmatic cases for testing the robustness of the DFT computational method

The excitation energies and transfer integrals for a G-G stack from a regular B-DNA structure are reported in Table 6. The upper part of the Table contains our results with different basis sets and different exchange-correlation functionals. The lower part of the Table contains results by other authors with different computational setups.

The basis set effects on the couplings are carefully tested by using the BHH functional. Due to the presence of the π -orbitals and the spatial gap between the nucleobases, the V_{IF} value for hole transfer is sensitive to the presence of triple-zeta, polarization, and diffuse functions in the basis sets for the heavy atoms, while their inclusion on hydrogen atoms has minor effects. The TZVP, cc-pVTZ, 6-311++g**, and 6-311++g(3df,3pd) basis sets turn out to yield comparable accuracy. One may claim that the transfer integral corresponding to the 6-311++g(3df,3pd) basis set is somewhat smaller than the V_{IF} values using the other three basis sets. Indeed, the difference is not substantial and is consistently found for all the systems under study. Consequently, all the conclusions of this work can be elicited from any of those basis sets. Nevertheless, we consider the value of 0.065 eV, resting on the use of the largest standard Pople-style basis set 6-311++g(3df,3pd), as our best estimate of V_{IF} for G-G. This basis set outdoes the diffuse basis set that has been successfully used by others⁷⁶ for the calculation of electron affinities and ionization potentials of single nucleotide bases. In addition, its use in conjunction with the B97-1 hybrid functional has been proposed⁷⁷ as an effective conventional DFT approach to the van der Waals interactions, which control the interactions between the nucleotides in DNA.^{43,78} Note, however, that our results in Table 1 show that the use of the first and last functionals of the Becke-97 series, as well as of the other commonly employed hybrid functionals B3LYP⁷⁹ and PBE0,⁸⁰ yields unusually large transfer integrals when used in conjunction with the 6-31* basis set, thus indicating an excessive delocalization of the valence electronic charge. We expect that this trend holds when more extended basis sets are

adopted. Therefore, the suggestion made by Becke and others⁷⁸ about the B97-1/6-311++g (3df,3pd) scheme as the best computational framework for describing van der Waals interactions within DFT cannot be extended to the calculation of transfer integrals, where the tails of some molecular orbitals, and in particular of the HOMO, play a crucial role. In these cases, a critical improvement comes from the use of the BHH functional.

Our V_{IF} value at the BHH/6-311++g(3df,3pd) level is close to the recent estimate from Ref. 50, but the value of the effective electronic coupling is here established at a different theoretical level, using full DFT. Within the theoretical framework of Ref. 50 the electronic coupling H_{IF} is computed by means of the self-consistent-charge density functional tight-binding method (SCC-DFTB), and V_{IF} is derived, in a single-orbital picture (hence, by making use of an effective Hamiltonian operator), from the Löwdin transformation⁸¹ $V_{\text{IF}} = H_{\text{IF}} - S_{\text{IF}} (H_{\text{II}} + H_{\text{FF}})^{-1} S_{\text{IF}}^T$. The resulting values of H_{IF} and V_{IF} differ significantly, and we report both quantities from Ref. 50 in Table 6. Our four best estimates of V_{IF} fall into the gap between these two values, which can also be expected on the basis of the fact that V_{IF} from the Löwdin transformation and H_{IF} are an underestimate and an overestimate, respectively, of the true transfer integral (V_{IF} can only be increased by quadratic terms in S_{IF} neglected in the Löwdin transformation⁸¹). An even larger difference between H_{IF} and V_{IF} is found in Ref. 73, where a full DFT fragment orbital (FO) approach is used. It is worth noting that the derivation of V_{IF} from H_{IF} through the Löwdin transformation (which preserves the simple form reported above for small enough overlaps) avoids the orthogonalization of the fragment orbitals, but can be affected by the approximations involved in the single-orbital framework. The relevant error is unpredictable, partly related to the fact that in the full-electron picture, where $H_{\text{II}} = E_{\text{I}}$ and $H_{\text{FF}} = E_{\text{F}}$, the two quantities H_{IF} and $S_{\text{IF}} (H_{\text{II}} + H_{\text{FF}})^{-1} S_{\text{IF}}^T$ are much larger than their difference and thus liable for significant errors once any approximation is introduced. Ultimately, our best DFT calculations provide a narrow range for the value of V_{IF} , against the large and variable gap between V_{IF} and H_{IF} coming from Refs. 50 and 73. Note also that the SCC-DFTB method inherits the shortcomings of DFT (due to the approximate character of any currently available exchange-correlation functional), and introduces additional approximations relative to the full self-consistent field DFT.⁸² In DFTB most of the electron-electron interactions, including exchange and correlation, are captured in a single-particle Hamiltonian. The missing interaction plays an important role in charged systems. In fact, the DFTB approach yields an excessive localization of the net charge in electronic systems involving π -orbitals.⁸² The approximate inclusion of the missing interaction in the improved SCC-DFTB method can essentially recover the delocalization of the excess charge (as in the systems of Ref. 82), but a residual error in the charge delocalization can still determine a considerable underestimation of the transfer integral between the stacked bases in the xGC-xGC system. While the safe statement of this conclusion requires further theoretical analysis of the relevant multi-electron effects, the arguments in this paragraph support our DFT method against the possibility that additional approximations involved by the SCC-DFTB method accidentally yield a better effective electronic coupling.

On the other hand, let us note that the best calculations in Ref. 50 include dynamical effects by means of MD-coupled simulations, which also allow to obtain MM charges to mimic polarization effects of the environment. Indeed, Eq. 1 has been also applied, in combination with MD, to significantly larger systems (up to 142 atoms).⁸³ Moreover, the same Eq. 1 can be easily implemented in the approximate DFT scheme of Ref. 50, as well as in any QM/MM computational schemes. The minor role of the backbone^{30,31} and the properties of a suitable environment for molecular electronics applications should be considered, in order to infer the actual influence of the external environment and the most convenient theoretical setup.

As shown in Table 6, our method gives a best value of the transfer integral for the G-G stack smaller than the HF values ensuing from diverse methods, namely an energy-splitting method,

³² the fragment charge difference (FCD) method,⁷⁴ and the generalized Mulliken-Hush (GMH)⁸⁴ method,⁷⁵ in the Koopmans' theorem⁸⁵ approximation (KTA). Note that the comparison between our results and those reported from HF calculations is not invalidated by the fact that the latter were obtained with a poorer basis set. In fact, while the inclusion of diffuse (in the Pople-style basis sets) and polarization basis functions turns out to have a remarkable effect within DFT, it is not crucial within HF schemes.^{29,32,74}

Some HF values were already demonstrated to be quantitatively imprecise by more recent studies with post-HF approaches.⁷⁵ Our value of V_{IF} from Eq. 1 at the BHH/6-311++g(3df, 3pd) level is comparable to post-HF values, in particular to that of the complete active space self-consistent (CASSCF) approach with an active space made up of 7 electrons within 8 π -orbitals.⁷⁵ In fact, the corresponding V_{IF} and ΔE_v differ by ~2% and ~3%, respectively. The matching with CASSCF(7,8) supports our DFT method, thus its application to the larger base stacks considered in the present work, where post-HF methods become unmanageable. However, one should be careful in taking CASSCF results in general as those of maximum achievable accuracy, and indeed our best value for V_{IF} proposes as an alternative benchmark. In fact, multi-reference schemes, such as CASSCF, take account of the electron correlation only in the active space, and have some theoretical disadvantages with respect to our DFT-based approach. For instance, in order to obtain quantitative accuracy, multi-reference treatments often need empirical decisions, such as the choice and the handling of the active orbitals.^{86,87} In addition, a proper description of electron correlation effects requires large basis sets that are affordable in a DFT approach but are often excessively expensive or unfeasible for post-HF methods such as CASSCF or CAS-PT2 (e.g., in Ref. 75 the 6-31g* basis set is used).

