Theoretical Studies of Short Polyproline Systems: Recalibration of a Molecular Ruler

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY A · APRIL 2009

Impact Factor: 2.69 · DOI: 10.1021/jp811395r · Source: PubMed

6 AUTHORS, INCLUDING:

Lena Dolgikh, PhD
Eli Lilly
12 PUBLICATIONS **171** CITATIONS
SEE PROFILE

Seonah Kim
National Renewable Energy Laboratory
26 PUBLICATIONS **354** CITATIONS
SEE PROFILE

Brent P Krueger
Hope College
29 PUBLICATIONS **1,410** CITATIONS
SEE PROFILE

Adrian Roitberg
University of Florida
164 PUBLICATIONS **7,517** CITATIONS
SEE PROFILE

Theoretical Studies of Short Polyproline Systems: Recalibration of a Molecular Ruler†

Elena Dolghih,‡,§ Wilfredo Ortiz,‡,§ Seonah Kim,‡,§ Brent P. Krueger,⊥ Jeffrey L. Krause,‡,§ and Adrian E. Roitberg*,‡,§

Quantum Theory Project, University of Florida, P.O. Box 118435, Gainesville, Florida 32611-8435, Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200, and Department of Chemistry, Hope College, Holland, Michigan 49422-9000

Received: December 24, 2008; Revised Manuscript Received: February 3, 2009

FRET experiments enable studies of the chemical and physical properties of individual molecules, which has long been a dream of chemists. However, these modern experimental techniques are still limited by the lack of information about the dynamic behavior of the fluorescent labels as well as by the use of dipole−dipole approximation even at short donor-to-acceptor distances. Our results help to suggest that these assumptions need to be carefully considered when designing experiments. We show that at short donor−acceptor separation, dipole−dipole approximation breaks down and Förster theory fails and cannot be used to obtain correct distances. We also explicitly demonstrate that dyes' linkers allow for a lot of flexibility in the fluorescent label orientation and position resulting in distances much shorter than assumed earlier.

Introduction

The ability of an electronically excited molecule (the donor, D) to transfer energy through space to a second molecule (the acceptor, A) was first observed by Perrin in the 1930s[1] and formally described by Förster in the 1940s.[2−4] This resonant energy transfer (RET) phenomenon has been used by structural biologists through fluorescence-detected resonant energy transfer (FRET) experiments, which are known by a variety of names,[5] since its effectiveness was demonstrated in the 1960s and 1970s by the Stryer group and others.[6−11]Chromophores of various types are attached to systems ranging from peptides and oligonucleotides to proteins and nucleic acids to membranes, and fluorescence data are used to infer distances, and hence structure, as well as relative distances and dynamic structural data.[12−17] This is particularly important for biological systems in solution, for which direct structural information is difficult to obtain. In particular, the last 5−10 years have seen a dramatic increase in the use of FRET as a number of laboratories have applied it to single molecules.[13,16,18−23]

At the basis of most analyses of FRET experiments is the familiar form of the Förster equation,[4]

$$k_{\mathrm{RET}} = 0.5291\frac{\kappa^2 Q_{\mathrm{D}} J(\lambda)}{\eta^4 N_{\mathrm{A}} \tau_{\mathrm{D}} R^6} \qquad (1)$$

in which k_{RET} is the rate constant of resonance energy transfer, κ^2 is an orientation factor between the donor and acceptor transition dipoles, Q_{D} is the quantum yield of the donor in the absence of the acceptor, τ_D is the lifetime of the donor in the absence of the acceptor, $J(\lambda)$ is the spectral overlap of the emission spectrum of the donor with the absorption spectrum of the acceptor (in units of M^{-1} cm^3), η is the index of refraction of the medium, N_A is Avogadro's number, and R is the distance between the centers of the donor and the acceptor transition dipoles (in cm). The advantage of Förster's equation is that all of the parameters can be determined from the properties of the donor and acceptor individually except for the two structural parameters, κ and R. While this formula has proven to be extremely useful, its derivation involves a number of assumptions that often limit the detail with which structural information can be extracted from FRET data. These approximations include the following:

(1) The coupling between the donor and acceptor is assumed to be in the weak regime, meaning that the excited states of the FRET system are approximately the same as those of the independent donor and acceptor (observationally this means that the absorption spectrum of the FRET system is simply the sum of the donor and acceptor absorption spectra).

(2) The donor is in an excited electronic state that is vibrationally relaxed.

(3) The coupling between the donor and acceptor is described by dipole−dipole interaction.

In most cases, two additional assumptions are employed:

(4) Both the donor and acceptor sample all possible relative orientations on a time scale that is rapid compared to the energy transfer time scale, in which case the average value of $\kappa^2 = {}^2/_3$.[24]

(5) The rates of other kinetic pathways in the system, k_{OP} (i.e., Q_{D}, J, and τ_{D}) remain constant on the time scale of the experiment allowing E_{ff}, the efficiency of energy transfer,

$$\langle E_{\mathrm{ff}} \rangle = \left\langle \frac{k_{\mathrm{RET}}}{k_{\mathrm{RET}} + k_{\mathrm{OP}}} \right\rangle \qquad (2)$$

to be averaged statically and expressed in terms of a simpler distance parameter

† Part of the "George C. Schatz Festschrift".
* Address correspondence to this author.
‡ Quantum Theory Project, University of Florida.
§ Department of Chemistry, University of Florida.
⊥ Department of Chemistry, Hope College.

10.1021/jp811395r CCC: $40.75
Published on Web 03/05/2009

$$\langle E_{\mathrm{ff}}\rangle = \frac{\langle k_{\mathrm{RET}}\rangle}{\langle k_{\mathrm{RET}}\rangle + k_{\mathrm{OP}}} = \left[1 + \left(\frac{\langle R^6\rangle}{{R_0}^6}\right)\right]^{-1} \quad (3)$$

where R_0, the Förster radius, is the distance at which the efficiency is 50%. Förster radii are tabulated for many D−A pairs,[5] allowing eq 3 to be used to model experimental data. It is important to point out that while the quantum yield of the donor in the absence of the acceptor is assumed to be constant, it has to be measured for the specific system to which the donor is attached since it can be quite different from the donor free in solution.

