Excitation Energy Transport in DNA modelled by Multi-Chromophoric Field-Induced Surface Hopping

Matthias Wohlgemuth and Roland Mitric*

Institut für Physikalische und Theoretische Chemie, Julius-Maximilians-Universität Würzburg, Emil-Fischer-Str. 42, 97074 Würzburg, Germany.

E-mail: roland.mitric@uni-wuerzburg.de

Supplementary Information

Analysis of the Transition density matrix

We have analyzed the character of the excitation and calculated the delocalization length (DL) and charge transfer character (CT) using the decomposition scheme of the transition density matrix (TDM) adopted from Plasser et al.[1–3] based on the work of Tretiak and co-workers[4]. A schematic representation of in the following derived quantities is given in Fig. S1. Briefly, the TDM expressed in atomic orbitals is decomposed into logical fragments and the CT number Ω_{AB} for two fragments A and B is calculated according to:

$$\Omega_{AB} = \frac{1}{2} \sum_{a \in A} \sum_{b \in B} (\mathbf{DS})_{ab} (\mathbf{SD})_{ab} + D_{ab} (\mathbf{SDS})_{ab} \tag{1}$$

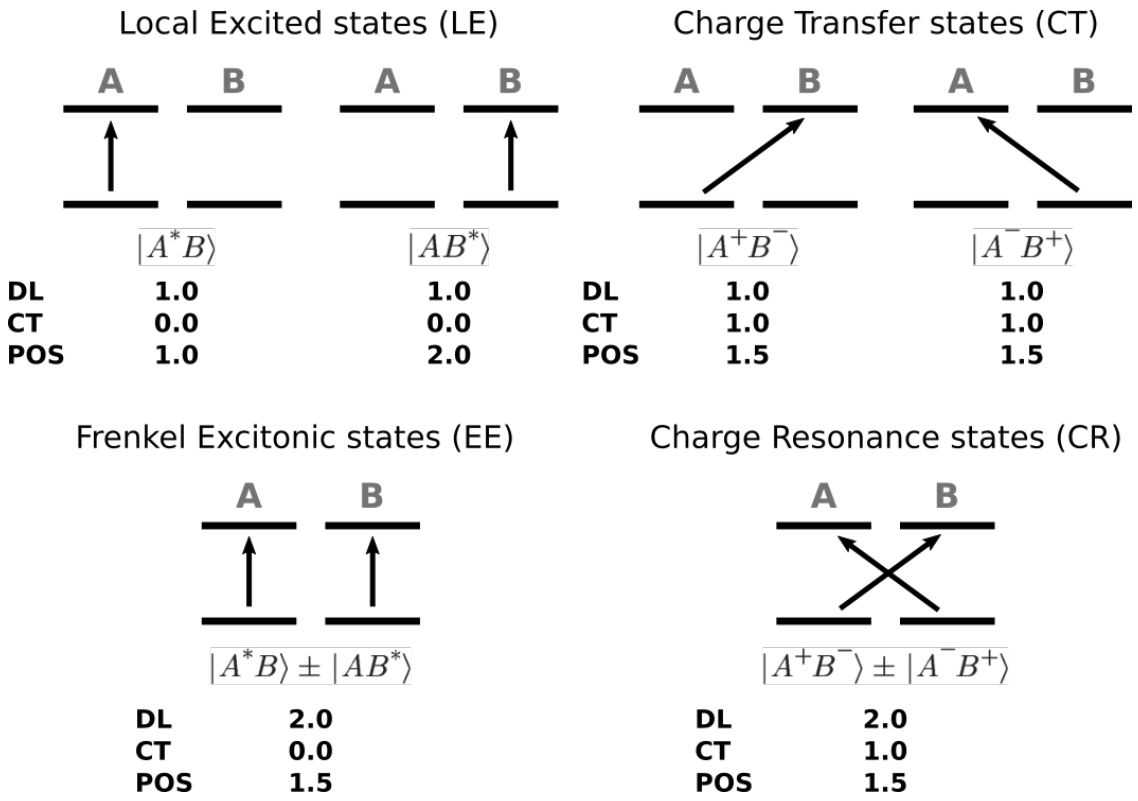

Figure S1: Schematic representation of derived quantities (DL, CT, POS) from transition density matrix analysis for a dimer A and B. The top row represents are localized electronic states being either local excitations (LE) in each monomer or charge transfer excitations (CT). The bottom row shows delocalized excitations being either excitonic excitations (EE) or charge resonance excitations (CR).

where D and S are transition density and overlap matrices (expressed in atomic orbitals), respectively, and the indices a and b are running over the atomic orbitals associated with the fragments. If not stated otherwise we consider the individual nucleobases as fragments. For details how to interpret these charge transfer numbers see Ref.[4–6]. These numbers give rise to an unique decomposition of the excited state into local and charge transfer contributions for the individual fragments. By normalization of the CT numbers by $\Omega = \sum_{AB} \Omega_{AB}$:

$$CT = \Omega^{-1} \sum_{A} \sum_{B \neq A} \Omega_{AB} \tag{2}$$

a the relative contribution of charge transfer character to an excitation is obtained. The CT values are ranging from 0, meaning no charge transfer character, to 1, corresponding to

pure charge transfer character. If not mentioned otherwise, we considered excitations with CT values above 0.3 as excitation with significant CT character. In order to obtain the total TDM for the partitioned QM system, the TDM of each sub-system is contracted with the eigenvectors obtained from diagonalizing the excitonic hamiltonian $\mathbf{H}_{MC}$. Note, that off-diagonal matrix elements representing transitions from one sub-system to another are thus zero and CT excitations between the sub-systems are not included. Therefore, the two charge transfer contributions in the S_4 and S_8 states present in the full QM system (Tab. S1) are absent in the partitioned system since the sub-system boundary is across the T9:A30 and A10:T31 base pairs.

