Monte-Carlo Diffusion-Enhanced Photon Inference: Distance Distributions and Conformational Dynamics in single-molecule FRET

Antonino Ingargiola,[1,a] Shimon Weiss,[1,b] and Eitan Lerner[1,2,c]
[1] *Department of Chemistry and Biochemistry, University of California Los Angeles, USA*
[2] *Department of Biological Chemistry, The Hebrew University, Jerusalem , Israel*

(Dated: August 4, 2018)

Abstract

Single-molecule Förster Resonance Energy Transfer (smFRET) is utilized to study the structure and dynamics of many bio-molecules, such as proteins, DNA and their various complexes. The structural assessment is based on the well-known Förster relationship between the measured efficiency of energy transfer between a donor (D) and an acceptor (A) dye and the distance between them. Classical smFRET analysis methods called photon distribution analysis (PDA) take into account photon shot-noise, D-A distance distribution and, more recently, interconversion between states in order to extract accurate distance information. It is known that rapid D-A distance fluctuations on the order of the D lifetime (or shorter) can increase the measured mean FRET efficiency and thus decrease the estimated D-A distance. Nonetheless, this effect has been so far neglected in smFRET experiments, potentially leading biases in estimated distances.

Here we introduce a PDA approach dubbed MC-diffusion-enhanced photon inference (MC-DEPI). MC-DEPI recolor smFRET experiments taking into account dynamics of D-A distance fluctuations, multiple interconverting states and photo-blinking. Using this approach, we show how different underlying conditions may yield an identical FRET histograms and how the additional information from fluorescence decays helps distinguishing between the different conditions. Then, we introduce a machine learning fitting approach for retrieving the D-A distance distribution, decoupled from the above-mentioned effects. We show that distance interpretation of smFRET experiments of even the simplest dsDNA is nontrivial and requires a decoupling the effects of rapid D-A distance fluctuations on FRET in order to avoid systematic biases in the estimation of the D-A distance distribution.

Keywords: Single-molecule FRET, Diffusion-enhanced FRET, PDA, fluorescence decays, conformational dynamics, distance distributions, Monte-Carlo, Bayessian optimization, machine learning

I. INTRODUCTION

Förster resonance energy transfer (FRET) is a phenomenon in which an electronically excited fluorophore transfers a fraction of its excitation energy non-radiatively to another fluorophore in the ground-state. The energy transfer occurs as long as the fluorescence spectrum of the first fluorophore, the donor (D), overlaps with the excitation spectrum of the second fluorophore, the acceptor (A), and as long as they are in close proximity. The efficiency of FRET depends on the sixth power of the distance between the donor and acceptor fluorophores[1]. This distance dependence gained FRET the term "molecular ruler" and indeed it has been used as such to resolve conformational information for inter-fluorophore distances in the range 2-10 nm on biological molecules since 1967[2]. Yet, such measurements at the ensemble level allowed retrieval of ensemble-averaged distances that may have been interpreted in many different ways.

The capability to measure FRET at the single molecule level allowed classification of molecules of the ensemble into classes with different FRET efficiencies, as well as identifying the dynamics of interconversion between them. Indeed, since single-molecule FRET (smFRET) was introduced in 1996[3], it has been used by many different labs to report conformational sub-populations and dynamics of a myriad of biological macromolecules, such as RNA, DNA, proteins and their complexes[4]. Each of these sub-populations of FRET efficiencies represents a time-average over a few milliseconds of either a single conformational state or of multiple different conformational states with transitions between them that occur faster than this timescale. Therefore each FRET sub-population entails the distance information on a conformational state. In the last few years smFRET has gradually become a tool used for retrieval of intramolecular distance information[4]. Performing such measurements on a biological molecule doubly labeled at many different pairs of atomic positions has opened the door for determination of the ensemble of structures that define different conformational states existing in the solution at ambient temperature[5–9]. However, there can still be many different distance-related interpretations of a FRET sub-population, hence the retrieval of distance information from smFRET sub-populations is far from being straightforward.

[a] Electronic mail: ingargiola.antonino@gmail.com
[b] Electronic mail: shiweiss@g.ucla.edu
[c] Electronic mail: eitanlerner1@g.ucla.edu; eitan.lerner@mail.huji.ac.il

For the task of standardizing the procedure of distance information retrieval from smFRET measurements, FRET between dyes attached to a simple rigid molecule such as dsDNA has been used[10,11]. However, even in such a simple molecule, the distance interpretation of an smFRET measurement may be complicated by: (1) the distribution of the inter-dye distance[12,13], (2) how fast does this distance change[14,15] and (3) how do photophysical processes that compete with FRET influence the outcome of the smFRET measurements[16–18].

A single conformational state can be described by a potential well depending on the D-A distance as a reaction coordinate[19] and on a D-A self-diffusion coefficient, which describes the distance dynamics in the potential well[19,20]. Therefore, a conformational state can be defined by the distance distribution at equilibrium $p_{Eq.}(r)$, and the inter-dye self-diffusion coefficient D. It is also known that distance changes occurring in the timescale of the D fluorescence decay or faster, lead to enhancement of the FRET efficiency relative the static case[15,19–22]. Additionally, different conformational states interconverting on 10 µs time scales (or faster), yield a single millisecond-averaged FRET peak. Finally, there are three well-known sources of bias in smFRET measurements: (1) the γ-factor, accounting for the imbalance between the detected D and A signal due to differences fluorescence quantum yields and detection efficiencies, (2) the donor spectral leakage into the acceptor detection channel and (3) acceptor photons due to direct excitation by the donor laser[23]. For all these reasons, different combinations of distance distribution and diffusion coefficient may yield the same millisecond-averaged FRET sub-population.

Additional experimental information is beneficial to circumvent these difficulties and to more accurately retrieve distance information. These may include fluorescence anisotropy decays and their usage in smFRET when using pulsed excitation and time-correlated single photon counting (TCSPC)[24–26], fluorescence correlation spectroscopy (FCS)[27–29] and other methods that rely on photon statistics such as probability distribution analysis (PDA)[30,31], burst variance analysis (BVA)[5,32,33] or two-channel kernel-based density distribution estimator (2CDE)[34]. In summary, a multi-parameter approach may produce enough experimental data to retrieve the underlying distance information accurately, decoupling the results from all other possible effects[26].

The richest information source is found in the detected photons themselves, the time intervals between them and their identity (donor or acceptor photons). Therefore, a photon-by-photon approach may retrieve the maximal information content[35]. In that context, a PDA approach consists in "re-coloring" the donor and acceptor photons of single-molecule bursts according to an underlying model. Re-coloring is performed by a Monte-Carlo (MC) simulation of photon numbers in each burst, generated by a Binomial distribution. Then, histograms from the re-colored bursts are compared to the the experimental ones. This approach has been extensively used to fit experimental histograms of FRET efficiency broadened beyond shot-noise by a single distance distribution[36] or by two (or more) FRET states (without specification of their underlying distance distributions[31,37]). Still, these treatments did not take into account the FRET-enhancement that occur due to picosecond-to-nanosecond D-A distance changes and other effects discussed above. Crucially, the single fixed FRET efficiency assigned to each burst in PDA, implicitly assumes a broadening of the FRET peak due to static heterogeneity, while, in most cases, it is more realistic to assume fast D-A distance fluctuations due to linker dynamics. Finally, while PDA has been used mainly for fitting experimental FRET histograms, a similar approach can be used to fit other experimental histograms derived from smFRET measurements.

In this work, we introduce MC-diffusion-enhanced photon inference (MC-DEPI), a photon-by-photon Monte-Carlo-based re-coloring approach to properly and accurately analyze smFRET experiments, taking into account D-A self diffusion and other effects (single or multiple interconverting states, photo-blinking, correction factors). In MC-DEPI, we model D-A self diffusion trajectories as a stochastic process with a characteristic relaxation time τ_{relax}. Instantaneous D-A distances can be computed at arbitrary time points using an Ornstein-Uhlenbeck (OU) process and the Gillespie direct-update formula[38]. Note that, while an OU process directly models Gaussian distributed D-A distances, other distance distribution can be evaluated using a simple transformation (refer to Eq. 15 in sub-section II B on the dynamics module of MC-DEPI). In MC-DEPI, D-A distances are first computed at each photon timestamp, considered as the D excitation time. Then we simulate the D de-excitation process leading to either a D or an A photon. For this purpose, we simulate D-A distance trajectories with time steps much smaller than D and A fluorescence lifetimes. The de-exitation process depends on the the simulated trajectories and FRET efficiencies. At the end, for each timestamp we obtain the photon color (either D or A) and the nanotime (time separation between dye excitation and photon emission). Thus, the simulated data yields FRET histograms as well as donor and acceptor fluorescence decays. In the final step, we use a sequential model-based optimization to find optimal parameters which fit the simulation to the experimental results.

This paper is organized as follows. We first discuss the dependence of FRET on D-A distances. Then, we introduce the theoretical framework behind MC-DEPI. Afterwards we show how a given FRET population may result from many different underlying conditions and how the additional information from other histograms derived from the data can be used to decipher which of the conditions describes best the underlying conformational state or states. Then, we show, as a proof of concept, the results of MC-DEPI-based analysis of nanosecond-alternating laser excitation (nsALEX)[24,25,39] smFRET measurements of the distance between two dyes attached

to a pair of bases in dsDNA. We illustrate how complex its distance-related interpretation may be and how MC-DEPI allows for an accurate retrieval of distance information. Finally, we discuss other possible uses of MC-DEPI in the analysis of more complex systems, whether due to complex photophysics or due to complex distance dynamics.

II. FRET THEORY AND MC-DEPI

In FRET experiments one measures E, the efficiency of transfer of excitation energy from a donor (D) to an acceptor (A) fluorophore. This efficiency depends on the D-A distance, r, according to the Förster relation (Eq. 1):

$$E = \left[1 + \left(\frac{r}{R_0}\right)^6\right]^{-1} \tag{1}$$

where R_0, known as the Förster radius, is the distance at which the donor excitation energy is transferred to the acceptor with 50% efficiency. R_0 depends on the spectral overlap between the D fluorescence spectrum, $F_D(\lambda)$, and the A extinction spectrum, $\epsilon_A(\lambda)$, the D fluorescence quantum yield, ϕ_D, the orientation factor of D and A fluorophores, κ^2, and the refractive index of the medium between them, n (Eq. 2).

$$R_0^6 = \frac{9\log 10}{128\pi^5 N_A}\frac{\kappa^2\phi_D}{n^4}\int F_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda \tag{2}$$

The orientation factor, κ^2, is a function of the angles θ_D or θ_A between the direction of the D or A dipoles and the line connecting the centers of the D and A dipoles, and the angle φ between the D dipole and the A dipole, (Eq. 3).

$$\kappa^2 = (\cos\varphi - 3\cos\theta_D\cos\theta_A)^2 \tag{3}$$

The dependence of E on the 6-th power of r makes FRET a spectroscopic ruler sensitive to distances in the range between $1/2R_0$ and $3/2R_0$. Therefore, for a D-A pair with R_0 of 60 Å, FRET will report accurately on distances in the range 30-90 Å.

The Jablonski diagrams depict the photophysical processes that occur following the excitation of the D fluorophore. After D is excited from the S_0 ground-state to the S_1 excited-state (with a rate $k_{D,ex}$), it will be de-excited back to S_0, $S_1 \rightarrow S_0$, with a rate, k_D, either radiatively yielding an emitted photon or nonradiatively releasing the energy as heat, or due to FRET with a rate k_{FRET}. Another source of de-excitation from S_1 is due inter-system crossing to the triplet state, $S_1 \rightarrow T_1$, also known as triplet blinking, with a rate $k_{Blinking}$. This transition is rare relative to the $S_1 \rightarrow S_0$ transitions, and

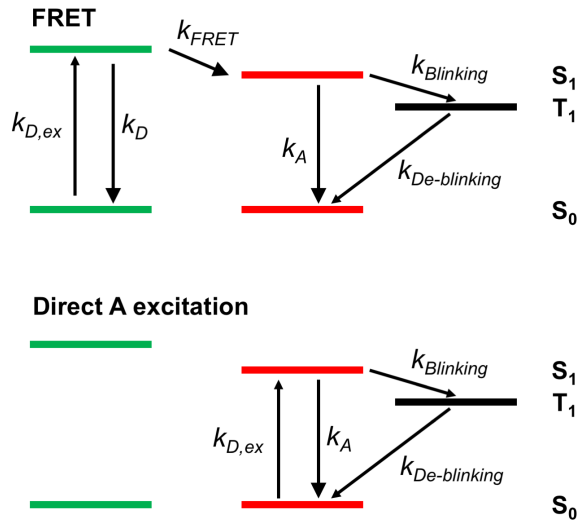

Figure 1. Jablonski diagrams of processes in smFRET leading to photon emission. Top - donor is excited at a rate $k_{D,ex}$ from ground state S_0 to the first singlet excited state S_1. The donor is de-excited from S_1 back to S_0 either through fluorescence, with a rate k_D (the sum of radiative and nonradiative de-excitation rates), or via FRET, with a rate k_{FRET} (Eq. 4). The latter leads to an excitation of the acceptor from S_0 to S_1. The acceptor is de-excited either back to S_0, with a rate k_A or to a triplet state, T_1, via inter-system crossing, with a rate $k_{Blinking}$. The slow transition from T_1 back to S_0 occurs with a rate $k_{De-blinking}$. If the acceptor is in a dark triplet state, it does not function as an acceptor for FRET, causing donor de-excitation only via k_D. Bottom - the acceptor may be excited directly by the excitation source intended for donor excitation.

hence is infrequent. However so is the $T_1 \rightarrow S_0$ transition (called here de-blinking), with a rate $k_{De-blinking}$. In summary, although blinking is rare, when it occurs it takes a long period of time to de-blink and in this time the fluorophore cannot be excited. Although triplet blinking can occur both in the donor and in the acceptor, triplet blinking of the donor does not result in photons. In this analysis we re-color existing photons, therefore processes that do not lead to the emission of a photon will be treated as occurring but not affecting the existing photons we re-color. On the other hand, triplet blinking of the acceptor lead to periods of time in which only the donor can be excited and FRET does not occur, therefore emitting just donor photons with nanotimes exponentially distributed according to k_D as the exponent. Therefore, only the triplet blinking process of the acceptor reproduces another source of photon re-coloring (in this case re-coloring of acceptor photons as donor ones).

In the Jablonski diagram that depicts the photophysical processes in a FRET measurement (Fig. 1, top), the acceptor can be excited via the FRET mechanism. However, at the excitation wavelength intended for donor excitation, acceptor excitation can also occur directly. The

fraction of direct acceptor excitation in well-designed smFRET experiments is small but non-negligible. If the acceptor is directly excited, the Jablonski diagram includes just the photophysical processes in the acceptor (Fig. 1, bottom).

For a given donor-acceptor distance, r, the rate of FRET, k_{FRET}, depends on the sixth power of r and of the Förster radius, R_0 (Eq. 4).

$$k_{FRET}(r) = k_D \left(\frac{R_0}{r}\right)^6 \tag{4}$$

The overall rate of donor de-excitation, $k_{D,FRET}$, is the sum of all possible de-excitation processes in the donor (See Fig. 1, top; Eq. 5).

$$k_{D,FRET}(r) = k_D \left(1 + \frac{R_0}{r}\right)^6 \tag{5}$$

Both E and $k_{D,FRET}$ have probabilistic meanings that are important for the derivation presented in this work. $k_{D,FRET}$ is related to the probability for donor de-excitation as in Eq. 6.

$$p(D^* \to D)(r) = 1 - e^{-k_{D,FRET}(r)\delta t} \tag{6}$$

When r is constant, Eq. 6 is valid for any $\delta t > 0$. E is the probability that donor de-excitation will occur due to FRET. Therefore, these two parameters are at the heart of simulating the donor excited-state survival dynamics in FRET (Fig. 2). The fundamental experimental observables in smFRET experiments are the identity of the detected photons (donor or acceptor photons), the absolute detection times (macrotime), and the detection time relative to the time in which the donor was excited (nanotime). The distribution of donor or acceptor nanotimes, also known as the donor or acceptor fluorescence decays, describe the characteristic times of the $S_1 \to S_0$ transition. These times can be retrieved from the solution of rate equations resulting from the Jablonski diagram in Fig. 1. Then, the de-excitation rates can be retrieved as the exponent in fits of the fluorescence decays to exponential functions. However, Eqs. 1, 5 describe the process of FRET for a constant r. What if r changes as a function of time, $r(t)$? Then, each distance, $r(t_i)$, introduces different $k_{D,FRET}$ and E values, and hence multi-exponential decays, where each exponent represents the contribution of a given distance value out of many others. Additionally rapid r dynamics introduces an overall decrease in donor nanotimes and increase FRET events[19,20]. r dynamics can be incorporated into analytical rate equations whose solution can be compared to experiments to find best-fit parameters . Such approach has been used in bulk FRET experiments[15,20,21] but, to the best of our knowledge, never in smFRET ones. However, this analytical approach becomes impractical as more complex photophysical schemes are added. Conversely, as shown in the next section, the Monte-Carlo scheme proposed here can be extended to describe more complex photophysics with minimal efforts.

