Characterizing Multiple Molecular States in Single-Molecule Multiparameter Fluorescence Detection by Probability Distribution Analysis

Stanislav Kalinin,* Suren Felekyan, Alessandro Valeri, and Claus A. M. Seidel*

Institut für Physikalische Chemie, Lehrstuhl für Molekulare Physikalische Chemie, Heinrich-Heine-Universität, Universitätsstraβe 1, Geb 26.32, 40225 Düsseldorf, Germany

Received: December 20, 2007

Probability distribution analysis (PDA) [M. Antonik et al., *J. Phys. Chem. B* **2006**, *110*, 6970] allows one to quantitatively analyze single-molecule (SM) data obtained in Förster resonance energy transfer (FRET) or fluorescence polarization experiments. By taking explicitly background and shot noise contributions into account, PDA accurately predicts the shape of one-dimensional histograms of various parameters, such as FRET efficiency or fluorescence anisotropy. In order to describe complex experimental SM-FRET or polarization data obtained for systems consisting of multiple non-interconverting fluorescent states, several extensions to the PDA theory are presented. Effects of brightness variations and multiple-molecule events are considered independently of the detection volume parameters by using only the overall experimental signal intensity distribution. The extended PDA theory can now be applied to analyze any mixture, by using any a priori model or a model-free deconvolution approach based on the maximum entropy method (MEM). The accuracy of the analysis and the number of free parameters are limited only by data quality. Correction of the PDA model function for the presence of multiple-molecule events allows one to measure at high SM concentrations to avoid artifacts due to a very long measurement time. Tools such as MEM and combined mean donor fluorescence lifetime analysis have been developed to distinguish whether extra broadening of PDA histograms could be attributed to structural heterogeneities or dye artifacts. In this way, an ultimate resolution in FRET experiments in the range of a few Ångström is achieved which allows for molecular Ångström optics distinguishing between a set of fixed distances and a distribution of distances. The extended theory is verified by analyzing simulations and experimental data.

1. Introduction

Data obtained in single-molecule Förster resonance energy transfer (SM-FRET) measurements are usually visualized as 1D histograms of the FRET efficiency or other related parameters.[1–13] These histograms are often fitted with a weighted sum of Gaussian distributions, for which the mean values are calculated. Under certain conditions, this approach to the analysis of SM-FRET data may provide correct mean FRET efficiencies,[14] which can be further used to calculate donor−acceptor (DA) distances.[15] However, experimental histograms contain a lot more information than just these mean values and allow one to reveal heterogeneities often found in biological systems. This information can be extracted by analyzing the width and the shape of FRET efficiency distributions, which is usually ignored in standard analysis.

The major challenge for quantitative analysis of SM data histograms is to separate shot-noise broadening from actual distributions of physical parameters. Even early attempts to predict the effect of shot-noise broadening[8,16] reveal, for instance, structural dynamics of syntaxin 1.[8] Recently, several quantitative methods have been developed, which extract useful information by analyzing photon statistics in SM experiments.[6,14,17–19] The probability distribution analysis (PDA) method[6] is intended mainly for the analysis of experimental distributions obtained from diffusing molecules. In PDA, both mean and width of a shot-noise broadened distribution are the functions of one parameter, for example, the mean FRET efficiency. PDA automatically accounts for shot-noise contribution and can unambiguously indicate whether any observed distribution is due to shot noise only or whether there must be a real distribution of the parameter of interest. However, in its present form, PDA has only a very limited potential to extract meaningful parameters of the individual states. This work shows how to handle various possible complications, which are typical for systems consisting of multiple fluorescent states. We also demonstrate how one can model underlying distributions of the parameters of interest in a rather general case. The ability to resolve multiple states is the main advantage of SM experiments, and thus, such an extension to PDA is clearly needed in order to be able to characterize individual states in a quantitative manner.

Intuitively, it might seem that in terms of the theory of PDA, such a modification is trivial. One may ask why the overall histogram obtained for a mixture of several species should not be just a sum of histograms calculated by using the reported PDA theory[6] and weighted by the fractions of the species. The main problem is that the theory of PDA makes use of the fact that photon distribution between color or polarization detection channels and the total signal intensity distribution, $P(S)$, are independent. The key idea of PDA is that the latter distribution is measurable and can be used as is, without much loss of relevant information (e.g., on FRET efficiency distributions). One of the important advantages of using the experimental $P(S)$ is the possibility to work with arbitrary long time windows (comparable with the diffusion time), which obviously contain large numbers of photons. This feature of PDA has proved to

* Corresponding authors. E-mail: stanislav.kalinin@uni-duesseldorf.de and cseidel@gwdg.de.

10.1021/jp711942q CCC: $40.75
Published on Web 06/21/2008

be especially helpful in resolving heterogeneous populations.[20] However, if the fluorescence intensity distribution is not the same for all species, the brightness and the photon distribution between color or polarization detection channels become correlated. In this work, we present an example when the use of the same (overall) fluorescence intensity distribution for all species produces completely erroneous results.

In contrast to PDA, fluorescence intensity distribution analysis (FIDA)[19,21] predicts the signal intensity distribution by modeling the shape of the detection volume. Therefore, the fluorescence intensity distribution for each species, at SM concentration, is obtained as an intermediate result, and the above-mentioned problem does not arise. The situation is similar for the rigorous treatment of photon statistics in SM fluorescence experiments, developed by Gopich and Szabo.[14] However, this theory requires the knowledge of the spatial intensity distribution, which in reality can strongly deviate from 3D Gaussian and depends on many experimental parameters that are difficult to control.[22–24] If we decide to stay only with the experimental signal intensity distribution, individual contributions of all species have to be extracted from that distribution.

This procedure is trivial only if the total brightness (i.e., signal sum of all detectors) of all species is the same or the individual fluorescence intensity distributions can be somehow measured independently. Equal brightness of all states was implicitly assumed when a Gaussian distribution of DA distances was fitted to experimental data obtained from a labeled DNA sample.[6] However, in reality, this is usually not the case because of different quantum yields of the dyes and unequal detection efficiencies of the donor and the acceptor detection channels. This problem has been pointed out by Nir et al. in their work on proximity ratio histogram (PRH) analysis,[18] which applies to bursts of varying duration. PRH analysis is based on principles similar as those of PDA and makes use of the experimental burst size distribution. Nir et al. suggested that the desired detection efficiencies can be "artificially achieved by misaligning one of the detectors or by discarding a constant fraction of the detected photons by one of the channels".[18] Although this idea may work well for pure FRET, it certainly does not cover all possible scenarios such as local dye quenching or the presence of several states, differing not only by the DA distance. Moreover, discarding photons leads to a loss of useful data and therefore increases the contribution of shot noise. In this work, we propose a general extension to PDA which takes into account brightness variations and recovers correct underlying values of FRET efficiencies or fluorescence anisotropies, as well as fractions of the species.

Another common problem of SM measurements on complex systems is a simultaneous observation of several molecules in the laser focus.[25,26] These multiple-molecule events appear in 1D histograms as a small mixed population between major peaks and can be easily misinterpreted as an individual population or as an indication of conformational dynamics. In this work, we propose an approach to correct the PDA model function for the presence of multiple-molecule events, which under certain conditions requires no additional model parameters. Exactly for the same reason as that for brightness correction, treatment of multimolecule events in PDA is not straightforward. In addition to what is done in FIDA, for example, one has to extract several individual fluorescence intensity distributions, which correspond to SM concentration, starting only from the experimental signal intensity distribution obtained for a concentrated mixture of several species. Correction for multiple-molecule events may help reduce the measurement time as well as the influence of laser instabilities and adsorption on the glass surface. It can also be applied when high concentrations cannot be avoided, for instance, when molecular complexes with moderate affinity are investigated.

The theory of PDA extended as discussed above allows one to test any underlying distribution against experimental data, which include shot-noise and background contributions. Complex underlying distributions are usually modeled as a number of discrete states or continuous distributions. An alternative approach to the modeling of these distributions is model-free deconvolution based on the maximum entropy method (MEM).[27–29] We demonstrate how MEM can be combined with PDA in order to extract unbiased underlying distributions from measurements on diffusing molecules. As applied to SM-FRET data, the combination of PDA and MEM naturally recovers distributions of DA distances. The proposed procedure is to a large extent similar to the approach introduced by Watkins et al.[17] for the analysis of photon traces of immobilized molecules. Applications of MEM to simulated and experimental data demonstrate its advantages and limitations and also reveal some general restrictions of the PDA method.

Recently, in many papers,[6,9,12,18] possible reasons for broadening of FRET efficiency distributions have been discussed, such as (1) optical misalignment of the green and red detection volumes, (2) photobleaching or blinking of the acceptor dye, (3) saturation of the acceptor dye, and (4) local quenching of the donor and/or acceptor dye, which is a very frequent cause. Here, we present a simple analysis tool in combination with 2D FRET analysis, which describes the correlation between mean donor lifetime distributions and FRET efficiencies. The use of these tools allows one to judge with high confidence whether the experimental broadening is caused by dye quenching.

2. Theory

2.1. Basic Theory of PDA. ***2.1.1. Single Species.*** The measured signal S consists of fluorescence (F) and background (B) photons. In this work, the values of S, F, and B are expressed in photon counts per time window of a fixed length (Δt). The signal is measured by two or more single-photon counting detectors and divided into two components (for example, green (G) and red (R) or parallel (||) and perpendicular ($\perp$)). The PDA method[6,20] starts from the calculation of the probability of observing a certain combination of photon counts in two detection channels 1 and 2, $P(S_1, S_2)$. For a single fluorescent state $P(S_1, S_2)$ is given by a product of independent probabilities.[6]

$$P(S_1, S_2) = \sum_{F_1+B_1=S_1;F_2+B_2=S_2} P(F)\, P(F_1, F_2|F) P(B_1)\, P(B_2) \tag{1}$$

In eq 1, $P(F)$ describes the fluorescence intensity distribution (i.e., the likelihood to observe F fluorescence photons in a time interval Δt). $P(B_1)$ and $P(B_2)$ represent the background intensity distributions, which are usually assumed to obey a Poisson distribution[25,30] with known mean intensities $\langle B_1 \rangle$ and $\langle B_2 \rangle$. $P(F_1,F_2|F)$ stands for the conditional probability of observing a particular combination of F_1 and F_2, provided that the total number of registered fluorescence photons is F. $P(F_1,F_2|F)$ can be expressed as a binomial distribution.[6,18]

$$P(F_1, F_2|F) = \frac{F!}{F_1!\,(F-F_1)!} {p_1}^{F_1} (1-p_1)^{F-F_1} \tag{2}$$

In eq 2, p_1 stands for the probability of a detected photon to be registered by the first detector (e.g., green in a FRET experiment

or parallel in an anisotropy experiment). The knowledge of $P(S_1, S_2)$ is sufficient to generate 1D histograms of any parameter, which can be expressed as a function of S_1 and S_2 (e.g., signal ratio S_1/S_2, FRET efficiency or fluorescence anisotropy).[20] Fitting of such histograms obtained for a single species requires only one floating parameter, p_1. In FRET experiments, the value of p_1 is unambiguously related to the FRET efficiency E according to eq 3.

$$p_1 = \left(1 + \alpha + \frac{E\Phi_{FA}}{(1-E)G\Phi_{FD}}\right)^{-1} \quad (3)$$

In eq 3, G stands for the ratio of the detection efficiencies, g_G and g_R, of the two detection channels ($G = g_G/g_R$), Φ_{FD} and Φ_{FA} are the fluorescence quantum yields of the donor and the acceptor, respectively, and α is the crosstalk from green donor signal into the red detection channel of acceptor.