Our theoretical approach is robust against basis set superposition errors (BSSE). In fact, by evaluating the BSSE through the counterpoise method,⁸⁸ we obtain an almost full cancellation of the BSSE on E_I and E_F . Ultimately, the overall effect of the BSSE on the transfer integral turns out to be almost negligible for small basis sets (e.g., 6-31g*), and completely negligible for large enough basis sets (e.g., TZVP).

Before turning to the A-A stack, let us make a further remark on the importance of the choice of the exchange-correlation functional for the DFT evaluation of transfer integrals. Waller and co-workers⁴³ obtained trends similar to ours in comparing the performance of different hybrid functionals for the evaluation of binding energies in π -stacked complexes, which are quite sensitive to the small electron density in the intermolecular region. They also showed⁴³ that the BHH functional is able to capture effects of subtle changes in electrostatic and dispersion forces, and gives results in good agreement with the best available data from post-HF calculations for a wide range of π -stacked systems, including DNA bases. They have argued⁴³ that the excellent performance of BHH might stem from a fortuitous cancellation of errors in the exchange and correlation energy terms that mimics the dispersion part of post-HF methods. Note also that, for all the π -stacks considered in this work, a substantial cancellation of errors incidental to van der Waals (in particular, dispersion) interactions can be reasonably expected when the energy difference ΔE_{IF} , entering Eq. 1, is calculated. Indeed, our careful tests suggest that the good performance of BHH for the systems studied in this work could ensue from an inherently good approximation of the exact exchange-correlation functional produced by the linear interpolation of the coupling-strength parameter in the adiabatic connection formula.³⁸ The above considerations on the BHH functional fall within the long-standing controversy about the validity of the currently available approximate DFT functionals. Indeed, our methodological checks offer a strong support to the use of DFT also to tackle problems that require the computation of transfer integrals. At least, BHH turns out to be an appropriate functional around the equilibrium separation between DNA (and xDNA) nearest neighbor nucleobases, which is very close to the stacking distance in the benzene dimer.⁴³ Note

also that, although in general DFT can yield an excessive delocalization of the valence charge due to spurious electron self-interaction⁸⁷ (as shown by the tests on various exchange-correlation functionals reported in Table 6), the comparison with the CASSCF method essentially corroborates the performance of the BHH functional. Possible residual effects of the electron self-interaction on the value of the effective electronic coupling may even decrease as a consequence of the nucleobase expansion, because the relevant charge is spread over a wider aromatic structure on each expanded base.

DNA base systems involving the A-A stack represent a problematic case for the evaluation of V_{IF} , as revealed by the spreading of published data over more than an order of magnitude (see the lower panel in Table 7). In addition, HF calculations³² indicate that the coupling between the adenines is slightly reduced by the presence of the adjacent thymines, while recent density functional tight-binding results⁵⁰ show an opposite trend. Our results for A-A stacks with the base step parameters of the regular B-DNA structure are reported in the upper panel of Table 7. The technical tests were done on a previously studied structure,⁷⁵ so to have a precise direct benchmark.

The relative difference between the V_{IF} values using accurate Pople-style basis sets is rather small and in line with the trend of the results in Table 6. As in the case of the G-G stack, the calculations done with the TZVP and cc-pVTZ basis sets, hence without diffuse functions, somewhat overestimate the electronic coupling, though still giving comparable results. Our estimates of V_{IF} for the structure $^A Q_{1i}$ are distributed around the CASSCF(7,8) value for the same structure⁷⁵ and are consistently larger than the HF values, therefore suggesting an excessive localization of the valence electron by the HF scheme. Note also that our calculations at the BHH/TZVP level do not converge to the true DFT ground state unless accurate enough convergence thresholds are used (as detailed in the Supporting Information, Figure S1). This happens because, at the given nuclear coordinates, the self-consistent field BHH/TZVP calculation explores a flat energy region before reaching the energy minimum. Out-of-convergence BHH/TZVP calculations yield ΔE_v and V_{IF} values comparable to those from multi-configuration calculations with the large (11,12) active space. We therefore argue that the CASSCF(11,12) and CAS-PT2(11,12) evaluations reported in Table 7 also may be affected by a computational issue. In fact, multi-configuration calculations can lead to several non-equivalent local minima whose number increases with the size of the active space,⁸⁹ thereby complicating the achievement of the global energy minimum. Indeed, the excitation energies and transition dipole moments shown in the Table 3 of ref ⁷⁵ do not allow us to guess about the achievement of the convergence in the active space for the CASSCF calculations, while the use of a larger active space is still unfeasible.

According to the above scenario, it can be reasonably expected that the correct value of V_{IF} for the $^A Q_{1i}$ A-A stack lies in the range between the CASSCF(7,8) and HF GMH-KTA estimates, where both the BHH/6-311++g** and BHH/6-311++g(3df, 3pd) values are located. Note also that the CASSCF(7,8) approach gives very similar values of V_{IF} for the G-G and A-A stacked dimers. Thus, if a different solution for the A-A stack can be obtained from the (11,12) active space, similar differences among the multi-configuration evaluations for the two base stacks would lead to a post-HF best value of V_{IF} probably close to our best estimate.

At the most accurate BHH/6-311++g(3df, 3pd) DFT level emerging from our tests, we notice a large spread of the V_{IF} values obtained for different configurational coordinates. The base step parameters are consistent in the set $\{^A Q_{ki}\}_{k=1,2,3}$, yet there are remarkable differences in the pertaining transfer integrals within the same set. Specifically, the $^A Q_{2i}$ and $^A Q_{3i}$ geometries correspond to out-of-plane displacements of the amino hydrogens on opposite adenines (Figure 3) of ~ 0.09 Å and ~ 0.19 Å, and to V_{IF} values of 0.032 eV and 0.010 eV, respectively. By inspecting the computed spin-orbitals, we note that significant portions of HOMO and LUMO

(Supporting Information) are localized around the amino nitrogens and are affected by the displacement of the bound hydrogens. This can have a strong effect on the hole transfer integral between the two flanking bases, as the latter is strictly related to the small tail of the HOMO on one of the two adenines. Based on the remarkable difference between $V_{\text{IF}}(^A Q_{2i})$ and $V_{\text{IF}}(^A Q_{3i})$ we infer that the Condon approximation is not fulfilled by the kind of nuclear mode shown in Figure 3, although the latter does not involve base step parameters, hence conformational changes. Consequently, the coupling between stacked adenines can be significantly affected by atomic displacements within a given conformation. This could also be at the origin of part of the discrepancies among different works reported in Table 7.

