Published in final edited form as:

J Phys Chem A. 2011 April 28; 115(16): 3997–4008. doi:10.1021/jp1095344.

Using molecular dynamics and quantum mechanics calculations to model fluorescence observables

Amy L. Speelman^a, Aurora Muñoz-Losa^b, Katie L. Hinkle^a, Darren B. VanBeek^a, Benedetta Mennucci^b, and Brent P. Krueger^a

^a Hope College, Department of Chemistry, Holland, MI 49423 (USA)

^b Dipartimento di Chimica e Chimica Industriale, Università degli Studi di Pisa, Via Risorgimento 35, 56126 Pisa (Italy)

Abstract

We provide a critical examination of two different methods for generating a donor-acceptor electronic coupling trajectory from a molecular dynamics (MD) trajectory and three methods for sampling that coupling trajectory, allowing the modeling of experimental observables directly from the MD simulation. In the first coupling method we perform a single quantum-mechanical (QM) calculation to characterize the excited state behavior, specifically the transition dipole moment, of the fluorescent probe, which is then mapped onto the configuration space sampled by MD. We then utilize these transition dipoles within the ideal dipole approximation (IDA) to determine the electronic coupling between the probes that mediates the transfer of energy. In the second method we perform a QM calculation on each snapshot and use the complete transition densities to calculate the electronic coupling without need for the IDA. The resulting coupling trajectories are then sampled using three methods ranging from an independent sampling of each trajectory point (the Independent Snapshot Method) to a Markov chain treatment that accounts for the dynamics of the coupling in determining effective rates. The results show that the IDA significantly overestimates the energy transfer rate (by a factor of 2.6) during the portions of the trajectory in which the probes are close to each other. Comparison of the sampling methods shows that the Markov chain approach yields more realistic observables at both high and low FRET efficiencies. Differences between the three sampling methods are discussed in terms of the different mechanisms for averaging over structural dynamics in the system. Convergence of the Markov chain method is carefully examined. Together, the methods for estimating coupling and for sampling the coupling provide a mechanism for directly connecting the structural dynamics modeled by MD with fluorescence observables determined through FRET experiments.

1 Introduction

The first application of fluorescence-detected resonant energy transfer (FRET) to study the structure of a biological system was conducted over four decades ago.¹ While FRET was used by a number of groups in the late 20th century, it has become extremely common in the most recent decade as it has been adopted by single-molecule spectroscopists as well as some high-throughput screening methods. Today FRET is applied to a wide variety of systems using a wide variety of fluorescence modalities ranging from femtosecond time-resolved fluorescence measurements of bulk samples^{2,3} to (essentially) steady-state measurements of single molecules.^{4–10} This wide range of methodology and application raises some concerns about the validity of various assumptions that underlie FRET. In particular, some experiments are drawing very quantitative conclusions from FRET measurements that have been performed in settings quite unlike those in which the assumptions were originally vetted in the 1970s and 80s.^{1,11–17} Thus, a number of groups have worked to provide a better understanding the range of validity for the basic FRET

assumptions^{18–32} as well as the impact of a number of parameters (e.g. fluctuations in dye brightness) on the experimental FRET observables, especially in single-molecule FRET (smFRET) experiments.^{22,33–40}

The most troubling FRET assumptions are connected with the motions of the donor (D) and acceptor (A) probe molecules. The most well-known is the assumption that both the D and A are free to rapidly sample all possible orientations. This isotropic assumption results in the orientation factor, κ (see below), averaging to $\langle \kappa^2 \rangle = 2/3$. When making the assumption of isotropic motion, one also assumes that the distance between the D and A is uncorrelated with the orientations of the probes.²⁸ Since these and other FRET assumptions relate to the motions of the D and A, classical molecular dynamics (MD) simulations can provide increased understanding of their validity. While there are a number of examples of MD simulations helping to inform FRET experiments (we cite just a few here),^{41–45} the connection is often at the level of an average distance or average structural parameter. While useful, these comparisons fail to take advantage of the detail present in both the experimental data and the simulation.

One exception is a recent paper from Unruh et al⁴⁶ in which a Markov chain sampling technique was used with the MD simulation results to generate modeled D fluorescence decays. These decays were fit and compared to fits of the experimental fluorescence decays. A similar sampling technique has been presented recently by Grubmüller and coworkers who examined FRET in a series of polyproline peptides and compared simulation results directly with experiment.⁴⁷ In the present work, we employ the Markov chain and other approaches to further explore how to computationally generate experimental observables directly from molecular simulations. While we focus on fluorescence in this work, we note that the methods examined here should also be applicable to other experimental techniques where structural fluctuations impact the observable.

We believe this type of exploration is of general importance because computational modeling may be able to help experiments push toward more quantitative results in more complicated systems. Simulations can provide critical detail about the validity of various assumptions that are required for analysis of the data. For FRET experiments in particular, directly modeling the experimental observables avoids many of the usual FRET assumptions and, therefore, allows careful comparison to experiment even in cases where the assumptions are invalid or simply where more molecular detail is needed. Thus, bridging the gap between experiment and simulation may provide an important step toward more quantitative analysis of FRET data.

Connecting simulation and experiment involves several difficulties, the first of which—obtaining an accurate and complete sampling of configuration space during the MD simulation—we will not attempt to address here. However the next two steps of the process will be examined. They involve translating the trajectory of structures into a trajectory of D-A electronic coupling strength and then translating the coupling trajectory into experimental observables.

We utilize two different methods to connect structure with D-A coupling. In the first we perform a single quantum mechanical (QM) calculation on each fluorescent probe to characterize the excited state behavior, specifically the transition dipole moment, and then map those transition dipoles onto the configuration space sampled by MD. The transition dipoles are then interacted through the ideal dipole approximation (IDA) to determine the electronic coupling between the probes. It is this coupling that mediates the electronic excitation transfer (EET) and is therefore essential to understanding FRET. In the second method we perform a QM calculation on each snapshot, using the linear response (LR)

formalism developed by Mennucci and coworkers^{23,29} to calculate the electronic coupling without need for the IDA. In either case, the result is a trajectory of the electronic coupling versus time which one would like to map onto the fluorescence behavior observed by experiment. In the sections that follow, we describe three methods for sampling from this trajectory of coupling strength with the goal of directly simulating the fluorescence observables relevant to a FRET experiment.

Three different systems are explored in these simulations (Figure 1). The first is based on the hen egg-white lysozyme (HEWL) protein, in which the D (7-dimethylamino-4-methylcoumarin — *DACM*) is attached via a succinimide linkage as in typical maleimide conjugation chemistry. The A (Eosin Y — *eosin*) binds noncovalently into a pocket meaning that its orientation is essentially fixed relative to the protein.^{48–52} Having the A bound in this way is similar to studies utilizing intrinsic fluorophores, GFP-related probes,⁵³ or any of the FAsH family of probes.⁵⁴ This HEWL system was used previously to examine the effects of deviation from the well-known $\langle \kappa^2 \rangle = 2/3$ approximation (see **Theory**) and of correlation between D/A orientation (κ^2) and separation (R_{DA}).²⁸ The second and third systems are based on a simple double-stranded DNA oligonucleotide. The sequence was chosen to provide a standard B-form DNA structure with the D (Pacific Blue — *PB*⁵⁵) always at the 5' end of one strand (which we term the 0 position) and the A (fluorescein) at either the 5' end of the antisense strand (the 0' position) or in the middle of the antisense strand (the 16' position). The two samples are referred to as Oligo0-0', which has roughly 25 nucleotides between the probes and Oligo0-16' with roughly 10 nucleotides between them. For each of the three systems we simulate a histogram of FRET efficiencies and the decay of D fluorescence for comparison with single-molecule FRET and bulk time-resolved fluorescence experiments, respectively.

