THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Indocyanine Dyes Approach Free Rotation at the 3′ Terminus of A-RNA: A Comparison with the 5′ Terminus and Consequences for Fluorescence Resonance Energy Transfer

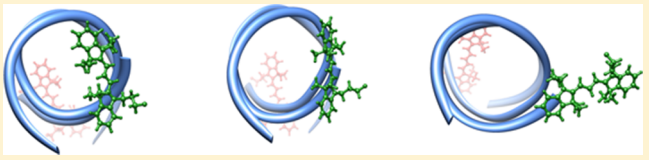
Peker Milas,[†,¶] Ben D. Gamari,[†,¶] Louis Parrot,[†] Brent P. Krueger,[‡] Sheema Rahmanseresht,[†] James Moore,[†] and Lori S. Goldner*[,†]

[†]Department of Physics, University of Massachusetts, Amherst, Amherst, Massachusetts

[‡]Department of Chemistry and Biochemistry, Hope College, Holland, Michigan

Supporting Information

ABSTRACT: Cyanine dyes are widely used to study the folding and structural transformations of nucleic acids using fluorescence resonance energy transfer (FRET). The extent to which FRET can be used to extract inter- and intramolecular distances has been the subject of considerable debate in the literature; the contribution of dye and linker dynamics to the observed FRET signal is particularly troublesome. We used molecular dynamics (MD) simulations to study the dynamics of the indocarbocyanine dyes Cy3 and Cy5 attached variously to the 3′ or 5′ terminal bases of a 16-base-pair RNA duplex. We then used Monte Carlo modeling of dye photophysics to predict the results of single-molecule-sensitive FRET measurements of these same molecules. Our results show that the average value of FRET depends on both the terminal base and the linker position. In particular, 3′ attached dyes typically explore a wide region of configuration space, and the relative orientation factor, κ^2, has a distribution that approaches that of free-rotators. This is in contrast to 5′ attached dyes, which spend a significant fraction of their time in one or more configurations that are effectively stacked on the ends of the RNA duplex. The presence of distinct dye configurations for 5′ attached dyes is consistent with observations, made by others, of multiple fluorescence lifetimes of Cy3 on nucleic acids. Although FRET is frequently used as a molecular "ruler" to measure intramolecular distances, the unambiguous measurement of distances typically relies on the assumption that the rotational degrees of freedom of the dyes can be averaged out and that the donor lifetime in the absence of the acceptor is a constant. We demonstrate that even for the relatively free 3′ attached dyes, the correlation time of κ^2 is still too long to justify the use of a free-rotation approximation. We further explore the consequences of multiple donor lifetimes on the predicted value of FRET.



INTRODUCTION

Fluorescence resonance energy transfer[1,2] (FRET) is widely used in exploring the global structure or structural transformations of nucleic acids.[3,4] Combined with single-molecule-sensitive techniques, FRET provides a method for direct observation of the conformational changes of DNA and RNA.[5−8] Although observation of gross changes in distance between the donor and acceptor dyes is straightforward, the quantitative interpretation of FRET data to extract structure is generally complicated by dye photophysics[9] and linker dynamics.[10] As a minimum, quantitative interpretation of FRET requires that the configuration and dynamics of the dyes be understood or modeled.

Indocarbocyanine dyes Cy3 and Cy5 are commonly used to label nucleic acids for FRET. The location and orientation of these dyes attached to DNA[11−13] and DNA/RNA hybrids[14] have been the subject of some discussion recently, as has the behavior of closely related sulfoindocarbocyanines.[15,16] Tethered to the 5′ end of a double-stranded A- or B-form helix, indocarbocyanines with structures and linkages as shown in Figure 1 are known to spend a significant fraction of their time "stacked" on the double-stranded (ds) nucleic acid.[11−13] Steric hindrance and other intramolecular interactions are known to prevent the dyes from exploring all orientational configurations.[14,17,18] Dye and dye-linker dynamics, as well as dye photophysics, can change significantly, depending on the details of the local chemical and physical environment, including, for example, the neighboring base and position of the dye linkage (3′ or 5′ termini here). Indeed, MD simulations have already demonstrated differences in the free energy landscape of Cy3 terminally attached to different bases.[13]

To explore the effects of linkage position and the terminal base on FRET for dye-labeled RNA, we used a combined molecular dynamics (MD)−Monte Carlo (MC) approach. Molecular dynamics simulations with explicit solvent were run for dye-labeled dsRNA to extract the trajectory of the dyes with picosecond resolution over hundreds of nanoseconds. These trajectories were then used as the basis of a Markov chain Monte Carlo simulation of FRET that models the fluorescence from the donor and acceptor dyes. Using the MC simulation,

Received: November 8, 2012
Revised: April 27, 2013
Published: June 25, 2013

ACS Publications © 2013 American Chemical Society

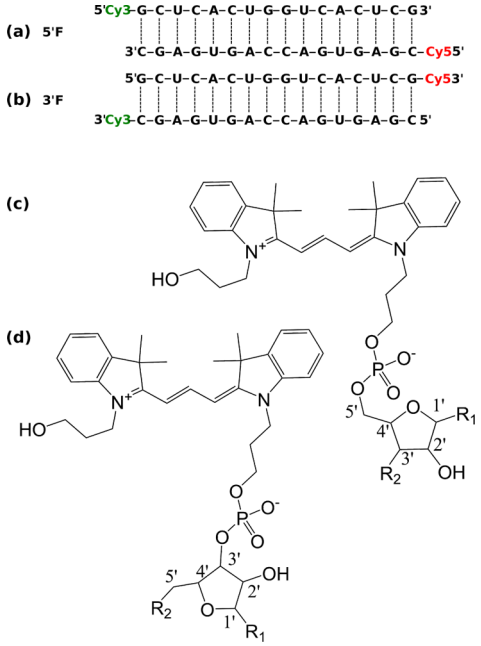


Figure 1. Duplexes, dyes, and linkers used in this study. (a) The 5′F duplex. 5′R is the same with dye positions swapped. (b) The 3′F duplex. 3′R is the same with dye positions swapped. (c) Dye attachment to the 5′ terminus. R1 is the terminal C or G, and R2 is the phosphate of the next nucleoside. (d) Dye attachment to the 3′ terminus. R1 is the terminal C or G and R2 is the phosphate of the next nucleoside.

we bridge the gap in time scales between molecular dynamics (MD) simulations (<1 μs) and typical measurements of FRET (time resolution > 100 μs). To explore the effect of dye dynamics on FRET, we first took the dyes to have single fluorescent lifetimes and, therefore, single quantum yields;[19,20] this assumption is commonly used in interpretation of FRET data. These results are compared with the more realistic case in which the donor dye is permitted to have multiple quantum yields (and corresponding fluorescent lifetimes) that depend on the configuration of the dye on the RNA. For Cy3 and Cy5 dyes terminally attached to RNA, as in Figure 1, we find that FRET depends on both the dye linkage position (i.e., 3′ or 5′ location) and the terminal base. For 5′ attachment, the dyes spend a significant fraction of their time in one or two states "stacked" on the double-stranded helix. For 3′ attached dyes the situation is quite different: in most cases, the dyes explore a wide range of orientations about their tether. In all cases, FRET predicted from the Monte Carlo model is shifted significantly from what might be expected from freely rotating dyes.