Discussion on the calculated transfer integrals and their geometry dependence for each base pair dimer

1. GC-GC—We showed in Table 1 that our best value of 0.075 eV for $V_{\text{IF}}(^{GC} Q_i)$ in the ideal GC-GC dimer is close to the SCC-DFTB value⁵⁰ and lies between the two HF estimates from Refs. 32 and 51. All of these values are significantly larger than the corresponding V_{IF} estimate at the INDO/S⁹⁰ level,¹⁸ while they are all reasonably similar to the rms-INDO/S in the presence of an adjacent nucleobase.¹⁸ Ref. 18 provides a possible order of magnitude for the transfer integral in a real GC-GC stack. On the other hand, Table 1 displays a wide range of V_{IF} values for the hole transfer in the ideal GC-GC stack. An analogous spreading of the results from different theoretical methods can be expected for any conformation of the GC-GC stack in a real B-DNA chain. This suggests the possibility of a significant systematic error in the rms effective electronic coupling whenever the averaging procedure applies to not accurate enough evaluations for the individual conformations. In this respect, significant advantages of our method come from the fact that it adopts a full-electron description and does not make direct computational use of the Löwdin transformation (not even in a multi-electron picture). In addition, Eq. 1 can give a good approximation beyond the two-state model.³⁴

Regardless of their relative accuracies, all the methods listed in Table 1 contribute to define a maximal range (with reference to both the method and the dependence on conformation changes) for the transfer integral in base pair dimers involving the G-G stack. In this respect, although the values from Refs. 18 and 51 of the coupling for the GC-GC stack differ significantly at a single-point level, they still represent valuable information, in relation to the comparison with the xGC-xGC system. This information is extended to the A-DNA stack in Ref. 51, where the maximum absolute changes of V_{IF} due to structural perturbations by variation of the base step parameters are quantified as 0.059 eV and 0.099 eV for A-DNA and B-DNA, respectively. For example, the value $V_{\text{IF}}(^{GC} Q_r) = 0.058$ eV here obtained for a stack from real A-DNA (see rightmost column in Table 1) is equal to the value in Ref. 51 at the ideal geometry. Such a perfect matching certainly has a fortuitous connotation, given the differences in the respective geometries and theoretical-computational methods. On the other hand, the comparison between our estimates of the effective coupling for the GC-GC stacks from A-DNA and B-DNA (whose geometries further differ for the structural perturbation of the second structure relative to the base step parameters for regular A-DNA) may still suggest the possibility of narrower V_{IF} ranges at variable conformations, having reference to the nature of the G base π -stacking and being partially robust relative to the kind of DNA structure.

2. xGC-xGC—As mentioned in the context of Table 2, the coordinate $^{GC} Q_{x1i}$ for the ideal xGC-xGC dimer is close to the transition state coordinate Q_t where the ET diabatic states come to degeneracy (namely, $\Delta E_{\text{IF}} = 0$ and consequently $\Delta E_v = 2V_{\text{IF}}$), thus maximizing the probability for the CT process^{17,19}. In fact, $\Delta E_{\text{IF}}(^{GC} Q_{x1i})$ is much smaller than $V_{\text{IF}}(^{GC} Q_{x1i})$ (see the upper panel in Table 2), so that $2V_{\text{IF}}(^{GC} Q_{x1i})$ is almost equal to $\Delta E_v(^{GC} Q_{x1i})$. At the transition state coordinate the expression of V_{IF} in Eq. 1 shows an eliminable discontinuity, since both its numerator (because of $\Delta E_{\text{IF}} = 0$) and its denominator (because $a = b$) vanish.³⁴

Thereby, the evaluation of the transfer integral requires an increasing computational accuracy when Q_t is approached. As an alternative, provided that the Condon approximation, which consists in neglecting the dependence of V_{IF} on Q ,⁹¹ is satisfied, V_{IF} can be evaluated at some coordinate farther from Q_t . Therefore, in order to assess the reliability of V_{IF} ($^{GC}Q_{x1i}$), the near degeneracy of the diabatic CT states, relative to their electronic coupling, has been removed through a geometry optimization on one xG, leading to the coordinate $^{GC}Q_{x2i}$. According to the outcome reported in Table 2, the relative difference between V_{IF} ($^{GC}Q_{x1i}$) and V_{IF} ($^{GC}Q_{x2i}$) is rather small, less than 5% at the BHH/6-311++g(3df, 3pd) computational level, despite the outstanding change in ΔE_{IF} . The small difference in V_{IF} can be attributed to a small rotation of the relaxed xG relative to the complementary C base, validates the transfer integral evaluation near the transition state coordinate and indicates the pertinence of the Condon approximation for changes in the nuclear coordinates that preserve the conformation (as defined by the base step parameters).

All the results presented in this work offer a good framework to explore the validity of the Condon approximation in various DNA and xDNA sequences and conformations. In this context, by using a full-electron computational scheme, we go beyond current understanding derived by single-particle HF methods^{29,92} on a few DNA systems.

3. AT-AT—Differently from the case of GC-GC stacks, our calculations indicate that in AT-AT stacks the Condon approximation is only partially valid, even within the same conformation. The perfect matching of the V_{IF} values at the ideal coordinates $^{AT}Q_{2i}$ and $^{AT}Q_{3i}$ reported in Table 3 expresses the fulfillment of the Condon approximation along a suitable reaction coordinate connecting $^{AT}Q_{2i}$ with $^{AT}Q_{3i}$. However, both V_{IF} ($^{AT}Q_{2i}$) and V_{IF} ($^{AT}Q_{3i}$) are different from V_{IF} ($^{AT}Q_{1i}$), even though $^{AT}Q_{2i}$ is closer to $^{AT}Q_{1i}$ than to $^{AT}Q_{3i}$ along the reaction coordinate ΔE_{IF} . This expresses a failure of the Condon approximation in ideal AT-AT stacks, which can be attributed to a distortion of the amino group in the relaxed portion of both $^{AT}Q_{2i}$ and $^{AT}Q_{3i}$. Such distortions are crucial for the efficiency of charge transfer because the amino groups bear significant portions of the tails of the most relevant spin-orbitals (Figure 4). The failure of the Condon approximation in AT-AT systems complicates the comparison of V_{IF} values in natural and expanded A-rich fragments, because of the high sensitivity to structural details. This point will be discussed further below.

4. xAT-xAT—The electronic couplings of both the ideal and real xAT-xAT stacks in Table 4 are similar to the one obtained for AT-AT in the real geometry $^{AT}Q_{2r}$. Table 4 also shows that ΔE_{IF} is smaller than V_{IF} at the ideal geometry $^{AT}Q_{x1i}$, which is an indication of the proximity to the transition state coordinate Q_t . Therefore, as explained above, we considered other ideal structures, obtained by moving away from Q_t . This was achieved through a geometry optimization on one expanded adenine, starting from the $^{AT}Q_{x1i}$ geometry. The ideal $^{AT}Q_{x2i}$ and $^{AT}Q_{x3i}$ structures were drawn from the first and the last relaxation step, respectively. $^{AT}Q_{x1i}$, $^{AT}Q_{x2i}$ and $^{AT}Q_{x3i}$ correspond to the same ideal conformation but have structural deformations with respect to each other. Although $^{AT}Q_{x1i}$ and $^{AT}Q_{x2i}$ are well separated along the reaction coordinate ΔE_{IF} , the respective V_{IF} values are identical within our accuracy at the explored BHH/cc-pVTZ computational level. This result validates the robustness of our approach even in close proximity of the transition state coordinate, and helps to establish the relevance of the Condon approximation within a given conformation for xAT-xAT fragments, after its assertion above for the xGC-xGC system. Therefore, the evaluation of V_{IF} is expected to be meaningful at any point of a two-dimensional configurational space, where the various conformations of a given nucleobase pair dimer are represented by different points along an axis, and the atomic configurations around a fixed conformation correspond to an axis orthogonal to the first one through the point representing the given conformation. The two axes describe an accepting and an inducing mode, respectively, according to the nomenclature of Ref. 64. An accepting mode is energetically coupled to the electron transition,

so that the ET probability is significant for a regime of nuclear configurations around the transition state coordinate, where the nuclear factor of the electron transfer rate has a maximum^{17,19} and the Condon approximation can be applied. Nevertheless, its applicability cannot be excluded out of the favored local regime of configurations. In fact, our results support its validity also out of that regime, significantly far from the set of nuclear coordinates that corresponds to the transition state in the given conformation. Hence, the transfer integral can be reliably evaluated at any coordinates pertaining to that conformation. On the contrary, along an inducing mode, the energies of the donor and acceptor states weakly depend on the configuration. As a consequence, the electron transfer is not limited to a local configuration regime and generally leads to a breakdown of the Condon approximation. Note that a single reaction coordinate, sufficient for describing the CT process, can be defined as an average over the inducing mode, therefore being related to a free energy concept.

