Statistics of Single-Molecule Detection

Jörg Enderlein,* David L. Robbins, W. Patrick Ambrose, Peter M. Goodwin, and Richard A. Keller

CST-1, MS M888, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Received: October 21, 1996; In Final Form: February 26, 1997[⊗]

A new mathematical approach for the calculation of the photon detection statistics in single-molecule detection experiments is presented. The approach is based on a path integral representation of the involved probability distributions. A numerical Monte Carlo sampling method for computing the involved path integrals is given. The general ideas are applied to the modeling of measured burst size distributions in single-molecule detection experiments with B-phycoerythrin.

1. Introduction

The detection of single fluorescent molecules in liquids at room temperature has made big advances since the first successful detection of a molecule labeled with multiple fluorophores by Hirschfeld in 1976.[1] Single-molecule detection (SMD) capabilities have improved to the point where several groups have demonstrated the detection of molecules containing a single fluorophore as the molecule transits a focused laser beam.[2−15] There are applications envisioned where detection of single molecules (not ensembles) is necessary; for example, single-molecule detection will be used in DNA sequencing,[16−20] sizing of DNA fragments,[21−23] genetic screening,[24,25] diagnostics,[26,27] the study of single-molecule chemical kinetics,[28,29] and the detection of minute amounts of substances.[30,31] Single-molecule detection is of fundamental importance since processes are revealed that are not observable in bulk measurements, for example, direct observations of the interconversion of chemical conformations and surprising single-quantum system behavior (emission and photobleaching behavior) in proteins with multiple chromophores.[33]

Detection of single molecules in liquids at room temperature is challenging because of several effects: the surrounding liquid and possible impurities cause high background signals, diffusion causes molecules to exit or miss the detection volume, and photobleaching limits the overall number of detectable photons from one molecule. Thus, for detecting single molecules under these circumstances and obtaining a maximum amount of information from single-molecule experiments, it is desirable to have a sufficiently complete model of the experiment at hand for predicting the experimental outcome and comparing it with the measured data. An efficient model will allow the experimental conditions to be modified in such a way that one achieves optimal detection results and gains information about physical parameters of the studied molecules.

In this paper we describe a general and first principles approach for modeling the photon detection statistics in single-molecule detection experiments in a fluid flow. Using a path integral approach provides a general framework for efficient computation of photon detection in a single-molecule experiment, independent of any particular starting assumptions about the experimental setup. The formulation is exact for individual molecules, and any experimental conditions can be incorporated.

2. Theory

2.1. Path Integral Formulation. In describing a single-molecule detection experiment, the main problem is to obtain an expression for the statistics of photon detection within a given interval of time. Knowing the statistics, one can derive any experimentally measurable characteristic of interest such as the autocorrelation function or the burst size distribution. In refs 34 and 35, approximate and exact analyses of single-molecule detection experiments were presented. In these works, the molecule diffusion was neglected and a rather complicated mathematical approach was used. Here, we present a direct mathematical approach that includes the effects of molecule diffusion and photobleaching. As will be demonstrated below, the most convenient tool for describing photon detection statistics is a path integral representation. In this approach, the probability distribution of detecting photons emitted by single molecules is given by an infinite sum over all possible paths a single molecule can follow. Thus, the path integral description is very close to the physics of SMD. Additionally, a path integral representation is an excellent starting point for designing efficient Monte Carlo simulations of the problem, as will be shown in subsection 2.2.

For concreteness, we will use a particular experimental geometry to describe the path integral method. The main experimental setup, used in this paper, is shown in Figure 1. Single molecules are injected from a small capillary into the sheath flow of a surrounding larger capillary. Downstream, a laser beam is focused into the capillary for exciting fluorescence from the molecules passing through the probe volume. The resulting bursts of fluorescence photons are detected by appropriate, large numerical aperture (NA) detection optics with the optical axis (*x*-axis) perpendicular to the flow direction (*z*-axis) and to the laser beam (*y*-axis).

The main conceptual problem in calculating the probability of detecting a given number of photons within a given interval of time is that the fluorescent molecules are generally photolabile and often photobleach before exiting the probe volume (see also refs 36−38). To show the origin of this difficulty, we have to look closely at the time evolution of the probability that a single molecule is not yet photobleached while passing through the laser beam. As long as no photon is detected, this probability decays exponentially in proportion to the total light received by the molecule. At the moment when a photon is detected, this probability jumps to a higher value, corresponding to the chance that the detected photon was emitted by the molecule

* To whom correspondence should be addressed.
[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1997.

S1089-5647(96)03261-0 CCC: $14.00

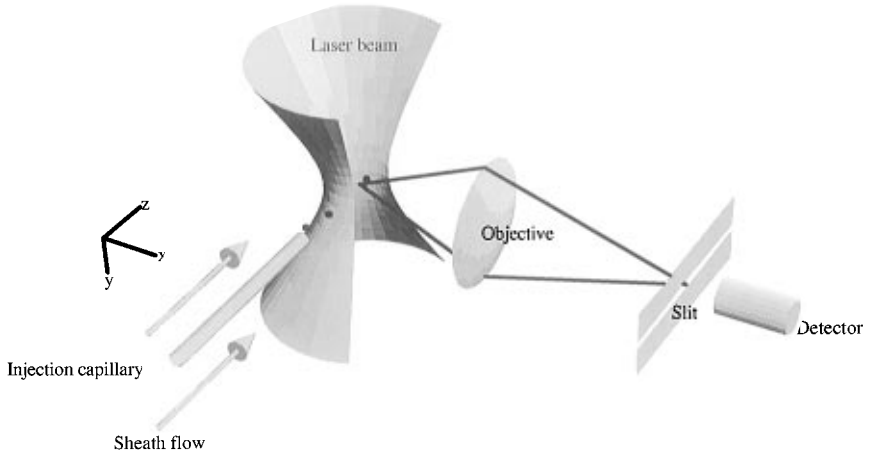

Figure 1. Sheath flow SMD experimental schematic.