The DL value, defined as the inverse partition ratio[7]:

$$DL = \Omega^2 \left[\sum_A \left(\sum_B \frac{\Omega_{AB} + \Omega_{BA}}{2} \right)^2 \right]^{-1}, \tag{3}$$

measures the distribution of the excitation over multiple fragments and thus ranges from 1 to the number of fragments, where A and B are the indices of the fragments. A DL value close to 1 represents a local transition where the excitation is located (mostly) at one fragment. At larger DL values the excitation is distributed over multiple nucleobases and thus represents either an exciton or a charge resonance state (if CT value is large as well).

Lastly, an average location of the excitation can be computed using[2]:

$$POS = (2\Omega)^{-1} \sum_A A \sum_B (\Omega_{AB}^2 + \Omega_{BA}^2) \tag{4}$$

which ranges from 1 to the number of fragments. The POS value essentially only makes sense if two fragments are considered, thus we only give this value in the main text to obtain a measure for the delocalization between the DNA strands.

Therefore, in order to allow assignment of an excitation to individual nucleobases, we follow the work of Tretiak and co-worker[4,8] defining a fractional transition density (FTD)

for the fragment A according to:

$$FTD = \frac{\sum_{a\in A}\sum_{b\in A} D_{ab}^2}{\sum_A \sum_{a\in A}\sum_{b\in A} D_{ab}^2} \tag{5}$$

where **D** is again the transition density matrix in the atomic basis. The FTD value, ranging from 0 to 1 for each fragment and summing to 1, provides the relative contribution of each fragment to the total transition density and thus allows easily the assignment of the transitions to the monomers. Due to the helical structure of the dsDNA we considered it challenging for the reader to realize the nature and location of excitations from volumetric transition density plots, thus we decided to use the FTD values to create a two dimensional surface plot. In order to visualize also the location of the created hole and the electron due to the excitation we modified the above approach by decomposing the TDM by singular value decomposition (SVD) in order to obtain natural transition orbitals[9] yielding small set of orbitals representing hole and electron, which are transformed in the atomic basis. From these orbitals we calculate the hole (**H**) and electron (**P**) density. Using Eq. (5) by inserting **H** and **P** instead of **D** we obtain two sets of FTD values for the location of hole and the electron, respectively.

The plots presented in the figures S2 and S3, showing the locations of excitations in the UV/Vis spectra discussed in the main text, are then obtained by using a schematic two-dimensional representation, assigning a specific position to each nucleobase. Then each FTD value is convolved with a two-dimensional Gaussian, where the height is determined by the FTD values, centered at the corresponding position of the nucleobase (we put a small offset, such that hole and electron FTD value do not cancel). The figures can be interpreted as follows: (i) if the hole and electron are located at the same nucleobase and only at one, the transition is a localized excitation (LE), (ii) if the hole and electron are on two or more nucleobases an exciplex is formed and (iii) and if electron and hole are on two different nucleobases the transition is of charge transfer character (CT).

Table S1: Summary and analysis of the 20 energetically lowest excited states obtained from the OM2/MM optimized geometry of $(dAdT)_{10} : (dAdT)_{10}$ dsDNA, where the 6 central base pairs ($(AT)_3 : (AT)_3$) are treated as a single QM-System and the remaining DNA, water and ions as MM part. [a]Corresponding assignment in the excitonically coupled 3 (AT) : (AT)-system (cf. Tab. S2). A double assignment indicates a combination of those states.