For either base-stacked or freely rotating dyes, the usual discussion of dye rotational freedom (or lack thereof) is obviated by the use of this MD/MC approach to modeling, since the dye positions are explicitly accounted for in the calculation of FRET. In the ideal dipole approximation, the efficiency of energy transfer from the donor to the acceptor dye is given by

$$E = \frac{1}{1 + \left(\frac{R}{R_F}\right)^6} \tag{1}$$

in which R is the distance between dyes, and the Förster radius, R_F, is given by[3,21]

$$R_F^{\,6} = \frac{9c^4 J \eta_D \kappa^2}{8\pi n^4} \tag{2}$$

In this expression, n is the solvent's refractive index, c is the speed of light, η_D is the quantum yield of the donor dye in the absence of the acceptor, and κ is an orientation factor defined below. The integral $J = \int f(\omega)\sigma(\omega)\omega^{-4}\, d\omega$ describes the spectral overlap of the donor emission and acceptor absorption: $f(\omega)$ is the fluorescence spectrum of the donor normalized such that its integral over all ω is 1; $\sigma(\omega)$ is the molecular cross section of the acceptor; and ω is the angular frequency (not the wavenumber). For convenience, we also define the following quantity,

$$R_0 = \sqrt[6]{\frac{2}{3\kappa^2}}\, R_F \tag{3}$$

which is the commonly used value of R_F evaluated with $\langle\kappa^2\rangle$ = 2/3 for freely rotating dyes.

To explore the effect of rotation alone, we initially held J and η_D constant, as is frequently assumed in FRET. To probe the effect of changes in the dyes' nonradiative processes that might accompany, for example, stacking on RNA, we added to this model values of η_D that changed with the configuration of the dye on the RNA in the MD simulation. The quantum yield η_D can be expressed as

$$\eta_D = \frac{k_{Dr}}{k_D} \tag{4}$$

where k_D, the decay rate of the donor excited state, is given by $k_D = (1/\tau_D) = k_{Dr} + k_{Dnr}$. Here τ_D is the excited state lifetime, k_{Dr} is the radiative decay rate of the excited state, and k_{Dnr} represents all other nonradiative decay processes. The orientation factor, κ, ranges from −2 to 2 and depends upon the relative orientations of the dye transition dipoles, whose unit vectors are denoted $\hat{\mu}_1$ and $\hat{\mu}_2$:

$$\kappa = (\hat{\mu}_1 \cdot \hat{\mu}_2) - 3(\hat{\mu}_1 \cdot \hat{R})(\hat{\mu}_2 \cdot \hat{R}) \tag{5}$$

Here, $\hat{R}$ is the unit vector along $\vec{R}$, the displacement from the donor dye to the acceptor dye. Most often, it is assumed that the dyes are freely rotating so that the orientational average, κ^2 = 2/3, can be used to estimate the rate of energy transfer. However, apparent orientational freedom is not by itself sufficient to justify the use of this average: correlations between R and κ^2 and long correlation times in either parameter (relative to τ_D, the lifetime of the donor) can modify FRET even when the relative orientation of the dyes has a distribution very similar to that of freely rotating dyes.[22,23] These correlations are naturally included in the MD/MC model presented here.

Förster transfer described by eq 1 results from an ideal dipole approximation that is known to work well for cyanine dyes separated by at least 2 nm if they are freely rotating and at least 5 nm if they are statically oriented.[24] In this study, all the dyes are moving to some extent, so the distance at which the ideal dipole approximation begins to lose validity lies somewhere between 2 and 5 nm. Given that dye separations are all roughly 5 nm in this work, the use of eq 1 is justified.

SIMULATION METHODS

MD Simulations of Dye-Labeled RNA. Molecular dynamics simulations were run to extract R and κ^2 trajectories needed for modeling FRET. Simulations were performed with

dx.doi.org/10.1021/jp311071y | J. Phys. Chem. B 2013, 117, 8649−8658

Amber 11 and AmberTools 1.4[25] using the FF99SB force field.[26] Each of the four different structures described in the caption of Figure 1 was modeled.

The 16-base-pair RNA duplex was prepared using the Nucleic Acid Builder (NAB) package from AmberTools 1.4. Indocarbocyanine dyes Cy3 and Cy5 and the linkers for the dyes are not among the default residues of the Amber package;[25] therefore, we created models of the trans isomer for both dyes and their carbon linkers in Protein Data Bank (PDB) format and calculated their minimum energy conformation with Firefly 7.1.G.[27] This was performed using a 6-31G(d) basis set and density functional theory with the B3LYP1 functional. Restrained electrostatic potential (RESP) point charges were calculated[28,29] for the Cornell et al.[30] force field using RED Tools vIII.3.[31] These were used to parametrize the point charges for each atom in the new residues.

The RNA, linkers, and dyes were combined into a single PDB structure. The dyes were initially oriented with the conjugated chain nearly parallel to the helical axis of the RNA, extending outward from the end of the RNA. Using XLEAP, 22 Na^+ ions, corresponding to roughly 159 mM, were added to the system to neutralize the phosphates. The system was solvated in a truncated octahedral box with a 2 nm buffer of TIP3P water. Nonbonded interactions were calculated using particle mesh Ewald molecular dynamics (PMEMD) for electrostatics and a 1 nm cutoff for van der Waals interactions. The Amber PMEMD.CUDA software was run on an nVidia GTX 480 graphical processing unit (GPU), which generated at most 10 ns per day.

Each duplex described in Figure 1 was equilibrated for 20 ns prior to beginning production MD; a comparison of short- and long-term fluctuations in energy was used as an equilibration criterion, as described by Van Beek et al.[22] Following equilibration, production runs of 300 ns were done under conditions of constant pressure and temperature using a Langevin thermostat. Snapshots were saved every 1 ps. Typical configurations of Cy3 on 5′ terminal G from the 5′F duplex simulation are shown in Figure 2. More configurations are presented in movies 1−3 of the Supporting Information.

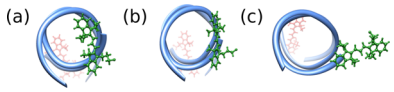


Figure 2. Typical snapshots of Cy3 attached to 5′ terminal G on the 5′F duplex:[32] (a) primary "base-stacked" configuration with r = 0.51 nm, θ = 49.27°, R = 4.94 nm, and κ^2 = 0.76; (b) unstacked configuration with r = 0.65 nm, θ = 47.02°, R = 4.71 nm, and κ^2 = 0.96; (c) wandering dye with r = 1.21 nm, θ = −168.03°, R = 5.49 nm, and κ^2 = 0.004. Here, r is the distance between the geometric centers of the terminal base pair and the conjugated chain of the dye, θ is the twist angle between the dye and the nearest base pair, and R and κ^2 are as defined in eqs 1 and 5.

Monte Carlo Modeling of FRET. To predict and model FRET, we used a Markov chain Monte Carlo model of a pair of interacting two-level dyes that utilized the MD trajectories for instantaneous values of the interdye distance R and the relative orientation represented by κ^2. The fluorescent and Förster processes accounted for in the Markov chain are shown in Figure 3. The rate of donor excitation used here, $k_{exc} = 10^5\ s^{-1}$, is within the range of experimentally accessible values, determined by the laser intensity and the absorption cross section of the donor. The rate of excited state decay of the

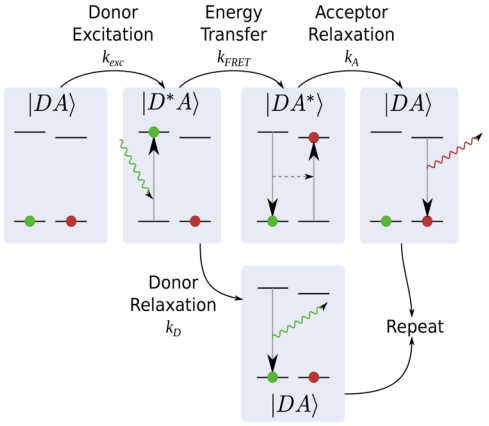


Figure 3. Markov chain diagram for the photophysical processes modeled here. Donor and acceptor relaxation can be radiative or nonradiative (not shown).

acceptor is given by $k_A = k_{Ar} + k_{Anr} = 1/\tau_A$ where k_{Ar} and k_{Anr} are the radiative and nonradiative decay rates of the acceptor, respectively. In the absence of the acceptor, k_D is known to be multivalued for Cy3 terminally attached to nucleic acids, an effect that is attributed to a nonradiative decay rate k_{Dnr}, or to a donor quantum-yield η_D, that depends on the dye's interaction with RNA.[33,34] Triplet, isomerization, and charge transfer processes occur on time scales significantly longer than the total MC simulation time and so could not be accounted for here; in many cases, it is possible to avoid the effect of these processes in experiments. We took J and the solvent refractive index, n, to have fixed values. The instantaneous rate of energy transfer, k_{FRET}, is given by[2]

$$k_{FRET} = k_D\left(\frac{R_F}{R}\right)^6 = \frac{E}{1-E}k_D \qquad (6)$$

where R_F is given in eq 2 and depends on κ^2 and η_D.