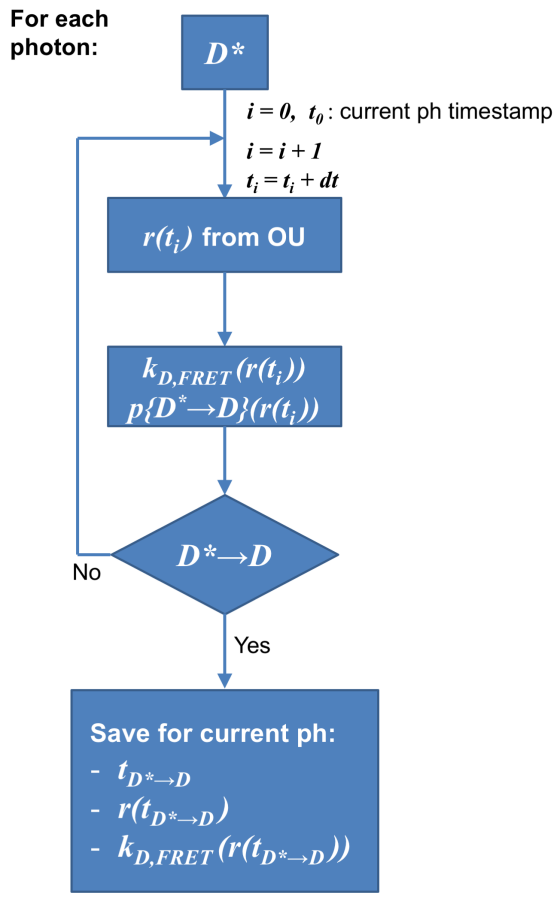

Figure 2. The core algorithm for the Monte-Carlo calculation of FRET per photon. Starting from t_0, immediately after excitation (the donor is excited, D^*), the donor-acceptor distance, r, is advanced in time-steps dt (usually of the TCSPC resolution) according to an Ornstein-Uhlenbeck (OU) process of diffusion in a potential well, to produce a distance trajectory, $r(t_i)$[38]. For each time step, the donor de-excitation rate and probability are calculated. Using the donor de-excitation probability the first Monte-Carlo step is to test whether at a given time step de-excitation occurred. If yes, the time iterations halt and the time of donor de-excitation, the donor-acceptor distance and the rate of donor de-excitation at this time are saved and will be used for Monte-Carlo steps in recoloring this photon.

A. MC-DEPI – photophysics module

As an alternative to using coupled rate equations to describe the Jablonski diagrams of Fig. 1, it is also possible to use Monte-Carlo simulations to simulate r dynamics and the photophysics. In this approach, we simulate photon IDs and nanotimes on top of existing photon macro-

times, namely re-coloring experimental photons. The diagram in Fig. 2 describes an MC-DEPI simulation per a given photon. The core algorithm starts at excitation time t_0, when the donor reaches the excited state S_1 (D^* being the excited donor). The trajectory of D-A distances, $r(t)$, is produced by a separate process that we will describe later (see sub-section II B on the dynamics module of MC-DEPI). The probability of donor de-excitation $p(D^* \rightarrow D)(r(t_i))$ is computed according to Eq. 6 for each distance, $r(t_i)$. Then, we test whether de-excitation occurred by comparing $p(D^* \rightarrow D)(r(t_i))$ to a uniformly random number between 0 and 1. If donor de-excitation did not occur, we advance to the next distance at time t_{i+1}. When donor de-excitation occurs, we save the current time $t_i = t_{D^* \rightarrow D}$ and the current D-A distance $r(t_{D^* \rightarrow D})$.

Using these three variables we move on in the re-coloring process (Fig. 3). Using Eq. 1 we calculate E. $E(r(t_{D* \rightarrow D}))$ is the probability that donor de-excitation occurred due to FRET, which leads to acceptor excitation, and eventually emission of an acceptor photon (in a simulation of re-coloring existing photons). Therefore, the MC assessment of FRET is a process of re-coloring photons as acceptor photons if FRET occurred, or as donor photons if not. However, $E(r(t_{D* \rightarrow D}))$ describes the probability for FRET to occur assuming it is the only process that leads to photon re-coloring. In an smFRET experiment there are other processes that dictate photon IDs: (i) direct acceptor excitation (Fig. 1, bottom), in which, with a probability d_T, a photon from the donor laser is absorbed by the acceptor fluorophore instead of from a donor; (ii) the leakage of a fraction Lk of donor photons in the acceptor channel; (iii) the γ factor bias caused by the different probabilities of detecting donor or acceptor photons, due to different fluorescence quantum yields and detection efficiencies; (iv) acceptor photoblinking (Fig. 1) causing periods of time that will include solely donor photons due to the acceptor being in a dark triplet state. The probability for acceptor blinking, $p_{Blinking}$, is constant and depends on the competition between the inter-system crossing $S_1 \rightarrow T_1$ and $S_1 \rightarrow S_0$ transitions (Eq. 7).

$$p_{Blinking} = \frac{k_{Blinking}}{k_A + k_{Blinking}} \tag{7}$$

Normally, in smFRET experiments, the measure E_{RAW} is corrected to obtain E. Here we use the inverse relation (Eq. 8) to compute $E_{RAW}(r(t_{D^* \rightarrow D}))$ from $E(r(t_{D^* \rightarrow D}))$ as a function of the three correction factors Lk, d_T and γ. Note that d_T is the ratio of the A and D absorption cross-sections at the donor-excitation wavelength (Eq. 8b). $E_{RAW}(r(t_{D^* \rightarrow D}))$ is essentially the probability of detecting an A photon with a given FRET efficiency and correction factors. Thus, we use $E_{RAW}(r(t_{D^* \rightarrow D}))$ to randomly select if a photon is labeled as D or A.

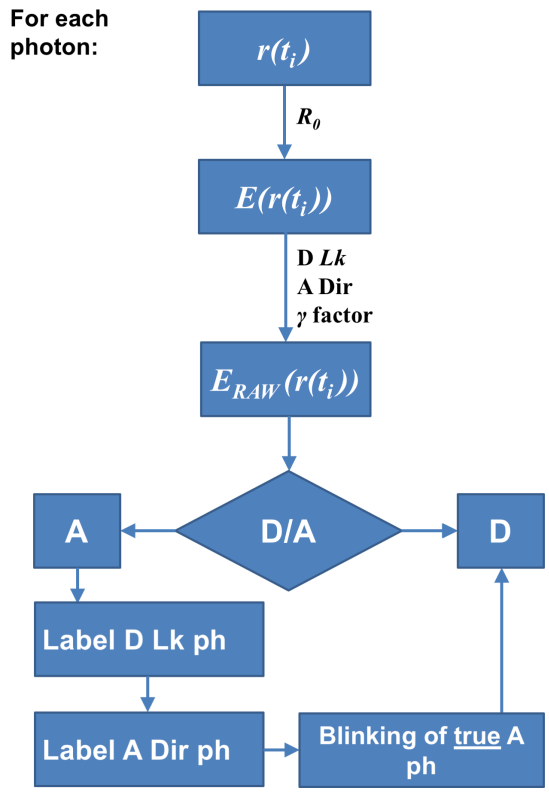

Figure 3. The photon re-coloring scheme per photon. Using the donor-acceptor distance $r(t_i)$ at donor de-excitation time saved in the previous step (see 2), we calculate the FRET efficiency $E(r(t_i))$ (Eq. 1). Then, using γ factor, the donor leakage Lk and the direct acceptor excitation d_T we obtain $E_{RAW}(r(t_i))$ (Eq. 8). $E_{RAW}(r(t_i))$ is the probability of detecting a photon in the acceptor channel due FRET, donor leakage, acceptor direct excitation and taking into account the bias introduced by the γ factor. Using $E_{RAW}(r(t_i))$ we randomly label photon as D or A. Next, we randomly select fractions of A photons as leakage and acceptor excitation photons. The remaining A photons are purely due to FRET. We assign A photons from FRET or direct acceptor excitation a nanotime that is the sum of the donor de-excitation time plus a time drawn from the acceptor fluorescence decay distribution. For D and donor-leakage photons, we set the nanotime to the donor de-excitation time. In a last Monte Carlo step, we simulate acceptor photo-blinking. Each *"true"* A photon (due to FRET or direct A excitation) falling during triplet blinking period is relabeled as a D photon. For these photons no FRET can occur, thus the nanotime is drawn from the intrinsic D fluorescence decay distribution. The intrinsic fluorescence decays for both D or A can be single- or multi-exponential.

$$E_{RAW} = \frac{E(\gamma - L_k) + L_k + d_T\gamma}{E(\gamma - L_k - 1) + L_k + d_T\gamma + 1} \tag{8a}$$

$$d_T = \frac{\sigma^A_{D_{ex}}}{\sigma^D_{D_{ex}}} \tag{8b}$$

Then, we further label A photons as caused by FRET, direct acceptor excitation or from donor leakage, using the Lk and d_T. We remember that only acceptor photons that are labeled as originating from FRET or from direct acceptor excitation are true acceptor-emitted photons. Finally, we simulate periods of acceptor triplet blinking, recoloring acceptor-emitted photons happening during triplet blinking as D photons. The duration of each A dark (blinked) state is drawn randomly from an exponential distribution with mean lifetime $1/k_{De-blinking}$. All the acceptor photons re-colored as D due to acceptor triplet blinking, are marked as "acceptor blinked".

In addition to encoding the photon IDs, we simulate photon nanotimes for different classes of photons as follows: (i) for purely donor photons (not "acceptor blinked"), the nanotime is simply the time of donor de-excitation, $t_{D^*\to D}$; (ii) for an acceptor photon due to FRET, the nanotime is the sum of $t_{D^*\to D}$ with a random number drawn from the intrinsic A fluorescence decay distribution; (iii) for an acceptor photon from direct acceptor excitation, the nanotime is a random number drawn from the intrinsic A fluorescence decay distribution, since it represents an acceptor photon without the addition of time spent in the donor; (iv) for an acceptor photon due to leakage of donor into the acceptor detection channel, the nanotime is $t_{D^*\to D}$, since it represents a donor-emitted photon that was simply detected in the "wrong" detection channel; finally (v) for a donor photon that has been marked as "acceptor blinked", the nanotime is a random number drawn from the intrinsic D fluorescence decay distribution, since the acceptor is absent, hence the donor in these times acts as a donor-only species. In the simplest case intrinsic D or A fluorescence decays are a exponential distributions with rates k_D or k_A. However, in practice, many organic dyes exhibit more complex decays requiring a multi-exponential model with two or more rates.

B. MC-DEPI – dynamics module

So far we described the photophysics module of MC-DEPI. Now we describe the simulation of trajectories of D-A distances $r(t)$. Consider a single-molecule burst. The burst is a time series of detected donor and acceptor fluorescence photons. The photon ID, macrotime and nanotime of each photon, described earlier, are accurately defined as:

1. Photon ID of photon i, ID_i, which depends on whether the ith detected photon in the burst was a donor or acceptor photon

2. Macrotime of photon i, $t_{macro,i}$, which is the photon detection time at a resolution of a few tens of nanoseconds for the ith photon in a burst (comparable to pulsed-excitation repetition time in pulsed-excitation smFRET)

3. Nanotime of photon i, $t_{nano,i}$, which is the photon detection time relative to the moment of excitation for the ith photon in a burst (in the resolution of a few ps)

The description of the detection time of the ith photon in a burst relative to the beginning of the measurement is $t_{macro,i}+t_{nano,i}$. This definition of $t_{macro,i}$ allows assuming it was the time at which the molecule that produced the detected photon was first excited.

We start by describing r dynamics, $r(t)$, for a single conformational state that is characterized by an equilibrium distribution of donor-acceptor distances, $p_{Eq.}(r)$, and distance changes dictated by diffusion in a potential well of the conformational state with a donor-acceptor self-diffusion coefficient, D. The relation between the potential well of the simulated conformational state, $U(r)$, and $p_{Eq.}(r)$ is given in Eq. 9.

$$p(r) \propto e^{-\frac{U(r)}{k_B T}} \tag{9}$$

where k_B is the Boltzmann constant and T is the absolute temperature. There are many models of $p_{Eq.}(r)$, including the Gaussian distribution, $p_G(r)$ (Eq. 10a), the skewed or radial Gaussian distribution, $p_{rG}(r)$ (Eq. 10b), and models that describe polymers such as the wormlike chain, $p_{WLC}(r)$ (Eq. 10c).

$$p_G(r) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}} \tag{10a}$$

$$p_{rG}(r) = c(r-r_0)^2 e^{-\frac{[(r-r_0)-\mu]^2}{2\sigma^2}}$$
$$p_{rG}(r \le r_0) = 0 \tag{10b}$$

$$p_{WLC}(r) = \frac{c(r-r_0)^2}{[1-(\frac{r-r_0}{L})^2]^{\frac{9}{2}}} e^{-\frac{9L}{8l_p}\cdot\frac{1}{1-(\frac{r-r_0}{L})^2}}$$
$$p_{WLC}(r \le r_0) = 0 \tag{10c}$$

where c is a normalization factor, μ and σ are related to the mean and the standard deviation of the distance (these are exactly the mean and standard deviation of the distance in the case of $p_G(r)$), L and l_p are the polymer overall length (the contour length) and the persistence length, respectively, and r_0 is a distance offset below which inter-dye distances do not occur. Models of distance distributions with an offset distance (Eqs. 10b and 10c) are a good approximation for conditions in which dyes are attached to flexible regions of a molecule connected by a rigid part. It is important to note that although $p_G(r)$ (Eq. 10a) describes the distribution of distances for a harmonic potential well (see Eq. 9), mathematically it can allow nonzero probabilities for negative distance values, which is of course unphysical. The other distance distribution models described here are mathematically defined to have nonzero probabilities only for distance values above the distance offset, r_0.

While the molecule crossed the detection volume, it had a specific time-trajectory of donor-acceptor distances, $r(t)$, where at each instance the distance depended on its relevant probability drawn from the distance distribution at equilibrium, $p_{Eq.}(r)$, on the distance at the time the previous photon, $i-1$, was detected and on the time-interval between the previous and the current distances, $t_i - t_{i-1}$. The dependence r in the trajectory, $r(t_i)$, on the previous distance, $r(t_{i-1})$, and on the time interval between them changes as a function of the donor-acceptor self-diffusion coefficient, D – the faster the diffusional change of r is, the faster $r(t)$ decays from $r(t_{i-1})$ to a distance that is randomly sampled from $p_{Eq.}(r)$. For a single-molecule burst, the macrotime of the first photon, $t_{macro,1}$, defines time zero for the single-molecule distance trajectory. Therefore the distance at time zero, $r(t_0)$, can be drawn randomly from $p_{Eq.}(r)$. Following a given simulated $r(t)$, the photon IDs and nanotimes can be simulated from excitation time that yielded the ith photon to detection time of the ith photon (from $t_{macro,i}$ to $t_{macro,i}+t_{nano,i}$) according to the description depicted above (Sub-section II A and Figs. 2 and 3). Afterwards, the photon IDs and nanotimes can be simulated from detection time of the ith photon to excitation time that yielded the $i+1^{th}$ photon ($t_{macro,i}+t_{nano,i}$ to $t_{macro,i+1}$). This way, the photon IDs and the nanotimes of the photon time-series are simulated in a photon-by-photon fashion without removing the experimental macrotimes. This is important since the density and number of macrotimes in a burst affect the shot-noise characteristics, the experimental brightness and consequently the time resolution for identification of conformational dynamics. We next move on to describe how to simulate the distance trajectory depending on the underlying $p_{Eq.}(r)$ and the self-diffusion coefficient, D.