#	**E (eV)**	**f**	**CT**	**DL**	**contribution (character)**	**eq. exc. state[a]**
1	4.60	0.138	0.05	1.06	97% A29 ($\pi\pi^*$)	S_1
2	4.65	0.203	0.08	1.29	91% A31 ($\pi\pi^*$)	S_2
3	4.71	0.041	0.04	1.31	88% A11 ($\pi\pi^*$)	S_3
4	4.74	0.059	0.30	2.00	52% A09 ($\pi\pi^*$); 27% A09→T10 (CT)	S_5
5	4.75	0.113	0.02	1.08	94% A13 ($\pi\pi^*$)	S_4
6	4.79	0.141	0.04	1.77	90% A33 ($\pi\pi^*$)	S_6
7	4.82	0.010	0.06	1.08	96% A29 ($n\pi^*$)	S_7
8	4.86	0.006	0.19	1.55	72% A09 ($n\pi^*$); 16% A09→T10 (CT)	S_9
9	4.87	0.005	0.04	1.44	83% A29 ($n\pi^*$); 12% A09 ($\pi\pi^*$)	S_8
10	4.89	0.004	0.11	1.97	71% A11 ($\pi\pi^*$); 11% T10 ($n\pi^*$)	$S_{10} + S_{11}$
11	4.90	0.001	0.02	1.09	96% A31 ($n\pi^*$)	S_{13}
12	4.91	0.004	0.13	1.59	88% A33 ($\pi\pi^*$)	S_{12}
13	4.92	0.011	0.11	2.10	58% A09 ($n\pi^*$); 19% A31 ($\pi\pi^*$)	$S_{14} + S_{15}$
14	4.93	0.010	0.11	2.75	47% A13 ($\pi\pi^*$); 26% A31 ($\pi\pi^*$)	$S_{14} + S_{15}$
15	4.94	0.017	0.12	1.91	68% A13 ($n\pi^*$); 16% A31 ($\pi\pi^*$)	$S_{14} + S_{15}$
16	4.95	0.006	0.09	2.51	72% A13 ($n\pi^*$)	$S_{14} + S_{15}$
17	4.96	0.070	0.06	3.51	63% A33 ($\pi\pi^*$); 10% T32 ($\pi\pi^*$)	S_{17}
18	4.99	0.166	0.24	3.39	41% T10 ($\pi\pi^*$); 21% T30 ($\pi\pi^*$)	S_{18}
19	5.03	0.435	0.02	1.90	51% T08 ($\pi\pi^*$); 38% T32 ($\pi\pi^*$)	S_{19}
20	5.04	0.144	0.57	2.87	41% A11→T30 (CT); 24% A11→A31 (CT); 17% T12 ($\pi\pi^*$)	-

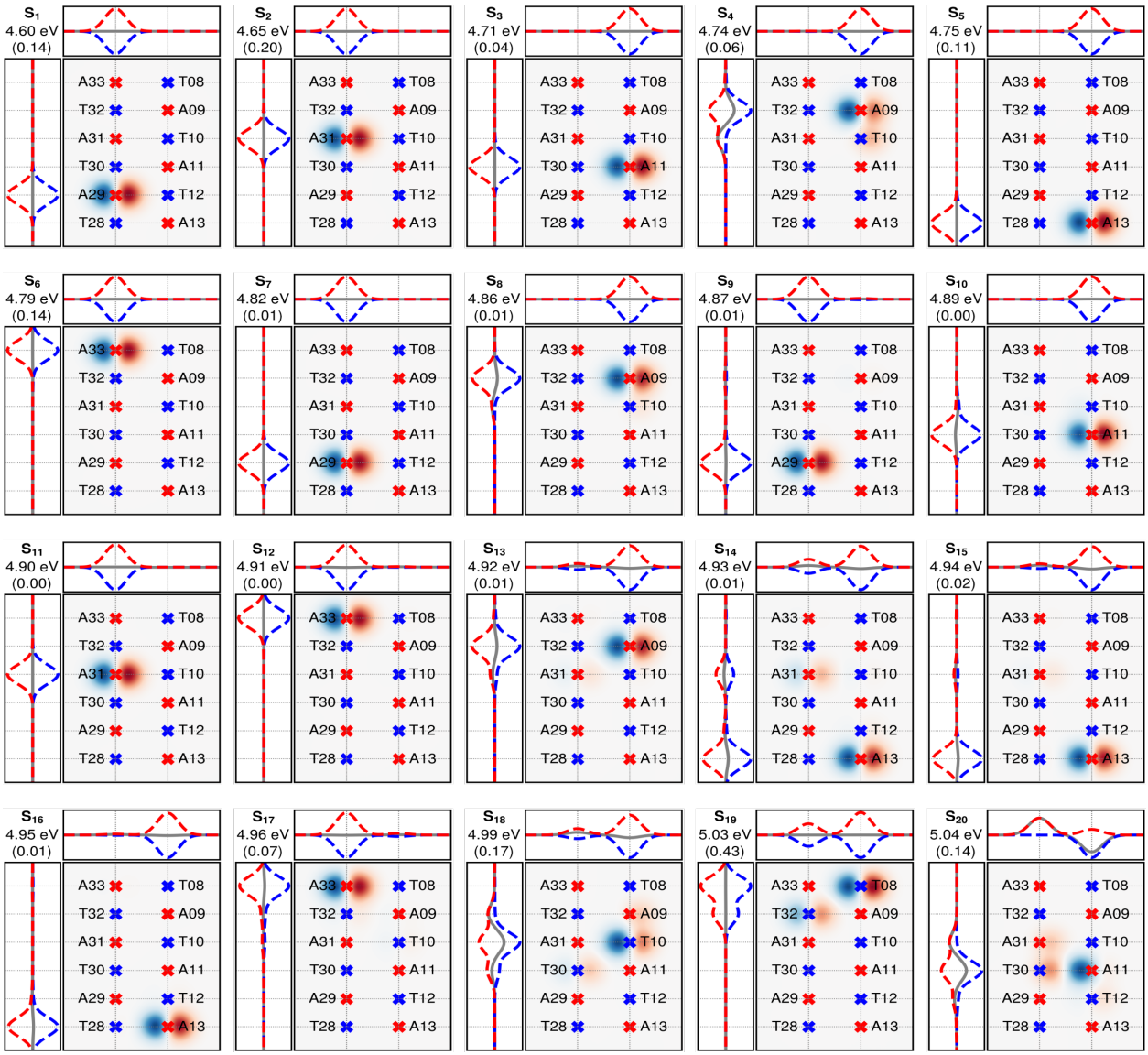

Figure S2: Graphical analysis in terms of the fractional transition density (FTD) of the 20 energetically lowest excited states obtained from the OM2/MM optimized geometry of $(\mathrm{dAdT})_{10} : (\mathrm{dAdT})_{10}$ dsDNA, where the QM region consists of the 6 central base pairs $((\mathrm{AT})_3 : (\mathrm{AT})_3)$. The textured area indicate the location (nucleobase) of the hole (blue) and electron (red) where the excitation is located. The left and top curves are the projection of these densities along the pair- and strand-axis, respectively.

