Determination of Intramolecular Distance Distribution Functions Using the "Spectroscopic Ruler". 1. Theoretical Feasibility

Guojun Liu

Department of Chemistry, University of Calgary, 2500 University Drive, NW, Calgary, Alberta, Canada T2N 1N4

Received May 27, 1992; Revised Manuscript Received August 25, 1992

ABSTRACT: The decay of the fluorescence intensity of an energy donor group, attached to one site of a macromolecule, due to dipole–dipole energy transfer to an acceptor group, attached to another site, can be analyzed to yield the distance distribution function between the two sites. In previous studies, a functional form for the distance distribution is usually assumed to facilitate data analysis to yield the floating parameters which determine the fine shape of the distribution function. In the present paper, a new procedure for treating the donor fluorescence intensity decay is proposed. Using the proposed data treatment procedure, one is now able to make an absolute determination of the distance distribution function. The successful use of such a procedure in recovering two types of distance distribution functions is demonstrated in this paper by analyzing computer-simulated decay curves.

I. Introduction

The "spectroscopic ruler" equation, derived by Förster[1] and verified experimentally first by Latt et al.[2] and Stryer et al.,[3] governs the dependence of rate constants k_{ET} for dipole–dipole electronic energy transfer on the separation distance R between an energetically excited donor D* and a ground-state acceptor A

$$k_{ET} = (R_0/R)^6/\tau_0 \tag{1}$$

In eq 1, $1/\tau_0$ is the decay rate constant of the excited donor group in the absence of energy transfer and R_0 is the critical energy transfer distance at which the rate constant for energy transfer is equal to the decay rate constant of the donor group. The magnitude of R_0 can be calculated using

$$R_0 = \left(\frac{9000 \ln 10 \kappa^2 \phi_D \mathbf{J}}{125\pi^5 n^4 N_A}\right)^{1/6} \tag{2}$$

where ϕ_D is the donor fluorescence quantum yield in the absence of energy transfer, n is the refractive index of the solvent at the wavelength of excitation, and N_A is Avogadro's constant. κ^2, the orientation factor, is defined by

$$\kappa^2 = \cos^2\alpha - 3\cos^2\beta\cos^2\gamma \tag{3}$$

where α, β, and γ are angles formed between the emission transition dipole of the donor and the absorption transition dipole of the acceptor, the emission transition dipole of the donor and the vector connecting the centers of the donor and the acceptor, and the absorption dipole of the acceptor and the distance vector. $\mathbf{J}$ is the overlap integral between the normalized fluorescence intensity, I_λ, of the donor and the extinction coefficient, ϵ_λ, of the acceptor at wavelength λ

$$\mathbf{J} = \int_0^\infty \lambda^4 I_\lambda \epsilon_\lambda \, d\lambda \tag{4}$$

where, by definition

$$\int_0^\infty I_\lambda \, d\lambda = 1 \tag{5}$$

The spectroscopic ruler equation has been widely used in determining intramolecular distances in proteins and enzymes[4] and in synthetic organic polymers.[5] In most of the studies, steady-state fluorescence techniques were used to obtain the efficiency of energy transfer from donors to acceptors. The energy transfer efficiency was then used to extract an ensemble-averaged separation distance between the donor and acceptor sites, provided that the distribution function for the distances between the donor and the acceptors is known.

Katchalski-Katzir et al.[6] were the first ones to use time-correlated single-photon-counting (TCSPC) techniques to study intramolecular energy transfer. They used a series of donor and acceptor end-labeled oligopeptides in an attempt to derive information regarding the end-to-end distance distribution function of the studied peptide chains. They showed that the fluorescence intensity of the donor group after δ-pulse excitation should decay according to

$$I(t) \propto \int_0^{R_{mx}} P(R) \exp[-(t/\tau_0) - (t/\tau_0)(R_0/R)^6] \, dR \tag{6}$$

provided that the medium used for energy transfer study is highly viscous and the end-labeled peptide is at high dilution in the medium. The high viscosity ensures that no significant end group diffusion occurs during the excitation lifetime of the donor group. The high dilution of the end-labeled sample is to make sure that intermolecular energy transfer is minimized.