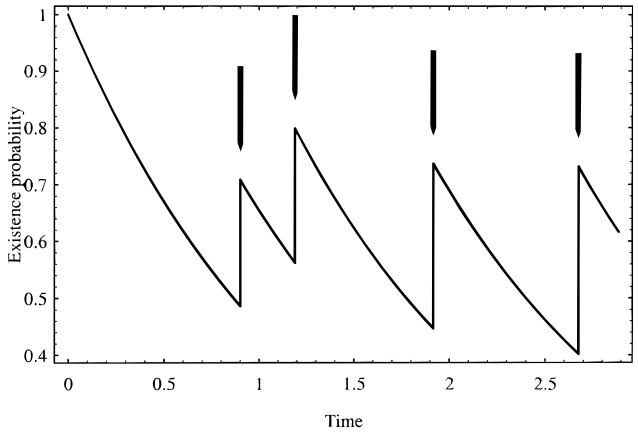

Figure 2. Model calculation of the existence probability of a single molecule due to photobleaching. Each arrow indicates the detection of a photon. The photobleaching rate was assumed to be 0.8 per time unit, the ratio of the probability of detecting a background photon to the probability of detecting a fluorescence photon was set to be 0.2 per time unit.

and not caused by the background, see Figure 2. As can be seen in this figure, after a photon detection event, the existence probability does not jump back to unity, which is due to a nonvanishing probability that the detected photon was caused by background. Thus, every detected photon will cause a jump in the probability that the molecule is not yet photobleached, and thus, this probability will depend on the complete history of the detection process.

In a classical sense, there is an analogy to the quantum-mechanical Schrödinger cat problem. In the Schrödinger cat problem, the cat is placed in a box with a device to release poison triggered by a single radioactive decay event. If we receive no information about the state of the cat, then it evolves into a mixed quantum-mechanical state. Any time we hear the cat meow, then the state of the cat collapses to the "live-state". Thus, the evolution of the quantum-mechanical state is similar to the existence probability of the single molecule—every time we detect a photon the existence probability suddenly changes. The difference between the quantum-mechanical case and the existence probability is, of course, that in the first case the real physical state of the system is changing, whereas in the latter case only our knowledge about the system changes.

The time evolution of the existence probability is crucial in the calculation of the probability to detect a given number of photons during a given interval of time. Because of photobleaching, it is not possible to derive a simple expression for the probability of detecting a given number of photons in a given interval of time, and especially there is no hope to obtain some kind of integral (Chapman−Kolmogorov) or differential (Fokker-Planck) equation (ref 39) for this detection probability.

The probability $P(N,t_b,t_a)$ of detecting N photons within a given time interval $\{t_a,t_b\}$ can be constructed in the following way. Since photobleaching is the main cause of computational difficulties, one divides all possible cases of molecule passages into two classes, according to whether or not the molecule is photobleached within the given interval of time $\{t_a, t_b\}$.

Let us first assume that the molecule is photobleached within the small interval δt exactly at time $t_{bl} < t_b$. Second, assume that the molecule is at position $\vec{r}_a$ at initial time t_a and at position $\vec{r}_{bl}$ at bleaching time t_{bl}. Third, one considers the case that the N photons are detected within small time intervals δt at the intermediate times $t_1, ..., t_N$; $t_a \le t_1 < t_2 < ... < t_N < t_{bl}$, when the molecule is at the intermediate positions $\vec{r}_1, \vec{r}_2, ..., r_N$. The probability for such a combined event is given by

$$\delta tV_{bl}(\vec{r}_{bl})K(t_{bl},t_N,\vec{r}_{bl},\vec{r}_N)\delta tV_f(\vec{r}_N)...K(t_2,t_1,\vec{r}_2,\vec{r}_1)\delta tV_f(\vec{r}_1) \times K(t_1,t_a,\vec{r}_1,\vec{r}_a)p_0(t_a,\vec{r}_a)\delta\vec{r}_a \quad (1)$$

where $p_0(t_a,\vec{r}_a)\delta\vec{r}_a$ is the probability to find a molecule in the small volume $\delta\vec{r}_a$ at position $\vec{r}_a$ at time t_a. $K(t_{i+1},t_i,\vec{r}_{i+1},\vec{r}_i)$ is the probability that no photon is detected and no photobleaching happens between t_i and t_{i+1}, if the molecule was at position $\vec{r}_i$ at time t_i and at position $\vec{r}_{i+1}$ at time t_{i+1} and $\delta tV_f(\vec{r})$ and $\delta tV_{bl}(\vec{r})$ are the probabilities to detect a photon and to photobleach the molecule within time δt, respectively, if the molecule is at position $\vec{r}$. Integrating this expression over all possible bleaching times $t_a \le t_{bl} \le t_b$, and over all possible intermediate times t_i, $t_a \le t_1 < t_2 < ... < t_N < t_{bl}$, and all positions $\vec{r}_a, \vec{r}_1, ..., \vec{r}_N, \vec{r}_{bl}$ yields the probability of detecting N photons within $\{t_a,t_b\}$ if the molecule is photobleached within this time interval.

To obtain the remaining contribution of molecules, which are photobleached at times $t_{bl} > t_b$, one has to add a corresponding integral over the expression

$$K(t_b,t_N,\vec{r}_b,\vec{r}_N)\delta tV_f(\vec{r}_N)...\delta tV_f(\vec{r}_2)K(t_2,t_1,\vec{r}_2,\vec{r}_1)\delta tV_f(\vec{r}_1) \times K(t_1,t_a,\vec{r}_1,\vec{r}_a)p_0(t_a,\vec{r}_a)\delta\vec{r}_a \quad (2)$$

Thus, it remains to find explicit expressions for p_0, K, V_f and V_{bl}. The path of the molecule through the detection volume is governed by the sheath flow and molecular diffusion. To start, let us consider the diffusion of a single molecule in a flowing solution. If the flow velocity is denoted by $\vec{v}$ and the diffusion constant of the molecule by D, then the probability to find the molecule at position $\vec{r} + \delta\vec{r}$ at time $t + \delta t$ when it was at position $\vec{r}$ at time t is given by

$$\frac{1}{(4\pi D\delta t)^{3/2}}\exp\left(-\frac{(\delta\vec{r}-\vec{v}\delta t)^2}{4D\delta t}\right) \tag{3}$$

