Published in final edited form as:

Anal Biochem. 1991 June ; 195(2): 243–254. doi:10.1016/0003-2697(91)90324-m.

Correction for Incomplete Labeling in the Measurement of Distance Distributions by Frequency-Domain Fluorometry

Joseph R. Lakowicz^{*,2}, Ignacy Gryczyński^{*}, Wiesław Wiczkowski^{*,3}, Józef Kuśba^{*,4}, Michael L. Johnson[†]

^{*}University of Maryland School of Medicine, Center for Fluorescence Spectroscopy, Department of Biological Chemistry, 660 West Redwood Street, Baltimore, Maryland 21201

[†]Department of Pharmacology, University of Virginia, Charlottesville, Virginia 22908

Abstract

Measurements of time-resolved fluorescence are now being used to recover conformational distributions of biological macromolecules. The fluorescence data of the donor are easily corrupted by incomplete labeling of the macromolecules by the acceptor. In the present paper we describe a general procedure to correct for incomplete acceptor labeling in the determination of distance distributions from frequency-domain measurements of the donor fluorescence decay kinetics. The method can also be used to determine the extent of acceptor labeling. Simulated data were used to determine the effect of incomplete labeling on resolution of the distance distribution and the effect on the recovered distributions if one fails to account for incomplete labeling by the acceptor. The expressions and implemented algorithm were verified using known mixtures of donor–control and donor–acceptor pair molecules, which simulated the presence of a donor population lacking the acceptor. Finally, we present data on the distance distributions between two labeled sites in myosin S1 (Cys-697 to Cys-707) where it was not possible to obtain complete labeling of the acceptor site.

Fluorescence resonance energy transfer (FRET)⁵ is widely used to measure the distances between sites on biological macromolecules (1–4). These measurements are possible because FRET is a through-space interaction which occurs over distances characteristic of the dimension of these molecules, 20–60 Å (5). More recently, the FRET measurements have been extended to the resolution of distance or conformational distributions of flexible molecules (6–8). These measurements are a continuation of the pioneering studies of Haas, Steinberg, and co-workers (9–12). Determination of distance distributions is possible from the time-resolved measurements because the distribution of distances results in a range of energy transfer rates, which in turn result in complex decay kinetics for the energy transfer donor. Consequently, the determination of distance distributions relies heavily on the measurement of complex decay kinetics on the nano- and picosecond timescales.

²To whom correspondence should be addressed.

³Permanent address: University of Gdańsk, Institute of Chemistry, Gdańsk, Sobieskiego 18, Poland.

⁴Permanent address: Technical University of Gdańsk, Department of Technical Physics and Applied Mathematics, 80-952 Gdańsk, Poland.

⁵Abbreviations used: FRET, fluorescence resonance energy transfer; FD, frequency domain; D, donor; A, acceptor; hw, half-width.

During the past decade there has been remarkable progress in the instrumentation used to measure time-resolved fluorescence, and these instruments are increasingly available to researchers in biochemistry and biophysics. In particular, the frequency-domain (FD) measurements (13–15) are gaining popularity because of their simplicity, high resolution of complex decay kinetics (16–19), and ability to measure picosecond decay times at gigahertz frequencies (20). In fact, it has recently been demonstrated that the frequency-domain FRET measurements allow resolution of both the distance distribution and the end-to-end diffusion coefficients of flexible molecules (21, 22). Such resolution is possible using just the donor decay kinetics, in spite of reports which suggest the contrary (23). This ability to provide experimental data about the conformational heterogeneity and dynamics of flexible molecules is likely to be of widespread usefulness in biophysical studies of protein dynamics, protein folding and assembly, protein-membrane interactions, and other functions of biopolymers which rely upon their segmented mobility.

In the present paper we address a problem encountered in many FRET measurements. In these measurements the desired information is contained in the donor decay kinetics and the effect of the acceptor on these kinetics. Frequently, it is not possible to obtain complete labeling with either donor (D) or acceptor (A). Incomplete labeling by the donors is not of concern because the unlabeled fraction is not seen in the donor measurements. However, donor-labeled molecules which lack an acceptor cause a serious distortion of the data. This is because the donor-alone decay kinetics are distinct from those of the D-A pairs and because the donor emission from these donor-alone molecules is more intense than those of the D-A pairs due to the absence of energy transfer quenching. Just a few percent of donor-alone molecules can result in an inability to fit the data to any rational distance distribution or in a distortion of the recovered distance distribution. Consequently, we developed a method to correct for the presence of a donor-alone component in the frequency-domain data. This method allows the extent of acceptor labeling to be determined from the data. Alternatively, increased resolution of the distance distribution is obtained if the extent of acceptor labeling is known from other data and this value held constant during the analysis.

THEORY

Distance distributions between fluorescent donors and acceptors can be recovered from the intensity decay kinetics of the donors. Energy transfer to nearby acceptors results in an altered donor decay due to the competing rate processes of emission and energy transfer. A range of donor-to-acceptor distances results in a range of energy transfer rates, which in turn results in a more complex decay for the donor. For most macromolecules the intensity decays of the donor, even in the absence of an energy transfer acceptor, are multiexponential:

$$I_D(t) = I_{0D} \sum_i \alpha_{Di} \exp(-t/\tau_{Di}). \quad [1]$$

In this expression α_{Di} are the relative initial intensities (at $t = 0$) of the particular exponential components which are characterized by the decay times τ_{Di} . The factors α_{Di} are usually

normalized so that $\sum_i \alpha_{Di} = 1$. I_{0D} is the intensity of the donor emission at time $t = 0$. The process of nonradiative energy transfer occurs in macromolecules labeled by donors and acceptors. In the most common case of dipole-dipole interaction the energy transfer rate depends on distance and is given by

