THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

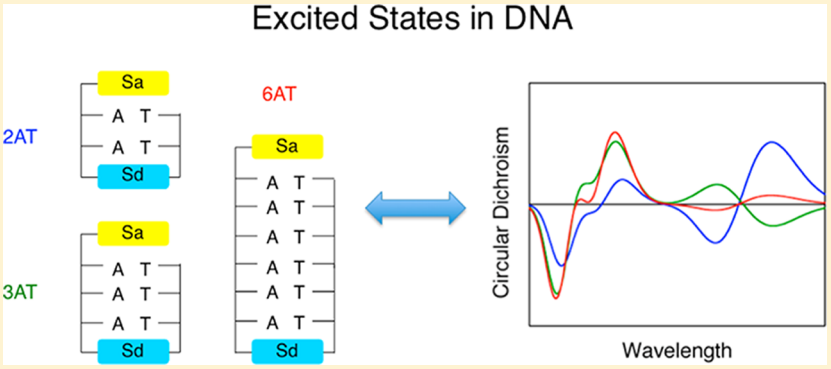
pubs.acs.org/JPCB

Effect of Structural Dynamics and Base Pair Sequence on the Nature of Excited States in DNA Hairpins

Sameer Patwardhan,[†,‡] Stefano Tonzani,[‡] Frederick D. Lewis,[‡] Laurens D. A. Siebbeles,[†] George C. Schatz,[‡] and Ferdinand C. Grozema*[,†]

[†]Opto-Electronic Materials Section, Department of Chemical Engineering, Delft University of Technology, Julianalaan 136, 2628 BL, Delft, The Netherlands

[‡]Department of Chemistry, Center for Nanofabrication and Molecular Self-Assembly, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208-3113, United States



ABSTRACT: In this article, a theoretical study of the electronic and spectroscopic properties of well-defined DNA hairpins is presented. The excited states in the hairpins are described in terms of an exciton Hamiltonan model, and the structural dynamics of the DNA model systems is explicitly taken into account by molecular dynamics simulations. The results show that the model reproduces the experimentally observed absorption and circular dichroism spectra accurately in most cases. It is shown that structural disorder leads to excited states that are largely localized on a single base pair, even for regular DNA sequences consisting only of AT base pairs. Variations in the base pair sequence have a significant effect on the appearance of the spectra but also on the degree of delocalization of the excited state.

1. INTRODUCTION

The absorption of UV light by DNA is responsible for a range of health effects, including photoaging and skin cancer.[1−3] The initial species that is the starting point for most DNA photochemistry, including often-occurring pyrimidine photodimerization, is the singlet excited state that is formed directly upon absorption of light. The nature of this initial excited state (i.e., whether it is localized on a single base or delocalized) has been a subject of considerable speculation.[4] Even for double-stranded model systems that consist of only adenine-thymine (AT) base pairs, it is not clear whether it is localized on a single base or delocalized over several bases.[5−8] Although the ground-state absorption spectrum of DNA closely resembles the sum of the spectra of the constituent bases, there is also clear evidence from circular dichroism spectra that there is considerable electronic coupling between neighboring bases.[9]

The emission spectrum of DNA model systems shows the dependence on the excitation wavelength and base-pair sequence, indicating the complex dynamics of excited-state relaxation.[10,11] The red-shifted fluorescence band, compared to individual bases, has been attributed to emission from a delocalized excited state. In contrast to this picture, a recent study shows the fluorescence spectrum in natural DNA that is very similar to that of a stoichiometric mixture of monomeric bases, arising from bright $\pi\pi^*$ states. However, the DNA fluorescence decays span five decades of time, with 98% of the photons being emitted at times longer than 10 ps.[12] This suggests the involvement of dark excimer-like or charge-separated states during the evolution of the initial state that is formed directly on excitation. The nonradiative deactivation pathway involving interbase proton transfer between Watson/Crick pairs has been suggested in localizing the excitation and then quenching it through intersystem crossing and charge transfer.[13]

Apart from the fundamental interest in the nature of excited states in DNA, there is also direct biological relevance. A major route to UV-induced damage is the formation of thymine (T-T) dimers. It has been shown that the formation of such T-T dimers depends considerably on the base pairs that surround the two neighboring thymine bases in the DNA stack.[14,15] Part

Received: July 19, 2012
Revised: August 13, 2012
Published: August 13, 2012

   dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

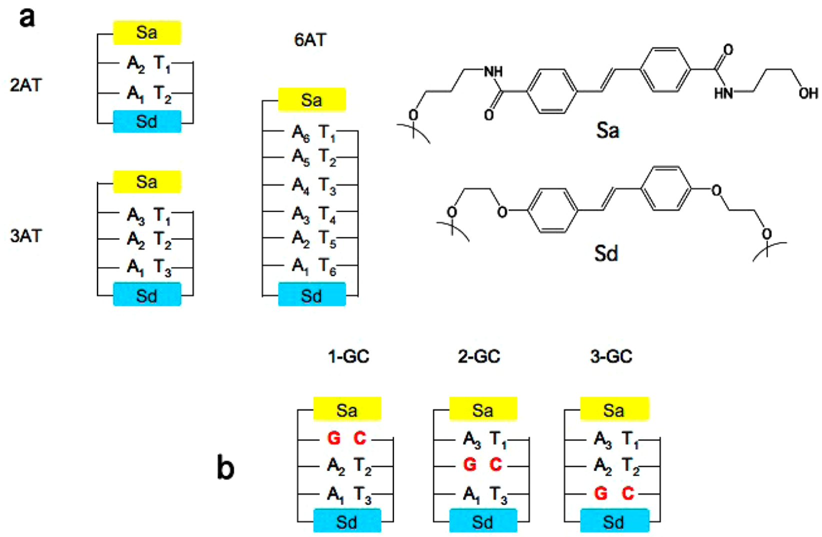


Figure 1. Schematic representation of Sd-linked hairpins containing a capping stilbene (Sa).

of the context dependence can be explained by the conformational properties because the precise orientation of the two thymines is important in determining the dimerization yield.[16,17] There are also indications that the neighboring bases considerably affect the electronic structure of the excited state that eventually results in thymine dimers.[18] Therefore, it is of considerable interest to study the electronic structure of excited states in DNA and their dependence on the base pair sequence.

In this article, a theoretical study of excited states in well-defined DNA model systems is described. The model systems consist of DNA hairpins containing up to six base pairs as schematically shown in Figure 1. The structure of these hairpins closely resembles that of regular B-type DNA.[19,20] This makes them good model systems for actual DNA while keeping the systems small enough to make detailed simulations possible.[21,22] The linker between the two strands, in this case, an Sd stilbene moiety (Figure 1) keeps the two strands together. This results in B-DNA-like structures, even for the shortest hairpins that would not be stable otherwise. At the other end, a "capping" stilbene Sa is appended and has been shown to stack onto the outer base pairs in a well-defined manner.[23] For these hairpins, the absorption spectra and circular dichroism spectra have been calculated while explicitly taking the structural fluctuations into account. It is shown that the base pair sequence significantly affects the excited-state electronic structure, which may be of importance for the formation of thymine dimers in DNA. Additionally, it is argued that a combined theoretical and experimental study of circular dichroism spectra of DNA model systems is a promising approach to studying the structural dynamics in these systems.

2. COMPUTATIONAL METHODOLOGY

In this work, exciton theory[10,13,24−28] has been used to investigate the excited states and optical properties of DNA hairpins. In Section 2.1, the exciton Hamiltonian is described and equations are provided for calculating the optical properties. The evaluation of diagonal and off-diagonal elements of the exciton matrix is discussed in Section 2.2. The underlying assumptions in the current methodology and their implications have been discussed in previous publications.[10,13,27]

2.1. Exciton Theory. The excited-state wave functions (Ψ^{exciton}) of DNA hairpins can be expressed, within the Frenkel exciton model, in terms of the excited states of constituent nucleobases and stilbenes.

$$\Psi^{\text{exciton}} = \sum_{m}^{N_{\text{mol}}} \sum_{i} C_{m,i} \Psi_m^i \tag{1}$$

Here, the localized wave function Ψ_m^i is the product of the molecular electronic wave functions on the individual nucleobases, where nucleobase m is in the ith excited state and all other nucleobases are in the ground state; $C_{m,i}$ is the corresponding coefficient. N_{mol} is the total number of nucleobases in the hairpin. The exciton Hamiltonian ($\hat{H}$) takes the form

$$\hat{H} = \sum_{m=1}^{N_{\text{mol}}} \sum_i \varepsilon_m^i \hat{B}_m^{i\dagger} \hat{B}_m^i + \frac{1}{2} \sum_{\substack{m,l=1 \\ m \neq l}}^{N_{\text{mol}}} \sum_{i,j} J_{ml}^{ij} (\hat{B}_l^{j\dagger} \hat{B}_m^i + \hat{B}_m^{i\dagger} \hat{B}_l^j) \tag{2}$$

where $\hat{B}_m^{i\dagger}$ and $\hat{B}_m^i$ are the creation and the annihilation operators of the ith excited state on molecule m; ε_m^i is the site energy; J_{ml}^{ij} is the exciton coupling between the ith excited state on molecule m and the jth excited state on molecule l. In a matrix representation, ε_m^i and J_{ml}^{ij} would be the diagonal and the off-diagonal matrix elements of the Hamiltonian matrix, respectively. The diagonal (site energy) terms consist of several contributions:

$$\varepsilon_m^i = \Delta\varepsilon_m^i + J_{mm}^i + \delta_m^i \tag{3}$$

The first term on the right ($\Delta\varepsilon_m^i$) represents the ith excitation energy of molecule m in vacuum. The second term (J_{mm}^i) is the displacement energy that accounts for the change in monomer excitation energy due to the presence of other molecules; see eq 5. The term δ_m^i takes into account the random fluctuations in excitation energy that are not directly caused by interactions between the nucleobases, for instance, molecular distortions and energy fluctuations due to the presence of solvent molecules and ions in the surrounding medium. It is sampled from a Gaussian distribution with standard deviation σ_δ.