The following Ornstein-Uhlenbeck (OU) stochastic process describes diffusional motion in a one-dimensional harmonic potential well (Eq. 11).

$$r'(t+dt) = r'(t) - \frac{1}{\tau_c} r'(t) + N\sqrt{Ddt} \tag{11}$$

where N is a unitary normally distributed random number (mean of 0, standard deviation of 1), $r'(t)$ is the distance time-series, dt is a positive infinitesimal time increment and τ_c is the dynamics relaxation time that depends on the self-diffusion coefficient, D, and the standard deviation , σ, of the underlying Gaussian distribution, according to Eq. 12.

$$\tau_c = \frac{2\sigma^2}{D} \tag{12}$$

This definition, however, requires using infinitesimally small time increments, dt, which makes its use impractical for the typically large time differences between consecutive photons in single-molecule bursts (with intervals in the µs). Gillespie developed a simple direct update formula that allows advancing the OU process in Eqs. 11 and 12 by arbitrary time intervals, Δt (Eq. 13[38]).

$$r'(t+\Delta t) = r'(t)e^{-\frac{\Delta t}{\tau_c}} + N\sqrt{\frac{D\tau_c}{2}\left(1 - e^{-\frac{2\Delta t}{\tau_c}}\right)} \tag{13}$$

Using this approach allows simulating the distance after a time interval Δt, assuming we know what was the previous distance for a Gaussian distribution with a μ=0 and σ=1. If $p_G(r)$ (Eq. 10a) is the simulated distance distribution and it is defined by other μ and σ values, each of the simulated distance values, r', can be converted to distance values, r, of the underlying distance distribution, which is in this case $p_G(r)$, according to Eq. 14.

$$r(r) = r'(t) \cdot \sigma + \mu \tag{14}$$

If other models of distance distributions are used, then each distance, r', can be mapped to the distance represented by the simulated distance distribution model, r, by considering that the probability of r' to occur is the same as the probability of r (Eq. 15a).

$$\int_{-\infty}^{r'} N(r',0,1)\, dr' = \int_0^r p_k(r)\, dr \tag{15a}$$

$$\int_{-\infty}^{r'} N(r',0,1)\, dr' = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{r'}{\sqrt{2}}\right)\right] \tag{15b}$$

where $p_k(r)$ is the simulated distance distribution that might be one of the models presented in Eq. 10 with specific simulated parameter values and k represents the name of the model. Note that the integral on the right hand-side of Eq. 15a is taken starting from a distance of 0, since it assumes the model of the simulated distance distribution is defined as a distribution that represents only positive distance values. However the Normal distribution to the left-hand side is symmetric around 0 and defined over the whole real number space. Additionally, an analytical solution to the term on the left (for the standard normal distribution) exists (Eq. 15b). If an analytical solution to the term on the right can be derived, it is preferable to use it explicitly. However if it does not have any known analytical solution, it should be numerically calculated. One can test this procedure to simulate a time-series with constant time steps and show that no matter what values of $p(r)$ and D are used, the autocorrelation of the resulting distance trajectory decays exponentially with a mean relaxation time, τ_c, and that the mean square displacement divided by the time, $\langle r^2\rangle(t)/t$, as a function of time, t, reaches a plateau with a value of the simulated D, as expected.

C. MC-DEPI: intra-lifetime diffusion

In time-correlated single photon counting (TCSPC) measurements photon detection times are collected in

time bins representing an array of possible discretized nanotimes, also known as TCSPC time bins. The size of the TCSPC bin, δt, is typically in picoseconds and defines the accuracy and uncertainty of nanotime recording. The number of TCSPC bins multiplied by the TCSPC bin size (the TCSPC time resolution) defines the maximal possible nanotime values in an experiment. For an experimental array of TCSPC time bins we produce $r'(t)$ using Gillespie's direct update formulas for the OU stochastic process (Eq. 13). We start by simulating $r'(t = \delta t)$ knowing what was $r'(t_0)$, which is basically the time of excitation, which we defined as $t_{macro,i}$. Then we compute $r'(t = 2\delta t)$, after a time interval of δt, knowing $r'(t = \delta t)$ that was calculated in the previous step. This stepwise calculation yields the simulated $r'(t)$ starting from the moment of excitation and throughout all possible TCSPC time bins, in jumps of δt. Then, using Eq. 14 or 15 (depending on the underlying simulated $p_{Eq.}(r)$) we map the time-series of distances, $r'(t)$, that follows a standard normal distribution to a time-series of distances, $r(t)$, that follows the simulated $p_{Eq.}(r)$, using Eq. 15. Next, we follow the steps given above (see subsection II A) and in Fig. 2 to calculate the donor de-excitation for the ith photon, $t_{D*\to D,i}$ and the simulated distance at that time, $r(t_{D*\to D,i})$. Then we follow the additional steps given above (see sub-section II A) and in Fig. 3 to define the photon ID and nanotime.

D. MC-DEPI: inter-timestamps diffusion

The previous step allowed proper simulation of the photon ID and nanotime of the ith photon, taking care of possible diffusion-enhanced FRET effects. However, for advancing the simulation to the macrotime of the $i+1^{th}$ photon, $t_{macro,i+1}$, the time of donor de-excitation for the ith photon, $t_{D^*\to D,i}$, and the simulated distance at that time, $r(t_{D^*\to D,i})$, is used. The distance that was simulated at $t_{macro,i} + t_{D^*\to D,i}$ and the time interval between this time and the $i + 1$ macrotime, $\Delta t = t_{macro,i+1} - (t_{macro,i} + t_{D^*\to D,i})$ will be used to simulate the distance at $t_{macro,i+1}$ for the $i + 1^{th}$ photon using Gillespie's direct update formulas (Eq. 13) and the distance mapping approach in Eqs. 14 and 15, (both to map distances r to r' and backwards from r' to r).

E. MC-DEPI workflow: the overall simulation

The simulation procedure described above will be performed on all bursts, where each burst represents the macromolecule under study that had a different initial distance, $r(t_{macro,1})$, randomly sampled from the underlying simulated equilibrium distance distribution $p_{Eq.}(r)$. Due to the burst separation being much larger than the distance fluctuation relaxation time, different bursts have independent initial distance distance dynamics, representing different molecules out of the ensemble that are not synchronized in time.

The simulation can also describes systems interconverting between more than one conformational state (each associated with a given $p_{Eq.}(r)$ and D). In this case, the current state is simulated for each photon using a Continuous-Time Markov Chain (CTMC) model. In the CTMC formalism, the probability of being in each state can be computed at arbitrary times knowing the transition matrix and the initial state probabilities. This property is used to randomly generate the current state for each photon. The theory has been described by Gopich and Szabo[40] with reference to single-molecule experiments or, for the general CTMC theory, in any statistics book treating stochastic processes[41]. See also the attached notebook "Continous-Time Markov Chain" for details on the formalism. Once the state of the photon is selected, the simulation proceeds as for the single state case, using the distance distribution and diffusion coefficient of the selected state.

F. MC-DEPI: fitting the experiment

Using MC-DEPI, we can compare a simulation of a given set of conditions ($p_{Eq.}(r)$, D, number of interconverting states and their interconversion rates) to the experimental results. Then, we can iterate the simulation until the simulated results matches the experimental ones. The most important observable to be compared is the FRET histogram, similarly to what is done in PDA. But, even when comparing only FRET histograms, MC-DEPI crucially differs from the PDA because it takes into account not only distance distributions but also the D-A self-diffusion which generates the diffusion-enhanced FRET effect. In addition, MC-DEPI includes a detailed description of acceptor triplet blinking. Moreover, differently from PDA, MC-DEPI is a photon-by-photon approach and can reproduce fluorescence decays and other histograms which can yield a more informative comparison with experiments[26,27,32,42,43]. In this work, we focus on comparing the FRET histograms and the donor and acceptor fluorescence decays. Beechem and Haas have shown that a global analysis of both the donor and acceptor fluorescence decays resulting from a time-resolved FRET measurement, increases the accuracy of the retrieved $p_{Eq.}(r)$ and D parameters[20].

In order to find parameters that best fit the experiment, we build a loss function (also called cost function) which is smaller the closer simulation is to the experiment. Due to the Monte Carlo nature of the simulation, the loss function has an intrinsic noise, i.e. multiple evaluation of the same point give different results). For this reason, performing a classical gradient-based optimization which requires a deterministic function is unfeasible. Instead, we use a Bayesian global optimization approach, where at each iteration a new simulation is performed and a new statistical model for the loss function is computed. The statistical model, also known as acquisition function,

learns a better approximation of the loss function at each iteration. By minimizing the acquisition function at each iteration, we find a new set of parameters which is chosen as the new point to be evaluated for the loss function. In this work, the acquisition function was computed via Gaussian process regression as implemented in the open source *scikit-optimize* python package[44].

As noted before, the loss function used in this paper is the sum of two components, one assessing the FRET histograms and the other assessing the fluorescence decays. The E component of the loss function is the mean square error of the simulated and experimental FRET histogram. The fluorescence decay component, conversely, is split in two sub-components one for D and one for A decay. For each decay we use a negative log-likelihood function, similar to what is used for fitting fluorescence decays using the maximum likelihood approach[45]. In order to combine different losses in a single loss function, we normalize each component by the standard deviation of the Monte Carlo noise. We compute the standard deviation empirically by repeating 100 MC-DEPI simulations with the same set of parameters but with different seeds for the random number generator. Finally, the global loss function is computed as the sum the different components divided by their standard deviation. For details of the derivation of the loss function see section A 1.

The excitation impulse response function (IRF) was taken into account by adding to each simulated nanotime a random number distributed as the experimental IRF distribution. The IRF is different for the D or A channel nanotimes.

III. MC-DEPI: AMBIGUITY IN FRET HISTOGRAMS

The most common representation of smFRET results is a histogram of FRET efficiency values of all the identified single-molecule bursts, better known as FRET histograms. These histograms show Gaussian-like sub-populations of bursts with common mean FRET efficiencies, known as FRET sub-populations. The mean FRET efficiency of a single FRET sub-population is a time-average of FRET dynamics caused by changes of donor-acceptor distances that occurred while the single molecule crossed the detection volume (typically in ms). The distance dynamics that is time-averaged and yields the corresponding FRET sub-population can be described relative to the conformational state characteristics, namely the equilibrium distance distribution, $p_{Eq.}(r)$, and the inter-dye self-diffusion coefficient, D. However, $p_{Eq.}(r)$ can have many different shapes, mean distances and widths, which affect the value of the mean FRET efficiency. Additionally, if distance dynamics occur in times comparable or faster than the donor lifetime, $1/k_{D,FRET}$, the combination of $p_{Eq.}(r)$ and D influences the value of the mean FRET efficiency[19,20].

Another important experimental parameter is the width of a FRET sub-population. These sub-populations will always have a minimal width caused by the calculation of FRET efficiency from bursts having a limited number of donor and acceptor detected photons and from the effect of background photons, to produce the better known shot-noise-limited FRET sub-population. However, widening beyond the shot-noise limit may occur either due to static heterogeneity (two molecular species with distinct FRET sub-populations that highly overlap) or dynamic heterogeneity (one molecular species that interconverts between multiple conformational states with distinct mean FRET efficiencies). Therefore, widening of a FRET sub-population beyond the shot-noise limit entails additional information.

As a special case, distinct FRET sub-populations may be interpreted as different molecular species having different mean FRET efficiencies. Alternatively, they may be interpreted as an outcome of a single molecular species capable of dynamically transitioning between distinct conformational states characterized by two different mean FRET efficiencies, but only if the timescale of the transitions is larger than the characteristic duration of the single-molecule bursts (which reports on the time it took the molecule to traverse the detection volume). However, if the timescale of transitions between the different conformational states is comparable to the single-molecule burst durations, many bursts will include donor and acceptor photons with different mean FRET efficiencies. This is because while the molecule crosses the detection volume, it interconverts multiple times between different conformational states with different mean FRET efficiencies. The outcome is a FRET histogram that includes a sub-population that *bridges* between the FRET sub-populations of the different conformational states. The faster the transition dynamics is relative to the single-molecule burst durations, the larger the amplitude of the *bridge* sub-population will be on the expense of a decrease in the amplitude of the sub-populations of the original conformational states[46,47]. Finally if the dynamics is more than ten times faster than single-molecule burst durations, the outcome will be a single shot-noise limited FRET sub-population with a mean FRET efficiency that equals the equilibrium weighted average of the mean FRET efficiencies of the underlying conformational states. Overall, single FRET sub-population may be interpreted in many different ways and not necessarily by a single conformational state with a mean FRET efficiency as that of the FRET sub-population. This point is very important in the debate on how to properly and accurately translate smFRET results into distance information, conformational states and their intricate dynamics.

To stress these points we used the MC-DEPI approach to simulate a multitude of different conditions that may lead to the same shot-noise limited FRET efficiency with $\langle E \rangle = 0.4$ (arbitrarily chosen). For this set of simulations we chose the following set of parameter values: donor and acceptor fluorescence characterized by a single lifetime

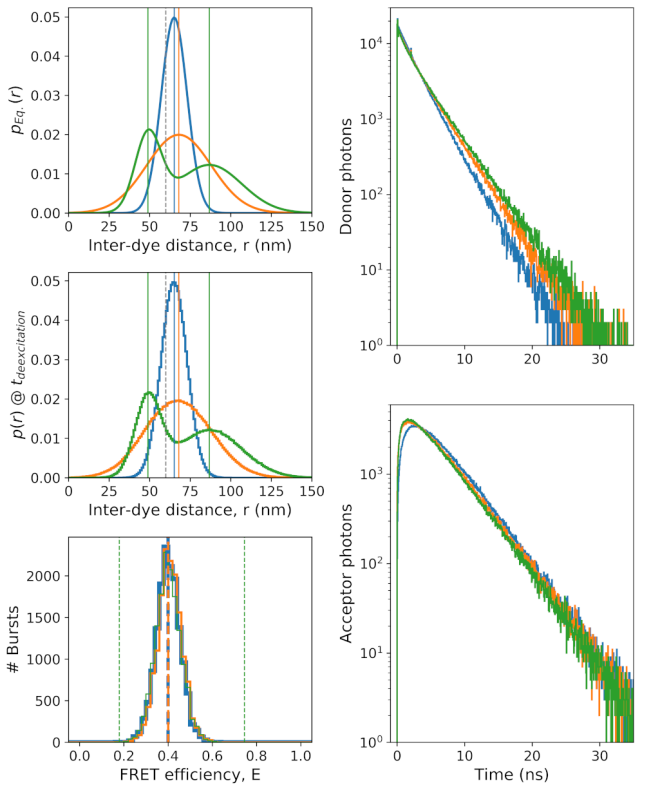

Figure 4. Three different conditions produce the same FRET histograms and the additional information lies in the shape of other histograms. Three different Equilibrium distance distributions, $p_{Eq.}(r)$, with different shapes (solid lines) and their corresponding distribution of distances at donor de-excitation time, $p(r)@t_{deexcitation}$, (stair plot). The mean of the underlying $p_{Eq.}(r)$ is indicated with vertical lines. This figure reports simulation results of Gaussian $p_{Eq.}(r)$ of a single conformational state that is narrow (blue), wide (red) and two conformational states that dynamically interconvert with a relaxation time of τ_r=5 µs (green). A dashed vertical gray line shows the value of R_0. Bottom-left: The FRET histogram of the three different conditions is the same. The dashed vertical lines show the mean FRET values of each state. Right: Donor (top) and acceptor (bottom) fluorescence decays, contain the additional information on the different conditions that yield the same FRET histograms.

component having τ_D = 3.8 ns, τ_A = 4 ns, R_0 = 60 Å and TCSPC bin widths of 10 ps. For the sake of simplicity the simulated fluorescence decays were not convoluted with an IRF. Simulations re-colored the existing photon macrotimes of single-molecule bursts in a measurement of a mixture of two dsDNA molecules doubly labeled with a donor, ATTO 550, and an acceptor, ATTO 647N, with 7 and 17 bp separations (which we name d7 and d17; additional information on the measurement and on the burst analysis is given the Materials and Methods appendix A). The simulated photon IDs and nanotimes are then used to plot the underlying $p_{Eq.}(r)$ (Fig. 4, Right-top), the simulated distribution of distances at the time of donor de-excitation, $p(r)@t_{deexcitation}$ (Fig. 4, Right-center),the simulated FRET histogram (Fig. 4, Right-bottom) and the simulated donor and acceptor fluorescence decays (Fig. 4, Right-top and bottom, respectively).