While fluorescence has been used to probe a number of systems in which assumptions 1 and 2 are questionable,[25−28] these systems generally exhibit different steady-state spectra than the individual chromophores (e.g., the photosynthetic reaction center), due to small values of R and, consequently, high rates of energy transfer. In contrast, the majority of applications of FRET to structural biology (including those addressed here) lie safely in the weak coupling regime and display energy transfer time scales slow enough to allow vibrational relaxation in the donor, which happens on a picosecond time scale or faster. The other assumptions, particularly (3) and (4), are more often suspect. While this has been pointed out by many authors,[12,29] little work has been done to assess the validity of these assumptions quantitatively.[30−33] Recent advancements in single-molecule fluorescence techniques, however, both sparked a renewed interest in more precise interpretation of experimental FRET data and allowed for the above assumptions to be tested. A system of choice for some of these studies became the classic dye-labeled polyproline system that was the subject of the original investigation of FRET, described as a molecular ruler.

In this influential, and highly cited, paper, Stryer and Haugland reported in 1967 an experimental study of the dependence of energy transfer on distance.[9] The authors synthesized a series of related molecules (see Figure 1) consisting of an α-naphthyl energy donor and a dansyl energy acceptor, connected to poly-L-proline spacers with 1 to 12 monomer units. Poly-L-proline was chosen because it is known to form oligomers that are rather rigid in structure.

To test the use of FRET as a spectroscopic ruler, the authors compared the experimentally measured energy transfer efficiency to the "known" donor−acceptor distance to see if the relationship between the two was well-described by Förster theory. The energy transfer efficiency was determined by measuring the steady-state excitation spectrum of the acceptor as a function of the number of spacers, n. The donor−acceptor separation was determined by estimating the end-to-end distance of the poly-L-proline spacers by using Cowan−McGavin crystallographic coordinates,[34,35] and the distance from the end of the spacer to the center of the chromophores by using molecular models and assuming the linkers are fully extended.

The results of the experiments showed that the efficiency of energy transfer was well-fit by a curve with an R^{-6} dependence (cf. Figure 4 from that work[9]), in excellent agreement with the predictions of the Förster model. (Though, as we note below, the x-axis of this figure is not well-determined.) On the basis of these results, Stryer and Haugland verified Förster's theory and suggested that energy transfer could serve as a spectroscopic ruler. That is, by measuring the efficiency of resonance energy transfer, the distance between the donor and the acceptor can be determined directly. This paper used the foundation of Förster theory to effectively create the field of FRET, which has been broadly active from the 1970s through today.

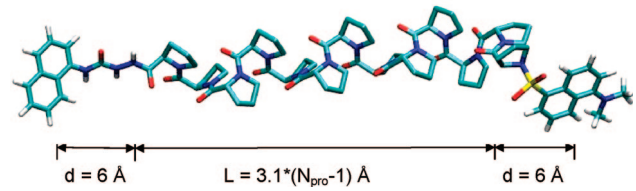

Figure 1. Cartoon of a polyproline structure, $n = 12$, with the donor and acceptor attached by linkers. Shown are the Cowan−McGavin distances and the extended linker lengths as assumed by Stryer and Haugland.[9]

Despite the success of this early work, there are multiple reasons to suspect that the poly-L-proline story is actually more complicated, even from the data presented in the original paper. Notably, the data at low n do not fall on the R^{-6} line (cf. Figure 5 from that work,[9] which does not include points $n = 1$, 2, or 3). In fact, in a problem in the Physical Chemistry textbook by Atkins and de Paula related to this experiment,[36] students are instructed to remove the first two points to improve the fit. Atkins and de Paula's justification is that the data at shorter distances, with efficiencies near 100%, are expected to be inaccurate. While it is the case that these high efficiencies are difficult to measure experimentally, these are also the points of shortest donor−acceptor distance where Förster theory becomes questionable due to possible failures of assumptions 3 and 4. If all of the data are plotted, the best-fit line has a slope closer to R^{-4} than R^{-6}. Another reason for suspicion is, as noted in the paper, the discrepancy in the Förster radius, R_0. This parameter, which corresponds to the distance at which the fluorescence efficiency equals 50%, was found from the fit to FRET data to be 34.6 Å.[9] However, using the standard method of calculating R_0 from tabulated spectroscopic parameters yields 27.2 Å.[9]

Several studies have investigated the polyproline system since then. Jacob et al.[37] used NMR and molecular dynamic (MD) studies to show that the donor−acceptor distance for poly-(L)-n-proline, $n = 6−8$, is smaller than expected from the rigid rod model of the system and suggested that flexible linkers of the two labels, nitroxide and tryptophan, are responsible. Watkins et al.,[38] using Alexa dyes, observed higher efficiencies then would be expected for $n = 8−24$, which, he suggested, could be explained by presence of a few *cis*-proline residues. They also argued that $\langle\kappa^2\rangle = {}^2/_3$ is a valid approximation due to rotational motions of the dyes being much faster than the interphoton timing. Sahoo et al.,[39] using NMR and molecular dynamics, showed that linkers for longer peptides result in backfolding of fluorophores onto the protein causing shorter donor−acceptor distances. They also showed that at shorter distances, R_0 and R are correlated and suggested that breakdown of dipole−dipole approximation could be responsible. Schuler et al.[40] looked at systems from $n = 6$ to 40 using Alexa fluorophores as labels. He found efficiencies lower than predicted by Förster at short distances and much higher than expected at longer distances. He suggested that the former can be explained by insufficient κ^2 averaging due to fast energy transfer rate, while ascribing the latter to the polyproline flexibility. In later studies on 20-L-proline, using experiment and theory, Best et al.[41] showed that flexibility of dye linkers plays an important role in experimental data interpretation and suggested that this effect could be responsible for Stryer and Haugland R_0 discrepancy. Finally, Doose et al.[42] used photo-induced electron transfer quenching by fluorescence to demonstrate that the presence of cis-isomers results in efficiencies much higher than expected. As a result of these studies, it became clear that the original picture of polyproline peptide

being a rigid rod with *all-trans*-proline residues and randomly rotating dyes attached by defined spacers is far from correct.