The distribution $P(F)$ in eq 1 is not directly measurable. One can overcome this problem by using a sufficiently high threshold $S_{min} \gg B$ (where $B = B_1 + B_2$) and approximate $P(F)$ by $P(S)$, which is directly obtained from the measurement (eq 4).[6]

$$P(S_1, S_2) = \sum_{F_1+B_1=S_1;F_2+B_2=S_2} P(S)\, P(F_1, F_2|S - B_1 - B_2) P(B_1)\, P(B_2) \quad (4)$$

The underlying distribution $P(F)$ can also be easily calculated by deconvolution from the total signal intensity distribution $P(S)$, which is given by

$$P(S) = P(F) \otimes P(B) \quad (5)$$

Details of the deconvolution procedure are described in ref 20. Most of the theoretical considerations presented in this work depend on the knowledge of the distribution $P(F)$; therefore, the deconvolution approach is preferred even when $S_{min} \gg B$ is fulfilled.

2.1.2. Multiple Species and Brightness Correction. Factorization in eq 1 is valid only when all four distributions are independent. In particular, it is assumed that the distributions $P(F)$ and $P(F_1,F_2|F)$ are uncorrelated. This is true only if a single fluorescent state is considered, or if the total brightness of all states is the same. In general, each state can be characterized by an individual FRET efficiency E_j or polarization anisotropy r_j, as well as by the brightness Q_j. Therefore, a general extension of eq 1 to the case of k different states requires the knowledge of individual fluorescence intensity distributions $P_j(F)$ ($j = 1,..., k$). Let us first assume that these distributions refer to the same concentration of the species and consider only SM events. The observed distribution $P(S_1, S_2)$ for $F > 0$ is then given by a sum of $P_j(S_1, S_2)$ terms, weighted by the concentrations c_j ($j = 1,..., k$):

$$P(S_1, S_2) = \sum_{j=1}^{k} \left[c_j \sum_{F_1+B_1=S_1;F_2+B_2=S_2} P_j(F)\, P_j(F_1, F_2|F)\, P(B_1)\, P(B_2) \right] = \sum_{F_1+B_1=S_1;F_2+B_2=S_2} \left(\sum_{j=1}^{k} [c_j P_j(F)\, P_j(F_1, F_2|F)]\, P(B_1)\, P(B_2) \right) \quad (6)$$

It is convenient to consider the point $F = 0$ and the convolution with background separately and to rewrite eq 6 as

$$P(F_1 = 0, F_2 = 0) = P(F = 0) \quad (7a)$$

$$P(F_1, F_2) = \sum_j [c_j P_j(F)\, P_j(F_1, F_2|F)] \qquad (F > 0) \quad (7b)$$

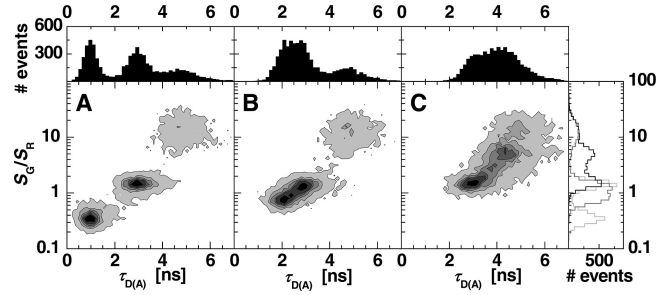

Figure 1. Three typical cases observed in FRET experiments on mixtures of two FRET-active species and one donor-only labeled species. Results of burstwise analysis of simulated data are presented in 2D histograms of G-to-R signal ratio, S_G/S_R, versus the donor fluorescence lifetime in the presence of the acceptor, $\tau_{D(A)}$. (A) All three species can be completely separated on the burst level or can be studied independently in a series of experiments. (B) The donor-only population can be separated, or an additional measurement on pure donor-only sample can be performed, but the two FRET species cannot be separated. (C) All three populations can be neither separated nor independently studied.

$$P(S_1, S_2) = \sum_{F_1+B_1=S_1;F_2+B_2=S_2} P(F_1, F_2)\, P(B_1)\, P(B_2) \quad (7c)$$

In eqs 6 and 7, $P_j(F_1,F_2|F)$ describes the probability of observing a particular combination of F_1 and F_2, for the jth species, provided that the total number of fluorescence photons is F, and $P(F)$ denotes the overall fluorescence intensity distribution.

Unfortunately, it is a rule rather than an exception that the brightness of different species is not the same. For instance, in FRET experiments the total brightness is determined by the fluorescence quantum yields of the donor and the acceptor and the corresponding detection efficiencies. Therefore, all fluorescence intensity distributions $\{P_j(F)\}$ have to be calculated or measured independently (cf. eqs 6 and 7). In the next sections, we show how one can calculate distributions $\{P_j(F)\}$ by starting only from the experimental $P(S)$, that is, by making no assumptions about the shape of the laser focus and the detection volume. In the following, applying eq 6 or eq 7 instead of eq 1 is referred to as brightness correction. The correction methods are described in Sections 2.2–2.6.

Depending on how well the species can be separated by a burst classification algorithm and/or what additional data are available, different procedures of obtaining individual fluorescence intensity distributions $\{P_j(F)\}$ can be applied. Three typical scenarios of SM multiparameter fluorescence detection (MFD) in FRET experiments are illustrated in Figure 1. With increasing complexity, Figure 1 shows typical 2D parameter histograms of S_G/S_R versus $\tau_{D(A)}$, where $\tau_{D(A)}$ stands for the donor lifetime in the presence of acceptor. In the next sections, we show how the distributions $\{P_j(F)\}$ can be obtained in each of these three cases, that is, (A) all species can be separated, (B) only donor-labeled species can be separated, and (C) all three populations overlap.

Please note that for cases B and C, it is of course not necessary that different species exhibit nonoverlapping peaks on the $\tau_{D(A)}$ versus S_G/S_R 2D histogram (Figure 1), because one can use also other dimensions of the fluorescence parameter data set.[7] Advanced SM analysis methods such as MFD [7,31] and ALEX [32–34] are of great help for separating individual species, simplifying further analysis by PDA.

2.2. Independent Measurements of $P_j(F)$ (Scaling of $P(F)$ with Concentration). In the simplest case of nonoverlapping populations (Figure 1A), 1D histograms of interest (e.g., FRET

efficiency distributions) can be generated separately for each species, and usually, there should be no need to apply PDA to the overall histogram. We briefly discuss this case to show how the individual fluorescence intensity distributions $P_j(F)$ depend on concentrations of the corresponding species, which is relevant if these distributions have been obtained from different measurements. At low (SM) concentrations, multiple-molecule events can be neglected. It then becomes convenient to consider the point $F = 0$ separately, because for $F > 0$, $P_j(F)$ is proportional to the concentration of the corresponding species, c_j. One can then express $P_j(F)$ as

$$P_j(F) = (1 - p_j^{\text{in}}) + p_j^{\text{in}} P_j(0\text{lin}) \qquad F = 0$$
$$P_j(F) = P_j(F) = p_j^{\text{in}} P_j(F\text{lin}) \qquad F > 0 \tag{8}$$

where p_j^{in} stands for the probability of finding a molecule of the jth kind in the observation volume and $P_j(F\text{lin})$ denotes the fluorescence intensity distribution, provided that a molecule of species j is situated in the laser focus. The exact value of p_j^{in} depends on how the observation volume is defined; however, for the analysis, it is sufficient that under SM conditions, p_j^{in} is directly proportional to the concentration of the jth species, c_j. Note that the values of $P_j(F = 0)$ enter only in $P(F_1 = 0, F_2 = 0)$, which is actually equal to $P(F = 0)$ (see eq 7a). Therefore, the values of individual $P_j(F = 0)$ are irrelevant (i.e., we are not interested in which exact species has emitted zero photons; only the overall $P(F = 0)$ is needed). For simplicity, we assume in the following that the point $F = 0$ is taken care of by setting $P(F_1 = 0, F_2 = 0) = P(F = 0)$ and do not consider this point explicitly. It is then clear from eq 8 that for $F > 0$, $P_j(F)$ can be simply rescaled by multiplying each point of this distribution by the value of the corresponding relative concentration c_j. In other words, the shape of the interesting part of $P_j(F)$ does not change with the concentration, as long as the concentrations are low and multiple-molecule events are rare.

The knowledge of the overall deconvoluted distribution $P(F)$ (cf. eq 5) allows us to avoid absolute concentrations $\{c_j\}$ (which are often irrelevant) and work with relative concentrations (i.e., fractions) of the species $\{x_j\}$ $\left(\sum_{j=1}^{k} x_j = 1\right)$. The individual distributions $P_j(F)$ must be then rescaled so that they obey eq 9.

$$\sum_{j=1}^{k} x_j P_j(F) = P(F) \tag{9}$$

The normalization constant for $\{P_j(F)\}$ can be easily obtained by requiring that the mean fluorescence signal, calculated from a set of $\{P_j(F)\}$, is equal to the experimental fluorescence signal $\langle F \rangle$ for the fluorescence counts F, (eq 10).

$$\langle F \rangle = \sum_{j=1}^{k} \sum_{F=0}^{\infty} F x_j P_j(F) = \sum_{F=0}^{\infty} F P(F) \tag{10}$$

Each $P_j(F)$ (for $F > 0$) is multiplied by the same factor so that eq 10 is satisfied. Thus, for a general k-state model, $2k - 1$ floating parameters are needed: k parameters describing mean FRET efficiencies E_j or fluorescence anisotropies r_j and $k - 1$ relative concentrations or fractions (x_1,..., x_{k-1}; $x_k = 1 - \sum_{j=1}^{k-1} x_j = 1$).

2.3. Calculating $P(F)$ from a Measurement on Another Species (Scaling with Brightness). In practice, it is not always possible to separate bursts from similar states (Figure 1B). However, very often, an independent measurement on the same molecule is available, in which a single fluorescent state can be observed. For instance, in FRET experiments, a donor-only (D-only) labeled sample is usually studied separately. For convenience, let us assign $j = 0$ to this reference state and set the relative brightness of this species, Q_0, to be equal to 1. The fluorescence intensity distribution for this species $P_0(F)$ allows one to calculate the corresponding brightness distribution, $P(q_i)$, as follows from the representation of $P_0(F)$ as a weighted sum of Poisson distributions.[21,25,35,36]

$$P_0(F) = \sum_i P(q_i) \frac{q_i^F \exp(-q_i)}{F!} \tag{11}$$

To describe the irradiance shells of the focal volume, we use a predefined set of brightness values $\{q_i\}$, consisting of 100−200 elements. Because we do not attempt to relate the brightness distribution $P(q_i)$ to any physical parameters of the detection volume, we adapt the coefficients $P(q_i)$ to fit the experimental distribution $P(S)$ (cf. eq 5). Therefore, we also do not need the assumption that fluorescence intensity emitted by a molecule is constant during a counting time interval. As we proposed before,[20] starting from the distribution $P_0(F)$ any other $P_j(F)$ can be derived, provided that (i) the relative brightness of the jth state, Q_j, is known, (ii) diffusion properties of species 0 and j are the same, and (iii) saturation effects are negligible or affect all species equally. Under these conditions,

$$P_j(F) = \sum_i P(q_i) \frac{(Q_j q_i)^F \exp(-Q_j q_i)}{F!} \tag{12}$$

The above requirements are, in general, reasonable. The brightness ratio Q_j/Q_0 can often be obtained from independent data (e.g., from lifetime distributions for anisotropy PDA) or by FIDA.[21] We refer to this case as manual brightness correction. In other cases, Q_j might be not known, but it is often directly related to the corresponding E_j or r_j. Thus, the calculation of the relative brightness Q_j can be built into a fitting procedure (automatic brightness correction) without introducing new free parameters in the analysis, as will be shown below. Please note that in this work, only relative brightness values are relevant. When these values are given in kiloHertz, they define the count rate which corresponds to a mean concentration of one molecule in the observation volume.[19,21]

The other two assumptions regarding diffusion and saturation effects are needed to justify the use of the same brightness distribution $P(q_i)$ for all species (eqs 11 and 12). The distribution $P(q_i)$ describes the probabilities that a molecule with a brightness of 1 takes a certain path through the observation volume, for which the mean number of emitted photons is q_i. It is clear that another (jth) species having the same diffusion coefficient would take the same path with the same probability $P(q_i)$, emitting however $Q_j q_i$ photons on average. Thus, the shape of $P(q_i)$ as it enters in eqs 11 and 12 is independent of the brightness of the molecule, being determined only by the shape of the observation volume and the molecular diffusion coefficient.