As a first example, we simulated three different conditions from conformational states modeled by Gaussian $p_{Eq.}(r)$ (Eq. 10a): (i) a single conformational state with a narrow $p_{Eq.}(r)$(μ=65.14 Å, σ=8 Å, τ_c=50 ns; Fig. 4, blue); (ii) a single state with a wide $p_{Eq.}(r)$ (μ=68.07 Å, σ=20 Å, τ_c=50 ns; 4, orange); and (iii) two conformational states with transition relaxation time of τ_r=5 µs (f_1=0.385, μ_1=49 Å, σ_1=8 Å, $\tau_c, 1$=50 ns, f_2=0.615, μ_2=86.81 Å, σ_2=20 Å, $\tau_c, 2$=50 ns; Fig. 4, green). Note that although these three conditions have very different $p_{Eq.}(r)$ (Fig. 4, Left-top), their FRET histograms turn out to be exactly the same, where all characterized by $\langle E\rangle$=0.4 and by a shot-noise limited width, even for the case of two conformational states (5 µs transition dynamics results in averaged-out FRET sub-population; Fig. 4, Left-bottom). It is important to note that in the simulated conditions that include FRET dynamics of τ_c=50 ns, the distribution of distances of molecules at the time in which donor was de-excited (Fig. 4, center-left) is the same as $p_{Eq.}(r)$ (Fig. 4, top-left; $\tau_c > \tau_D$, the donor lifetime). However, the donor and acceptor fluorescence decays have different shapes (Fig. 4, Center-top and -bottom, respectively). Therefore for these conditions, the analysis of the fluorescence decays is essential for distinguishing between the three different underlying conditions.

Next we use MC-DEPI to simulate the results of FRET dynamics of a single conformational state modeled by a Gaussian (Eq. 10a), having the same width (σ=20 Å) with increasing mean distance (μ=68.07, 70.95, 74.5 and 79.28 Å) and decreasing dynamics relaxation time, τ_c (τ_c=50, 8, 2.5 and 0.8 ns; Fig. 5, Left-top, blue, orange, green and red, respectively), where at 8 and 2.5 ns, the dynamics time is comparable to the donor fluorescence lifetime (4 ns) and at 0.8 ns, it is faster than the donor lifetime. In the simulation, the combination of these values correspond to D with values of 16, 100, 320 and 1000 Å^2/ns (using the transformation in Eq. 12). These simulation conditions were chosen to reproduce a FRET histogram with $\langle E\rangle$=0.4 (Fig. 5, Left-bottom). Without taking into account the effect of diffusion enhancement on FRET, the increase in distance in these simulations should have yielded a corresponding decrease in $\langle E\rangle$ (Eq. 1; Fig. 5, Left-bottom, vertical dashed lines). However, if FRET dynamics occurs in times comparable or faster than the donor lifetime (τ_D=4 ns in this case), FRET events from shorter distances represented in $p_{Eq.}(r)$ are enhanced due to the rapid dynamics of r and the higher $k_{D,FRET}$ and p_D at lower r values (Eqs. 5, 6, respectively). Indeed the distances at the time in which donor de-excitation occurred (Fig. 5, Left-top) are the shorter distances represented in $p_{Eq.}(r)$ (Fig. 5, Left-center): the smaller the FRET dynamics relaxation time, τ_c, relative to the donor fluorescence lifetime, τ_D=4

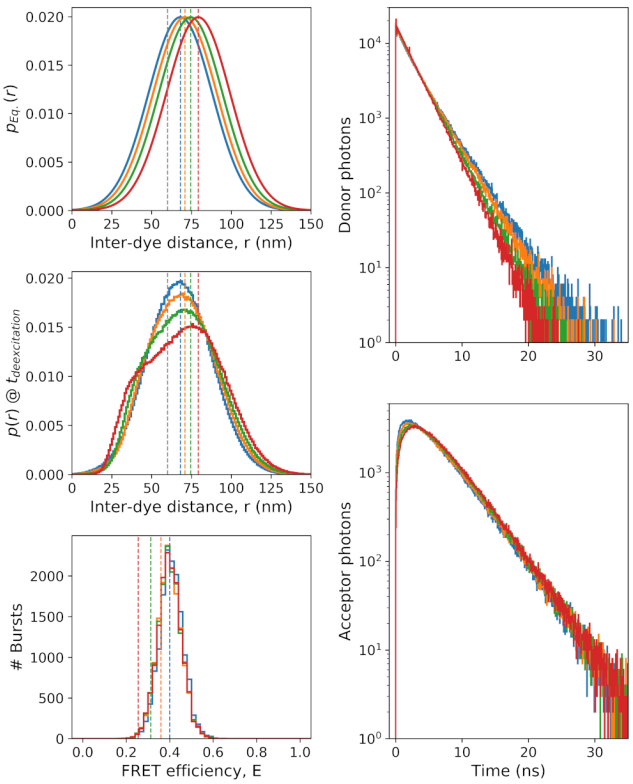

Figure 5. Diffusion-enhanced FRET that lead to the same FRET histogram. Distance dynamics in the timescale (τ_c=50, 8, 2.5 and 0.8 ns as blue, orange, green and red, respectively) of the donor fluorescence lifetime (τ_D=4 ns) lead to the same FRET histogram (Left-bottom) for four different single Gaussian $p_{Eq.}(r)$ with the same width and different mean distances (Left-top). $p(r)@t_{deexcitation}$ (stair plot) is different than $p_{Eq.}$ (Center-left) due to the effect of enhancement of donor de-excitation at shorter times from shorter distances. The mean of the underlying $p_{Eq.}(r)$, without the effect of diffusion-enhanced FRET, is indicated with vertical lines. A dashed vertical gray line shows the value of R_0. The dashed vertical lines show the mean FRET values of each state. Right: Donor (top) and acceptor (bottom) fluorescence decays, contain the additional information on the different conditions that yield the same FRET histograms.

ns, the larger is the deviation between the distribution of distances at the time of donor excitation and at the time of its de-excitation. Therefore, the simulation conditions allow rapid FRET dynamics to compensate for the increase in distances represented in $p_{Eq.}(r)$, and to yield the same FRET histograms. If the outcome of these different conditions is the same single FRET subpopulation, how can one distinguish between these different conditions? The shapes of the donor fluorescence decays are different (Fig. 5, Right-top). In the case of a wide distance distribution as the one simulated here, if $\tau_c > \tau_D$, the donor fluorescence decay will be multi-exponential where each lifetime component represents a different distance out of $p_{Eq.}(r)$ (Fig. 5, Right-top, blue). However, from that point on, the smaller τ_c is compared to τ_D, more FRET events occur from shorter distances than from longer ones, which leads to a higher weight of the smaller lifetime components until when $\tau_c < \tau_D$, the donor fluorescence decay becomes mono-exponential (Fig. 5, Right-top, red). This also affects the acceptor fluorescence decay because it depends on the donor fluorescence decay (Fig. 5, Right-bottom). Therefore, the different shapes of the donor and acceptor fluorescence decays have the additional information that can help in distinguishing between the different conditions simulated here.

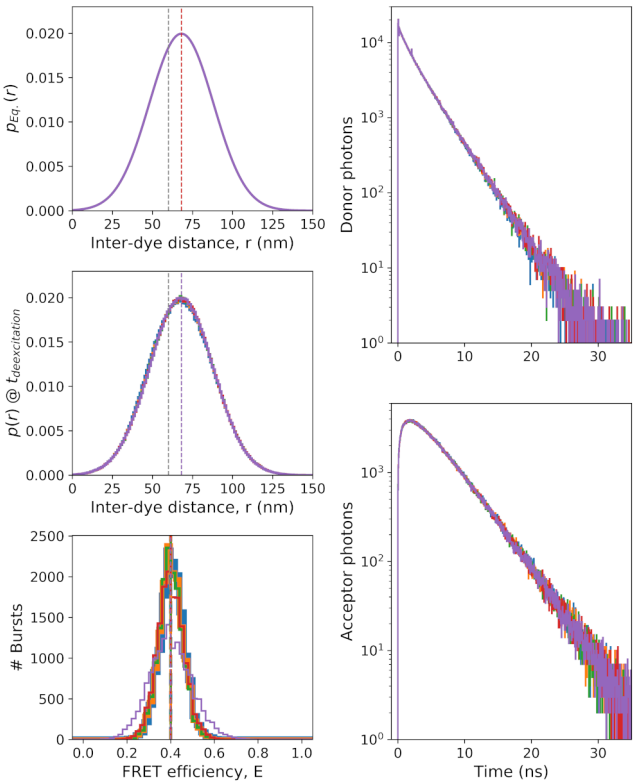

Figure 6. Slow dynamics in a single conformational state. Distance dynamics in the timescale (τ_c=50 ns, 200 ns, 1 µs, 10 µs and 100 µs as blue, orange, green, red and magenta, respectively) that is slower than the donor fluorescence lifetime (τ_D=4 ns) lead to the same FRET histogram (Left-bottom) for the same single Gaussian $p_{Eq.}(r)$ (Left-top) as long as τ_c is much faster than the inter-photon times (the FRET histogram for τ_c=100 µs gets wider). $p(r)@t_{deexcitation}$ (stair plot) is the same as $p_{Eq.}(r)$ (Center-left) because dynamics slower than donor lifetime does not lead to large changes of distances between the times of donor excitation and de-excitation. The mean of the underlying $p_{Eq.}(r)$ is indicated with vertical lines. A dashed vertical gray line shows the value of R_0. The dashed vertical lines show the mean FRET values of each state. Right: donor (top) and acceptor (bottom) fluorescence decays, do not contain the additional information on the different conditions that yield the same FRET histograms.

In cases where the sole differences arise from FRET dynamics that take more time compared to τ_D, the flu-

orescence decays are not expected to include information additional to the FRET histograms. Here we used MC-DEPI to simulate exactly the same Gaussian $p_{Eq.}(r)$ (μ=68.07 Å, σ=20 Å) with FRET dynamics occurring at different timescales, all larger than τ_D. The values of simulated τ_c were 50 ns, 200 ns, 1 µs, 10 µs and 100 µs (Fig. 6, blue, orange, green, red and magenta, respectively). These conditions simulate dynamics of the donor-acceptor distance as diffusion in a 1D harmonic potential well. While this is a perfectly smooth description of a 1D potential well, a physical potential well may be rugged due to the existence of multiple conformational states separated by energetic barriers smaller than k_BT. Such slow dynamics may arise from increased viscosity or due to the ruggedness of the potential well. Still, it is important to mention that although we simulate such dynamics, the slowest τ_c that was reported in the scientific literature to describe dynamics in a single conformational state had values of a few µs for single-stranded DNA in viscous media[48]. In most cases such slow dynamics characterized the transition dynamics between conformational states, rather than within a single state. These simulation conditions of the same $p_{Eq.}(r)$ (Fig. 6, Left-top), produce the same distribution of distances at the time of donor de-excitation (Fig. 6, Left-center) and a single FRET sub-population with $\langle E \rangle$=0.4 (Fig. 6, Left-bottom). Note, however that in the case where τ_c=100 µs (Fig. 6, Left-bottom, magenta), the FRET histogram becomes wider than the shot-noise limited width of the FRET histograms in the case of faster FRET dynamics. This occurs because a large number of inter-photon times are smaller than τ_c. This means that changes in distances occur slower than the time between consecutive photons. If FRET dynamics occur faster than inter-photon times, the donor-acceptor distance at the time of the detection of each photon in the burst can be randomly samples from $p_{Eq.}(r)$, no matter what was the time interval between this photon and the distance at the time of the previous photon. This means that the memory of the distance between consecutive photons is lost. If, however, FRET dynamics occur slower than inter-photon times, the distance at the time a photon was detected also has dependence on the distance at the time of the previous photon was detected. In this case, there will be photons with time intervals shorter than τ_c that will retain the memory of the distance from the time of the previous photon. Therefore, the additional width in the FRET sub-population qualitatively hints on additional information about dynamics.Yet, the fluorescence decays of the same $p_{Eq.}(r)$ with FRET dynamics slower than τ_D are expected to have exactly the same shape and indeed they are (Fig. 6, Right-top and -bottom).

Another set of conditions that may lead to a FRET histogram with a single shot-noise limited FRET sub-population is when there are actually more than a single conformational state and the relaxation time of the transitions between them, τ_r, is slow compared to the donor fluorescence lifetime but much faster than the time the

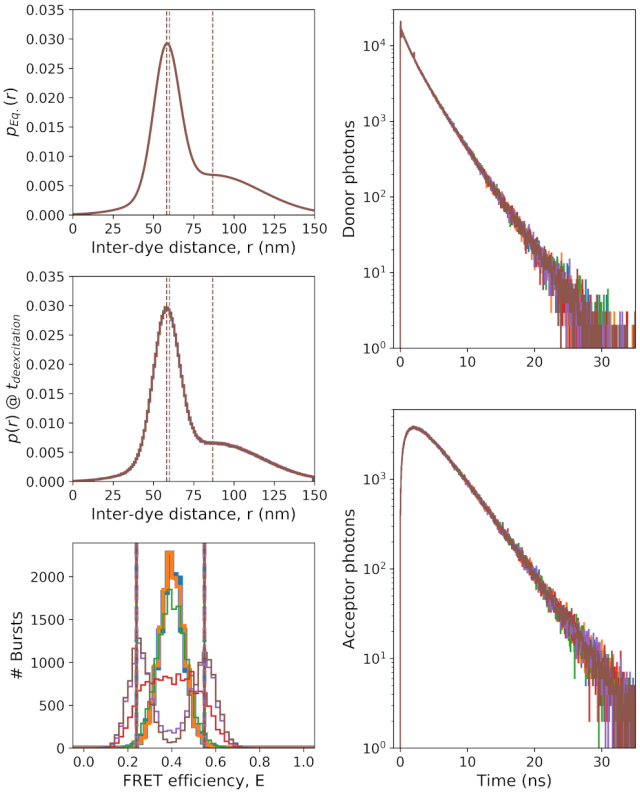

Figure 7. Slow dynamics between two conformational states. Conformational interconversion dynamics in the timescale (τ_r=1 µs, 10 µs, 100 µs, 1 ms, 10 ms and 100 ms as blue, orange, green, red, magenta and brown, respectively) that is slower than the donor fluorescence lifetime (τ_D=4 ns) lead to the same shot-noise limited single population FRET histogram (Left-bottom) for same single Gaussian $p_{Eq.}(r)$ (Left-top) as long as τ_r is much faster than the inter-photon times (the FRET histogram for τ_r=100 µs gets wider and at τ_r=1 ms and above splits into two FRET sub-populations). $p(r)@t_{deexcitation}$ (stair plot) is the same as $p_{Eq.}(r)$ (Left-center), because dynamics slower than donor lifetime does not lead to large changes of distances between the times of donor excitation and de-excitation. The means of the underlying $p_{Eq.}(r)$ are indicated with vertical lines. A dashed vertical gray line shows the value of R_0. The dashed vertical lines show the mean FRET values of each state. Right: Donor (top) and acceptor (bottom) fluorescence decays, do not contain the additional information on the different conditions that yield the same FRET histograms.

single molecules traverse through the detection volume. Here we used MC-DEPI to simulate exactly the same two conformational states as two Gaussian $p_{Eq.}(r)$ (f_1=0.5, μ_1=58.2 Å, σ_1=8 Å, $\tau_{c,1}$=50 ns, f_2=0.5, μ_2=86.81 Å, σ_2=30 Å, $\tau_{c,2}$=50 ns) with transition relaxation time, τ_r= 1 µs, 10 µs, 100 µs, 1 ms, 10 ms nd 100 ms (Fig. 7, top-left, blue, orange, green, red, magenta and brown, respectively). In this simulation, as in the simulation of the single state slow dynamics, the distribution of distances at the time of donor excitation (Fig. 7, Left-top) is identical to the one at the time of donor-de-excitation (Fig. 7,

Left-center), mainly since both the conformational state dynamics and the transition dynamics are slower than the timescale of the donor lifetime ($\tau_c, \tau_r >> \tau_D$). This, of course, promises that the different simulated conditions will produce identical donor and acceptor fluorescence decays (Fig. 7, Right-top and -bottom, respectively). As expected, when the transition dynamics occur in times slower than the single-molecule burst durations (τ_r = 100, 10 ms) the FRET histogram includes two FRET sub-populations that are well-separated (Fig. 7, Left-bottom, brown and magenta, respectively). when the transition dynamics occur in times comparable to burst durations (τ_r=1 ms; Fig. 7, Left-bottom, red), a large portion of the bursts include multiple transitions between the two states. Therefore, the FRET efficiency values of these bursts are between the values of the mean FRET efficiency of the two sub-populations. From that point, the faster the transition dynamics is, the more bursts will include more frequent transitions between the two states, and produce a FRET histogram with a single averaged-out sub-population. Using the above-mentioned conditions at τ_r=10 µs or faster this averaged-out FRET sub-population is characterized by $\langle E \rangle$=0.4 and a shot-noise limited width. Additional experimental information is required to identify that this single FRET sub-population actually represents a time-average of two conformational states with distinct FRET characteristics. The shape of the donor and acceptor fluorescence decays (Fig. 7, Right-top and -bottom, respectively) may serve as additional experimental information distinguish between the possibility for a single conformational state and the case of two conformational state, but moving from the former to the latter has to be justified experimentally. One common justification follows statistical inference rules. According to this approach, one moves from a simple model of a single conformational state to a more complex model of two interconverting conformational states only if the former fails to be compared properly to the experimental results.