The exciton coupling J^{ij}_{ml} and displacement energy J^{i}_{mm} can be written in terms of molecular ground state (ψ^{g}_{k}) and excited-state (ψ^{j}_{k}) wave functions and the Coulomb operator.

$$J^{ij}_{ml} = \langle \psi^{i}_{m}\psi^{g}_{l}|\hat{V}_{ml}|\psi^{g}_{m}\psi^{j}_{l}\rangle \tag{4}$$

$$J^{i}_{mm} = \sum_{l\neq m}^{N_{mol}} (\langle \psi^{i}_{m}\psi^{g}_{l}|\hat{V}_{ml}|\psi^{i}_{m}\psi^{g}_{l}\rangle - \langle \psi^{g}_{m}\psi^{g}_{l}|\hat{V}_{ml}|\psi^{g}_{m}\psi^{g}_{l}\rangle) \tag{5}$$

The Coulomb operator, V_{ml}, between electrons of molecules m and l is given by

$$\hat{V}_{ml} = \frac{1}{4\pi\varepsilon_0}\sum_{i}^{n_m}\sum_{j}^{n_l}\frac{e^2}{(r_i - r_j)} \tag{6}$$

where r_i and r_j are the positions of the ith and jth electrons on molecule m and l, respectively. n_m is the total number of electrons on molecule m.

Eigenvectors $\tilde{C}^k$ and eigenvalues E_k of the collective excited states are obtained by diagonalizing the exciton matrix. This information can be used to obtain optical properties such as absorption and circular dichroism (CD) spectra. The dipole strength of the kth eigenstate (D_k) that determines the absorption spectrum is given by

$$\begin{aligned} D_k &= |\sum_{m}^{N_{mol}}\sum_{i} C^{k}_{m,i}\vec{\mu}^{i}_{m}|^2 \\ &= \sum_{m}^{N_{mol}}(C^{k}_{m,i})^2|\vec{\mu}^{i}_{m}|^2 + \sum_{\substack{m,n \\ m\neq n}}^{N_{mol}}\sum_{i,j} C^{k}_{m,i}C^{k*}_{n,j}(\vec{\mu}^{i}_{m}\cdot\vec{\mu}^{j}_{n}) \end{aligned} \tag{7}$$

Here, $\vec{\mu}^{i}_{m}$ is the transition dipole vector of the ith excited state on molecule m. To calculate the circular dichroism (CD) spectra, the rotational strength of the kth eigenstate (R_k) has to be evaluated. This is given by

$$R_k = -\frac{\pi}{2\lambda_k}\sum_{\substack{m,n \\ m\neq n}}^{N_{mol}}\sum_{i,j} C^{k}_{m,i}C^{k*}_{n,j}(\vec{X}_{mn}\cdot\vec{\mu}^{i}_{m}\times\vec{\mu}^{j}_{n}) \tag{8}$$

where λ_k is the excitation wavelength associated with the kth eigenstate and $\vec{X}_{mn}$ is the displacement vector between the centers of molecules m and n. The centers are determined by the weighted average of the corresponding transition densities.[10]

In the exciton model calculations, N molecular excited states are combined to give N delocalized excitonic states of the system. The dipole strength (absorption spectrum) and rotational strength (CD spectrum) are calculated using eqs 7 and 8, respectively, which results in a spectrum containing N discrete absorption lines. To make a comparison with the experimental spectra, these should be broadened to take structural and environmental fluctuations into account. This can be achieved by simple Gaussian broadening with a standard deviation of σ_{broad}, but a more insightful way is to explicitly include disorder into the exciton Hamiltoninan. It should be noted that these two approaches to account for disorder do not necessarily give similar results.[27]

The degree of delocalization of the exciton can be characterized by the inverse participation ratio L_k, also called the exciton delocalization length, which gives an indication of the number of coherently coupled chromophores for the kth eigenstate (or the spatial extent of the excited state) and is defined as[10,27]

$$L_k = \frac{1}{\sum_{m}^{N_{mol}}\sum_{i}(C^{k}_{m,i})^4} \tag{9}$$

2.2. Evaluation of Parameters. The focus of the present work is the lowest excited states of DNA hairpin model systems. Therefore, only the lowest-energy states of the constituent chromophores have been included in the exciton matrix. Neglecting higher-energy states and dark states has little influence on the excited-state properties because they are only weakly coupled to the low-energy states.[27] In previous studies by Markovitsi et al., the inclusion of the two lowest-energy excited states for purine bases and one excited state of pyrimidine bases has been proposed on the basis of the optical spectra of nucleobases in water solution.[10,29−31]

An accurate estimate of the site energies (ε) and exciton couplings (J) is crucial to describing the excited states correctly. To determine the site energies, the excited-state energies of individual molecules ($\Delta\varepsilon$) in the absence of other molecules have to be determined first (eq 3). All nucleobases absorb within a narrow frequency range. For this reason, the optical properties of DNA strands cannot be used to estimate the absorption energies ($\Delta\varepsilon$) and oscillator strengths of constituent nucleobases. However, the optical spectra of nucleobases in solution do not give a reliable estimate of the excitation energies and oscillator strengths because solvent polarization effects and the presence of different tautomers complicate the assignment of the transitions.

A more reliable way to obtain $\Delta\varepsilon$ is from gas-phase spectroscopy data. Laser-desorption techniques in combination with ultrafast spectroscopy have been used to measure gas-phase optical spectra.[32] However, the interpretation and assignment of peaks have been very controversial because of the presence of several nucleobase tautomers that are simultaneously detected by these techniques. Therefore, a direct comparison of the gas-phase data with theoretical calculations is not always possible. Among all nucleobases, guanine has been studied most extensively. The resonance-enhanced multiphoton ionization (REMPI) spectrum of guanine was first reported by De Vries et al. in 1999, who demonstrated the presence of several tautomers in the gas phase.[33,34] From subsequent experiments, it was concluded that the most stable 9-H keto conformation could not be observed because of the very short lifetime of its excited states, and four rare conformations were instead assigned to the observed spectra.[35−40]

The excited-state energies of the four rare conformations of guanine obtained from REMPI spectra are best described by the adiabatic excitation energies calculated with the complete active-space second-order perturbation theory (CASPT2).[41] The conformation that prevails in DNA, according to these calculations, has an estimated adiabatic excitation energy of ∼4 eV. Literature values for the excitation energies of nucleobases obtained from CASPT2 calculations are summarized in Table 1. The calculated values are compared to experimental values obtained in water solution that were used in previous exciton studies because of the lack of reliable gas-phase and theoretical values.[10] On the basis of the excitation energies obtained from CASPT2, only the lowest excited state is included for thymine and cytosine, but the two lowest excited states are included for guanine. For adenine, only the second excited state is included

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

Table 1. Excitation Energies and the Corresponding Oscillator Strengths (f) of the Lowest Excited States of Nucleobases Deduced from the Experimental Absorption Spectra in Water Solution and Theoretical (CASPT2) Calculations

	energy (eV)/f (exp[10])		energy (eV)/f (CASPT2)	
	excited state 1	excited state 2	excited state 1	excited state 2
adenine	4.56/0.05	4.82/0.19	5.16/0.004[47]	**5.35/0.175**[47]
thymine	4.66/0.24		**4.89/0.167**[48]	
cytosine	4.57/0.21		**4.50/0.065**[49]	
guanine	4.56/0.19	5.01/0.21	**4.93/0.158**[50]	**5.77/0.145**[50]

because the lowest excited state has zero oscillator strength from the ground state and hence will not couple with neighboring chromophores.

