

Circular Dichroism Spectra of DNA Hairpins Studied by the Green Function Method

D. Balamurugan,[†] Frederick D. Lewis,[‡] and Alexander L. Burin^{*,†}*Department of Chemistry, Tulane University, New Orleans, Louisiana 70118, and Department of Chemistry, Northwestern University, Evanston, Illinois 60208-3113**Received: July 4, 2006; In Final Form: February 12, 2007*

The circular dichroism (CD) spectrum of a large biological molecule represented as an array of interacting chromophores is investigated. The configuration-averaged Green function formalism is developed to describe the CD and absorption spectra. The perturbation theory expansion is derived for absorption and CD spectra in the case of strong interaction of chromophores with their environment (solvent and/or internal dynamics) compared to their interaction with each other. We apply this formalism to study CD spectra of DNA hairpins.

I. Introduction

Optical properties of biological macromolecules (proteins, DNA) are sensitive to molecular structure and function in living cells.^{1–11} Therefore, the optical methods, including infrared¹² and ultraviolet¹³ photoabsorption, circular dichroism (CD),^{2,14} hole burning spectroscopy,¹⁵ and fluorescence resonance energy transfer (FRET),^{16–18} are widely used to study biological systems. Optical methods have an advantage compared to other methods (e.g., X-rays or mass spectrometry) that they permit the investigation of biological molecules under native conditions of a living cell. On the other hand, in contrast with X-rays they do not provide direct information about molecular structure. Therefore, results of optical investigation of a biological molecule depend crucially on a theoretical model used to interpret the data.

In this paper we perform a theoretical analysis of a CD spectrum in the regime of a strong interaction of molecular excitations with their environment. Circular dichroism (CD) spectroscopy is a very powerful method compared to other optical studies. In contrast with absorption it provides a direct information about the relative orientation of chromophore groups within the large molecule.^{1–3,5,14} This allows time-resolved measurements of structural and conformational changes of the molecule possibly during its participation in biological processes. In contrast with FRET, the CD method does not need fluorescent chromophores. For molecules like DNA the main contribution to a CD spectrum comes from the interaction of base pairs with each other rather than the individual base contributions. Therefore, a CD spectrum is very sensitive to the DNA structure. Recent CD investigation of capped DNA hairpins¹⁴ has demonstrated a relevance of this method to define a relative orientation of chromophores in hairpins consisting of up to 11 base pairs with end-capping stilbene dicarboxamide (Sa) groups (see Figure 1). Only poly-A–poly-T base pair sequences will be considered in our paper because instead of exciton formation the charge separation takes place in sequences containing GC pairs due to the guanine oxidation.¹⁴

The theory of circular dichroism expresses its intensity through its rotational strength given by the dot product of the

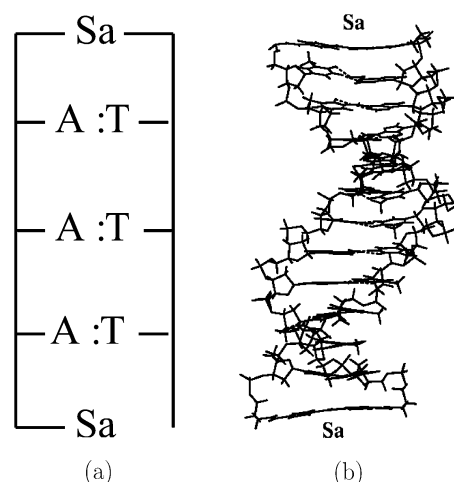


Figure 1. Schematic and molecular models of stilbene dicarboxamide (Sa) end-capped DNA hairpins containing n (A–T) base pairs: (a) schematic model with $n = 3$ AT base pairs and (b) molecular model with $n = 11$ pairs. In the molecular model, Sa chromophores are labeled and AT base pairs are not labeled for clarity.

transition dipole moment μ and the transition magnetic moments m

$$\text{CD} \propto \mathbf{m} \cdot \mu \quad (1)$$

On the basis of this definition, it is possible to calculate the CD intensity for the given molecular transition. Numerical implementation of the electronic structure theory is capable of calculating the rotational strength eq 1 for a variety of molecules.¹⁹ However, this numerical approach is good for small molecules, while in molecules like DNA the main contribution to a CD spectrum comes from the interaction between different base pairs (chromophores). Indeed, each DNA base pair has an essentially planar structure. The transition dipole moment of a planar molecule belongs to its plane, while its magnetic moment is perpendicular to the plane. Therefore, the circular dichroism vanishes for planar base pairs because $\mathbf{m} \cdot \mu = 0$ (eq 1), and only base pair interactions can lead to a nonzero effect. The advantage of the system is that the chromophore interaction contains information about the molecular structure.²

Arrays of interacting chromophores contain a very large number of atoms, so they cannot be studied directly using ab

* Corresponding author. E-mail: aburin@tulane.edu.

[†] Tulane University.

[‡] Northwestern University.

initio methods,¹⁹ and therefore the separation of a molecule into an array of weakly coupled chromophores is necessary to make an adequate approach to the problem. In addition, this approximation¹⁹ ignores the interaction with the solvent, while our study takes it explicitly into account.

To our knowledge all existing approaches initially ignore the chromophore–environment interaction adding it as an artificial spectrum broadening effect at the very end of the consideration.^{2,5,13} Such approaches are well justified, when the interaction w with the environment is weak compared to the interaction V between chromophores

$$w \ll V \quad (2)$$

In particular, our approximation is different from the one developed by Woody and co-workers.⁵ We do not make the Gaussian expansion for the CD spectrum shape but make such an assumption for the absorption spectrum only. Our theory predicts that the CD spectrum shape is different from the Gaussian shape assumed for the absorption only. This shape is expressed through the Gaussian spectrum using the principal value of the singular integral (see eq 33). We believe that this approach is more justified, when the perturbation theory with respect to the weak exciton coupling is applicable.

Biological molecules always have a highly polar environment of water having a dielectric constant $\epsilon = 80$. Therefore, there is no reason to expect chromophore–solvent interactions to be weak. Since the electronic interaction of two Sa groups in DNA hairpin decreases rapidly with the number n of DNA base pairs separating them (as $1/n^3$, see Figure 1; cf. eq 5), it is clear that for sufficiently large n the interaction between two Sa chromophores becomes small compared to any reasonable interaction with the environment. In this regime we can no longer treat chromophores as being identical, and their interaction as well as their CD spectrum is essentially off-resonant in contrast with the approach taken in refs 2 and 13. Therefore, the case of the strong environmental interaction requires the use of a different theoretical method. Below we develop the general analysis of molecules possessing arbitrarily strong chromophore–environment interactions employing the configuration-averaged Green function technique. This method has been developed to describe excitations in disordered systems,^{20,21} where it allows a straightforward calculation of physical values of interest (e.g., electrical or thermal conductivity) averaged over various configurations of system disorder. In our case averaging will be performed over different configurations of an Sa group environment, leading to a wide dispersion of chromophore excitation energies and a broad absorption spectrum.