In eq 6, $P(R)$ is the end-to-end distance distribution function of the peptide chains studied; $P(R)\,dR$ is the probability of finding chains which possess an end-to-end distance between R and $R + dR$; the integration over R goes from zero to R_{mx}, the maximum end-to-end distance of a chain. It is obvious from eq 6 that the key to the knowledge of the end-to-end distance distribution function of an oligopeptide chain is buried in the decay curve of the donor group.

Although eq 6 is derived for the case of end-labeled peptide chains, it is, however, applicable to distance distributions of any other origin. One can, for example, place a donor group at one end of a polymer chain and then an acceptor group x segments away from the donor group and y segments away from the other end of the chain. In this case, $P(R)$ is the distance distribution function between the labeled end and segment x.

When the donor decay data for end-labeled peptides were treated, Katchalski-Katzir et al.[6] assumed the following form for the end-to-end distance distribution

0024-9297/93/2226-1144$04.00/0

function $P(R)$:

$$P(R) = 4\pi R^2 \exp[-a(R-b)^2]$$

where a and b were fitting parameters. The values of a and b were determined by fitting experimentally obtained decay curves using eq 6. In their case the excitation pulse was not sufficiently narrow to be approximated as a δ-pulse and due account was given to the finite excitation pulse width in data treatment.

The data treatment procedure of Katchalski-Katzir et al. has been adopted for use by other researchers. Lackowicz et al.,[7] for example, used the data treatment procedure for revealing information about the distance distribution between sites in proteins in their native and denatured states.

The disadvantage of the data treatment procedure of Katchalski-Katzir et al. is that one needs preknowledge of the form of the distance distribution function of donor and acceptor sites. In cases when such preknowledge is absent or when one is asked to make an absolute determination of the distance distribution function of donor and acceptor sites, the Katchalski-Katzir procedure fails. Furthermore, a satisfactory fit of experimental data by making use of a preassumed shape for the distance distribution function only indicates that the assumed distribution is compatible with the experimental data. It does not demonstrate that this distribution is unique.

In this paper, an alternative procedure for the treatment of fluorescence donor decay data is proposed. In the proposed treatment method, no a priori functional forms are assumed for $P(R)$. Instead, a computer program used is asked to make an absolute determination of $P(R)$ from donor fluorescence intensity decay data. In section II, the theoretical rationale for the suggested data treatment is presented. In section III, computational algorithms are presented and discussed. The theoretical feasibility of such a treatment procedure to recover $P(R)$ is demonstrated by analyzing computer-simulated decay curves in section IV. The paper finishes by drawing some conclusions in section V.

II. Theoretical Consideration

Further Comments on Equation 6. In addition to the high-viscosity and high-dilution assumptions which have been mentioned, the derivation of eq 6 also assumed that the fluorescence of a donor group attached to one end of an oligopeptide chain in the absence of energy transfer decayed single exponentially and that the acceptor group at the other end did not introduce any additional quenching mechanism other than the quenching of donor fluorescence via dipole–dipole energy transfer. Furthermore, it was assumed that before energy transfer takes place, the donor and acceptor groups have experienced many different relative orientations due to Brownian rotation of the donor and acceptor groups so that the relative orientation factor κ^2 assumes a statistical value of $^2/_3$.