Table S2: Summary and analysis of the 20 energetically lowest excited states obtained from the OM2/MM optimized geometry of $(dAdT)_{10} : (dAdT)_{10}$ dsDNA, where the QM-System consists of 3 sub-systems containing two base pairs each (3 (AT) : (AT)). The sub-systems are then coupled using the transition dipole approximation (TDA). [a] Excitonic wavefunction obtained from diagonalization of the excitonic hamiltonian.

#	**E (eV)**	**f**	**CT**	**DL**	**contribution (character)**	**wavefunction**[a]
1	4.60	0.163	0.03	1.03	98% A29 ($\pi\pi^*$)	1.0 $\|0\ 0\ 1\rangle$
2	4.68	0.255	0.11	1.31	91% A31 ($\pi\pi^*$)	1.0 $\|0\ 1\ 0\rangle$
3	4.72	0.095	0.05	1.55	77% A11 ($\pi\pi^*$); 11% A13 ($\pi\pi^*$)	0.9 $\|0\ 2\ 0\rangle$ − 0.3 $\|0\ 0\ 2\rangle$
4	4.75	0.049	0.01	1.06	95% A13 ($\pi\pi^*$)	1.0 $\|0\ 0\ 2\rangle$
5	4.77	0.033	0.02	1.49	72% A09 ($\pi\pi^*$); 14% A33 ($\pi\pi^*$)	1.0 $\|1\ 0\ 0\rangle$
6	4.80	0.198	0.03	1.53	75% A33 ($\pi\pi^*$); 20% A09 ($\pi\pi^*$)	1.0 $\|2\ 0\ 0\rangle$
7	4.81	0.011	0.04	1.02	99% A29 ($n\pi^*$)	1.0 $\|0\ 0\ 3\rangle$
8	4.86	0.028	0.03	1.17	88% A29 ($\pi\pi^*$); 11% T28 ($\pi\pi^*$)	1.0 $\|0\ 0\ 4\rangle$
9	4.88	0.005	0.07	1.04	99% A09 ($n\pi^*$)	1.0 $\|3\ 0\ 0\rangle$
10	4.89	0.010	0.05	2.28	44% A11 ($\pi\pi^*$); 40% A13 ($n\pi^*$); 13% T10 ($\pi\pi^*$)	0.8 $\|0\ 3\ 0\rangle$ + 0.6 $\|0\ 0\ 5\rangle$
11	4.89	0.018	0.04	2.03	60% A13 ($n\pi^*$); 30% A11 ($\pi\pi^*$)	− 0.6 $\|0\ 3\ 0\rangle$ + 0.8 $\|0\ 0\ 5\rangle$
12	4.90	0.004	0.14	1.64	87% A33 ($\pi\pi^*$)	1.0 $\|4\ 0\ 0\rangle$
13	4.91	0.001	0.02	1.02	99% A31 ($n\pi^*$)	1.0 $\|0\ 4\ 0\rangle$
14	4.93	0.006	0.02	1.20	84% A13 ($n\pi^*$); 16% T12 ($n\pi^*$)	1.0 $\|0\ 0\ 6\rangle$
15	4.93	0.006	0.06	1.63	61% A09 ($n\pi^*$); 18% A33 ($n\pi^*$)	1.0 $\|5\ 0\ 0\rangle$
16	4.94	0.079	0.08	1.72	77% A31 ($\pi\pi^*$); 11% A11 ($\pi\pi^*$)	1.0 $\|0\ 5\ 0\rangle$
17	4.96	0.008	0.06	2.02	67% A33 ($\pi\pi^*$); 18% A09 ($n\pi^*$); 12% T32 ($n\pi^*$)	1.0 $\|6\ 0\ 0\rangle$
18	4.97	0.450	0.05	2.46	58% T10 ($\pi\pi^*$); 23% T30 ($\pi\pi^*$); 13% A11 ($n\pi^*$)	1.0 $\|0\ 6\ 0\rangle$
19	5.01	0.490	0.02	1.98	50% T08 ($\pi\pi^*$); 36% T32 ($\pi\pi^*$)	1.0 $\|7\ 0\ 0\rangle$
20	5.03	0.358	0.04	1.81	61% T28 ($\pi\pi^*$); 28% T12 ($\pi\pi^*$)	1.0 $\|0\ 0\ 7\rangle$