In the following sections we first provide some background on the theory used to describe FRET. Then the computational methods used to sample configuration space, to follow the excited state behavior, and the three methods used to sample the coupling strength trajectory are described. In **Results and Discussion** we use the IDA-derived coupling trajectory to characterize the methods individually and then compare their results, using two modeled experimental observables—the single-molecule FRET (smFRET) histogram and the time-resolved fluorescence decay of the FRET donor (trFRET). Finally, we compare the results of the IDA and LR calculations of coupling and discuss how we believe simulation can be used in the future for comparison to FRET experiments.

2 Theory

The key event in a FRET experiment is the transfer of electronic excitation from the D to the A. This EET process was described theoretically by Förster in the 1940's.^{56,57} His description has been extremely successful, with the vast majority of FRET papers using some form of the Förster equation:

$$k_{EET} = \frac{9000(\ln 10)\kappa^2\varphi_D}{128\pi^5 N_A n^4 R_{DA}^6 \tau_D} \int_0^\infty d\tilde{\nu} N f_D(\tilde{\nu}) \epsilon_A(\tilde{\nu}) \tilde{\nu}^{-4} \quad (1)$$

where κ is the orientation factor, $\kappa = \hat{\mu}_D \cdot \hat{\mu}_A - 3(\hat{\mu}_D \cdot \hat{R}_{DA})(\hat{\mu}_A \cdot \hat{R}_{DA})$, $\vec{\mu}_i$ are transition dipole moments ($\hat{\mu}_i$, unit vectors), $\hat{R}_{DA}$ is the unit vector along the line connecting the centers of the D and A, R_{DA} is the distance between the D and A, n is the index of refraction, N_A is Avogadro's number, and φ_D and τ_D are the D fluorescence quantum yield and lifetime in the absence of the A. The integral is known as the spectral overlap, which involves the normalized emission spectrum of the D, Nf_D , and the absorption spectrum of the acceptor,

ε_A , integrated in this case over wavenumbers, $\tilde{\nu}$. We point the reader to several excellent reviews for further information.^{32,56,58–62} Note that there are two contributions to the right hand side of Eq. 1, the electronic coupling between the D and A that enables the transfer of energy and the requirement that energy be conserved, which limits energy transfer to those D/A pairs with matching ground to excited state energy gaps. This separation of the coupling and conservation terms can be seen more clearly by backing up to a Golden Rule expression for EET

$$k_{EET} = \frac{2\pi}{\hbar} |V|^2 J_{DA} \quad (2)$$

where V is the electronic coupling between the donor and acceptor and J_{DA} is the overlap of the Frank-Condon envelopes that are responsible for the line shapes of the D and A transitions. Förster blended the V and J_{DA} terms to derive Eq. 1, which has proven extremely useful experimentally. He also made several assumptions, the most critical to the current work being the IDA. That is, Förster assumed that the electronic coupling that mediates EET could be described by the Coulombic interaction between D and A transition dipoles, i.e. that $V \approx V_{Coul} \approx V_{dip-dip}$. While we will utilize the IDA for most of the results presented here, we will also utilize the more rigorous linear response (LR) method^{23,29} (described below) to examine limitations of the IDA.

In addition to the IDA, most researchers make three further assumptions:^{28,59,63,64} i) that all of the spectral parameters (ϕ_D , τ_D , $f_D(\tilde{\nu})$, and $\varepsilon_A(\tilde{\nu})$) are the same for all D-A pairs in the ensemble as well as constant with respect to structural/solvent fluctuations, ii) that D and A sample orientation space isotropically such that $\langle \kappa^2 \rangle = 2/3$, and iii) that fluctuations in κ and R_{DA} are not correlated. We will also utilize the first assumption, which is primarily expressed by a constant spectral overlap term. In contrast, we will carefully address the second and third assumptions. These depend entirely on the structural dynamics of the D and A and can therefore be estimated through computational modeling of the structural fluctuations of the system.

3 Methods

The above theoretical considerations suggest that successful modeling of a FRET experiment requires: a) a time-ordered ensemble of structural configurations that represent the sample, b) a measure of the electronic coupling between the D and A probe molecules for all (or perhaps some subset) of those configurations, and c) an effective way to sample the resulting trajectory of electronic coupling and translate it into the experimental observables of interest. In this work we use classical MD simulations for a). We employ QM calculations in two ways for b) depending on whether we utilize the IDA or the LR approach. Our primary focus is evaluating three different methods for c), which are described below.

3.A MD simulations

For the HEWL system, MD was performed as described previously²⁸ using the AMBER suite of software.^{65,66} Briefly, a crystal structure (PDB ID: 1BWH) was used as the starting coordinates for the protein. Residue 47 was changed from threonine to cysteine and bonded to DACM via a succinimide group. Eosin was pulled into the binding site using restraints during a short MD run. The Cornell et al. force field⁶⁷ was used for the protein; dye and linker parameters will be described below. Explicit solvent (with a buffer size of 20 Å) and counterions were added yielding a total system size of roughly 55,000 atoms. Particle mesh

Ewald^{68,69} was used for electrostatic interactions, a 9 Å cutoff for Van der Waals interactions, SHAKE was applied to all bonds with hydrogen,⁷⁰ and the timestep was 2 fs. Equilibration of the system was determined by seeing that the drifts of total energy and system density (measured by the slopes of linear fits to the last 25 ps of equilibration) were < 5% of the short-term fluctuations (measured by the standard deviation). MD was run in the isothermal-isobaric (NPT) ensemble with temperature and pressure coupling constants of 10 ps. The eosin remained bound without restraints throughout the 60 ns of production simulation. Snapshots were saved every 200 fs.

Double-stranded DNA systems were built in canonical B-form starting structures and MD was performed using AMBER. The oligo (sense) sequence is D-GGTCATCTCAGCAAAGTCCAGTCGG where the D- indicates the donor dye and C6 linker connected to the phosphate-sugar backbone such that they are the zeroth residue. One of the antisense oligo strand sequences is A-CCGACTGGACTTTGCTGAGATGACC where the A- indicates the acceptor dye and linker connected to the phosphate-sugar backbone at the 0' position. The other antisense oligo sequence is CCGACTGGACTTTGCT*GAGATGACC where the T* indicates a modified thymine with linker and A dye attached to the nucleotide itself at the 16' position (see Fig 1). Thus in the Oligo0-0' system the D and A are 25 nucleotides apart and in the Oligo0-16' system the dyes are much closer, roughly 10 nucleotides apart. Dyes were attached by either extending one of the strands with an additional residue (for 5' attachment) or by replacing a nucleotide with a dye-linker-nucleotide unit (for internal attachment). The ff99 force field^{71,72} was used to describe the DNA with linker and dye parameters described below. Dyes were initially placed with their linkers in an extended conformation. Energetically unfavorable conformations were relieved through brief minimization in implicit solvent. Solvent was modeled implicitly using the Generalized Born (GB) model II of Onufriev, Bashford, and Case⁷³ (igb=5) with mbondi2 radii and counterion screening equivalent to a salt concentration of 0.1 M. A 12 Å cutoff was used for nonbonded interactions, SHAKE was applied to bonds with hydrogen atoms,⁷⁰ Langevin dynamics (with a collision frequency of 2ps^{-1}) was used to regulate the temperature, and the timestep was 1 fs. Equilibration was determined by seeing that the drifts in total energy and GB solvation energy (measured by the slopes of linear fits to the last 25 ps of equilibration) were small (< 5%) relative to their short-term fluctuations (measured by the standard deviation). Roughly 350 ns of production simulation was run for each of the 0-0' and 0-16' oligos with snapshots saved every 100 fs.