In practice, it is not unusual that the diffusion properties of all species in a series of experiments are similar. In particular, in FRET experiments, D-only and DA species differ only by the acceptor dye, which is usually much smaller than the macromolecule of interest and therefore has a negligible effect on the diffusion coefficient. A similar situation occurs when, for example, binding of a small ligand to a macromolecule is investigated.

Saturation effects (e.g., triplet formation) are more difficult to account for. These effects may change the effective shape of the observation volume,[24] possibly to a different extent depending on the fluorescence properties of the species. For example, in FRET experiments, the population of the triplet state of the

donor dye depends on its fluorescence lifetime, which in turn depends on the distance to the acceptor.[12] Moreover, at high excitation powers, the assumption of equal G and R observation volumes can be violated because of different triplet behavior of the donor and the acceptor dyes. Therefore, in real measurements, saturation effects must be controlled by selecting appropriate excitation intensities.

The proposed dependence of fluorescence intensity distribution on brightness (eq 12) is strictly valid only if the whole photon trace is analyzed. Burstwise analysis involves the use of a burst search algorithm, which is usually empirical and may change the shape of all $\{P_j(F)\}$ differently, depending on the mean brightness. One can overcome this problem to some extent by including adjacent background photons into fluorescence bursts (cf. ref 37). Instead of selecting bursts from the molecules of interest, one can also exclude bursts due to unwanted species and analyze the remaining data as a whole photon trace. Because less-bright bursts are usually not selected, this procedure cannot completely suppress the fluorescence of unwanted species; however, the final contribution of these species to the total signal is not expected to be significant.

2.4. Calculation of $P_j(F)$ in a General Case (Simultaneous Deconvolution). It is also possible to obtain the individual fluorescence intensity distributions $\{P_j(F)\}$ even if none of the species of interest can be studied separately (Figure 1C). In this case, one needs a guess for relative brightness and concentrations of all species, which can be, for instance, intermediate results of a fitting procedure. By assuming that all concentrations are low and eq 8 is valid and by combining eqs 9 and 12, one obtains

$$P(F) = \sum_{j=1}^{k} x_j P_j(F) = \sum_i P(q_i) \sum_{j=1}^{k} x_j \frac{(q_i Q_j)^F \exp(-q_i Q_j)}{F!} \tag{13}$$

Eq 13 is valid under the same conditions as those for eq 12, that is, when the diffusion coefficient is the same for all species, saturation effects are negligible, and a whole photon trace is considered. At this point, we assume that the values of $\{Q_j\}$ and $\{x_j\}$ are either known or represent the result of a fitting iteration. Therefore, eq 13 contains exactly the same set of free parameters as the original eq 11; that is, only the distribution $P(q_i)$ has to be calculated, which can be done in the same way as that for a single species.[20] From $P(q_i)$, any $P_j(F)$ is easily obtained according to eqs 11 and 12. This method is relatively time-consuming, because deconvolution has to be performed before each fit iteration.

In the next two sections, we demonstrate that in many typical cases, deconvolution of $\{P_j(F)\}$ can be combined with fitting procedure, without introducing extra unknown parameters.

2.5. Brightness Correction in FRET Experiments. So far, we have assumed that the relative brightness of each species Q_j is known, which is needed to calculate the distributions $\{P_j(F)\}$ according to eq 12 or by deconvolution (eq 13). Alternatively, the relative brightness of each species can be expressed as a function of the corresponding mean FRET efficiency E_j, that is, automatically calculated before each fit iteration. In this case, the fitting procedure has to be modified so that before each iteration, the set of $\{P_j(F)\}$ is also updated by using current values of $\{Q_j\}$ and $\{x_j\}$ where applicable. By assuming that all brightness variations are only due to FRET, one can show that Q_j is related to E_j according to (see Appendix)

$$Q_j \propto 1 - E_j\left(1 - \frac{\Phi_{FA}}{G\Phi_{FD}(1+\alpha)}\right) \tag{14}$$

For convenience, eq 14 is normalized so that the brightness of D-only species is equal to 1.

2.6. Brightness Correction in Anisotropy Experiments. Anisotropy values often vary because of local mobility of the dyes or local quenching. If only one of these processes is responsible for all anisotropy variations, it is possible to find a relationship between the mean anisotropy values r_j and the mean brightness Q_j, that is, to implement automatic brightness correction. Here, it is important to distinguish between experimental brightness, which is proportional to the fluorescence signal measured by two detectors and should be used in eq 12, and the theoretical one, which is related to the fluorescence emitted in all spatial directions. In the first case (i.e., different dye mobility), the theoretical total intensity does not change (for details, see Appendix); however, the observed brightness $(F_{||} + F_{\perp})$ might be polarization-dependent. The mean measured brightness is then given by

$$Q_j \propto 1 + \frac{r_j(3Gl_2 - 3l_1 + 2 - G)}{1+G} \tag{15}$$

In eq 15, l_1 and l_2 denote the correction factors that describe depolarization by the microscope objective,[38] and $G = g_{\perp}/g_{||}$. For realistic values of G, l_1, and l_2, the above-mentioned effect is relatively weak. For instance, if $G = 1$, $l_1 = 0.03$, and $l_2 = 0.04$, the anisotropy change from 0 to 0.4 leads only to ~20% increase in the measured brightness.

For the case of local quenching, we assume that all species exhibit the same single rotational correlation time. The anisotropy then obeys the Perrin equation.[15,39] Moreover, it is often reasonable to assume that the fluorescence quantum yield is proportional to the fluorescence lifetime. One then obtains (for details, see Appendix)

$$Q_j \propto \left(1 + \frac{r_j(3Gl_2 - 3l_1 + 2 - G)}{1+G}\right)\left(\frac{r_0}{r_j} - 1\right) \tag{16}$$

In eq 16, r_0 denotes the fundamental fluorescence anisotropy,[15,39] which is a property of the dye and can be precisely measured in ensemble experiments. Let us note that according to eq 16, for low anisotropies, Q_j becomes very large, which may cause numerical instabilities when using eq 16 in practice.

Such automatic brightness correction in PDA is not limited to these few common cases that we have discussed. Whenever one can find a relationship between the mean brightness and other parameters characterizing a fluorescent state, expressions similar to eqs 14–16 can be derived and included into the PDA model function to account for brightness effects, without introducing any unnecessary free parameters.

2.7. Model-Free Deconvolution. It has been shown that the PDA approach can unambiguously separate shot-noise broadening from the distribution of relevant physical parameters.[6] The underlying distribution can be modeled either with a fixed mean FRET efficiency (single distance) or alternatively by distributed FRET efficiencies (distribution of distances, e.g., Gaussian [6,18]). In practice, continuous distributions are usually modeled as a large number of discrete states, with the corresponding fractions being described by a function with a small number of parameters (e.g., a Gaussian distribution with known mean and width). In this way, PDA is a perfect tool to compare between experimental shot-noise broadened distributions and models which might include a certain number of static fluorescent states and continuous distributions. However, the opposite is not true: it

is clear that, in general, the underlying distribution cannot be unambiguously recovered by PDA because of limited data quality.

In addition to data fitting with a predefined model, a model-free deconvolution of the underlying distributions is possible, which is usually associated with the MEM.[27–29,40] Here, we present an extension to the PDA method inspired by the work of Watkins et al.,[17] who proposed to use MEM to extract distributions of distances from photon traces obtained from immobilized molecules. We adapt these ideas to diffusing molecules while using the PDA framework to account for shot-noise and background contributions.

The MEM is believed to extract the most unbiased distributions that satisfactory describe the experimental data.[27–29,40] Instead of minimization of reduced chi-squared values (χ_r^2), the following function is maximized (eq 17).

$$\Theta = \nu s - \chi_r^2 \tag{17}$$

In eq 17, ν is a constant, and s is an entropy-like function. [28]

$$s = \sum_j \left(X_j - M_j - X_j \log \frac{X_j}{M_j} \right) \tag{18}$$

where $X = \{X_j\}$ is the distribution of the parameter of interest (e.g., DA distance) and $M = \{M_j\}$ is the initial model for this distribution. The role of the factor ν in eq 17 is addressed differently by several authors. A theoretical approach is presented by Gull and Skilling.[40] However, in many other works, the value of ν is adjusted to maximize the entropy or the value of Θ, while the value of χ_r^2 is still statistically reasonable (see, for example, refs 17 and 27). In particular, Watkins et al.[17] proposed that χ_r^2 must be within $1 \pm 1/(n)^{1/2}$, where in our case, n would be the number of PDA histogram bins.

In our experience, somewhat more reproducible results (at least when applied in combination with PDA) can be obtained by requiring that, for a given data set, χ_r^2 shall not exceed its minimal possible value by more than $1/(n)^{1/2}$ (or any value proportional to $1/(n)^{1/2}$ which is believed to be statistically significant). This condition seems to be reasonable, because even for a perfect fit, χ_r^2 is not necessarily equal to 1 but is described by a distribution with a mean of 1 and a variance of $2/n$,[41] and therefore, in many cases, $\chi_r^2 < 1 + 1/(n)^{1/2}$ cannot be fulfilled. In reality, the minimal possible value of χ_r^2 can be even higher because of experimental artifacts, a few of which are discussed in refs 6 and 18. Thus, in this work, the value of ν was manually adjusted so that χ_r^2 exceeded its minimal value (usually corresponding to $\nu = 0$) by $1/(n)^{1/2}$, which effectively determines the width of the MEM distribution of X.

Another possible source of uncertainties in MEM is the initial model function M. The final result (X) can also vary (although usually not dramatically) depending on which parameter is assumed to be uniformly distributed a priori. The DA distance R (as used in this work) seems to be a good choice, but there is no obvious reason to prefer this parameter to the FRET efficiency, for example.

Maximization of Θ (eq 17) was performed as proposed by Vinogradov and Wilson.[42] The resulting distribution of DA distances was converted into a distribution of FRET efficiencies and further used to generate a shot-noise broadened PDA model histogram according to eq 1 or eq 7.

2.8. Multiple-Molecule Events. In this section, we show how the PDA model function can be corrected for the presence of multiple-molecule events in a SM measurement. As we will see, for burstwise data, an empirical probability of multimolecule events has to be introduced, whereas no extra free parameters are required if whole photon traces are analyzed.