In this paper we use molecular dynamics simulations along with accurate treatment of the donor−acceptor coupling to address the free rotation of the dyes as well as dipole−dipole approximations for polyproline systems (assumptions 3 and 4). Results of simulations of simple polyproline peptides are compared with Förster theory as well as classic work by Stryer and Haugland. In particular, we use the transition density cube (TDC) method,[30] described below, to calculate the donor−acceptor coupling. The TDC method is valid at all interchromophore distances as opposed to traditional Förster theory, which breaks down at small R. We also compare estimated efficiencies directly with corresponding distances. Note that in the experimental work by Stryer and Haugland, as with the vast majority of FRET measurements, the interchromophore distance was not known or measured, but *modeled*. That is, they measured the efficiency of energy transfer (the y-axis), but could not measure independently the distance (the x-axis). We propose a simple and general computational approach that can be used to test the validity of assumptions 3 and 4 prior to or in concert with experiments. More importantly, this methodology allows for quantitative processing of FRET experiments, as we illustrate for the case of dye-labeled polyproline, even in situations in which assumptions 3 and 4 are found to be invalid.

Background and Theory

The foundation for eq 1 is the Golden Rule, which after applying the first two assumptions from above can be written as

$$k_{\mathrm{RET}} = \frac{2\pi}{\hbar}|V_{\mathrm{DA}}|^2 J(\varepsilon) \tag{4}$$

or,

$$k_{\mathrm{RET}} = \frac{1}{\hbar^2 c}|\tilde{V}_{\mathrm{DA}}|^2 J(\tilde{\nu}) \tag{5}$$

where V_{DA} is the electronic coupling between the donor and the acceptor in energy units, and $\tilde{V}_{\mathrm{DA}}$ is the coupling (in units of cm^{-1}). J is the spectral overlap, which represents the conservation of energy. In eq 5, J has units of cm and is derived from donor emission and acceptor absorption spectra that have been normalized on a cm^{-1} scale.[43] Note that since J is determined directly from experimental spectra, it incorporates, to a certain extent, the effects of dielectric screening.[44]

While a number of mechanisms make significant contributions to the total D−A coupling at various length scales ranging from orbital-overlap-dependent mechanisms at a few Å[45] to radiative mechanisms at many meters,[46] the dominant contribution at most relevant biological distances is the Coulombic coupling[14,47]

$$V_{\mathrm{DA}} \approx V_{\mathrm{Coul}} = \left\langle \psi^*_{\mathrm{D}}\psi_{\mathrm{A}}\left|\frac{1}{R}\right|\psi_{\mathrm{D}}\psi^*_{\mathrm{A}}\right\rangle \tag{6}$$

where Ψ_{D} and Ψ_{A} are wave functions of the donor and the acceptor respectively in ground and excited (*) states, while R is the donor−acceptor distance.

If the D−A separation is large compared to the size of the transition dipoles, one may safely use only the first term of a multipole expansion to approximate the Coulombic coupling as the transition-dipole transition-dipole coupling, i.e., incorporate assumption 3,

$$V_{\mathrm{Coul}} \approx V_{\mathrm{dip-dip}} = \frac{5042\kappa|\mu_{\mathrm{D}}||\mu_{\mathrm{A}}|}{\eta^2 R^3} \tag{7}$$

where μ_{D} and μ_{A} are transition dipole moments of the donor and the acceptor, respectively. In his seminal work, Förster did this, allowing him to then incorporate the A transition dipole magnitude into J and arrive at eq 1. (Note that $J(\lambda)$ in eq 1 has units of $\mathrm{M}^{-1}\ \mathrm{cm}^3$.) While eq 1 is widely used and quite useful, it precludes a detailed consideration of the direct mechanism of coupling between the D and A. Here, we will use eq 5, which separates the coupling and spectral overlap terms.

While Förster's application of assumption 3 has generally been successful, no clear guidelines exist to determine a priori whether a particular system of interest has a separation that is large enough to justify use of the dipole−dipole expression. For instance, many systems exhibit relatively strong electronic coupling (though still in the "weak coupling regime") despite rather large donor−acceptor separations.[30,48] Therefore, the applicability of eq 1 should be questioned in many modern applications. In principle, one may simply include higher order terms from the multipole expansion to improve eq 7:

$$V_{\mathrm{Coul}} = V_{\mathrm{dip-dip}} + V_{\mathrm{dip-quad}} + V_{\mathrm{quad-quad}} + V_{\mathrm{dip-oct}} + V_{\mathrm{quad-oct}} + V_{\mathrm{oct-oct}} + \ldots \tag{8}$$

However, in practice, the higher order terms change sign and the series converges slowly. In addition, the relevant quantity is the square of eq 8, which introduces cross terms, making the multipole expansion unwieldy.

Several methods have been proposed to calculate the total Coulombic coupling accurately,[49−51] including some that represent the Coulombic coupling essentially exactly.[30,32,33] Here we use the transition density cube (TDC) method[30] to estimate the total D−A coupling for each of a series of system configurations determined through molecular dynamics (MD) simulation.

In this method, quantum mechanical calculations are used to determine the ground and excited state wave functions of a molecule of the donor and the acceptor. These wave functions are integrated into densities on a discrete three-dimensional grid to yield the TDC of the molecule

$$M_N(r) = \int_s \psi_N \psi^*_N \,\mathrm{d}s\,\mathrm{d}r \tag{9}$$

where N refers to the donor or the acceptor, in its ground or (*) excited state. This TDC can then be reoriented to overlay with the D and A position at any given time, t, in the MD simulation. The total interaction between the donor TDC and an acceptor TDC oriented in an analogous manner gives the full Coulombic coupling (in the limit of finely grained TDCs, for a given QM model chemistry) from that configuration, where the sum is over the D and A grids

$$V_{\mathrm{TDC}} = \sum_{i\cdot j} \frac{M_{\mathrm{D}}(i)M_{\mathrm{A}}(j)}{4\pi\varepsilon_0 r_{ij}} \tag{10}$$

The TDCs are reoriented and V_{Coul} (and $V_{dip\text{-}dip}$) calculated for each snapshot from the simulation, yielding a distribution of D−A interactions representing the effects of structural dynamics in the system. Thus, assumption 3 can be evaluated directly for the distribution of structures that is relevant to a particular system under consideration.