Dye and linker residues were parameterized by analogy with the Cornell et al. force field⁶⁷ and the generalized AMBER force field (GAFF)⁷⁴ with the exception one atom type. All of the dyes contain a heterocyclic (aromatic) oxygen atom that is not present in either force field. Complete description of the aromatic oxygen (type OA) parameter derivation as well as the charges and atom types for the DACM and eosin probes are given in VanBeek et al.,²⁸ while charges and atom types for Pacific Blue (PB) and fluorescein are provided as supplementary information to this article.

3.B QM calculations

The production MD simulations provided a trajectory of snapshots for each system. For each snapshot an estimate of the D-A electronic coupling that mediates the transfer of energy was needed. We estimated the coupling in two ways: 1) using the transition-dipoles of the D and A (the IDA), and 2) using the full transition densities of the D and A.

For the transition-dipole coupling, $V_{dip-dip}$, we use a single QM calculation on an optimized structure of the fluorophore to determine the magnitude and orientation of the transition dipole within the molecular framework of the molecule. The two dipoles are then mapped

onto each snapshot from the MD simulation and their interaction is determined through the standard dipole-dipole approach

$$V_{dip-dip} = \frac{\kappa |\vec{\mu}_D| |\vec{\mu}_A|}{4\pi\epsilon_0 n^2 R_{DA}^3}$$

where all the involved quantities have been defined in Eq.1. The transition dipoles were determined using time-dependent density functional theory (TD-DFT)^{75,76} employing the B3LYP functional and the 6-31G(d) basis set.^{77,78} Note that in this case the magnitudes of the D and A transition dipoles are fixed, while the orientations and distance change through structural fluctuations estimated via the MD simulation. This implementation accounts for motions of the fluorophores and the resulting effects on κ and R_{DA} within the limitations of the IDA.

We and a number of other groups have sought to improve on the IDA with a variety of methods that capture more detail from the transition density than simply its dipole moment.^{18-21,23,29,30,32,79} In particular, methods that utilize the full D and A transition densities have been coded into the popular packages Gaussian^{23,80} and Q-Chem.^{30,81} The former is termed the Linear Response (LR) method and is what we utilize in this work. Details of the LR calculation are given elsewhere,^{23,29} though we present a brief overview here. The most significant contribution to the LR coupling is the Coulombic interaction of the D and A transition densities

$$V_{Coul} = \int d\vec{r}_D \int d\vec{r}_A \frac{\rho_A^{T*}(\vec{r}_A) \rho_D^T(\vec{r}_D)}{4\pi\epsilon_0 |\vec{r}_D - \vec{r}_A|} \quad (3)$$

in which the $\rho_i^T(\vec{r}_i)$ are the transition densities that retain the three-dimensional structural information of the molecules. In contrast, when using the IDA, the details of the individual D and A transition densities are first averaged into the transition dipoles, such that those details are lost, and then the interaction is determined. In the LR method, the full structural details of the individual D and A transition densities are accounted for during calculation of the interaction. Thus, fluctuations in the transition orientation and magnitude relative to the molecular framework are accounted for. In addition, the total LR coupling includes the Coulombic interaction above (Eq. 3) as well as exchange and correlation contributions (though the correlation contribution must also be supported by the quantum mechanical method, e.g. with TD-DFT) and overlap interactions:

$$V_{LR} = \int d\vec{r}_D \int d\vec{r}_A \rho_A^{T*}(\vec{r}_A) \left(\frac{1}{4\pi\epsilon_0 |\vec{r}_D - \vec{r}_A|} + g_{xc}(\vec{r}_A, \vec{r}_D) \right) \rho_D^T(\vec{r}_D) - \omega_{D/A} \int d\vec{r}_D \int d\vec{r}_A \rho_A^{T*}(\vec{r}_A) \rho_D^T(\vec{r}_D) \quad (4)$$

where $\omega_{D/A}$ is the common transition energy of both monomers and g_{xc} is the exchange-correlation operator. Note that the exchange-correlation and overlap contributions are expected to be negligible for the systems examined here because they only contribute when the D and A are very close together; nonetheless these contributions are included in the LR formalism and were evaluated. In this work, we determine the transition densities using TD-DFT employing the B3LYP functional^{75,76} and the 6-31G(d) basis set.^{77,78} We showed previously that this model chemistry yields transition densities that are consistent with much higher level calculations for the types of fluorescent probes used here.³¹

Whether the IDA or LR approaches were used, the result is a trajectory of D-A electronic coupling strength as a function of time. The primary focus of this article is determination of how to properly sample these trajectories to yield a realistic model of fluorescence observables. The mechanics of three methods are described below with comparison of the results to follow in **Results and Discussion**.

3.C Trajectory Sampling Methods

The most straightforward approach is the *independent snapshot method (ISM)* in which each snapshot from the trajectory is treated as an independent member of the population. Thus, one assumes a D-A coupling trajectory with n snapshots contains n independent members of the complete population of coupling values that are possible. These n values can then be used to generate, for example, the intensity of the time-resolved decay of D fluorescence, $I(t)$, for comparison to bulk trFRET measurements or a histogram of FRET efficiencies, E_i , for comparison to smFRET data through

$$I(t) = \sum_{i=1}^n e^{-t(k_{EET,i} + \tau_D^{-1})} \quad (5)$$

and

$$E_i = \frac{R_{0,\kappa}^6}{R_{0,\kappa}^6 + \frac{R_{DA,i}^6}{\kappa_i^2}} \quad (6)$$

respectively, where τ_D is the fluorescence lifetime of the D in the absence of FRET, $k_{EET,i}$ is determined through

$$k_{EET,i} = \frac{1}{\tau_D} \kappa_i^2 \left(\frac{R_{0,\kappa}}{R_{DA,i}} \right)^6 \quad (7)$$

and $R_{0,\kappa}$ is the traditional Förster radius, R_0 , for the D and A with the $\langle \kappa^2 \rangle = 2/3$ value removed such that

$$R_{0,\kappa} = \left(\frac{3}{2} \right)^{\frac{1}{6}} R_0 \quad (8)$$

As will be shown below, there are several drawbacks to the ISM. The most obvious is that it assumes each D-A coupling point to be independent, while each MD snapshot results directly from the previous one. Though we skip some number of MD snapshots between each D-A coupling determination there is clearly some degree of correlation between neighboring coupling values. More importantly, there is no consideration of how the dynamics of the D-A coupling trajectory affect the observed rate constants.⁸² For example, if a certain configuration represents a very small D-A coupling and if that configuration is only occupied for very short periods of time, then that configuration is not of much

importance in determining the FRET efficiency of the real system. While no EET will occur from that configuration, the system will visit other, higher coupling configurations before the D leaves the excited state. Thus, snapshots with low coupling values should be of little importance if they are transiently occupied. The ISM, however, provides them with equal weight in Equations 5 and 6.

The *windowing method* provides some measure of accounting for transiently occupied states, by attempting to average them out. It is essentially the same as the ISM except that a number of snapshots are averaged together in an attempt to reduce the impact of any single (possibly outlying) snapshot. Thus, a D-A coupling trajectory of n points is reduced to n/w points where w is the window size.