Furthermore, if the position dependent laser intensity within the capillary is given by the function $I(\vec{r})$ (photons per area per time), then the probability to detect a fluorescence photon within the sufficiently small time interval δt is given by

$$\delta t V_{\mathrm{f}}(\vec{r}) = \delta t \eta(\vec{r})\Phi_{\mathrm{f}}\sigma I(\vec{r}) \tag{4}$$

where $\eta(\vec{r})$ is the position dependent collection and detection efficiency of the detection optics and electronics, Φ_{f} is the fluorescence quantum yield, and σ is the absorption cross section of the molecule, respectively. Here, the intensity is assumed low and background counts are neglected. We will consider nonzero background at the end of this subsection. Similarly, the probability of photobleaching the molecule within that time interval δt is equal to

$$\delta t V_{\mathrm{bl}}(\vec{r}) = \delta t \Phi_{\mathrm{bl}}\sigma I(\vec{r}) \tag{5}$$

where now Φ_{bl} denotes the photobleaching quantum yield.

By combining the previous ideas, the probability of *not* detecting a fluorescence photon and *not* photobleaching the molecule within the finite time interval $\{t_{\mathrm{a}},t_{\mathrm{b}}\}$ is given by

$$\left(\prod_{k=1}^{M-1}\int\frac{\mathrm{d}^3\vec{r}_k}{(4\pi D\delta t)^{3/2}}\right)\exp\left[-\sum_{k=0}^{M}\left\{\frac{[(\vec{r}_{k+1}-\vec{r}_k)/\delta t-\vec{v}]^2}{4D}+V_{\mathrm{f}}\left(\frac{\vec{r}_k+\vec{r}_{k+1}}{2}\right)+V_{\mathrm{bl}}\left(\frac{\vec{r}_k+\vec{r}_{k+1}}{2}\right)\right\}\delta t\right] \tag{6}$$

whereby $\vec{r}_0 \equiv \vec{r}_{\mathrm{a}}$ is the supposed initial position and $\vec{r}_M \equiv \vec{r}_{\mathrm{b}}$ is the final position of the molecule, M is given by $(t_{\mathrm{b}}-t_{\mathrm{a}})/\delta t$, and the approximation $1-\delta t V(\vec{r}) \approx \exp(-\delta t V(\vec{r}))$ was employed. Letting the time interval δt tend to zero and M to infinity results in the following path integral (see, e.g., refs 40 and 41) for K:

$$K(t_{\mathrm{b}},t_{\mathrm{a}},\vec{r}_{\mathrm{b}},\vec{r}_{\mathrm{a}}) = \int_{(\vec{r}_{\mathrm{a}},t_{\mathrm{a}})\sim(\vec{r}_{\mathrm{b}},t_{\mathrm{b}})} D\vec{r}(t)\times \exp\left[-\int_{t_{\mathrm{a}}}^{t_{\mathrm{b}}}\mathrm{d}t\left(\frac{[\dot{\vec{r}}(t)-\vec{v}]^2}{4D}+V_{\mathrm{f}}[\vec{r}(t)]+V_{\mathrm{bl}}[\vec{r}(t)]\right)\right] \tag{7}$$

where $\int_{(\vec{r}_{\mathrm{a}},t_{\mathrm{a}})\sim(\vec{r}_{\mathrm{b}},t_{\mathrm{b}})} D\vec{r}(t)$ denotes the integration over all possible paths starting at position $\vec{r}_{\mathrm{a}}$ at time t_{a} and ending at position $\vec{r}_{\mathrm{b}}$ at time t_{b}, and a dot over a symbol denotes its differentiation with respect to time.

Let us now introduce the following generating function $Z_\epsilon(t_{\mathrm{b}},t_{\mathrm{a}})$

$$Z_\epsilon(t_{\mathrm{b}},t_{\mathrm{a}}) = \int_{t_{\mathrm{a}}}^{t_{\mathrm{b}}}\mathrm{d}t_{\mathrm{bl}}\int D\vec{r}(t)V_{\mathrm{bl}}[\vec{r}(t_{\mathrm{bl}})]\times \exp\left[-\int_{t_{\mathrm{a}}}^{t_{\mathrm{bl}}}\mathrm{d}t\left(\frac{[\dot{\vec{r}}(t)-\vec{v}]^2}{4D}+\epsilon V_{\mathrm{f}}[\vec{r}(t)]+V_{\mathrm{bl}}[\vec{r}(t)]\right)\right]\times p_0[t_{\mathrm{a}},\vec{r}(t_{\mathrm{a}})] + \int D\vec{r}(t)\exp\left[-\int_{t_{\mathrm{a}}}^{t_{\mathrm{b}}}\mathrm{d}t\left(\frac{[\dot{\vec{r}}(t)-\vec{v}]^2}{4D}+\epsilon V_{\mathrm{f}}[\vec{r}(t)]+V_{\mathrm{bl}}[\vec{r}(t)]\right)\right]p_0[t_{\mathrm{a}},\vec{r}(t_{\mathrm{a}})] \tag{8}$$

whereby the path integration is now extended over all possible paths with arbitrary initial and end points. Then, by direct comparison with the above expressions 1 and 2, one can see that the probability of detecting N photons within time interval $\{t_{\mathrm{a}},t_{\mathrm{b}}\}$ is given by

$$P(N,t_{\mathrm{b}},t_{\mathrm{a}}) = \lim_{\epsilon\to 1}\frac{(-1)^N}{N!}\frac{\mathrm{d}^N Z_\epsilon(t_{\mathrm{b}},t_{\mathrm{a}})}{\mathrm{d}\epsilon^N} \tag{9}$$