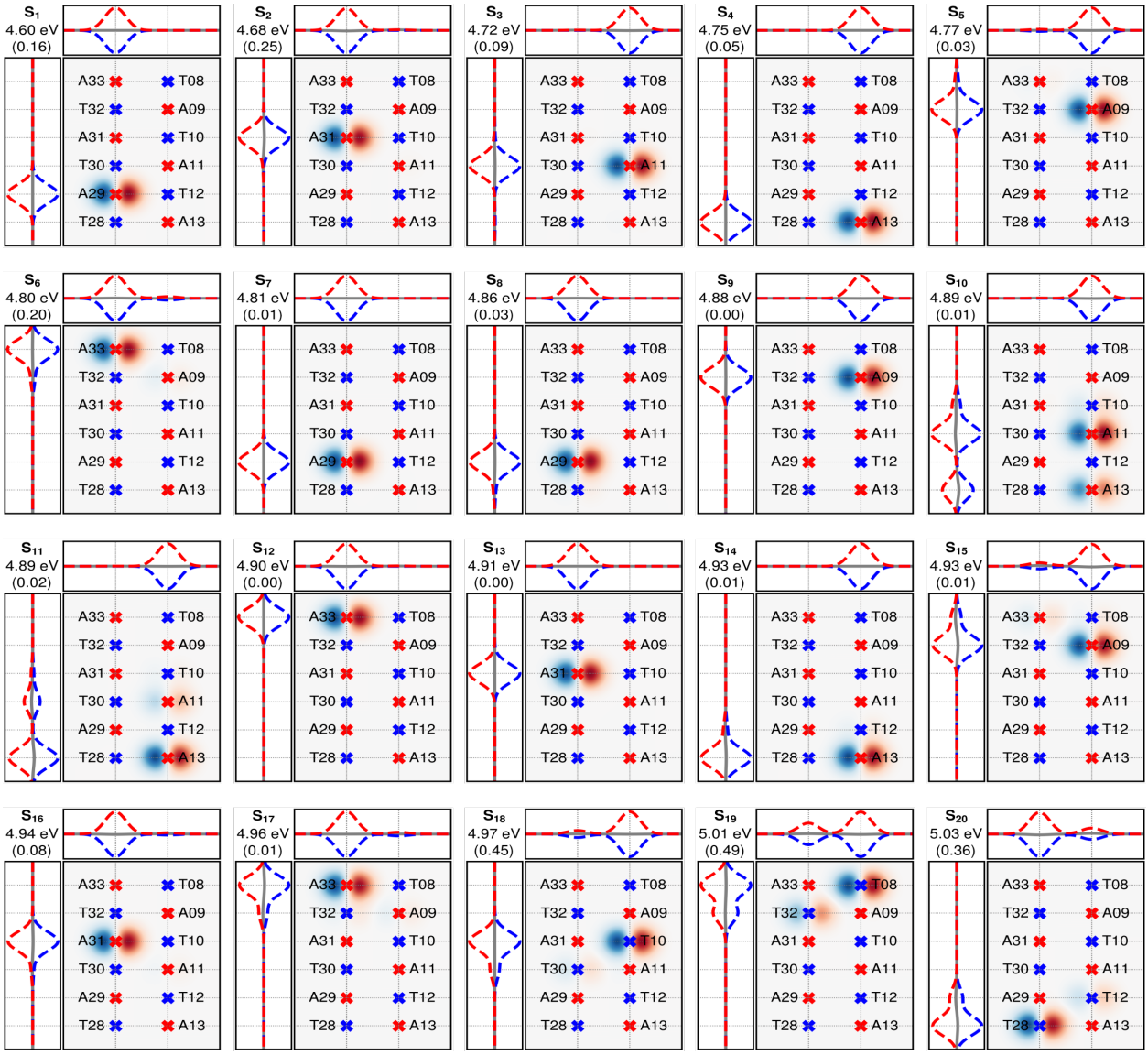

Figure S3: Graphical analysis in terms of the fractional transition density (FTD) of the 20 energetically lowest excited states obtained from the OM2/MM optimized geometry of $(dAdT)_{10} : (dAdT)_{10}$ dsDNA, where the QM region consists of 3 sub-systems containing two base pairs each (3(AT) : (AT)). The sub-systems are coupled using the transition dipole approximation (TDA). The textured area indicates the location (nucleobase) of the hole (blue) and electron (red) where the excitation is located. The left and top curves are the projection of these densities along the pair- and strand-axis, respectively.