The main idea of the correction procedure is the same as that in FIDA methods,[19,21] that is, to calculate a convolution of a number (m) of $P'(F_1, F_2)$ terms, where $P'(F_1, F_2)$ describes the probability of observing a particular combination (F_1, F_2) for single molecules and $1/m$ is equivalent to dilution factor of the original sample. The convolution can be calculated with the help of generating functions[14,19,21] or via 2D discrete Fourier transform (DFT, cf. ref 43).

$$P(F_1, F_2) = \underbrace{P'(F_1, F_2) \otimes \ldots \otimes P'(F_1, F_2)}_{m \text{ times}} = \mathrm{DFT}^{-1}([\mathrm{DFT}(P'(F_1, F_2))]^m) \tag{19}$$

In eq 19, DFT^{-1} denotes the inverse DFT. Let us first consider the case when the brightness of all species is similar and eq 1 can be used. Unlike in FIDA, in PDA, $P'(F_1, F_2)$ cannot be easily calculated because the corresponding SM fluorescence intensity distribution $P'(F)$ is not known and the distribution $P(F)$ is available only for a concentrated sample. Therefore, one has to find a distribution $P'(F)$ which would yield the experimental total $P(F)$ after m convolution steps. We can again make use of the DFT.

$$P'(F) = \mathrm{DFT}^{-1}([\mathrm{DFT}(P(F))]^{1/m}) \tag{20}$$

In eq 20, $P'(F)$ is a fluorescence intensity distribution for a sample, diluted m times as compared to the original mixture, for which the corresponding distribution is $P(F)$. Substitution of $P'(F)$ into eq 1 yields $P'(F_1, F_2)$. It is clear that by choosing sufficiently high values of m, contribution of multiple-molecule events to $P'(F_1, F_2)$ can be reduced to any desired level. In this work, the value of m is selected so that $P'(F = 0) > 0.99$, which roughly corresponds to the SM detection concentration, for example, the mean number of molecules in the observation volume of $N_{FCS} = 0.0015$ (for $Q = 100$ kHz).

This procedure can be extended to the case of different $P_j(F)$, provided that a reference (single-state, e.g., D-only) sample is available. First, each $P_j(F)$ has to be calculated according to eq 12 and then rescaled according to eqs 8 and 10 to include the fraction of the jth species. Before applying eq 20, a proper value of $P_j(F = 0)$ must be assigned (obviously, $P_j(F = 0) = 1 - \sum_{F>0} p_j(F)$). When all distributions $\{P_j'(F)\}$ are calculated (eq 20), $P'(F_1, F_2)$ can be obtained according to eq 7b. $P(F_1, F_2)$ is then calculated (see eq 19) and convoluted with background (eq 7c) to yield $P(S_1, S_2)$.

To our knowledge, all commonly used burst search algorithms tend to find bursts with high photon numbers; that is, multi-molecule events are preferentially selected. Therefore, the proposed correction for multiple-molecule events is strictly valid only if complete photon traces are analyzed. In a case of burstwise analysis, the contribution of multimolecule events becomes difficult to predict. At relatively low concentrations, it should be sufficient to consider only single- and double-molecule events with certain weights. Unfortunately, we have found no obvious relationship between these weights which would lead to the best fit and any parameters of our burst search algorithm. Thus, if a precise correction for multimolecule events is required, the whole photon trace should be analyzed.

3. Methods

3.1. Chemicals. Rhodamine 110 (Rh110) was purchased from Radiant Dyes (Germany) and used as received. Rhodamine

6G bound to a double-stranded DNA (Rh6G-dsDNA) was prepared as described elsewhere.[44]

3.2. SM Fluorescence Measurements. The experiments were performed by using a MFD technique, as described elsewhere.[8,9,45] Briefly, SM fluorescence detection was performed by using a confocal epi-illuminated microscope with excitation by an active mode-locked Ar^+ laser (73.5 MHz, 150 ps, Inova Sabre, Coherent, Palo Alto, CA) at 476.5 nm. The linearly polarized beam was focused into solution with a 60 × 1.2 water immersion objective (UPlan Apo, Olympus, Hamburg, Germany). The excitation power in the focus of the microscope objective was measured by a power meter (Fieldmaster FM-2, Coherent). The diameter of the focus in the *xy*-plane was about 1.2 μm, as estimated by fluorescence correlation spectroscopy (FCS) from diffusion time of Rh110 molecules ($t_D = 0.26$ ms) by using diffusion constant from the literature. Fluorescence detection was performed with the same objective by using a confocal pinhole ($\varnothing = 100$ μm) which results in a detection volume of ~3 fl. The fluorescence signal was separated from the laser light by a dichroic beamsplitter (Q485DCLP, AHF, Analysentechnik, Tübingen, Germany) and further divided into its parallel (||) and perpendicular ($\perp$) components by a polarizing beamsplitter cube (VISHT11, Gsänger, Planegg, Germany). The color range of detection channels was selected by interference filters (green HQ520/66 and red HQ630/60, AHF Analysentechnik, Tübingen, Germany). The photons were detected by four single photon avalanche diodes (Micro Photon Devices, Bolzano, Italy) coupled to PC-based time-correlated single-photon counting (TCSPC) modules (modified SPC 132, Becker and Hickel GmbH, Berlin, Germany). Experiments on Rh6G-dsDNA were performed as described in ref 44.

3.3. Simulations of SM Fluorescence Experiments. Simulations of SM measurements were performed by using the Brownian dynamics approach[21,46–48] with a few modifications as described elsewhere.[20] The maximal brightness was taken to be 100 kHz, which was similar to typical experimental values measured with our setup. The diffusion time of all species was 1 ms, and the simulation time step was 0.005 ms. Poisson-distributed background was added when necessary. The mean number of molecules in the focus (defined as in FCS[49]) was 0.01 unless stated otherwise. All simulations were performed on a 2.2 GHz PC; the speed was about 20 000 events/s. The data were stored in our regular data format[50] for standard analyses.

3.4. Data Analysis. The recorded sequence of photon events is used to compute the intensity trace of equal non-interleaving time windows (or time bins); the duration of the time window is $\Delta t = 1$ ms unless otherwise stated. The data are presented as 1D histograms of any ratiometric or normalized parameter of interest as described.[6,20] Calculation of $P(S_1, S_2)$ for the case of multiple states is conveniently performed via $P(F_1, F_2)$ (eq 7): convolution with the background distribution, which is usually the most time-consuming step, is performed only once irrespectively of the number of fluorescent states considered. The fitting procedure may however need a modification if the values of $\{Q_j\}$ are not known a priori. Then, it becomes necessary to recalculate the fluorescence intensity distributions $\{P_j(F)\}$ before each iteration. To summarize, for the case of SM-FRET data, the fitting procedure can be described as follows.

Step 1. Read data and calculate experimental $P(S_1, S_2)$. Generate a 1D histogram of choice from the experimental data. Deconvolute total $P(F)$. Suggest an initial guess for concentrations and FRET efficiencies or model parameters (e.g., the mean and the width of a Gaussian distribution of DA distances).

Step 2A. (All species are separated; cf. Figure 1A) Load previously calculated $\{P_j(F)\}$, or

Step 2B. (D-only species can be separated; cf. Figure 1B) load previously calculated $P(q_i)$, obtained from a D-only sample, for example (cf. Section 2.3). Calculate $\{Q_j\}$ for species 1 to j by using the latest $\{E_j\}$ (eq 14) and then $\{P_j(F)\}$ according to eq 12, or

Step 2C. (All populations overlap; cf. Figure 1C) calculate $\{Q_j\}$ and deconvolute $\{P_j(F)\}$ (eq 13) by using the latest values of $\{E_j\}$ and $\{x_j\}$.

Step 3. Calculate 2D model distribution $P(S_1, S_2)$ (eq 7). Generate a 1D model histogram and compare with the experimental data.

Step 4. Use an optimization algorithm (e.g., Levenberg–Marquardt) to find a new set of model parameters (e.g., $\{E_j\}$ and $\{x_j\}$). Repeat steps 2–4 until convergence is reached.

A general k-states model includes $2k - 1$ floating parameters: k mean FRET efficiencies or anisotropies and $k - 1$ relative concentrations (or fractions). Quality of the fits is judged by χ_r^2 values and weighted residuals plots.

4. Results and Discussion

4.1. Experimental Verification of Brightness Scaling Algorithm. The dependence of the fluorescence intensity distribution on the relative brightness (eq 12) is essential for all brightness correction procedures described in Sections 2.3–2.6. Although eq 12 itself is a direct consequence of the well-known Mandel's formula[35] (eq 11), deconvolution of the brightness distribution $P(q_i)$ from noisy data introduces numerical instabilities, which could make derived distributions (eq 12) unusable. We show that this is not the case by verification of eq 12 on experimental data, by considering also saturation effects that are not explicitly taken into account in eqs 11 and 12.

For this, we measured the fluorescence of Rh110 dye at SM concentration, at two different excitation intensities. The excitation power as measured in the focus of the microscope objective was 741 and 195 μW, which corresponds to the mean irradiance $I_0/2(H) = 65.5$ kW/cm^2 and $I_0/2(L) = 17.2$ kW/cm^2, respectively.[51] The brightness of Rh110 at these excitation intensities was $Q_H = 90.3$ kHz and $Q_L = 46.9$ kHz, as determined by FIDA,[19] and subscripts H and L refer to high and low excitation power, respectively. Note that the brightness ratio (0.52) is significantly different from the power ratio (~0.26), which indicates strong saturation.

Case 1: Different Detection Efficiencies. Because the signal is registered by four detectors, a lower apparent brightness can be simply achieved by disregarding photons with perpendicular polarization with respect to the polarization of excitation light, yielding the distribution $P_H(F_{||})$ and the corresponding brightness $Q_{||,H} = 46.2$ kHz. Figure 2A shows this distribution together with the distribution $P_H(F)$, scaled according to eq 12 ($Q_{||,H}/Q_H = 0.51$). It is clear that in this case, a good agreement between the experimental and calculated fluorescence intensity distributions is observed in the whole range of photon numbers, which proves that the proposed procedure of calculating unknown fluorescence intensity distributions according to eqs 11 and 12 is valid.

Case 2: Nonlinear Saturation. To study the influence of optical saturation, we compared the experimental fluorescence intensity distribution measured at lower power, $P_L(F)$, with that derived according to eq 12, starting from the distribution $P_H(F)$ measured at higher power. The relevant fluorescence intensity distributions calculated by deconvolution from the corresponding signal intensity distributions $P(S)$ are presented in Figure 2B.

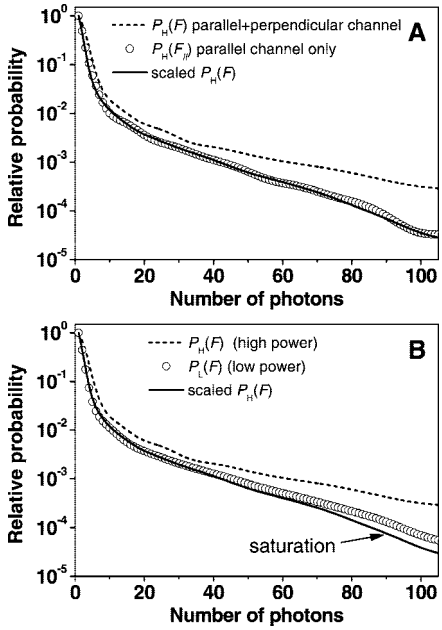

Figure 2. Dependence of fluorescence intensity distributions on the mean brightness. The dashed line on both plots shows the fluorescence intensity distribution of Rh110, $P_H(F)$, measured at 741 μW at the objective, detected by all four channels. (A) Experimental distribution $P_H(F_{||})$ of Rh110 measured at 741 μW (only photons with parallel polarization are counted, open circles) and distribution obtained from $P_H(F)$ by using eq 12 (solid line, $Q_{||H}/Q_H = 0.51$). (B) Comparison of the experimental distribution $P_L(F)$ measured at 195 μW (open circles) and the distribution obtained according to eq 12 (solid line, $Q_L/Q_H = 0.52$), starting from the experimental $P_H(F)$ measured at 741 μW.