The last assumption about the average κ^2 value of $^2/_3$ appears difficult to achieve experimentally. First, on one hand, one needs a highly viscous solvent to ensure that donor and acceptor groups do not undergo significant translational motion in space during the excitation lifetime of the donor group; i.e., the root-mean-square distance traveled by the donor or the acceptor group is much smaller than R_0. On the other hand, one requires that the viscosity is low enough to allow the rotational freedom of the donor and acceptor groups. Considering the case of an end-labeled flexible polymer chain, the relative position change of the end groups is caused by the rotations of each segment of the chain. The slowing down of the relative diffusional motion between the ends is the result of the slowing down of segmental rotation. Thus, the finding of a solvent which ensures that the latter is still significant and the former motion is insignificant is difficult. Then, even if such a solvent can be found, the assumed relation, $\kappa^2 = {}^2/_3$, may not be valid for those donor and acceptor pairs which are close in space, because energy transfer is very fast for those pairs. The large energy transfer rate constant makes it impossible for those groups to experience different relative orientations before energy transfer takes place.

The aforementioned difficulties can, however, be overcome by making use of donor and acceptor groups with degenerate transitions.[8] It has been reported that the use of a transition metal ion with a purportedly triply degenerate transition as the acceptor reduced the uncertainty in the κ^2 value to $\pm 12\%$ of $^2/_3$.[9] The use of a donor and acceptor pair, both with purportedly triply degenerate transitions, has been believed to removed all uncertainty associated with the κ^2 value of $^2/_3$.[10]

The derivation of eq 6 was based on the following argument. After the excitation of the donor group by δ-pulse radiation, the intensity of donor emission decreases (1) due to the self-deactivation with rate constant $1/\tau_0$ and (2) due to the transfer of energy from D* to A with rate constant k_{ET} as defined by eq 1. For the fraction $P(R)\,dR$ of oligopeptide chains which have an end-to-end distance between R and $R + dR$, the donor fluorescence intensity decays according to

$$I(R,t) \propto P(R) \exp[-(t/\tau_0) - (t/\tau_0)(R_0/R)^6]\,dR \qquad (7)$$

For a system of chains of various end-to-end distances, the fluorescence intensity decays according to eq 6.[6]

Approximation of Equation 6 by an Exponential Series. Regardless of the functional form of $P(R)$, the integration of eq 6 can be approximated by a series of exponential terms. That is, the whole integration range, 0 to R_{mx}, can be divided into N equivalent subintervals of increment $\delta R = R_{mx}/N$, where N is, for example, ~ 100. At subinterval i, $R_i = (i - {}^1/_2)\,\delta R$ as follows from the trapezoidal rule. It is further defined

$$a_i = P(R_i)\,\delta R \qquad (8)$$

and

$$1/\tau_i = 1/\tau_0 + (1/\tau_0)(R_0/R_i)^6 \qquad (9a)$$

or

$$\tau_i = \tau_0 \frac{R_i^6}{R_0^6 + R_i^6} \qquad (9b)$$

Then eq 6 becomes

$$I(t) \propto \sum_{i=1}^{N} a_i \exp(-t/\tau_i) \qquad (10)$$

where τ_i can be calculated from eq 9b and are known quantities once R_{mx} and N are known; $I(t)$ is the intensity of fluorescence emission from the donor group at time t after δ-pulse excitation and can be monitored experimentally; the a_i's, proportional to $P(R_i)$, are unknown.

Convoluted Form of Equation 6. The derivation of eq 6 is based on the δ-pulse assumption. In reality, an excitation pulse may not be extremely short-lived. Thus, the distribution in the intensity of an excitation source as a function of time needs to be taken into account. Denoting the intensity of an excitation source at time t as $L(t)$ and

the theoretically calculated experimental emission intensity as $I_{th}(t)$, $I_{th}(t)$ and the donor fluorescence decay law $I(t)$ defined by eq 6 are related by[11]

$$I_{th}(t) \propto \int_0^t L(s)\, I(t-s)\, ds \tag{11}$$

The $I(t-s)$ term given by eq 6 can be approximated by a sum of exponential terms as has been discussed

$$I(t-s) \approx \sum_{i=1}^{N} a_i \exp[-(t-s)/\tau_i] \tag{12}$$

Inserting eq 12 into eq 11 leads to

$$I_{th}(t) = \sum_{i=1}^{N} a_i \int_0^t L(s) \exp[-(t-s)/\tau_i]\, ds \tag{13}$$