According to the above picture, we can say that $^{AT}Q_{x1i}$ and $^{AT}Q_{x2i}$ are essentially coincident along the inducing mode and appreciably separated along the accepting coordinate. On the contrary, the $^{AT}Q_{x3i}$ geometry involves a slight deviation from the idealized conformation, i.e., a roll of $\sim 1.6^\circ$. As a consequence, although the difference between $\Delta E_{\text{IF}}(^{AT}Q_{x2i})$ and $\Delta E_{\text{IF}}(^{AT}Q_{x3i})$ is negligible, $V_{\text{IF}}(^{AT}Q_{x3i})$ is appreciably smaller than $V_{\text{IF}}(^{AT}Q_{x2i})$, yet of the same order of magnitude. In other words, $^{AT}Q_{x2i}$ and $^{AT}Q_{x3i}$ are appreciably separated along the inducing mode, while quite close along the accepting mode. Moreover, the difference between $V_{\text{IF}}(^{AT}Q_{x2i})$ and $V_{\text{IF}}(^{AT}Q_{x3i})$ may be expected on the basis of other reports on similar structural changes in natural DNA dimers.^{74,93}

Transfer integrals of AT-AT: role of the nuclear configuration and shortcomings of the one-particle two-state model

The more complicated picture ensuing from the DNA systems with natural adenines on the one hand denotes a strong dependence of V_{IF} upon some critical features of the nuclear configuration, on the other hand highlights some shortcomings of the one-electron two-state model. Nevertheless, our full-electron calculations confirm the expectations from Ref. 32 that the effective electronic coupling is barely affected by the presence of the thymines. On the contrary, according to Ref. 50, V_{IF} in the ideal conformation is slightly larger for the stack that includes the complementary bases, but unexpectedly corresponds to a negligible overlap between the reactant and product states, as can be deduced by the equality of V_{IF} and H_{IF} in the Table 3 therein.

The scenario is further complicated by the outcome of our full-electron V_{IF} calculations on the real AT-AT structures of Figures 2b–c. The $^{AT}Q_{1r}$ geometry fits into a covalently cross-linked B-DNA duplex relevant to the study of repair mechanisms, which shows a dramatic widening of the major groove, though without disruption of the Watson-Crick base pairing.⁵⁶ $^{AT}Q_{2r}$ and $^{AT}Q_{3r}$ differ from the ideal structure significantly more than $^{AT}Q_{1r}$ and belong to a nicked DNA molecule (representing a substrate for repair enzymes), which adopts a B-DNA conformation despite the absence of a phosphate group.⁵⁷ $V_{\text{IF}}(^{AT}Q_{1r})$ is quite small and similar to the effective coupling in the ideal stacks. $V_{\text{IF}}(^{AT}Q_{2r})$ is larger by one order of magnitude and very close to the rms value in Ref. 18. $V_{\text{IF}}(^{AT}Q_{3r})$ has the same order of magnitude as $V_{\text{IF}}(^{AT}Q_{2r})$ and is even larger. The difference in the effective coupling between the $^{AT}Q_{2r}$ and $^{AT}Q_{3r}$ geometries can be traced back to the larger tilt in the latter (Figure 2c), which causes a different stacking of the two adenines, while the huge difference between these two transfer integrals and $V_{\text{IF}}(^{AT}Q_{1r})$ deserves further investigation. The occurrence of distorted structures that enable an efficient CT channel between stacked adenines depends on conformation and flexibility of the B-DNA system under consideration. Moreover, the data presented in Table 3 indicate a wide range of V_{IF} , which is comparable to the maximum range of 0.099 eV from

Ref. 51 and demands a particularly careful sampling of the configurational space in evaluating the rms transfer integral of a real stack.

Within the two-state/single-particle model, the analysis of the above scenario rests on the spin orbitals illustrated in Figure 4. In particular, the distribution of the (valence) charge in the adiabatic ground and first excited states is usually obtained from HOMO and LUMO, respectively. They form an orthonormal orbital basis to represent the localized HOMOs of the diabatic ET states, where the transferring electron charge is assumed to be essentially localized before and after the charge transfer process. Moreover, in the nonadiabatic ET regime HOMO and LUMO are expected to each substantially overlap with one of the localized HOMOs. As shown by Figure 4, in the ${}^{AT}Q_{1r}$ geometry, as well as in the ideal structures, the HOMO is mainly localized on a thymine (i.e., T₁ in Figure 4a) and the LUMO is mainly localized on the opposite adenine (A₂). Hence, correspondingly appropriate spin orbitals representing the diabatic states should be mainly localized on the same two nucleobases. However, our full-electron calculations on the donor and acceptor groups lead to diabatic states localized on the adenines. The above mentioned model could be recovered by considering the HOMO-1 of the ground electronic state, localized on A₁, in place of its HOMO. As a matter of fact, while the LUMO of the overall charged system is always localized on one adenine, as expected on the basis of the A and T ionization potentials,²⁰ the presence of the electron hole lowers the potential energy of the electron charge on the opposite adenine, and thus the HOMO can also be localized on the thymine of the acceptor base pair. This corresponds to an inversion of the level ordering expected for the single AT pair, which is the case for the ideal AT-AT stacks. On the other hand, in Q_{1r} the localization of the HOMO on T₁ is also favored by the distorted geometry of the methyl group, with a H $\ddot{C}$ H angle of about 77°, which raises the energy of the outer spin orbital localized on T1. In fact, we find a HOMO localized on A₁ after geometry optimization on the methyl group, so that the two-state/single-particle model can be used in its standard way. However, our full-electron calculations reveal that the methyl relaxation and the consequently rearranged level scheme do not correspond to an increase in the transfer integral and that, essentially, the overall CT process proceeds always between the adenines.

In summary, our first principles full-electron calculations on the AT-AT and xAT-xAT stacks point to an increase of the average coupling in the AT-AT dimers upon A size expansion, although significantly distorted AT-AT stacks (${}^{AT}Q_{3r}$) relevant to practical applications, can lead to effective electronic couplings comparable to or even larger than the couplings between adjacent xAT pairs. A future work with transfer integral evaluations on MD snapshots is desirable for conclusive corroboration of the above point. However, it is supported by the following considerations: (i) strongly distorted AT-AT stacks (possibly characterized by step parameters prohibitive for the more stable structures with the expanded adenines) still show effective transfer integrals of the same order of magnitude as found in xAT-xAT stacks. (ii) The spreading and stabilization of the valence charge on the size-extended nucleobases relative to the natural ones is favored by a decrease in the electronic energy (see also Ref. 94), irrespective of the π -stacking conformation. (iii) The structural stability around the idealized geometry is expected to give a strong decrease in the transfer integral variance relative to the much more flexible natural DNA.