Of the assumptions listed above, point 4, $\langle\kappa^2\rangle = {}^2/_3$, has received the most careful attention. In particular, work by Dale and co-workers,[52−54] and by van der Meer,[5,24] provides a means by which limits on the possible values of κ^2 can be determined. However, the error bars are often broad and the techniques, while often cited, are used infrequently. The TDC method allows evaluation of orientation-dependence at each snapshot, allowing an examination of the validity of the $\langle\kappa^2\rangle = {}^2/_3$ assumption for the particular dye and peptide system under study. It also allows addressing directly the possibility that there may be a correlation between several of the factors in eq 1. For instance, if κ and R are not independent of each other, then $\langle\kappa^2R^{-6}\rangle \neq \langle\kappa^2\rangle\langle R^{-6}\rangle$ and the average value of κ^2 is irrelevant, regardless of the validity of assumption 4.[55]

Computational Methods

Our molecular dynamics simulations of the poly-L-proline oligomers were performed with the AMBER 8.0[56] software package. Two force fields, *amber94* and general amber force field (*gaff*), were used to model the peptide and the dyes, respectively. Parameters and geometries of the dyes can be found in the Supporting Information. Electrostatic solvent effects were simulated by using the generalized Born model with cutoff for nonbonded interactions set to 12 Å. The friction effect of the solvent was simulated with Langevin dynamics, using friction coefficient γ of 1.0 ps^{-1}. Bonds involving hydrogens were constrained to their equilibrium values by using the SHAKE algorithm and a time step of 2 fs was used. After energy minimization of the initial systems, 50-ns production runs at 300 K were performed for each system. Configurations from the dynamics were saved every 5 ps.

Calculations of TDC elements were performed in Q-Chem,[57] using the Hartree−Fock level of theory and the 6-31G* basis set. Excited states of the donor and the acceptor were computed by using the configuration interaction singles (CIS) method with CIS excited state roots set to 4. The number of cube elements was set to 50 000 after test runs with larger values indicated convergence.

Results and Discussion

We performed large-scale molecular dynamics simulations of the dye-substituted poly-L-proline system in Figure 1. Our results show that the end-to-end distances not including the dyes (i.e., PRO1Cα−PRO*n*Cα) scaled with 3.0 Å, close to the expected crystallographic length of 3.1 Å per proline. When the data were fitted to the worm-like chain (WLC) model[58] (Figure 2), the persistence length was consistently larger than the contour length indicating that the systems is indeed a rigid rod for the short distances studied here. Note that "rigid" is still a relative definition here since for example for $n = 14$, the end-to-end distance varies from 30 to 45 Å, which, once weighted by the $1/R^6$ term results in very asymmetric efficiency distributions. Within the WLC model, fit of the data (Figure 2, inset) indicates that the persistence length of the polyproline system has to be above 70 Å, which supports the estimates of it being in the 90 to 130 Å range.[41]

When the motions of the dyes are included, by measuring the distance from the center of the donor to the center of the acceptor, the distributions change dramatically (see Figure 3). The dyes do not simply add a well-defined spacer between the end of the poly-L-proline and the center of the transition moments of the dyes as was assumed by Stryer and Haugland.[9] Rather, the inherent flexibility of the fluorophores linkers broadens the distributions considerably and the distance between the donor and acceptor does not scale with n in a simple way. The inset of Figure 3 shows the distance distribution for the $n = 2$ system, over the first and second halves of the simulation. They overlap almost perfectly, showing that the simulation is converged. The importance of the flexibility of the linkers on donor-to-acceptor distances has been pointed out by Best et al.[41] in a recent study where conformational dynamics of two Alexa fluorophores attached to 20-L-proline has been investigated. Our simulations, conducted on polyproline $n < 20$ and a different set of dyes, not only confirm the importance of the effect but also demonstrate it explicitly for a series of related systems. Thus, Figure 4 illustrates the variation in κ and R due to flexibility of linkers. It suggests that the acceptor samples a significant fraction of orientation space which, in turn, suggests that assumption 4 may be valid. Figure 5 further shows that $\langle\kappa^2\rangle$ does seem to converge around ${}^2/_3$ (albeit with large standard

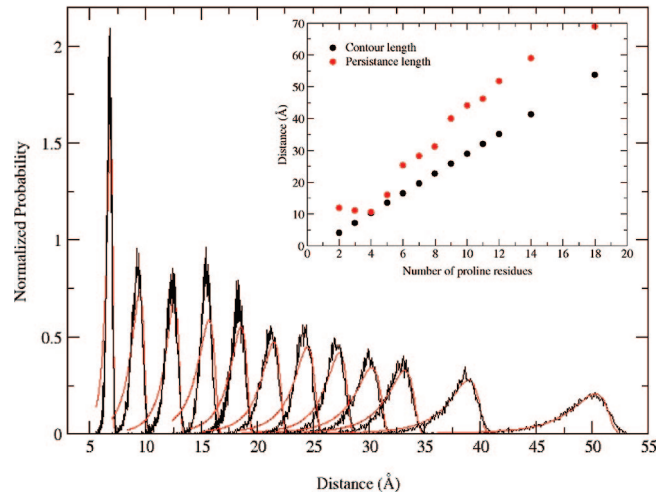

Figure 2. Cα−Cα distance distributions between the terminal proline residues during the polyproline−dye molecular dynamics. In black lines we show the MD distributions, and in red the fit to a WLC model. The inset shows the fitted values of contour length (black circles) and persistence length (red circles) vs. the number of proline residues.

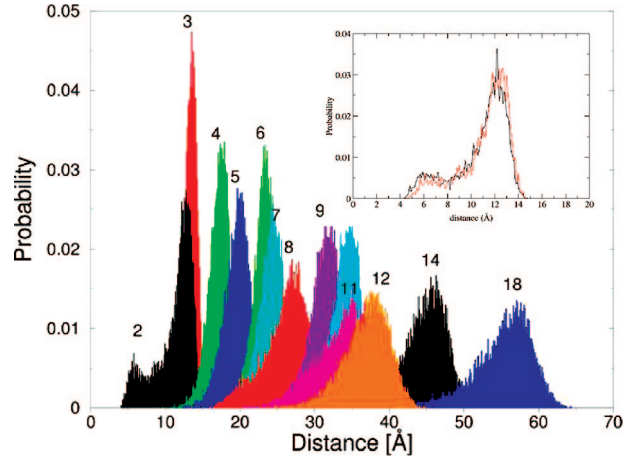

Figure 3. Center of dansyl to center of naphthyl distance distributions during the polyproline−dye molecular dynamics. The inset shows the distance distribution for the $n = 2$ system for the first and second half of the simulation.