The integral of eq 13 can be solved numerically using the trapezoidal rule. Denoting the time interval between two adjacent data points or channels as Δt, the signal intensity in channel j or at time $t = j\Delta t$ is[11]

$$I_{th}(j) = \sum_{i=1}^{N} a_i y(i,j) \tag{14}$$

where $y(i,j)$ can be calculated using[11]

$$y(i,j) = \exp(-j\Delta t/\tau_i) \sum_{k=1}^{j} \{\exp(k\Delta t/\tau_i)\, L(k) + \exp[(k-1)\Delta t/\tau_i]\, L(k-1)\}(\Delta t/2) \tag{15}$$

Values $I_{th}(j)$ are, in principle, the fluorescence intensities which should be observed experimentally in the absence of experimental error. In reality, $I_{ex}(j)$ are observed and $I_{ex}(j)$ values and eq 14 are used to recover a_i values.

Solution of Simultaneous Linear Equations for a_i. Suppose that decay curve $I_{ex}(t)$ consists of $\mathcal{N}$ data points. It is immediately obvious that eq 14 is the general expression for $\mathcal{N}$ simultaneous linear equations with N unknown a_i values. The solution of N unknowns from $\mathcal{N}$ simultaneous equations is a trivial problem if $\mathcal{N} \geq N$. However, this would be the case only if the time resolution of the instrument is extremely high so that it can resolve the smallest τ_i, i.e., τ_1, and the data precision is also extremely high so that error in each data point is negligible. In reality, data precision and instrument resolution are not extremely high, and solution of a_i by solving simultaneous linear equations is impractical.

Minimization of χ^2 for a_i. In practice, a_i are solved by minimizing the χ^2 function defined by

$$\chi^2 = [1/(\mathcal{N}_2 - \mathcal{N}_1 + 1)] \sum_{j=\mathcal{N}_1}^{\mathcal{N}_2} \frac{[I_{ex}(j) - I_{th}(j)]^2}{I_{ex}(j)} \tag{16}$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ are the first and the last data points of a decay curve being fitted, $(\mathcal{N}_2 - \mathcal{N}_1 + 1)$ is the total number of data points used for curve fitting, $I_{ex}(j)$ and $I_{th}(j)$ are the fluorescence intensity experimentally observed and that calculated using eqs 14 for the jth data point. Obviously, the smaller the χ^2, the better is the agreement between an experimentally obtained curve and one which is theoretically calculated.

The feasibility of recovering the coefficients of N exponential terms correctly, where N is as large as 200, from minimizing χ^2 functions by a technique termed the exponential series method (ESM) has been demonstrated by work of Ware and co-workers.[12-14] Bronchon et al.[15,16] used the maximum entropy method (MEM), a supposedly more robust version of ESM, for recovering a_i values. In using the MEM, a function defined as the Shannon–Jaynes entropy is maximized subject to the constraint that the χ^2 function is first minimized and then maintained close to the minimal value. Ware and co-workers showed that MEM is not as effective as ESM when used for recovering a_i values from fluorescence quenching data of donor molecules in the presence of energy acceptors.[12] In this paper, data will be analyzed by using EMS only.

Values of χ^2 are quadratic functions of a_i values. The minimization of a quadratic function with N variables has been extensively studied in the past.[17,18] Various algorithms for function minimization are available from the International Mathematical and Statistical Library (IMSL). However, the use of those routines can be very restrictive because one cannot manipulate those routines, e.g., the slowing down of those routines to decrease the size of searching steps, at will. A commercial program specifically written for treating fluorescence decay data using ESM can be purchased from Photon Technology International, Inc. (PTI). The restriction here is again the fact that one has to accept whatever is provided in the package. The program can only be run on an IBM compatible personal computer, which is speed-limited and thus reduces the number of exponential terms which one can use for data treatment. For these reasons, we have developed our own program for ESM analysis. Since the algorithm of the ESM program from PTI has not been described anywhere in detail, the mathematics and the algorithm of our program are presented here in some detail in the hope that interest researchers can conveniently produce their own version of the program.