Simulations begin at $t = 0$ with the molecular configuration in the first frame of the MD trajectory and the dyes both in the ground state ($|DA\rangle$). The MC model then steps sequentially through the MD trajectory in increments of Δt = 1 ps. Subincrements of Δt = 100 fs were also used to check for numerical accuracy; no significant differences were found.

The wait time, τ, for transitions from ground into the donor-excited state ($|D^*A\rangle$) is exponentially distributed assuming an excitation rate k_{exc}. From $|D^*A\rangle$, the system has a probability, at each time step, of either undergoing energy transfer (with rate k_{FRET}), decaying radiatively (with rate k_{Dr}), or decaying nonradiatively (with rate k_{Dnr}). If the system undergoes FRET, it ends up in $|DA^*\rangle$, from which state it can undergo radiative or nonradiative decay of the acceptor back to the ground state. The output of the MC simulation was photon emission times recorded on the radiative $|D^*A\rangle \rightarrow |DA\rangle$ and $|DA^*\rangle \rightarrow |DA\rangle$ transitions.

To determine when the system leaves an excited state, a Bernoulli variate ("true" or "false," for example) is drawn at each time step with weight $p = \Delta t\ \Sigma_i k_i$, where the sum is carried out over potential destination states. If "true" is drawn, the system leaves the excited state, and a destination state, i, will be drawn where $p_i = k_i/\Sigma_j k_j$.

When the end of the MD trajectory is reached, the simulation returns to the beginning of the trajectory. This periodic extension of the MD simulation results in an unphysical sampling in those cases that the system is excited

dx.doi.org/10.1021/jp311071y | J. Phys. Chem. B 2013, 117, 8649−8658

upon reaching the end of the trajectory. However, with τ_D = 1 ns,[14,18,33,35] and 300 ns trajectories, this occurs no more than once for every 100 000 photons. Note that concurrent excitation of the dyes is unlikely in the weak excitation limit typical in most experiments, so this process is omitted from the model.

Throughout this work, the quantum yield of the acceptor is taken to be the same as the average quantum yield of the donor. This makes sense because the quantum yield of the acceptor does not modify Förster energy transfer, but it does appear in the experimentally relevant fluorescence detection-correction factor $\gamma = (\chi_A\eta_A)/(\chi_D\eta_D)$, where χ_A (χ_D) is the acceptor (donor) channel collection efficiency and η_A (η_D) is the acceptor (donor) quantum yield.[36] By equating the donor and acceptor quantum yields and omitting collection efficiency from the model, we effectively set γ = 1, which makes the MC result for $\langle E\rangle$ directly comparable with γ-corrected FRET data.

For modeling FRET with a single-valued donor fluorescent lifetime, we take R_0 =5.8 nm. This value is within the range of 5.6–6.5 nm that has been reported in the literature for this dye pair.[11,37−39] When multiple donor lifetimes are modeled, the value of R_0 changes as η_D changes during a trajectory. In this case, we take R_0 = 5.8 nm at the population-weighted average value of η_D.

EXPERIMENTAL MATERIALS AND METHODS

Labeled RNA was purchased from Integrated DNA Technologies (IDT). Samples were prepared at a final concentration of 100 pM in HEPES−NaOH at pH 7.8 with 100 mM NaCl and 5 mM $MgCl_2$. Also included were 15 nM protocatechuate-3,4-dioxygenase (PCD) and 5 mM protocatechuic acid (PCA) as an enzymatic oxygen-quenching system.[40] Single-molecule-sensitive FRET measurements were accomplished using a confocal microscope with 514 nm excitation from an argon−krypton laser at 50 μW. Donor and acceptor channel photons were detected using homemade photon timing and laser control circuitry based on an FPGA[41] and two avalanche photodiodes (PerkinElmer SPCM-AQRH-15). All FRET data are fully corrected for background, crosstalk, and the detection collection factor γ. Further details regarding sample preparation, data acquisition and data analysis are in the Supporting Information.

RESULTS AND DISCUSSION

MD Simulations. Trajectories from the MD simulation of the 5′F duplex are shown in Figure 4; trajectories for the 5′R, 3′F, and 3′R duplexes are found in the Supporting Information, Figures 1−3. The distance between the geometric centers of the dye molecules, R, is shown in Figure 4a. The orientation factor κ^2, shown in Figure 4b, is calculated by approximating the transition dipoles to be parallel to the conjugated bond structure between the two indole-like moieties. The instantaneous value of energy transfer efficiency, shown in Figure 4c, is calculated from R and κ^2 using eqs 1, 3, and 2. To the right of each trajectory in Figure 4 is a histogram of the plotted values.

Immediately evident from these trajectories is that there are changes in R and κ^2 at time scales both faster and slower than the average fluorescence lifetime of the donor, $\tau_D \approx 1$ ns.[14,18,33,35] In some cases, structures fluctuating about a particular R or κ^2 persist for tens of nanoseconds; much shorter fluctuations are also evident. The existence of correlations at many time scales makes it difficult to extract a correlation time

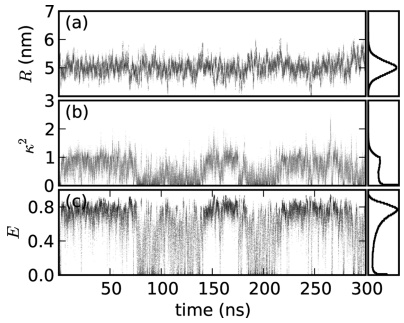


Figure 4. Trajectories and histograms for 5′F duplex. (a) The distance, R, between the geometric centers of the two dyes; a histogram of R values is on the right. (b) The orientation factor, κ^2; a histogram of κ^2 values is on the right. (c) Instantaneous values of E; a histogram of E values is on the right. Here $\langle R\rangle$ = 5.01 nm, $\langle\kappa^2\rangle$ = 0.57, and $\langle E\rangle$ = 0.580. The Pearson coefficient between κ^2 and R is −0.173. To avoid plotting 300 000 points, each figure is a two-dimensional histogram 500 bins wide and 300 tall using grayscale where white corresponds to 0 occurrences and black corresponds to ≥50 occurrences in (a), ≥100 occurrences in (b), and ≥30 occurrences in (c). In all cases, ≤0.1% of the nonzero pixels are saturated.

from a finite MD trajectory; autocorrelation decays for R and κ^2 can be found in the Supporting Information. Cross correlations between R and κ^2 are also evident in some cases. For the 5′F duplex whose trajectory is shown in Figure 4, R and κ^2 have a modest anticorrelation (Pearson's correlation coefficient) of −0.173, whereas these two parameters appear uncorrelated for the reverse 5′ configuration (Pearson coefficient of −0.008, Supporting Information Figure 1). For the 3′F and 3′R duplexes, the Pearson coefficients are 0.037 and −0.133, respectively (Supporting Information Figures 2, 3). It is worth noting that the correlation of R and κ^2 for the 5′F and 3′R duplexes appears to be due to the relatively static position of Cy5 on cytosine, whether 3′ or 5′ attached (discussed below, Figure 5a and 6a). In these cases, changes in R and κ^2 are due primarily to excursions of Cy3, which accounts for the correlation between them.