$$k_{DAi}(r) = \frac{1}{\tau_{Di}} \left(\frac{R_0}{r} \right)^6, \quad [2]$$

where R_0 is the Forster radius and r is the donor-to-acceptor distance. For the case of a donor-acceptor pair this additional deactivation process requires the donor-alone decay functions $\alpha_{Di} \exp(-t/\tau_{Di})$ to be modified to the form to be modified to the form $\alpha_{Di} \exp\{-[\tau_{Di}^{-1} + k_{DAi}(r)]t\}$. We assumed here that energy transfer is the only mechanism of donor quenching, hence the α_{Di} are unchanged by the presence of acceptor. Also, we assumed that the rates of energy transfer scale according to the decay times of each component in the decay (Eq. [2]). While we do not have experimental proof for this scaling of the transfer rates, Eq. [2], it is known that for single exponential decays the transfer rates are proportional to the reciprocal decay time (5). The distance r depends on the conformation of the macromolecule. Because of the very large number of different conformations it is reasonable to introduce the probability function $P(r)$ which describes the distribution of D-A distances in the investigated system. The function $P(r)$ is defined so that $P(r)dr$ is the probability that a given macromolecule has the donor-to-acceptor distance r belonging to the interval $(r, r + dr)$. In general the distance cannot be less than a certain minimal value $r_{\min}$ and greater than a certain maximal value $r_{\max}$. We parameterize the function $P(r)$ assuming a Gaussian model

$$P(r) = \begin{cases} \frac{1}{Z} \exp\left[-\frac{1}{2} \left(\frac{r - R_{av}}{\sigma}\right)^2\right] & \text{for } r_{\min} \leq r \leq r_{\max}, \\ 0 & \text{elsewhere,} \end{cases} \quad [3]$$

where Z is the normalization factor

$$Z = \int_{r_{\min}}^{r_{\max}} \exp\left[-\frac{1}{2} \left(\frac{r - R_{av}}{\sigma}\right)^2\right] dr. \quad [4]$$

The average distance and standard deviation of the untruncated Gaussian function are R_{av} and σ , respectively. The standard deviation is related to the half-width of the distribution (hw, full width at half-maximum height) by $hw = 2.35\sigma$. The intensity decay of an ensemble of donor-labeled macromolecules which are also labeled by acceptor is therefore given by

$$I_{\text{DA}}(t) = I_{0\text{DA}} \int_{r_{\min}}^{r_{\max}} P(r) \sum_i \alpha_{\text{Di}} \exp \left[-\frac{t}{\tau_{\text{Di}}} - \frac{t}{\tau_{\text{Di}}} \left(\frac{R_0}{r} \right)^6 \right] dr, \quad [5]$$

where $I_{0\text{DA}}$ denotes the emission intensity of the donors at $t = 0$. If the only mechanism of donor quenching is energy transfer, and if inner filter effects are negligible, then one expects $I_{0\text{DA}} = I_{0\text{D}}$. We note that in lifetime measurements the total intensity often is not measured. Hence, the values of $\sum \alpha_{\text{Di}}$ are normalized to unity, and the factor $I_{0\text{D}}$ and $I_{0\text{DA}}$ are ignored.

In frequency-domain fluorometry, the measured quantities are the frequency-dependent phase shift (ϕ_ω) and modulation (m_ω). For any given intensity decay and distance distribution these values can be calculated (c) from the sine (N_ω) and cosine (D_ω) transforms of the intensity decay function

$$\phi_{c\omega} = \arctan(N_\omega/D_\omega), \quad [6]$$

$$m_{c\omega} = (N_\omega^2 + D_\omega^2)^{1/2}. \quad [7]$$

The sine and cosine transforms are given by

$$N_{\omega\text{DA}} = \frac{1}{J_{\text{DA}}} \int_{r_{\min}}^{r_{\max}} P(r) \sum_i \frac{\alpha_{\text{Di}} \omega \tau_{\text{DAi}}^2}{1 + \omega^2 \tau_{\text{DAi}}^2} dr, \quad [8]$$

$$D_{\omega\text{DA}} = \frac{1}{J_{\text{DA}}} \int_{r_{\min}}^{r_{\max}} P(r) \sum_i \frac{\alpha_{\text{Di}} \tau_{\text{DAi}}}{1 + \omega^2 \tau_{\text{DAi}}^2} dr, \quad [9]$$

$$J_{\text{DA}} = \int_{r_{\min}}^{r_{\max}} P(r) \sum_i \alpha_{\text{Di}} \tau_{\text{DAi}} dr. \quad [10]$$

For clarity we note that the decay times in Eqs. [8]–[10] are dependent upon the D-to-A distance and are related to the donor-alone decay times by

$$\frac{1}{\tau_{\text{DAi}}} = \frac{1}{\tau_{\text{Di}}} + \frac{1}{\tau_{\text{Di}}} \left(\frac{R_0}{r} \right)^6. \quad [11]$$

The parameters describing the distance distribution are estimated from the experimental (ϕ_ω and m_ω) values of the phase shift and the modulation by using a nonlinear least-squares fitting procedure (24) in which the reduced χ^2 ratio is minimized

$$\chi_R^2 = \frac{1}{v} \sum_{\omega} \left(\frac{\phi_{\omega} - \phi_{c\omega}}{\delta\phi} \right)^2 + \frac{1}{v} \sum_{\omega} \left(\frac{m_{\omega} - m_{c\omega}}{\delta m} \right)^2, \quad [12]$$

where v is the number of degrees of freedom and $\delta\phi$ and δm are the experimental uncertainties in the measured phase and modulation values. For the measurement of distance distributions the parameters varied to minimize χ_R^2 are R_{av} and hw . Occasionally, the extent of labeling by acceptor (L , see below) is also a fitted parameter. The parameters which describe the donor-alone decay (α_{Di} and τ_{Di}) are determined by an independent measurement and analysis of the donor-alone frequency response. These α_{Di} and τ_{Di} values are held fixed during the distance distribution analysis.

Incomplete Labeling by Acceptor

Suppose now that the acceptor is present at a fractional labeling L of less than unity. The acceptor-free donor molecules are unquenched by energy transfer; hence, the observed donor emission contains a component which is due to the donor-alone molecules. Additionally, these acceptor-free donors contribute to the observed fluorescence in excess of their molar proportion because of their higher relative quantum yield. In order to calculate the distance distribution one must take into consideration this nonstoichiometric labeling (under labeling) of the molecule by acceptor.