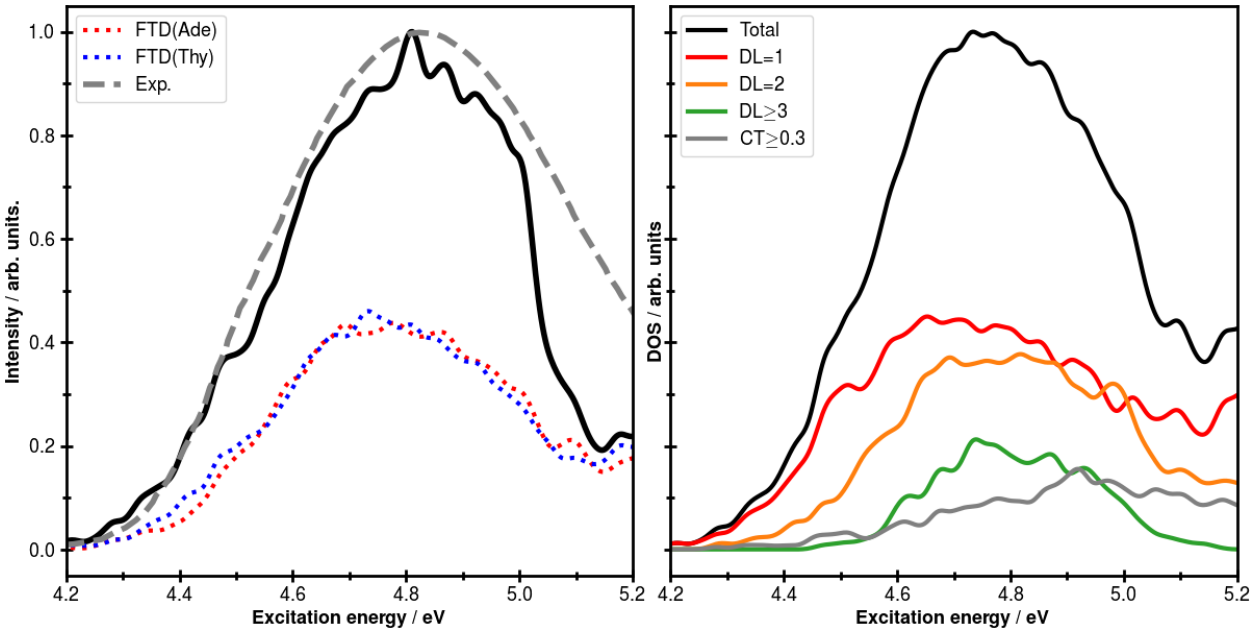

Figure S4: First absorption band of the UV/Vis spectrum averaged over geometries obtained form NVT-MD simulations of the $[(\mathrm{AT}):(\mathrm{AT})]_3$ trimer embedded in $(\mathrm{dAdT})_{10}:(\mathrm{dAdT})_{10}$ dsDNA solvated by water. Left panel: Oscillator strengths convoluted by a Lorentzian (black line, width 0.015 eV) and the contributions of excitations in Adenine (red) and Thymine (blue) to the transition density. The gray dotted lines indicate the absorption spectrum of $(\mathrm{dAdT})_{10}:(\mathrm{dAdT})_{10}$ in aqueous buffered solutions reproduced from Fig. 12 of Ref.[10]. Right panel: Total density of states (black) and decomposition into localized states (blue, DL = 1), delocalized states where the excitation is delocalized over two (DL = 2) or more (DL $\geq$ 3) bases, and states with charge transfer contributions (CT $\geq$ 0.3).

Fitting Parameters

The parameters of Tab. S3 are obtained from by fitting the following model functions to the populations of Fig. 3 of the main text:

$$P_1(t) = B\exp\left(-\frac{t}{\tau_2}\right) + Y_0 \tag{6}$$

$$P_2(t) = B\left[C\exp\left(-\frac{t}{\tau_2}\right) + (1-C)\exp\left(-\frac{t}{\tau_3}\right)\right] + Y_0 \tag{7}$$

$$P_{1+2}(t) = A\exp\left(-\frac{t}{\tau_1}\right) + B\left[(1-C)\exp\left(-\frac{t}{\tau_2}\right) + C\exp\left(-\frac{t}{\tau_3}\right)\right] + Y_0 \tag{8}$$

Table S3: Parameters obtained from fitting the graphs in Fig. 3 of the main text. Parameters marked as "-" have not been fitted. [a]Model functions used to fit the curves are given by Eqs. (6), (7) and (8).[b]Average life time obtained from the weighted sum of the individual components.

Curve	**Model**	**A**	**B**	**C**	$\mathbf{Y}_0$	τ_1 / **fs**	τ_2 / **ps**	τ_3 / **ps**	$\langle\tau\rangle$ / **ps**[b]	$\mathbf{R}^2$
1-S_0	2	-	0.66	0.53	-	-	0.70	74.75	35.5	0.98
S_L	1+2	-0.85	0.31	0.43	-	48	0.13	2.06	0.89	0.92
S_E	1	-	0.31	-	0.11	-	0.02	-	0.02	0.15
S_1	1+2	-0.84	0.21	0.33	-	68	0.18	1.95	0.88	0.88

Fractional transition density analysis

In order to analyze the location of the excited states in time-dependent terms we calculated average FTD values along the trajectories, which is shown in Fig. S5 (again normalized to the number of trajectories in the excited state at a given time). We have used two different patterns for selecting the fragments: (i) using the nucleobases adenine (red curve) and thymine (blue curve) and (ii) using the two strands of the dsDNA (green and orange curve). Since we are dealing with a heteropolymorphic duplex and both strands are therefore built from the same (but complementary) sequence of alternating adenines and thymines, both strands are excited equally by the laser pulse. This symmetry is not broken along the simulations and both strands keep on average equally excited, while on the single trajectory level the location of the excitation is frequently oscillating from one strand to the other. This is also consistent with the results from Fig. 4B of the main text, since either the excitation is localized at one strand, or at later times delocalized over both strands. Both of these cases result in a FTD of 0.5 for each of the strands, since in the localized case about half of the trajectories exhibit excitation on strand A and half on strand B, while in the delocalized case the excitation is equally distributed over strands A and B.