In the case of a nonzero background all of the above considerations remain valid, except for the following replacement in eq 8:

$$\int_{t_{\mathrm{a}}}^{t_{\mathrm{bl,b}}}\mathrm{d}t\,\epsilon V_{\mathrm{f}}[r(t)] \to \epsilon\tau_{\mathrm{b}}V_{\mathrm{bg}} + \int_{t_{\mathrm{a}}}^{t_{\mathrm{bl,b}}}\mathrm{d}t\,\epsilon V_{\mathrm{f}}[r(t)] \tag{10}$$

where V_{bg} is the mean background rate, and τ_{b} is the burst duration. This burst duration is itself a function of the molecule's path, its bleaching time, and is dependent also on the burst sifting algorithm used for data processing. Especially this latter dependency complicates the exact incorporation of background effects into model calculations. In cases with a narrow burst duration distribution, a good approximation is making τ_{b} in eq 10 equal to the mean burst duration. Then the main effects of the background are a stretching of the BSD to higher burst sizes and an increase in its width. Another problem is that a major background source in many realistic experiments is fluctuating fluorescence from contaminants. Without detailed knowledge about the nature of these contaminants, a model cannot account for their influence on the BSD.

2.2. Monte Carlo Sampling. In general, it is impossible to find an analytical expression for $Z_\epsilon(t_{\mathrm{b}},t_{\mathrm{a}})$ and thus for $P(N,t_{\mathrm{b}},t_{\mathrm{a}})$. The most convenient numerical method to calculate $P(N,t_{\mathrm{b}},t_{\mathrm{a}})$ is by means of a Monte Carlo simulation. The basic difficulty in calculating eq 8 is that the path integral runs over an infinite number of paths. The idea of a Monte Carlo calculation is to choose these paths randomly and then to perform the remaining integrations numerically. After sampling over a large number of different paths one expects to gain a sufficiently precise approximation of the true value of $P(N,t_{\mathrm{b}},t_{\mathrm{a}})$.

The course of the calculation will then be as follows. The time interval $\{t_{\mathrm{a}},t_{\mathrm{b}}\}$ is divided into discrete but sufficiently small time steps Δt, so that the laser light intensity and detection efficiency that the moving molecule experiences does not change significantly within time Δt. For the discretized times $\{t_{\mathrm{a}}, t_{\mathrm{a}}+\Delta t, t_{\mathrm{a}}+2\Delta t, \ldots, t_{\mathrm{b}}\}$ and fixed initial position $\vec{r}_0 \equiv \vec{r}_{\mathrm{a}}$, a given number Q of random paths $\{\vec{r}_0, \vec{r}_1, \ldots, \vec{r}_{(t_{\mathrm{b}}-t_{\mathrm{a}})/\Delta t}\}$ are generated by means of the probability distribution

$$p(\vec{r}_{k+1}-\vec{r}_k) = \frac{1}{[4\pi D\Delta t]^{3/2}}\exp\left[-\frac{(\vec{r}_{k+1}-\vec{r}_k-\vec{v}\Delta t)^2}{4D\Delta t}\right] \tag{11}$$

For these paths, the integrals occurring in eq 8 are approximated by the following sums

$$V_{\mathrm{bl,f}}(m\Delta t) = \sum_{k=0}^{m}V_{\mathrm{bl,f}}(\vec{r}_k)\Delta t \tag{12}$$

Using the abbreviation $T = t_{\mathrm{b}} - t_{\mathrm{a}}$, the values

$$\frac{1}{Q}\sum_{k=0}^{T/\Delta t-1}\Delta t\, V_{\mathrm{bl}}(\vec{r}_{k+1})V_{\mathrm{f}}^{\,N}(k\Delta t)\exp[-V_{\mathrm{f}}(k\Delta t)-V_{\mathrm{bl}}(k\Delta t)]\times p_0(t_{\mathrm{a}},\vec{r}_0) + \frac{1}{Q}V_{\mathrm{f}}^{\,N}(T)\exp[-V_{\mathrm{f}}(T)-V_{\mathrm{bl}}(T)]p_0(t_{\mathrm{a}},\vec{r}_0) \tag{13}$$

are added to the values of $P(N,t_{\mathrm{b}},t_{\mathrm{a}})$. By repeating this process a sufficient number of times and for all possible initial positions

$\vec{r}_0$, the true value of $P(N,t_b,t_a)$ can be approximated with arbitrary precision. Of course, for approximating the integrals of eq 8, higher order methods than the simple sums in the last two expressions can be used. The Monte Carlo method described above can be extremely efficient, especially if $p_0(t,\vec{r})$ is very localized and diffusion is small. Then a small number of paths per initial point $\vec{r}_0$ can be enough to obtain a very precise result.

2.3. Example: Burst Size Distribution. The results of eqs 8 and 9 are directly applicable to the calculation of the burst size distribution (BSD) in a SMD experiment. Let us define a burst as the number of detected fluorescence photons within time T, when the molecule starts at $t = 0$ at a fixed distance before the detection region, and T is chosen sufficiently large so that the molecule has completely passed the detection region after time T. In this case, eqs 8 and 9 are directly applicable. Consider the special case of negligible diffusion compared to the sample stream flow. Letting the diffusion constant D tend to zero in eq 8, the path integral representation transforms into

$$Z_\epsilon(T) = \int_0^T \mathrm{d}t_{bl} \int \mathrm{d}\vec{r}_0 V_{bl}[\vec{r}_0 + \vec{v}t_{bl}] \times \exp[-\int_0^{t_{bl}} \mathrm{d}t(\epsilon V_f(\vec{r}_0 + \vec{v}t) + V_{bl}(\vec{r}_0 + \vec{v}t))] p_0(\vec{r}_0) + \int \mathrm{d}\vec{r}_0 \exp[-\int_0^T \mathrm{d}t(\epsilon V_f(\vec{r}_0 + \vec{v}t) + V_{bl}(\vec{r}_0 + \vec{v}t))] p_0(\vec{r}_0) \quad (14)$$