Since the same donor molecules are emitting in the presence and absence of acceptor, the preexponential factors α_{Di} , or equivalently the values of I_{0D} and I_{0DA} , are the same for both species. The observed intensity $I(t)$ of the mixture of the nonlabeled and acceptor-labeled macromolecules is given by

$$I(t) = I_D^m(t) + I_{DA}^m(t), \quad [13]$$

where the m indicates the decays which are observed from the mixture of donor-alone and donor-acceptor molecules. Recalling that the intensities at $t = 0$ (I_{0D} and I_{0DA}) are assumed to be unaffected by energy transfer one obtains

$$I_{0D}^m = (1 - L)I_0, \quad [14]$$

$$I_{0DA}^m = LI_0, \quad [15]$$

where $I_0 = I(t = 0)$ irrespective of the presence or absence of acceptor. Introducing Eqs. [1] and [5] into Eq. [13] and making use of Eqs. [14] and [15] we obtain

$$I(t) = (1 - L)I_0 \sum_i \alpha_{Di} \exp(-t/\tau_{Di}) + LI_0 \int_{r_{\min}}^{r_{\max}} P(r) \sum_i \alpha_{Di} \exp\left[-\frac{t}{\tau_{Di}} - \frac{t}{\tau_{Di}} \left(\frac{R_0}{r}\right)^6\right] dr. \quad [16]$$

The frequency-domain expressions for the observed emission are

$$N_{\omega} = \frac{1}{J} [(1 - L)J_D N_{\omega D} + LJ_{DA} N_{\omega DA}], \quad [17]$$

$$D_{\omega} = \frac{1}{J} [(1 - L)J_D D_{\omega D} + LJ_{DA} D_{\omega DA}], \quad [18]$$

$$J = (1 - L)J_D + LJ_{DA}, \quad [19]$$

where $N_{\omega DA}$, $D_{\omega DA}$, and J_{DA} are given by Eqs. [8]–[10], and

$$N_{\omega D} = \frac{1}{J_D} \sum_i \frac{\omega \alpha_{Di} \tau_{Di}^2}{1 + \omega^2 \tau_{Di}^2}, \quad [20]$$

$$D_{\omega D} = \frac{1}{J_D} \sum_i \frac{\alpha_{Di} \tau_{Di}}{1 + \omega^2 \tau_{Di}^2}, \quad [21]$$

$$J_D = \sum_i \alpha_{Di} \tau_{Di}. \quad [22]$$

Equations [20]–[22] are also the usual expressions used to analyze the frequency-domain data in terms of a multiexponential intensity decay (24, 25). During analysis of the frequency-domain data for incompletely-labeled samples the variable parameters are R_{av} , hw , and the fractional labeling (L). If known from other measurements, L can be held constant during the analysis.

It is important to relate the fractional labeling (L) to the fractional intensities of the donor-alone (f_D) and donor-acceptor (f_{DA}) populations. Using our former notation, the quantities f_D and f_{DA} can be defined as

$$f_D = \frac{\int_0^{\infty} I_D^m(t) dt}{\int_0^{\infty} I(t) dt}, \quad [23]$$

and

$$f_{\text{DA}} = \frac{\int_0^\infty I_{\text{DA}}^m(t) dt}{\int_0^\infty I(t) dt}. \quad [24]$$

Equations [23] and [24] together with Eqs. [1], [5], and [13]–[16] yield

$$f_{\text{D}} = \frac{1 - L}{1 - L + Lq_{\text{DA}}} \quad [25]$$

$$f_{\text{DA}} = \frac{Lq_{\text{DA}}}{1 - L + Lq_{\text{DA}}},$$

where

$$q_{\text{DA}} = \int_{r_{\text{min}}}^{r_{\text{max}}} P(r) \frac{1}{1 + (R_0/r)^6} dr. \quad [26]$$

The quantity q_{DA} has a meaning of quantum yield of the donor-acceptor population relative to the quantum yield of the donor-alone population. The numerical value of q_{DA} can be obtained as an independent experimental observable, or it can be calculated from the recovered distance distributions.

EXPERIMENTAL PROCEDURES

Fluorescence lifetimes were determined from the frequency response of the donor emission using the instrument described previously (15). The cavity-dumped output of a synchronously pumped dye laser was used to generate a laser pulse train with a repetition rate of 3.79 MHz and a pulse width of about 5 ps, which was then frequency-doubled. This source is intrinsically modulated to many gigahertz and used to directly excite the samples. The intensity decays were measured using rotation-free magic angle polarized conditions. For measurements of TU2D the laser dye was rhodamine 6G, frequency-doubled to 290 nm, with the donor emission isolated using a 340-nm interference filter. For measurements of myosin S1 the laser dye was pyridine 2, frequency-doubled to 355 nm, with a 480 nm emission filter.

For the least-squares analysis, the random errors in the phase angle ($\delta\phi$) and the modulation (δm) were assumed to be 0.2° and 0.005, respectively. These same amounts of random error were added to the simulated phase and modulation values, as described previously for simulated multiexponential decays (25). The data analysis programs are written in Fortran 77 and run on Silicon Graphics computers (4D/70 or GTX120). These programs are available at the Center for Fluorescence Spectroscopy library and can be accessed via modem or ethernet.

Synthesis of TU2D is described elsewhere (27). Myosin S1 was labeled on Cys-697 (SH₁) by IAEDANS and on Cys-707 (SH₂) by DDPM, as described previously (28). The degree of SH₁ and SH₂ labeling was determined by absorbance measurements (29).

RESULTS

Simulations of the Extent of Labeling on the Resolution of Distance Distributions

We used simulated frequency-domain data to evaluate the effects of incomplete acceptor labeling on the form of the data and the recovered distance distributions. For these simulations we used assumed parameters which are typical of our experimental systems: an unquenched donor decay time of 7 ns (single exponential in this case), a Forster distance of 25 Å, a mean distance of 20 Å, and a half-width of 15 Å. Of course, it is possible to perform simulations for a range of distance distributions and decay times. However, the objective of the present simulations is to illustrate how incomplete labeling by acceptor affects the frequency-domain data for typical distributions and decay times, to illustrate the effects on the distance distribution if donor-alone molecules are present but not accounted for in the analysis, and to demonstrate that it is possible and practical to correct for incomplete labeling by acceptor.