In agreement with experimental findings of Ref. [11] adenine and thymine contribute equally to the excitation at $t = 0$. However, in the following few 200 fs the population is transferred from thymine to adenine until a steady excitation level of thymine at ∼20% and consequently of adenine at ∼80% is reached (cf. Fig. S5 of the main text). This finding is consistent with

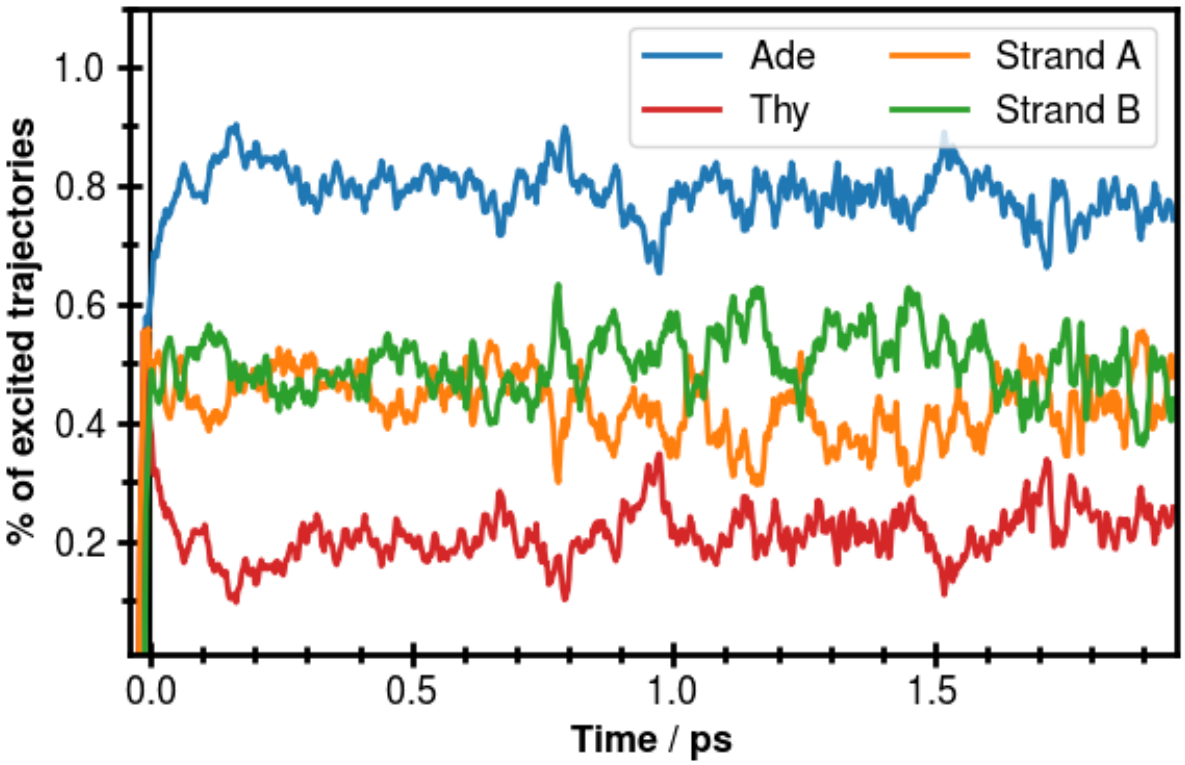

Figure S5: Averaged fractional transition density (FTD) for all adenines (red, Ade) and all thymines (blue, Thy), or strand A (orange) and strand B (green), respectively, showing equal amount of excitation in both strands and a dominant excitation in adenine.