Finally, acceptor photoblinking can lead to FRET dynamics, between times in which both donor and acceptor photons are being emitted (the FRET species) and others where the acceptor is in a dark triplet state for long periods of time, hence only donor photons are emitted and with nanotimes dictated just by the donor intrinsic de-excitation processes, with a rate k_D. we simulated a set of conditions that may lead to a FRET histogram with a single shot-noise limited FRET sub-population as long as the characteristic times the acceptor spends in the triplet state, $\tau_{triplet}$, are shorter than the inter-photon times. The simulated conditions were a single Gaussian $p_{Eq.}(r)$ (μ=65 Å, σ=20 Å, τ_c=50 ns) with blinking probability, $p_{Blinking}$=0.05 (Eq. 7), and $\tau_{triplet}$= 5 µs, 250 µs, 1 ms and 5 ms (Fig. 8, Left-top, orange, green, red and magenta, respectively). For comparison, we also simulated another single Gaussian $p_{Eq.}(r)$ without acceptor photoblinking, that still lead to a shot-noise limited FRET sub-population with $\langle E \rangle$=0.4 (Fig. 8, Left-top, blue). In this simulation the distribution of distances at the time of donor excitation (Fig. 8, Left-top) is identical to the one at the time of donor-de-excitation (Fig. 8, center-left), mainly since the acceptor blinking dynamics is slower than the timescale of the donor lifetime

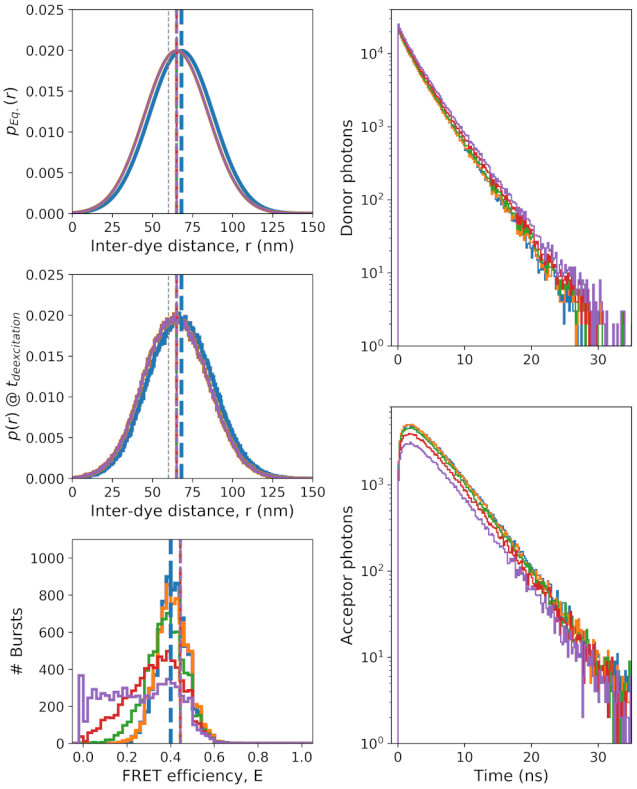

Figure 8. Acceptor blinking dynamics. Acceptor triplet lifetimes in the timescale ($\tau_{triplet}$=5 µs, 250 µs, 1 ms, 1 ms and 5 ms as blue, orange, green, red and magenta, respectively) with blinking probability, $p_{Blinking}$=0.05 (Eq. 7) lead to the same shot-noise limited single population FRET histogram (Left-bottom) for the same single Gaussian $p_{Eq.}(r)$ (Left-top, orange, green, red and magenta) as long as $\tau_{triplet}$ is much faster than the inter-photon times (the FRET histogram for $\tau_{blinking}$=250 µs gets wider and at $\tau_{blinking}$=1 ms and above becomes smeared from $\langle E \rangle$=0.4 dan towards $\langle E \rangle$=0) (when the acceptor is blinked, FRET cannot occurs, hence the donor functions as a donor-only species, hence $\langle E \rangle$=0 while the acceptor is in the triplet state). $p(r)@t_{deexcitation}$ (stair plot) is the same as $p_{Eq.}(r)$ (Left-center), because dynamics slower than donor lifetime do not lead to large changes of distances between the times of donor excitation and de-excitation. The mean distance for the underlying $p_{Eq.}(r)$ is indicated with vertical lines. A dashed vertical gray line shows the value of R_0. The dashed vertical lines show the mean FRET values of each state. Right: Donor (top) and acceptor (bottom) fluorescence decays of the different simulated conditions have slightly different shapes, hence contain additional information on the different conditions that yield the same FRET histograms. Shown in blue are conditions with which no acceptor triplet blinking yields a shot-noise limited FRET sub-population with $\langle E \rangle$=0.4.

$(1/k_{Blinking}, \tau_{triplet} >> \tau_D)$. The donor fluorescence decays show slightly different shapes, due to a fraction of the donor photons that originates from a donor-only species when the acceptor is in the triplet state and does not function as an acceptor of FRET (Fig. 8, Right-top). The acceptor fluorescence decays have the same shapes but different amplitudes, since acceptor triplet blinking results in more donor photons on the expense of acceptor photons, since there are long periods in which the acceptor cannot be excited (when it is in the triplet state) and due to the fact that the simulation re-colors a constant amount of photons (Fig. 8, Right-bottom). When acceptor blinking dynamics occur in times slower than the single-molecule burst durations ($\tau_{blinking}$ = 5, 1 ms) the FRET histogram becomes smeared from $\langle E \rangle$=0.4 to $\langle E \rangle$=0 (Fig. 8, Left-bottom, magenta and red, respectively). when acceptor blinking dynamics occur in times comparable to burst durations but slower than the inter-photon times ($\tau_{blinking}$=250 µs; Fig. 8, Left-bottom, green), a large portion of the bursts include multiple blinking transitions. Therefore, the FRET sub-population becomes wider than shot-noise limited width. However when $\tau_{triplet}$=5 µs, the FRET sub-population is characterized by a shot-noise limited width (Fig. 8, Left-bottom, orange). Additional experimental information is required to identify that this single FRET sub-population actually represents a time-average of between a FRET species with $\langle E \rangle > 0.4$ and a donor-only species with $\langle E \rangle = 0$. The shape of the donor fluorescence decays (Fig. 8, Right-top) may serve as a starting point to distinguish between the possibility for a single conformational state and the case of triplet blinking, but moving from the former to the latter has to be justified experimentally.

In summary, the comparison of different experimental histograms (FRET histograms and fluorescence decays) to their MC-DEPI simulated counterparts can serve as a better approach for retrieving the underlying conformational states and their dynamics.

IV. MC-DEPI: THE CASE OF DOUBLE-STRANDED DNA

Next, we show how using MC-DEPI simulations we are capable of analyzing smFRET experimental results. For this we performed smFRET measurements of two dsDNA constructs labeled with the same pair of donor and acceptor dyes (ATTO 550 and ATTO 647N, respectively). In one molecule we name d7, the dyes were separated by 7 base-pairs (bp), and in the other we name d17, the dyes were separated by 17 bp. We performed nanosecond alternating laser excitation (nsALEX) smFRET measurements, on freely-diffusing labeled dsDNA molecules, allowing us to: (i) gain the photon ID, its macrotime and its nanotime, for each detected photon; (ii) gain detected photons with interphoton times in the microseconds timescale; and (iii) separate between molecular species with fluorescently active dyes and others where one of the dyes has photobleached. Using a series of control measurements and analyses, we also calculated for each labeled dsDNA molecule the values of R_0, donor fluorescence lifetime components, donor fluorescence quantum yields, and some of the correction factors required for accurate smFRET analysis (for more details on the experiments, please see the Materials and Methods appendix A).

First, we analyzed the experimental results of the d7 molecules. We performed a global fit of MC-DEPI simulation results,trying many different conditions, to both the FRET histogram and to the donor and acceptor fluorescence decays. We used the MC-DEPI framework taking into account dsDNA as a single conformational state. In this case, we used a Gaussian distance distribution to describe the conformational state. Additionally, we included the possibility of rapid donor-acceptor self-diffusion to introduce FRET enhancement. We also included acceptor photo-blinking as a possibility. In fitting experimental results to modeled results, the residuals (the difference between the former and the latter) is usually assessed. However, it is important to understand that different simulations of the same parameters induce results that are slightly different from each other. This introduces an intrinsic dispersion. It is therefore important to understand that the residuals of a given fit are considered satisfactory already if they are comparable to the intrinsic dispersion.

The results of the fitting procedure are shown in Fig. 9. One can see how τ_c=0.9 ns changes the distance distribution from what it is in equilibrium (at excitation time; Fig. 9 Left-top) to what it is at donor de-excitation (Fig. 9, Left-center). The fit to the FRET histogram (Fig. 9, Left-bottom) and to the donor and acceptor fluorescence decays (Fig. 9, Right-top and -bottom, respectively) are shown (Fig. 9, gray - experimental; blue-fitted) following a fit to MC-DEPI simulations with a single Gaussian $p_{Eq.}(r)$ with diffusion and acceptor photo-blinking. The best fit parameters of $p_{Eq.}(r)$ are μ=32 Å and σ=15 Å. Using the recently introduced FRET-restrained positioning and screening (FPS) tool[49], we calculated the values of μ and σ expected assuming each orientation in the overall dye available volume is equally probable (Fig. 9, Right-top, inset). The values were μ=35.7 Å and σ=10.2 Å. Note, however, that the assumptions that this calculation take, may lead to different values. This is because the probability of each orientation in the dye available volume is not necessarily equally probably. Additionally, calculation of the expected FRET using FPS assumes the dye explore all orientations in the dye available volume rapidly relative to the dye fluorescence lifetime. This may not be true in our case. The best fit diffusion relaxation time was τ_c=0.93 ns, which is in the order of the donor fluorescence lifetime (a 0.94 fraction with 4.02 ns and a 0.06 fraction with 0.37 ns). This value together with the value of σ translate into a donor-acceptor diffusion coefficient of D=484 Å^2/ns (using the transformation in Eq. 12). A re-

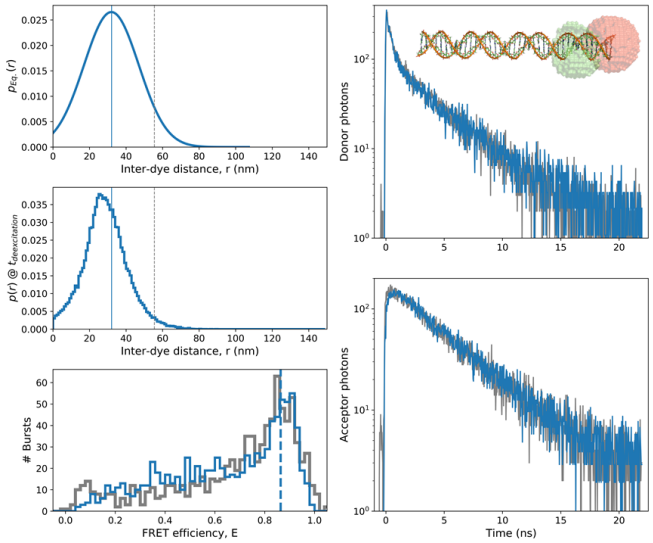

Figure 9. Fitting results to measurements of dsDNA with donor and acceptor separated by 7 bp. nsALEX smFRET measurements of a dsDNA labeled with ATTO 550 and ATTO 647N as donor and acceptor dyes, separated by 7 bp (d7 molecule) were taken. The experimental results are shown in grey and the best fit MC-DEPI simulation of a Gaussian distance distance distribution is shown in blue. The best fit parameters are a single Gaussian $p_{Eq.}(r)$ with μ=32 Å and σ=15 Å, together with distance dynamics τ_c=0.93 ns corresponding to D=484 Å^2/ns (using the transformation in Eq. 12). Additionally, the best fit parameter values of the acceptor photoblinking were $\tau_{triplet}$=2.04 ms with blinking probability $p_{Blinking}$=0.0295. The single Gaussian $p_{Eq.}(r)$ is shown (Left-top). Note that the distribution of distances at the time of donor de-excitation, $p(r)@t_{deexcitation}$, is different than the one at the time of excitation (in equilibrium), $p_{Eq.}(r)$, mainly at the long distance range, due to the diffusion-enhanced effect. These results fit well with the FRET histogram (Left-bottom) and with the donor and acceptor fluorescence decays (Right-top and bottom, respectively). Note that here the simulated fluorescence decays are shown after convolution with the experimental IRFs. The depiction of the dsDNA molecule with donor and acceptor dyes separated by 7 bp (shown are the dyes available volumes in green and red, respectively), is shown in the inset of the Right-top panel.

laxation time of 0.93 ns is well within the dye depolarization times reported for these dyes in the literature[50–53]. Indeed, FRET is higher than what it should have been without taking into account the diffusion-enhanced effect. However, FRET could have been higher without the balancing effect of acceptor photo-blinking (which reduces FRET values). Without including acceptor photo-blinking, the width of the FRET histogram turns out to be too narrow compared to the experimental one. Additionally, the low FRET tail shown in the experimental results (Fig. 9, Left-bottom, grey) cannot be explained otherwise.

The acceptor photo-blinking best-fit parameters are Photo-blinking probability, $p_{Blinking}$=0.0295 (which is simply the inter-system crossing efficiency) and an acceptor photo-blinked lifetime, $\tau_{triplet}$=2.04 ms.

Next we show the best fit results of MC-DEPI simulations for the same dsDNA sample, only this time labeled with donor and acceptor dyes separated by 17 bp (a d17 molecule). This sample includes exactly the same DNA sequence as in the d7 molecule, the same dyes and the same measurement conditions, hence we do not expect to get different acceptor photo-blinking parameters (same photo-blinking probability and same acceptor photo-blinked lifetime). Therefore, we will use the same acceptor photo-blinking parameters found in the fit to the d7 sample, as constants in the fit to the d17 sample.

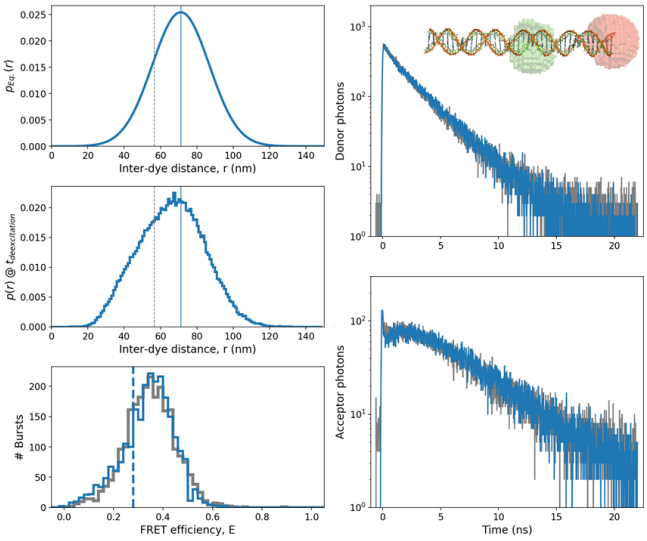

Figure 10. Fitting results to measurements of dsDNA with donor and acceptor separated by 17 bp. nsALEX smFRET measurements of a dsDNA labeled with ATTO 550 and ATTO 647N as donor and acceptor dyes, separated by 17 bp (d17 molecule) were taken. The experimental results are shown in grey and the best fit MC-DEPI simulation of a Gaussian distance distance distribution is shown in blue. The best fit parameters are a single Gaussian $p_{Eq.}(r)$ with μ=71.1 Å and σ=15.7 Å, together with distance dynamics τ_c=1.21 ns corresponding to D=407 Å^2/ns (using the transformation in Eq. 12). The single Gaussian $p_{Eq.}(r)$ is shown (Left-top). Note that the distribution of distances at the time of donor de-excitation, $p(r)@t_{deexcitation}$, is different than the one at the time of excitation (in equilibrium), $p_{Eq.}(r)$ due to the diffusion-enhanced effect. These results fit well with the FRET histogram (Left-bottom) and with the donor and acceptor fluorescence decays (Right-top and bottom, respectively). Note that here the simulated fluorescence decays are shown after convolution with the experimental IRFs. The depiction of the dsDNA molecule with donor and acceptor dyes separated by 17 bp (shown are the dyes available volumes in green and red, respectively), is shown in the inset of the Right-top panel.