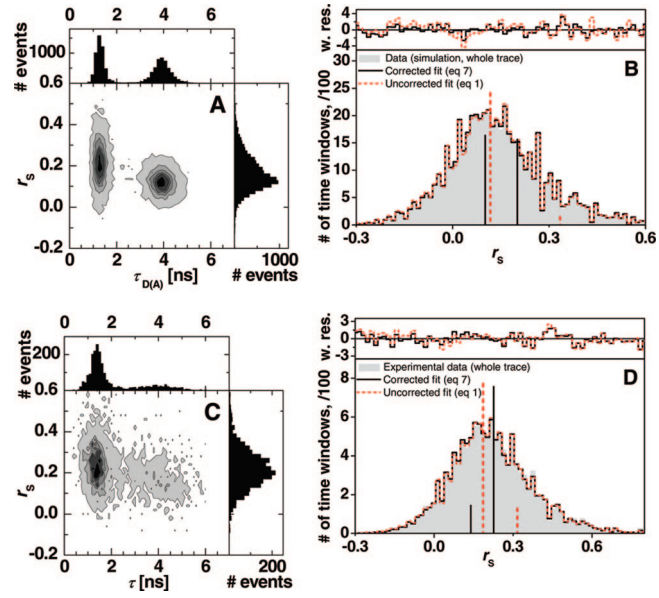

Figure 3. (A) 2D scatter-corrected fluorescence anisotropy (r_S) versus $\tau_{D(A)}$ histogram showing the results of burstwise analysis of simulated mixture of two static states, for which $r_1 = 0.1$ (50%), $r_2 = 0.2$ (50%), $Q_1 = 100$ kHz, and $Q_2 = 33.7$ kHz. The corresponding fluorescence lifetimes are 4 and 1.29 ns, respectively. $\langle B_{||} \rangle = 1.7$, $\langle B_{\perp} \rangle = 1.3$, and $t_D = 1$ ms. The population exhibiting lower anisotropy contains ~46% of all fluorescence bursts. To simplify burst search, we assumed that quenching of state 2 is due to FRET and added red photons, $Q_{red} = Q_1 - Q_2 = 66.3$ kHz. (B) 1D histogram of scatter-corrected fluorescence anisotropy obtained from the whole photon trace of the same data set (gray area). PDA fits with (solid black line) and without (dashed red line) brightness correction are shown. Fitting of the PDA model, corrected for brightness effects (eq 7), yields $r_1 = 0.102$ (53.3%) and $r_2 = 0.206$ (46.7%); $\chi_r^2 = 1.12$. Uncorrected fit (i.e., assuming the same $P(F)$ for both species) yields $r_1 = 0.117$ (94.7%) and $r_2 = 0.334$ (5.7%); $\chi_r^2 = 2.24$. Fixing true values and applying uncorrected model results in $\chi_r^2 = 22.79$. (C) Scatter-corrected fluorescence anisotropy (r_S) versus fluorescence lifetime histogram showing the results of burstwise analysis of Rh6G-labeled dsDNA fluorescence. The population exhibiting lower anisotropy contains ~24% of all fluorescence bursts. (D) 1D histogram of scatter-corrected fluorescence anisotropy obtained from the same data set (gray area). PDA fits with (solid black line) and without (dashed red line) brightness correction are shown. Fitting of the PDA model, corrected for brightness effects (eq 7), yields $r_1 = 0.140$ (16.0%) and $r_2 = 0.227$ (84.0%); $\chi_r^2 = 0.69$. Uncorrected fit yields $r_1 = 0.186$ (85.7%) and $r_2 = 0.317$ (14.4%); $\chi_r^2 = 0.83$.

Clearly, at low photon numbers, the scaled distribution $P_H(F)$ (eq 12) closely resembles the distribution $P_L(F)$. However, at higher photon numbers, small but systematic deviations can be observed, which can be expected because of triplet formation.[25] Nevertheless, the S_G/S_R signal ratio histogram obtained from Rh110 data can be satisfactorily fitted by using the scaled $P_H(F)$, providing $\chi_r^2 = 1.11$ and $\langle F_G/F_R \rangle = 10.97$. Common PDA fit (eq 1) of the same data yields $\chi_r^2 = 0.93$ and $\langle F_G/F_R \rangle = 10.91$ (data not shown). These results show that fluorescence intensity distributions can be scaled according to eq 12 for the whole range of excitation intensities relevant for SM fluorescence experiments.

It is important to emphasize that unlike in FIDA, the assumption of $\Delta t \ll t_D$ is not required: for Rh110, $t_D = 0.26$ ms and $\Delta t = 1$ ms. This fact allows us to work with large time windows and therefore high photon numbers, significantly improving the ability of PDA to resolve multiple species.[20]

4.2. Effect of Brightness Variations on PDA. ***4.2.1. Brightness Correction.*** At first, we verify the proposed brightness correction methods (sections 2.2–2.4) by applying them to simulated data which closely mimic experimental data. The analysis of simulations, for which the true parameters are exactly known, is performed in order to characterize stability of the analysis and search for systematic errors. Later, we apply PDA to real experimental data. We also show that in both cases, brightness variations could produce results that dramatically deviate from the expected (or simulated) values, if they are not taken into account by using appropriate brightness correction methods (eqs 6 and 7).

To test the proposed brightness correction methods, a system has been simulated which consists of two states, 1 and 2, where one state is partially quenched (Figure 3A) so that $Q_1 = 100$ kHz is reduced to $Q_2 = 33.7$ kHz. Because of local quenching, the two states have different fluorescence anisotropies: $r_1 = 0.1$ and $r_2 = 0.2$. The corresponding fluorescence lifetimes of these two states are 4 and 1.29 ns, respectively, and the rotational correlation time (ρ) is the same for both states ($\rho_1 = \rho_2 = 1.43$ ns). In the simplest case, we can obtain $P_1(F)$ and $P_2(F)$ independently (Section 2.2; cf. Figure 1A). Then, we calculate the distribution $P_2(F)$ according to eq 12 (brightness correction with reference $P(F)$; cf. Figure 1B) by assuming that the first state can be studied independently and $P_1(F)$ is known. Finally, both individual fluorescence intensity distributions can be simultaneously deconvoluted by using eq 13.

Figure 3A shows the results of burstwise analysis of the simulated data in a scatter-corrected fluorescence anisotropy (r_S) versus $\tau_{D(A)}$ 2D frequency histogram. From Figure 3A, it is clear that two populations with $r_1 \approx 0.1$ and $r_1 \approx 0.2$ are present, which have approximately equal concentrations. A histogram of scatter-corrected fluorescence anisotropy (r_S)[20,52] generated from the same data set is presented in Figure 3B. If, however, this histogram is analyzed by the PDA model which takes no brightness effects into account (eq 1), $r_1 = 0.117$ (94.7%) and

TABLE 1: Comparison of Brightness Correction Methods as Applied to Simulated and Experimental Data Sets (Figure 3)

brightness correction/analysis method	Q_2/Q_1	r_1	r_2	χ_r^2
Simulated Data, $r_1 = 0.1$ (50%) and $r_2 = 0.2$ (50%) (Figure 3A,B)				
none, single state	not needed	0.125 (100%)		5.51
none, two states		0.117 (94.7%)	0.334 (5.7%)	2.24
		0.1 (50%) (fixed)	0.2 (50%) (fixed)	22.79
none, 1/10 data	not needed	0.117 (93.9%)	0.29 (6.1%)	1.12
individual $P_j(F)$	not needed	0.101 (53.7%)	0.206 (46.3%)	1.16
with reference $P(F)$ (eq 12)	0.337 (manual, fixed)	0.102 (53.3%)	0.206 (46.7%)	1.13
	automatic, eq 16	0.102 (52.3%)	0.204 (47.7%)	1.13
Simultaneous deconvolution (eq 13)	0.337 (manual, fixed)	0.101 (51.2%)	0.201 (48.8%)	1.11
	automatic, eq 16	0.103 (53%)	0.208 (47%)	1.10
Experimental Data (Figure 3C,D)				
none	not needed	0.186 (85.7%)	0.317 (14.4%)	0.83
simultaneous deconvolution (eq 13)	0.35 (manual, fixed)	0.132 (9.8%)	0.220 (90.2%)	0.84
	automatic, eq *16*	0.140 (16.0%)	0.227 (84.0%)	0.69
burst size distribution (from ref 44)		lower (18%)	higher (82%)	
selected data $\tau > 2.65$ ns	not needed	0.144		1.00
selected data $\tau < 2.65$ ns	not needed		0.214	1.01

$r_2 = 0.334$ (5.7%) are obtained, which is completely inconsistent with visual analysis of 2D histogram in Figure 3A. To prove that these results are not simply due to instabilities, we used the correct fixed values in the fit and obtained $\chi_r^2 = 22.79$, which indicates that eq 1 is inappropriate.

For an improved PDA analysis, let us first assume that the relative brightness of the second state is known (manual brightness correction). For instance, in the case shown in Figure 3A, one can estimate it from the fluorescence lifetime distributions ($Q_2/Q_1 \approx 1.29$ ns/4 ns $= 0.323$). Alternatively, the brightness values can be estimated by FIDA[21] if the difference in brightness is significant. In our case, the FIDA method yields $Q_1 = 98.4$ kHz and $Q_2 = 33.7$ kHz, which is in excellent agreement with the values used in the simulation.

The best fit of PDA model function, in which the brightness effects are taken into account by calculating $P_2(F)$ from reference $P_1(F)$ according to eqs 11 and 12, is also shown in Figure 3B. The recovered model parameters are in a good agreement with the values used in the simulations: $r_1 = 0.102$ (53.3%) and $r_2 = 0.206$ (46.7%) ($\chi_r^2 = 1.12$). All other brightness correction methods, that is, using precalculated individual $P_1(F)$ and $P_2(F)$ and simultaneous deconvolution of these distributions by using eq 13, produce very similar results (Table 1).

For the analysis of S_G/S_R distributions in FRET measurements, the influence of brightness effects is usually less dramatic. Strong correlation between the FRET efficiency and brightness may produce poor fits and wrong calculated fractions of the species. However, when the major peaks are well resolved, the mean positions of these peaks (i.e., mean FRET efficiencies) are usually not much affected by brightness variations and can be recovered with a reasonable accuracy (data not shown).

4.2.2. Automatic Brightness Correction. In the previous subsection, we have assumed that the values of Q_1 and Q_2 are known, which might not always be possible for a real system. Now, we show how automatic calculation of relative brightness of each state as described in Sections 2.5 and 2.6 can be applied. In other words, the values of Q_1 and Q_2 are updated depending on the intermediate values of r_1 and r_2 for each fit iteration. For the systems shown in Figure 3, eq 16 is relevant.

Table 1 summarizes the conclusions of the previous subsection, as well as the results obtained by automatic brightness correction, that is, when the value of Q_2/Q_1 is calculated according to eq 16.