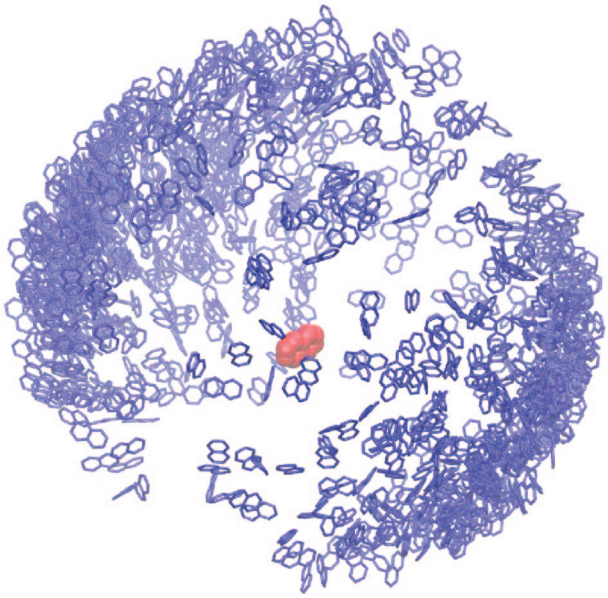

Figure 4. Overlapped snapshots form the MD run for polyproline-12 with attached dyes. They have been rmsd fitted to the donor dye (naphthyl, in red vdw spheres). The dansyl acceptor dye is shown in blue lines. All acceptor positions are shown from the perspective of the donor.

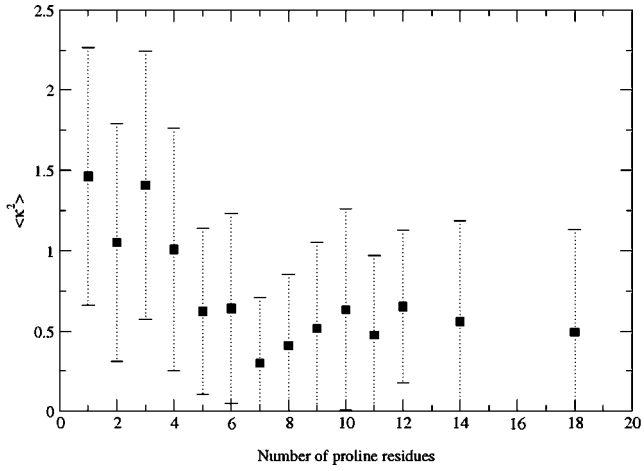

Figure 5. Average κ^2 as a function of number of proline residues in the system. Error bars are shown as one standard deviation.

deviations) for most of the systems with $n > 4$, coinciding with those data points used by Stryer and Haugland for their fit of the distance dependence.

Of additional interest is not simply the way in which the dyes sample orientation space, but also how this sampling may be correlated with the sampling of D−A distance. To test the independence of κ and R, we computed the ratio between the static and dynamic limits of averaging κ^2 and R^{-6}

$$\xi = \frac{\langle\kappa^2\rangle\langle R^{-6}\rangle}{\langle\kappa^2 R^{-6}\rangle} \tag{11}$$

for each proline system. The results, as shown in Figure 6, indicate that κ and R are strongly correlated for the smaller systems. Analysis of molecular dynamics snapshots provides an explanation. For example, a value of ξ of ~3 for $n = 2$ corresponds to a bimodal distance distribution (Figure 3) and reflects two conformational states visited by α-naphthyl (Figure

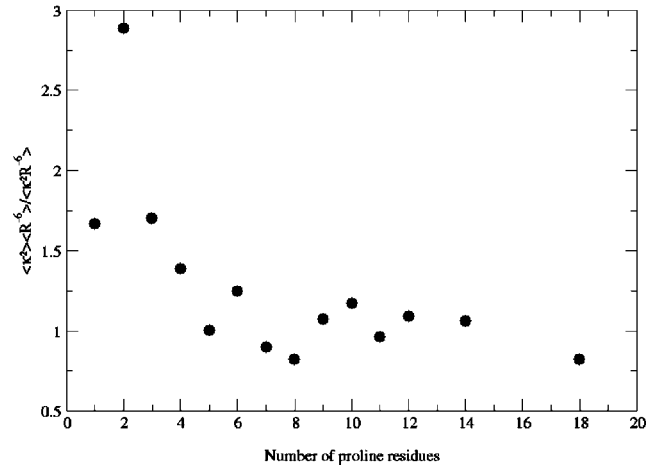

Figure 6. Ratio of static average to dynamic average of κ^2 and R as a function of number of proline residues in the system.