Simulated frequency responses for the assumed life-time and distribution are shown in Fig. 1. In these simulations we varied the extent of labeling from 0 to 100%. The presence of an energy transfer acceptor results in a shift of the response toward higher frequencies, which is of course due to the donor decay being shortened by energy transfer. It is evident from the frequency response for no acceptor ($L = 0$) and for complete acceptor labeling ($L = 1$) that there is a modest change in shape of the curve due to the presence of acceptor (Fig. 1). This increased dispersion in the frequency response is due to the presence of acceptor over a range of distances. If the acceptors were present only at a single unique distance, then the frequency response would be shifted on the frequency axis but there would be no change in shape. However, it is important to note that for any given experiment one does not know whether there is a distribution of distances, or if there is complete labeling by acceptor. Additionally, the presence of 10% of the donor population without acceptor ($L = 0.9$) results in a shifted frequency response, but without a dramatic change in shape. Hence, it is easy to imagine that incomplete labeling will not necessarily be obvious from the experimental data and that the unrecognized presence of donors without acceptor could distort the recovered distance distributions.

It is important to consider the effect of a donor-alone population on the apparent or recovered distance distributions. This effect was determined by analyzing these simulated data with L less than unity, with the assumption of complete labeling by acceptor ($L = 1.0$). Two examples of these apparent distributions are shown as an insert in Fig. 1, with a more detailed summary in Table 1. For our assumed parameters, an acceptor labeling of $L = 0.9$ is sufficient to result in a significant distortion of the distribution. If the acceptor labeling is decreased to $L = 0.75$ the apparent distribution bears little resemblance to the actual distribution. At first glance it is perhaps surprising that the recovered distributions show increased probability on the short distance side of the maximum, as intuition would suggest that donors without acceptors would show increased probabilities at the longer distances. This apparent contradiction is probably due to the relatively small donor intensity for these closely spaced D-A pairs and due to our use of a Gaussian distribution. The longer-lived donor-alone emission, which is present in the data, is accounted for the increased probability

for $r > R_{av}$ in the recovered distribution. This increased probability for $r > R_{av}$ is accompanied by a larger probability increase for $r < R_{av}$, but these donors are strongly quenched by energy transfer and contribute little to the frequency response. The results of the analyses with $L = 0$ illustrate that the recovered distance distributions can be seriously in error if labeling by acceptor is incomplete and if this is not accounted for in the analysis.

It is also of interest to determine whether the presence of incomplete acceptor labeling could be determined from the goodness-of-fit. For this purpose we fitted the simulated data for incomplete labeling ($L < 1.0$) under the assumption that labeling was complete ($L = 1.0$).

The results of these analyses (Table 1) show that χ_R^2 becomes unacceptably elevated (fourfold) if just 5% of the molecules lack acceptor ($L = 0.95$). For lower values of L the fits are still worse; for example, at $L = 0.5$ the χ_R^2 value is 25-fold higher. This means that L is not a highly correlated parameter with R_{av} or hw , and goodness-of-fit can not be compensated by adjusting the distance distribution. However, it is important to recognize that one can easily accept χ_R^2 values near 4 when experimenting with unknown systems.

Such values can be rationalized as being due to a variety of causes, such as a non-Gaussian distribution and effects of the acceptor on the donor besides energy transfer, to name a few. Hence, these results indicate that unlabeled fractions of 10% can go undetected, but their presence will result in a significant distortion of the recovered distance distribution. The data in Table 1 also show that the fraction of the total emission resulting from the D-A pairs decreases more rapidly than the mole fraction due to the higher donor yields in the absence of acceptor.

The nature of reporting experimental results often precludes the presentation of inadequate or inconclusive experiments. However, our initial attempts to recover distance distributions revealed high sensitivity to the extent of acceptor labeling. In a variety of experiments with labeled proteins were unable to fit our data to the Gaussian distance distribution model. The fits frequently did not converge, and when they did the distributions were similar to that shown for $L = 0.75$ (Fig. 1). These efforts consumed a considerable amount of time for repeated measurements and analyses. In retrospect we now recognize that these difficulties were probably the effect of incomplete acceptor labeling, and we now exercise greater control over sample purity. The positive side of these failed experiments is that the distance distribution analysis is highly sensitive to sample purity, and reasonable parameters do not result from the analysis of data from impure samples.

Importantly, one can recover the correct distance distribution, in the presence of a significant donor-only population, if one accounts for this population during the analysis. This is illustrated in Table 1 which shows that an essentially correct distribution was recovered down to $L = 0.5$, at which point the D-A pairs contributed only 21% to the total donor emission ($f_{DA} = 0.21$). If the extent of labeling is held fixed, which is conceivable since L may be known from other measurements, then the distributions can be recovered to $L = 0.25$ and possibly lower. As will be described below, this result is overly optimistic, and a more practical limit appears to be $L = 0.5$. Of course, the actual lower limit for L will depend upon the precise value of R_0 , R_{av} , and hw . Nonetheless, these results demonstrate that it is

possible to correct for the unlabeled fraction and to recover distances distribution for such preparations.

We were surprised by the relatively small uncertainties predicted in the distance distributions for the acceptor-poor samples (Table 1). Hence, we examined the χ_R^2 surfaces for the parameters characterizing these simulations, L , R_{av} , and hw (Fig. 2). These surfaces are constructed by holding one parameter fixed at the value indicated on the x -axis, followed by minimization of χ_R^2 by variation of the other parameters. This analysis is thought to account for all correlation between the parameters. While the uncertainties listed in Table 1 are also expected to account for correlation (30, 31), we have noticed in these studies and others that the χ_R^2 surfaces provide larger estimated uncertainties. We believe that construction of the χ_R^2 surfaces is the most rigorous method to estimate the uncertainties in the parameter values, including the effects of correlation between the parameters.