In this paper the method of configuration-averaged Green functions is used to calculate CD spectra of capped DNA hairpins^{14,13} (Figure 1). Disorder is characterized using the absorption spectrum within the same frequency range. The hairpins considered in the present study have 1–4 AT base pairs (cf. Figure 1). For a larger number of base pairs the DNA–Sa interactions can no longer be ignored. This consideration is beyond the scope of this paper. Since there is no remarkable difference between absorption and CD spectra of hairpins and dumbbells,¹⁴ the results are applicable also to dumbbells.²³ Our results show that CD spectra of hairpins calculated using the configuration-averaged Green function method are in good agreement with experiment, and our method thus permits the characterization of the shape and the strength of a CD spectrum.

The paper is organized as follows. In section II, we derive the general expression for a CD spectrum of interacting

chromophores coupled to their environment using the configuration-averaged Green function technique. In section III, this method is employed to calculate the CD spectrum of Sa–AT_{*n*}–Sa ($n = 1$ –4) hairpins in the spectral domain of Sa group absorption. Finally, conclusions are given.

II. Absorption and CD Spectra of Interacting Chromophores

A. Model. Optical properties of a molecule are defined by its electronic excitations. To study these excitations in a large molecule, it is convenient to separate it into smaller parts called chromophores, which are coupled by a relatively weak interaction. For instance, in DNA each chromophore can be represented by a DNA base pair.

Optical excitations of a biological molecule containing N interacting chromophores can be described by the Frenkel exciton Hamiltonian in the Heitler–London approximation (see e.g. refs 2, 6, and 22):

$$\hat{H} = \sum_{i=1}^N \epsilon_i n_i + \frac{1}{2} \sum_{i \neq j}^N V_{ij} (a_i^\dagger a_j + a_j^\dagger a_i) \quad (3)$$

Here indices i and j enumerate chromophores, and $n_i = a_i^\dagger a_i$ stands for the population of the excited-state of the chromophore i , where $n_i = 0$ in the ground state and $n_i = 1$ in the excited state. Operators a_i^\dagger and a_i describe a creation and an annihilation of the excited state of the chromophore i , whose excitation energy is ϵ_i . The excitation energy ϵ_i depends on the particular chromophore and on environment-induced fluctuations of the chromophore energy. These fluctuations are caused by both internal fluctuations of the molecular structure (vibrations, etc.) and fluctuations of the solvent configuration in the vicinity of the chromophore. In this paper we treat all fluctuations classically for the sake of simplicity although the vibrations of light atoms under certain conditions can require quantum mechanical analysis. We assume that energy fluctuations are described by the joint distribution function

$$P(\epsilon_1, \epsilon_2, \dots, \epsilon_N) \quad (4)$$

The term V_{ij} in eq 3 stands for the resonant interaction of chromophores with each other leading to sharing the optical excitation between them. We assume that V_{ij} is a dipole–dipole interaction, which can be expressed in terms of the transition dipole moments μ_i, μ_j , and a vector \mathbf{R}_{ij} connecting them as

$$V_{\text{dip}} = \frac{\mu_i \mu_j - 3(\mathbf{n} \mu_i)(\mathbf{n} \mu_j)}{\epsilon_m R_{ij}^3}, \quad \mathbf{n} = \frac{\mathbf{R}_{ij}}{R_{ij}} \quad (5)$$

where ϵ_m is the medium optical dielectric constant taken at the frequency equal to the photon frequency $\omega = E/c$. Equation 5 has been successfully used in ref 13 to describe the CD intensity in capped DNA hairpins.

Collective exciton modes of the molecule having a given environment configuration $\{\epsilon_i\}$, $i = 1, \dots, N$ can be found solving a stationary Schrödinger equation

$$E_\alpha c_{\alpha i} = \epsilon_i c_{\alpha i} + \sum_j V_{ij} c_{\alpha j} \quad (6)$$

for a single excitation wave function

$$|\alpha\rangle = \sum_i c_{\alpha i} |i\rangle \quad (7)$$

The notation $|i\rangle$ describes the molecular state where the i th chromophore is excited ($n_i = 1$), while other $N - 1$ chromophores are in their ground states. We will consider the interaction of light with the molecule as the superposition of interactions with N independent collective states $|\alpha\rangle$ characterized by their transition dipole moments

$$\mu_\alpha = \sum_i c_{\alpha i} \mu_i \quad (8)$$

B. Absorption and CD Spectra. The interaction of a large chiral molecule (like DNA) with left- and right-circularly polarized light is different. For instance, the molecule having a right-handed helix shape (cf. Figure 1) might interact more strongly with right-circularly polarized light than with left-circularly polarized light because transition dipole moments of chromophores rotate along the molecule in correlation with the light polarization. This difference results in circular dichroism in the differential absorption of left- and right-circularly polarized light.

A brief derivation of the circular dichroism spectrum in terms of the Green functions of the model eqs 3 and 4 can be given as follows. We assume that our molecule is made of almost fully planar chromophores (like DNA). An isolated planar chromophore cannot by itself contribute to the CD spectrum because its transition magnetic moment \mathbf{m} is perpendicular to the chromophore plane while its transition dipole moment μ belongs in the plane. Therefore, their dot product, which defines the rotation strength (eq 1, ref 2), is zero. Under this assumption one can ignore the dispersion of the electric field within the single chromophore and the interaction Hamiltonian of the molecule with the resonant left- or right-circularly polarized light propagating along the z -axis can be represented as

$$V_E^{l,r} = -F_0 \sum_{j=1}^N \mu_i (a_i^+ + a_i) \times (\mathbf{e}_x \cos(\mathbf{q} \cdot \mathbf{R}_i - \omega t + \rho) \pm \mathbf{e}_y \sin(\mathbf{q} \cdot \mathbf{R}_i - \omega t + \rho)) \quad (9)$$

where \mathbf{q} is the wavevector of light and \mathbf{R}_i is the coordinate of the center of the chromophore i (we assume that the origin $\mathbf{R} = 0$ is located in the center of the whole molecule). The phase factor φ is a random number between 0 and 2π , and F_0 stands for the external electric field amplitude of the light wave. One can always assume that the size of the molecule is much smaller than the optical wavelength so that $qR_i \ll 1$. The difference in absorption of the right- and left-circularly polarized lights shows up in the first order in qR . Therefore, in order to investigate the circular dichroism, one should expand the interaction Hamiltonian eq 9 with respect to the parameter qR . This yields

$$V_E^{l,r} = -F_0 \sum_{i=1}^N \mu_i (a_i^+ + a_i) \times (\mathbf{e}_x (\cos(-\omega t + \varphi) - (\mathbf{q} \cdot \mathbf{R}_i) \sin(-\omega t + \varphi)) \pm (\mathbf{e}_x (\sin(-\omega t + \varphi) + (\mathbf{q} \cdot \mathbf{R}_i) \cos(-\omega t + \varphi))) \quad (10)$$