The *Markov chain* method is an alternative averaging scheme that utilizes a Monte Carlo style sampling of the dynamics of the D-A coupling trajectory. At each trajectory point, probabilities are determined for four possible fates of excitation energy residing on the D: radiative decay of the D, nonradiative decay of the D, EET from D to A, and remaining in the D excited state. A random number is chosen and compared with the probabilities to indicate one of the four events. If remaining in the excited state is indicated, the chain continues by moving to the next trajectory point, calculating new probabilities, and selecting a new random number. If one of the three terminal events is indicated, that is recorded along with the time of the chain end point relative to the starting point. This is repeated with each trajectory point serving as a chain start point for some specified number of iterations, providing a distribution of terminal events for each start point. End-effects are avoided by stopping the procedure the first time a chain reaches the end of the trajectory without terminating. (For the simulations conducted here, less than 1 % of each trajectory failed to serve as a chain start point.) Probabilities of the terminal events were determined through

$$\begin{aligned} P_{rad} &= 1 - e^{-k_{rad}\Delta t} \\ P_{nonrad} &= 1 - e^{-k_{nonrad}\Delta t} \\ P_{EET,i} &= 1 - e^{-k_{EET,i}\Delta t} \end{aligned} \quad (9)$$

where

$$\begin{aligned} k_{rad} &= \frac{\varphi_D}{\tau_D} \\ k_{nonrad} &= \frac{1}{\tau_D} - k_{rad} \end{aligned} \quad (10)$$

and Δt is the time between trajectory points. The probability of remaining in the D excited state is simply $1 - (P_{rad} + P_{nonrad} + P_{EET,i})$. Note that only $P_{EET,i}$ varies with the trajectory in this work. P_{rad} and P_{nonrad} are based on average (tabulated) values of φ_D and τ_D . In principle, D excited state properties determined through QM calculations could be used to estimate fluctuations in φ_D and τ_D at each point in the trajectory, giving $\varphi_{D,i}$ and $\tau_{D,i}$, though this was not attempted in the current work. Modeling trFRET and smFRET data with this method requires the input of just three parameters (beyond the D-A electronic coupling): φ_D , τ_D , and R_0 . For the PB-Fluorescein pair these parameters are $\varphi_D = 0.78$, $\tau_D = 2.86$ ns, and $R_0 = 55$ Å. For the DACM-eosin pair these parameters are $\varphi_D = 0.42$, $\tau_D = 3.1$ ns, and $R_0 = 45$ Å.

After completion of all of the Markov chains, the D fluorescence decay was estimated by constructing a histogram of the delay times between chain initiation and chain termination

for each chain that terminated in a D radiative decay event. This is conceptually similar to acquisition of a time-correlated single photon counting decay profile.^{83,84} A FRET efficiency, E_i , was determined for each point in the trajectory by calculating the fraction of Markov chains initiated at that point that terminated in EET. These efficiencies were then gathered in a histogram to model experimental smFRET results.

Convergence of the modeled trFRET and smFRET data were tested carefully against both the number of iterations (Markov chains per trajectory point) and the spacing of the points in the D-A coupling trajectory. These data are given in Supplementary Information. Our conclusions from the convergence testing are that 500 (or more) iterations per point and a point spacing of 2 ps (or less) are sufficient to robustly capture quantitative details in both types of data. We note that the limitation on point spacing may be simply due to inadequate sampling because the total number of Markov chains is reduced as the number of trajectory points is reduced. Thus larger trajectory point spacings (e.g. 5 or 10 ps) may be adequate to capture the necessary dynamics if the trajectory were long enough. In this work we used 1000 iterations and a point spacing of 0.2 ps.

4 Results and Discussion

We first present comparison of the three methods used to sample the D-A coupling trajectory, followed by comparison of the dipole-dipole and LR methods used to determine the D-A coupling. Finally, implications of these results on the modeling of experimental FRET data are discussed.

Before beginning analysis of the MD simulations, we illustrate differences between the three sampling methods by applying them to two different model datasets. In the first dataset, the EET rate fluctuates rapidly between states of high and low rate. The values used in the trajectory are a high EET rate of 0.99 ns^{-1} , a low EET rate of 0.01 ns^{-1} and a τ_D of 2 ns. The rate changed from low to high every 10 ps such that the system explores both states many times during the natural lifetime of the excited state. The results are shown in terms of simulated smFRET histograms in Figure 2. The ISM yields a bimodal histogram with spikes at low and high FRET efficiency, whereas both the window (100 ps) and Markov methods have a single peak at a FRET efficiency of 0.5. Because the window time (100 ps) and the D excited state lifetime (2 ns) are both significantly longer than the timescale of the fluctuations in rate (10 ps), both methods completely average over the dynamics, yielding a single peak corresponding to the average EET rate. In contrast the ISM, which does no averaging, fails to account for the dynamics of the system. In the second model, the EET rate remains low for an extended period (480 ps) and then briefly transitions to the high rate (20 ps) before returning low. The system does sample both states several times during the excited state lifetime (2 ns), but not during the window time (100 ps). This model is analogous to a system of two probes with long lifetimes freely diffusing in solution, studied by several groups in the 1960s and 70s.^{11,85,86} The probes spend most of their time far apart (the low rate state) and briefly come near each other allowing the EET rate to spike. As shown in Figure 3, the ISM histogram again is bimodal corresponding to 96 % of the time spent with a low rate and 4 % of the time with a high EET rate. Because the window time is now shorter than the fluctuation time, its results are similar to the ISM. In contrast, the Markov method has a single peak at a modest FRET efficiency, matching what is expected from experiment.

These examples highlight the fact that the ISM does not account for the effects of changing EET rates on the FRET observables. If the D probe samples a variety of states with different EET rates during its excited state lifetime, then the ISM method fails. The windowing method does average over fast fluctuations, but it cannot adapt to account for dynamics that

occur at multiple timescales. The Markov method, however, properly accounts for the competing kinetic pathways in the system, averaging over the dynamics in a realistic way. Thus the impact of fluctuations in EET rate on the experimental observables is captured only by the Markov method.

Keeping those results in mind, we now apply the three methods to the MD simulations. Figure 4 shows both smFRET and trFRET results for the lysozyme simulation. The ISM (snapshot) results in a trFRET decay with a relatively long lifetime and a smFRET histogram with a spike at 0.99, a peak at 0.91, and a long tail extending to low FRET efficiency, including a small peak at 0. The window results yield progressively faster trFRET decays as the window size increases (1, 5, 10, 20, 50, and 100 ps are shown in the figure). As might be expected, the 10 ps Window results in a smFRET histogram with reduced probability at the two extremes relative to the ISM (except the spike at 0.99, which is essentially the same). The peak near 0.90 is slightly shifted to lower efficiency and the small peak at zero efficiency is eliminated, while the probability is increased through much of the tail region. The Markov chain results in a trFRET decay that is faster than even the longest window size and the smFRET histogram is qualitatively different. The tail at low efficiency has been completely eliminated (probabilities are exactly 0 below 0.69 efficiency) resulting in significantly more probability in the peak near 0.85 efficiency. However, the spike at 0.99 is very similar in all three methods. This spike arises because the D probe remains in a high EET state for an extended period of the MD simulation, from the 39th to 46th ns (cf. Figure 7). This state is similar to the structure shown in Figure 1 (family C) of VanBeek et al.²⁸ except that the D is even closer to the A and the state is much longer lived. Because effectively no dynamics occur during this ≈ 7 ns period and the EET rate is very high, it impacts all three methods with the same spike at essentially unit FRET efficiency.

The Oligo0–16' results, shown in Figure 5, are similar to lysozyme without the high efficiency spike. The ISM gives the slowest trFRET decay with larger window sizes giving uniformly faster decays and the Markov chain giving the fastest decay. The ISM smFRET histogram again shows a long tail toward low efficiency with a small peak at 0. The window result has a reduced low efficiency tail and small shift of the peak toward lower efficiency. For the Markov chain the tail is completely eliminated and the high efficiency peak contains significantly more probability, though it does not appear to be shifted relative to the window results. In fact, the Markov results retain significantly more probability than the window method at very high efficiency—more similar to the ISM.

While both lysozyme and Oligo0–16' exhibit high FRET efficiencies, the Oligo0–0' results, shown in Figure 6, are significantly different displaying very low FRET efficiency (less than 0.05). In this extreme case the ISM method results in a clearly unphysical smFRET histogram sharply peaked at zero with a tail extending to higher efficiency. Note that the bin width is 0.005 in Figure 6. The window method (100 ps) yields a peak at 0.02 with a similar tail extending to higher efficiency. The Markov model again has the most well-defined peak, this time centered near 0.035, with a minimal tail.