The absorption spectra of the stilbene linker (Sd) and capping stilbene (Sa) are distinct from those of nucleobases. The maximum of the lowest absorption band of both stilbenes is at ca. 330 nm, and these absorption bands do not overlap with the low-energy bands of the nucleobases.[42,43] As a result of this difference in excited-state energies, the excited states of stilbenes and nucleobases are weakly coupled. Therefore, the delocalized exciton states on the DNA stacks are not affected by the precise values of site energies and oscillator strengths of Sa and Sd. Moreover, because the spectra of Sa and Sd are similar for isolated molecules in solution and those incorporated in the hairpins,[44] the experimental excitation energy (3.71 eV) and transition dipole moment (6.7 D) for Sa are used in the present study.[45] These values are close to the corresponding values of 4 eV and 7.2 D for unsubstituted stilbene obtained using CASPT2 calculations.[46]

The exciton coupling (J_{ml}) and displacement energy (J_{mm}) can be calculated using the atomic transition densities q_i^t, the excited-state atomic charge densities q_i^e, and the ground state atomic charge densities q_i^g associated with molecular excited states.

$$J_{ml} = \frac{1}{4\pi\varepsilon_0}\sum_{i}^{N_m}\sum_{j}^{N_l}\frac{q_i^t q_j^t}{(R_i^m - R_j^l)} \tag{10}$$

$$J_{mm} = \frac{1}{4\pi\varepsilon_0}\sum_{l\neq m}^{N_{mol}}\sum_{i}^{N_m}\sum_{j}^{N_l}\left[\frac{q_i^e q_j^g}{(R_i^m - R_j^l)} - \frac{q_i^g q_j^g}{(R_i^m - R_j^l)}\right] \tag{11}$$

where R_i^m is the position of atom i on molecule m and N_m is the total number of atoms on molecule m. The evaluation of atomic (transition) densities is performed using singly excited configuration interaction (CIS) calculations based on an INDO ground-state wave function, as described previously.[26,27] Although the direction of the transition dipole moments is also obtained from these calculations, it is convenient to evaluate it using the sum of atomic position vectors weighted by atomic transition densities for arbitrarily oriented molecules in the hairpins.[27]

2.3. Molecular Dynamics Simulations. Atomistic molecular dynamics simulations were performed to account for the effect of structural fluctuations on the excited-state properties. The classical MD simulations were performed on all DNA hairpins in solution at 300 K using the Amber 8 package[51] as described in previous publications.[21,45] After equilibration, a production MD simulation of 2 ns was performed. From these simulations, 400 snapshots were extracted with a time interval of 5 ps. This information was used to investigate the excited-state properties as a function of time as well as the average properties.

3. RESULTS AND DISCUSSION

A schematic diagram of the structures of the studied hairpins is provided in Figure 1. The dependence of the parameters that determine the excited-state and optical properties of the AT hairpins (Figure 1a) on structural fluctuations is discussed in Section 3.1. Subsequently, in Sections 3.2 and 3.3, the above parameters are used to calculate the optical properties and excited-state electronic structure of these hairpins. The effect of introducing GC into hairpins consisting of three base pairs (Figure 1b) is considered in Section 3.4. The same and analogous hairpins have previously been used to study excited-state properties and charge transfer experimentally.[52]

3.1. Structural Dependence of the Site Energy. The site energy in the exciton Hamiltonian description corresponds to the excitation energy of a single chromophore in the presence of the electrostatic environment of its neighbors. It is composed of several contributions (eq 3): the excitation energy in vacuum ($\Delta\varepsilon$), the shift, J_{mm}, due to the neighboring electrostatic environment, and a random disorder term, δ_m^i, due to structural deformations or solvent effects sampled from a Gaussian distribution with a standard deviation of σ_δ. This means that $\Delta\varepsilon$ is constant whereas J_{mm} fluctuates in time. Using the snapshots from the MD simulations that were discussed in Section 2, variations of the sum $\Delta\varepsilon + J_{mm}$, were calculated as a function of time. The dependence of $\Delta\varepsilon + J_{mm}$ on time and its

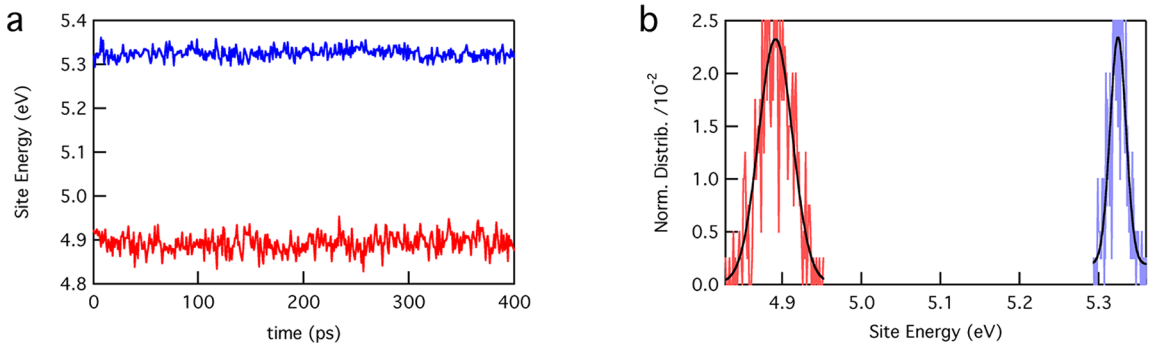


Figure 2. (a) Temporal fluctuations in the site energy, $\Delta\varepsilon + J_{mm}$, for the excited states of middle thymine (red) and adenine (blue) nucleobases in hairpin 3AT. (b) Corresponding distribution of the site energies.

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

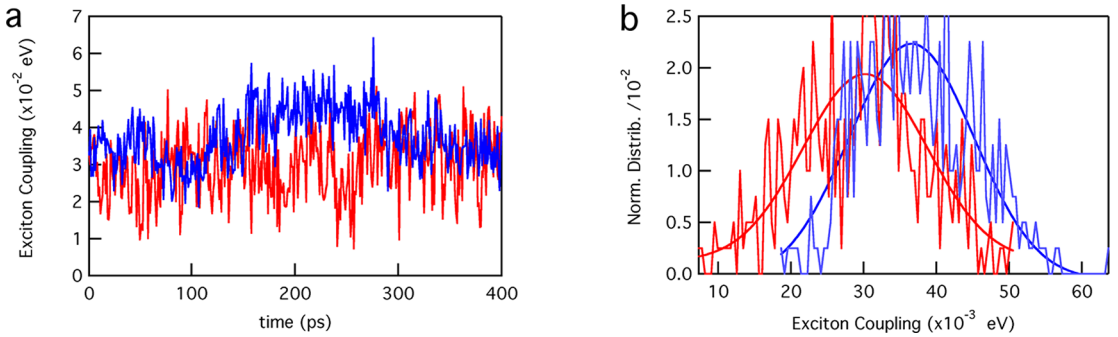


Figure 3. (a) Temporal fluctuations in the exciton couplings, J_{ml}, between neighboring thymines (red) and adenines (blue) shown in hairpin 3AT. (b) Corresponding distributions of J_{ml}.