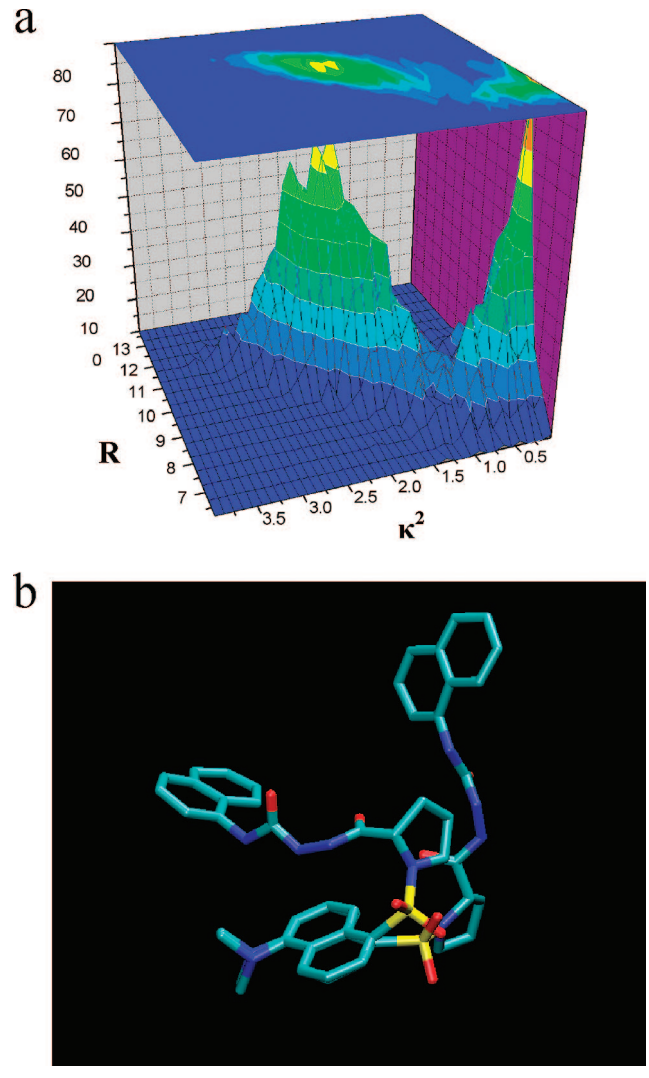

Figure 7. (a) Correlation between orientation and distance for the $n = 2$ system. (b) Two conformational clusters for naphthyl in the $n = 2$ system.

7). As the number of proline residues increases, so does the conformational freedom of the linkers and the correlation drops off. This demonstrates that κ and R can be assumed to be independent for $n > 6$.

Assumption 3 is investigated in Figure 8, which shows the ratio between the FRET rates calculated by using dipole−dipole

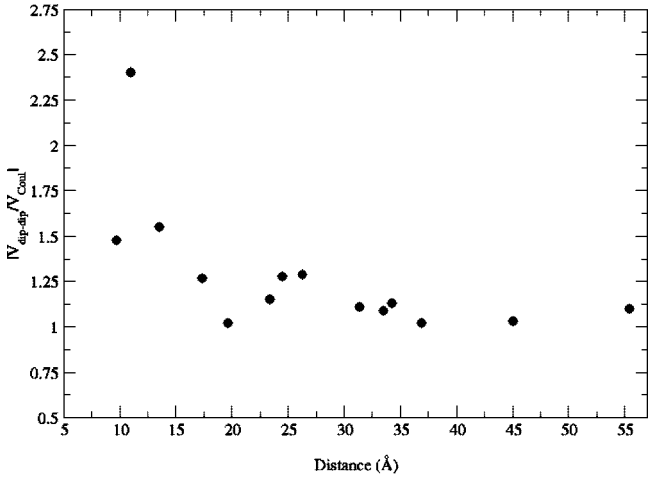

Figure 8. Ratio of $|V_{\text{dip-dip}}/V_{\text{Coul}}|$ couplings as function of average distances from the MD runs.

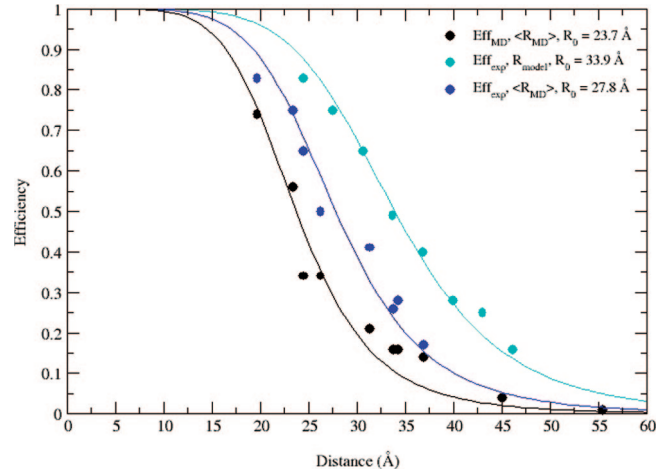

Figure 9. FRET efficiency vs. distance and R^{-6} fits: experimental efficiency vs. modeled distance used by Stryer and Haugland (light blue), experimental efficiency vs. average MD distance (dark blue), and MD efficiency vs. MD distance (black).

approximation and full Coulombic coupling. As one can see from the figure, for systems with $n < 5$, the ratio of $|V_{\text{dip-dip}}/V_{\text{Coul}}|$ is larger than 1 indicating that dipole−dipole approximation at short distances results in overestimation of the rate of energy transfer. This clearly illustrates that use of Förster theory at short distances is not correct. Note that for distances larger than 30 Å, corresponding to $n > 9$, the agreement between $V_{\text{dip-dip}}$ and V_{Coul} is reasonably good, though errors greater than 10% are observed. However, for distances less than 30 Å, which is $\sim R_0$ in this case, assumption 3 is suspect with errors of more than 20% being common. For systems with $n < 5$, errors are generally greater than 50%. Thus, for the majority of oligomers studied here one must be cautious in the use of assumption 3. Note that a 20% error in the coupling strength becomes a 44% error in k_{RET} and that this same 20% error in coupling leads to a 6% error in determining R from a measured value of k_{RET} (because $R \propto k_{\text{RET}}^{1/6}$).

We would argue, that the breakdown of dipole−dipole approximation, and not the lack of orientational averaging, can explain lower efficiencies at shorter distances in Schuler et al. studies.[40] In this case, the sizes of the donor and the acceptor, 7 to 12 Å,[40] are significant in comparison to their separation of, for example, 24 Å corresponding to an poly(8)proline system. As one can see in Figure 8, for the naphthyl−dansyl pair, dipole−dipole approximation results in overstimation of the rate of energy transfer for $n = 1$, 2, and 3.

On the basis of this observation, we can define "comparable distance" as a distance three times that of the probe length or less. This should be of particular importance for donor−acceptor pairs with short Förster radii, such as, for example, the tryptophan−pyrene pair ($R_0 = 28$ Å) and the tryptophan−heme ($R_0 = 29$ Å) where acceptor lengths are 7 and 12 Å, respectively.[59] In a recent study, Sahoo et al. measured FRET between donor−acceptor pairs with R_0 of 10 Å.[60] Their distances, derived from efficiencies using the Förster formula, were consistently larger than average distances from molecular dynamics simulations of the same system. While they ascribed this discrepancy to force field parameters, we believe that that breakdown of the dipole−dipole approximation is the main cause.