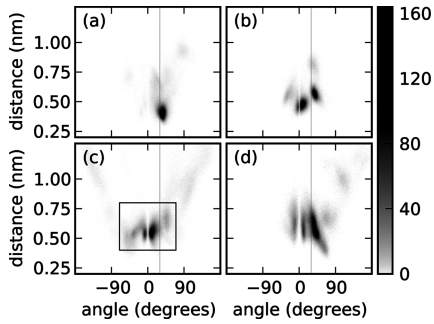


Figure 5. Two-dimensional histogram of distance r and angle θ for 5′ attached dyes. r and θ are defined in the text and describe the distance between and relative orientation of the dye and nearest base pair. (a) Cy5 on 5′ terminal C, (b) Cy5 on 5′ terminal G, (c) Cy3 on 5′ terminal C, and (d) Cy3 on 5′ terminal G. Panels a and d are taken from the 5′F duplex simulation; b and c are from the 5′R duplex simulation. The vertical line at 30° corresponds to the helical twist of A-RNA. The boxed region in panel c runs from −70° to 70° and 0.40 to 0.80 nm; it contains 89% of the population. Note that to bring out sparse populations, the color scale is proportional to the population raised to the power 0.6.

dx.doi.org/10.1021/jp311071y | J. Phys. Chem. B 2013, 117, 8649−8658

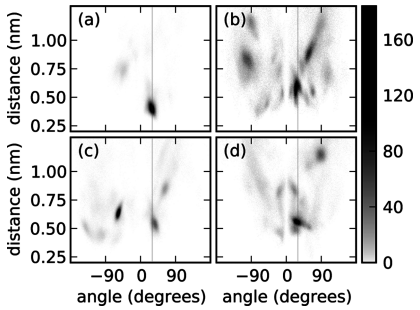


Figure 6. Two-dimensional histogram of distance r and angle θ for 3′ attached dyes. r and θ are defined in the text and describe the distance between and relative orientation of the dye and nearest base pair. (a) Cy5 on 3′ terminal C, (b) Cy5 on 3′ terminal G, (c) Cy3 on 3′ terminal C, and (d) Cy3 on 3′ terminal G. Panels a and d are taken from the 3′R duplex simulation; b and c are from the 3′F duplex simulation. The vertical line at 30° corresponds to the helical twist of A-RNA. Note that to bring out sparse populations, the color scale is proportional to the population raised to the power 0.6.

Although there are significant structural fluctuations that can be seen in the R and κ^2 trajectories of Figure 4, the 5′ attached dyes spend a significant fraction of their time effectively stacked on or near the end of the RNA at a distance roughly twice the axial rise between base pairs. A two-dimensional histogram of dye configurations on the 5′ terminal base is shown in Figure 5. The angle θ in Figure 5 and 6 is given by arccos $(\vec{Z}_d \cdot \vec{Z}_b)$, where $\vec{Z}_d$ is the direction along the conjugated carbon chain in the dye, pointing toward the free end of the dye. $\vec{Z}_b$ points along the adjacent base-pair axis, from the dye-attached nucleotide to the opposite nucleotide. It is defined by connecting the two C1′ atoms of the purine and complementary pyrimidine nucleotides. For dyes stacked on the end of the RNA, θ is therefore a twist angle. The distance r is defined to be between the geometric centers of the conjugated chain and the base pair. These definitions of the twist angle θ and distance r are similar to those in use elsewhere.[42]

The effect of stacking is particularly obvious for Cy5 attached to 5′ terminal C (Figure 5a), which spends the majority of its time with an average angular twist from the last base pair that is very close to the 30° helical twist of A-RNA (gray line in the figure). When attached to a 5′ terminal G (Figure 5b), the situation changes; there are now two locations that might be considered "stacked", one of which has θ very near zero degrees. Note that the peak near zero splits to avoid overlap of dye and base π-orbitals that occurs at 0°. Cy3 on 5′ terminal G (Figure 5d) also shows both peaks, although they show a larger distribution in both r and θ. For Cy3 on 5′ terminal C (Figure 5c), a population is shifted to the zero degree peak, with a narrower distribution in r. Cy5 has only very rare excursions to large angles, whereas Cy3 exhibits occasional excursions to much higher angles and a wider range of distances. The boxed region in Figure 5c represents 89% of the population. This is consistent with the work of Iqbal et al.,[12] who found that FRET on a DNA/RNA hybrid is commensurate with dyes that spend 12% of their time freely rotating and is otherwise found confined to a configuration that has lateral rotations within a 42° half-width half-maximum distribution. Our simulations suggest that Cy3 is responsible for most of the free rotation and that the stacked configuration consists of several distinct states.

Although the stacked configurations near $\theta = 30°$ are still in evidence, comparison of Figures 6 and 5 shows that cyanine dyes have substantially more configurational freedom when connected to the 3′ terminus. From Figure 6b, it is evident that Cy5 on 3′ terminal G explores a wide range of configurations in the course of the simulation. Cy3 shows a similarly large range of configurations on both 3′ terminal C (Figure 6c) and 3′ terminal G (Figure 6d). In all cases, the 3′ attachment appears to preclude the strong structural peaks on either side of $\theta = 0°$ that are favored for most of the 5′ attached dye structures. If so, then the loss of the peaks near $\theta = 0°$ may account for some of the additional rotational freedom of 3′ attached dyes. Of the 3′ configurations studied here, only Cy5 attached to a 3′ terminal C (Figure 6a), shows little evidence of configurational freedom.

The greater rotational freedom of these 3′ attached dyes is reflected in their κ^2 distributions, which are shown in Figure 7.

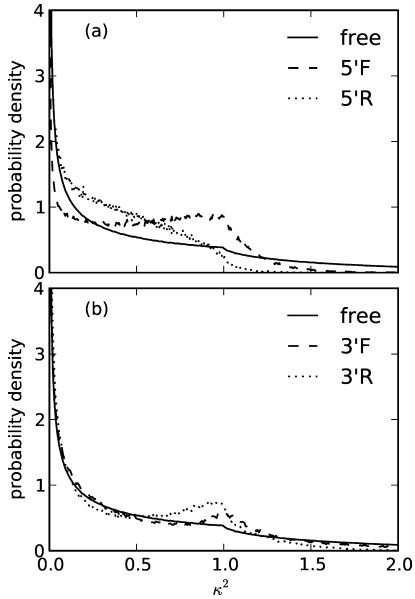


Figure 7. (a) Distribution of κ^2 for 5′F and 5′R duplexes for $0 < \kappa^2 < 2$. (b) Distribution of κ^2 for 3′F and 3′R duplexes. In both cases, the distribution of κ^2 for freely rotating dyes is shown for comparison.

With dyes attached to the 5′ end of the RNA, Figure 7a shows significant deviations from the freely rotating distribution for all values of κ^2. The κ^2 distributions for the 3′ attached dyes are much more similar to the freely rotating distribution, differing only in the peak region around $\kappa^2 = 1$ and in the tail at high κ^2. Despite the differences in the distributions, the average κ^2 values are similar for 3′ and 5′ attachment with $\langle\kappa^2\rangle = 0.57$ and 0.36 for the 5′F and 5′R duplexes respectively, and $\langle\kappa^2\rangle = 0.53$ and 0.43 for the 3′F and 3′R duplexes. All four systems yield $\langle\kappa^2\rangle$ values that are below the freely rotating value of 2/3. This is at least partly due to the fact that high values of κ^2 correspond to the case in which the conjugated chain of the dyes would need to be aligned along the helical axis: most of these configurations are sterically forbidden. This accounts for both the slightly low value of $\langle\kappa^2\rangle$ and the smaller populations (compared with the free-rotation distribution) at $\kappa^2 > 1.5$.

Monte Carlo Model and Consequences for FRET. Because the MD simulations discussed above show fluctuations in R and κ^2 occurring at time scales that are both faster and slower than τ_D (Figure 4 and Supporting Information Figures 1−3, 8, and 9) and because there are correlations between R and κ^2 in the 5′F and 3′R duplexes, there is no simple approximation of FRET that can be used here. Instead, we used

dx.doi.org/10.1021/jp311071y | J. Phys. Chem. B 2013, 117, 8649−8658

the Monte Carlo model discussed above to predict FRET histograms in the four systems of Figure 1.