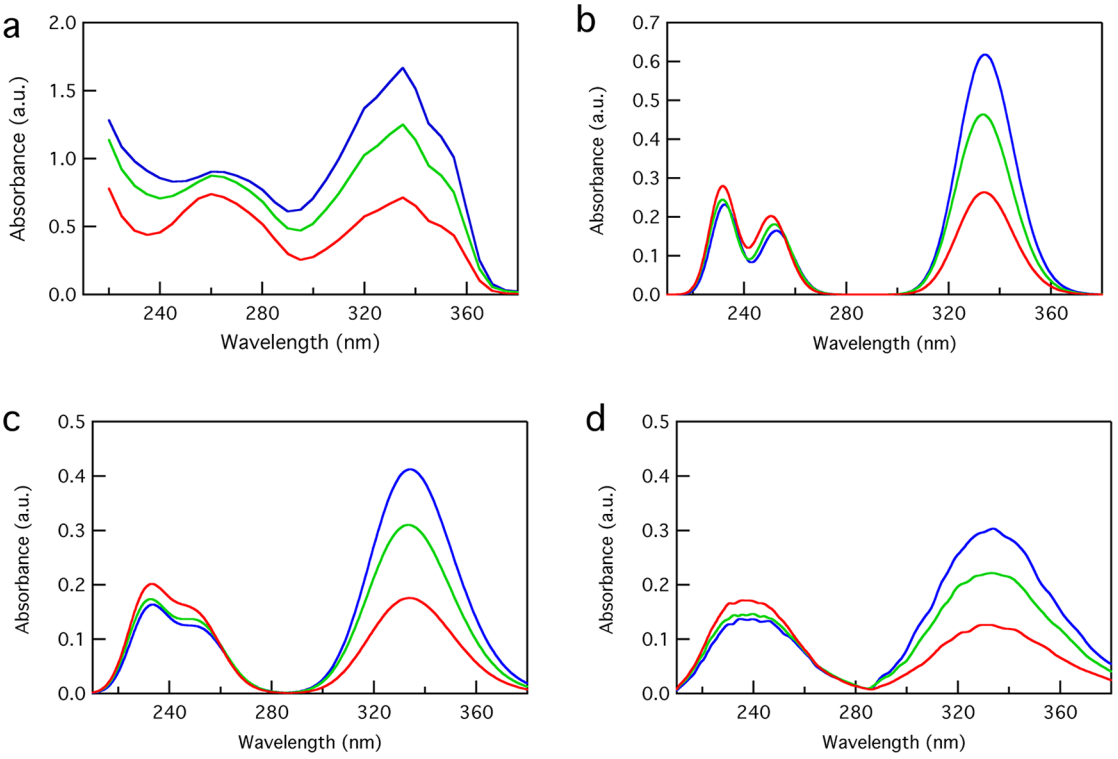


Figure 4. Absorption spectra for DNA hairpins in arbitrary units, normalized with the number of nucleobases, for 2AT(blue), 3AT(green), and 6AT(red): (a) experimental spectra;[44] (b, c) calculated spectra broadened with Gaussians with standard deviations of σ_{broad} = 0.12 and 0.18 eV, respectively; and (d) calculated spectra with disorder in site energies of σ_δ = 0.25 eV.

distribution are shown in Figure 2 for the 3AT hairpin that consists of three adenine-thymine base pairs. The distribution of the site energies (in the absence of random disorder) is significantly broader for thymine than it is for adenine. The distributions are well described by a Gaussian function with standard deviations of σ = 0.021 eV for thymine and σ = 0.014 eV for adenine. This difference in σ is due to differences in the atomic charge distributions, as explained in Section 2. The values for J_{mm} are on the order of 0.01 eV; that is, they are two orders of magnitude lower than the corresponding excitation energies ($\Delta\varepsilon$). Similar results have been obtained for nucleobases in other hairpins at different positions, which shows that the distribution of site energies and their fluctuations can be sampled independently for each nucleobase.

The width of the experimental absorption spectra of isolated nucleobases and DNA duplexes in solution indicates site-energy fluctuations of the order of 0.1 eV, which is an order of magnitude higher than the fluctuations in J_{mm} calculated here. These large fluctuations in the excitation energy of the individual bases ($\Delta\varepsilon$) are largely caused by structural fluctuations inside the bases. In previous theoretical work on charge transfer in DNA hairpins, it was shown that such structural fluctuations as obtained from MD simulations result in distributions of energies of the highest occupied molecular orbital with a width of 0.15 eV.[21] Because there is an indirect relation between fluctuations in the orbital energies and the excitation energy, it seems likely that fluctuations on the order of 0.1 eV can be attributed to structural deformations of the bases. From this, it can be concluded that the energy shifts caused by structural fluctuations (σ_δ) are considerably larger than those due to interaction with neighboring bases (i.e., the displacement energy, J_{mm}). These results are in agreement with previous studies on DNA double helices.[10,29−31]

3.2. Structural Dependence of the Exciton Coupling.

The off-diagonal elements of the exciton Hamiltonian, or the exciton coupling matrix elements (J_{ml}), are responsible for the

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

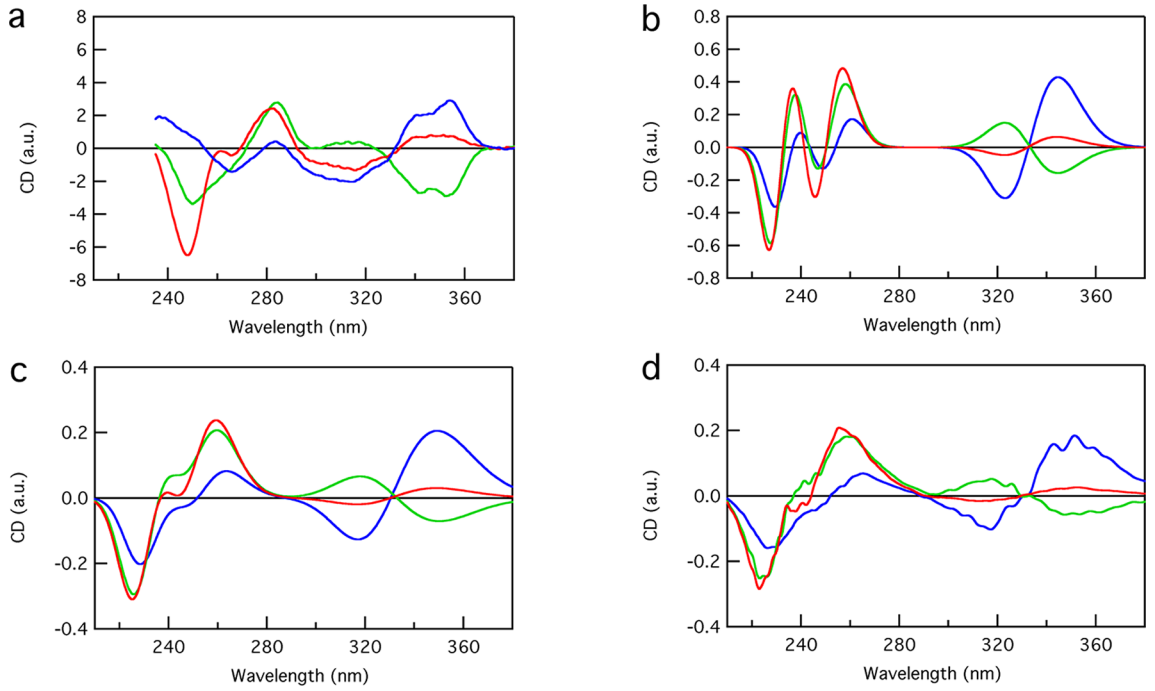


Figure 5. CD spectra for DNA hairpins (blue) 2AT, (green) 3AT, and (red) 6AT: (a) experimental spectra; (b, c) calculated spectra broadened with Gaussians with standard deviations of σ_{broad} = 0.12 and 0.18 eV, respectively; and (d) calculated spectra with a disorder in site energies of σ_δ = 0.25 eV.

mixing of the localized excited states on the individual bases, which may lead to delocalized excitonic states. The exciton coupling between neighboring thymines and neighboring adenines was calculated for the snapshots obtained from the MD simulations described above. In Figure 3, these exciton coupling matrix elements are plotted as a function of time and their distribution is shown. The distribution was found to be close to Gaussian for both cases with average values of 0.030 and 0.036 eV for the AA and TT couplings, respectively. Moreover, the width of the distributions (σ = 0.085 eV) is of the same order of magnitude as the absolute value of the coupling, which means that the structural dynamics will markedly influence the degree of delocalization of the excited states. Further calculations have shown that the exciton coupling between nucleobases on different strands (1×10^{-3}–7×10^{-3} eV) is even smaller than the intrastrand coupling. Therefore, if delocalized excitonic states exist, then they are likely to extend over only one strand of DNA and not across the hydrogen bonds to the other strand.

3.3. Absorption and CD Spectra of AT Hairpins. Using the parameters discussed above and the methodology outlined in Section 2, we can calculate the optical absorption and CD spectra for the hairpins. The calculated optical absorption spectra for the three AT-containing hairpins are shown in Figure 4, together with the experimental spectra. In all absorption spectra, two or three distinct absorption bands can be discriminated, depending on the way in which the site-energy disorder is included. (See below.) The lowest-energy band at around 330 nm is due to the absorption of stilbenes Sa and Sd. At higher energy of around 240 nm, there is an absorption feature due to nucleobases adenine and thymine. The overall agreement between the experimental (Figure 4a) and calculated spectra (Figures 4b–d) is found to be reasonable. In the calculated spectra, the band due to the nucleobases is shifted ~25 nm to the blue, which is attributed to the limited description of polarization effects in the model.

As discussed in Section 2.1, the disorder included in the calculations by averaging over the different structural conformations from the MD simulations is rather limited because only contributions due to the direct interaction between bases are taken into account (through J_{ml} and J_{mm}, see eqs 4 and 5). Comparison with experimental data shows that this is not sufficient. In the calculated spectra in Figure 4b,c, the additional disorder is accounted for by convoluting the calculated spectra (without random fluctuations in site energies, σ_δ = 0) with Gaussians having standard deviations of σ_{broad} = 0.12 and 0.18 eV, respectively, and averaging over 400 conformations obtained from MD simulations.