Distance distribution parameters and labeling χ_R^2 surfaces are shown in Fig. 2. These surfaces show that all parameters (R_{av} , hw , and L) are well determined from the data for degrees of labeling of 0.50 or greater. That is, the values of χ_R^2 are significantly dependent upon the values of all three parameters and there do not appear to be local minima where the least-squares analyses could become trapped. Below $L = 0.50$ the χ_R^2 surfaces become flat, indicating that the parameter values are poorly determined. In these analyses the extent of acceptor labeling L was regarded to be a variable parameter. However, this value can often be determined from other chemical or spectroscopic data. If L is known, and held fixed during the analysis, then the resolution of R_{av} and hw is considerably improved. This is evident from the χ_R^2 surfaces for R_{av} and hw found when L is held fixed (Fig. 2, dashed lines) and from the lower parameter uncertainties for fixed values of L (Table 1).

Effect of Incomplete Acceptor Labeling on the Distance Distribution of a Donor-Acceptor Pair TU2D

In order to verify our theory and test the implemented algorithm, we used known mixtures of a donor-control (TMA) and donor-acceptor (TU2D) molecules (Fig. 3). In these molecules the donor is indole, which emits near 340 nm, and the acceptor is the dansyl moiety. The effects of end-to-end diffusion on energy transfer were avoided by using the viscous solvent propylene glycol at 5°C. Variable amounts of incomplete labeling by acceptor were simulated by adding increasing amounts of the donor to the D-A pair. This results in an increase of donor absorption (Fig. 3), as well as donor emission (not shown), due to presence of free donors in the solution.

Phase and modulation data for the TU2D donor-acceptor system with adjusted incomplete labeling by acceptor (0.947 and 0.494) are shown in Fig. 4. The solid lines show the fits to the data using the known values of L . The dashed line shows the frequency response of the donor alone ($L = 0$). The dotted line shows the best fit to the data obtained for the system with labeling adjusted to 0.494 when analyzed with L fixed at 1.0. This results in systematic

deviations, which are shown () in the lower panels of Fig. 4, where χ_R^2 is elevated from 1.3 to The inability to fit the data with L fixed at 1.0 indicates that one can detect the presence of an unlabeled fraction by an inadequate fit. It should be noted that both R_{av} and hw are variable parameters in this fit with $L = <1.0>$. If R_{av} and/or hw were fixed, then χ_R^2 would be still larger and the unlabeled fraction easier to detect.

The distance distribution analyses for TU2D solutions containing the donor-alone molecules are summarized in Table 2. As expected from the simulations, we recovered acceptable values of R_{av} , hw , and L for values of $L = 0.494$ and higher. The force-fit analyses with $L = 1$ give several-fold larger values of χ_R^2 and unreasonable values of the distribution parameters. As the extent of labeling decreases, the uncertainties increase in the recovered values of R_{av} , hw , and L . The uncertainties appear to be somewhat elevated at $L = 0.494$ and substantially elevated for lower values of L , suggesting that the distance distribution parameters are poorly determined for $L < 0.5$. If the value of L is held fixed at these known value, the R_{av} and hw are determined with less uncertainty.

We examined the χ_R^2 surfaces of the fitted parameters (L , R_{av} , and hw) for the experimental systems with high ($L = 0.947$) and low ($L = 0.494$) adjusted labeling (Fig. 5). As expected, the resolution of the system with low labeling ($L = 0.494$) was much poorer than for higher degrees of labeling ($L = 0.947$). However, improved resolution of distance distribution parameters (R_{av} and hw) for the system with low labeling can be obtained if analysis is performed with L fixed at the known value (Fig. 5, bottom, dashed lines). The importance of correcting for incomplete labeling measurements is clearly seen from the recovered distance distributions in Fig. 6. For the system with labeling adjusted to 0.715, the distance distribution recovered using the correction for incomplete labeling is in good agreement with that obtained for a fully labeled system (Fig. 6, dotted line, Table 2), and the recovered value of $L = 0.709$ is close to the expected value ($L = 0.715$). In contrast, the distribution recovered for this system without the correction for incomplete labeling ($L = 1$) is grossly in error (Fig. 6, dashed line) and χ_R^2 is 10-fold higher than for the fit with the correction (Table 2).

Effect of Macromolecular Conformation on the Recovered Distance Distribution and Labeling

The conformation of proteins depend on many factors such as temperature, pH, or presence of metal ions, and the average distance between a donor and acceptor (labeled at a specific site) can be significantly different in different conformational states. It is important to consider the influence of R_{av} on the resolution of the labeling and distribution parameters.

For simulations we used $\tau_D = 20$ ns (this value is close to lifetime of dansyl or IAEDANS), $hw = 15$ Å, $R_0 = 30$ Å, and $L = 0.9$ and 1.0. These values are comparable to those found for myosin S1 labeled with IAEDANS and the acceptor DPPM (28). In these studies it was not possible to obtain complete labeling by the acceptor. Myosin S1 is thought to undergo a contraction in the presence of MgATP. Hence, three frequency responses for $L = 0.9$ and R_{av}

= 20, 30, and 40 Å are shown in Fig. 7 (solid line). The shape of the frequency response for $R_{av} = 20$ Å (higher energy transfer) shows more deviation from the single exponential decay response of the donor (dashed line) than the shape obtained for $R_{av} = 30$ Å and for $R_{av} = 40$ Å. Also, the difference between frequency responses predicted for complete labeling ($L = 1$, dotted line) and incomplete labeling ($L = 0.9$, solid line) is more visible for higher energy transfer. This is due to the higher relative intensity of the donor-alone molecules as compared to the more closely spaced donor-acceptor pairs. We expect to obtain better resolution in distance distribution parameters and labeling for shorter average distance (higher energy transfer). The distance distribution analyses for the simulated data with different values of R_{av} are presented in Table 3. The distance distribution resolution depends on the average distance and is superior for the shorter D-A distance. This is seen visually in Fig. 8 which shows the χ_R^2 surfaces for the parameters. Whereas the resolutions obtained for shorter average distances (20 and 30 Å) are satisfactory, the χ_R^2 surfaces obtained for longer average distance (40 Å, dashed and dotted line) are flat and indicate very poor resolution in distance distribution parameters and labeling.