Table 1 makes it clear that all brightness correction methods provide similar results. We have found no indications for severe instabilities introduced by any of these approaches, although the initial guess parameters are quite distant from the expected values $r_1 = 0$ (50%) and $r_2 = 0.27$ (50%). The accuracy of the methods based on the calculation of several unknown fluorescence intensity distributions is comparable with that of the approach which makes use of individual $\{P_j(F)\}$, directly obtained in a separate series of experiments. Simultaneous deconvolution of all $\{P_j(F)\}$ (eq 13) is clearly the most general approach, but it is also the slowest and is not compatible with our realization of MEM, which will be discussed in Section 4.4. Brightness correction with reference $P(F)$ (eq 12) does not require time-consuming computations and should be also often applicable. Automatic correction, that is, iterative calculation of $\{Q_j\}$ from other fit parameters, introduces no noticeable systematic errors or instabilities as compared to independent calculation of $\{Q_j\}$ but usually needs more iterations to converge. Another important outcome of the analyses of the simulated system is that sometimes, a PDA model function, not corrected for brightness variations (eq 1), might produce a reasonably good fit but completely wrong results. As shown in Table 1, χ_r^2 values do not always indicate that the model function is wrong, especially when data quality is limited by a low number of events (see example 1/10 data; $\chi_r^2 = 1.12$).

4.2.3. Application to Experimental Data. The dye Rhodamine 6G bound to the 5′ end of a double-stranded DNA exhibits multiple conformational states,[44] which are characterized by different fluorescence lifetimes and anisotropies (Figure 3C). In the present work, we use a two-state model, which is sufficient to fit the anisotropy data, although other methods indicate that the brighter subpopulation additionally consists of two states.[44] We are not able to resolve these two states solely by anisotropy PDA; however, we can show that brightness-corrected PDA provides at least qualitatively reasonable results, whereas the simple model (eq 1)[6,20] does not. Experimental histograms of scatter-corrected anisotropy r_S and PDA fits are presented in Figure 3D, and the results of anisotropy analyses are summarized in Table 1. It is clear that using the same fluorescence intensity distribution $P(F)$ for all states leads to severe errors in estimating both anisotropies and fractions of the two states. On the other hand, applying brightness correction yields reasonable fractions and mean anisotropies of individual conformations of Rh6G, in agreement with independent data.[44] The brightness values Q_1 and Q_2 can be estimated from the ratio of the mean fluorescence lifetimes of the two states ($Q_2/Q_1 \approx 0.35$; manual brightness correction) or automatically

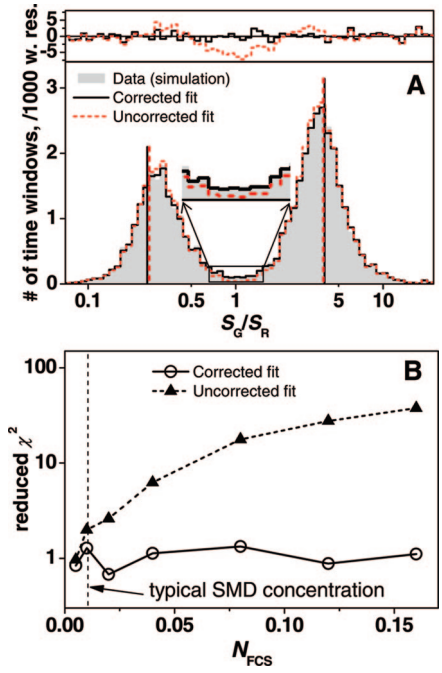

Figure 4. Effect of multiple molecule events on the quality of PDA fits. (A) S_G/S_R ratio histograms for the total FCS concentration of 0.04 (gray area) shown together with corrected PDA model (solid black line; $E_1 = 0.799$ (40.1%),$E_2 = 0.200$ (59.9%), $\chi_r^2 = 1.13$) and uncorrected PDA model (red dashed line; $E_1 = 0.794$ (40.0%), $E_2 = 0.203$ (60.0%), $\chi_r^2 = 6.22$). (B) χ_r^2 plotted versus the mean number of molecules in the focus (N_{FCS}), for the same values of E_1 and E_2 and relative concentrations. Simulation parameters: $E_1 = 0.8$ (40%); $E_2 = 0.2$ (60%); $Q_1 = Q_2 = 100$ kHz; $\langle B_G \rangle = 2$; $\langle B_R \rangle = 1.2$; $t_D = 1$ ms. The vertical dashed line indicates a typical SM concentration.

according to eq 16. These two brightness correction methods provide similar mean anisotropies of the individual conformations of Rh6G (Table 1), which also agree well with the values calculated by PDA of preselected data (Table 1).

To summarize, in its present form, the PDA method not only is able to detect heterogeneities but also accurately extracts relevant parameters of individual states. In the presented example with simulated data (Figure 3A,B), the parameters of two similar states have been extracted from one broad distribution with an accuracy of 3−4% or better. Application of PDA to experimental data reveals the presence of at least two conformations of Rh6G-DNA, which is not obvious from visual analysis of 1D anisotropy histograms (Figure 3C, right subplot). Recovered parameters of these states are in reasonable agreement with independent data.

4.3. Correction for Multiple-Molecule Events. In experiments with freely diffusing single molecules, it is usually not possible to completely avoid multimolecule events. The only way to reduce distortions of experimental SM histograms due to a small fraction of multiple-molecule events is to perform SMD experiments at very low concentrations (mean number of molecules in the focus $N_{FCS} \approx 0.001$), which necessarily implies much longer measurement times and huge data amounts (mostly scatter photons are recorded). Instead, one can use the correction procedure presented in Section 2.8 (eqs 19 and 20), as illustrated in this section.

An example showing the effect of multiple-molecule events on PDA histograms is presented in Figure 4. The simulated system models a mixture of two static FRET states, with $E_1 = 0.8$ (40%; $S_G/S_R = 0.25$) and $E_2 = 0.2$ (60%; $S_G/S_R = 4$). Already at a mean number of molecules in the focus of $N_{FCS} = 0.02$ (under our threshold conditions for burst selection, approximately 3−4 bursts per second at $Q = 100$ kHz), high reduced χ_r^2 values ($\chi_r^2 > 2$) can be easily misinterpreted as meaningful and motivate the use of unnecessarily complex models (e.g., broadened states or conformational dynamics). The deviations at $S_G/S_R \approx 1$ between the purely SM PDA model function and the simulated data can be clearly seen (Figure 4A). A good fit can be achieved by the multimolecule-correction procedure (eqs 19 and 20; $m = 10$), yielding $\chi_r^2 = 1.13$. The extracted values of $E_1 = 0.799$ and $E_2 = 0.20$ and the corresponding fractions are also in excellent agreement with the simulated values (see Figure 4 legend).

The extent to which multiple molecule events influence the fit quality depends on how well the major peaks are separated. Unlike brightness effects, which should be taken care of mainly when overlapping peaks are considered (cf. Figure 3), multiple-molecule events are especially pronounced for distant peaks (e.g., a high FRET and a D-only population) and result mainly in poor fit quality. However, the extracted values of FRET efficiencies are not strongly affected, although the results of the corrected fit are much closer to the simulated values (see legend of Figure 4). Correction procedure works at least up to $N_{FCS} = 0.16$ (~20 bursts/s), as shown in Figure 4B. At this concentration, the probability of multimolecule events is 0.084,[25] and the uncorrected two-state model yields $\chi_r^2 = 37.74$.

4.4. Model-Free Deconvolution Approach and its Limitations. In this section, we investigate the properties of MEM as applied to SM-FRET data on diffusing molecules. We do not consider explicitly simultaneous model-free deconvolution and brightness correction, because combining MEM with the reference $P(F)$ method described in Section 2.3 is trivial. However, our MEM algorithm[42] depends on the assumption that the analyzed distribution (i.e., 1D PDA histogram) is a linear combination of a set of precalculated distributions. The general deconvolution method (eq 13) introduces cross correlation between the fractions x_j and distributions $P_j(F)$ and therefore would require a major modification of the MEM algorithm. Therefore, we limit ourselves to the case when at least one distribution $P_0(F)$ can be measured separately (see Figure 1A,B), although combination of MEM with eq 13 is in principle possible. In the following, we analyze several FRET-distance scenarios by MEM (the analysis of corresponding FRET efficiency histograms is shown in the Supporting Information).

A simulation of a single FRET state with $R = R_0 = 50$ Å ($E = 0.5$) and the results of MEM deconvolution are presented in Figure 5A. Model-free deconvolution extracts a somewhat broader distribution with a standard deviation of ~0.25 Å (Figure 5B). In practice, this value sets the limit of broadening that can be detected by PDA. It is clear that the value of this intrinsic broadening depends on the data quality. The data sets shown in Figure 5 contain approximately 10^5 useful 1 ms time windows with a minimum photon number of 20, which roughly corresponds to a typical measurement time of 1−2 h at $N_{FCS} = 0.01$.

Figure 5B shows the case of two states, for which the DA-distances differ by 2 Å (4% of the Förster radius $R_0 = 50$ Å). A model with two distances (eq 7) can be fitted to these data, yielding $R_1 = 49$ Å (52.5%), $R_2 = 50.9$ Å (47.5%), and $\chi_r^2 = 1.04$. Interestingly, the shot-noise broadened distribution can be also satisfactorily fitted by assuming a Gaussian distribution of distances with a mean $\langle R \rangle = 50$ Å and a standard deviation $\sigma_R = 1.02$ Å ($\chi_r^2 = 1.37$). The dip in the MEM model function indicates that the two-state model is more appropriate than a Gaussian distribution, because any distribution smoother

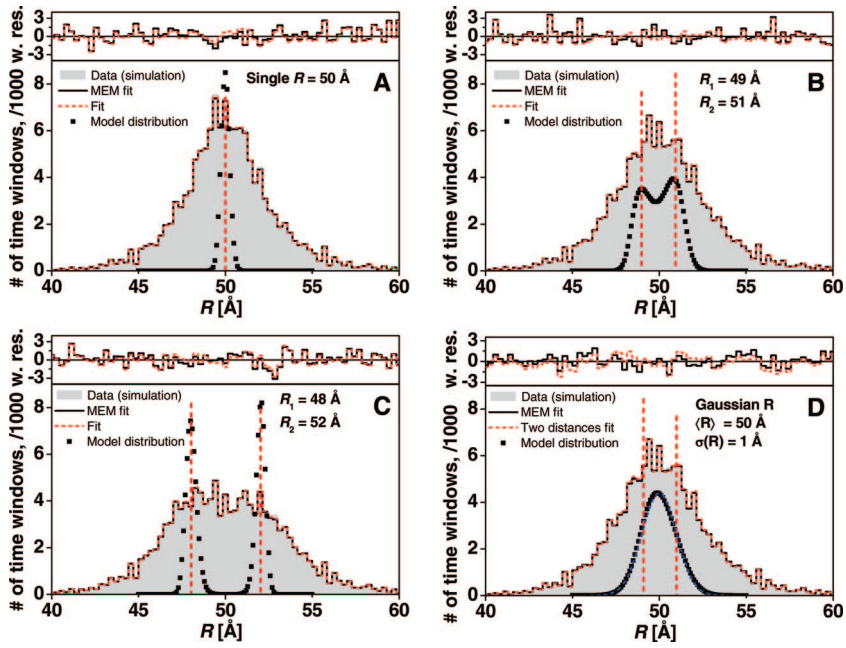

Figure 5. Model-free analyses of various FRET efficiency distributions by PDA. Histograms of simulated data are presented as gray areas, MEM distributions are shown as black squares, and the corresponding PDA model functions are represented by solid black line. One- or two-state models and the corresponding fit functions are shown as red dotted lines, and the solid blue line shows the simulated distribution. Weighted residuals are displayed above each plot. The following cases are considered. (A) One state with $R = 50$ Å. (B) Two states with $R_1 = 49$ Å and $R_2 = 51$ Å (50% each). (C) Two states with $R_1 = 48$ Å and $R_2 = 52$ Å (50% each). (D) Gaussian distribution of distances with $\langle R \rangle = 50$ Å and the standard deviation $\sigma_R = 1$ Å. Common parameters: $R_0 = 50$ Å, $\langle B_G \rangle = 2$, and $\langle B_R \rangle = 1.2$.

than that shown in Figure 5B must produce a statistically significant increase in the χ_r^2 value. It would be difficult to draw this conclusion just by comparing the variances of the underlying distributions of distances, which are essentially the same for the two-state and the Gaussian models. Note also that the peak positions of the distribution obtained by MEM closely match the expected values. If the two states are separated by 4 Å (8% of R_0), the peaks of the distance distribution recovered by MEM do not overlap anymore (Figure 5C). It can be unambiguously stated that at least two states are present. No unimodal distribution of distances can provide satisfactory fit quality.