A more physical way to include the additional disorder in the calculated spectra is by explicitly including disorder in the site energies in the exciton matrix. For each of the 400 snapshots from the MD simulation, 50 different realizations of the site energy were sampled. The disorder in site energies σ_δ was varied until a reasonable agreement with the experimental spectra was reached. The absorption spectra obtained by including random fluctuations in the site energies with a standard deviation of σ_δ = 0.25 eV is shown in Figure 4d. Comparison with the experimental spectra in Figure 4a shows that the individual bands due to the adenine and thymine cannot be discerned, indicating that a site-energy disorder of ~0.25 eV is reasonable.

The experimental and calculated CD spectra for hairpins 2AT, 3AT, and 6AT are shown in Figure 5. In Figure 5b,c, the spectra convoluted with Gaussians with standard deviations of σ_{broad} = 0.12 and 0.18 eV, respectively, are shown. In the case of smaller broadening, two distinct positive bisignate curves can be identified at 233 and 250 nm. These correspond to the intrastrand excitonic bands of adenine and thymine at 254 and 233 nm that are shown in Figure 4b. On increasing the width of

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447–11458

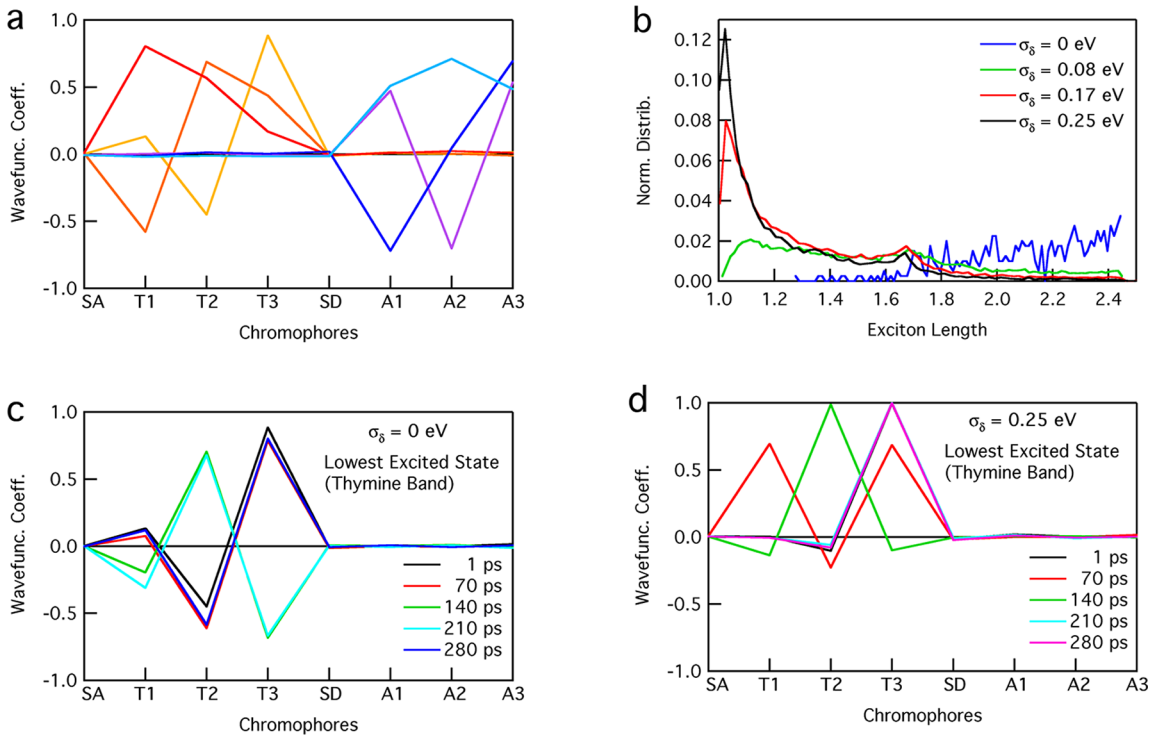


Figure 6. (a) Excited-state wave functions in hairpin 3AT for one MD snapshot and $\sigma_\delta = 0$, thymine band (red, left side), and the adenine band (blue, right side). (b) Average exciton delocalization length for excitons localized on thymines for different values of σ_δ. (c) Time dependence of the lowest-energy wave function for $\sigma_\delta = 0$ eV. (d) Time dependence of the lowest-energy wave function for $\sigma_\delta = 0.25$ eV.

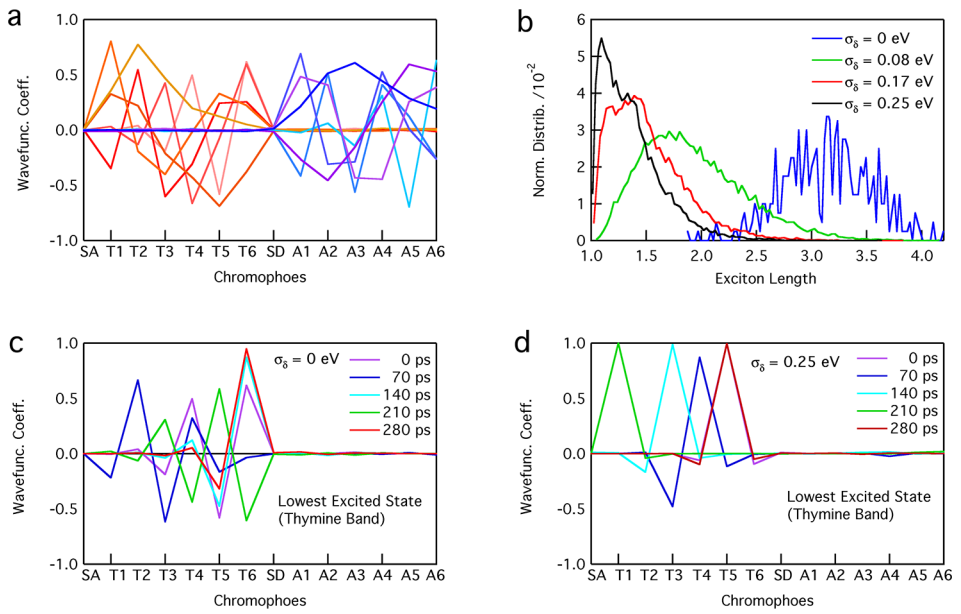


Figure 7. (a) Excited-state wave functions in hairpin 6AT for one MD snapshot and $\sigma_\delta = 0$, the thymine band (red, left side), and the adenine band (blue, right side). (b) Average exciton delocalization length for excitons localized on thymines for different values of σ_δ. (c) Time dependence of the lowest-energy wave function for $\sigma_\delta = 0$ eV. (d) Time dependence of the lowest-energy wave function for $\sigma_\delta = 0.25$ eV.

the Gaussian broadening to $\sigma_{broad} = 0.18$ eV, the two bisignate curves merge to give a single bisignate curve. In this context, it is interesting that for hairpins containing a small number of base pairs (<4) a single bisignate curve appears in the experimental CD spectra, whereas a shoulder can clearly be identified on the positive peak for larger hairpins.[52] This can be attributed to the formation of a more rigid structure when the number of base pairs in the hairpin is increased. As discussed above for the absorption spectra, a more direct way to include the effect of disorder is by including it in the site energies in the exciton Hamiltonian. The CD spectra calculated by including random disorder in site energies ($\sigma_\delta = 0.25$ eV) is shown in Figure 5d.

Overall, the correspondence between calculated and experimental spectra is found to be quite good. In the region around 330 nm where the stilbenes absorb, the dependence on

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

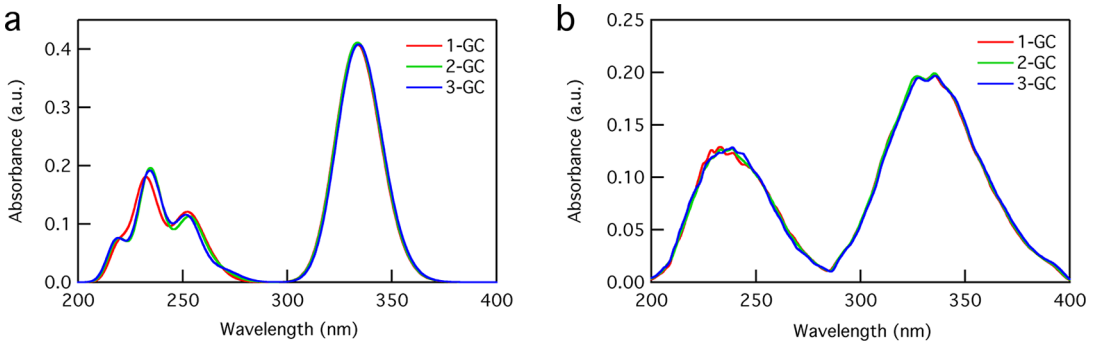


Figure 8. Absorption spectra for GC-containing DNA hairpins with (a) broadening after calculation (σ_{broad} = 0.12 eV) and (b) energy disorder in the site energy (σ_δ = 0.25 eV).