Distance Distributions in Myosin Subfragment S1 with and without MgADP

From previous studies (28, 29) it is known that fluorescence resonance energy transfer between a donor (IAE-DANS) linked to SH₁ (Cys-697) of myosin S1, and an acceptor (DDPM) attached to SH₂ (Cys-707), is significantly enhanced upon addition of MgADP, indicating a decrease in the SH₁-SH₂ distance of about 7 Å induced by the binding of the nucleotide. We studied the distance distributions in this experimental system, where it was not possible to obtain complete labeling of the acceptor site. Phase and modulation data for the myosin S1 system, without (●) and with (○) MgADP, are presented in Fig. 9. The solid lines show the best fit to the data, and the dashed line shows the frequency response predicted for the donor alone. The distance distribution analyses are summarized in Table 4. These results are comparable to the above simulations where we varied the average D-A distance. The extent of acceptor labeling and half-width of the Gaussian distribution are almost the same with and without MgADP, while the average distance in the presence of the nucleotide is 6.2 Å smaller than without the nucleotide. We note that the extent of labeling recovered from our algorithm is in excellent agreement with that found from absorbance measurements, $L = 0.95$ (28). The resolution in all recovered parameters, illustrated by χ_R^2 surfaces (Fig. 10), is superior in the presence of MgADP, and the uncertainties are also smaller in the presence of nucleotide (Table 4). This improved resolution at higher degrees of transfer (smaller R_{av}) is consistent with the above simulations. Of course, as the average distance decreases still further, one can expect a loss of resolution due to the low donor intensity from the D-A pairs. The data could not be fitted well using $L = 1$ (Table 4 and Fig. 9, dotted line), especially in the presence of MgADP where χ_R^2 is elevated about 15-fold.

The recovered distance distributions between Cys-697 and Cys-707 of myosin S1 in the presence and absence of MgADP are presented in Fig. 11. Whereas the average distance shows a significant shift in presence of nucleotide, the half-width remains almost the same. This MgADP-induced conformational change in myosin S1 results in a substantial decrease

in distance from Cys-697 to Cys-707, but the remarkable flexibility of the polypeptide chain between these two labeled sites does not change. It should be emphasized that it was essential to perform the correction for the unlabeled fraction in order to obtain these results. Without this correlation one finds that MgADP results not only in a contraction of the protein, but also in a disordering of the protein, as is seen from the distribution recovered for an assumed value of $L = 1.0$. We note that our previously published results (29) are not in error due to underlabeling because the correction procedure was applied to the S1 data.

DISCUSSION

Time-resolved energy transfer measurements between donor and acceptor on macromolecules can be used to determine the specific distances and flexibility of the polymer chains. Frequently, it is not possible to obtain complete labeling of the acceptor site; that is, some fraction of the macromolecules are labeled only by the donor. Ignoring even a small amount of incomplete acceptor labeling in the data analysis can result in a poor goodness-of-fit and more importantly, unrealistic distance distribution parameters. To avoid misinterpretation of the data, one should use correction procedures which account for incomplete labeling by acceptor. While the resolution of the distance distribution is decreased by the presence of donor-alone molecules, the simulations and experimental studies of mixtures of the donor-control and donor-acceptor pair show that it is possible and practical to recover correct values of the labeling and distance distribution parameters if L is higher than 0.5. For lower values of L the distribution parameters can be reasonable recovered only if L is kept constant at the known value.

Acknowledgments

This work was supported by grants from the National Institutes of Health (GM 35154) and from the National Science Foundation (DMB-8804931), with additional support from the National Science Foundation for instrumentation (DMB-8511065) and the Center for Fluorescence Spectroscopy (DIR-8710401). J.R.L. and W.W. acknowledge support from the Medical Biotechnology Center at the University of Maryland. The authors thank Professor Herb Cheung for use of the myosin S1 data, which was obtained under NIH Grant AR-31239, and Ms. Peggy Eis for her careful editing of this manuscript.

REFERENCES

1. Stryer L (1978) *Annu. Rev. Biochem* 47, 819–846. [PubMed: 354506]
2. Steinberg IZ (1971) *Annu. Rev. Biochem* 40, 83–114. [PubMed: 4331120]
3. Lakowicz JR (1983) *Principles of Fluorescence Spectroscopy*, pp. 303–309, Plenum, New York.
4. Cheung H (1990) in *Fluorescence Spectroscopy: Principles and Applications*, (Lakowicz JR, Ed.), Plenum, New York, in press.
5. Forster T (1948) *Ann. Phys. (Leipzig)* 2, 55–75.
6. Lakowicz JR, Johnson ML, Wicz W, Bhat A, and Steiner RF (1987) *Chem. Phys. Lett* 138, 587–593. [PubMed: 31839681]
7. Lakowicz JR, Gryczynski I, Cheung HC, Wang CK, Johnson ML, and Joshi N (1988) *Biochemistry* 27, 9149–9160. [PubMed: 3242618]
8. Lakowicz JR, Gryczynski I, Wicz W, Laczko G, Prendergast FG, and Johnson ML (1990) *Biophys. Chem* 36, 99–115. [PubMed: 2207280]
9. Haas E, Wilchek H, Katchalski-Katzir E, and Steinberg IZ (1975) *Proc. Natl. Acad. Sci. USA* 72, 1807–1811. [PubMed: 1057171]
10. Amir D, and Haas E (1986) *Biopolymers* 25, 235–240. [PubMed: 2420384]