A special case occurs if the collection and detection efficiency is approximately constant, $\eta(\vec{r}) \approx \eta_0$, over the detection region (defined by the intersection of the sample stream with the laser beam). In this case, eq 14 can be simplified further. Then, $V_f(\vec{r}) = \eta_0(\Phi_f/\Phi_{bl})V_{bl}(\vec{r})$, and the integration over t_{bl} can be done explicitly, yielding

$$Z_\epsilon(T) = \int \mathrm{d}\vec{r}_0 \frac{1 + \epsilon\lambda \exp[-(1+\epsilon\lambda)V_{bl}(T,\vec{r}_0)]}{(1+\epsilon\lambda)} p_0(\vec{r}_0) \quad (15)$$

where the abbreviations $\lambda = \eta_0\Phi_f/\Phi_{bl}$ and

$$V_{bl}(T,\vec{r}_0) = \int_0^T \mathrm{d}t\, V_{bl}(\vec{r}_0 + \vec{v}t) \quad (16)$$

were used. For a high probability of photobleaching during time $T(V_{bl}(T,\vec{r}_0) \gg 1)$ but sufficiently large λ ($\lambda \gg 1$), one obtains from eq 15 via eq 9 for small N

$$\ln P(N,T) \approx -\ln\lambda - N/\lambda \quad (17)$$

Thus, the BSD decays exponentially with the decay constant proportional to the ratio $1/\lambda$ for small numbers of detected photons.

3. Application to Experiment: Detection of Single B-Phycoerythrin Molecules

3.1. Experimental Setup and Data Processing. The experimental apparatus used to collect the B-phycoerythrin data is similar to that previously used for DNA fragment sizing.[22,21] Excitation was accomplished with a continuous wave Ar^+ laser operated at 514.5 nm. The output was attenuated with a polarizer/half-wave plate assembly to 0.5 mW and focused to a 20 μm ($1/e^2$ diameter) circular spot at the center of a 250 × 250 μm^2 square bore sheath flow cuvette (NSG Precision Cell, Inc.). Fluorescence was collected at 90° with a 40X, 0.85 NA microscope objective (Nikon Fluor) and spatially filtered with a 400 μm slit located at the image plane of the microscope objective. The probe volume thus defined was approximately 3 pl. Light passing through the slit was spectrally filtered with a 30 nm band-pass filter centered at 575 nm (575DF30 Omega Optical). The filtered light was focused on the 200 × 200 μm^2 active area of a photon counting avalanche photodiode (SPCM-200-PQ C.D. 2027, EG&G Optoelectronics Canada) using a 10X microscope objective. The sheath fluid was ultrapure water and was delivered to the flow cell using gravity feed. The sheath volumetric flow rate was adjusted to give transit times through the probe volume in the range of 0.4−6 ms. B-Phycoerythrin dissolved in 1X Dulbecco's phosphate buffered saline at a concentration of 1.4×10^{-10} M was forced through a capillary (o.d. 90 μm, i.d. 20 μm, Polymicro Technologies) via a pressure differential. B-Phycoerythrin eluted from the capillary tip, which was positioned inside the square bore cuvette approximately 100 μm upstream of the probe volume. Sample delivery rates were adjusted such that, on average, less than one B-phycoerythrin molecule occupied the probe volume.

Photoelectron pulses from the photodiode were amplified, conditioned with a constant fraction discriminator, and counted with a multichannel scalar (MCS) PC card (Oxford Instruments). An IDL (Interactive Data Language, Research Systems, Inc.) program was used to analyze the data. Photon bursts from B-phycoerythrin were sifted from the data, and the burst areas were integrated and histogrammed.

3.2. Model Calculations. In our calculations of the BSD, we assumed a Gaussian profile of the laser beam

$$I(x,z) = \frac{2P_0}{\pi w^2} \exp\left[-2\frac{x^2+z^2}{w^2}\right] \quad (18)$$

where w denotes the beam waist radius, and P_0 is the total laser power (in photons per time unit). We neglect the divergence of the beam, assuming that the extent of the detection region along the laser beam is sufficiently small. The collection efficiency function $\eta(\vec{r})$ was calculated as described in ref 42. Furthermore, the following parameter values were used throughout all calculations: diffusion constant $D = 43\ \mu m^2/s$, absorption cross section $\sigma = 5.4 \times 10^{-7}\ \mu m^2$, fluorescence quantum yield $\Phi_f = 0.98$, total laser power $P_0 = 1.29 \times 10^{15}\ s^{-1}$, laser beam waist radius $w = 10\ \mu$m, width of the slit image in the object space $d = 10\ \mu$m, maximum collection efficiency $\eta_0 = 1.3 \times 10^{-2}$, and distance of the injection point from the laser beam $z_{inj} = -100\ \mu$m. The sample stream diameter at the detection volume was mainly determined by the lateral diffusion of the molecules during the time between injection and detection. The time step used in the Monte Carlo simulation was $\Delta t = 50\ \mu$s. The integration time T was always chosen large enough so that the molecules always passed through the detection region. For every simulation, 10^4 paths were sampled as described in section 2.2. One simulation took around 15 min on a Pentium processor 133-MHz PC. The simulation program was written with *Matlab* and can be requested from the authors.[43]

There remained six unknown parameters which had to be adjusted: the photobleaching quantum yield Φ_{bl}, the flow velocity v (assumed to be uniform over the detection region), the position of the sample stream axis x_0 and y_0, the hydrodynamically focused sample stream radius in the absence of diffusion R, and the acceleration parameter κ (see Appendix). Fortunately, the different parameters influence different features of the BSDs, facilitating their adjustment to the complete set of the measured distributions. We analyzed five different BSDs measured at different (but unknown) flow velocities. For all five measurements, the parameters were the same except the flow velocity and the transverse position of the injection capillary. Before every measurement, the transverse position of the injection capillary was readjusted to achieve highest count rates in the avalanche photodiode.