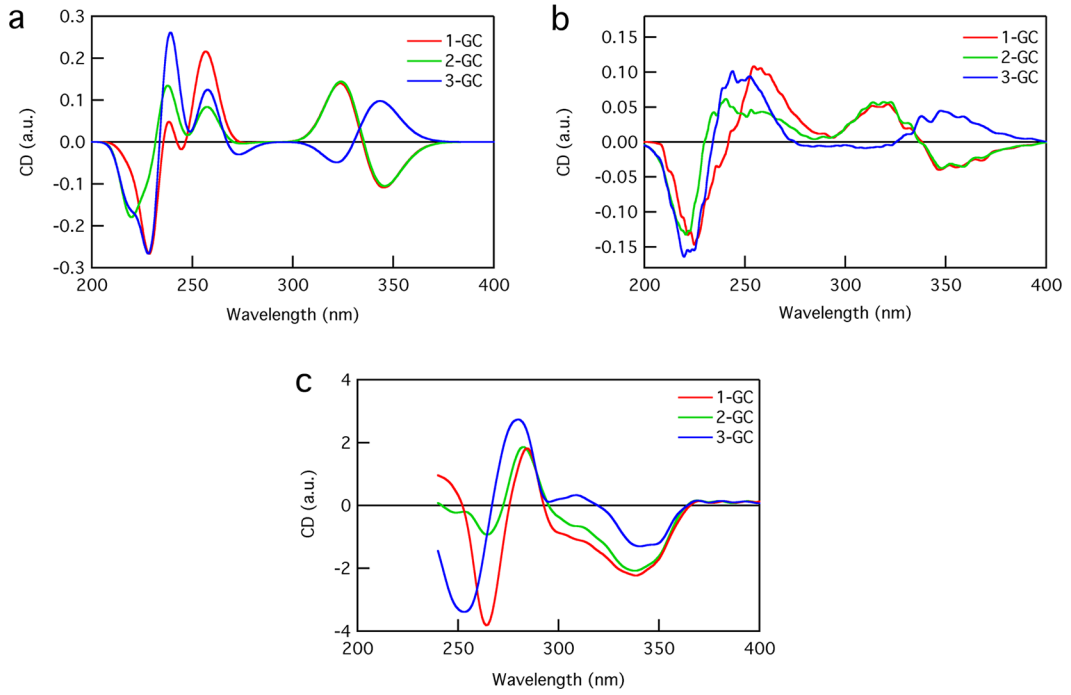


Figure 9. CD spectra for GC-containing DNA hairpins with (a) broadening after calculation (σ_{broad} = 0.12 eV) and (b) energy disorder in the site energy (σ_δ = 0.25 eV). (c) Experimental spectra for these hairpins.

the number of AT base pairs in the hairpin is nicely reproduced. The bisignate feature in the CD spectrum strongly depends on the mutual orientation of the two stibenes.[52] The sign of this feature changes twice on increasing the number of AT base pairs from 2 to 6, and the amplitude decreases with increasing distance. Also, in the nucleobase region of the spectra, the CD spectrum is reasonably reproduced, although there is a slight shift toward high energy, as discussed above for the absorption spectrum. The good description of the spectrum with a site-energy disorder illustrates that the model that is used here is very reasonable and that it may also be used to study the nature of the excited states in more detail.

3.4. Nature of the Excited States. One of the issues in the discussions on the photophysics of DNA is the nature of the initially excited state (i.e., whether it is delocalized or localized on a single base[4−8]). The advantage of theoretical work is that it gives not only the spectra that can be compared to experimental results but also the wave functions of the excited states. From these, insight is obtained into the nature of these excited states. The excited-state wave functions that were calculated for hairpins 3AT and 6AT for a single conformation from the MD simulations are plotted in Figures 6a and 7a, respectively. The figures do not include the two lowest-energy states because these are localized on stilbenes. In these calculations, no additional site-energy disorder (σ_δ) was included. It is clear that in the absence of additional site-energy disorder the excited states are almost fully delocalized along a single strand, either on thymines or on adenines. This can be attributed to two factors: the site energies of thymines and adenines are significantly different (~0.46 eV) and the interstrand exciton coupling is small compared to intrastrand exciton coupling. The thymine exciton bands are lower in energy than adenine bands because of the lower site energies for the thymine bases.

The degree of delocalization can be quantified by calculating the exciton delocalization length, L, as given by eq 9. In Figures 6b and 7b, the exciton delocalization length is shown for different amounts of additional site-energy disorder for hairpins 3AT and 6AT, respectively. The value of L was calculated by averaging over all excited states that are localized on the thymine bases and is therefore an average for all exciton states on the thymines. The random fluctuations in the site energies

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

were included as described above. In the absence of additional site-energy disorder, the excited states are rather delocalized, extending over about three thymine bases for both hairpins 3AT and 6AT. The relatively broad distribution of L shows that the disorder due to the displacement energy J_{mm} and exciton coupling has a considerable effect on the distribution of the exciton state. On increasing the amount of site-energy disorder, the average of the distribution shifts to shorter exciton lengths but the distribution remains rather broad, especially for 6AT.

The effect of site-energy disorder (σ_δ) is also clearly observed in the wave function of the exciton states. The wave function for the lowest exciton on the thymine bases is plotted as a function of time with and without disorder (σ_δ = 0 and 0.25 eV) in Figures 6c and 7c and 6d and 7d, respectively. In the absence of disorder, the lowest exciton is always delocalized over different thymines, although the position fluctuates in time. When disorder is introduced, the exciton becomes almost completely localized on a single thymine, with small amplitudes on the neighboring bases.

3.5. Effect of Base Pair Sequence on Spectra and Nature of Excited States. As discussed in the Introduction, the UV-induced photodimerization of thymine bases in DNA has been found to be very sensitive to the base pair sequence.[16] The presence of a GC base pair in the vicinity of a TT step appears to be particularly important.[53,54] In this section, the effect of the presence of a GC base pair on the properties of the excited states of AT hairpins is studied. Three hairpins are considered that were obtained by replacing one of the AT base pairs in 3AT by a GC base pair (Figure 1b).

The calculated absorption and CD spectra of the hairpins are shown in Figure 8, both for broadening after the calculation (Figure 8a) and for the inclusion of site-energy disorder (Figure 8b). The calculated absorption spectra are very similar for all hairpins and do not change much on the inclusion of GC base pairs. The opposite is true for the calculated CD spectra shown in Figure 9a,b, where distinct differences between different GC-containing hairpins are observed. In hairpin 3-GC, the sign of the bisignate curve in the stilbene region is opposite to that in the other two GC-containing hairpins, which is surprising considering the same helicity of all hairpins. To explain this, the distributions of exciton couplings between stilbenes for different hairpins are plotted in Figure 10. This figure shows that a subtle difference in the structural conformation of 3-GC changes the sign of exciton coupling between stilbenes. As a result, the sign of the bisignate curve in the calculated CD spectrum is reversed. These results have important implications for the interpretation of optical properties of stilbene-capped DNA hairpins where the stilbenes have been used not only for the stability of the hairpins but more importantly as a ruler for the helicity, conformational flexibility, and length of the hairpins.[17,44,45]

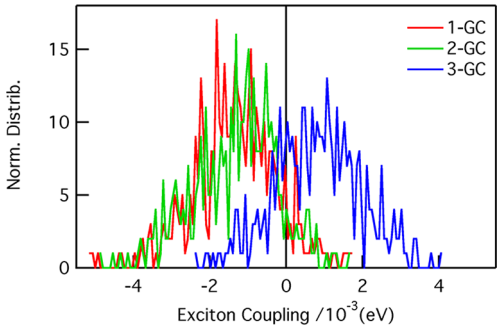


Figure 10. Distribution of exciton coupling between the stilbenes in GC-containing hairpins.

Differences are also observed in the nucleobase region of the CD spectra, although the overall shape is similar for all of the GC-containing hairpins. The adenines and thymines in 2-GC are separated by one GC base pair whereas they are adjacent neighbors in 1-GC and 3-GC. As a result, the exciton coupling between the two adenines and the two thymines is smaller for 2-GC, resulting in a lower CD intensity. The presence of one GC base pair (in the absence of site-energy disorder) makes only a small contribution to the CD intensity because it is weakly coupled to the nucleobases in their respective strands as a result of differences in site energies (Table 1). A comparison of the calculated CD spectra with the experimental spectra for the GC-containing hairpins shows that the opposite sign of the bisignate curve for 3-GC is not reproduced well. This is attributed to the very subtle nature of the occurrence of these bisignate features, with slight changes in the average structure having a strong effect on the overall shape of the CD spectra. Because of the strong dependence of the CD spectra on the structural degrees of freedom, it can be used as a good indication of the accuracy of the molecular dynamics force field used. Although the force field used here gives good overall agreement for most hairpins, the mismatch for 3-GC shows that there is some room for improvement in the force field.