As observed for the two model systems, the ISM tends to yield smFRET histograms that are unphysical, putting too much emphasis on snapshots that are at the extremes of very high or very low FRET efficiency. In contrast, the Markov method appropriately averages these snapshots with others. Even in the Oligo0–0' system that is dominated by low efficiency configurations and where the magnitude of the fluctuations in the EET rate is small, the ISM yields a histogram in which the low efficiency states are weighted unphysically high. Configurations with modest FRET efficiencies that are visited regularly are likely to result in an EET event and therefore should outweigh the very low efficiency configurations. The window method provides an improvement because it averages over these moderate EET

states. However, the Markov method accounts for the competing kinetic pathways in the system, giving experimentally viable observables across all FRET efficiencies and all timescales of fluctuations.

The different impact of rate fluctuations in the three methods is illustrated in Figure 7, which gives the trajectories of FRET efficiencies for each. Because every snapshot yields an 'observed' FRET efficiency, the ISM exhibits massive fluctuations and the presence of the zero efficiency states is obvious. These fluctuations are moderated in the window data and the elimination of the zero-efficiency peak is clear, though the fluctuations are still quite large. The Markov trajectory is much more moderated as it removes fluctuations on both the low and high efficiency extremes. This ability of the Markov method to account for changes in EET rate is illustrated in Figure 8, which shows a trajectory of $k_{EET,i}$ for each snapshot overlaid with the number of visits each snapshot received during the Markov chain simulation. Because both k_{rad} and k_{nonrad} are fairly small for PB and DACM, snapshots with low k_{EET} rarely terminate, passing the D excitation onto the next snapshot. Thus, a stretch of snapshots with low EET rates tend to have a uniform and large number of visits. The first snapshot with high k_{EET} that follows is likely to lead to a terminal EET event, such that later snapshots tend to have fewer visits. Thus, the Markov chain method incorporates, not just information about the k_{EET} of the snapshot itself, but also those that precede it. In addition, this is not done in a blanket averaging fashion as in the window method, but in a dynamic fashion determined by the trajectory itself. A similar plot of the number of visits for either the ISM or window methods would simply be a flat line. Thus, the Markov chain method provides a realistic sampling of the D-A coupling trajectory, enabling modeling of fluorescence observables for comparison to experiment.

While we have confidence in the Markov method for sampling the trajectory of D-A coupling values, a number of questions remain about how best to generate those D-A coupling values in the first place. One of the largest of those questions is the viability of the ideal dipole approximation (IDA).^{31,79} We have investigated this here by using both a transition dipole-transition dipole treatment of the coupling and also the LR method described above. The electronic structure calculations needed for the LR method are taxing, and therefore an exhaustive investigation of even just one of the simulation systems was not possible. We choose several short portions of the Oligo0–16' and lysozyme trajectories that exhibited high D-A coupling. All regions showed significant differences between the IDA and LR coupling values. For the following discussion we will focus on the longest of those, a 20 ns portion of the Oligo0–16' trajectory.

A comparison of the resulting k_{EET} values are shown in Figure 9. The average LR rate, 5.85 ns^{-1} , is indicated on the y-axis showing that the IDA rate differs from the V_{LR} rate by a large amount. In fact, the average difference in rate is 9.44 ns^{-1} , meaning that the $V_{dip-dip}$ coupling provides an EET rate that is, on average, a factor of 2.6 times too high. This is illustrated in Figure 10, which shows trFRET and smFRET data modeled by applying the Markov chain method to the QM LR couplings as well as the simpler $V_{dip-dip}$ model used for the data presented in Figures 4 through 8. It is clear from Figure 10 that overestimation of the D-A coupling by the $V_{dip-dip}$ approximation has a large impact on the resulting fluorescence observables. Both the trFRET and the smFRET data show large deviations in V_{LR} rates versus $V_{dip-dip}$ rates. This is particularly noteworthy in the trFRET data which showed only modest dependence on sampling method. One should bear in mind that the above analysis was conducted on the 20 ns of Oligo0–16' trajectory most likely to show large differences between the LR and IDA approaches. However, it is clear that use of the IDA can lead to significant errors in situations where FRET efficiencies are high and that caution should be exercised when comparing FRET efficiency to distance in these situations. Muñoz-Losa and coworkers examined the failure of the IDA in detail for a limited set of

idealized orientations for several commonly-used organic dyes. They found that the IDA performs poorly for each orientation out to distances of 50 Å and possibly more.³¹ However, averaging over many orientations (κ values) leads to cancellation of error such that the IDA appears to perform well down to D-A distances of roughly 20 Å. The 20 ns of simulation examined here, which demonstrate poor performance of the IDA, exhibit D-A distances ranging from roughly 20 to 30 Å in agreement with the conclusions of that earlier work. Note that the effectiveness of the IDA has not been determined for other parts of the Oligo0–16' trajectory with larger D-A distances.

5 Conclusion

We have provided a critical examination of two different methods for generating a D-A electronic coupling trajectory from a MD trajectory and three methods for sampling that coupling trajectory. Together, these methods provide a mechanism for directly connecting the structural dynamics modeled by MD with fluorescence observables determined through FRET experiments.

For the three sampling methods, we found that the Markov chain method, because it accounts for fluctuations in the EET rate, provided more physically sensible results than either the ISM or windowed methods. This held true for three different systems exhibiting average FRET efficiencies ranging from quite high (~ 0.90) to very low (~ 0.05). In these cases the ISM method yielded smFRET histograms that were qualitatively different, with high asymmetry and long tails. The window method yielded results that lay somewhere between Markov and ISM, but more closely resembled the ISM results.

For the two coupling methods, we found that in regions of high D-A coupling, the IDA method strongly overestimated the coupling strength. For a limited 20 ns region of trajectory in the Oligo0–16' system, the IDA overestimated the EET rate by a factor of 2.6 leading to a heavily shifted smFRET histogram.

As FRET experiments push toward more quantitative results in more complicated systems, computational modeling may be able to help in two ways. First, simulations can provide critical detail about the validity of assumptions required for analysis of the data. Second, directly modeling the experimental observables avoids many of the usual FRET assumptions and, therefore, allows careful comparison to experiment even in cases where the assumptions are invalid or simply where more molecular detail is needed. The information presented here strengthens the ability of computation to provide support for detailed analysis of FRET experiments.

6 Supporting material

As referenced previously, charges and atom types for the PB and fluorescein probes as well as two figures showing the dependence of Markov chain sampling on the number of iterations and snapshot spacing are available at <http://www.acs.org>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

BPK acknowledges funding support from the NIH (GM082100), NSF-MRI (0520704), NSF-REU (0243828), ACS-PRF (42390-GB6), and by a grant from the Howard Hughes Medical Institute through the Undergraduate Science Education Program. DBV was supported by the Beckman Foundation. Some calculations were performed on TeraGrid resources (MCA07S034) and at the Hope College Computational Science and Modeling Laboratory.

BM acknowledges financial support from Gaussian Inc. In addition, BPK thanks Helmut Grubmüller, Martin Höfling, and Nathan Tintle for helpful discussions and Paul Frybarger for creation of some of the analysis scripts.