The absorption of light by the sample is usually characterized by the absorbance coefficient α describing the decrease in a transmitted light intensity with the sample thickness t ($I_{tr} \sim e^{-\alpha t}$). This coefficient can usually be very well approximated by the kinetic formula²⁴

$$\alpha = n\sigma \quad (11)$$

where n is the density of absorbing molecules and σ is the cross section for absorption of light by the single molecule. The

absorption cross section σ is defined as $\sigma = I/j$, where I is the rate of the energy absorption by the single molecule and j is the energy flux of the incident wave. The energy absorption rate can be expressed using the Fermi golden rule as²⁴

$$I_{l,r}(E) = E \frac{\pi}{3\hbar} \sum_\alpha \mu_\alpha^2 F_0^2 \delta(E_\alpha - E) \quad (12)$$

The energy flux j is given by the absolute value of the Poynting vector $j = \sqrt{\epsilon_s} c F_0^2 / (4\pi)$, where $\epsilon_s \sim 1.78$ is a solvent (water) dielectric constant at optical frequencies. The expression for the energy absorption rate is written in zeroth order in a small parameter qR_{ij} , and it does not depend on the light polarization. Thus, we get²⁴

$$\alpha = n\sigma = E \frac{4\pi^2}{3\sqrt{\epsilon_s} \hbar c} \sum_\alpha \mu_\alpha^2 \delta(E_\alpha - E) \quad (13)$$

It is convenient to express the absorption using Green functions of exciton, which are defined as

$$g_{ij}(E) = \sum_{\alpha=1}^N \frac{c_{\alpha i}^* c_{\alpha j}}{(E_\alpha - E - i\delta)}, \quad \delta \rightarrow 0 \quad (14)$$

Using the definition of the transition dipole moment (eq 8) the identity $\text{Im}(1/(x - i\delta)) = \pi\delta(x)$ at $\delta \rightarrow 0$ (see e.g. refs 21 and 25), and assuming wave functions $c_{\alpha i}$ to be real, one can express the absorption coefficient (eq 13) as

$$\alpha = n\sigma = \frac{4\pi n E}{3\sqrt{\epsilon_s} \hbar c} \sum_{i,j} \mu_i \mu_j \text{Im}(g_{ij}(E)) \quad (15)$$

In the first order in qR one can obtain the correction term to the absorption spectrum, which results in the circular dichroism, i.e., the difference in absorptions of right- and left-circularly polarized light waves (eq 10). This term defines the efficiency of the circular dichroism for a single molecule

$$\text{CD}(E) = \alpha_l(E) - \alpha_r(E) = \frac{4\pi n E^2}{\hbar^2 c^2} \sum_{i,j} (\mathbf{e}_z \mathbf{R}_{ij}) \times ((\mu_i \mathbf{e}_x)(\mu_j \mathbf{e}_y) - (\mu_j \mathbf{e}_x)(\mu_i \mathbf{e}_y)) \text{Im}(g_{ij}(E)) \quad (16)$$

where we have expressed the wavevector q through the photon energy E as $q = \sqrt{\epsilon_s} E / (c\hbar)$.

One should notice that eq 16 has been already averaged over the random phase ϕ of the light field (see eq 9). Averaging of this result over the random angle between the electric field and the molecular orientation can be made reexpressing eq 16 as

$$\text{CD}(E) = \frac{4\pi n E^2}{\hbar^2 c^2} \sum_{i,j} (\mathbf{e}_z \mathbf{R}_{ij}) (\mathbf{e}_z \cdot (\mu_i \times \mu_j)) \text{Im}(g_{ij}(E))$$

Then averaging over the random orientation of the molecule with respect to the light propagation direction, \mathbf{e}_z can be made using the identity $\langle (\mathbf{e}_z \mathbf{a})(\mathbf{e}_z \mathbf{b}) \rangle = \mathbf{ab}/3$, and it leads to the factor 1/3 in the final expression

$$\text{CD}(E) = \frac{4\pi n E^2}{3\hbar^2 c^2} \sum_{i,j} \text{Im}(g_{ij}(E)) \mathbf{R}_{ij} \cdot (\mu_i \times \mu_j) \quad (17)$$

TABLE 1: Fit Parameters Used in the Gaussian Function Superposition $P(E) = \sum_{i=1}^N [A_i E/w_i (2\pi)^{1/2}] \exp[-(E - \epsilon_{0i})^2/2w_i^2]$

n	1	2	3	4
ϵ_0	3.74	3.70, 4.13	3.51, 3.73, 4.16	3.49, 3.65, 3.85, 4.36
w	0.206	0.19, 0.34	0.06, 0.18, 0.35	0.06, 0.11, 0.19, 0.38
A	2.60	2.37, 1.22	0.25, 2.17, 1.16	0.31, 0.85, 1.82, 0.84

The Green function $g_{ij}(E)$ is computed for the given sequence of chromophore energies $\{\epsilon_i\}$, $i = 1, 2, \dots, N$, so it is more appropriate to express it as $g_{ij}(E, \{\epsilon_i\})$. To find the integrated response of the large number of molecules, one should average this function over various environment configurations (fluctuations of ϵ_i). The configurationally averaged Green function $G_{ij}(E)$ is defined by the integration of the Green function for the given configuration $\{\epsilon_i\}$ over the distribution of configurations (eq 4)

$$G_{ij}(E) = \langle g_{ij}(E) \rangle$$

$$= \int_{-\infty}^{\infty} d\epsilon_1 \dots \int_{-\infty}^{\infty} d\epsilon_N P(\epsilon_1, \dots, \epsilon_N) g_{ij}(E, \{\epsilon_i\}) \quad (18)$$

Substituting this configurationally averaged Green function into eqs 15 and 16, we get

$$\alpha = \frac{4\pi n E}{3\sqrt{\epsilon_s} \hbar c} \sum_{i,j} \text{Im}(G_{ij}(E)) \mu_i \mu_j \quad (19)$$

for the absorption spectrum and

$$\text{CD}(E) = \frac{4\pi n E^2}{3\hbar^2 c^2} \sum_{i,j} \text{Im}(G_{ij}(E)) \mathbf{R}_{ij} \cdot (\mu_i \times \mu_j) \quad (20)$$

for the circular dichroism spectrum.