Consequences of Rotation. To investigate the effects of rotational motion separate from changes in lifetime and quantum yield, we first made a simplifying approximation in which the nonradiative processes are assumed to be independent of dye configuration, so that k_D is single-valued and constant throughout the simulation. For this study, we used a radiative lifetime for Cy3 of τ_{Dr} = 2.0 ns as determined by Sanborn et al.[16] We further set η_D = 0.5, so that τ_D = 1 ns, which is an intermediate value for Cy3 on nucleic acids.[14,18,33,35] Below, we consider other values of η_D. To facilitate comparison with γ-corrected data, the quantum-yield of the acceptor was set equal to that of the donor, as discussed in the Methods section on Monte Carlo modeling of FRET. The lifetime of the acceptor is irrelevant to the outcome of the model, so for convenience we set τ_A = 1 ns.

As described in the Methods section, the output of the MC simulation was donor and acceptor photon arrival times. For purposes of constructing a FRET histogram, photons generated by the simulation were binned into arrival time intervals (bins) with a time width, Δt, chosen such that the mean number of photons per bin was $\langle N_A + N_D \rangle$ = 100. Here, N_A (N_D) refers to the number of acceptor (donor) photons in a particular bin, and $\langle N_A + N_D \rangle$ = 100 is typical for an experimental single molecule trajectory. From the binned list of photon counts, we computed a corresponding list of ratios, $E = (N_A/(N_A + N_D))$. This FRET trajectory was then used to construct the histogram shown in Figure 8b and 8d for 5′ and 3′ attached dyes, respectively. For comparison, histograms of the instantaneous value of E from the corresponding MD trajectory are shown in Figure 8a and c. The MC simulations were run until a total of ~10^6 photons were generated, giving a run-to-run standard deviation in the value of $\langle E \rangle$ of 0.0005.

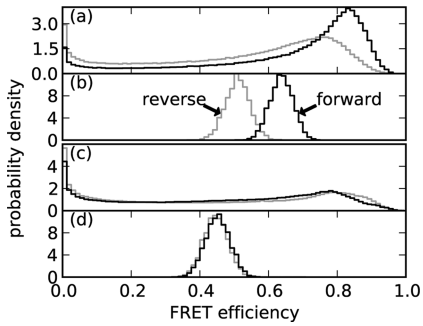


Figure 8. (a) Histogram of instantaneous E from the MD trajectory of the 5′F (black) and 5′R (gray) duplexes; $\langle E_{inst} \rangle$ = 0.580 and 0.447, respectively. (b) MC predicted distribution of E for the 5′F and 5′R duplexes; $\langle E \rangle$ = 0.638 and 0.509, respectively. (c) Histogram of instantaneous E from the MD trajectory of the 3′F and 3′R duplexes; $\langle E_{inst} \rangle$ = 0.402 and 0.392, respectively. (d) MC predicted distribution of E for the 3′F and 3′R duplexes; $\langle E \rangle$ = 0.452 and 0.445, respectively.

The donor quantum yield for Cy3 on DNA has been measured between 0.16 and 0.39.[14,16,34,43] Since quantum yields are difficult to determine with accuracy, we take a modeling approach and consider the change in $\langle E \rangle$ with values of of η_D that are both higher and lower than those found in the literature. The results for η_D = 0.5 are reported above in Figure 8. For η = 0.1, the values are shifted by only −3.4% to −4.8%; the results are given at the far right in Table 1. Note that for all entries in Table 1, τ_{Dr} = 2.0 ns and R_0 = 5.8 nm. Changes in η_D are assumed to come from changes in k_{Dnr}.

In the case that the fluctuations in R and κ^2 are separately either faster or slower than the donor lifetime, analytical expressions derived by Gopich and Szabo[44,45] can be used to estimate $\langle E \rangle$. Comparison of the MC predicted values for $\langle E \rangle$ with those of the analytic expressions gives insight into the effect of the observed correlations in R and κ on FRET; the various results are given in Table 1. These approximations require as input the distributions for R and/or κ^2; we obtain these distributions from our MD simulation. As should be expected, the limit of slow rotation and slow distance fluctuation ("slow−slow" in Table 1) reproduces the average of instantaneous FRET from the MD simulation alone (given in the captions of Figure 4 and Supporting Information Figures 1−3). It is also interesting to note that substitution of $R = \langle R \rangle$ and $\kappa^2 = \langle \kappa^2 \rangle$ into eqs 1 and 3 gives a result within 0.01 of the "fast−slow" approximation, a result that occurs in this approximation when the distance fluctuations are small on the scale of R_0. From Table 1, it is evident that the results for $\langle E \rangle$ from the MC model fall in a gap between approximations that assumes fast and slow orientational motion. Note that the assumption of slow or fast changes in distance is relatively unimportant, especially for the 5′F and 5′R systems, which should not be surprising, given the relatively small fluctuations in R for those molecules. Still, the assumption of fast fluctuations in R always gives higher FRET than the assumption of slow fluctuations, as has been discussed elsewhere.[46] It is also clear that the 3′R duplexes give results for $\langle E \rangle$ that are very close to the slow-rotation limit, despite their relative rotational freedom.

Although there is no simple analytic expression for estimating FRET when fluctuations occur on time scales both shorter and longer than τ_D, it is possible to approximate $\langle E \rangle$ using integrals over short sections of the MD trajectory,[44] as described in the supplement of ref 47. We compared the results of the MC model to this integration, sampling 50 000 starting

Table 1. Mean Values of R and κ^2 from the MD Simulations, along with $\langle E \rangle$ Predicted Using Both Analytic Approximations[44,45] and the MC Model Presented Here[a]

			FRET efficiency					
			orientation−distance				Monte Carlo	
duplex	$\langle R \rangle$ (nm)	$\langle \kappa^2 \rangle$	fast−fast	fast−slow	slow−fast	slow−slow	η = 0.5	η = 0.1
5′F	5.01	0.57	0.69	0.67	0.59	0.58	0.64	0.62
5′R	5.08	0.36	0.56	0.55	0.46	0.45	0.51	0.49
3′F	5.50	0.53	0.56	0.53	0.42	0.40	0.45	0.43
3′R	5.34	0.43	0.56	0.52	0.42	0.39	0.44	0.42

[a]For all approximations and simulations, γ = 1 and R_0 = 5.8 nm.

dx.doi.org/10.1021/jp311071y | J. Phys. Chem. B 2013, 117, 8649−8658

points and taking the integration out to 15 ns. We found no significant differences between the two methods: agreement in $\langle E \rangle$ was better than 0.001 in all cases, and the difference could be attributed to sampling error.

Consequences of Multiple Fluorescence Lifetimes. It is quite natural to include in these simulations changes in the quantum yield of the donor, η_D, that may be associated with changes in the configuration of the dye. Multiple fluorescent lifetimes are consistently found for Cy3 on DNA.[12,18] These include a fast component, near 0.3 ns, that is close to that of Cy3 in solution and that is identified with dyes that are free to undergo excited state isomerizations.[48] Excited state isomerization introduces a nonradiative decay pathway that increases k_{Dnr}, thereby lowering the fluorescence lifetime and quantum yield.[48] Longer lifetimes are associated with Cy3 that is more tightly confined, for example, by interactions with the DNA, which reduces the isomerization rate.

Of the many measurements of Cy3 lifetime on nucleic acids,[14,18,33,35] the most relevant here was for Cy3 attached to a 5′ terminal cytosine on an RNA/DNA duplex,[14] which were consistent with three component fluorescence decays with τ_{D1} = 1.77 ns, τ_{D2} = 0.96 ns, and τ_{D3} = 0.31 ns.[14] Most of the steady-state fluorescence signal was associated with τ_{D1} and τ_{D2}, and only 5% of the intensity was attributed to τ_{D3}.[14] In the absence of lifetime measurements for Cy3 on the RNA modeled here, we use these values for τ_{D1}, τ_{D2}, and τ_{D3} to explore the consequences of multiple lifetimes on FRET in the 5′R duplex, which also contains a Cy3 attached to a terminal cytosine.