Review of Mathematics for Function Minimization Using the Conjugated Gradient Method. Conjugated gradient method locates the minimum of function χ^2 along conjugated search directions[17] in the N-dimensional a_i space. Suppose that after k search steps, a function minimization program has arrived at a point represented by the end point of a position vector $\mathbf{a}_k$ (called point $\mathbf{a}_k$ hereafter) from an initial point $\mathbf{a}_0$. The program then checks whether point $\mathbf{a}_k$ is the desired minimum based upon criteria provided for judging the locus of a minimum. If the desired χ^2 minimum has not been located, it will make the next search for a new point $\mathbf{a}_{k+1}$. The $\mathbf{a}_{k+1}$ point is chosen by minimizing function χ^2 along a line which is parallel to search direction $\mathbf{s}_k$ and passes through the point $\mathbf{a}_k$. The search direction $\mathbf{s}_k$ is conjugated to the previous line search direction $\mathbf{s}_{k-1}$.[17]

One way of constructing conjugated search directions is to use the following formula:

$$\mathbf{s}_k = -\mathbf{g}^k + h_k \mathbf{s}_{k-1} \qquad k = 0, 1, 2, \ldots \tag{17}$$

where $\mathbf{g}_k$ is the gradient vector of χ^2 at point $\mathbf{a}_k$. The scalar quantity h_k is calculated using

$$h_k = 0, \quad \text{if } k = 0 \tag{18a}$$

$$h_k = (\mathbf{g}_k)^2/(\mathbf{g}_{k-1})^2, \quad \text{if } k > 0 \tag{18b}$$

The N components of $\mathbf{g}_k$, i.e., partial derivatives $\partial\chi^2/\partial a_i$, can be easily shown from eqs 16 and 14 to be

$$\partial\chi^2/\partial a_i = -[2/(\mathcal{N}_2 - \mathcal{N}_1 + 1)] \sum_{j=\mathcal{N}_1}^{\mathcal{N}_2} \frac{I_{ex}(j) - I_{th}(j)}{I_{ex}(j)} y(i,j) \tag{19}$$

The position vector $\mathbf{p}_k$ of any point on the line which is parallel to $\mathbf{s}_k$ and passes through point $\mathbf{a}_k$ can be expressed as the sum of position vector $\mathbf{a}_k$ and direction

vector $\mathbf{s}_k$

$$\mathbf{p}_{k+1} = \mathbf{a}_k + t_k\mathbf{s}_k \tag{20}$$

where t_k is a scalar variable. The $(k+1)$th point in the minimization of χ^2 in a_i space is chosen using

$$\mathbf{a}_{k+1} = \mathbf{a}_k + t_k^*\mathbf{s}_k \tag{21}$$

where t_k^* is the t_k value which minimizes χ^2 along the search line. The value of t_k^* is determined by taking the partial derivative of χ^2 with respect to t_k and setting $\partial\chi^2/\partial t_k$ to zero. The expression for calculating t_k^* is[17]

$$t_k^* = -\frac{\mathbf{s}_k^{\mathrm{T}}\mathbf{g}_k}{\mathbf{s}_k^{\mathrm{T}}\mathbf{B}\mathbf{s}_k} \tag{22}$$

where $\mathbf{s}_k^{\mathrm{T}}$ denotes the transpose of vector $\mathbf{s}_k$; $\mathbf{B}$ is the curvature or Hessian matrix for χ^2.[17] Components $B(i,j)$ or $\partial^2\chi^2/\partial a_i\partial a_j$ of the Hessian matrix can be shown from eqs 14 and 19 to be

$$B(i,j) = [2/(\mathcal{N}_2 - \mathcal{N}_1 + 1)]\sum_{k=\mathcal{N}_1}^{\mathcal{N}_2}\frac{y(i,k)\,y(j,k)}{I_{\mathrm{ex}}(k)} \tag{23}$$

where $B(i,j)$ are independent of the values of a_i.