C. Perturbation Theory Approach. We will use perturbation theory to evaluate the Green function (eq 18) and the absorption and CD spectra (eqs 19 and 20). The Green function $g_{ij}(E)$ (eq 14) satisfies the following equation:

$$(\epsilon_i + i\eta - E)g_{ij}(E) + \sum_{k \neq i} V_{ik} g_{kj}(E) = \delta_{ij}$$

$$\eta \rightarrow 0^+ \quad (21)$$

Rearranging the terms in eq 21, one can formally express the Green function as

$$g_{ij}(E) = \frac{\delta_{ij}}{E - \epsilon_i - i\eta} + \sum_k \frac{V_{ik}}{E - \epsilon_i - i\eta} g_{kj}(E) \quad (22)$$

As was discussed in the Introduction, we assume that the width of the distribution of fluctuating energy exceeds the interaction V between different chromophores. Indeed, according to our analysis (see Table 1) of the Sa spectrum broadening w by the solvent, one can estimate the strength of the chromophore–solvent interaction using the width of Gaussian fit $w > 0.05$ eV for all absorption lines. The interaction of two Sa chromophores can be calculated using the Sa group transition dipole moment estimate $\mu \sim 5$ D made using the time-dependent density functional theory.^{19,27} The strongest interaction between Sa groups takes place in Sa–AT–Sa system where the separation of two Sa chromophores can be estimated as 0.68 nm. Then even ignoring screening, one can estimate the dipolar interaction of two Sa groups as 0.05 eV, which is still smaller than the characteristic interaction with the solvent.

Then we can restrict our study to the first nonvanishing contribution in V , i.e., the first order in V . As we will see below, the zeroth-order term diagonal in chromophore indices i and j defines the absorption of light, while it does not contribute to the CD spectrum. The expansion to the first order in V can be obtained taking the zeroth-order approximation $g_{ij}^R \approx \delta_{ij}/(E - \epsilon_i - i\eta)$ and substituting it to the right-hand side of eq 22. This yields

$$g_{ij}(E) \approx \frac{\delta_{ij}}{E - \epsilon_i - i\eta} + \frac{V_{ij}}{(E - \epsilon_i - i\eta)(E - \epsilon_j - i\eta)} \quad (23)$$

In the next step we find the imaginary part of the retarded Green function (eq 23) and average it over random energies ϵ_i . For this purpose we introduce the distributions $P_i(\epsilon_i)$ and the joint distribution $P_{ij}(\epsilon_i, \epsilon_j)$ for chromophores i and j , which can be obtained integrating the distribution (eq 4) over all remaining energies. Then the configurationally averaged Green function can be represented as

$$\text{Im}G_{ij}(E) \approx \pi \delta_{ij} P_i(E) + \pi V_{ij} \text{PV} \left(\int_{-\infty}^{+\infty} d\epsilon_i \frac{P_{ij}(\epsilon_i, E)}{E - \epsilon_i} + \int_{-\infty}^{+\infty} d\epsilon_j \frac{P_{ij}(E, \epsilon_j)}{E - \epsilon_j} \right) \quad (24)$$

where PV stands for the principal value of the singular integrals.

The first term in eq 24 contributes to the absorption, which can be evaluated using the definition provided in eq 19 as

$$\alpha(E) = \frac{4\pi n E}{3\sqrt{\epsilon_s} \hbar c} \sum_{i,j=1}^N \mu_i \mu_j \text{Im}(G_{ij}(E)) \approx \frac{4\pi^2 n E}{3\sqrt{\epsilon_s} \hbar c} \sum_{i,j=1}^N \mu_i^2 P_i(E) \quad (25)$$

This expression establish the relationship between distributions of single chromophore absorption energies $P_i(E)$ (for chromophore i) and the absorption spectrum.

The second, off-diagonal term in the Green function (eq 24) contributes to the CD spectrum. This contribution can be finally expressed as

$$\text{CD} \approx \frac{8\pi^2 n E^2}{3\hbar^2 c^2} \sum_{i,j} \mathbf{R}_{ij} \cdot (\mu_i \times \mu_j) V_{ij} \text{PV} \left(\int_{-\infty}^{+\infty} d\epsilon_i \frac{P_{ij}(\epsilon_i, E)}{E - \epsilon_i} + \int_{-\infty}^{+\infty} d\epsilon_j \frac{P_{ij}(E, \epsilon_j)}{E - \epsilon_j} \right) \quad (26)$$

III. Low-Energy CD Spectrum of Sa-End-Capped DNA Hairpins

A. CD Spectrum of DNA Hairpins at “Low” Energies. In this section we use the Green function formalism to study CD spectra of Sa-end-capped DNA hairpins (Figure 1). Absorption of light by the DNA hairpins is caused by both AT pairs and Sa groups. However, the absorption of DNA is noticeable only in the domain of high energies $E > 4.3$ eV, while only Sa groups absorb light in the energy domain between 3.3 and 4.3 eV (see Figure 2). It is easier to model the absorption and CD spectra in this low-energy domain where one can ignore the DNA contribution. Also, perturbation theory is relevant for Sa groups because they are separated from each other by the AT bridge, so that their interaction is really weak. Therefore, in this paper we restrict our consideration to absorption and circular dichroism

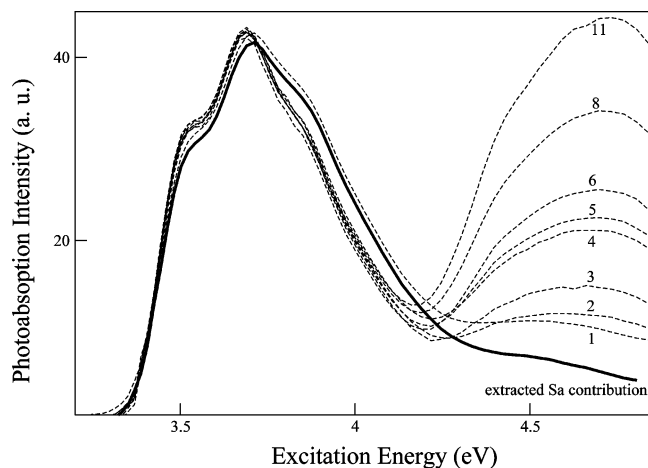


Figure 2. Experimental photoabsorption spectrum of the Sa-(AT)_n-Sa ($n = 1-6, 8, 11$) hairpins^{13,14} and the extrapolated Sa absorption spectrum (solid line, see text for details). The Sa-(AT)_n-Sa hairpin spectrum is labeled by the number of AT base pairs n .

in the low-energy domain of Sa absorption and ignore the DNA contribution.