Figure 5D illustrates the case of a Gaussian distribution of distances ($\langle R \rangle = 50$ Å; standard deviation $\sigma_R = 1$ Å; approximated by 12 discrete states), which is especially interesting to compare with the case of $R_1 = 49$ Å and $R_2 = 51$ Å (Figure 5B). A two-state model with $R_1 = 49.1$ Å (52.3%) and $R_2 = 51$ Å (47.7%) can be very well fitted to the observed distribution ($\chi_r^2 = 0.70$; Figure 5D). The underlying distributions have exactly the same variance, and there is no visible difference in shot-noise broadened distributions between these cases. Nevertheless, MEM is able to detect such a subtle difference and shows that in the latter case (Figure 5D), a smooth distribution is sufficient to fit the data. The extracted distribution of distances very closely resembles the simulated one (extracted $\sigma_R = 1.03$ Å), as shown in Figure 5D.

The analyses of simulated data reveal several limitations of PDA, which are common for many methods dealing with noisy data. In general, it is not always possible to unambiguously distinguish between several discrete states and a continuous underlying distribution. MEM can to some extent help choose the most appropriate model or at least determine the minimal number of states simply by visual analysis of deconvoluted distributions. However, the only way to resolve similar molecular states is to improve the data quality significantly. Exactly the same problem has been pointed out by Ware et al.[53,54] regarding the analysis of TCSPC data, which also obey Poisson statistics. However, there is an important difference between PDA and lifetime measurements in terms of recovered distributions of DA distances that these two methods can provide. In contrast to lifetime distributions, PDA histograms are not affected by fast (nanoseconds−microseconds) processes,[14] which in particular include local dye motions and even rotations.[55] Clearly, these processes are averaged out on the millosecond time scale. Thus, PDA perfectly complements bulk[56] and SM subensemble[3] lifetime measurements, making it possible to study slow processes and static heterogeneities selectively. If both fast dynamic and quasi-static heterogeneities affect the DA distance distribution, the width of these distributions recovered by lifetime analysis should be broader than that obtained by PDA.

4.5. Correlation between FRET Efficiencies and Fluorescence Lifetimes of the Donor. It has been shown that extra (i.e., beyond the shot noise) broadening of experimental FRET efficiency (E) histograms can be due to numerous artifacts and therefore must be interpreted with caution. For instance, poor overlap between green and red observation volumes[6,18] and significant population of the acceptor triplet state[12] may also contribute to the width of FRET efficiency distributions. Thus, it is often not obvious to what extent the observed extra broadening can be attributed to molecular structure, that is, real distance distributions, or to artifacts. Independent information on structural heterogeneities can be obtained by analyzing the shape of the distribution of donor lifetimes, calculated for individual bursts or time windows (tw), which is not sensitive to most of the optical and photophysical artifacts. To avoid confusion with distributions of fluorescence lifetimes, which one can extract from bulk or subensemble TCSPC data,[3,56] we will refer to this distribution as tw-lifetime distribution. In SM experiments with small photon numbers, tw-lifetimes are usually determined by a maximum likelihood estimator (MLE).[57,58] If the fluorescence decay is multi-exponential, MLE yields a mean lifetime, which is close to the fluorescence-weighted mean lifetime.[9] As mentioned before, the distribution of lifetimes as

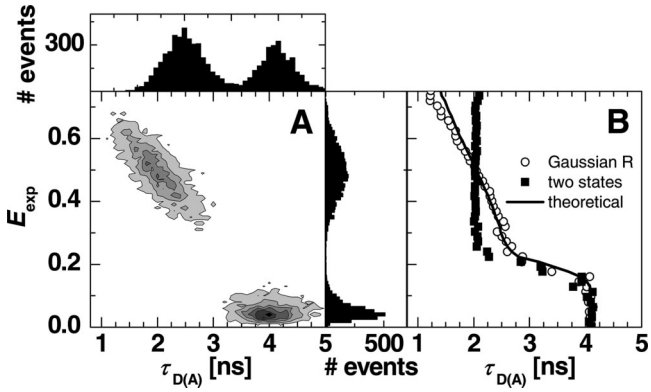

Figure 6. (A) 2D S_G/S_R signal ratio versus $\tau_{D(A)}$ histogram representing results of burstwise analysis of simulated data, mimicking a mixture of two types of molecules, FRET active and a D-only species. FRET species exhibits static Gaussian distribution of distances with $\langle R \rangle$ = 50 Å and a standard deviation σ_R = 3 Å. FRET population clearly shows a correlation between the S_G/S_R signal ratio and the donor lifetime. (B) Mean donor lifetimes calculated for tws falling into particular bins of proximity ratio histogram., plotted as a function of the proximity ratio. The distribution of tw-lifetimes obtained for static Gaussian distribution of distances (open circles; see also Figure 6A) agrees with the theoretical one (solid line) calculated according to eqs 21 and 22. If the PRH is broadened because of complex acceptor dye photophysics, for example, each state exhibits a single fluorescence lifetime (filled squares). Simulation parameters: $\langle B_G \rangle = 2$; $\langle B_R \rangle = 1.2$; R_0 = 50 Å; t_D = 1 ms.

obtained from ensemble (or subensemble) data contains information on fast dynamics (e.g., dye motions), whereas the width of tw-lifetime distributions is primarily determined by the shot noise and not by such fast processes. However, rigorous analysis of the shape of these distributions is a very challenging task because of the complex behavior of lifetime fitting algorithms. It might turn out to be easier to analyze the shape of mean TAC channel number distribution, as it is done in fluorescence intensity−lifetime distribution analysis.[59] Quantitative theory of tw-lifetime distributions is beyond the scope of this work.

Fortunately, for the qualitative tests performed here, the complete theory of the tw-lifetime distribution is not needed. The main idea is that real distance distributions (including those due to conformational dynamics) produce not only extra broad FRET efficiency histograms but also distributions of tw-lifetimes, which must be correlated with E distributions. In extreme cases, such a correlation is clearly visible on 2D burst frequency histograms of proximity ratio ($E_{exp} = S_R/(S_G+S_R)$) versus donor lifetime $\tau_{D(A)}$ (Figure 6A). Similar histograms have been introduced by Rothwell et al.[9] to distinguish FRET-related donor quenching from local dye quenching. Alternatively, one can compare the donor lifetimes calculated for bursts, which fall into left and right parts of E histogram, as proposed by Merchant et al.[60] However, such visual analyses can be often subjective. Here, we propose a simple approach to the analysis of the mean donor lifetimes calculated for each bin of a PDA histogram, which can be easily calculated theoretically and compared with the experimental data.

Let us assume that a PDA histogram of FRET efficiency $P(E_1),..., P(E_n)$ is obtained, which can be fitted with a distance model $P(R_1),..., P(R_m)$. For each distance R_j (j = 1,..., m), a shot-noise broadened PDA histogram $P(E_1|R_j),..., P(E_n|R_j)$ can be calculated separately (with or without brightness correction). To obtain the theoretical mean value of the donor lifetime for each value of E, the lifetime must be averaged over the model distance distribution

$$\langle \tau_{D(A)}(E_i) \rangle = \sum_{j=1,...,m} \tau_{D(A)}(R_j)\, P(R_j)\, P(E_i|R_j) \qquad (21)$$

where $\tau_{D(A)}(R_j)$ corresponds to the donor lifetime of a FRET pair, for which the DA distance is equal to R_j. The value of $\tau_{D(A)}(R_j)$ can be calculated according to eq 22[15]

$$\tau_{D(A)}(R_j) = \tau_{D(0)} \frac{R_j^6}{R_0^6 + R_j^6} \qquad (22)$$

where $\tau_{D(0)}$ is the fluorescence lifetime of the donor dye in the absence of acceptor and R_0 is the Förster radius. The only requirement for using eq 21 is that the lifetime fitting algorithm provides unbiased mean values in the whole range of interest and for all photon numbers. For the MLE, this requirement is fulfilled.[58] We have demonstrated our approach on FRET efficiency distributions; however, it is clear that eqs 21 and 22 are valid also for other FRET-related parameters, such as the signal ratio S_G/S_R or the proximity ratio $S_R/(S_G + S_R)$.

Examples of $\langle \tau_{D(A)} \rangle$ versus proximity ratio plots are presented in Figure 6B, which shows that these plots indeed allow one to judge on the origin of extra broadening of FRET distributions. Distributions of DA distances result in a clear correlation between the mean FRET efficiency and the donor lifetime (open circles, Figure 6B), which can be predicted according to eqs 21 and 22. Small deviations in the range of short lifetimes can be attributed to difficulties in separating fluorescence decays from the scatter contribution.[61] If it is believed that extra broadening of FRET efficiency histograms is due to an artifact, $\tau_{D(A)}(R_j)$ is not expected to obey eq 22 (filled squares). Deviations of the mean $\langle \tau_{D(A)} \rangle$ values from the theory, which are clearly visible in Figure 6B, indicate that extra broadening is likely due to acceptor dye quenching or optical artifacts, for example. In other words, these effects do not produce a correlation between the proximity ratio and the donor lifetime, and for each subpopulation, $\tau_{D(A)}$ shows no dependence on the proximity ratio (Figure 6B).

5. Concluding Remarks

The theory of PDA has been extended to take into account various possible complications that may arise when several fluorescent states contribute to measured shot-noise broadened distributions. The improved PDA theory can be applied to analyze any mixture, by using any a priori model or model-free deconvolution approach. The accuracy of the analysis and the number of free parameters are limited only by data quality. The general case of brightness variations is considered without need to discard photons. In most of the typical cases, brightness correction does not require any additional knowledge of the system or new model parameters and can be built into a fitting procedure. Correction of the PDA model function for the presence of multiple-molecule events allows one to measure at above SM concentrations when necessary and to avoid artifacts due to long measurement time. PDA is now able to quantitatively analyze complex FRET efficiency and anisotropy distributions with multiple states and different brightness values. Tools such as MEM and combined mean donor fluorescence lifetime analysis have been developed to distinguish between states with fixed distances and a distribution of distances.

Acknowledgment. S.K. is grateful to the Alexander von Humboldt Foundation for financial support. This study was supported by the Deutsche Forschungsgemeinschaft (DFG) via the SFB 590 Inherent and adaptive differentiation and the Volkswagen foundation Grant I/78 837.

Appendix: On the Derivation of Eqs 14–16

Equation 14

The FRET efficiency is given by

$$E = \frac{F_A/\Phi_{FA}}{F_D/\Phi_{FD} + F_A/\Phi_{FA}} \quad \text{(A1)}$$

where F_D and F_A are the fluorescence intensities of the donor and the acceptor, respectively. These intensities are related to the observed green and red signals according to eq A2

$$F_G = g_G F_D \propto (1-E)\Phi_{FD} g_G \quad \text{(A2a)}$$

$$F_R = g_R F_A + \alpha g_G F_D \propto E\Phi_{FA} g_R + \alpha F_G \quad \text{(A2b)}$$

By assuming for convenience that the brightness of D-only species is equal to 1, the relative brightness of a FRET species ($F_G + F_R$) can be expressed as

$$Q = \frac{F_G + F_R}{Q(E=0)} = \frac{(1-E)\Phi_{FD} g_G (1+\alpha) + E\Phi_{FA} g_R}{\Phi_{FD} g_G (1+\alpha)} \quad \text{(A3)}$$

Eq A3 is equivalent to eq 14.