III. Computational Techniques

All programs are written in Fortran. Programs are compiled and executed on an IBM RISC system/6000 Model 320 workstation. A typical run for decay curve generation takes fractions of a second. The speed of the ESM program varies depending on factors such as the number of data points $(\mathcal{N}_2 - \mathcal{N}_1 + 1)$ to be fitted and the number of exponential terms N used in curve fitting. For $N = 100$, $(\mathcal{N}_2 - \mathcal{N}_1 + 1) = 965$, the ESM program using the conjugated gradient algorithm run at its full speed locates the χ^2 minimum in less than 10 s.

Generation of Decay Curves. Fluorescence decay curves are generated using eqs 14 and 15. Our program starts by prompting for input of an N value, typically 1000 for decay curve generation, and parameters related to $P(R)$. The inputted $P(R)$ function and eq 8 are then used to calculate a_i values. Afterward, R_0 and τ_0 are read in and τ_i is calculated using eq 9b. The $L(s)$ profile used for decay curve generation in this paper is that acquired on a TCSPC system using a picosecond laser as the excitation source. $L(s)$ usually consists of 1024 data points. The time range T used for registering the decay in $I_{\mathrm{ex}}(t)$ is then $T = 1024 \times \Delta t$, where Δt is the time increment between two adjacent channels. The value of Δt in this study is set to either 0.0400 or 0.0430 ns, because it is the setting which we use for time-resolved fluorescence studies of fluorene and anthracene end-labeled poly(methyl methacrylate) samples, which will be the topic of discussion for the second paper in the sequence on this topic.[19]

Numerical solution of eqs 14 and 15 is achieved by making use of a standard routine from the monograph by O'Connor and Phillip.[11] Decay curves generated using eq 14 are smooth and free of statistical experimental errors. They are made to resemble experimental data by adding synthetic noise to each data point drawn from a Gaussian probability distribution of zero mean and variance equal to the fluorescence counts $I_{\mathrm{ex}}(t)$ at that point. The generation of Gaussian noise is achieved by making use of an IMSL subroutine supported on the IBM workstation.

Recovery of $P(R)$ from Analysis of Synthesized Decay Curves Using ESM. Examination of eqs 17–22 indicates that a program which minimizes χ^2 in a_i space using the conjugated gradient method requires the following components: (1) a decay curve $I_{\mathrm{ex}}(t)$ which is to be fitted together with the intensity profile $L(t)$ of the excitation source; (2) a starting point or the initial guess of a_i values; (3) τ_i values which will enable the calculation of $y(i,j)$ and $B(i,j)$; (4) a subroutine for the calculation of components of $\mathbf{g}_k$, provided point $\mathbf{a}_k$ is given; (5) a subroutine for the calculation of χ^2; (6) the main program which calculates h_k, constructs $\mathbf{s}_k$, and calculates t_k^*.

Based on the foregoing analysis, the algorithm of our program goes as follows:

1. A file containing the donor fluorescence decay data and the profile of the excitation source is read in.

2. Values of N, R_0, τ_0, R_1, and R_x are inputted, where R_1 and R_x are the smallest and largest fitting distances for $P(R)$, respectively. The value of N, although it can go as large as the workstation can handle, is typically chosen to be smaller than 200, because the use of larger N numbers would not change the shape of recovered $P(R)$ but needs much longer computational time.

3. τ_i values are calculated using eq 9b.