We will use our theory to interpret experimental data for both absorption and circular dichroism spectra. These data include the shapes of both spectra and the relation between them. Our analysis starts with expression of both spectra in terms of the distribution function $P(E)$ of absorption energies for the Sa group. The distribution $P(E)$ is caused by fluctuations of hairpin environment. Using this expression in eq 25 and taking into account that we have two Sa chromophores, one can express the absorption spectrum as

$$\alpha(E) \approx \frac{8\pi^2 n E \mu^2 P(E)}{3\sqrt{\epsilon_s} \hbar c} = CEP(E)$$

$$C = \frac{8\pi^2 n \mu^2}{3\sqrt{\epsilon_s} \hbar c} \quad (27)$$

where $\mu \sim 5$ D is the transition dipole moment of the Sa chromophore and the constant C is introduced to separate the spectrum shape term from the term expressing the quantitative spectral intensity.

Similarly, one can evaluate the CD spectrum. Since two Sa chromophores are identical, we can use the symmetry property $P(\epsilon_1, \epsilon_2) = P(\epsilon_2, \epsilon_1)$ for the distribution of fluctuations of their energies and simplify eq 26 to the form

$$CD \approx \frac{16\pi^2 n E^2}{3\hbar^2 c^2} V_{12} \mathbf{R}_{12} \cdot (\mu_1 \times \mu_2) PV \int d\epsilon \frac{P(E, \epsilon)}{E - \epsilon} \quad (28)$$

Moreover, Sa chromophores are located far from each other and hence their configurations defined by the probability function $P(\epsilon_1, \epsilon_2)$ are weakly correlated. Therefore, we set

$$P_2(\epsilon_1, \epsilon_2) = P(\epsilon_1) P(\epsilon_2) \quad (29)$$

where $P(\epsilon_1)$ is the distribution of excitation energies for a single chromophore (see eq 27). Using this single chromophore distribution, the CD spectrum can be reexpressed as

$$CD(E) \approx \frac{16\pi^2 n E^2}{3\hbar^2 c^2} V_{12} \mathbf{R}_{12} \cdot (\mu_1 \times \mu_2) P(E) PV \int d\epsilon \frac{P(\epsilon)}{E - \epsilon} \quad (30)$$

For future analysis it is convenient to express a CD spectrum as the product of independent parts to be compared to the experiment separately. One can rewrite it as

$$CD(E) = C_1 S(\theta_n, R_n) I_{CD}(E)$$

$$C_1 = \frac{8\pi^2 n \mu^4}{3\sqrt{\epsilon_s} \hbar^2 c^2 a^2} \quad (31)$$

where the constant C_1 is the common adjustable factor between theory and experiment and $a = 3.4$ Å is the period of the DNA sequence. The factor $S(\theta_n, R_n)$ expresses the rotational strength dependence on the distance R_n between Sa chromophores and their twisting angle θ_n (cf. eq 5)

$$S(R, \theta) = \frac{\sqrt{\epsilon_s} 2 V_{12} a^2 \mathbf{R}_{12} \cdot (\mu_{01} \times \mu_{02})}{\epsilon_m \mu^4} = \frac{\sqrt{\epsilon_s} a^2 \sin(2\theta)}{\epsilon_m R^2} \quad (32)$$

Note that this expression is valid only when both transition dipole moments μ_{01} and μ_{02} belong to the parallel planes perpendicular to the vector connecting two chromophores. Also, only dipole-dipole interaction is considered. In the final expression we omitted the dependence of the spectrum on the absolute value of the transition dipole moments, assuming them to be constants and including them into the adjustable factor. We believe that environmental interaction is much weaker than the internal interaction within the chromophore so it cannot change transition dipole moments significantly.

The last term expresses the shape of the CD spectrum as

$$I_{CD}(E) = E^2 P(E) PV \int d\epsilon \frac{P(\epsilon)}{E - \epsilon} \quad (33)$$

In sections III.B and III.C we compare the shape of the absorption and CD spectra, while in section III.D the factor $S(R, \theta)$ responsible for the CD rotational strength (eq 32) is discussed. Finally, in section III.E we compare the theoretical estimate for the dimensionless ratio of CD absorption intensities expressed through the ratio of constants C_1 and C (eqs 27 and 31)

$$\frac{C_1}{C} = \frac{\mu^2}{a^2 \hbar c} \approx 0.72 \times 10^{-3} \quad (34)$$

This ratio is much less than unity because the circular dichroism is induced by the weak relativistic interaction of order of $e^2/(\hbar c)$.

B. Definition of the Energy Distribution Function in Terms of the Absorption Spectrum. To describe the CD spectrum shape of DNA hairpins (eq 33), one has to define the distribution $P(\epsilon)$ for fluctuations of the excitation energy ϵ of Sa chromophores. It is important that within our approach this distribution defines the intensity of absorption by the stilbene chromophore (eq 27). Therefore, one can replace the value $EP(E)$ with the experimental absorption intensity for Sa groups (Figure 2).

The absorption measured experimentally contains not only the Sa contribution but also the DNA contribution. We need to extract the Sa contribution alone. To get rid of the DNA contribution, one can use the experimental absorption data¹⁴ for DNA hairpins with $n = 1, 2, 3, \dots, 11$ base pairs. Then the experimental absorption spectrum $\alpha_{\text{tot}}(E, n)$ can be approximately represented as the sum of two contributions from Sa end groups and n contributions from AT pairs

$$\alpha_{\text{tot}}(E, n) = 2\alpha_{\text{Sa}}(E) + n\alpha_{\text{AT}}(E) \quad (35)$$

In our study the AT pair absorption has been taken as $\alpha_{\text{AT}}(E) = (I_{\text{tot}}(E, 11\text{-mer}) - I_{\text{tot}}(E, 8\text{-mer}))/3$. Maximum available values of n were chosen in order to minimize the end effect. Then the average Sa group absorption has been expressed as

$$\alpha_{\text{Sa}}(E) = \frac{\sum_n (\alpha_{\text{tot}}(E, n) - n\alpha_{\text{AT}}(E))}{2 \cdot 8} \quad (36)$$

where the sum in the numerator of eq 36 is taken over all eight experimental values of n (see Figure 2). Modifications of this extrapolation procedure do not change our result remarkably. The extrapolated absorption spectrum of an Sa group is shown in Figure 3.