V. Conclusions

This work presents accurate DFT evaluations of the effective electronic couplings in natural and size-expanded DNA base stacks. The implementation of Eq. 1 in a DFT scheme that makes use of the BHH functional and sufficiently large basis sets allows us to make a reliable comparison between electron transfer integrals in DNA and xDNA. More generally, we offer an efficient method for the study of CT in π -stacked systems, especially in the presence of large aromatic structures, for which post-HF methods are unfeasible.

The calculations on natural systems clarify unsettled issues in the literature, such as the huge spread of previous results on the A-A stack. The present theoretical development is important in relation to the possibility of detecting single-mismatches in DNA sequences, as well as for the comparison with the effective electronic couplings in xDNA systems. Our calculations on the size-expanded stacks show a significant increase of the transfer integrals between augmented bases at an absolute (for xG stacks) or, at least, average (for xAT stacks) level. As a matter of fact, our current results leave open the possibility of a remarkable increase of the transfer integral in xAT sequences. In general, the indications for a faster CT in xDNA than in natural DNA will be assessed after a suitable conformational sampling on the explored sequences and possibly other sequences. Note that the improvement of CT along xDNA duplexes is not trivially expected exclusively on the basis of their increased structural stability²³ and was not found by SCC-DFTB calculations.⁵⁰ It can be potentially exploited for future development of programmable DNA-based nanoelectronics, which requires a more rapid charge transport than in natural DNA duplexes,^{12d} as well as in genetic applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Giacomo Fiorin, Ben Levine, Russel DeVane, Grace Brannigan, Miguel Fuentes-Cabrera, Bobby Sumpter, Anna Garbesi and Andrea Ferretti for helpful discussions and/or technical support. This work was funded by the NIH (grant number GM 067689), and by the EC through projects “DNA-based Nanowires” (contract IST-2001-38951) and “DNA-based Nanodevices” (contract FP6-029192). Computing time was provided by CMM (UPenn, Philadelphia, PA); CINECA (Bologna, IT); ORNL (Oak Ridge, TN) and NERSC (Berkeley, CA) through project CNMS2008-016.

References

1. Boon EM, Livingston AL, Chmiel NH, David SS, Barton JK. *Proc Natl Acad Sci USA* 2003;100:12543. [PubMed: 1455969]
2. DeRosa MC, Sancar A, Barton JK. *Proc Natl Acad Sci USA* 2005;102:10788. [PubMed: 16043698]
3. Heller A. *Faraday Discuss* 2000;116:1. [PubMed: 11197472]
4. Porath D, Cuniberti G, Di Felice R. *Top Curr Chem* 2004;237:183.
5. Endres RG, Cox DL, Singh RRP. *Rev Mod Phys* 2004;76:195.
6. Zwolak M, Di Ventra M. *Rev Mod Phys* 2008;80:141.
7. Boon EM, Ceres DM, Drummond TG, Hill MG, Barton JK. *Nat Biotechnol* 2000;18:1096. [PubMed: 11017050]
8. Drummond TG, Hill MG, Barton JK. *Nat Biotechnol* 2003;21:1192. [PubMed: 14520405]
9. Giese B, Amaudrut J, Köhler AK, Spormann M, Wessely S. *Nature* 2001;412:318. [PubMed: 11460159]
10. Grozema FC, Tonzani S, Berlin YA, Schatz GC, Siebbles LDA, Ratner MA. *J Am Chem Soc* 2008;130:5157. [PubMed: 18324767]
11. Giese B, Wessely S. *Angew Chem Int Ed* 2000;39:3490.
12. (a) Bhattacharya PK, Barton JK. *J Am Chem Soc* 2001;123:8649. [PubMed: 11535068] (b) Boone E, Schuster GB. *Nucleic Acids Res* 2002;30:830. [PubMed: 11809898] (c) Takada T, Fujitsuka M, Majima T. *Proc Natl Acad Sci USA* 2007;104:11179. [PubMed: 17592151] (d) Osakada Y, Kawai K, Fujitsuka M, Majima T. *Proc Natl Acad Sci USA* 2006;103:18072. [PubMed: 17116889]
13. Seeman NC, Lukeman PS. *Rep Prog Phys* 2005;68:237.
14. (a) de Pablo PJ, Moreno-Herrero F, Colchero J, Gómez-Herrero J, Herrero P, Baró AM, Ordejón P, Soler JM, Artacho E. *Phys Rev Lett* 2000;85:4992. [PubMed: 11102169] (b) Gomez-Navarro C, Moreno-Herrero F, de Pablo PJ, Colchero J, Gómez-Herrero J, Baró AM. *Proc Natl Acad Sci USA* 2002;99:8484. [PubMed: 12070346] (c) Braun E, Eichen Y, Sivan U, Ben-Yoseph G. *Nature*

- 1998;391:775. [PubMed: 9486645] (d) Storm AJ, van Noort J, de Vries S, Dekker C. Appl Phys Lett 2001;79:3881.
15. (a) Porath D, Bezryadin A, De Vries S, Dekker C. Nature 2000;403:635. [PubMed: 10688194] (b) Xu B, Zhang P, Li X, Tao N. Nano Lett 2004;4:1105. (c) Cohen H, Nogues C, Naaman R, Porath D. Proc Natl Acad Sci USA 2005;102:11589. [PubMed: 16087871] (d) Van Zalinge H, Schiffrin DJ, Bates AD, Haiss W, Ulstrup J, Nichols RJ. ChemPhysChem 2006;7:94. [PubMed: 16345118] (e) Kang N, Erbe A, Scheer E. New J Phys 2008;10:023030.
16. (a) Nitzan A. J Phys Chem A 2001;105:2677. (b) Nitzan A, Ratner MA. Science 2003;300:1384. [PubMed: 12775831] (c) Berlin YA, Ratner MA. Radiat Phys Chem 2005;74:124.
17. Marcus RA, Sutin N. Biochim Biophys Acta 1985;811:265.
18. Voityuk AA. J Chem Phys 2008;128:115101. [PubMed: 18361616]
19. Marcus RA. J Chem Phys 1956;24:966.1956;24:979. *ibid*
20. Steenken S, Jovanovic SV. J Am Chem Soc 1997;119:617.
21. Liu H, Gao J, Lynch SR, Saito YD, Maynard L, Kool ET. Science 2003;302:868. [PubMed: 14593180]
22. Gao J, Liu H, Kool ET. Angew Chem, Int Ed 2005;44:3118.
23. Liu H, Lynch SR, Kool ET. J Am Chem Soc 2004;126:6900. [PubMed: 15174859]
24. Lynch SR, Liu H, Gao J, Kool ET. J Am Chem Soc 2006;128:14704. [PubMed: 17090058]
25. Fuentes-Cabrera M, Sumpter BG, Wells JC. J Phys Chem B 2005;109:21135. [PubMed: 16853737]
26. Varsano D, Garbesi A, Di Felice R. J Phys Chem B 2007;111:14012. [PubMed: 18034470]
27. Lu XJ, Olson WK. Nucl Acids Res 2003;31:5108. [PubMed: 12930962]
28. Delaney S, Barton JK. J Org Chem 2003;68:6475. [PubMed: 12919006]
29. Rösch N, Voityuk AA. Top Curr Chem 2004;237:37.
30. Priyadarshy S, Risser SM, Beratan DN. J Phys Chem 1996;100:17678.
31. Tong GSM, Kurnikov IV, Beratan DN. J Phys Chem B 2002;106:2381.
32. Voityuk AA, Jortner J, Bixon M, Rösch N. J Chem Phys 2001;114:5614.
33. Newton MD. Chem Rev 1991;91:767.
34. Migliore A, Corni S, Di Felice R, Molinari E. J Chem Phys 2006;124:064501.
35. (a) Wu Q, Van Voorhis T. Phys Rev A 2005;72:024502. (b) Wu Q, Van Voorhis T. J Chem Theory Comput 2006;2:765. (c) Wu Q, Van Voorhis T. J Phys Chem A 2006;110:9212. [PubMed: 16854035]
36. Bylaska, EJ.; de Jong, WA.; Govind, N.; Kowalski, K.; Straatsma, TP.; Valiev, M.; Wang, D.; Apra, E.; Windus, TL.; Hammond, J., et al. NWChem, A Computational Chemistry Package for Parallel Computers, Version 5.1. Pacific Northwest National Laboratory; Richland, Washington 99352-0999, USA: 2007. (b) Kendall RA, Aprà E, Bernholdt DE, Bylaska EJ, Dupuis M, Fann GI, Harrison RJ, Ju J, Nichols JA, Nieplocha J, et al. Computer Phys Comm 2000;128:260.
37. Becke AD. J Chem Phys 1993;98:1372.
38. (a) Harris J, Jones RO. J Phys F 1974;4:1170. (b) Gunnarsson O, Lundqvist BI. Phys Rev B 1976;13:4274. (c) Langreth DC, Perdew JP. 1977;15:2884. *ibid* (d) Harris J. Phys Rev A 1984;29:1648.
39. Nowak, MJ.; Lapinski, L.; Kwiatkowski, JS.; Leszczynski, J. Computational Chemistry: Reviews of Current Trends. Leszczynski, J., editor. Vol. 2. World Scientific; Singapore: 1997. p. 140-216.
40. In the present work, we do not deal with the issue of the problematic geometry optimization by pure DFT, treated in: Kurita N, Danilov VI, Anisimov VM. Chem Phys Lett 2005;404:164.
41. Dreuw A, Weisman JL, Head-Gordon M. J Chem Phys 2003;119:2943.
42. Liao M, Lu Y, Scheiner S. J Comput Chem 2003;24:623. [PubMed: 12632477]
43. Waller MP, Robertazzi A, Platts JA, Hibbs DE, Williams PA. J Comput Chem 2006;27:491. [PubMed: 16444702]
44. Robertazzi A, Platts JA. J Phys Chem A 2006;110:3992. [PubMed: 16539422]
45. Anderson JA, Tschumper GS. J Phys Chem A 2006;110:7268. [PubMed: 16737279]
46. Miura M, Aoki Y, Champagne B. J Chem Phys 2007;127:084103-1. [PubMed: 17764225]
47. Yu W, Liang L, Lin Z, Ling S, Haranczyk M, Gutowski M. J Comput Chem 2009;30:589. [PubMed: 18711717]