Although many different degrees of freedom might be used to assign configurational states associated with these three lifetimes, we started here with the simplest parameter that we could identify, namely, the distance between the dye and the terminal base pair as defined previously for Figure 5 and 6. For the 5′R duplex, the histograms of dye-to-base distance, shown in Figure 9, are reasonably fit by a Gaussian mixture model[49]

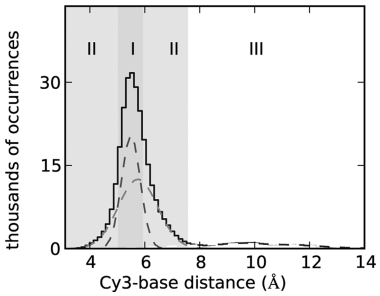


Figure 9. Histogram of the distance between Cy3 and the 5′ terminal base pair for the 5′R duplex. The parameters resulting from this fit are the following: stacked state I (dashed dark gray curve) mean $\langle r \rangle$ = 5.49 nm, standard deviation $\sigma = (\langle (r - \langle r \rangle)^2 \rangle)^{1/2}$ = 0.12 nm, weight w = 0.39; stacked state II (dashed light gray curve) $\langle r \rangle$ = 5.75 nm, σ = 0.51 nm, w = 0.50; unstacked state (dashed black curve) $\langle r \rangle$ = 9.79 nm, σ = 4.87 nm, w = 0.11.

with three Gaussians, which we associate with three different configurational states of the dye, labeled I, II, and III. A maximum likelihood inference was used to assign the states shown in Figure 9. Since the fluorescence lifetime of cyanine dyes increases with the decrease in isomerization that occurs in more rigid or confining environments,[16,50,51] we associate the narrowest distribution (I), which is closest to base-stacked, with the longest lifetime. The broadest distribution (III), which also has the largest average distance between the dye and base pair, is associated with the shortest lifetime. Consistent with the results of Iqbal et al.,[14] the latter population is also by far the smallest. Complete Gaussian mixture fit results are given in the caption of Figure 9.

To include multiple lifetimes in the MC model, each step in the MD simulation was assigned one of the states I, II, or III as described above. For state I, $k_{D1} = 1/\tau_{D1}$; for state II, $k_{D2} = 1/\tau_{D2}$; and for state III, $k_{D3} = 1/\tau_{D3}$. Again assuming that the underlying radiative rate is given by τ_{Dr} = 2.0 ns,[16] the quantum yields were calculated using eq 4 to be η_{D1} = 0.89, η_{D2} = 0.48, and η_{D3} = 0.16. Using these quantum yields, R_0 was also assigned appropriately for each state, subject to the constraint that R_0 = 5.8 nm at the mean quantum yield of $\langle \eta_D \rangle$ = 0.68. That this value of $\langle \eta_D \rangle$ is higher than might be expected[14,16,34,43] underscores the need for better determinations of quantum yield and radiative lifetime in future work, but it does not prevent us from exploring the general consequences of multiple lifetimes on FRET histograms.

With a value of $\langle \eta_D \rangle$ = 0.68, and using multiple lifetimes as discussed above, we find $\langle E \rangle$ = 0.5095 from the MC model of the 5′R duplex. If we instead use the same value of η_D but assume only a single fluorescent lifetime given by $\tau_D = \eta_D/k_{Dr}$ = 1.36 ns, we find $\langle E \rangle$ = 0.5139. Unlike the effect of including rotational dynamics, which give a substantial shift in $\langle E \rangle$, the inclusion of three separate lifetimes makes only a small change (−0.0044) in the calculated value of FRET. In addition, because the changes in dye configuration seen in the MD simulation are faster than the interphoton time, we would not expect to see any broadening due to the inclusion of multiple lifetimes. In all cases, the resulting distributions are homogeneous with widths that are shot-noise-limited.[52]

Comparison with Data. Single-molecule-sensitive solution FRET data for the 5′R duplex, corrected for background, cross-talk, and γ are shown in solid gray in Figure 10. Data are

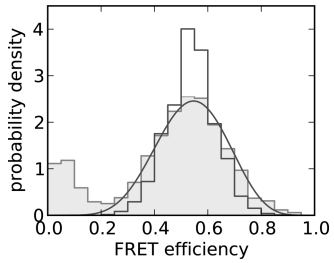


Figure 10. A comparison of data and model. The model used here includes the effect of multiple lifetimes associated with different configurational states of the dye, although the result is nearly indistinguishable from that of a single-lifetime model at the same $\langle \eta_D \rangle$ = 0.68. FRET histogram from data is shown filled with gray. The smooth solid black line is a best-fit to the data of a beta function, giving $\langle E \rangle$ = 0.540 ± 0.003. FRET histogram from the model, with R_0 = 5.9 nm at $\langle \eta_D \rangle$ = 0.68, is also shown in a solid black line. Data have been corrected for cross-talk, background, and gamma as described in the text.

acquired on molecules freely diffusing in solution at a concentration of 100 pM. The details of data analysis are discussed in the Supporting Information. To construct the FRET histogram, photon arrival times in each of two channels (donor and acceptor) were histogrammed into 5 ms bins, with bursts of photons indicating the presence of a labeled RNA molecule in the detection volume. Bins with more than 20

dx.doi.org/10.1021/jp311071y | J. Phys. Chem. B 2013, 117, 8649−8658

photons (sum of both channels) were used in calculations of FRET. This threshold resulted in an average of 32 photons per bin. The FRET efficiency distribution was fit to a mixture of two beta distributions. The component corresponding to the FRET peak is shown in Figure 10 as a continuous solid black line. The lower peak corresponds to donor-only molecules, and it is not of interest here. From the FRET data, we find $\langle E \rangle$ = 0.540 ± 0.003 from 1713 above-threshold bins, higher than the prediction using η_D = 0.68 and R_0 = 5.8 nm, above. Data taken on the 5′F duplex was very similar, with only a small positive shift in FRET: $\langle E \rangle$ = 0.576 ± 0.008 from 335 above-threshold bins. The uncertainty in both cases is the standard error of the mean of the distribution; it does not include the larger contribution arising from the uncertainty in γ.

From this work, it is clear that measurements of the τ_D's and τ_{Dr} or $\langle \eta_D \rangle$, specific to each molecule, are critical to the success of these models. For example, for the 5′R duplex, bringing the model into agreement with the data requires modifying η_D or R_0 or both. Since η_D is already considerably higher than might be expected and would need to be further raised to bring the model into better agreement with the data, we ran the MC simulation at various values of R_0 and found that for R_0 = 5.9 nm, $\langle E \rangle$ = 0.536 (with 27 000 bins in the simulation, the statistical uncertainty is negligible), which brings the model into agreement with the data.

To model the FRET histogram of Figure 10, we binned the simulated data with an average of 32 photons per bin. The resulting shot-noise-limited histogram is shown as a dark gray outline in the figure. The width of the data histogram, with a standard deviation of 0.128, is larger than what would be expected from shot noise,[44] for which the standard deviation is 0.099. Broadening in Cy3-Cy5 FRET histograms has been previously reported and attributed to long-lived states of the acceptor dye that are not accounted for in this model.[53]

DISCUSSION AND CONCLUSIONS

With notable exceptions,[11,12,18] almost all analysis of FRET from Cy dyes on nucleic acids assumes that the dyes are freely rotating on a time scale that is fast compared to τ_D. Indeed, it is now common to extract distances or distance distributions from FRET data in this limit. Here, we have shown that steric hindrance and "sticking" of cyanine dyes on RNA makes the slow-rotation approximation a somewhat better choice, particularly for 3′ attached dyes. However, in general, the "sticky" rotational behavior of the dyes is best dealt with using a model that includes the dye dynamics to generate model trajectories and then applying either a Monte Carlo or integral technique[44,47] to predict or model FRET.