Abbreviations

A	acceptor
D	donor
DACM	7-dimethylamino-4-methylcoumarin (a donor probe)
EET	electronic excitation transfer
FRET	fluorescence-detected resonance energy transfer
HEWL	hen egg-white lysozyme
IDA	ideal dipole approximation
LR	linear response (as opposed to IDA)
MD	molecular dynamics
PB	Pacific Blue
QM	quantum mechanic(al/s)
smFRET	single-molecule FRET
TD-DFT	time-dependent density functional theory
trFRET	time-resolved fluorescence decay of the FRET donor

References

1. Stryer L, Haugland RP. Energy Transfer: A Spectroscopic Ruler. *Proc Natl Acad Sci USA*. 1967; 58:719–726. [PubMed: 5233469]
2. Millar DP. Time-resolved fluorescence methods for analysis of DNA-protein interactions. *Methods Enzymol*. 2000; 323:442–459. [PubMed: 10944763]
3. Yang M, Millar DP. Fluorescence resonance energy transfer as a probe of DNA structure and function. *Methods Enzymol*. 1997; 278:417–444. [PubMed: 9170325]
4. Deniz AA, Laurence TA, Dahan M, Chemla DS, Schultz PG, Weiss S. Ratiometric Single-molecule Studies of Freely Diffusing Biomolecules. *Annu Rev Phys Chem*. 2001; 52:233–253. [PubMed: 11326065]
5. Ha T. Single-Molecule Fluorescence Resonance Energy Transfer. *Methods*. 2001; 25:78–86. [PubMed: 11558999]
6. Barkai E, Jung Y, Silbey R. Theory of single-molecule spectroscopy: Beyond the ensemble average. *Annu Rev Phys Chem*. 2004; 55:457–507. [PubMed: 15117260]
7. Johnson, CK.; Slaughter, BD.; Unruh, JR.; Price, ES. Fluorescence Probes of Protein Dynamics and Conformations in Freely Diffusing Molecules: Single-molecule resonance energy transfer and time-resolved fluorescence methods. In: Geddes, CD.; Lakowicz, JR., editors. *Reviews in Fluorescence*. Vol. 3. Springer; 2006. p. 239-260.
8. Mukhopadhyay S, Deniz AA. Fluorescence from Diffusing Single Molecules Illuminates Biomolecular Structure and Dynamics. *J Fluoresc*. 2007; 17:775–783. [PubMed: 17641956]
9. Deniz AA, Mukhopadhyay S, Lemke EA. Single-molecule biophysics: at the interface of biology, physics and chemistry. *J R Soc Interface*. 2008; 5:15–45. [PubMed: 17519204]
10. Roy R, Hohng S, Ha T. A practical guide to single-molecule FRET. *Nat Methods*. 2008; 5:507–516. [PubMed: 18511918]
11. Steinberg IZ, Katchalski E. Theoretical Analysis of the Role of Diffusion in Chemical Reactions, Fluorescence Quenching, and Nonradiative Energy Transfer. *J Chem Phys*. 1968; 48:2404–2410.

12. Haugland RP, Yguerabide J, Stryer L. Dependence of the Kinetics of Singlet-Singlet Energy Transfer on Spectral Overlap. *Proc Natl Acad Sci USA*. 1969; 63:23–30. [PubMed: 16591747]
13. Dale RE, Eisinger J. Intramolecular distances determined by energy transfer. Dependence on orientational freedom of donor and acceptor. *Biopolymers*. 1974; 13:1573–1605.
14. Dale, RE.; Eisinger, J. Polarized Excitation Energy Transfer. In: Chen, RF.; Edelhoch, H., editors. *Biochemical Fluorescence: Concepts*. Vol. 1. Marcel Dekker; 1975. p. 115-284.
15. Stryer L. Fluorescence Energy Transfer as a Spectroscopic Ruler. *Annu Rev Biochem*. 1978; 47:819–846. [PubMed: 354506]
16. Dale RE. Fluorescence depolarization and orientation factors for excitation energy transfer between isolated donor and acceptor fluorophore pairs at fixed intermolecular separations. *Acta Phys Polon*. 1978; A54:743–756.
17. Beardsley K, Cantor CR. Studies of Transfer RNA Tertiary Structure by Singlet-Singlet Energy Transfer. *Proc Natl Acad Sci USA*. 1970; 65:39–46. [PubMed: 5263760]
18. Krueger BP, Scholes GD, Fleming GR. Calculation of Couplings and Energy-Transfer Pathways between the Pigments of LH2 by the *ab Initio* Transition Density Cube Method. *J Phys Chem B*. 1998; 102:5378–5386. 9603.
19. Tretiak S, Middleton C, Chernyak V, Mukamel S. Bacteriochlorophyll and Carotenoid Excitonic Couplings in the LH2 System of Purple Bacteria. *J Phys Chem B*. 2000; 104:9540–9553.
20. Jordanides XJ, Scholes GD, Fleming GR. The mechanism of energy transfer in the bacterial photosynthetic reaction center. *J Phys Chem B*. 2001; 105:1652–1669.
21. Wong KF, Bagchi B, Rossky PJ. Distance and Orientation Dependence of Excitation Transfer Rates in Conjugated Systems: Beyond the Förster Theory. *J Phys Chem A*. 2004; 108:5752–5763.
22. Schröder GF, Grubmüller H. Maximum likelihood trajectories from single molecule fluorescence resonance energy transfer experiments. *J Chem Phys*. 2003; 119:9920–9924.
23. Iozzi MF, Mennucci B, Tomasi J, Cammi R. Excitation energy transfer (EET) between molecules in condensed matter: A novel application of the polarizable continuum model (PCM). *J Chem Phys*. 2004; 120:7029–7040. [PubMed: 15267604]
24. Ortiz W, Krueger BP, Kleiman VD, Krause JL, Roitberg AE. Energy Transfer in the Nanostar: The Role of Coulombic Coupling and Dynamics. *J Phys Chem B*. 2005; 109:11512–11519. [PubMed: 16852410]
25. Jean JM, Krueger BP. Structural Fluctuations and Excitation Transfer between Adenine and 2-Aminopurine in Single-Stranded Deoxytrinucleotides. *J Phys Chem B*. 2006; 110:2899–2909. [PubMed: 16471900]
26. Srinivas G, Bagchi B. Effect of Orientational Motion of Mobile Chromophores on the Dynamics of Förster Energy Transfer in Polymers. *J Phys Chem B*. 2001; 105:9370–9374.
27. Beierlein FR, Othersen OG, Lanig H, Schneider S, Clark T. Simulating FRET from Tryptophan: Is the Rotamer Model Correct? *J Am Chem Soc*. 2006; 128:5142–5152. [PubMed: 16608350]
28. VanBeek DB, Zwier MC, Shorb JM, Krueger BP. Fretting about FRET: Correlation Between κ and R . *Biophys J*. 2007; 92:4168–4178. [PubMed: 17384068]
29. Curutchet C, Scholes GD, Mennucci B, Cammi R. How solvent controls electronic energy transfer and light harvesting: Toward a quantum-mechanical description of reaction field and screening effects. *J Phys Chem B*. 2007; 111:13253–13265. [PubMed: 17973520]
30. Hsu C-P, You Z-Q, Chen H-C. Characterization of the Short-Range Couplings in Excitation Energy Transfer. *J Phys Chem C*. 2008; 112:1204–1212.
31. Muñoz-Losa A, Curutchet C, Krueger BP, Hartsell LR, Mennucci B. Fretting about FRET: Failure of the Ideal Dipole Approximation. *Biophys J*. 2009; 96:4779–4788. [PubMed: 19527638]
32. Scholes GD. Long-range Resonance Energy Transfer in Molecular Systems. *Annu Rev Phys Chem*. 2003; 54:57–87. [PubMed: 12471171]
33. Gopich IV, Szabo A. Photon counting histograms for diffusing fluorophores. *J Phys Chem B*. 2005; 109:17683–17688. [PubMed: 16853263]
34. Gopich IV, Szabo A. Theory of photon statistics in single-molecule Förster resonance energy transfer. *J Chem Phys*. 2005; 122:014707.