To evaluate the impact of assumptions 3 and 4 from the perspective of a FRET experiment, we would like to relate to experimental parameters that can be trusted, that is the RET efficiency that is measured for a system with a particular number of proline residues. Our recalibrated molecular ruler is shown in Figure 9 where both the TDC data and the original Stryer and Haugland data are plotted as RET efficiency versus distance. Note that theoretical efficiencies were calculated by using the Förster formula and only for systems $n > 4$ where dipole−dipole approximation is valid (see Figure 8). Both curves scale with nearly an R^{-6} dependence (the best fit for the theoretical curve is 5.8), but are shifted in R with respect to each other. The new ruler implies relatively large differences in estimates of R from a given RET efficiency. For example, a measured fluorescence efficiency of 20% would predict an R of 45 Å in the original ruler, while the recalibrated ruler predicts an R of 37 Å. This large shift between the two plots is mainly due to the flexibility of the linkers and dyes. Recall that the model employed by Stryer and Haugland assumed both that each proline introduced a fixed additional distance and that the linker/dye units simply added an additional fixed distance. Note in the figure that once experimental distances are plotted versus distances obtained from MD, R_0 becomes 27.8 Å, which is in excellent agreement with the value calculated by using spectroscopic parameters, 27.2 Å.[9] This flexibility of the linker and dye portion of the molecules explains why the R_0 value "measured" by Stryer and Haugland (34.6 Å) was so different than the tabulated value (27.2 Å). When we fit efficiencies calculated based on MD simulations, the R_0 value is 23.7 Å, which is in a good agreement with the tabulated value as well. Clearly for intermediate efficiencies, there are significant discrepancies between the simulations and experiment. Note that this region is the most sensitive to small changes in oligomer structure, which results in relatively large uncertainty ranges in each simulated point. Another cause for the difference between the two R_0 could be attributed to the possible presence of *cis*-proline residues in the experiment. Both Doose et al.[42] and Best et al.[41] demonstrated that the presence of even a few cis-residues could result in a much larger range of both distances and orientations sampled by the dyes. The structural dynamics introduced into the RET efficiency by the linker and dye portion of the molecule highlights a key issue in the interpretation of FRET experiments.

In this analysis, we have treated each configuration as though it were a separate experiment, effectively ignoring the time information present in the MD simulation. This introduces a small error into our simulated efficiencies, because we have ignored dynamic contributions to the rate of energy transfer.[61] However, Jean and Krueger[62] estimated this dynamic effect to be small in MD simulations of small ssDNA oligomers.

Conclusion

FRET experiments, particularly single-molecule experiments, have emerged as a powerful tool in the laboratory. Such experiments enable studies of the chemical and physical properties of individual molecules, which has long been a dream of chemists. However, even these modern experimental techniques are still limited by the assumptions that have been traditionally made in analysis of FRET data. Our simulation results help to suggest that these assumptions, especially the limitations of the IDA (assumption 3) and the effects of structural dynamics (assumption 4), need to be carefully considered when designing experiments. In particular, fluorescent probes should be far enough apart (greater than 30 Å for the α-naphthyl and dansyl probes) and should either be on very flexible linkers or, if the structure of the system is well-known, on very rigid linkers. If the system of study has dynamics of its own that are of interest, then the problem may be very complicated and simulations will likely be needed to understand the result.

Acknowledgment. This work was supported in part by the Department of Energy through grant DE-FG02- 02ER45995 (J.L.K. and A.E.R.), a University of Florida Alumni Fellowship (ED), Research Corporation Award CC5479 (B.P.K.), the Petroleum Research Fund, administered by the American Chemical Society (B.P.K.), and the Towsley Foundation (B.P.K.). Computational resources were provided by the Large Allocations Resource Committee through grants TG-MCA05S010 (A.E.R.) and TGCHE050058T (B.P.K.). The authors acknowledge the University of Florida High-Performance Computing Center for providing computational resources and support that have contributed to the research results reported within this paper (http://hpc.ufl.edu) as well as the Hope College Computational Science and Modeling Laboratory (http://www.hope.edu/academic/csm/) funded, in part, by the Howard Hughes Medical Institute. Additional computing resources were provided by the National Science Foundation-Major Research Instrumentation program (B.P.K.). We honor and thank George Schatz for many years of guidance, support and insights.