The failure of the fast, free-rotation approximation for 5′ attached dyes is perhaps not surprising, given that the κ^2 distribution from the MD simulation is quite different from what would be expected for free rotation. However, for 3′ attached dyes, we find considerable configurational freedom in three of four cases studied (Figure 6). This freedom results in a distribution of κ^2 that approximates that of freely rotating dipoles for the two combinations of 3′ attached dye-pairs studied here. This is not sufficient to justify the use of the free-rotation approximation because the correlation times for this free rotation are on the order or longer than τ_D (Supporting Information Figure S9). The MC approach used here and the integral approach described by Best et al.[47] naturally include fluctuations at all time scales modeled in the MD trajectory.

The distributions of κ^2 and the dynamics of the dyes were also found to depend on the terminal base[13] in a way that is consistent with the work of Iqbal et al.,[12] who showed dramatic oscillation in FRET as a function of duplex length due to base stacking. The nucleic acids used in that work had Cy3 and Cy5 attached to a 5′-terminal cytosine. Of those modeled here, this is the configuration for which we would expect the greatest effect of stacking and the largest correlation between R and κ^2 because only one structure dominates for cyanine dyes on 5′-terminal C (Figure 5a, c). We have not modeled cyanines on A or U, but note that Spiriti et al.[13] predict that stacking interactions are weakest for Cy3 attached to T on B-DNA. They further point out that this might be attributed to the methyl group on thymine, which is absent in RNA because T is replaced by U. In general, the dynamics of the dye and the notion that FRET distributions are sensitive to the terminal base are consistent with experimental studies that have demonstrated a change in dye photophysics with terminal base or DNA sequence.[33,34]

The shift in $\langle E \rangle$ upon the inclusion of multiple lifetimes (and therefore, quantum yields) associated with the different dye configurations is shown to be small so long as the population-weighted average quantum yield is used in place of three separate quantum yields. The small negative shift observed in the value of $\langle E \rangle$ is well out of the sampling error of the model but approximately equal to the standard error on the data. However, the situation should be expected to change if there are longer correlations in the dye configuration, either donor or acceptor, than can be observed with MD simulations. If the relevant correlation time approaches the interphoton time (typically between 1 μs and 1 ms), then it should be expected that, in addition to a shift in $\langle E \rangle$, the FRET histogram will be heterogeneous (e.g., broadened beyond shot noise). Indeed, longer-lived states of Cy5 have been discussed by Kalinin et al.[53] and might account for the excess width of the experimental FRET histogram.

A comparison of data and model for the 5′R duplex yielded an estimate of R_0 = 5.9 nm at $\langle \eta_D \rangle$ = 0.68. Even if the experimental FRET histogram was shot-noise-limited, this estimate needs to be viewed with some caution. Given the sensitivity of cyanine dyes to both linker position and adjacent base, it is clear that accurate estimates of R_0 require measurements of fluorescence lifetime(s) and average quantum yield for the donor dye in the specific configuration of each modeled molecule. Such measurements are complicated by the very short lifetime of fully solvated Cy3, but are planned for the near future.

Finally, it is worth noting that there are parameters other than lifetime, quantum yield, R, and κ^2 that might change with the dye configuration, including the overlap integral J or the local refractive index. That the dye spectra shift with environment is well known, but these shifts are often small, and barring ground-state isomerization, the resulting change in the overlap integral probably does not approach the large changes in quantum yield that must result from substantial changes in τ_D with dye configuration.

ASSOCIATED CONTENT

Supporting Information

Movies and time series plots showing trajectories of all four RNA systems, predicted FRET histograms for 3′ systems, autocorrelation functions of interdye distance and orientation factor for all four systems, description of experimental methods

and setup, analysis of experimental data. This information is available free of charge via the Internet at http://pubs.acs.org

AUTHOR INFORMATION

Corresponding Author

*E-mail: lgoldner@physics.umass.edu.

Author Contributions

¶P.M. and B.D.G. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Molecular graphics images from MD trajectories were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081). The authors thank Irina Gopich, and P.M. would like to thank Ricardo B. Metz, Ross Walker, and Scott M. Auerbach for helpful discussions. This work was funded by NSF MCB-0920139. B.K. was supported by NSF-MRI CHE 1039925 and NSF-RUI CHE-1058981.

REFERENCES

(1) Förster, T. Energiewanderung Und Fluoreszenz. *Naturwissenschaften* **1946**, *6*, 166−175.

(2) Förster, T.; Sinanoglu, O. *Modern Quantum Chemistry*; Academic Press: New York, 1965; Chapter Delocalized excitation and excitation transfer, pp 93−160.

(3) Clegg, R. M. Fluorescence Resonance Energy-Transfer and Nucleic-Acids. *Methods Enzymol.* **1992**, *211*, 353−388.

(4) Lilley, D. M. J.; Wilson, T. J. Fluorescence Resonance Energy Transfer as a Structural Tool for Nucleic Acids. *Curr. Opin. Chem. Biol.* **2000**, *4*, 507−517.

(5) Wennmalm, S.; Edman, L.; Rigler, R. Conformational Fluctuations in Single DNA Molecules. *Proc. Natl. Acad. Sci.* **1997**, *94*, 10641−10646.

(6) Ha, T.; Zhuang, X. W.; Kim, H. D.; Orr, J. W.; Williamson, J. R.; Chu, S. Ligand-Induced Conformational Changes Observed in Single RNA Molecules. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9077−9082.

(7) Grunwell, J. R.; Glass, J. L.; Lacoste, T. D.; Deniz, A. A.; Chemla, D. S.; Schultz, P. G. Monitoring the Conformational Fluctuations of DNA Hairpins Using Single-Pair Fluorescence Resonance Energy Transfer. *J. Am. Chem. Soc.* **2001**, *123*, 4295−4303.

(8) Zhuang, X. W.; Kim, H.; Pereira, M. J. B.; Babcock, H. P.; Walter, N. G.; Chu, S. Correlating Structural Dynamics and Function in Single Ribozyme Molecules. *Science* **2002**, *296*, 1473−1476.

(9) Chung, H. S.; Louis, J. M.; Eaton, W. A. Distinguishing between Protein Dynamics and Dye Photophysics in Single-Molecule FRET Experiments. *Biophys. J.* **2010**, *98*, 696−706.

(10) Sindbert, S.; Kalinin, S.; Nguyen, H.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Mueller, S.; Seidel, C. A. M. Accurate Distance Determination of Nucleic Acids via Förster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity. *J. Am. Chem. Soc.* **2011**, *133*, 2463−2480.

(11) Norman, D. G.; Grainger, R. J.; Uhrin, D.; Lilley, D. M. J. Location of Cyanine-3 On Double-Stranded DNA: Importance for Fluorescence Resonance Energy Transfer Studies. *Biochemistry* **2000**, *39*, 6317−6324.

(12) Iqbal, A.; Wang, L.; Thompson, K. C.; Lilley, D. M. J.; Norman, D. G. The Structure of Cyanine 5 Terminally Attached to Double-Stranded DNA: Implications for FRET Studies. *Biochemistry* **2008**, *47*, 7857−7862.

(13) Spiriti, J.; Binder, J. K.; Levitus, M.; van der Vaart, A. Cy3-DNA Stacking Interactions Strongly Depend on the Identity of the Terminal Basepair. *Biophys. J.* **2011**, *100*, 1049−1057.

(14) Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J. Orientation Dependence in Fluorescent Energy Transfer Between Cy3 and Cy5 Terminally Attached to Double-Stranded Nucleic Acids. *Proc. Natl. Acad. Sci.* **2008**, *105*, 11176−11181.