The different CD spectra of the hairpins imply differences in their excited-state properties. The exciton delocalization lengths (L) of the thymine band are plotted as a function of time in Figure 11 for hairpins 1-GC, 2-GC, 3-GC, and 2AT in the absence of site-energy disorder. The values do not exceed 2, which indicates that excitations in the thymine exciton band are delocalized primarily over the two neighboring thymines. For 1-GC, 3-GC, and 2AT, the value of L varies widely between 1 and 2, whereas excitations in 2-GC are primarily localized on a single base (thymine). This is due to the weak exciton coupling of thymines that are separated by GC in hairpin 2-GC, as discussed above.

The effect of the introduction of site-energy disorder on the exciton delocalization length (L) of the thymine band for 3-GC is plotted in Figure 12a. Contrary to the situation in AT hairpins, the presence of site-energy disorder broadens the distribution of L rather than making it narrower. Values between 2 and 3 start occurring because of the partial delocalization of thymine excited states on cytosine. The site energies of cytosine and thymine are 4.5 and 4.89 eV, respectively. The disorder in site energies, which is on the order of the difference in their site energies, is responsible for the reduction in their energy difference for several realizations. This results in the intermittent occurrence of strongly coupled states delocalized over two thymines and a cytosine. The delocalization length increases as the site-energy disorder is increased and becomes constant when the width of the disorder distribution (σ_δ) is close to the difference in the site energies for thymine and cytosine. In Figure 12b, the distributions of L are plotted at σ_δ = 0.25 eV for 1-GC, 2-GC, 3-GC, and 3AT. The distributions are very similar for all hairpins except for 2-GC, which exhibits slightly less delocalized excitations on average.

The above results show that diagonal disorder reduces the distribution of L for chromophores that are close in energy, whereas it increases the distribution of L for chromophores that

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

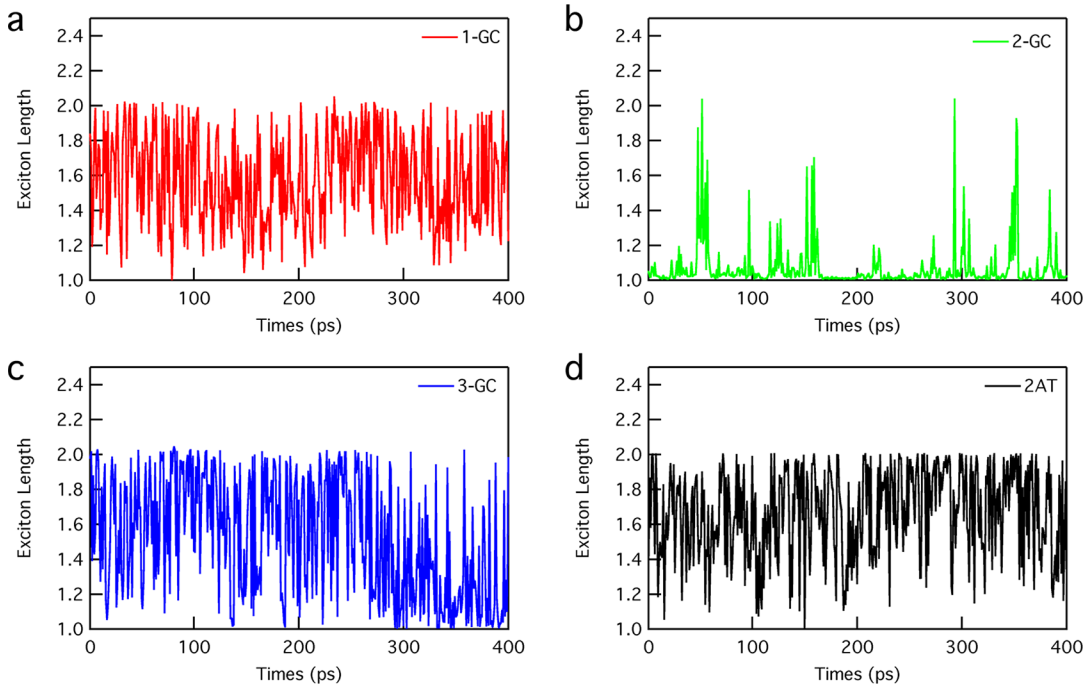


Figure 11. Delocalization as a function of time for GC-containing hairpins in the absence of site-energy disorder.

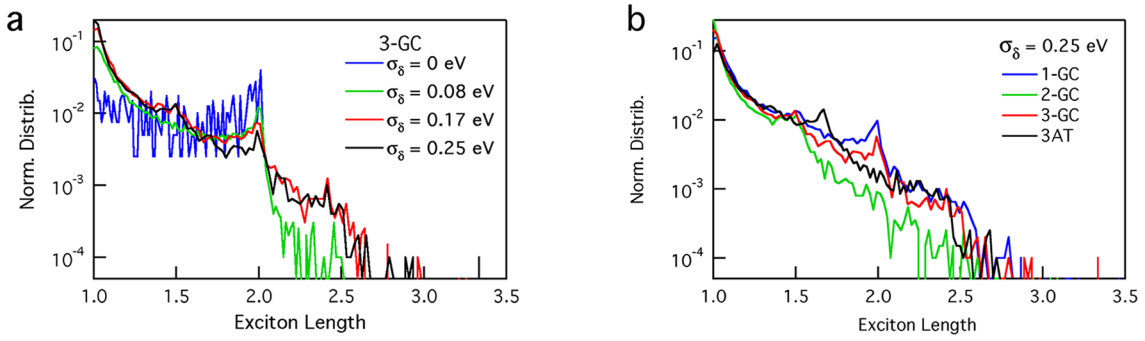


Figure 12. (a) Distributions of the exciton delocalization length L for the thymine band at different levels of site-energy disorder (σ_δ) for 3-GC. (b) Distribution of L at σ_δ = 0.25 eV for different hairpins.

are energetically separated. The mixing of exciton bands associated with different nucleobases can play a crucial role in excited-state dynamics and photoinduced reactions, such as the formation of the thymine dimer. It has been shown experimentally that the photodimerization of thymine−thymine steps in DNA depends to a large extent on the nucleobases in the environment.[16] This context dependence can, to some extent, be attributed to differences in the conformation of the thymine−thymine step induced by the surrounding bases.[17,55] However, it cannot be excluded that the electronic structure of the excited state also has a distinct effect on its being prone to TT dimerization. In this respect, it is of interest to examine the exciton coupling between the cytosine and its neighboring thymine for hairpins 1-GC and 3-GC. For hairpin 3-GC, this exciton coupling, averaged over all 400 conformations from the MD simulations, is 2×10^{-2} eV, which is very similar to the coupling between two thymines in the same hairpin (1.9×10^{-2} eV). For 1-GC with cytosine on the 3′ side of the thymines, the exciton coupling between cytosine and thymine is 1.4×10^{-4} eV, which is 2 orders of magnitude smaller than in 3-GC, where the cytosine is located on the 5′ side. These large differences in exciton coupling can result in considerable differences in the degree of delocalization in the hairpins, which is likely to have a pronounced effect on the photodimerization. It should be noted that all calculations presented here are based on MD simulations for the ground state of the hairpins. It cannot be excluded that structural rearrangements take place upon excitation, which may also have an effect on the TT dimerization efficiency.

4. SUMMARY AND CONCLUSIONS

The excited-state and optical properties of DNA hairpins have been investigated in the framework of a Frenkel exciton model, accounting explicitly for structural dynamics by including structural information from molecular dynamics simulations.

The agreement between the calculated absorption and circular dichroism (CD) spectra and experiments is good for hairpins containing only adenine-thymine base pairs. It is shown that it is possible to rationalize the optical properties in terms of various interbase excitonic interactions and to assign exciton bands associated with nucleobases and stilbenes in the observed spectra. Intrastrand exciton states are formed in the hairpins that correspond to adenine and thymine strands. The excited-state properties are sensitive to molecular distortion and

dx.doi.org/10.1021/jp307146u | J. Phys. Chem. B 2012, 116, 11447−11458

structural dynamics of the hairpin. Because of the disorder in site energy, the excitons become largely localized with the intermittent occurrence of delocalized states, which are believed to play a role in the photodimerization of thymines.

The effect of the presence of a GC base pair on the properties of AT hairpins is studied in order to shed light on its possible influence on the photodimerization of thymines. The energy difference between thymine and cytosine is such that in the absence of disorder the excited states do not delocalize over both types of bases. The introduction of additional site-energy disorder leads to the intermittent delocalization of the excited state over both adenines and thymines. The calculations show that the exciton coupling between thymine and cytosine are strongly dependent on the mutual position of the bases. The exciton coupling for a cytosine on the 5′ side of the two thymines is shown to be 2 orders of magnitude larger than for a cytosine on the 3′ end. This will certainly have an effect on the excited-state delocalization in DNA and on the dependence of the photodimerization of thymines on the local base pair environment.