The results of the fitting procedure are shown in Fig. 10. One can see how τ_c=1.2 ns changes the distance distribution from what it is in equilibrium (at excitation time; Fig. 10 Left-top) to what it is at donor

de-excitation (Fig. 10, Left-center). The fit to the FRET histogram (Fig. 10, Left-bottom) and to the donor and acceptor fluorescence decays (Fig. 10, Right-top and -bottom, respectively) are shown (Fig. 10, gray - experimental; blue-fitted) following a fit to MC-DEPI simulations with a single Gaussian $p_{Eq.}(r)$ with diffusion and acceptor photo-blinking. The best fit parameters of $p_{Eq.}(r)$ are μ=71.1 Å and σ=15.7 Å. Note that for these $p_{Eq.}(r)$ values, if diffusion-enhanced FRET would have been neglected, the mean FRET efficiency would have been lower than what it is with it (Fig. 10, Left-bottom, blue dashed vertical line versus peak FRET population, respectively). Fitting these results without taking into account diffusion-enhanced FRET would have yielded a distance distribution with a significantly shorter mean distance.

Using the recently introduced FRET-restrained positioning and screening (FPS) tool[49], we calculated the values of μ and σ expected assuming each orientation in the overall dye available volume is equally probable (Fig. 10, Right-top, inset). The values were μ=65.3 Å and σ=10.8 Å. The best fit diffusion relaxation time was τ_c=1.21 ns, which is in the order of the donor fluorescence lifetime (a 0.94 fraction with 3.93 ns and a 0.06 fraction with 0.45 ns). This value together with the value of σ translate into a donor-acceptor diffusion coefficient of D=407 Å^2/ns (using the transformation in Eq. 12). A relaxation time of 1.21 ns is well within the dye depolarization times reported for these dyes in the literature[50–53]. As in the best fit results of d7, also in d17 FRET is higher than what it should have been without taking into account the diffusion-enhanced effect. However, FRET could have been higher without the balancing effect of acceptor photo-blinking (which reduces FRET values). Without including acceptor photo-blinking, the width of the FRET histogram turns out to be too narrow compared to the experimental one.

In summary, using MC-DEPI simulations we were able to fit nsALEX smFRET experimental results of d7 and d17 molecules. Doing so we found that dye linker dynamics, leading to rapid donor-acceptor distance changes (in the range from hundreds of picoseconds to a few nanoseconds) exists and affects FRET results as so did the existence of acceptor photo-blinking. We assume the distance distribution at equilibrium, $p_{Eq.}(r)$, that describes dye movements solely due to dye and linker movements, can be described as a Gaussian function. Using this assumption we find that the mean donor-acceptor distances are different than can be calculated using FPS by 10%. Additionally, we find that the standard deviation of the donor-acceptor distance is larger than the one calculated by FPS, by $\sim$ 5 Å.

V. DISCUSSION

A. Benchmarking on smFRET of dsDNA molecules - not so straightforward

As one can understand, the analysis of smFRET data to yield meaningful and precise distance information is not so straightforward. First, as the simulations have shown, a single FRET population can have many underlying interpretations, all of which are valid as long as not proven differently with additional data from the experiment. In the recent years a part of the single-molecule community has put efforts into standardizing smFRET as a tool for accurate retrieval of distance information[11]. They have done so by measuring the same dsDNA doubly-labeled samples across different laboratories, benchmarking on the rigidity and known structure of dsDNA. The study has shown comparisons of the donor-acceptor apparent distance, calculated directly from the peak FRET efficiency (found by fitting the FRET histogram) using Eq. 1. The authors of the paper assume (i) the dye rotational rate is much faster than the rate of donor de-excitation due to FRET; (ii) the donor-acceptor distance changes (by diffusion) much slower than the donor de-excitation time. Although these assumptions are explicitly expressed, they are not absolutely valid even for molecules such as dsDNA molecules. The assumption that changes in r occur much slower than the donor fluorescence lifetime means that each molecule that was excited had a specific value of r at the moment of excitation that has not changed until the donor was de-excited, a few hundreds of picoseconds to nanoseconds afterwards. While in many cases this serves as a useful approximation, in other cases this approximation does not hold any more. For instance, when using smFRET to measure the distance between donor and acceptor fluorophores labeling a rigid dsDNA molecule, this assumption may break. The organic dyes used in smFRET are large and are connected via long linkers. Just the rotational dynamics of the dyes may introduce distance changes, since rotational dynamics are manifested as changes in dye angles, which after multiplication by the dye size yield changes of the positions of the center of the dyes in space. Since the mean size of these dyes from their attachment atoms to dsDNA bases and until the center of the fluorophore is in the range 10-20 Å, such rotational dynamics may introduce large distance changes, that can yield distance distributions with standard deviations in the range 8-10 Å (calculation performed using the FRET-restrained positioning and screening software[49]). The rotational correlation times of such dyes were assessed in several works from analyses of fluorescence anisotropy decays[50–53]. The typical rotational correlation times of these dyes are in the range 0.3-1 ns, and the typical fluorescence lifetime of these dyes is in the range 1-4 ns. Therefore, in these cases, r changes in the timescale of the fluorescence lifetime or slightly faster. In summary, the effect of 0.3-1 ns rotational dynamics of large dyes

(with 1-4 ns fluorescence lifetimes) with large linkers on top of a rigid molecule , lead to diffusion-enhanced FRET that has to be taken into account in the distance interpretation of smFRET measurements of molecules as simple as dsDNA ($\sim 10\%$ differences in the mean D-A distance compared to when not taking into account the effect of diffusion-enhanced FRET).

Additionally, proper description of the underlying distance distribution and acceptor photo-blinking complicate the interpretation of smFRET data on molecules even as simple as dsDNA and have to be properly handled. Finally, even rigid molecules, such as dsDNA, that are always treated as having a single conformational state, may be explained as showing rapid dynamics between different conformational states. Although the dsDNA is indeed rigid, if one (or more) of the dyes tend to stack at the termini of the duplex for long times[54,55], this already introduced states in which the distribution of distances between the centers of the donor and acceptor dyes is different than when the dyes are free to rotate. Therefore there can be an additional layer of conformational dynamics between two conformational states having close distance distributions. Hence conformational dynamics has to be taken into account in the analyses of smFRET experiments, if the trials to interpret the data as a single conformational state fail.

B. Different types of smFRET experiments that can be analyzed by the DEPI approach

All diffusion-based smFRET measurements produce detected photons, where the photon macrotimes and photon IDs are recorded. smFRET measurements based on continuous-wave (cw) excitation do not produce the photon nanotimes. This, however, does not mean the MC-DEPI approach cannot be used to analyze such experimental results, Still it will be very hard if not impossible to distinguish between different combinations of $p(r)$ and D values. Additionally, in cw-based smFRET, the only histograms that are available for comparison with the simulated photons are the FRET histogram. For MC-DEPI analyses of cw-based smFRET measurements, both the donor and acceptor fluorescence lifetimes can serve either as additional free fitting parameters, or can be assumed to have values as the ones reported in the literature.

Additionally, different types of diffusion-based smFRET analyses can help narrow down the search for the parameter value range that yield simulations that fit well with the experimental results. The analysis of fluorescence correlation functions may allow resolving the relaxation times of the conformational changes, τ_c. This can help in better identifying the correlation term from FRET dynamics, to resolve τ_c. A pulsed-interleaved excitation (PIE; also known as nanosecond alternating laser excitation, nsALEX) allows separating single-molecule bursts of molecules carrying just a donor or just an acceptor. This can help identify the exact donor and acceptor fluorescence lifetimes from the same experiment.

Schuler and co-workers have introduced an approach that combines cw excitation with high time resolution using TCSPC to retrieve the FRET-related fluorescence correlation functions that range from seconds to picoseconds. This approach, better known as nsFCS, allowed identifying FRET dynamics at times that are inaccessible in the conventional setups for measuring detecting fluorescence from single molecules[29]. The analysis according to the MC-DEPI approach can compare simulated FRET histograms and nsFCS correlation curves. In the case of nsFCS, fluorescence decays are not recorded as TCSPC histograms. Nevertheless, the nsFCS correlation functions include the anti-bunching process that can be modeled as analogous to the fluorescence decays, assuming the excitation rate was low[56].

Finally, the work presented here was based on the distance analysis of FRET assuming the orientational dynamics does not contribute to changes in FRET. Dye rotational dynamics introduce dynamics both in r and κ^2. Gopich and Szabo have shown that in case of a constant r and rotational correlation times that are five times smaller than donor fluorescence lifetime, rotational dynamics introduce only minute deviations of the FRET efficiency dependence on κ^2[22]. The ratio between the dye rotational correlation time and fluorescence lifetime describes experimental values very well for typical organic dyes used in smFRET[50–53]. Only at ratios larger than 0.5 this deviation becomes significant. Overall, this means we can safely assume the contribution of rotational dynamics to κ^2 dynamics is negligible and dynamics-enhanced FRET is introduced mostly due to distance dynamics. It is important to mention that this assessment was based on the assumption that the probability density function of κ^2 is the one that yields a mean of $2/3$, which is introduced by assuming both donor and acceptor experience rotations in all possible θ and ϕ angles. In most cases not all θ and ϕ angles are accessible by the dyes, which may or may not introduce a different mean κ^2 value.

A multi-parameter fluorescence detection (MFD) smFRET experiment involves recording single-molecule photons from four different detection channels: donor or acceptor and also parallel and perpendicular polarizations per each donor/acceptor channel[26]. This allows calculating not only donor and acceptor fluorescence decays but also the associated fluorescence anisotropy decays. Parameters extracted from analysis of fluorescence anisotropy decays have direct links to the dyes' orientational dynamics and to the boundary conditions of such dynamics in space, usually treated as a cone in which the dye wobbles[57]. The results of analyses of the rotational dynamics of the dyes from fluorescence anisotropy decays have direct links to, κ^2, through geometrical considerations. If we were to mimic donor and acceptor fluorescence decays through simulating the photon IDs and nanotimes from first principles and distance dynamics,

simulating fluorescence anisotropy decays is also possible from first principles and rotational dynamics, which also affect the instantaneous values of κ^2, hence the values of the donor de-excitation rate constants, $k_{D,FRET}$, and FRET efficiencies, E[58,59].

C. MC-DEPI: beyond FRET

MC-DEPI allows to simulate, separately, the dynamics and the photophysics, and then to combine them. In the case presented here, the dynamics of a single donor-acceptor distance was simulated and then used in the framework of the FRET photophysics for proper re-coloring of photons. The algorithm for photon re-coloring was laid down as flow charts (Figs. 2 and 3) describing the basic transitions in the Jablonski diagram (Fig. 1) and their associated dependence on the instantaneous donor-acceptor distance at each time step of the dynamics. Therefore the framework of MC-DEPI can allow implementation of other schemes for use in analyses of complex experiments.

As a first example, one can combine FRET with protein-induced fluorescence enhancement (PIFE)[60]. In PIFE a dye that functions as a molecular rotor upon excitation is used. A common dye used in PIFE experiments is Cy3 that exhibits trans-cis isomerization after it is excited mostly in the trans isoform. However, while de-excitation of Cy3 from its trans isoform mostly results in emission of a photon, de-excitation from the cis isoform is almost absolutely nonradiative. Nevertheless, when Cy3 is physically restricted, (e.g. by exclusion of a nearby bound protein), it becomes more fluorescent due to the inhibition of the trans-to-cis transition, leading to increase in its de-excitation from the trans isoform[17]. Additionally, PIFE can be treated as another molecular ruler. While the FRET ruler reports on donor-acceptor distances in the range 3-9 nm, PIFE may report on distances of Cy3 from the surface of a bound protein in the range 0-3 nm[17,60]. Therefore, one may simulate the dynamics of the PIFE-measured distance and then employ a series of calculations and Monte Carlo steps to employ the photophysics of PIFE in each step. In reference to FRET, the combination of FRET and PIFE can be employed by simulating the dynamics of the FRET distance and the PIFE distance and at each time step the donor de-excitation is evaluated. The evaluation of the de-excitation event is now dependent not only on the intrinsic fluorescence and on FRET (see Fig. 1) but also on the excited-state trans-cis isomerization rate, where the excited-state cis isoform is treated as tightly-coupled to the cis ground-state[17]. It is important to remind that dynamics of both types of distances need not be simulated independently. One may think of a more comprehensive model that describes changes in FRET and PIFE distances as correlated.

This way of thinking can be employed also to design an analytical framework to analyze multi-color multi-distance FRET measurements using MC-DEPI. Overall, the logical separation of the dynamics module from the photophysical module allows to use MC-DEPI as a versatile tool for analysis of a variety of different complex experiments with complex photophysical schemes and multiple reaction or conformational coordinates.

VI. SOFTWARE AND DATA AVAILABILITY

The the core MC-DEPI recoloring simulations are implemented in the open source `depi` python package (https://github.com/OpenSMFS/depi). The notebooks used for this paper are available on GitHub (https://github.com/tritemio/mcdepi2018-paper-analysis). The experimental data files are available on Figshare (https://doi.org/10.6084/m9.figshare.6931271).

VII. ACKNOWLEDGEMENTS

We would like to thank Dr. Xavier Michalet for fruitful discussions. This work was supported by the NIH under grant numbers GM069709 (to Shimon Weiss) and GM095904 (to Xavier Michalet); NSF under grant number MCB-1244175 (to Shimon Weiss).

REFERENCES

[1] T. Förster, Annalen der Physik **437**, 55 (1948).
[2] L. Stryer and R. P. Haugland, Proceedings of the National Academy of Sciences of the United States of America **58**, 719 (1967).
[3] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, and S. Weiss, Proceedings of the National Academy of Sciences **93**, 6264 (1996).
[4] E. Lerner, T. Cordes, A. Ingargiola, Y. Alhadid, S. Chung, X. Michalet, and S. Weiss, Science **359**, eaan1133 (2018).
[5] E. Lerner, A. Ingargiola, and S. Weiss, The Journal of Chemical Physics **148**, 123315 (2018).
[6] J. Ma, I. S. Yanez-Orozco, S. R. Adariani, D. Dolino, V. Jayaraman, and H. Sanabria, JoVE (Journal of Visualized Experiments) , e55623 (2017).
[7] M. Beckers, F. Drechsler, T. Eilert, J. Nagy, and J. Michaelis, Faraday Discussions **184**, 117 (2015).
[8] T.-O. Peulen, O. Opanasyuk, and C. A. Seidel, The Journal of Physical Chemistry B (2017), 10.1021/acs.jpcb.7b03441.
[9] J. J. McCann, L. Zheng, D. Rohrbeck, S. Felekyan, R. Kühnemuth, R. B. Sutton, C. A. M. Seidel, and M. E. Bowen, Proceedings of the National Academy of Sciences **109**, 15775 (2012).
[10] A. Sali, H. M. Berman, T. Schwede, J. Trewhella, G. Kleywegt, S. K. Burley, J. Markley, H. Nakamura, P. Adams, A. M. J. J. Bonvin, W. Chiu, M. D. Peraro, F. Di Maio, T. E. Ferrin, K. Grünewald, A. Gutmanas, R. Henderson, G. Hummer, K. Iwasaki, G. Johnson, C. L. Lawson, J. Meiler, M. A. Marti-Renom, G. T. Montelione, M. Nilges, R. Nussinov, A. Patwardhan, J. Rappsilber, R. J. Read, H. Saibil, G. F. Schröder, C. D. Schwieters, C. A. M. Seidel, D. Svergun, M. Topf, E. L. Ulrich, S. Velankar, and J. D. Westbrook, Structure **23**, 1156 (2015).
[11] B. Hellenkamp, S. Schmid, O. Doroshenko, O. Opanasyuk, R. Kühnemuth, S. R. Adariani, A. Barth, V. Birkedal, M. E. Bowen, H. Chen, T. Cordes, T. Eilert, C. Fijen, M. Götz,

G. Gouridis, E. Gratton, T. Ha, C. A. Hanke, A. Hartmann, J. Hendrix, L. L. Hildebrandt, J. Hohlbein, C. G. Hübner, E. Kallis, A. N. Kapanidis, J.-Y. Kim, G. Krainer, D. C. Lamb, N. K. Lee, E. A. Lemke, B. Levesque, M. Levitus, J. J. McCann, N. Naredi-Rainer, D. Nettels, T. Ngo, R. Qiu, C. Röcker, H. Sanabria, M. Schlierf, B. Schuler, H. Seidel, L. Streit, P. Tinnefeld, S. Tyagi, N. Vandenberk, K. R. Weninger, B. Wünsch, I. S. Yanez-Orozco, J. Michaelis, C. A. M. Seidel, T. D. Craggs, and T. Hugel, arXiv:1710.03807 [physics, q-bio] (2017), arXiv:1710.03807 [physics, q-bio].