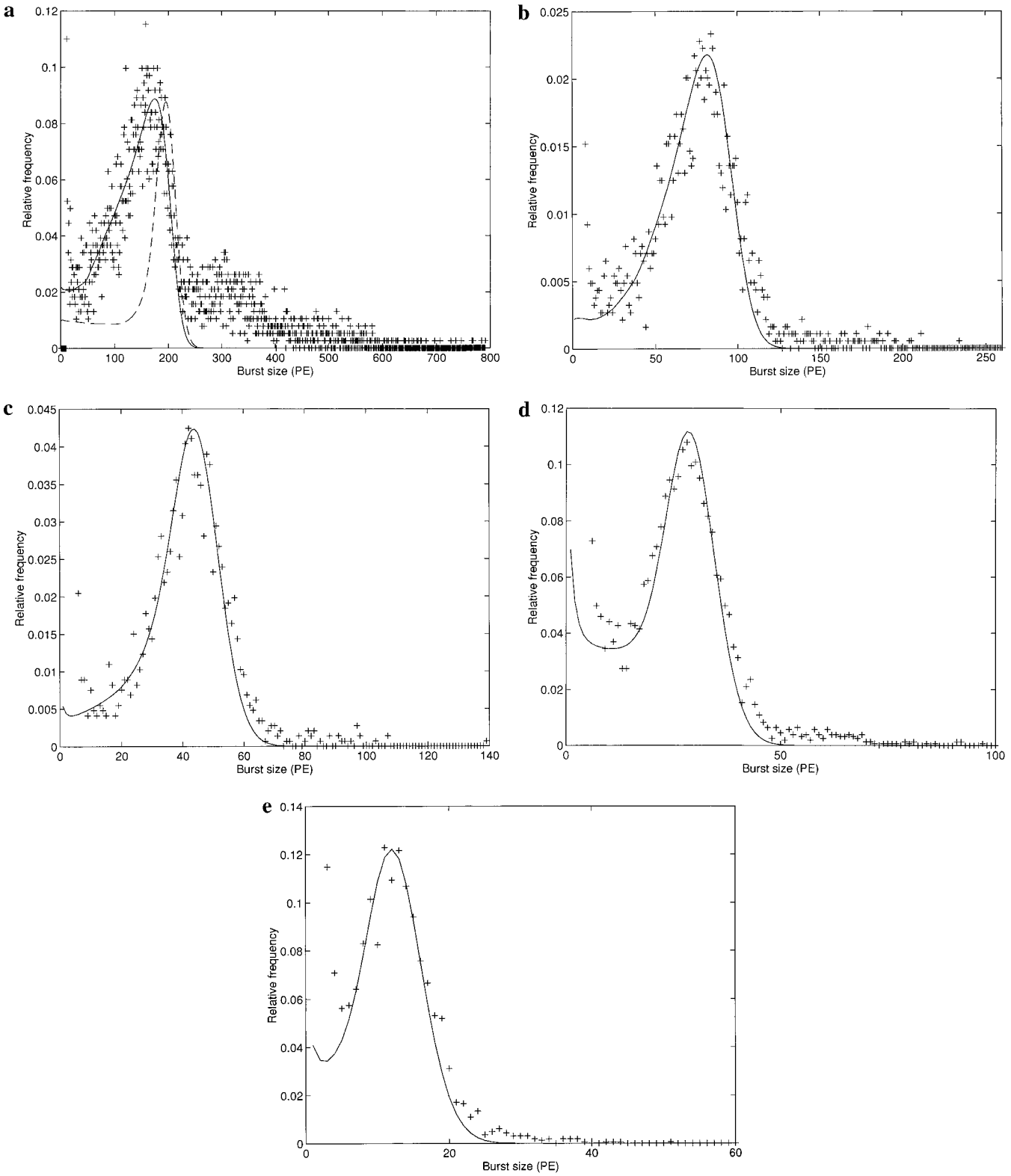

Figure 3. (a) Measured BSD (crosses) and simulation result (solid line) for $v = 3.5$ mm/s. For comparison, a simulation result with $x_0 = 0$ and $y_0 = 0$ is shown (dashed line). (b) Measured BSD (crosses) and simulation result (solid line) for $v = 7.4$ mm/s. (c) Measured BSD (crosses) and simulation result (solid line) for $v = 15.0$ mm/s. (d) Measured BSD (crosses) and simulation result (solid line) for $v = 23.5$ mm/s. (e) Measured BSD (crosses) and simulation result (solid line) for $v = 53.0$ mm/s.

The measured BSDs together with the simulation results are shown in Figures 3a−e. The values used for the photobleaching quantum yield, focused sample stream radius and acceleration parameter were $\Phi_{bl} = 1.4 \times 10^{-5}$, $R = 0.1$ μm, $\kappa = 175$ s^{-1}, and were kept unchanged for all five curves. The photobleaching quantum yield is in good agreement with literature values.[44,45] The best values found for the flow velocity v and sample stream axis shift $\{x_0,y_0\}$ for each curve are given in Table 1. The deviations between measurement and simulation in the high burst size region are due to the simultaneous crossings of two or more molecules through the detection region, which our present model does not account for. The deviations in the low burst size range are due to background noise in the measured data. In the calculations, background was neglected ($V_{bg} = 0$), since it effects the BSD only in the limit of small bursts.

TABLE 1: Flow Velocity Values and Sample Stream Axis Shifts Used for Modeling the Measured BSDs

curve no.	flow velocity v (mm/s)	shift x_0 (μm)	shift y_0 (μm)
1	3.5	3.2	1.5
2	7.4	2.8	2.5
3	15.0	1.4	3.6
4	23.5	0.0	4.6
5	53.0	0.0	4.5

As noted above, the different parameters effect the shape of the BSD in different ways: the flow velocity v determines mainly the position of the high burst size edge of the BSD and the value of the photobleaching quantum yield Φ_{bl} has its biggest influence in the low burst size region. The values of the sample stream diameter R and the acceleration parameter κ influence the shape of the BSD at its maximum. The most unexpected result of the model calculation was the fact that the overall width of the BSD is mostly determined by the lateral shift $\{x_0,y_0\}$ of the sample stream. A dramatic change of the BSDs shape can be seen in Figure 3a, where the result of a calculation with $x_0 = 0$ and $y_0 = 0$ is included. It should be noted that jitter of the sample stream position can have a similar effect on the BSD. It is also possible that aberrations in the collection optics broaden the BSD.