Supporting Information Available: Tables of the dye force field parameters and charges. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Perrin, F. *Ann. Phys.* **1932**, *17*, 283–314.
(2) Förster, T. *Naturwissenschaften* **1948**, *33*, 166–175.
(3) Förster, T. *Ann. Phys.* **1948**, *2*, 55–75.
(4) Förster, T. In *Modern Quantum Chemistry*; Sinanoglu, O., Ed.; Academic Press, Inc.: New York, 1965; Vol. 3, pp 93−137.
(5) van der Meer, B. W.; Coker, G. I.; Chen, S.-Y. In *Resonance Energy Transfer: Theory and Data*; VCH Publishers: New York, 1994.
(6) Haugland, R. P.; Yguerabide, J.; Stryer, L. *Proc. Natl. Acad. Sci. U.S.A.* **1969**, *63*, 23–30.
(7) Hillel, Z.; Wu, C.-W. *Biochemistry* **1976**, *15*, 2105–2113.
(8) Stryer, L. *Annu. Rev. Biochem.* **1978**, *47*, 819–846.
(9) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 719–726.
(10) Vanderkooi, J. M.; Ierokomas, A.; Nakamura, H.; Martonosi, A. *Biochemistry* **1977**, *16*, 1262–1267.
(11) Wu, C.-W.; Stryer, L. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 1104–1108.
(12) Selvin, P. R. *Nat. Struct. Biol.* **2000**, *7*, 730–734.
(13) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743–747.
(14) Scholes, G. D. *Annu. Rev. Phys. Chem.* **2003**, *54*, 57–87.
(15) Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Lilley, D. M. J. *Proc. Nat. Acad. Sci. U.S.A.* **1993**, *90*, 2994–2998.
(16) Jia, Y.; Talaga, D. S.; Lau, W. L.; Lu, H. S. M.; DeGrado, W. F.; Hochstrasser, R. M. *Chem. Phys.* **1999**, *247*, 69–83.
(17) Kulzer, F.; Orrit, M. *Annu. Rev. Phys. Chem.* **2004**, *55*, 585–611.
(18) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3670–3675.
(19) Deniz, A. A.; Laurence, T. A.; Beligere, G. S.; Dahan, M.; Martin, A. B.; Chemla, D. S.; Dawson, P. E.; Schultz, P. G.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5179–5184.
(20) Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6264–6268.
(21) Ha, T.; Ting, A. Y.; Liang, J.; Caldwell, W. B.; Deniz, A. A.; Chemla, D. S.; Schultz, P. G.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 893–898.
(22) Rhoades, E.; Gussakovsky, E.; Haran, G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3197–3202.
(23) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528–1533.
(24) van der Meer, B. W. In *Resonance Energy Transfer*; John Wiley & Sons, Inc.: New York, 1999; pp 151−172.
(25) Bradforth, S. E.; Jimenez, R.; van Mourik, F.; van Grondelle, R.; Fleming, G. R. *J. Phys. Chem.* **1995**, *99*, 16179–16191.
(26) Xie, X.; Du, M.; Mets, L.; Fleming, G. R. In *Time-Resolved Laser Spectroscopy in Biochemistry III*; SPIE: Bellingham, WA, 1992; pp 690−706.
(27) Caro, C. D.; Visschers, R. W.; van Grondelle, R.; Völker, S. J. *J. Phys. Chem.* **1994**, *98*, 10584–10590.
(28) Scholes, G. D.; Jordanides, X. J.; Fleming, G. R. *J. Phys. Chem. B* **2001**, *105*, 1640–1651.
(29) Weiss, S. *Science* **1999**, *283*, 1676–1683.
(30) Krueger, B. P.; Scholes, G. D.; Fleming, G. R. *J. Phys. Chem. B* **1998**, *102*, 5378–5368.
(31) Srinivas, G.; Bagchi, B. *J. Phys. Chem. B* **2001**, *105*, 9370–9374.
(32) Wong, K. F.; Bagchi, B.; Rossky, P. J. *J. Phys. Chem. A* **2004**, *108*, 5752–5763.
(33) Beierlein, F. R.; Othersen, O. G.; Lanig, H.; Schneider, S.; Clark, T. J. *J. Am. Chem. Soc.* **2006**, *128*, 5142–5152.
(34) Sasisekharan, V. *Acta Crystallogr.* **1959**, *12*, 903–909.
(35) Shimmel, P. R.; Flory, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 52–59.
(36) Atkins, P. W.; de Paula, J. In *Physical Chemistry*, 7th ed.; W. H. Freeman and Co.: New York, 2002.
(37) Jacob, J.; Baker, B.; Bryant, G. R.; Cafiso, D. S. *Biophys. J.* **1999**, *77*, 1086–1092.
(38) Watkins, L. P.; Chang, H.; Yang, H. *J. Phys. Chem. A* **2006**, *110*, 5191–5203.
(39) Sahoo, H.; Roccatano, D.; Hennig, A.; Nau, W. M. *J. Am. Chem. Soc.* **2007**, *129*, 9762–9772.
(40) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.
(41) Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18964–18969.
(42) Doose, S.; Neuweiler, H.; Barsch, H.; Sauer, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17400–17405.
(43) Krueger, B. P.; Scholes, G. D.; Jimenez, R.; Fleming, G. R. *J. Phys. Chem. B* **1998**, *102*, 2284–2292.
(44) Scholes, G. D.; Curutchet, C.; Mennucci, B.; Cammi, R.; Tomasi, R. *J. Phys. Chem. B* **2007**, *111*, 6978–6982.
(45) Scholes, G. D.; Harcourt, R. D. *J. Chem. Phys.* **1996**, *104*, 5054–5061.
(46) Berberan-Santos, M. N.; Pereira, E. J. N.; Martinho, J. M. G. In *Resonance Energy Transfer*; John Wiley & Sons, Inc.: New York, 1999; pp 108−149.
(47) Scholes, G. D. In *Resonance Energy Transfer;* John Wiley & Sons, Inc.: New York, 1999; pp 212−243.
(48) Ortiz, W.; Krueger, B. P.; Kleiman, V. D.; Krause, J. L.; Roitberg, A. E. *J. Phys. Chem. B* **2004**, *109*, 11512–11519.
(49) Chang, J. C. *J. Chem. Phys.* **1977**, *67*, 3901–3909.
(50) Golebiewski, A.; Witkowski, A. *Rocz. Chem.* **1959**, *33*, 1443.
(51) Sauer, K.; Cogdell, R. G.; Prince, S. M.; Freer, A.; Isaacs, N. W.; Scheer, H. *Photochem. Photobiol.* **1996**, *64*, 564–576.
(52) Dale, R. E. *Acta Phys. Pol.* **1978**, *A54*, 743–756.
(53) Dale, R. E. *Biopolymers* **1974**, *13*, 1573–1605.
(54) Dale, R. E.; Eisinger, J. In *Biochemical Fluorescence: Concepts*; Marcel Dekker: New York, 1975; Vol. 1, pp 115−284.
(55) VanBeek, D. B.; Zwier, M. C.; Shorb, J. M.; Krueger, B. P. *Biophys. J.* **2007**, *92*, 4168–4178.
(56) Case, D. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Homak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California, San Francisco, CA. 2004.
(57) Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Leychenko, V.; O'Neill,

D. P.; Distasio, R. A.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herber, J. M.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynoy, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

(58) Thirumalai, D.; Ha, B. Y.In *Theoretical and mathematical models in polymer research*; Grosberg, A., Ed.; Academia: New York, 1988.

(59) Wu, P.; Brand, L. *Anal. Biochem.* **1994**, *218*, 1–13.

(60) Sahoo, H.; Roccatano, D.; Zacharias, M.; Nau, W. M. *J. Am. Chem. Soc.* **2006**, *128*, 8118–8119.

(61) Troisi, A.; Nitzan, A.; Ratner, M. A. *J. Chem. Phys.* **2003**, *119*, 5782–5788.

(62) Jean, J. M.; Krueger, B. P. *J. Phys. Chem. B* **2006**, *110*, 2899–2909.

JP811395R