We have performed a Gaussian fit of the Sa absorption spectrum with the DNA background employing up to four Gaussian functions. Use of the Gaussian fit is in the spirit of the Marcus classical theory²⁶ (see also the discussion by Tinoco and co-workers¹) of a solvent interaction with electronic transitions. The parameters of the fit are defined in Table 1. The Gaussian fits are shown in Figure 3 together with the Sa absorption spectrum. We see from Figure 3 that the fit using a single Gaussian function accounts for the Sa absorption spectrum only near its maximum but cannot describe the tails. The fit by two Gaussians looks much better. The approximation of the absorption spectrum using 3 or 4 Gaussian functions gives some minor improvement in approaching the experimental data, including the weak shoulders seen near the first maximum at $E \approx 3.5$ eV. Considering the fit of the absorption spectrum, one can distinguish two contributions characterized by energies $E_1 \sim 3.7$ eV and $E_2 \sim 4.3$ eV. The first contribution has an intensity about twice as large as the second one, and the width of the first contribution is about 2 times smaller than that of the second one (see Table 1). The increase of the number of Gaussians used to fit the data is equivalent to some improvement in the description of the first low-energy maximum. The nature of this complicated absorption spectrum is not clear. Possibly there exist two configurations of the Sa group characterized by different excitation energies and different coupling strengths of this excitation to the solvent. These two configurations can be responsible for two different contributions to the absorption. Additional structural features of the low-energy peak obtained using three and four Gaussians can be due to electron–vibration interactions.

C. Investigation of the CD Spectrum Shape. Generally, the numerical evaluation of the principal value of any integral has a problem with convergence. This problem can be avoided if $P(\epsilon)$ has a Gaussian form. Then one can use the following identity²⁸ to evaluate the principal value of the integral in the CD expression

$$\text{PV} \int_{-\infty}^{+\infty} d\epsilon \frac{\exp[-b\epsilon^2]}{\epsilon - \epsilon_1} = \sqrt{\frac{\pi}{b}} \int_0^{+\infty} dt \sin[\epsilon_1 t] \exp(-t^2/4b) \quad (37)$$

The latter nonsingular integral can be easily evaluated numerically.

The CD spectra calculated using eqs 30 and 37 with the different number of Gaussian $n = 1, 2, 3, 4$ are shown in Figure 4 together with the experimental spectrum of Sa–(AT)₁–Sa. We see from Figure 4 that the calculated spectrum using a single Gaussian has low-energy band which is positive and high-energy band which is negative. The energy splittings between the

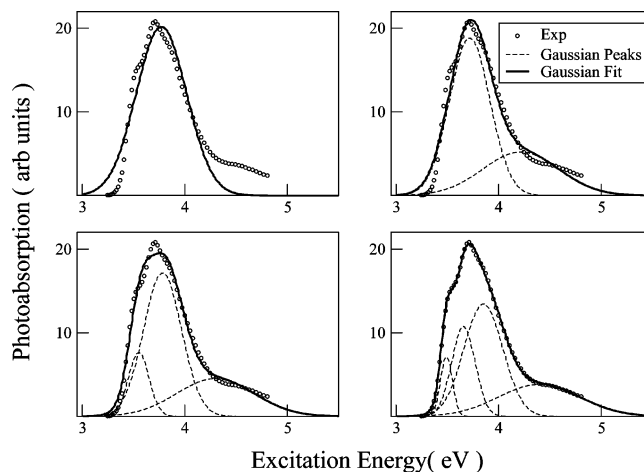


Figure 3. Fit of the photoabsorption spectrum of Sa chromophore using variable number of Gaussian functions. Figure panels from left to right correspond to 1, 2, 3, and 4 Gaussians, respectively (see Table 1 for definitions of all fits). The open circle represents the experimental absorption spectrum of Sa chromophore including the extracted DNA background. The broken curves represent the Gaussian functions used in the fit, and the continuous curve represents the Gaussian fit.

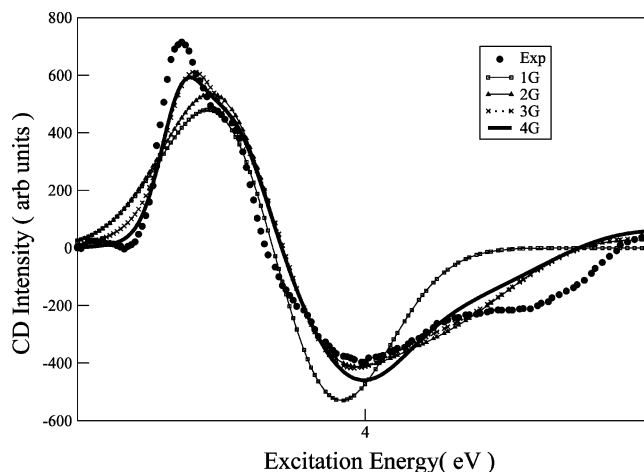


Figure 4. Comparison of the experimental CD spectrum^{13,14} (filled circle) of Sa–AT–Sa with the spectrum calculated using a variable number of Gaussians (G) used to obtain the spectrum is indicated in the figure legend.

positive and negative bands are in a good agreement with the experiment. However, the negative band has stronger intensity than the positive band. This is due to the multiplication factor E^2 in the CD expression (cf. eq 30). This contrasts with the experiment, where the negative band has lower intensity than the positive band. Therefore, the single Gaussian calculation could not completely describe the experimental CD spectrum.