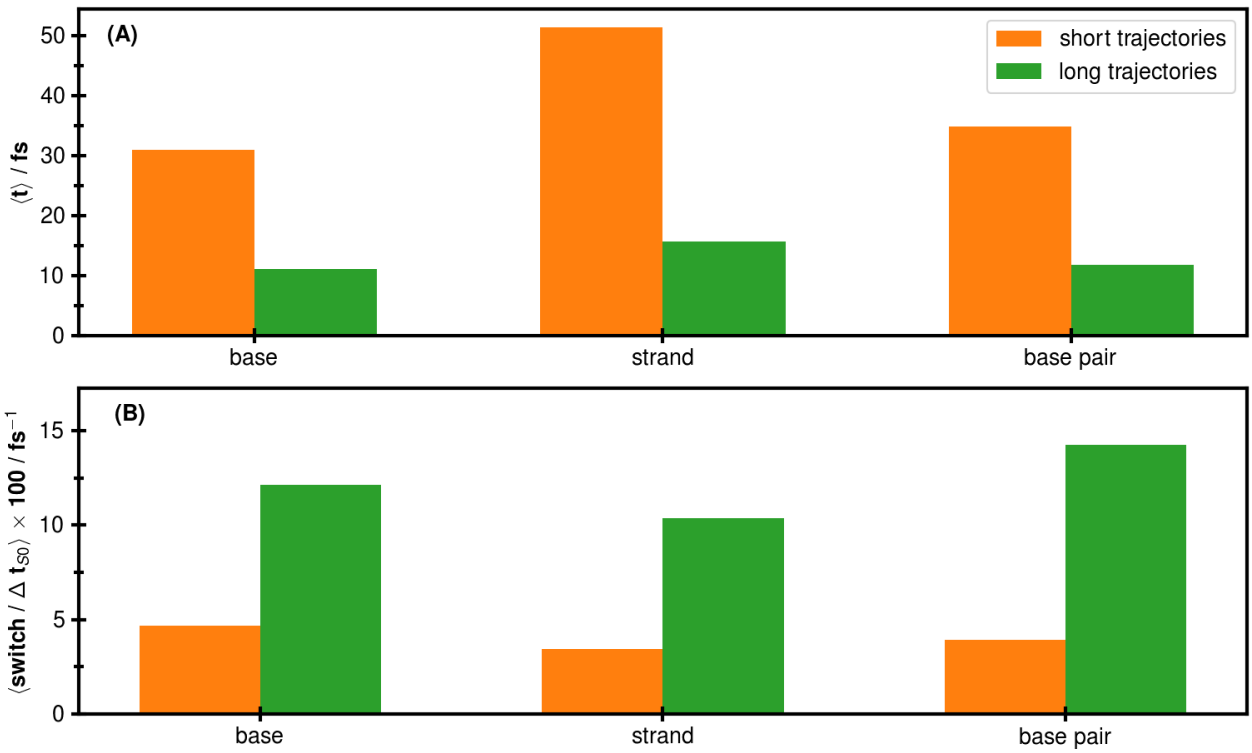

Figure S6: Statistical analysis of FTD, where only the maximum FTD value is considered for averaging. The trajectories are divided into short trajectories (orange), which decay to the ground-state within the first 800 fs, and long trajectories (green) which reside for more than 800 fs in the excited state. **(A)** Average duration an excitation resides at a time dominantly (max FTD > 0.8) on one nucleobase (left), within a base pair (middle) and within one of the strands (right). **(B)** Averaged rate for changing the location of the excitation with respect to a single base (left), horizontally between the strands (middle) and vertically along the stacked base pairs (right).

the short time constant found for the S_L population in Fig. 3 of the main text, which is identical to the experimentally determined lifetime of dAMP[12]. Together with the findings that (i) the DL value (cf. Fig. 4A of the main text) is increasing with time towards the delocalization of the excitation and (ii) from the POS value (cf. Fig. 4B of the main text) it becomes evident that the amount of delocalization between the strands increases, it can be concluded that the energy transport is dominantly conducted through adenine.

In Fig. S6A we show the average time an excitation resides at a single nucleobase (left), within a strand or within a pair of nucleobases. We divided the trajectories in 2 classes, as described in the main text. One group is classified as short and decays within 800 fs to the

ground state. The second group is classified as long and stays more than 800 fs in the excited state. Excitations in short trajectories are in average for 31 fs at a time located on a single nucleobase before switching to another, which is two times as longer than for the long living trajectories (11 fs). The same trend is found for the excitation being located within a base pair (35 fs vs. 11 fs) and a single strand (51 fs vs. 16 fs). The large difference in the latter case indicates that in the early stages of the excited state dynamics the excitation energy is dominantly transported within the strand, while for longer living trajectories inter-strand excitation energy transport is becoming more important (cf. Figs. 4B and 5 of the main text).

The rates of changing the location of an excitation from one base to another are shown in Fig. S6B (left), which support the previous findings. The nucleobase on which the excitation is dominantly localized, is changed in average 12 times per 100 fs in long trajectories, compared to 5 switches per 100 fs for short trajectories (Fig. S6B). The same trend holds for switching from one strand to the other (0.10 fs^{-1} vs. 0.03 fs^{-1}) and transport of the excitation within the stack (0.14 fs^{-1} vs. 0.04 fs^{-1}).

References

(1) Nogueira, J. J.; Plasser, F.; González, L. *Chem. Sci.* **2017**, *8*, 5682–5691.

(2) Plasser, F.; Aquino, A. J.; Hase, W. L.; Lischka, H. *J. Phys. Chem. A* **2012**, *116*, 11151–11160.

(3) Plasser, F.; Lischka, H. *J. Chem. Theory Comput.* **2012**, *8*, 2777–2789.

(4) Tretiak, S.; Mukamel, S. *Chem. Rev.* **2002**, *102*, 3171–3212.

(5) Luzanov, A.; Zhikol, O. *Int. J. Quantum Chem.* **2010**, *110*, 902–924.

(6) Luzanov, A.; Prezhdo, O. *Int. J. Quantum Chem.* **2005**, *102*, 582–601.

(7) Bouvier, B.; Gustavsson, T.; Markovitsi, D.; Millié, P. *Chem. Phys.* **2002**, *275*, 75–92.

(8) Fernandez-Alberti, S.; Makhov, D. V.; Tretiak, S.; Shalashilin, D. V. *Phys. Chem. Chem. Phys.* **2016**, *18*, 10028–10040.

(9) Martin, R. L. *J. Chem. Phys.* **2003**, *118*, 4775–4777.

(10) Bouvier, B.; Dognon, J.-P.; Lavery, R.; Markovitsi, D.; Millié, P.; Onidas, D.; Zakrzewska, K. *J. Phys. Chem. B* **2003**, *107*, 13512–13522.

(11) Kwok, W.-M.; Ma, C.; Phillips, D. L. *J. Phys. Chem. B* **2009**, *113*, 11527–11534.

(12) Markovitsi, D.; Sharonov, A.; Onidas, D.; Gustavsson, T. *ChemPhysChem* **2003**, *4*, 303–305.