4. Components of $\mathbf{a}_0$ are all assumed equal; i.e., $a_1 = a_2 = \ldots = a_N = 1/N$ at point $\mathbf{a}_0$.

5. Matrix $y(i,j)$ is constructed using eq 15.

6. Matrix $B(i,j)$ is constructed using eq 23.

7. Function χ^2 is minimized via the following steps: (a) establish the starting point $\mathbf{a}_k$, where $k = 0, 1, 2, \ldots$; (b) call χ^2 subroutine to calculate χ^2 at point $\mathbf{a}_k$ using eq 16 (if the χ^2 value is satisfactory, go to 8); (c) call gradient subroutine to calculate components of $\mathbf{g}_k$ using eq 19; (d) calculate h_k using eq 18 and construct search direction $\mathbf{s}_k$; (e) calculate t_k^* using eq 22 and calculate a_i values for the $(k+1)$th iteration; (f) return to (a).

8. Save R_i values and values of a_i and χ^2 at the last point in a_i space. Save weighted residues of fitting and calculate autocorrelation between weighted residues.

Since τ_i in using the ESM program are so chosen that $R_{i+1} - R_i$ is constant for all i, a_i recovered is directly proportional to $P(R_i)$ as has been argued previously.

Steepest Descent Algorithm. A program which minimizes χ^2 using the steepest descent method can be easily obtained by slightly modifying the conjugated gradient algorithm. In using the steepest descent method, the $(k+1)$th search direction is always the same as the negative gradient direction; i.e., $\mathbf{s}_k = -\mathbf{g}_k$. That is, the h_k value of eq 17 is always zero.

Algorithm Stopping Criterion. The change in χ^2, $\Delta\chi^2$, for two consecutive iterations has been used as the criterion for determining whether the minimal χ^2 has been located. That is, if the $\Delta\chi^2$ value from two consecutive iterations is smaller than a preassigned cut-off $\Delta\chi^2$ value denoted as $\Delta\chi^2_0$, the program saves R_i and a_i values and stops. For the fitting of decay curves generated using eq 14 in the absence of added noise, the choice of $\Delta\chi^2_0$ is immaterial as long as it is less than 10^{-4} and larger than a computer's floating point precision. For noise-added decay curves, the choice of a $\Delta\chi^2_0$ value is tricky. On one hand, a too small $\Delta\chi^2_0$ value will tend to overfit the data, i.e., the introduction of artificial oscillations in a_i values when plotted against τ_i. On the other hand, a too large $\Delta\chi^2_0$ will not guarantee the correct recovery of a_i values. Fortunately, between a too large and a too small $\Delta\chi^2_0$, there is a range of $\Delta\chi^2_0$ values, the use of which will yield an almost invariant $P(R)$. The $\Delta\chi^2_0$ value range increases with the precision of the decay data fitted. The more the counts in the peak channel, the larger the usable $\Delta\chi^2_0$ value range is. For a decay curve with peak channel counts larger than 10^6, the usable $\Delta\chi^2_0$ value range is typically

between 10^{-2} and 10^{-7} if t_k^* calculated from eq 22 is multiplied by 0.50 before its use in eq 21 to slow down the search for minimum at each iteration of the conjugated gradient algorithm.

The best fit to a noise-added $I_{ex}(t)$ curve using various versions of ESM programs should have χ^2 values close to 1.0. If noise is not added, χ^2 have values typically smaller than 0.001.

Comparison of Different Algorithms. The conjugated gradient method is extremely efficient and can locate the minimum in typically less than 100 iterations. The program is extremely powerful when used for treating noise-free decay curves. However, due to the giant searching steps it takes, the conjugated gradient algorithm can easily cross the boundary between what is an acceptable fit and what is an excessive fit, i.e., an overfit which introduces artificial oscillations in a_i values, when used for treating a noise-added decay curve. The boundary between what is an acceptable and an excessive fit can be made "wider" or more "visible" by slowing down the program. Instead of using t_k^* in eq 21 to calculate $\mathbf{a}_{k+1}$, one can each time replace t_k^* by a reduced value, e.g., $0.5t_k^*$.

The steepest descent algorithm is extremely inefficient. It seems to be less efficient than the ESM program purchased from PTI, although no exact comparison was made between the two programs due to the fact that they were run on different computers. The advantage of the steepest descent algorithm is that it has fewer chances of overfitting a decay curve.