35. Lee NK, Kapanidis AN, Wang Y, Michalet X, Mukhopadhyay J, Ebricht RH, Weiss S. Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophys J*. 2005; 88:2939–2953. [PubMed: 15653725]
36. Nir E, Michalet X, Hamadani KM, Laurence TA, Neuhauser D, Kovchegov Y, Weiss S. Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments. *J Phys Chem B*. 2006; 110:22103–22124. [PubMed: 17078646]
37. Kalinin S, Felekyan S, Antonik M, Seidel CAM. Probability Distribution Analysis of Single-Molecule Fluorescence Anisotropy and Resonance Energy Transfer. *J Phys Chem B*. 2007; 111:1025310262.
38. Gopich IV, Szabo A. Single-molecule FRET with diffusion and conformational dynamics. *J Phys Chem B*. 2007; 111:12925–12932. [PubMed: 17929964]
39. Kalinin S, Sisamakakis E, Magennis SW, Felekyan S, Seidel CAM. On the Origin of Broadening of Single-Molecule FRET Efficiency Distributions beyond Shot Noise Limits. *J Phys Chem B*. 2010; 114:6197–6206. [PubMed: 20397670]
40. Sisamakakis E, Valeri A, Kalinin S, Rothwell PJ, Seidel CAM. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods Enzymol*. 2010; 475:455–514. [PubMed: 20627168]
41. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA. Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc Natl Acad Sci USA*. 2007; 104:1528–1533. [PubMed: 17251351]
42. Schuler B, Lipman EA, Eaton WA. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*. 2002; 419:743–747. [PubMed: 12384704]
43. Schuler B, Lipman EA, Steinbach PJ, Kumke M, Eaton WA. Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *Proc Natl Acad Sci USA*. 2005; 102:2754–2759. [PubMed: 15699337]
44. Margittai M, Widengren J, Schweinberger E, Schröder G, Felekyan S, Hausteine E, König M, Fasshauer D, Grubmüller H, Jahn R, Seidel CAM. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proc Natl Acad Sci USA*. 2003; 100:15516–15521. [PubMed: 14668446]
45. Woźniak AK, Schröder G, Grubmüller H, Seidel CAM, Oesterhelt F. Single-molecule FRET measures bends and kinks in DNA. *Proc Natl Acad Sci USA*. 2008; 105:1833718342.
46. Unruh JR, Kuczera K, Johnson CK. Conformational Heterogeneity of a Leucine Enkephalin Analogue in Aqueous Solution and Sodium Dodecyl Sulfate Micelles: Comparison of Time-Resolved FRET and Molecular Dynamics Simulations. *J Phys Chem B*. 2009; 113:14381–14392. [PubMed: 19780516]
47. Hoefling M, Lima N, Seidel CAM, Schuler B, Grubmüller H. Structural Heterogeneity and Quantitative FRET Efficiency Distributions of Polyprolines through a Hybrid Atomistic Simulation and Monte Carlo Approach. *PLoS ONE*. 2011 in press.
48. Kepka AG, Grossweiner LI. Photodynamic Inactivation of Lysozyme by Eosin. *Photochem Photobiol*. 1973; 18:49–61. [PubMed: 4795549]
49. Baugher JF, Grossweiner LI, Lewis C. Intramolecular energy-transfer in lysozyme-eosin complex. *J Chem Soc Faraday II*. 1974; 70:1389–1398.
50. Chang MC, Cross AJ, Fleming GR. Internal Dynamics and Overall Motion of Lysozyme Studied by Fluorescence Depolarization of the Eosin Lysozyme Complex. *J Biomol Struct Dyn*. 1983; 1:299–318. [PubMed: 6086059]
51. Cross AJ, Fleming GR. Influence of inhibitor binding on the internal motions of lysozyme. *Biophys J*. 1986; 50:507–512. [PubMed: 3756301]
52. Jordanides XJ, Lang MJ, Song X, Fleming GR. Solvation Dynamics in Protein Environments Studied by Photon Echo Spectroscopy. *J Phys Chem B*. 1999; 103:7995–8005.
53. Shaner NC, Steinbach PA, Tsien RY. A Guide to Choosing Fluorescent Proteins. *Nature Methods*. 2005; 2:905–909. [PubMed: 16299475]
54. Adams SR, Campbell RE, Gross LA, Martin BR, Walkup GK, Yao Y, Llopis J, Tsien RY. New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *J Am Chem Soc*. 2002; 124:6063–6076. [PubMed: 12022841]

55. Sun W-C, Gee KR, Haugland RP. Synthesis of novel fluorinated coumarins: Excellent UV-light excitable fluorescence dyes. *Bio Med Chem Lett*. 1998; 8:3107–3110.
56. Förster, T. Delocalized Excitation and Excitation Transfer. In: Sinanoglu, O., editor. *Modern Quantum Chemistry*. Vol. III. Academic Press, Inc; 1965. p. 93-137.
57. Förster T. Intermolecular energy transference and fluorescence. *Ann Physik*. 1948; 2:55–75.
58. Wu P, Brand L. Resonance energy transfer: methods and applications. *Anal Biochem*. 1994; 218:1–13. [PubMed: 8053542]
59. Van Der Meer, BW.; Coker, GI.; Chen, SYS. *Resonance Energy Transfer: Theory and Data*. VCH Publishers, Inc; 1994.
60. Clegg, RM. Fluorescence Resonance Energy Transfer. In: Wang, XF.; Herman, B., editors. *Fluorescence Imaging and Microscopy*, Wiley - Interscience, volume 137 of *Chemical Analysis*. 1996. p. 179-252.
61. Andrews, DL.; Demidov, AA. *Resonance Energy Transfer*. John Wiley & Sons; 1999.
62. Selvin PR. The renaissance of fluorescence resonance energy transfer. *Nat Struct Biol*. 2000; 7:730–734. [PubMed: 10966639]
63. Dale RE, Eisinger J, Blumberg WE. The orientational freedom of molecular probes: The orientation factor in intramolecular energy transfer. *Biophys J*. 1979; 26:161–194. [PubMed: 262414]
64. Wu P, Brand L. Orientation Factor in Steady-State and Time-Resolved Resonance Energy Transfer Measurements. *Biochem*. 1992; 31:7939–7947. [PubMed: 1510980]
65. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Merz, KM.; Wang, B.; Pearlman, DA.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, J.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, JW.; Ross, WS.; Kollman, PA. AMBER 8. University of California; San Francisco, San Francisco, CA: 2004.
66. Case, D.; Darden, T.; Cheatham, TE., III; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, RC.; Zhang, W.; Merz, K.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvy, I.; Wong, KF.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D.; Seetin, M.; Sagui, C.; Babin, V.; Kollman, P. AMBER 10. University of California; San Francisco, San Francisco, CA: 2008.
67. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A Second Generation Force Field For the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J Am Chem Soc*. 1995; 117:5179–5197.
68. Darden T, York D, Pedersen L. Particle mesh Ewald: An $N \log N$ method for Ewald sums in large systems. *J Chem Phys*. 1993; 98:10089–10092.
69. York DM, Darden TA, Pedersen LG. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *J Chem Phys*. 1993; 99:8345–8348.
70. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of cartesian equations of motion of a system with constraints – molecular dynamics of n-alkanes. *J Comput Phys*. 1977; 23:327–341.
71. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem*. 2000; 21:1049–1074.
72. Hornak V, Abel R, Okur A, Strockbine B, Roitberg AE, Simmerling C. Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins*. 2006; 65:712–725. [PubMed: 16981200]
73. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins*. 2004; 55:383–394. [PubMed: 15048829]
74. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and Testing of a General Amber Force Field. *J Comput Chem*. 2004; 25:1157–1174. [PubMed: 15116359]
75. Stratmann RE, Scuseria GE, Frisch MJ. An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules. *J Chem Phys*. 1998; 109:8218–8224.