48. Csontos J, Palermo N, Murphy R, Lovas S. *J Comput Chem* 2008;29:1344. [PubMed: 18172837]
49. Takano Y, Taniguchi T, Isobe H, Kubo T, Morita Y, Yamamoto K, Nakasuji K, Takui T, Yamaguchi K. *J Am Chem Soc* 2002;124:11122. [PubMed: 12224960]
50. Kuba T, Woiczikowski PB, Cuniberti G, Elstner M. *J Phys Chem B* 2008;112:7937. [PubMed: 18543986]
51. Ivanova A, Shushkov P, Rösch N. *J Phys Chem A* 2008;112:7106. [PubMed: 18613656]
52. Clowney L, Jain SC, Srinivasan AR, Westbrook J, Olson WK, Berman HW. *J Am Chem Soc* 1996;118:509.
53. Gao YG, Robinson H, Wang AH. *Eur J Biochem* 1999;261:413. [PubMed: 10215851]
54. Varsano D, Di Felice R, Marques MAL, Rubio A. *J Phys Chem B* 2006;110:7129. [PubMed: 16599476]
55. Fuentes-Cabrera M, Zhao X, Kent PRC, Sumpter BG. *J Phys Chem B* 2007;111:9057. [PubMed: 17650925]
56. da Silva MW, Bierbrver RG, Wilds CJ, Noronha AM, Colvin OM, Miller PS, Gamcsik MP. *Bioorg Med Chem* 2005;13:4580–4587. [PubMed: 15953553]
57. Aymani J, Coll M, van der Marel GA, van Boom JH, Wang AHJ, Rich A. *Proc Natl Acad Sci USA* 1990;87:2526–2530. [PubMed: 2320572]
58. Warshel A. *J Phys Chem* 1982;86:2218–2224.
59. An ideal conformation is marked by the base step parameters used to generate it with nucleic acid builders. A real conformation is specified by the name of the pdb file where it is extracted from.
60. Troisi A, Orlandi G. *J Phys Chem B* 2002;106:2093.
61. Nakatani K, Dohno C, Saito I. *J Am Chem Soc* 1999;121:10854.
62. Jortner J, Bixon M, Voityuk AA, Rösch N. *J Phys Chem A* 2002;106:7599.
63. Bixon M, Jortner J. *Adv Chem Phys* 1999;106:35.
64. Troisi A, Nitzan A, Ratner MA. *J Chem Phys* 2003;119:5782.
65. Lewis FD, Liu J, Weigel W, Retting W, Kurnikov IV, Beratan DN. *Proc Natl Acad Sci USA* 2002;99:12536. [PubMed: 12228728]
66. Bixon M, Giese B, Wessely S, Langenbacher T, Michel-Beyerle ME, Jortner J. *Proc Natl Acad Sci USA* 1999;96:11713. [PubMed: 10518515]
67. Berlin YA, Burin AL, Ratner MA. *J Am Chem Soc* 2001;123:260. [PubMed: 11456512]
68. Bixon M, Jortner J. *J Am Chem Soc* 2001;123:12556. [PubMed: 11741420]
69. Conwell EM, Rakhmanova SV. *Proc Natl Acad Sci USA* 2000;97:4556. [PubMed: 10758150]
70. Liu CS, Hernandez R, Schuster GB. *J Am Chem Soc* 2004;126:2877. [PubMed: 14995205]
71. Conwell EM, Bloch SM, McLaughlin PM, Basko DM. *J Am Chem Soc* 2007;129:9175. [PubMed: 17585762]
72. Kuba T, Elstner M. *J Phys Chem B* 2009;113:5653. [PubMed: 19331336]
73. Senthikumar K, Grozema FC, Fonseca Guerra C, Bickelhaupt FM, Lewis FD, Berlin YA, Ratner MA, Siebbeles LDA. *J Am Chem Soc* 2005;127:14894. [PubMed: 16231945]
74. Voityuk AA, Rösch N. *J Chem Phys* 2002;117:5607.
75. Blancafort L, Voityuk AA. *J Phys Chem A* 2006;110:6426. [PubMed: 16706397]
76. Wetmore SD, Boyd RJ, Eriksson LA. *Chem Phys Lett* 2000;322:129.
77. (a) Becke AD. *J Chem Phys* 1997;107:8554. (b) Hamprecht FA, Cohen AJ, Tozer DJ, Handy NC. 1998;109:6264. *ibid*
78. (a) Johnson ER, DiLabio GA. *Chem Phys Lett* 2006;419:333. (b) Johnson ER, Becke AD. 2006;432:600. *ibid*
79. (a) Becke AD. *J Chem Phys* 1993;98:5648. (b) Lee C, Yang W, Parr RG. *Phys Rev B* 1988;37:785.
80. Adamo C, Barone V. *J Chem Phys* 1998;110:6158.
81. Löwdin PO. *J Chem Phys* 1950;18:365.
82. Niehaus TA, Di Carlo A, Frauenheim Th. *Org Electr* 2004;5:167.
83. Migliore A, Corni S, Di Felice R, Molinari E. *J Phys Chem B* 2007;111:3774. [PubMed: 17388538]