11. Amir D, and Haas E (1986) *Biochemistry* 26, 2162–2175.
12. Haas E (1986) in *Photophysical and Photochemical Tools in Polymer Science*, (Winnik MA, Ed.), pp. 316–341, Reidel, Holland.
13. Gratton E, and Limkeman M (1983) *Biophys. J* 44, 315–324. [PubMed: 6661490]
14. Lakowicz JR, and Maliwal BP (1985) *Biophys. Chem* 21, 61–78. [PubMed: 3971026]
15. Lakowicz JR, Laczko G, and Gryczynski I (1986) *Rev. Sci. Instrum* 57, 2499–2506.
16. Alcala JR, Gratton E, and Prendergast FG (1987) *Biophys. J* 51, 587–596. [PubMed: 3580485]
17. Alcala JR, Gratton E, and Prendergast FG (1987) *Biophys. J* 51, 597–604. [PubMed: 3580486]
18. Lakowicz JR, Joshi NB, Johnson ML, Szmazinski H, and Gryczynski I (1987) *J. Biol. Chem* 262, 10,907–10,910.
19. Joshi N, Johnson ML, Gryczynski I, and Lakowicz JR (1987) *Chem. Phys. Lett* 135, 200–207.
20. Laczko G, Gryczynski I, Gryczynski Z, Wicz W, Malak H, and Lakowicz JR (1990) *Rev. Sci. Instr.* 61, 2331–2337.
21. Lakowicz JR, Kusba J, Wicz W, Gryczynski I, and Johnson ML (1990) *Chem. Phys. Lett* 173, 319–326.
22. Lakowicz JR, Kusba J, Wicz W, Gryczynski I, Szmazinski H and Johnson ML (1991) *Biophys. Chem* 39, 79–84. [PubMed: 2012836]
23. Beechem JM, and Haas E (1989) *Biophys. J* 55, 1225–1236. [PubMed: 2765658]
24. Bevington PR (1969) *Data Reduction and Error Analysis for Physical Sciences*, p. 336, McGraw-Hill, New York.
25. Lakowicz JR, Gratton E, Laczko G, Cherek H, and Limkeman M (1984) *Biophys. J* 46, 463–477. [PubMed: 6498264]
26. Gratton E, Lakowicz JR, Maliwal B, Cherek H, Laczko G, and Limkeman M (1984) *Biophys. J* 46, 479–486. [PubMed: 6498265]
27. Lakowicz JR, Wicz W, Gryczynski I, Fishman M, and Johnson ML, in preparation.
28. Cheung HC, Gryczynski I, Malak H, Wicz W, Johnson ML, and Lakowicz JR (1990) *Biophys. Chem*, in press.
29. Cheung HC, Gonsoulin F, and Garland F (1985) *Biochem. Biophys. Acta* 832, 52. [PubMed: 2932161]
30. Johnson ML (1983) *Biophys. J* 44, 101–106. [PubMed: 6626675]
31. Johnson ML, and Frasier SG (1985) in *Methods in Enzymology*, (Hirs CHW, and Timasheff SN, Eds.), Vol. 117, pp. 301–342, Academic Press, San Diego.

FIG. 1.

Simulated frequency responses for the distance distribution model for unlabeled(---) and 100% labeled(—) donor-acceptor systems. The parameters used for the simulations were $\tau_D = 7$ ns, $R_0 = 25$ Å, $R_{av} = 20$ Å and $hw = 15$ Å. Frequency responses for intermediate labelings ($0 < L < 1$) are between the dashed and solid lines. The insert shows the influence of incomplete labeling on the distance distribution recovered with the assumption $L = 1$.

FIG. 2. χ^2_R surfaces for the parameters recovered from the distance distribution model (Eq. [16]) for the donor-acceptor systems with various degrees of labeling. During this χ^2_R analysis one of the three fitting parameters (L , R_{av} , hw) was kept constant at the fixed value and the other two were varied to minimize χ^2_R . The dashed lines show χ^2_R surfaces for the 50% labeled system analyzed with L fixed at 0.5. The dotted lines show the upper values of χ^2_R consistent with the χ^2_R values being due to random errors in 67% of multiple measurements.

FIG. 3. Absorption spectra of the TMA donor-control and TU2D donor-acceptor mixtures. The values of *L* refer to the molar proportion of TMA and TU2D, and A is the absorption spectrum of the acceptor, dansylamide.

FIG. 4.

Phase and modulation data for the TU2D donor-acceptor system containing unlabeled donor (TMA). The dashed line (---) shows the frequency response of the donor TMA. The lower panels show the deviations with $L = 0.494$ (●) and with L fixed at $L = \langle 1.0 \rangle$ (○).

FIG. 5.
 χ^2_R surfaces obtained for TU2D donor-acceptor system with low and high extents of labeling. The dashed lines show χ^2_R surfaces for $L = 0.494$ analyzed with L fixed at 0.494. The dotted lines indicate the highest values of χ^2_R consistent with random noise in 67% of repetitive measurements.

FIG. 6.

Donor-to-acceptor distance distribution recovered from the measurements of an incompletely labeled system. The molar proportion of TU2D is thought to be 0.715 (Table 2), and the recovered value was $L = 0.709$. The dashed line shows the distribution obtained from the data analysis with L fixed at 1.0. The dotted line shows the distribution recovered for a completely labeled system.

FIG. 7. Simulated frequency responses for the distance distribution model with different average distances (R_{av}). For these simulations we used $\tau_D = 20 \text{ ns}$, $R_0 = 30 \text{ Å}$, and $hw = 15 \text{ Å}$.

FIG. 8.

χ^2_R surfaces for the parameters obtained from the distance distribution model for systems with different average distances (---, $R_{av} = 20 \text{ \AA}$; —, $R_{av} = 30 \text{ \AA}$; -•-, $R_{av} = 40 \text{ \AA}$). For these simulations $\tau_D = 20 \text{ ns}$, $R_0 = 30 \text{ \AA}$, $hw = 15 \text{ \AA}$, and $L = 0.9$. The dotted lines indicate the upper values of χ^2_R consistent with random noise in 67% of repetitive measurements.

FIG. 9.

Phase and modulation data for myosin subfragment S1 with (●) and without (○) MgADP. The dotted lines (·····) show the best fit with the degree of acceptor labeling held fixed at $L = \langle 1.0 \rangle$.

FIG. 10. χ^2_R surfaces obtained for myosin subfragment S1 in the presence (---) and in absence (—) of MgADP. The dotted lines indicate the highest values of χ^2_R consistent with random noise in 67% of repetitive measurements.

FIG. 11. Distance distributions (Cys-697 to Cys-707) in myosin subfragment S1 in the presence and absence of MgADP. The dashed lines show the distributions obtained from the analyses with L fixed at 1.0.