AUTHOR INFORMATION

Corresponding Author

*E-mail: f.c.grozema@tudelft.nl.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by a VENI grant from The Netherlands Organization for Scientific Research (NWO) and by the National Science Foundation (CHE-0628130). We thank Dr. Arun K. Thazhathveetil for providing the CD spectra of GC hairpins.

REFERENCES

(1) Dumaz, N.; Drougard, C.; Sarasin, A.; Daya-Grosjean, D. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10529−10533.
(2) Kraemer, K. H. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11−14.
(3) Taylor, J. S. *J. Chem. Educ.* **1990**, *67*, 835−841.
(4) Fiebig, T. *J. Phys. Chem. B* **2009**, *113*, 9348−9349.
(5) Bucharov, I.; Wang, Q.; Raytchev, M.; Trifonov, A.; Fiebig, T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4794−4797.
(6) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977−2019.
(7) Crespo-Hernández, C. E.; Cohen, B.; Kohler, B. *Nature* **2005**, *436*, 1141−1144.
(8) Markovitsi, D.; Talbot, F.; Gustavsson, T.; Onidas, D.; Lazzarotto, E.; Marguet, S. *Nature* **2006**, *441*, E7−E7.
(9) Johnson, W. C.; Tinoco, J., Jr. *Biopolymers* **1969**, *7*, 727−749.
(10) Bouvier, B.; Gustavsson, T.; Markovitsi, D.; Millié, Q. *Chem. Phys.* **2002**, *275*, 75−92.
(11) Rist, M.; Wagenknecht, H.-A.; Fiebig, T. *Chem. Phys. Chem* **2002**, *3*, 704−707.
(12) Vayá, I.; Gustavsson, T.; Miannay, F. o.-A.; Douki, T.; Markovitsi, D. *J. Am. Chem. Soc.* **2010**, *132*, 11834−11835.
(13) Bittner, E. R.; Czader, A. In *Energy Transfer Dynamics in Biomaterial Systems*; Burghardt, I., May, V., Micha, D. A., Bittner, E. R., Eds.; Springer-Verlag: Berlin, 2009; Vol. *93*, pp 103−126.
(14) Becker, M. M.; Wang, J. C. *Nature* **1984**, *309*, 682−687.
(15) Neelakandan, P. P.; Pan, Z.; Hariharan, M.; Lewis, F. D. *Photochem. Photobiol. Sci.* **2012**.
(16) Hariharan, M.; Lewis, F. D. *J. Am. Chem. Soc.* **2008**, *130*, 11870−11871.
(17) Hariharan, M.; Mccullagh, M.; Schatz, G. C.; Lewis, F. D. *J. Am. Chem. Soc.* **2010**, *132*, 12856−12858.
(18) Becker, M. M.; Wang, Z. *J. Mol. Biol.* **1989**, *210*, 429−438.
(19) Tuma, J.; Tonzani, S.; Schatz, G. C.; Karaba, A. H.; Lewis, F. D. *J. Phys. Chem. B* **2007**, *111*, 13101−13106.
(20) Lewis, F. D.; Liu, X.; Wu, Y.; Miller, S. E.; Wasielewski, M. R.; Letsinger, R. L.; Sanishvili, R.; Joachimiak, A.; Tereshko, V.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 9905−9906.
(21) Grozema, F. C.; Tonzani, S.; Berlin, Y. A.; Schatz, G. C.; Siebbeles, L. D. A.; Ratner, M. A. *J. Am. Chem. Soc.* **2008**, *130*, 5157−5166.
(22) Grozema, F. C.; Tonzani, S.; Berlin, Y. A.; Schatz, G. C.; Siebbeles, L. D. A.; Ratner, M. A. *J. Am. Chem. Soc.* **2009**, *131*, 14204−14205.
(23) Zhang, L. G.; Zhu, H. H.; Sajimon, M. C.; Stutz, J. A. R.; Siegmund, K.; Richert, C.; Shafirovich, V.; Lewis, F. D. *J. Chin. Chem. Soc.* **2006**, *53*, 1501−1507.
(24) Didraga, C.; Knoester, J. *J. Lumin.* **2003**, *102*, 60−66.
(25) Linnanto, J. M.; Korppi-Tommola, J. E. I. *Photosynth. Res.* **2008**, *96*, 227−245.
(26) Markovitsi, D.; Germain, A.; Millié, P.; Lécuyer, P.; Gallos, L. K.; Argyrakis, P.; Bengs, H.; Ringsdorf, H. *J. Phys. Chem.* **1995**, *99*, 1005−1017.
(27) Patwardhan, S.; Sengupta, S.; Würthner, F.; Siebbeles, L. D. A.; Grozema, F. C. *J. Phys. Chem. C* **2010**, *114*, 20834−20842.
(28) Spano, F. C. *J. Am. Chem. Soc.* **2009**, *131*, 4267−4278.
(29) Bouvier, B.; Dognon, J.; Lavery, R.; Markovitsi, D.; Millié, P.; Onidas, D.; Zakrzewska, K. *J. Phys. Chem. B* **2003**, *107*, 13512−13522.
(30) Emanuele, E.; Markovitsi, D.; Millié, P.; Zakrzewska, K. *Chem. Phys. Chem.* **2005**, *6*, 1387−1392.
(31) Emanuele, E.; Zakrzewska, K.; Markovitsi, D.; Lavery, R.; Millié, P. *J. Phys. Chem. B* **2005**, *109*, 16109−16118.
(32) de Vries, M.; Hobza, P. *Annu. Rev. Phys. Chem.* **2007**, *58*, 585−612.
(33) Nir, E.; Grace, L.; Brauer, B.; de Vries, M. *J. Am. Chem. Soc.* **1999**, *121*, 4896−4897.
(34) Nir, E.; Imhof, P.; Kleinermanns, K.; de Vries, M. S. *J. Am. Chem. Soc.* **2000**, *122*, 8091−8092.
(35) Chin, W.; Mons, M.; Dimicoli, I.; Piuzzi, F.; Tardivel, B.; Elhanine, M. *Eur. Phys. J. D* **2002**, *20*, 347−355.
(36) Choi, M.; Miller, R. *J. Am. Chem. Soc.* **2006**, *128*, 7320−7328.
(37) Mons, M.; Dimicoli, I.; Piuzzi, F.; Tardivel, B.; Elhanine, M. *J. Phys. Chem. A* **2002**, *106*, 5088−5094.
(38) Mons, M.; Piuzzi, F.; Dimicoli, I.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2006**, *110*, 10921−10924.
(39) Nir, E.; Janzen, C.; Imhof, P.; Kleinermanns, K.; de Vries, M. S. *J. Chem. Phys.* **2001**, *115*, 4604−4611.
(40) Zhou, J.; Kostko, O.; Nicolas, C.; Tang, X. N.; Belau, L.; de Vries, M. S.; Ahmed, M. *J. Phys. Chem. A* **2009**, *113*, 4829−4832.
(41) Chen, H.; Li, S. *J. Phys. Chem. A* **2006**, *110*, 12360−12362.
(42) Lewis, F. D.; Wu, T.; Liu, X.; Letsinger, R. L.; Greenfield, S. R.; Miller, S. E.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2000**, *122*, 2889−2902.
(43) Lewis, F. D.; Wu, Y.; Liu, X. *J. Am. Chem. Soc.* **2002**, *124*, 12165−12173.
(44) Lewis, F.; Wu, Y.; Zhang, L.; Zuo, X.; Hayes, R.; Wasielewski, M. *J. Am. Chem. Soc.* **2004**, *126*, 8206−8215.
(45) Lewis, F.; Zhang, L.; Liu, X.; Zuo, X.; Tiede, D.; Long, H.; Schatz, G. *J. Am. Chem. Soc.* **2005**, *127*, 14445−14453.
(46) Molina, V.; Merchán, M.; Roos, B. O. *J. Phys. Chem. A* **1997**, *101*, 3478−3487.
(47) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691−8696.
(48) Serrano-Perez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B* **2007**, *111*, 11880−11883.
(49) Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108−8109.
(50) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *J. Am. Chem. Soc.* **2008**, *130*, 2473−2484.

(51) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2309−2309.
(52) Lewis, F. D. *Pure Appl. Chem.* **2006**, *78*, 2287−2295.
(53) Holman, M. R.; Ito, T.; Rokita, S. E. *J. Am. Chem. Soc.* **2007**, *129*, 6−7.
(54) Kundu, L. M.; Linne, U.; Marahiel, M.; Carell, T. *Chem.—Eur. J.* **2004**, *10*, 5697−5705.
(55) Mccullagh, M.; Hariharan, M.; Lewis, F. D.; Markovitsi, D.; Douki, T.; Schatz, G. C. *J. Phys. Chem. B* **2010**, *114*, 5215−5221.