84. (a) Cave RJ, Newton MD. Chem Phys Lett 1996;249:15. (b) Cave RJ, Newton MD. J Chem Phys 1997;106:9213.
85. Koopmans T. Physica 1934;1:104.
86. Jensen, F. Introduction to computational chemistry. Vol. 2. John Wiley & Sons, Ltd; Chichester: 2007. p. 153-159.
87. Koch, W.; Holthausen, MC. A Chemist's Guide to Density Functional Theory. Vol. 2. Wiley-VCH Verlag GmbH; New York: 2000. p. 157
88. van Duijneveldt FB, van Duijneveldt-van de Rijdt JGCM, van Lenthe JH. Chem Rev 1994;94:1873.
89. Angeli C, Calzado CJ, Cimiraglia R, Evangelisti S, Maynau D. Mol Phys 2003;101:1937.
90. (a) Zerner MC, Loew GH, Kirchner RF, Mueller-Westerhoff UT. J Am Chem Soc 1980;102:589. (b) Ridley J, Zerner M. Theor Chim Acta 1973;32:111. Bacon, AD. MS Thesis. University of Ghelph; 1976.
91. Kuznetsov, AM.; Ulstrup, J. Electron Transfer in Chemistry and Biology. John Wiley & Sons; New York: 1999.
92. Voityuk AA, Rösch N. J Phys Chem B 2002;106:3013.
93. For example, according to the results of Ref. 74, after a roll of $+5^\circ$, the electronic coupling in the GC-GC dimer changes by 50%. By assuming that V_{IF} varies linearly with the roll (for small enough changes of this step parameter), approximately in the same way for different stacks, the change obtained for the xAT-xAT system can be reasonably expected.
94. Kurnikov IV, Tong GSM, Madrid M, Beratan DN. J Phys Chem B 2002;106:7.

Fig. 1. Idealized structures of the following dimers: (a) GC-GC ($^{GC}Q_i$) (b) xGC-xGC ($^{GC}Q_{xi}$), (c) AT-AT ($^{AT}Q_i$), and (d) xAT-xAT ($^{AT}Q_{xi}$). The natural DNA stacks are built with the base step parameters of the regular B-DNA. The stacks with expanded guanine (xG) and adenine (xA) use the average structural parameters from Ref. 24.

Fig. 2.

(a) $^{AT}Q_{1i}$ and $^{AT}Q_{2i}$ geometries of the ideal AT-AT stack, with the base step parameters of the regular B-DNA. Geometry optimizations on one A base and one AT pair give the structures represented in gray and blue, respectively. (b) Real B-DNA AT-AT stack from a covalently cross-linked duplex (geometry $^{AT}Q_{1r}$ in the main text). (c) Real AT-AT stacks from a nicked B-DNA duplex. The geometries $^{AT}Q_{2r}$ and $^{AT}Q_{3r}$ are depicted in red and blue, respectively.

Fig. 3.
A-A structures in Table 7: coordinates from Ref. 75 (blue), ${}^4Q_{2i}$ (green), and ${}^4Q_{3i}$ (red).

Fig. 4. HOMO-1 (gray), HOMO (cyan) and LUMO (pink) in the AT-AT geometries (a) $^{AT}Q_{1i}$, (b) $^{AT}Q_{1r}$, and (c) $^{AT}Q_{3r}$, computed at the BHH/6-311++g(3df, 3pd) level.

Fig. 5.
HOMO (cyan) and LUMO (pink) in the ideal xAT-xAT geometry, computed at the BHH/6-311 ++g(3df, 3pd) level.

Table 1

Upper Panel: Transfer Integral Values, in eV, in the Ideal GC-GC Stack from B-DNA ($V_{\text{IF}}(^{\text{GC}}Q_i)$) and in a Real GC-GC Stack from A-DNA ($V_{\text{IF}}(^{\text{GC}}Q_r)$), Calculated by Eq. 1. ^{a,b} Lower Panel: Selected Reference Values.

	Method	$^{\text{GC}}Q_i^a$	$^{\text{GC}}Q_r^b$
		V_{IF}	V_{IF}
Our work	6-311++g**	0.084	0.072
	6-311++g(3df,3pd)	0.075	0.058
	cc-pVTZ	0.085	0.067
	TZVP	0.088	0.075
Reference works	SCC-DFTB ⁵⁰	0.072 (0.101)	
	KTA, HF/6-31g* ³²	0.093	
	FCD, KTA, HF/6-31g* ⁵¹	0.063	0.058
	QM/MD, INDO/S ¹⁸	0.037 (ideal)	
		0.080 (rms)	

^aThe ideal GC-GC stack is built through the 3DNA software package,²⁷ by using experimental idealized atomic coordinates,⁵² namely with the step parameters of regular B-DNA: rise=3.38 Å, twist= 36°.

^bThe real A-DNA-like GC-GC stack is drawn from the PDB file 440d⁵³ and relaxed as described elsewhere.⁵⁴

Table 2

Vertical Excitation Energies (ΔE_v), Diabatic State Energy Differences (ΔE_{IF}), and Transfer Integrals (V_{IF}), in eV, Calculated by Eqs. 1–3, in the xGC-xGC Stack with Ideal Base Step Parameters, at the GCQ_{x1i} and GCQ_{x2i} Nuclear Coordinates.^a

Basis set	GCQ_{x1i}			GCQ_{x2i}		
	ΔE_v	ΔE_{IF}	V_{IF}	ΔE_{IF}	ΔE_{IF}	V_{IF}
6-31g*				0.589		0.243
6-311g*				0.592		0.219
6-311g**				0.594		0.219
6-311++g**	0.452	0.107	0.219	0.586		0.204
6-311++g(3df,3pd)	0.424	0.109	0.205	0.586		0.196
cc-pVTZ	0.445	0.113	0.215	0.590		0.202
TZVP	0.460	0.110	0.224	0.589		0.206

^aThe nuclear coordinates GCQ_{x1i} and GCQ_{x2i} are given in the Supporting Information. Their relation to the transition state coordinate is commented in the Discussion section.

Table 3

Upper Panel: Diabatic State Energy Differences (ΔE_{IF}) and Transfer Integrals (V_{IF}), in eV, Calculated by Eqs. 1–2, in Ideal (Left: ${}^{\text{AT}}Q_{Ii}$, ${}^{\text{AT}}Q_{2i}$, ${}^{\text{AT}}Q_{3i}$) and Real (Right: ${}^{\text{AT}}Q_{Ii}$, ${}^{\text{AT}}Q_{2i}$, ${}^{\text{AT}}Q_{3i}$) AT-AT Stacks. Lower Panel: Selected Reference Values of V_{IF} .