The CD spectrum calculated using $n = 2$ Gaussian functions reproduces the experimental data much better (Figure 4). The shape of the CD spectrum can be used to describe hairpins containing $n = 1$ and 2 AT base pairs, as shown in Figure 5. The proportionality constant C_0S (see eq 31) has been defined from the requirement of the optimum agreement of theoretical and experimental CD spectra at low energies $E < 3.7$ eV. The relationships between the theoretical predictions and the experimental results for this proportionality constant are discussed in section III.D. Here we consider the shape of the spectrum. One should notice the qualitative difference of theoretical and experimental CD spectra for $n > 2$. On our opinion these deviations are caused by the interaction of Sa groups with DNA excitations affecting the CD spectrum. The relative strength of

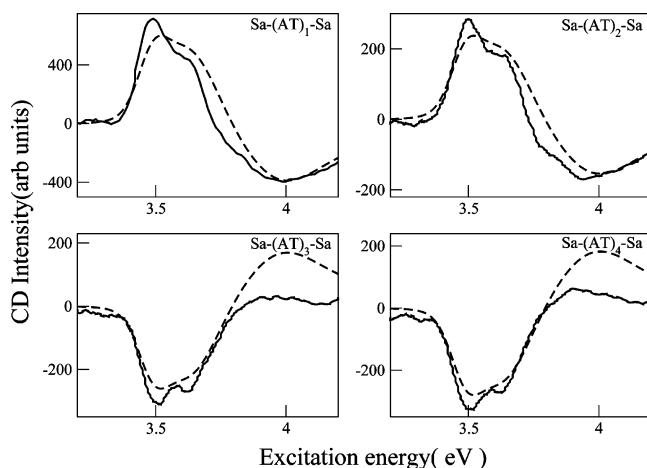


Figure 5. Calculated CD spectra of Sa-(AT)_n-Sa ($n = 1-4$) vs the experimental spectra. The continuous and broken curves represent the experimental and theoretical results, respectively

TABLE 2: Optimized Structure Parameters and MD Structure Parameters¹³ of Sa-(AT)_n-Sa ($n = 1-4$) Hairpins^a

hairpin	our values		MD	
	R	θ	R	θ
Sa-(AT) ₁ -Sa	(7.2)	(33)	7.2 ± 0.3	33 ± 12
Sa-(AT) ₂ -Sa	(11.7)	58	11.7 ± 0.7	59 ± 13
Sa-(AT) ₃ -Sa	(14.3)	120	14.3 ± 0.7	102 ± 18
Sa-(AT) ₄ -Sa	(17.7)	135	17.7 ± 0.6	131 ± 15

^a Values in parentheses were not determined independently but were used in fitting. Angle θ is expressed in degrees.

this interaction increases in comparison with the direct interaction of Sa groups with increasing the distance separating them. Clearly at $n \geq 3$ this interaction cannot be neglected.

D. Analysis of the Rotational Strength. To perform a comparison of the theory predictions for rotational strength eq 32 with the experiment, one should define the angle θ between the transition dipole moments of two Sa groups and the distance r separating them for different number of AT base pairs in the hairpin. All these data including average structural parameters of Sa-AT_n-Sa hairpins and their fluctuations have been reported in a recent molecular dynamics (MD) simulations.¹³ We included them into Table 2.

To fit the experimental data, we have chosen the constant C_1 (eq 31) in the way that the average distance and the twisting angle defined in the MD simulations for the simplest sequence Sa-AT-Sa provides the best fit of the experimental data. Then for other sequences Sa-(AT)_n-Sa with $n = 2, 3, 4$ we used the same constants C_0 and C_1 and define the parameter R_n as the average distance between Sa groups from the MD simulations.¹³ The remaining parameter, i.e., the twisting angle θ_n , has been defined from the requirement of the best agreement of our theory and the experiment. The data used to fit the experiment are shown in Table 2. It is clear that these data agree very well with the MD simulations.

E. Quantitative Relationship of CD and Absorption Intensities. In this section we perform the simple quantitative comparison of our theory with the experiment. The dimensionless ratio of the circular dichroism and absorption intensities at the first maximum of the CD intensity at energy $E \sim 3.5$ eV for the simplest sequence Sa-AT-Sa is given by²⁹

$$\left(\frac{CD(E_{\max})}{\alpha(E_{\max})} \right)_{\text{exp}} \approx 3.62 \times 10^{-4} \quad (38)$$

On the other hand, the theoretical prediction for this ratio can be expressed using eqs 27 and 31 as

$$\left(\frac{CD(E_{\max})}{\alpha(E_{\max})} \right)_{\text{th}} \approx \frac{C_1}{C} \frac{\sqrt{\epsilon_s} a^2 \sin(2\theta)}{\epsilon_m R^2} E_{\max} \text{PV} \int_{-\infty}^{\infty} \frac{P(\epsilon) d\epsilon}{\epsilon - E_{\max}} \quad (39)$$

where the energy E_{\max} is chosen to maximize the integral. Using the simple single Gaussian function approach for $P(\epsilon)$ (see Table 1) and applying eq 37, one can evaluate the overall integral as

$$E_{\max} \text{PV} \int_{-\infty}^{+\infty} \frac{P(\epsilon) d\epsilon}{\epsilon - E_{\max}} \approx 0.77 \frac{E_{\max}}{w} \quad (40)$$

Using the single Gaussian data fit from Table 1 and angular and distance parameters from Table 2, we can express the theoretical estimate of the ratio of CD and absorption intensities at energy corresponding to the maximum CD intensity as

$$\left(\frac{CD(E_{\max})}{\alpha(E_{\max})} \right)_{\text{th}} \approx \frac{2 \times 10^{-3} \sqrt{\epsilon_s}}{\epsilon_m} \quad (41)$$

The agreement between experimental and theoretical estimates (eqs 38 and 41) can be reached if we assume that $\epsilon_m/\sqrt{\epsilon_s} \sim 5$. This is definitely the overestimate for the dielectric constant of DNA, which is of order of 2,³⁰ so that $\epsilon_m/\sqrt{\epsilon_s} \sim 1.5$. Therefore, the theoretical prediction differs from the experiment by the factor of 3. Theory predicts the larger CD intensity than the one observed experimentally. This discrepancy can be due to fluctuations in molecular orientations (see Table 2), which reduces the maximum CD intensity. Also at relatively small distance between two Sa groups in Sa-AT-Sa sequence the dipolar approach can be no longer valid. The interaction with DNA bases can also affect the overall intensity. These problems require more accurate analysis which is beyond the scope of the present paper devoted to the shape of CD spectrum which is not very sensitive to fluctuations and particular exciton interaction mechanism.

IV. Conclusions

In this work we have studied the spectrum of the circular dichroism for large biological molecules in the regime where the interaction of electronic excitations with the environment is stronger than the coupling between these electronic excitations. The molecule has been represented by an array of interacting chromophores, and the interaction between them and the environment was introduced as disordering in chromophore excitation energies. The CD spectrum has been expressed in terms of the Green function in an ensemble of interacting chromophores. For the regime of a weak chromophore coupling compared to the solvent interaction, we have obtained the expression for the CD spectrum using perturbation theory. This result has been applied to the CD spectrum of DNA hairpins in the low-energy domain of Sa group absorption spectra. We have obtained a reasonable interpretation of the experimental data. The generalization of this theory to the circular dichroism at higher energies, where the absorption by DNA is significant, is straightforward and is under development.