(15) Urnavicius, L.; McPhee, S. A.; Lilley, D. M. J.; Norman, D. G. The Structure of Sulfoindocarbocyanine 3 Terminally Attached to dsDNA Via a Long, Flexible Tether. *Biophys. J.* **2012**, *102*, 561−568.

(16) Sanborn, M. E.; Connolly, B. K.; Gurunathan, K.; Levitus, M. Fluorescence Properties and Photophysics of the Sulfoindocyanine Cy3 Linked Covalently to DNA. *J. Phys. Chem. B* **2007**, *111*, 11064−11074.

(17) Lilley, D. M. J. *Methods in Enzymology*, 1st ed.; Elsevier B.V.: Amerdam, 2009; Vol. *469*; pp 159−187.

(18) Ouellet, J.; Schorr, S.; Iqbal, A.; Wilson, T. J.; Lilley, D. M. J. Orientation of Cyanine Fluorophores Terminally Attached to DNA Via Long, Flexible Tethers. *Biophys. J.* **2011**, *101*, 1148−1154.

(19) Speelman, A. L.; Muñoz Losa, A.; Hinkle, K. L.; VanBeek, D. B.; Mennucci, B.; Krueger, B. P. Using Molecular Dynamics and Quantum Mechanics Calculations to Model Fluorescence Observables. *J. Phys. Chem. A* **2011**, *115*, 3997−4008.

(20) Hoefling, M.; Lima, N.; Haenni, D.; Seidel, C. A. M.; Schuler, B.; Grubmueller, H. Structural Heterogeneity and Quantitative FRET Efficiency Distributions of Polyprolines Through a Hybrid Atomistic Simulation and Monte Carlo Approach. *PLOS One* **2011**, *6*, e19791.

(21) Novotny, L.; Hecht, B. *Principles of Nano-Optics*; Cambridge University Press: Cambridge, 2006.

(22) VanBeek, D. B.; Zwier, M. C.; Shorb, J. M.; Krueger, B. P. Fretting About FRET: Correlation Between Kappa and R. *Biophys. J.* **2007**, *92*, 4168−4178.

(23) Gopich, I. V.; Szabo, A. Single-Molecule FRET with Diffusion and Conformational Dynamics. *J. Phys. Chem. B* **2007**, *111*, 12925−12932.

(24) Munoz-Losa, C., A.; Curutchet; Krueger, B. P.; Hartsell, L. R.; Mennucci, B. Fretting About FRET: Failure of the Ideal Dipole Approximation. *Biophys. J.* **2009**, *96*, 4779−4788.

(25) Case, D.; et al. *AMBER 11*; University of California, San Francisco: San Francisco, 2010.

(26) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *PROTEINS: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.

(27) Granovsky, A. A. Firefly version 7.1.G. http://classic.chem.msu.su/gran/firefly/index.html.

(28) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(29) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. Application of the Multimolecule and Multiconformational RESP Methodology to Biopolymers: Charge Derivation for DNA, RNA, and Proteins. *J. Comput. Chem.* **1995**, *16*, 1357−1377.

(30) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. a. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(31) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. The R.E.D. Tools: Advances in RESP and ESP Charge Derivation and Force Field Library Building. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821−39.

(32) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera-A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605−12.

(33) Harvey, B. J.; Perez, C.; Levitus, M. DNA Sequence-Dependent Enhancement of Cy3 Fluorescence. *Photochem. Photobiol. Sci.* **2009**, *8*, 1105−10.

(34) Harvey, B. J.; Levitus, M. Nucleobase-Specific Enhancement of Cy3 Fluorescence. *J. Fluoresc.* **2009**, *19*, 443−448.
(35) Dietrich, A.; Buschmann, V.; Muller, C.; Sauer, M. Fluorescence Resonance Energy Transfer FRET and Competing Processes in Donor-Acceptor Substituted DNA Strands: A Comparative Study of Ensemble and Single-Molecule Data. *Rev. Mol. Biotechnol.* **2002**, *82*, 211−231.
(36) Ha, T. J.; Ting, A. Y.; Liang, J.; Caldwell, W. B.; Deniz, A. A.; Chemla, D. S.; Schultz, P. G.; Weiss, S. Single-Molecule Fluorescence Spectroscopy of Enzyme Conformational Dynamics and Cleavage Mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 893−898.
(37) Yasuda, R.; Masaike, T.; Adachi, K.; Noji, H.; Itoh, H.; Kinosita, K. The ATP-Waiting Conformation of Rotating F-1-ATPase Revealed by Single-Pair Fluorescence Resonance Energy Transfer. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9314−9318.
(38) Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T. J. Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy. *Biophys. J.* **2004**, *86*, 2530−2537.
(39) Rasnik, I.; Myong, S.; Cheng, W.; Lohman, T. M.; Ha, T. DNA-Binding Orientation and Domain Conformation of the E-Coli Rep Helicase Monomer Bound to a Partial Duplex Junction: Single-Molecule Studies of Fluorescently Labeled Enzymes. *J. Mol. Biol.* **2004**, *336*, 395−408.
(40) Aitken, C. E.; Marshall, R. A.; Puglisi, J. D. An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments. *Biophys. J.* **2008**, *94*, 1826−1835.
(41) Gamari, B. D.; Zhang, D.; Buckman, R. E.; Milas, P.; Denker, J. S.; Chen, H.; Hongmin, L.; Goldner, L. S. Inexpensive Electronics and Software for Photon Statistics and Corerlation Spectroscopy. *Am. J. Phys.* **2013**, in press.
(42) Olson, W.; Bansal, M.; Burley, S.; Dickerson, R.; Gerstein, M.; Harvey, S.; Heinemann, U.; Lu, X.; Neidle, S.; Shakked, Z.; Sklenar, H.; Suzuki, M.; Tung, C.; Westhof, E.; Wolberger, C.; et al. A Standard Reference Frame for the Description of Nucleic Acid Base-Pair Geometry. *J. Mol. Biol.* **2001**, *313*, 229−237.
(43) Malicka, J.; Gryczynski, I.; Fang, J.; Kusba, J.; Lakowicz, J. Photostability of Cy3 and Cy5-Labeled DNA in the Presence of Metallic Silver Particles. *J. Fluoresc.* **2002**, *12*, 439−447.
(44) Gopich, I. V.; Szabo, A. Theory of Single-Molecule FRET Efficiency Histograms. *Adv. Chem. Phys.* **2012**, *146*, 245−297.
(45) Gopich, I. *Private communication*, 2012.
(46) Schuler, B.; Lipman, E.; Steinbach, P.; Kumke, M.; Eaton, W. Polyproline and the "Spectroscopic Ruler" Revisited with Single-Molecule Fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754−2759.
(47) Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A. Effect of Flexibility and Cis Residues in Single-Molecule FRET Studies of Polyproline. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18964−18969.
(48) Aramendia, P.; Negri, R.; Sanroman, E. Temperature Dependence of Fluorescence and Photoisomerization in Symmetrical Carbocyanines − Influence of Medium Viscosity and Molecular Structure. *J. Phys. Chem.* **1994**, *98*, 3165−3173.
(49) Pedregosa, F.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(50) Cooper, M.; Ebner, A.; Briggs, M.; Burrows, M.; Gardner, N.; Richardson, R.; West, R. Cy3B: Improving the Performance of Cyanine Dyes. *J. Fluoresc.* **2004**, *14*, 145−50.
(51) Buettner, A. V. Radiationless Transitions in Cyanine Dyes. *J. Chem. Phys.* **1967**, *46*, 1398.
(52) Gopich, I.; Szabo, A. Theory of Photon Statistics in Single-Molecule Förster Resonance Energy Transfer. *J. Chem. Phys.* **2005**, *122*, 014707−014725.
(53) Kalinin, S.; Sisamakis, E.; Magennis, S. W.; Felekyan, S.; Seidel, C. A. M. On the Origin of Broadening of Single-Molecule FRET Efficiency Distributions Beyond Shot Noise Limits. *J. Phys. Chem. B* **2010**, *114*, 6197−6206.