As can be seen from Figures 3a−e, the model calculations are in good agreement with the measured BSDs. Currently, the weakest point of the BSD modeling is the rather simple assumptions concerning the hydrodynamics of the fluid flow (Appendix). Future work is required to improve our understanding of this point, which is essential for calculating the initial distribution $p_0(\vec{r})$.

4. Conclusions

Some remarks need to be added with concern to the assumptions made about the photophysics of the fluorescent molecules. In the present paper, no attention was paid to effects such as saturation of the excited state, triplet-state shelving, or more complicated nonlinear photobleaching mechanisms. However, the general path integral approach to the calculation of photon statistics is not affected by the concrete assumptions about the single molecule photophysics, and it is straightforward to extend the presented approach to include such photophysical effects.

The path integral approach works well not only for modeling SMD experiments in flow, but also for obtaining autocorrelation functions in fluorescence correlation spectroscopy, as was shown in refs 46 and 47. It can be applied also to the calculation of SMD statistics when more than just a single molecule can be present at the same time within the detection volume. This will be presented elsewhere.

Acknowledgment. We thank Lawrence R. Pratt (LANL, Los Alamos National Laboratory) for valuable discussions. J.E. greatly acknowledges the support of the German Academic Exchange Service, granting him his stay at LANL. J.E. also thanks Lloyd M. Davis (University of Tennessee Space Institute) for first attracting his attention to the problem of photon detection statistics in flow experiments and also thanks Edgar Klose (GMBU e.V. Berlin) for the great support of his work.

Appendix

Initial Molecule Distribution $p_0(\vec{r})$. The molecules are eluted from the injection capillary at position $z = z_{inj}$. The laser beam is located at $z = 0$. In our model calculations, we calculate molecule trajectories starting at an intermediate plane $z = z_0 < -3w$. At this distance from the laser beam, the light intensity is negligibly small. Here we estimate the molecule distribution in this plane. First of all, we assume that the molecules are always sufficiently near the axis of the outer capillary so that we can neglect any effects of the nonconstant velocity profile across the outer capillary cross section. The correct description of the hydrodynamics of the sample stream is rather complicated.[48] Thus, we use a simplified schematic model: at the injection point, the molecules are assumed to be uniformly distributed on a circle with radius R, and the sample stream at that point is assumed to have negligible velocity. The acceleration of the sample stream in the sheath flow is described by the following simple equation:

$$\frac{dv}{dt} = \kappa(v_0 - v) \tag{B1}$$

where v_0 is the final velocity corresponding to the maximum velocity of the Poiseuille flow in the capillary, and κ is a constant acceleration parameter. Thus, we are assuming that the sample stream is accelerated proportional to the difference between its actual velocity and its final velocity and that diffusion along the capillary axis can be neglected when compared with the sample stream flow. Thus, we have

$$v(t) = v_0[1 - \exp(-\kappa t)] \tag{B2}$$

and for the traveled distance $z - z_{inj}$:

$$\frac{z - z_{inj}}{v_0} = t - \frac{1}{\kappa}[1 - \exp(-\kappa t)] \tag{B3}$$

Inserting into the last equation for z the value z_0 defines the traveling time t_0 of the molecule between the injection point and the plane $z = z_0$.

Assuming that the initial distribution of molecules at the injection point $z = z_{inj}$ is uniform over a circle with radius R, and that the molecules diffuse laterally during their travel from this point to the plane $z = z_0$, the initial distribution p_0 is finally given by

$$p_0(x,y,z) = \frac{\delta(z - z_0)}{4\pi^2 R^2 D t_0}\int_0^R dr\, r \int_0^{2\pi} d\phi \times \exp\left[-\frac{(x - x_0 - r\cos\phi)^2 + (y - y_0 - r\sin\phi)^2}{4Dt_0}\right] \tag{B4}$$

Here, we took into account also that the sample stream axis is shifted by the vector $\{x_0,y_0\}$ against the center axis of the flow system, which was a significant parameter for a correct description of the measured data.

The proposed model is certainly oversimplified, since it does not take into account the real hydrodynamics of the sample stream nor such effects as variation of the sample stream diameter during the acceleration process. Thus, the involved parameters R and κ have a rather empirical character.

References and Notes

(1) Hirschfeld, T. *Appl. Opt.* **1976**, *15*, 2965.
(2) Shera, E. B.; Seitzinger, N. K.; Davis, L. M.; Keller, R. A.; Soper, S. A. *Chem. Phys. Lett.* **1990**, *174*, 553.
(3) Soper, S. A.; Shera, E. B.; Martin, J. C.; Jett, J. H.; Hahn, J. H.; Nutter, H. L.; Keller, R. A. *Anal. Chem.* **1991**, *63*, 432.
(4) Soper, S. A.; Davis, L. M.; Shera, E. B. *J. Opt. Soc. Am. B* **1992**, *9*, 1761.
(5) Whitten, W. B.; Ramsey, J. M.; Arnold, J. M.; Bronk, B. V. *Anal. Chem.* **1991**, *63*, 1027.