[12]H. Höfig, M. Gabba, S. Poblete, D. Kempe, and J. Fitter, Molecules **19**, 19269 (2014).

[13]S. Sindbert, S. Kalinin, H. Nguyen, A. Kienzler, L. Clima, W. Bannwarth, B. Appel, S. Müller, and C. A. M. Seidel, Journal of the American Chemical Society **133**, 2463 (2011).

[14]M. H. Jacob, R. N. Dsouza, I. Ghosh, A. Norouzy, T. Schwarzlose, and W. M. Nau, Journal of Physical Chemistry B **117**, 185 (2012).

[15]A. Grupi and E. Haas, Journal of Molecular Biology **405**, 1267 (2011).

[16]W. J. A. Koopmans, A. Brehm, C. Logie, T. Schmidt, and J. van Noort, Journal of Fluorescence **17**, 785 (2007).

[17]E. Lerner, E. Ploetz, J. Hohlbein, T. Cordes, and S. Weiss, The Journal of Physical Chemistry B **120**, 6401 (2016).

[18]X. Kong, E. Nir, K. M. Hamadani, and S. Weiss, Journal of the American Chemical Society **129**, 4643 (2007).

[19]E. Haas and I. Z. Steinberg, Biophysical Journal **46**, 429 (1984).

[20]J. M. Beechem and E. Haas, Biophysical Journal **55**, 1225 (1989).

[21]S. Nag, B. Sarkar, M. Chandrakesan, R. Abhyanakar, D. Bhowmik, M. Kombrabail, S. Dandekar, E. Lerner, E. Haas, and S. Maiti, Physical Chemistry Chemical Physics **15**, 19129 (2013).

[22]I. V. Gopich and A. Szabo, in *Single-Molecule Biophysics: Experiment and Theory, Volume 146*, Advances in Chemical Physics, Vol. 146, edited by T. Komatsuzaki, M. Kawakami, S. Takahashi, H. Yang, and R. J. Silbey (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2011) pp. 245–297.

[23]N. K. Lee, A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright, and S. Weiss, Biophysical Journal **88**, 2939 (2005).

[24]B. K. Müller, E. Zaychikov, C. Bräuchle, and D. C. Lamb, Biophysical journal **89**, 3508 (2005).

[25]T. A. Laurence, X. Kong, M. Jäger, and S. Weiss, Proceedings of the National Academy of Sciences of the United States of America **102**, 17348 (2005).

[26]E. Sisamakis, A. Valeri, S. Kalinin, P. J. Rothwell, and C. A. M. Seidel, in *Methods in Enzymology*, Vol. 475 (Elsevier Inc., 2010) 1st ed., pp. 455–514.

[27]T. Torres and M. Levitus, The Journal of Physical Chemistry B **111**, 7392 (2007).

[28]H. Sahoo and P. Schwille, Chemphyschem **12**, 532 (2011).

[29]D. Nettels, A. Hoffmann, and B. Schuler, The Journal of Physical Chemistry B **112**, 6137 (2008).

[30]S. Kalinin, S. Felekyan, M. Antonik, and C. A. M. Seidel, The Journal of Physical Chemistry B **111**, 10253 (2007).

[31]Y. Santoso, J. P. Torella, and A. N. Kapanidis, ChemPhysChem **11**, 2209 (2010).

[32]J. P. Torella, S. J. Holden, Y. Santoso, J. Hohlbein, and A. N. Kapanidis, Biophysical journal **100**, 1568 (2011).

[33]N. C. Robb, T. Cordes, L. C. Hwang, K. Gryte, D. Duchi, T. D. Craggs, Y. Santoso, S. Weiss, R. H. Ebright, and A. N. Kapanidis, Journal of Molecular Biology **425**, 875 (2013).

[34]T. E. Tomov, R. Tsukanov, R. Masoud, M. Liber, N. Plavner, and E. Nir, Biophysical Journal **102**, 1163 (2012).

[35]I. V. Gopich and A. Szabo, The Journal of Physical Chemistry B **113**, 10965 (2009).

[36]M. Antonik, S. Felekyan, A. Gaiduk, and C. A. M. Seidel, The Journal of Physical Chemistry B **110**, 6970 (2006).

[37]S. Kalinin, A. Valeri, M. Antonik, S. Felekyan, and C. A. M. Seidel, The Journal of Physical Chemistry B **114**, 7983 (2010).

[38]D. T. Gillespie, Physical Review E **54**, 2084 (1996).

[39]X. Michalet, S. Weiss, and M. Jäger, Chemical Reviews **106**, 1785 (2006).

[40]I. V. Gopich and A. Szabo, in *Theory and Evaluation of Single-Molecule Signals* (WORLD SCIENTIFIC, 2008) pp. 181–244.

[41]M. Liao, *Applied Stochastic Processes* (Chapman and Hall/CRC, 2013).

[42]H. D. Kim, G. U. Nienhaus, T. Ha, J. W. Orr, J. R. Williamson, and S. Chu, Proceedings of the National Academy of Sciences **99**, 4284 (2002).

[43]T. A. Laurence, Y. Kwon, E. Yin, C. W. Hollars, J. A. Camarero, and D. Barsky, Biophysical Journal **92**, 2184 (2007).

[44]T. Head, MechCoder, G. Louppe, I. Shcherbatyi, fcharras, Z. Vinícius, cmmalone, C. Schröder, nel215, N. Campos, T. Young, S. Cereda, T. Fan, rene-rex, K. K. Shi, J. Schwabedal, carlosdanielcsantos, Hvass-Labs, M. Pak, SoManyUsernamesTaken, F. Callaway, L. Estève, L. Besson, M. Cherti, K. Pfannschmidt, F. Linzberger, C. Cauet, A. Gut, A. Mueller, and A. Fabisch, "Scikit-optimize/scikit-optimize: V0.5.2," (2018).

[45]T. A. Laurence and B. A. Chromy, Nature Methods **7**, 338 (2010).

[46]I. V. Gopich and A. Szabo, The Journal of Physical Chemistry B **114**, 15221 (2010).

[47]R. Tsukanov, T. E. Tomov, Y. Berger, M. Liber, and E. Nir, The Journal of Physical Chemistry B **117**, 16105 (2013).

[48]T. Uzawa, R. R. Cheng, K. J. Cash, D. E. Makarov, and K. W. Plaxco, Biophysical Journal **97**, 205 (2009).

[49]S. Kalinin, T. Peulen, S. Sindbert, P. J. Rothwell, S. Berger, T. Restle, R. S. Goody, H. Gohlke, and C. A. M. Seidel, Nature Methods **9**, 1218 (2012).

[50]J. R. Unruh, G. Gokulrangan, G. H. Lushington, C. K. Johnson, and G. S. Wilson, Biophysical Journal **88**, 3455 (2005).

[51]M. E. Sanborn, B. K. Connolly, K. Gurunathan, and M. Levitus, The Journal of Physical Chemistry B **111**, 11064 (2007).

[52]N. Di Fiori and A. Meller, Biophysical journal **98**, 2265 (2010).

[53]V. Kudryavtsev, M. Sikor, S. Kalinin, D. Mokranjac, C. A. M. Seidel, and D. C. Lamb, ChemPhysChem **13**, 1060 (2012).

[54]J. Ouellet, S. Schorr, A. Iqbal, T. J. Wilson, and D. M. J. Lilley, Biophysical Journal **101**, 1148 (2011).

[55]O. Kroutil, I. Romancová, M. Šíp, and Z. Chval, The Journal of Physical Chemistry B **118**, 13564 (2014).

[56]D. Nettels, I. V. Gopich, A. Hoffmann, and B. Schuler, Proceedings of the National Academy of Sciences **104**, 2655 (2007).

[57]K. Kinosita, A. Ikegami, and S. Kawato, Biophysical Journal **37**, 461 (1982).

[58]M. Hoefling and H. Grubmüller, Computer Physics Communications **184**, 841 (2013).

[59]M. Hoefling, N. Lima, D. Haenni, C. A. M. Seidel, B. Schuler, and H. Grubmüller, PLOS ONE **6**, e19791 (2011).

[60]E. Ploetz, E. Lerner, F. Husada, M. Roelfs, S. Chung, J. Hohlbein, S. Weiss, and T. Cordes, Scientific Reports **6**, 33257 (2016).

[61]A. Ingargiola, E. Lerner, S. Chung, F. Panzeri, A. Gulinatti, I. Rech, M. Ghioni, S. Weiss, and X. Michalet, PLOS ONE **12**, e0175766 (2017).

[62]F. Panzeri, A. Ingargiola, R. R. Lin, N. Sarkhosh, A. Gulinatti, I. Rech, M. Ghioni, S. Cova, S. Weiss, and X. Michalet, in *Proceedings of SPIE*, Vol. 8590 (San Francisco, CA, USA, 2013) p. 85900D.

[63]A. Ingargiola, T. Laurence, R. Boutelle, S. Weiss, and X. Michalet, Biophysical Journal **110**, 26 (2016).

Appendix A: Materials and Methods

Nanosecond alternating laser excitation (nsALEX) smFRET measurements of two 40 bp dsDNA samples with two different inter-dye were carried out. In these samples,

the acceptor dye, ATTO 647N (ATTO-TEC, GmbH) labeled the 5'-end and the donor dye, ATTO 550 (ATTO-TEC, GmbH) internally labeled bases in the complementary strand separated by either 7 or 17 bp from the base to which the acceptor was attached (d7 and d17 samples, respectively; additional information in Ingargiola et al.[61]).

nsALEX smFRET measurements allowed us to acquire the ID, detection macrotime and nanotime for each detected photon. nsALEX measurements also allowed us to separate the sub-population of bursts that did not have a fluorescently-active acceptor (donor-only species) from ones that had fluorescence from both the donor and the acceptor (FRET species). The laser alternation period in nsALEX (the inverse of the lasers' repetition rate, 50 ns) allowed recording the fluorescence decays after acceptor excitation and after donor excitation, separately. Additionally, since the alternation period was in nanoseconds, fluorescence correlation curves starting at 1 µsdid not include the contribution from the alternation. nsALEX smFRET measurements were performed on a home-built single-molecule fluorescence setup described elsewhere[62], with the following differences: (i) the laser used for donor excitation was a pulsed picosecond diode laser (LDH-P-FA-530L, PicoQuant GmbH, Berlin, GmbH) with an output wavelength of 532 nm; (ii) the repetition rate of both pulsed picosecond diode lasers, used for alternating excitation of the donor and the acceptor dyes, were triggered by two separate picosecond diode laser drivers (PDL 800-B, PicoQuant, GmbH, Berlin, GmbH) at a repetition rate of 20 MHz; (iii) the two laser pulses were electronically interleaved by 22 ns; (iv) the signal from the pulsed diode laser used for acceptor excitation (output wavelength of 635 nm; LDH-P-635, PicoQuant, Berlin, GmbH) was fed as SYNC for the TCSPC card; (iv) the single photon avalanche photodiodes (SPADs) used for both donor and acceptor fluorescence detection were both red-enhanced SPADs with a small detection area (50 μm)[62]. Aside from these differences, the fluorescence emissions of the donor and the acceptor were separated using a dichroic long pass filter with the dividing edge at 624 nm (FF624-Di01, Semrock, Rochester, NY, USA). Emission that passed through the donor channel was selected with a 580/60 nm band-pass filter (FF01-580/60-25-d, Semrock, Rochester, NY, USA); the emission passed to the acceptor channel was selected by a 690/50 nm band-pass filter (ET690/50m, Chroma, Bellows Falls, Vermont, USA). The measurements presented were performed at an approximate depth of 50 µm inside the sample with a total acquisition time of several an hour per measured sample. The photon IDs, macrotimes and nanotimes were saved in the Becker and Hickl acquisition file, which was them converted into the general Photon-HDF5 file format[63].

Data was analyzed as in Lerner et al.[5]. Briefly, first the background (BG) rate was calculated for each photon stream (donor excitation donor emission, DexDem; donor excitation acceptor emission, DexAem, acceptor excitation acceptor emission, AexAem) for each 60 s of the measurement. Then single-molecule bursts were identified by as groups of consecutive photons where each subgroup of m (=10) consecutive photons had a photon rate larger than F (=6) times the overall BG rate. Then, the ratiometric proximity ratio and stoichiometries were calculated for each burst. In the next step, single-molecule bursts were selected according to two sets of criteria: (i) for donor-only species: bursts that had overall more than 20 photons from all streams, with stoichiometry values larger than 0.9 and proximity ratio values smaller than 0.1; (ii). for FRET species: bursts with more than 30 photons arising from donor excitation and more than 30 photons arising from acceptor excitation.

The photon IDs of the bursts were used for calculating the ratiometric FRET efficiency values of bursts and to assemble FRET histograms . The nanotimes were histogrammed to produce fluorescence decays for different photon streams (DexDem, DexAem, AexAem) and molecular species (donor-only and FRET). The mean of all nanotime of each photon stream in a burst were used to construct the mean nanotime histograms.

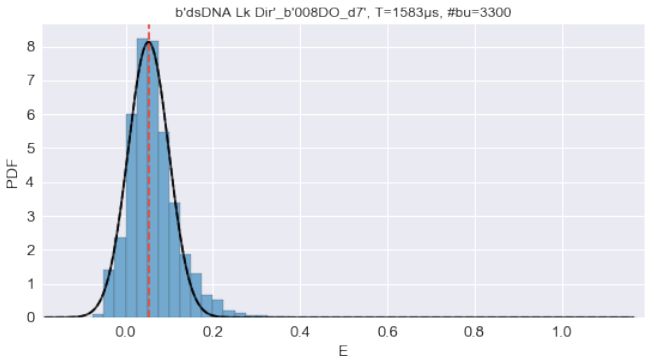

Figure 11. Assessment of the donor leakage probability, Lk, via alternating laser excitation (ALEX) measurement of a donor-only dsDNA molecule, having just the ATTO 550 dye attached intenrally to to the DNA base in the d7 sample. The peak value of the FRET histogram is larger than zero, to indicate the fraction of photons detected in the acceptor channel that originate from donor emission that leaked into the acceptor detection channel. The peak FRET value is used in the calculation of the Lk fraction.

The un-corrected experimental histograms were compared with the same histograms constructed from the recoloring simulations. These simulations represent ideal results without the effects of donor fluorescence leakage into the acceptor detection channel (Lk) and the direct excitation of the acceptor by the laser intended for donor excitation (d_T). Therefore the simulated histograms were un-corrected for these factors so that they can be compared with the un-corrected experimental histograms (see Eq. 8). The values of the correction factors acquired by Ingargiola et al.[61]) for the same samples were Lk=0.1 and d_T=0.06 for both d7 and d17 samples. ALEX measurements of donor-only d7 and d17 samples gave rise to peak FRET values of 0.052 and 0.058, corresponding to

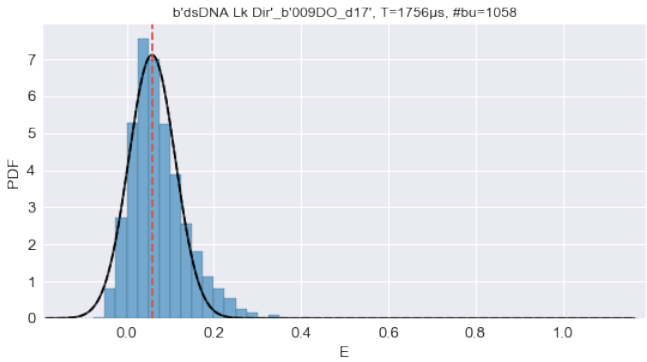

Figure 12. Assessment of the donor leakage probability, Lk, via alternating laser excitation (ALEX) measurement of a donor-only dsDNA molecule, having just the ATTO 550 dye attached internally to to the DNA base in the d17 sample. The peak value of the FRET histogram is larger than zero, to indicate the fraction of photons detected in the acceptor channel that originate from donor emission that leaked into the acceptor detection channel. The peak FRET value is used in the calculation of the Lk fraction.