Equation 15

The fluorescence intensities measured by two polarization channels, $F_{||}$ and $F_{\perp}$, are given by

$$F_{||} = g_{||}(I_{||}(1-l_1) + l_1 I_{\perp}) \quad \text{(A4a)}$$

$$F_{\perp} = g_{\perp}(l_2 I_{||} + (1-l_2) I_{\perp}) \quad \text{(A4b)}$$

In eq A4, $I_{||}$ and $I_{\perp}$ denote the true intensities of the polarized components, which would be measured with a perfect setup. In reality, the fluorescence is partly depolarized by the microscope objective, as described by introducing the correction factors l_1 and l_2 in eq A4.[38] The difference in the detection efficiencies of the two channels is taken into account by introducing the instrumental G factor ($G = g_{\perp}/g_{||}$).

The total brightness $Q = F_{||} + F_{\perp}$ has to be expressed as a function of the fluorescence anisotropy (eq A6), that is,

$$Q = F_{||} + F_{\perp} \propto I_{||} - l_1 I_{||} + l_1 I_{\perp} + G(I_{\perp} + l_2 I_{||} - l_2 I_{\perp}) \quad \text{(A5)}$$

$$r = \frac{I_{||} - I_{\perp}}{I_{||} + 2I_{\perp}} \quad \text{(A6)}$$

Eq A6 is equivalent to the well-known expressions (eq A7)

$$I_{||} = I_0(1+2r)/3 \quad \text{(A7a)}$$

$$I_{\perp} = I_0(1-r)/3 \quad \text{(A7b)}$$

where I_0 denotes the total fluorescence intensity ($I_0 = I_{||} + 2I_{\perp}$). As assumed in the first part of Section 2.6, the total intensity does not change; that is, I_0 is the same for all species and can be omitted. Substitution of eq A7 into eq A5 and further normalization to $Q(r=0) = 1$ yields eq 15.

Equation 16

By rearranging Perrin equation,[15] we obtain

$$\frac{\tau}{\rho} = \frac{r_0}{r} - 1 \quad \text{(A8)}$$

In addition, the brightness is proportional to the fluorescence quantum yield, which is by our assumption proportional to the fluorescence lifetime. If the rotational correlation time ρ is the same for all species, the right-hand part of eq A8 becomes proportional to the total brightness and should thus appear as an additional factor in eq 16 as compared to eq 15.

References and Notes

(1) Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6264–6268.
(2) Deniz, A. A.; Laurence, T. A.; Dahan, M.; Chemla, D. S.; Schultz, P. G.; Weiss, S. *Annu. Rev. Phys. Chem.* **2001**, *52*, 233–253.
(3) Laurence, T. A.; Kong, X. X.; Jager, M.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17348–17353.
(4) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743–747.
(5) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.
(6) Antonik, M.; Felekyan, S.; Gaiduk, A.; Seidel, C. A. M. *J. Phys. Chem. B* **2006**, *110*, 6970–6978.
(7) Widengren, J.; Kudryavtsev, V.; Antonik, M.; Berger, S.; Gerken, M.; Seidel, C. A. M. *Anal. Chem.* **2006**, *78*, 2039–2050.
(8) Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Haustein, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15516–15521.
(9) Rothwell, P. J.; Berger, S.; Kensch, O.; Felekyan, S.; Antonik, M.; Wöhrl, B. M.; Restle, T.; Goody, R. S.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1655–1660.
(10) Ha, T.; Hohng, S.; Joo, C.; Ha, T. *Biophys. J.* **2004**, *87*, 1328–1337.
(11) Ha, T. *Methods* **2001**, *25*, 78–86.
(12) Vogelsang, J.; Doose, S.; Sauer, M.; Tinnefeld, P. *Anal. Chem.* **2007**, *79*, 7367–7375.
(13) Tinnefeld, P.; Sauer, M. *Angew. Chem., Int. Ed.* **2005**, *44*, 2642–2671.
(14) Gopich, I.; Szabo, A. *J. Chem. Phys.* **2005**, 014707.
(15) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 2nd ed.; Kluwer Academic/ Plenum Publishers: New York, 1999.
(16) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T. J.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3670–3675.
(17) Watkins, L. P.; Chang, H. Y.; Yang, H. *J. Phys. Chem. A* **2006**, *110*, 5191–5203.
(18) Nir, E.; Michalet, X.; Hamadani, K. M.; Laurence, T. A.; Neuhauser, D.; Kovchegov, Y.; Weiss, S. *J. Phys. Chem. B* **2006**, *110*, 22103–22124.
(19) Kask, P.; Palo, K.; Fay, N.; Brand, L.; Mets, Ü.; Ullmann, D.; Jungmann, J.; Pschorr, J.; Gall, K. *Biophys. J.* **2000**, *78*, 1703–1713.
(20) Kalinin, S.; Felekyan, S.; Antonik, M.; Seidel, C. A. M. *J. Phys. Chem. B* **2007**, *111*, 10253–10262.
(21) Kask, P.; Palo, K.; Ullmann, D.; Gall, K. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13756–13761.
(22) Hess, S. T.; Webb, W. W. *Biophys. J.* **2002**, *83*, 2300–2317.
(23) Enderlein, J.; Gregor, I.; Patra, D.; Fitter, J. *Curr. Pharm. Biotechnol.* **2004**, *5*, 155–161.
(24) Gregor, I.; Patra, D.; Enderlein, J. *ChemPhysChem* **2005**, *6*, 164–170.
(25) Fries, J. R.; Brand, L.; Eggeling, C.; Köllner, M.; Seidel, C. A. M. *J. Phys. Chem.A* **1998**, *102*, 6601–6613.
(26) Orte, A.; Clarke, R.; Balasubramanian, S.; Klenerman, D. *Anal. Chem.* **2006**, *78*, 7707–7715.
(27) Skilling, J.; Bryan, R. K. *Mon. Not. R. Astr. Soc* **1984**, *211*, 111–124.
(28) Livesey, A. K.; Skilling, J. *Acta Crystallogr., Sect. A* **1985**, *41*, 113–122.
(29) Brochon J. C. *Methods of Enzymology*; Academic Press: New York, 1994; Vol. 240, pp 262−311.
(30) Qian, H.; Elson, E. L. *Biophys. J.* **1990**, *57*, 375–380.
(31) Kudryavtsev, V.; Felekyan, S.; Wozniak, A. K.; König, M.; Sandhagen, C.; Kühnemuth, R.; Seidel, C. A. M.; Oesterhelt, F. *Anal. Bioanal. Chem.* **2007**, *387*, 71–82.
(32) Kapanidis, A. N.; Lee, N. K.; Laurence, T. A.; Doose, S.; Margeat, E.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8936–8941.
(33) Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S. *Biophys. J.* **2005**, *88*, 2939–2953.
(34) Lee, N. K.; Kapanidis, A. N.; Koh, H. R.; Korlann, Y.; Ho, S. O.; Kim, Y.; Gassman, N.; Kim, S. K.; Weiss, S. *Biophys. J.* **2007**, *92*, 303–312.
(35) Mandel, L.; Sudarshan, E. C. G.; Wolf, E. *Proc. Phys. Soc.* **1964**, *84*, 435–444.
(36) Rigler, R.; Mets, Ü. *SPIE Proc.* **1992**, *1921*, 239–248.
(37) Laurence, T. A.; Kwon, Y.; Yin, E.; Hollars, C. W.; Camarero, J. A.; Barsky, D. *Biophys. J.* **2007**, *92*, 2184–2198.
(38) Koshioka, M.; Saski, K.; Masuhara, H. *Appl. Spectrosc.* **1995**, *49*, 224–228.
(39) Valeur, B. *Molecular Fluorescence: Principles and Applications*; Wiley-VCH Verlag: Weinheim, 2002.

(40) Gull, S. F.; Skilling, J. *Quantified Maximum Entropy MemSys5 User's Manual*; Maximum Entropy Data Consultants Ltd.: Suffolk, UK, 1999; Vol. 5[1.2].

(41) Soong, T. T. *Fundamentals of Probability and Statistics for Engineers*; John Wiley & Sons: West Sussex, England, 2004.

(42) Vinogradov, S. A.; Wilson, D. F. *Appl. Spectrosc.* **2000**, *54*, 849–855.

(43) Saffarian, S.; Li, Y.; Elson, E. L.; Pike, L. J. *Biophys. J.* **2007**, *93*, 1021–1031.

(44) Neubauer, H.; Gaiko, N.; Berger, S.; Schaffer, J.; Eggeling, C.; Tuma, J.; Verdier, L.; Seidel, C. A. M.; Griesinger, C.; Volkmer, A. *J. Am. Chem. Soc.* **2007**, *129*, 12746–12755.

(45) Kühnemuth, R.; Seidel, C. A. M. *Single Mol.* **2001**, *2*, 251–254.

(46) Enderlein, J.; Robbins, D. L.; Ambrose, W. P.; Goodwin, P. M.; Keller, R. A. *J. Phys. Chem. B* **1997**, *101*, 3626–3632.

(47) Laurence, T. A.; Kapanidis, A. N.; Kong, X. X.; Chemla, D. S.; Weiss, S. *J. Phys. Chem. B* **2004**, *108*, 3051–3067.

(48) Dix, J. A.; Hom, E. F. Y.; Verkman, A. S. *J. Phys. Chem. B* **2006**, *110*, 1896–1906.

(49) Magde, D.; Elson, E. L.; Webb, W. W. *Phys. Rev. Lett.* **1972**, *29*, 705–708.

(50) Felekyan, S.; Kühnemuth, R.; Kudryavtsev, V.; Sandhagen, C.; Becker, W.; Seidel, C. A. M. *Rev. Sci. Instrum.* **2005**, *76*, 083104−1083104−14.

(51) Eggeling, C.; Widengren, J.; Rigler, R.; Seidel, C. A. M. *Anal. Chem.* **1998**, *70*, 2651–2659.

(52) Schaffer, J.; Volkmer, A.; Eggeling, C.; Subramaniam, V.; Striker, G.; Seidel, C. A. M. *J. Phys. Chem. A* **1999**, *103*, 331–336.

(53) James, D. R.; Ware, W. R. *Chem. Phys. Lett.* **1985**, *120*, 455–459.

(54) James, D. R.; Ware, W. R. *Chem. Phys. Lett.* **1986**, *126*, 7–11.

(55) Isaksson, M.; Norlin, N.; Westlund, P. O.; Johansson, L. B. A. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1941–1951.

(56) Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1807–1811.

(57) Brand, L.; Eggeling, C.; Zander, C.; Drexhage, K. H.; Seidel, C. A. M. *J. Phys. Chem. A* **1997**, *101*, 4313–4321.

(58) Maus, M.; Cotlet, M.; Hofkens, J.; Gensch, T.; De Schryver, F. C.; Schaffer, J.; Seidel, C. A. M. *Anal. Chem.* **2001**, *73*, 2078–2086.

(59) Palo, K.; Brand, L.; Eggeling, C.; Jäger, S.; Kask, P.; Gall, K. *Biophys. J.* **2002**, *83*, 605–618.

(60) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528–1533.

(61) Eggeling, C.; Fries, J. R.; Brand, L.; Günther, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1556–1561.

JP711942Q