TABLE 1

Distance Distribution Parameters Recovered from Simulated Data^a

Input		Recovered			
<i>L</i>	<i>f_{DA}</i>	<i>R_{av}</i> (Å)	<i>hw</i> (Å)	<i>L</i>	χ^2_R
1.00	1.00	20.02 (0.05) ^b	15.05 (0.29)	1.002 (0.004)	1.1
		20.02 (0.05)	14.94 (0.11)	<1.00> ^c	1.1
0.95	0.82	19.94 (0.06)	14.74 (0.30)	0.944 (0.005)	1.0
		19.92 (0.06)	15.15 (0.12)	<0.95>	1.0
		19.72	18.63	<1.00>	3.6
0.90	0.74	19.88 (0.06)	15.72 (0.35)	0.911 (0.006)	0.9
		19.91 (0.06)	15.00 (0.12)	<0.90>	1.0
		19.32	22.39	<1.00>	4.7
0.75	0.46	19.94 (0.07)	15.07 (0.41)	0.751 (0.006)	0.8
		19.94 (0.07)	15.02 (0.12)	<0.75>	0.8
		15.55	42.31	<1.00>	9.9
0.50	0.21	19.87 (0.12)	15.02 (0.71)	0.499 (0.010)	0.8
		19.87 (0.12)	15.07 (0.16)	<0.50>	0.8
		0.07	297.1	<1.00>	24.7
0.25	0.07	19.80 (0.24)	13.63 (1.46)	0.234 (0.012)	0.8
		19.79 (0.29)	15.63 (0.35)	<0.25>	0.8
		579.5	212.2	<1.00>	16.1
0.10	0.03	19.45 (0.93)	14.62 (5.58)	0.096 (0.020)	1.2
		19.41 (0.98)	15.61 (1.11)	<0.10>	1.2
		301.0	93.0	<1.00>	7.5

^aThe parameter values used in simulations were: $\tau_D = 7$ ns, $\alpha = 1.0$, $R_0 = 25$ Å, $R_{av} = 20$ Å, $hw = 15$ Å, $\delta\phi = 0.2^\circ$, and $\delta_m = 0.005$.

^bUncertainties as estimated for correlated parameters, as described by Johnson (30,31).

^cThe fractional labeling was held fixed at the value in the angular brackets, < >.

TABLE 2
Distance Distribution Parameters for TU2D in Propylene Glycol at 5°C with Adjusted Labeling ($R_0 = 25.7 \text{ \AA}$)

Adjusted		Recovered			
L	f_{DA}	$R_{av} \text{ (\AA)}$	$hw \text{ (\AA)}$	L	χ^2_R
1.0	1.00	19.50 (0.11)	17.16 (0.45)	1.003 (0.004)	0.8
		19.58 (0.07)	16.53 (0.14)	<1.0> ^a	0.9
0.947	0.86	19.53 (0.14)	17.11 (0.63)	0.950 (0.007)	1.4
		19.69 (0.09)	16.03 (0.17)	(0.947)	1.5
		18.72	21.16	<1.0>	2.3
0.715	0.37	19.55 (0.21)	15.44 (0.95)	0.709 (0.013)	1.1
		19.48 (0.13)	15.85 (0.15)	<0.715>	1.0
		0.01	71.43	<1.0>	11.0
0.494	0.22	18.68 (1.15)	19.13 (3.43)	0.541 (0.044)	1.2
		19.69 (0.25)	15.54 (0.21)	<0.494>	1.3
		713.5	496.4	<1.0>	18.2
0.236	0.17	13.92 (186.2)	75.8 (156.3)	0.759 (1.100)	1.3
		20.74 (0.52)	13.65 (0.49)	<0.236>	1.4
		641.5	207.3	<1.0>	12.7
0.089	0.03	10.97 (332.5)	106.65 (1,224.5)	0.429 (3.068)	1.3
		21.69 (1.04)	11.01 (1.30)	<0.089>	1.3
		288.9	88.47	<1.0>	5.1

^aValues of angular brackets < > were held fixed at the indicated values.

TABLE 3
Distance Distribution Parameters Recovered from Simulated Data for Conformational Changes of Macromolecules^a

Input			Recovered			χ^2_R
R_{sw} (Å)	L	f_{DA}	R_{sw} (Å)	hw (Å)	L	
20	0.9	0.57	19.90 (0.10)	15.17 (0.28)	0.901 (0.003)	0.7
		0.57	19.93 (0.05)	15.07 (0.07)	<0.90> ^b	0.7
			6.33	37.27	<1.0>	19.7
30	0.9	0.81	<20.0>	<15.0>	<1.0>	893.9
		0.81	29.92 (0.21)	14.50 (0.50)	0.896 (0.016)	0.8
			29.97 (0.08)	14.61 (0.17)	<0.90>	0.8
40	0.9	0.88	31.24	17.49	<1.0>	1.4
		0.88	<20.0>	<15.0>	<1.0>	38.1
			38.88 (3.22)	13.85 (3.92)	0.839 (0.181)	0.8
		0.88	40.00 (0.48)	15.12 (1.08)	<0.90>	0.8
			42.16	17.31	<1.0>	0.9
			<20.0>	<15.0>	<1.0>	2.4

^aThe parameters used for the simulations were $\tau_D = 20$ ns, $R_0 = 30$ Å, hw = 15 Å, $\delta\phi = 0.2^\circ$, and $\delta_m = 0.005$.

^bValues in angular brackets < > were held fixed at the indicated values.

Distance Distribution between Cys-697 and Cys-707 in Myosin Subfragment-1 in the Absence and Presence of MgADP

TABLE 4

Conditions ^a	f_{DA}	$R_{av}(\text{\AA})$	$hw(\text{\AA})$	L	χ^2_R
Buffer	0.85	29.64 (0.15)	13.39 (0.63)	0.935 (0.015)	1.6
	0.85	26.94 (0.05)	13.38 (0.21)	<0.935 ^b	1.5
MgADP		27.58	15.96	<1.0>	2.2
	0.72	20.73 (0.08)	12.76 (0.32)	0.940	1.8
	0.72	20.73 (0.06)	12.73 (0.11)	<0.940>	1.7
		18.71	20.13	<1.0>	14.6

Note. The values of R_0 were determined to be 29.3 Å in buffer and with MgADP (28).

^a All samples were in 60 mM KCl, 30 mM Tris, pH 7.5. The myosin S1 concentration was 30 μM, MgCl2 was 1 mM, and ADP was 100 μM.

^b Values in angular brackets < > were held fixed at the indicated values.