(6) Soper, S. A.; Mattingly, Q. L.; Vegunta, P. *Anal. Chem.* **1993**, *65*, 740.
(7) Wilkerson, C. W.; Goodwin, P. M.; Ambrose, W. P.; Martin, J. C.; Keller, R. A. *Appl. Phys. Lett.* **1993**, *62*, 2030.
(8) Lee, Y. H.; Maus, R. G.; Smith, B. W.; Winefordner, J. D. *Anal. Chem.* **1994**, *66*, 4142.
(9) Rigler, R.; Widengren, J.; Mets, Ü. In *Fluorescence Spectroscopy*; Wolfbeis, O. S., Ed.; Springer: Berlin, 1993; pp 13−24.
(10) Mets, Ü.; Rigler, R. *J. Fluorescence* **1994**, *4*, 259.
(11) Nie, S.; Chiu, D. T.; Zare, R. N. *Science* **1994**, *266*, 1018.
(12) Nie, S.; Chiu, D. T.; Zare, R. N. *Anal. Chem.* **1995**, *67*, 2849.
(13) Li, L.-Q.; Davis, L. M. *Appl. Opt.* **1995**, *34*, 3208.
(14) Goodwin, P. M.; Affleck, R. L.; Ambrose, W. P.; Jett, J. H.; Johnson, M. E.; Martin, J. C.; Petty, J. T.; Schecker, J. A.; Wu, M.; Keller, R. A. In *Computer Assisted Analytical Spectroscopy*; Brown, S. D., Ed.; John Wiley and Sons: New York, 1996; p 61.
(15) Keller, R. A.; Ambrose, W. P.; Goodwin, P. M.; Jett, J. H.; Martin, J. C.; Wu, M. *Appl. Spectrosc.* **1996**, *50*, 12A.
(16) Jett, J. H.; Keller, R. A.; Martin, J. C.; Marrone, B. L.; Moyzis, R. K.; Ratliff, R. L.; Seitzinger, N. K.; Shera, E. B.; Stewart, C. C. *J. Biomol. Struct. Dyn.* **1991**, *7*, 301.
(17) Harding, J. D.; Keller, R. A. *Trends Biotechnol.* **1992**, *10*, 55.
(18) Ambrose, W. P.; Goodwin, P. M.; Jett, J. H.; Johnson, M. E.; Martin, J. C.; Marrone, B. L.; Schecker, J. A.; Wilkerson, C. W.; Keller, R. A.; Haces, A.; Shih, P.-J.; Harding, J. D. *Ber. Bunsen-Ges. Phys. Chem.* **1993**, *97*, 1535.
(19) Goodwin, P. M.; Schecker, J. A.; Wilkerson, C. W.; Hammond, M. L.; Ambrose, W. P.; Jett, J. H.; Martin, J. C.; Marrone, B. L.; Keller, R. A.; Haces, A.; Shih, P.-J.; Harding, J. D. *Proc. SPIE−Int. Soc. Opt. Eng.* **1993**, *1891*, 127.
(20) Goodwin, P. M.; Affleck, R. L.; Ambrose, W. P.; Demas, J. N.; Jett, J. H.; Martin, J. C.; Reha-Krantz, L. J.; Semin, D. J.; Schecker, J. A.; Wu, M.; Keller, R. A. *Exp. Tech. Phys.* **1995**, *41*, 279.
(21) Goodwin, P. M.; Johnson, M. E.; Martin, J. C.; Ambrose, W. P.; Marrone, B. L.; Jett, J. H.; Keller, R. A. *Nucleic Acids Res.* **1993**, *21*, 803.
(22) Petty, J. T.; Johnson, M. E.; Goodwin, P. M.; Martin, J. C.; Jett, J. H.; Keller, R. A. *Anal. Chem.* **1995**, *67*, 1755.
(23) Huang, Z.; Petty, J. T.; O'Quinn, B.; Longmire, J. L.; Brown, N. C.; Jett, J. H.; Keller, R. A. *Nucleic Acids Res.* **1996**, *24*, 4202.
(24) Castro, A.; Shera, E. B. *Anal. Chem.* **1995**, *67*, 3181.
(25) Castro, A.; Shera, E. B. *Appl. Opt.* **1995**, *34*, 3218.
(26) Eigen, M.; Rigler, R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5740.
(27) Rigler, R. *J. Biotechnology* **1995**, *41*, 177.
(28) Xue, Q.; Yeung, E. S. *Nature* **1995**, *373*, 681.
(29) Wang, J.; Wolynes, P. *Phys. Rev. Lett.* **1995**, *74*, 4317.
(30) Chen, D. Y.; Adelhelm, K.; Cheng, X. L.; Dovichi, N. J. *Analyst* **1994**, *119*, 349.
(31) Chen, D. Y.; Dovichi, N. J. *Anal. Chem.* **1996**, *68*, 690.
(32) Edman, L.; Mets, Ü.; Rigler, R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6710.
(33) Wu, M.; Goodwin, P. M.; Ambrose, W. P.; Keller, R. A. *J. Phys. Chem.* **1996**, *100*, 17406.
(34) Köllner, M. *Appl. Opt.* **1993**, *32*, 806.
(35) Enderlein, J. *Appl. Opt.* **1995**, *34*, 514.
(36) Mathies, R. A.; Peck, K.; Stryer, L. *Anal. Chem.* **1990**, *62*, 1786.
(37) Soper, S. A.; Nutter, H. L.; Keller, R. A.; Davis, L. M.; Shera, E. B. *Photochem. Photobiol.* **1993**, *57*, 972.
(38) Enderlein, J. *Proc. SPIE−Int. Sr. Opt. Engl.* **1994**, *Vol. 2136*, 226.
(39) Gardiner, C. W. *Handbook of Stochastic Methods*; Springer: Berlin, 1985.
(40) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill: New York, 1965.
(41) Kleinert, H. *Path Integral in Quantum Mechanics, Statistics and Polymer Physics*, 2nd ed.; World Scientific: Singapore, 1994.
(42) Enderlein, J.; Ambrose, W. P. The optical collection efficiency function in single molecule detection experiments. *Appl. Opt.*, in press.
(43) A *Matlab* program code of the Monte Carlo sampling can be obtained from the authors by sending an E-mail to j.enderlein@tbx.berlinet.de. *Matlab* is a registered trademark by MathWorks Inc.
(44) Mathies, R. A.; Stryer, L. *Applications of Fluorescence in Biomedical Sciences*; Taylor, D. L., Waggoner, A. S., Murphy, R. F., Lanni, F., Birge, R. R., Eds.; Alan R. Liss, Inc.: New York, 1986; pp 129−140.
(45) White, J. C.; Stryer, L. *Anal. Biochem.* **1987**, *161*, 442.
(46) Enderlein, J. *Exp. Tech. Phys.* **1995**, *41*, 211.
(47) Enderlein, J. *Phys. Lett. A* **1996**, *221*, 427.
(48) Kachel, V.; Fellner-Feldegg, H.; Menke, E. *Flow Cytometry and Sorting*; Melamed, M. R., Lindmo, T., Mendelsohn, M. L., Eds.; Wiley-Liss: New York, 1990; pp 27−44.