Acknowledgment. We thank Xiaobing Zuo, Mark Ratner, and George Schatz for useful discussions and comments. A.B. and D.B. acknowledge David Beratan and Boris Shklovskii for giving us the opportunity to stay and work with their groups at Duke University and the University of Minnesota during the

natural disaster in the city of New Orleans. This work is supported by the NSF CRC program, Grant No. 0628092.

References and Notes

- (1) Tinoco, I. *Adv. Chem. Phys.* **1962**, *4*, 113. Tinoco, I. *J. Am. Chem. Soc.* **1964**, *86*, 297. Tinoco, I.; Bradley, D. F.; Woody, R. W. *J. Chem. Phys.* **1963**, *38*, 1317.
- (2) *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, F. D., Ed.; Plenum Press: New York, 1996; p 738. Lightner, D. A.; Gurst, J. E. *Organic Conformational Analysis and Stereochemistry from Circular Dichroism Spectroscopy*; Wiley-VCH: New York, 2000; 487 pp.
- (3) Goldbeck, R. A.; Sagle, L.; Kim-Shapiro, D. B.; Flores, V.; Kligler, D. S. *Biochem. Biophys. Res. Commun.* **1997**, *235*, 610. Abramavicius, D.; Mukamel, S. J. *J. Chem. Phys.* **2006**, *124*, Art. No. 034113. Abramavicius, D.; Mukamel, S. J. *J. Chem. Phys.* **2005**, *122*, Art. No. 134305.
- (4) (a) Wouters, F. S.; Verveer, P. J.; Bastiaens, P. I. H. *Trends Cell Biol.* **2001**, *11*, 203. (b) Moerner, W. E. *J. Chem. Phys.* **2002**, *117*, 10925.
- (5) Hsu, M.-C.; Woody, R. W. *J. Am. Chem. Soc.* **1971**, *93*, 3515. Liu, Z.; Chen, K.; Ng, A.; Shi, Z.; Woody, R. W.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2004**, *126*, 15141.
- (6) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 2nd ed.; Plenum Publishing Corp.: New York, 1999; p 662. *Circular Dichroism. Principles and Applications*, 2nd ed.; Berova, N.; Woody, R. W.; Nakanishi, K., Eds.; Wiley: New York, 2000.
- (7) Murphy, C. J.; Arkin, M. R.; Jenkins, Y.; Ghatlia, N. D.; Bossman, S. H.; Turro, N. J.; Barton, J. K. *Science* **1993**, *262*, 1025.
- (8) Schuster, G. B.; Landman, U. *Top. Curr. Chem.* **2004**, *236*, 139.
- (9) Huber, R.; Fiebig, T.; Wagenknecht, H. A. *Chem. Commun.* **2003**, *15*, 1878.
- (10) Lewis, F. D.; Zhu, H. H.; Daublain, P.; Fiebig, T.; Raytchev, M.; Wang, Q.; Shafirovich, V. *J. Am. Chem. Soc.* **2006**, *128*, 791.
- (11) Lewis, F. D.; Wu, T.; Zhang, Y.; Letsinger, R. L.; Greenfield, S. R.; Wasielewski, M. R. *Science* **1997**, *277*, 673.
- (12) Khajepour, M.; Dashnau, J. L.; Vanderkooi, J. M. *Ann. Biochem.* **2006**, *348*, 40. Rubtsov, I. V.; Kumar, K.; Hochstrasser, R. M. *Chem. Phys. Lett.* **2005**, *402*, 439.
- (13) Lewis, F. D.; Zhang, L. G.; Liu, X. Y.; Zuo, X. B.; Tiede, D. M.; Long, H.; Schatz, G. C. *J. Am. Chem. Soc.* **2005**, *127*, 14445.
- (14) Lewis, F. D.; Wu, Y.; Zhang, L.; Zuo, X.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2004**, *126*, 8206.
- (15) Ponkratov, V. V.; Friedrich, J.; Vanderkooi, J. M.; Burin, A. L.; Berlin, Y. A. *J. Low Temp. Phys.* **2004**, *137*, 289.
- (16) Mitra, R. D.; Silva, C. M.; Youvan, D. C. *Gene* **1996**, *173*, 13.
- (17) Clapp, A. R.; Medintz, I. L.; Mauro, J. M.; Fisher, B. R.; Bawendi, M. G.; Mattoussi, H. *J. Am. Chem. Soc.* **2004**, *126*, 301.
- (18) Wimley, W. C.; White, S. H. *Biochemistry* **2000**, *39*, 4432.
- (19) Gaussian 03, Revision C.02: Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Wallingford, CT, 2004.
- (20) van Rossum, M. C. W.; Nieuwenhuizen, T. M. *Rev. Mod. Phys.* **1999**, *71*, 313.
- (21) Polishchuk, I. Y.; Maksimov, L. A.; Burin, A. L. *Phys. Rep.* **1997**, *288*, 205.
- (22) Chernyak, V.; Zhang, W. M.; Mukamel, S. *J. Chem. Phys.* **1998**, *109*, 9587.
- (23) Zhang, L.; Long, H.; Schatz, G. C.; Lewis, F. D. Manuscript in preparation.
- (24) Bernath, P. F. *Spectra of Atoms and Molecules*; Oxford University Press: New York, 1995; pp 18–21.
- (25) Landau, L. D.; Lifshitz, E. M. *The Classical Theory of Fields*; Addison-Wesley: Reading, MA, 1951.
- (26) Marcus, R. A. *Rev. Mod. Phys.* **1993**, *65*, 599.
- (27) The calculations details will be published elsewhere.
- (28) This identity can be proved expressing the principal value of the integral in eq 37 as $PV \int_{-\infty}^{+\infty} d\epsilon \{ \exp[-b\epsilon^2]/(\epsilon - \epsilon_1) \} = \lim_{\delta \rightarrow 0} \text{Re}(\int_0^{+\infty} dt \int_{-\infty}^{+\infty} d\epsilon \exp[-b\epsilon^2] \exp(i(\epsilon - \epsilon_1 + i\delta)t))$. Then evaluating first the integral with respect to t one can get the result eq 37.
- (29) Zuo, X., private communication.
- (30) Katsura, S.; Hirano, K.; Matsuzawa, Y.; Yoshikawa, K.; Mizuno, A. *Nucleic Acids Res.* **1998**, *26*, 4943.