Method	${}^{\text{AT}}Q_{Ii}$ ^a		${}^{\text{AT}}Q_{2i}$		${}^{\text{AT}}Q_{3i}$		${}^{\text{AT}}Q_{Ii}$ ^b		${}^{\text{AT}}Q_{2i}$ ^c		${}^{\text{AT}}Q_{3i}$ ^c	
	ΔE_{IF}	V_{IF}	ΔE_{IF}	V_{IF}	ΔE_{IF}	V_{IF}	V_{IF}	V_{IF}	V_{IF}	V_{IF}	V_{IF}	V_{IF}
6-311++g**			0.140		0.302	0.009	0.008	0.005	0.054		0.100	
6-311++g(3df,3pd)	0.066	0.017	0.137	0.008	0.300	0.003	0.003	0.003	0.059	0.091		
cc-pVTZ			0.140	0.012	0.305	0.011	0.004	0.062	0.096			
TZVP			0.141	0.010	0.304	0.008	0.006	0.059	0.106			
SCC-DFTB ⁵⁰			V_{IF}									
KTA, HF/6-31g** ³²			0.019 (0.019) ^d									
QM/MD, INDO/S ¹⁸			0.026									
			0.063									

^a An AT-AT stack of regular B-DNA is built by means of the 3DNA package with B-DNA step parameters.²⁷ ${}^{\text{AT}}Q_{Ii}$, ${}^{\text{AT}}Q_{2i}$ and ${}^{\text{AT}}Q_{3i}$ are obtained by different atomic relaxations of this 3DNA-built dimer. Specifically, the ${}^{\text{AT}}Q_{Ii}$ geometry is obtained after hydrogen relaxation, while ${}^{\text{AT}}Q_{2i}$ and ${}^{\text{AT}}Q_{3i}$ are obtained by geometry optimization on one AT pair and one adenine, respectively.

^b ${}^{\text{AT}}Q_{Ii}$ is drawn from the PDB file 1XCI.⁵⁶

^c ${}^{\text{AT}}Q_{2i}$ and ${}^{\text{AT}}Q_{3i}$ are derived from the Nucleic Acid Database file BDL021.⁵⁷

^d The value in parentheses pertains to the quantity H_{IF} .

Table 4
Diabatic State Energy Differences (ΔE_{IF}) and Transfer Integrals (V_{IF}), in eV, Calculated by Eqs. 1–2, in Ideal (Left: ${}^{AT}Q_{x1i}$, ${}^{AT}Q_{x2i}$, ${}^{AT}Q_{x3i}$) and Real (Right: ${}^{AT}Q_{x1r}$) xAT-xAT Stacks.

Basis set	${}^{AT}Q_{x1i}$ ^a		${}^{AT}Q_{x2i}$		${}^{AT}Q_{x3i}$		${}^{AT}Q_{x1r}$	
	ΔE_{IF}	V_{IF}	ΔE_{IF}	V_{IF}	ΔE_{IF}	V_{IF}	ΔE_{IF}	V_{IF}
6-311++g**			0.213	0.052	0.218	0.045	0.058	
6-311++g(3df,3pd)			0.212	0.050	0.215	0.040	0.060	
cc-pVTZ	0.037	0.054	0.213	0.054	0.222	0.043	0.056	
TZVP			0.213	0.056	0.219	0.046	0.061	

^aThe ideal stack ${}^{AT}Q_{x1i}$ is built through the 3DNA software,²⁷ by exploiting the average structural parameters in Ref. 24. ${}^{AT}Q_{x2i}$ and ${}^{AT}Q_{x3i}$ are derived from different steps of a geometry optimization (see the main text).

^b ${}^{AT}Q_{x1r}$ is drawn from Ref. 23, as described by some of us elsewhere.²⁶

Table 5

Summary of the calculated transfer integrals (best estimates, in eV) from Tables 1–4.

	GC-GC	xGC-xGC	AT-AT	xAT-xAT
Ideal	0.075	0.205	0.008–0.017	0.040–0.054
Real	0.058 (A-DNA)	---	0.003 0.059–0.091	0.060

Table 6

Upper Panel: Vertical Excitation Energy (ΔE_v), Transfer Integral (V_{IF}), and Transfer Integral Using the CDFT Diabatic States (V_{IFCDFT}), in eV, in the Ideal G-G Nucleobase Stack,^a Calculated by our DFT Implementation of Eqs. 1–3. Lower Panel: Selected Reference Values of V_{IF} , from DFT, HF and post-HF calculations.

	Method	GQ_i ΔE_v	V_{IF}	V_{IFCDFT}
Our work with the BHH exchange-correlation functional	6-31g*	0.498	0.141	0.129
	6-31g* + BSSE	0.492	0.139	
	6-311g*	0.467	0.103	
	6-311g**	0.464	0.102	
	6-311++g**	0.441	0.075	
	6-311++g(3df,3pd)	0.426	0.065	
	cc-pVTZ	0.439	0.076	
	TZVP	0.446	0.076	0.069
	TZVP + BSSE	0.446	0.076	
Our work with other exchange-correlation functionals	6-31g*, PBE0			0.619
	6-31g*, B97-3			0.627
	6-31g*, B3LYP			0.726
	6-31g*, B97			0.751
Other DFT-based works	SCC-DFTB ⁵⁰		0.061 (0.087) ^b	
	FO-DFT, TZ2P in ADF ⁷³		0.053 (0.119)	
HF-based works	KTA, HF/6-31g* ³²		0.084	
	FCD, KTA, HF/6-31g* ⁷⁴		0.081	
	GMH, KTA, HF/6-31g* ⁷⁵	0.472	0.083	
Post-HF-based work	CASSCF(7,8), 6-31g* ⁷⁵	0.414	0.067	
	CASSCF(11,12), 6-31g*	0.370	0.049	
	CAS-PT2(11,12), 6-31g*	0.392	0.051	

^aOur DFT values use the BHH functional, unless otherwise specified. Nuclear coordinates are taken from Ref. 75.

^bThe values in parentheses pertain to the quantity H_{IF} .

Table 7

Upper Panel: Vertical Excitation Energy (ΔE_v) and Transfer Integral (V_{IF}), in eV, in Ideal A-A Nucleobase Stacks, Calculated by Eqs. 1–3.^a Lower Panel: Selected Reference Values.

Method	ΔE_v	${}^a Q_{1i}$ [${}^b Q_{ii}$] V_{IF}	${}^a Q_{2i}$ [${}^b Q_{2i}$] V_{IF}	${}^a Q_{3i}$ [${}^b Q_{3i}$] V_{IF}
6-311++g**	0.128	0.056		
6-311++g(3df,3pd)	0.118	0.050 [0.018]	0.032 [0.010]	0.010 [0.009]
cc-pVTZ	0.156	0.071		
TZVP	0.165	0.077		
TZVP, less accurate convergence thresholds	0.062	0.001		
SCC-DFTB ⁵⁰		0.008 (0.013) ^c		
FO-DFT, TZ2P in ADF ⁷³		0.004 (0.038) ^c		
KTA, HF/6-31g** ³²		0.030		
GMH, KTA, HF/6-31g** ⁷⁵	0.078	0.030		
CASSCF(7,8), 6-31g** ⁷⁵	0.144	0.066		
CASSCF(11,12), 6-31g**	0.047	0.002		
CAS-PT2(11,12), 6-31g**	0.097	0.004		

^a ${}^a Q_{1i}$ has the same nuclear coordinates as in Ref. 75, so that both our transfer integrals and our vertical excitation energies can be compared to HF and post-HF data. The nuclear coordinates of the heavy atoms in ${}^a Q_{2i}$ (${}^b Q_{3i}$) are taken from Ref. 75, the hydrogens of the second (first) adenine (in the direction 5' to 3') are relaxed.

^b ${}^b Q_{1i}$, ${}^a Q_{2i}$ and ${}^a Q_{3i}$ are an alternative set of ideal coordinates for A-A stacks, extracted directly from the AT-AT stacks ${}^a Q_{1i}$, ${}^a Q_{2i}$ and ${}^a Q_{3i}$ of Table 3, respectively, and not further relaxed. These two sets of coordinates differ only for the presence or absence of thymines, while they have identical adenines. Thus, they are suitable to evaluate the role of thymines in charge transfer processes.

^cThe values in round parentheses in the lower panel pertain to the quantity H_{IF} .