76. Casida ME, Jamorski C, Casida KC, Salahub DR. Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold. *J Chem Phys.* 1998; 108:4439–4449.
77. Hariharan PC, Pople JA. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor Chim Acta.* 1973; 28:213–222.
78. Hariharan PC, Pople JA. The effect of d-functions on molecular orbital energies for hydrocarbons. *Chem Phys Lett.* 1972; 16:217–219.
79. Chang JC. Monopole effects on electronic excitation interactions between large molecules. I. Application to energy transfer in chlorophylls. *J Chem Phys.* 1977; 67:3901–3909.
80. Frisch, MJ.; Trucks, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Montgomery, JJA.; Vreven, T.; Kudin, KN.; Burant, JC.; Millam, JM.; Iyengar, SS.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, GA.; 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, JE.; Hratchian, HP.; Cross, JB.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, RE.; Yazyev, O.; Austin, AJ.; Cammi, R.; Pomelli, C.; Ochterski, JW.; Ayala, PY.; Morokuma, K.; Salvador, GAVP.; Dannenberg, JJ.; Zakrzewski, VG.; Dapprich, S.; Daniels, AD.; Strain, MC.; Farkas, O.; Malick, DK.; Rabuck, AD.; Raghavachari, K.; Foresman, JB.; Ortiz, JV.; Cui, Q.; Baboul, AG.; Clifford, S.; Cioslowski, J.; Stefanov, BB.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, RL.; Fox, DJ.; Keith, T.; Al-Laham, MA.; Peng, CY.; Nanayakkara, A.; Challacombe, M.; Gill, PMW.; Johnson, B.; Chen, W.; Wong, MW.; Gonzalez, C.; Pople, JA. Gaussian 03. Gaussian, Inc; Wallingford, CT: 2004.
81. Shao Y, Fusti-Molnar L, Jung Y, Kussmann J, Ochsenfeld C, Brown ST, Gilbert AT, Slipchenko LV, Levchenko SV, O'Neill DP, DiStasio RA Jr, Lochan RC, Wang T, Beran GJ, Besley NA, Herbert JM, Lin CY, Voorhis TV, Chien SH, Sodt A, Steele RP, Rassolov VA, Maslen PE, Korambath PP, Adamson RD, Austin B, Baker J, Byrd EFC, Daschel H, Doerksen RJ, Dreuw A, Dunietz BD, Dutoi AD, Furlani TR, Gwaltney SR, Heyden A, Hirata S, Hsu C-P, Kedziora G, Khalliulin RZ, Klunzinger P, Lee AM, Lee MS, Liang W, Lotan I, Nair N, Peters B, Proynov EI, Pieniazek PA, Rhee YM, Ritchie J, Rosta E, Sherrill CD, Simmonett AC, Subotnik JE, Woodcock HL III, Zhang W, Bell AT, Chakraborty AK, Chipman DM, Keil FJ, Warshel A, Hehre WJ, Schaefer HF III, Kong J, Krylov AI, Gill PM, Head-Gordon M. Advances in quantum chemical methods and algorithms in the Q-Chem 3.0 program package. *Phys Chem Chem Phys.* 2006; 8:3172–3191. [PubMed: 16902710]
82. Troisi A, Nitzan A, Ratner MA. A rate constant expression for charge transfer through fluctuating bridges. *J Chem Phys.* 2003; 119:5782–5788.
83. O'Connor, DV.; Phillips, D. Time-Correlated Single Photon Counting. Academic Press; 1984.
84. Becker, W. Advanced Time-Correlated Single Photon Counting Techniques. Springer; 2005.
85. Thomas DD, Carlsen WF, Stryer L. Fluorescence Energy Transfer in the Rapid-Diffusion Limit. *Proc Natl Acad Sci USA.* 1978; 75:5746–5750. [PubMed: 16592590]
86. Elkana Y, Feitelson J, Katchalski E. Effect of Diffusion on Transfer of Electronic Excitation Energy. *J Chem Phys.* 1968; 48:2399–2404.
87. Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics. *J Molec Graphics.* 1996; 14:33–38.

Figure 1.

(Center) Structures of the four probe molecules with linkers investigated in this work, along with (left) cartoon representations of the three individual oligonucleotides and (right) a snapshot of the lysozyme system. In the lysozyme snapshot (image created with VMD87), the protein (gray) is shown with stick bonds while the DACM (purple/dark gray) and eosin (orange/light gray) are shown as Van der Waals spheres.

Figure 2.

Comparison of the three sampling methods for a contrived dataset in which the data alternate rapidly between a high EET rate state and a low EET rate state. The Markov Chain method appropriately averages over the fluctuations as does the Window method. However, the IS method fails to account for the dynamics of the trajectory, giving a bimodal histogram. In this example the natural lifetime of the donor is 2 ns, the high EET rate is 0.99 ns^{-1} , the low EET rate is 0.01 ns^{-1} , and the rate changes every 10 ps.

Figure 3.

Comparison of the three sampling methods for a contrived dataset in which the system remains in a low EET rate state for a long period (480 ps) followed by a briefly occupied high EET rate state (20 ps), similar to what might occur in the case of free diffusion. In this case both the IS and Window methods yield bimodal FRET efficiency histograms while the Markov Chain method yields the appropriate singly-peaked histogram. As in Fig. 2, the natural lifetime of the donor is 2 ns, the high EET rate is 0.99 ns^{-1} , and the low EET rate is 0.01 ns^{-1} .

Figure 4.

Comparison of modeled data for the three sampling methods using the lysozyme trajectory.

Left: smFRET data. For clarity only every fourth point is shown with a symbol. The Markov data are exactly zero below 0.69 efficiency. Note that 100ps window data (not shown) are significantly noisier than the 10ps window; the shape continues the progression from IS to the 10ps window with the primary peak shifting to lower efficiency (roughly 0.85) and getting broader at the expense of some population at very low efficiency (lowest nonzero probability is at an efficiency of 0.25). Bin spacing is 0.02. **Right:** trFRET data. Six different window sizes are shown: 1, 5, 10, 20, 50, and 100 ps. They vary uniformly from top to bottom—the shortest window is closest to IS and the longest to Markov.

Figure 5.

Comparison of modeled data for the three sampling methods using the Oligo0–16' trajectory. **Left:** smFRET data. For clarity only every fourth point is shown with a symbol. Bin spacing is 0.02. **Right:** trFRET data. Six different window sizes are shown: 1, 5, 10, 20, 50, and 100 ps. They vary uniformly from top to bottom—the shortest window is closest to IS and the longest to Markov.

Figure 6. Comparison of modeled smFRET data for the three sampling methods using the Oligo0-0' trajectory. For clarity only every third point is shown with a symbol. Bin spacing is 0.005. Because the efficiencies are low, the donor trFRET decays differ only marginally from the donor-only decay and are not shown.

Figure 7.

Trajectory of relative EET efficiencies from the lysozyme simulation. Efficiencies are shown for the ISM (black), Window 10 ps (light gray), and Markov chain (red/medium gray) methods.

Figure 8.

Trajectory of the relative transition dipole-transition dipole EET rate (black) from the lysozyme simulation along with the number of times each frame of the trajectory was visited during the Markov sampling (red). Low visits show that a frame slightly earlier in the trajectory was a frequent source of a terminal EET event.

Figure 9.

Difference in EET rates from the LR and the IDA treatments of the coupling, shown as $(k_{EET,IDA,i} - k_{EET,LR,i})$. A 20 ns subset of the Oligo0–16' trajectory was used; note the range of the x-axis. The average value of the linear response rate over this range (5.85 ns^{-1}) is indicated with a large tick on the y-axis. Note that 90.8 % of the snapshots exhibit LR $k_{EET,i}$ values between 0 and 15 ns^{-1} . The average value of the rate difference is 9.44 ns^{-1} . Both the LR and IDA values were taken from QM calculations on each snapshot, such that fluctuations in the calculated transition magnitude affect both rates equally.

Figure 10.

Modeled trFRET and smFRET data using the Markov model and a subset of the Oligo0–16' trajectory (from 80–100 ns) as obtained within the QM-LR approach and the simpler transition dipole-transition dipole approximation.