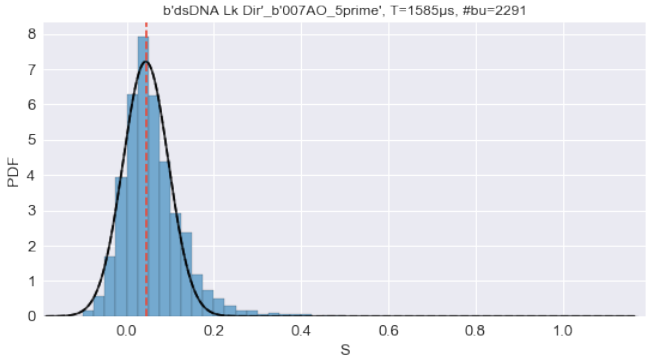

Figure 13. Assessment of the direct acceptor excitation probability, d_T, via alternating laser excitation (ALEX) measurement of an acceptor-only dsDNA molecule, having just the ATTO 647N dye attached to its 5'-end. The peak value of the stoichiometry histogram is larger than zero, to indicate the fraction of acceptor excitation directly by the excitation source intended for donor excitation. The peak stoichiometry value is used in the calculation of the d_T fraction.

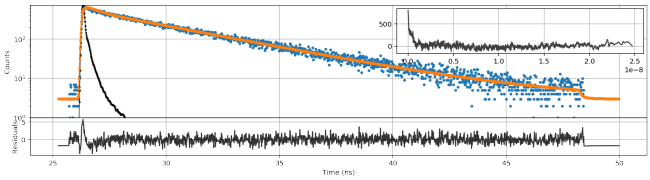

Figure 14. The best fit results of a bi-exponential function to the donor fluorescence decay of the donor-only species in the nsALEX measurement of the d7 dsDNA sample.

Lk values of 0.055 and 0.062, respectively (Figs. 11 and 12. In summary these were the values of the Lk factor and not 0.1 as reported previously for these samples[61].

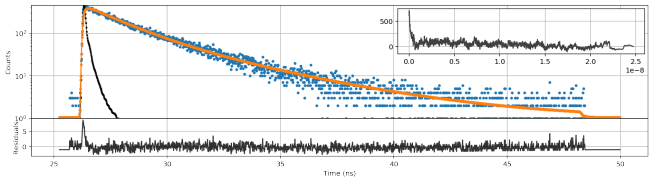

Figure 15. The best fit results of a bi-exponential function to the donor fluorescence decay of the donor-only species in the nsALEX measurement of the d17 dsDNA sample.

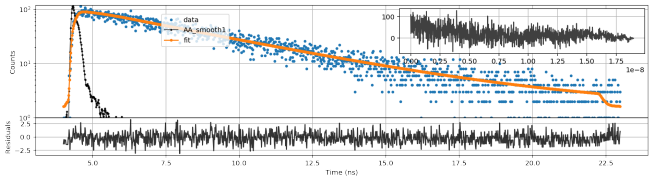

Figure 16. The best fit results of an exponential function to the acceptor fluorescence decay after excitation of the acceptor in the nsALEX measurement for the d7 dsDNA sample.

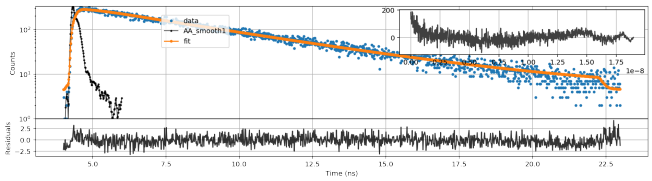

Figure 17. The best fit results of an exponential function to the acceptor fluorescence decay after excitation of the acceptor in the nsALEX measurement for the d17 dsDNA sample.

As for the assessment of d_T values, we performed ALEX measurements of an acceptor-only sample. The value of the peak stoichiometry, S_{peak}, was 0.044 (Fig. 13). The value of d_T equals $S_{peak}/1 - S_{peak}$ multiplied by the β factor elucidated from analysis of ALEX measurements introduced by Lee et al.[23]. The results of this analysis for ALEX measurements of d7 and d17 samples yielded β=1.32 and γ=0.81. Using this value of β, d_T=0.060 exactly as reported previously for this sample[61]. Regarding un-correcting the simulated histograms for the γ factor, the approach by Lee et al.[23] assumes that the ratio of the acceptor and donor fluorescence quantum yields in both samples is the same. The ratio of fluorescence quantum yields is the same as the ratio of fluorescence lifetimes. The donor fluorescence decays of the donor-only bursts were best fit with a bi-exponential function with 3.92 and 0.45 ns lifetime components with 0.93 and 0.07 fractions, respectively, for both the d7 and d17 samples (Figs. 14 and 15). The donor amplitude-weighted average fluorescence lifetimes of the d7 and d17 samples were 3.69 and 3.68 ns. The acceptor-only fluorescence decays for the d7 and d17 samples in the nsALEX measurement were best fit with an exponential function with lifetimes of 1.10 and 1.09 ns, respectively (Figs. 16 and 17). The ratio of the acceptor lifetime and the mean donor life-

time for both d7 and d17 samples were similar (1.10 and 1.09, respectively). This means that the assumption of the same γ factor in both samples is justified. Another way to retrieve an estimation of the γ factor is by using Eq. A1:

$$\gamma = \frac{\frac{1-\langle E\rangle}{\langle E\rangle}}{\frac{1-\langle E^*\rangle}{\langle E^*\rangle}} \tag{A1}$$

where $\langle E^*\rangle$ is the peak FRET efficiency of a FRET subpopulation after correction for Lk and d_T and $\langle E\rangle$ is the value of the mean FRET efficiency extracted using Eq. A2.

$$\langle E\rangle = 1 - \frac{\langle \tau_{D,FRET}\rangle}{\langle \tau_D\rangle} \tag{A2}$$

and the mean lifetimes are extracted from a fit of sum of exponentials to the donor fluorescence decays in the presence of FRET and in the absence of it (donor-only species). The d17 sample was fitted with a sum of three exponents function, after fixing two exponents to be equal those of the donor-only values, to account for possible donor-only fraction coming from acceptor triplet blinking (Fig. 18). The donor amplitude-weighted fluorescence lifetime of sample d17 was 2.49 ns, which together with the average lifetime of the donor-only species gave rise to $\langle E\rangle$=0.32. This value, together with $\langle E^*\rangle$=0.27 gave rise to γ=0.78 (using Eq. A1). Performing this procedure for the d7 sample was very hard due to limited amount of donor photons (high FRET). This value was very close to the value found using the approach by Lee et. al. Therefore the values of the correction factors are Lk=0.06, d_T=0.06 and γ=0.78.

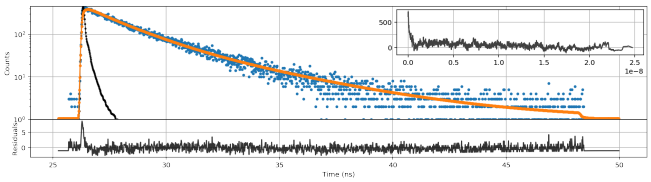

Figure 18. The best fit results of a sum of three exponents function to the donor fluorescence decay of the FRET species in the nsALEX measurement of the d17 dsDNA sample.

To facilitate the distance-to-FRET conversions, one needs to know the value of R_0. The value documented for the ATTO 550 donor and ATTO 647N acceptor is 65 Å. We performed a set of steady-state fluorescence measurements to determine the value of R_0 for these dyes when attached to the dsDNA molecules, using Eq. 2. First we assessed the donor fluorescence quantum yields (QYs) of the donor by comparing the area under the fluorescence spectra of the the donor in the d7 and d17 samples and compared them to the area under the fluorescence spectrum of a reference dye, Cy3B, with a documented QY

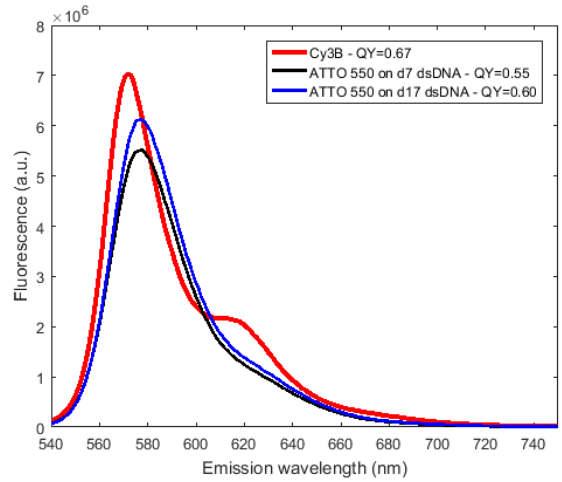

Figure 19. The differences in the fluorescence quantum yield (QY) of the ATTO 550 donor. Donor fluorescence spectra of ATTO 550 on top of dsDNA (in d7 and d17 samples in black and blue, respectively) were recorded and compared to the spectrum of free Cy3B as a reference with a known QY value of 0.67. Comparing the areas below the spectra yields QY values of 0.55 and 0.60 for the donor in the d7 and d17 samples, respectively

.

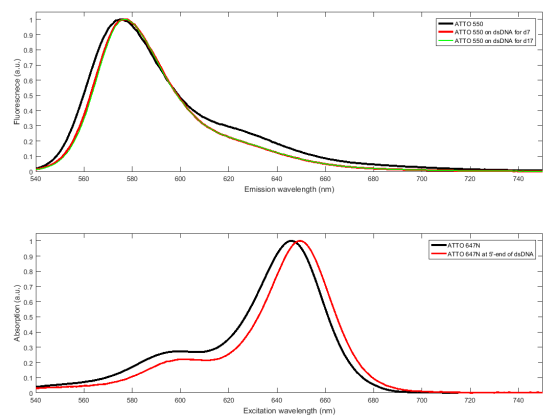

Figure 20. The differences in the overlap integral of ATTO 550 and ATTO 647N donor-acceptor FRET pair. Top - donor fluorescence spectra of ATTO 550 on top of dsDNA (in d7 and d17 samples in red and green, respectively) are blue-shifted relative to the spectrum of the free dye (black). Bottom - acceptor extinction spectra of ATTO 647N on top of dsDNA (in d7 and d17 samples in red and green, respectively) are red-shifted relative to the spectrum of the free dye (black).

of 0.67 (Fig. 19)[5]. These spectra were recorded after excitation at λ=532 nm, which is the wavelength of donor excitation. Additionally, these spectra were divided by the optical density of that sample at this wavelength. The procedure yielded QY values of 0.55 and 0.60 for the d7 and d17 samples, respectively. These values are

smaller than the value of the QY for ATTO 550 free in solution (0.80). Next we recorded the donor fluorescence spectra and the acceptor absorption spectrum (Fig. 20; in both d7 and d17 samples, the acceptor labeled the same position, the 5'-end of the dsDNA, and the dsDNA in both samples had the same sequence[61]). The donor fluorescence spectra showed a blue shift relative to the spectrum of ATTO 550 free in solution. Additionally, the acceptor extinction spectrum showed a red shift relative to the spectrum of ATTO 647N free in solution. In summary, the values of the FRET overlap integral of the d7 and d17 samples were 2.87 and 3.06 $10^{15} M^{-1} cm^{-1} nm^4$. Additionally, the refractive index, n, was measured by an Abbe refractometer and was 1.334. Taking all of these results and assuming the value of $\langle \kappa^2 \rangle$ was 2/3 yielded R_0 values of 55.5 and 56.5 Å for the d7 and d17 samples, respectively Repeating the steady-state measurements and re-calculating the values of R_0 yielded similar values within a deviation of 0.5 Å. These values deviate by 1 nm from the values of R_0 that was calculated using the steady-state spectra of the dyes free in solution.

All simulated fluorescence decays were convoluted with the experimentally-acquired IRFs in the following manner: (i) a probability density function of the IRF was produced (it was normalized so that the area underneath it equaled 1); (ii) the simulated nanotime after IRF convolution was calculated as the addition of a random number (sampled from the probability density function of the IRF calculated in step i) to the simulated nanotime before IRF convolution; (iii) the mean time of the IRF was subtracted from all IRF-convoluted nanotimes acquired in step ii.

1. The loss function

A MC-DEPI simulation maintain the experimental timestamps while simulating each photon's IDs and nanotimes. Thus, it allows to generate FRET histograms, fluorescence decays or any other observable computed from smFRET data (BVA, FCS, etc.). In order to compare simulations with experiments we build a loss function, expressing the distance between a specific simulation and the experimental data. In this work we focus using FRET histograms and fluorescence decays as basis for the comparison.

The loss function $\mathcal{L}(\theta; \mathbf{D})$ is function of vector θ input parameters and of the experimental data $\mathbf{D}$. We split the parameters in fixed θ_c and varying θ_v. The former are estimated before the simulation and include the intrinsic D and A fluorescence lifetimes, correction factors, background level. The varying parameters include all the unknown physical parameters that we want to estimate from the data. Choosing which parameter is fixed and which is varying depends on the amount of available information for each specific sample.

The loss function $\mathcal{L}(\theta; \mathbf{D})$ will contain many components, one for each observable to be compared. In this work we have two components: FRET histograms $\mathcal{L}_E(\theta; \mathbf{D})$ and the fluorescence decays $\mathcal{L}_{FL}(\theta; \mathbf{D})$. For FRET component, we used the sum of squared residuals between simulated and experimental FRET histogram (Eq. A3).

$$\mathcal{L}_E(\theta; \mathbf{D}) = \sum_i (E_i - E_i)^2 \tag{A3}$$

For the loss function of each fluorescence decay we used the negative log-likelihood function of the decay histograms assuming Poisson statistics in each bin.

$$\mathcal{L}_{FL,X}(\theta; \mathbf{D}) = \sum_i \lambda_i - k_i \log \lambda_i \tag{A4}$$

where λ_i and k_i are the simulated and experimental fluorescence decay values in TCSPC bin i and X is either D or A. The derivation of Eq. A4 is reported below.

The mass-function of a Poisson distribution is:

$$p(k|\lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{A5}$$

Each bin of a TCSPC histogram is a Poisson variable. Thus, here, the Poisson rate λ is the simulated value and k the experimental value of the bin. Given an estimated rate λ, the likelihood of one bin with counts k is $\mathcal{L}(k|\lambda) = p(k|\lambda)$. The log-likelihood is:

$$\log p(k|\lambda) = k \log \lambda - \lambda - \log(k!) \tag{A6}$$

Considering n TCSPC bins, the likelihood function is product of the likelihood in each bin:

$$\mathcal{L}(k|\lambda) = \prod_i p(k_i|\lambda_i) \tag{A7}$$

and the log-likelihood is:

$$\log \mathcal{L}(k|\lambda) = \sum_i \log p(k_i|\lambda_i) = \sum_i k_i \log \lambda_i - \lambda_i - \log(k_i!) \tag{A8}$$

The term $\{\lambda_i\}$ represents the fluorescence decay values in each bin resulting from an MC-DEPI simulation. The term $\log(k_i!)$, instead, depends only the experimental fluorescence decays and is constant when varying θ.

The loss function for each fluorescence decay is chosen as the negative log-likelihood function, removing terms not varying with θ (Eq. A9).

The loss function $\mathcal{L}_{FL}$ is computed from the loss function for D and A decays as follows:

$$\mathcal{L}_{FL} = \sigma_{FL,D}^{-1} \mathcal{L}_{FL,D} + \sigma_{FL,A}^{-1} \mathcal{L}_{FL,A} \tag{A9}$$

where $\sigma_{FL,D}$ and $\sigma_{FL,A}$ are the standard deviation of $\mathcal{L}_{FL,D}$ and $\mathcal{L}_{FL,A}$ when keeping both data and MC-DEPI parameters constant (Monte Carlo noise).

Similarly, the total loss function $\mathcal{L}(\theta; \mathbf{D})$ is computed from the two components as follows:

$$\mathcal{L} = \sigma_{FL}^{-1}\, \mathcal{L}_{FL} + \sigma_{E}^{-1}\, \mathcal{L}_{E} \tag{A10}$$