IV. Results and Discussion

In this section, the feasibility for the successful use of the data treatment procedure described in the previous section in recovering $P(R)$ for a few practical situations will be illustrated. First, the possibility of carrying out time-resolved fluorescence studies of donor and acceptor labeled proteins to reveal conformations of proteins is demonstrated. Then, the use of the technique for determining end-to-end distance distribution functions of flexible polymer chains is discussed.

Conformations of Proteins. In previous energy transfer studies, proteins in their native state have been assumed to be rigid and thus sites of interest are assumed to be separated by single distances. In reality, proteins are never at rest and fluctuations in their structure take place constantly.[20] The difference between protein structural fluctuations and those of a flexible polymer chain or denatured proteins is in magnitude, and structural fluctuations in native proteins are obviously much less in degree than denatured proteins. If the structure of a protein fluctuates, it is expected that at some instant in time the same two sites of different molecules might have different separation distances because the fluctuation of protein molecules cannot be all synchronized. In fact, a distance distribution is expected for the two sites. The width of the distribution function tells us to what degree the distance between the two sites fluctuates. If two sites of interest in a protein can be labeled with energy donors and acceptors, the fluorescence technique proposed in this paper will be able to recover the distribution function of distances between the two sites. If the distribution function of distances between a few pairs of sites of a protein can be determined, a detailed picture as to how and how much the structure of a protein fluctuates can be obtained. Also, the current technique can be used to reveal the conformational change of proteins from their native states to denatured states upon the addition of denaturants as was done by Lakowicz et al.[7]

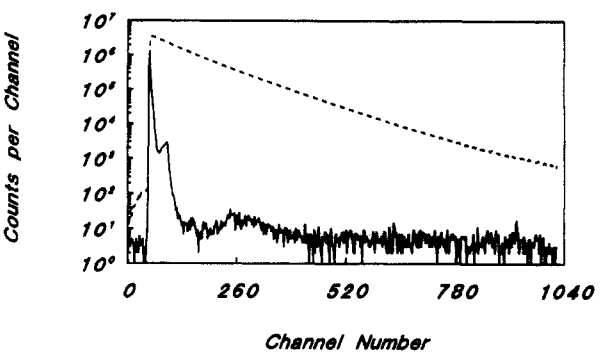

Figure 1. Donor decay curve (- - -) generated using eqs 14, 15, 6, and 24 and the following parameters: $R_a = 25$ Å, $\sigma = 3$ Å, $R_0 = 25$ Å, $\tau_0 = 6.65$ ns, and $\Delta t = 0.0400$ ns. The solid curve represents a laser profile acquired experimentally.

Lakowicz et al. monitored the conformational change of melittin upon the addition of water, a denaturant, to methanol containing α-helical melittin. As the water content increased, the magnitude of structural fluctuations increased as was revealed by using a spectroscopic ruler consisting of an intrinsic tryptophan group as the donor and an attached dansyl group as the acceptor. In recovering the distance distribution function, $P(R)$ was assumed to possess the following functional form

$$P(R) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\{-(1/2)[(R - R_a)/\sigma]^2\} \qquad (24)$$

where R_a and σ, fitting parameters in their treatment, are the average separation distance and the standard deviation in the distance distribution, respectively.

Suppose that $P(R)$ given by eq 24 is indeed the distance distribution function sought. The current data treatment procedure is able to recover $P(R)$ without resorting to assumptions about the functional form.

Shown in Figure 1 is a noise-added decay curve generated from eqs 14, 6, and 24 using the following parameters: $R_a = 25$ Å, $\sigma = 3$ Å, $R_0 = 25$ Å, $\Delta t = 0.040$ ns, and $\tau_0 = 6.65$ ns. Fitting of the decay curve using our ESM program yielded the lifetime distribution shown in Figure 2a. The autocorrelation function between weighted residues of fitting[11] is presented in Figure 2b. The autocorrelation function is completely randomized, which suggests that the recovered a_i values are statistically acceptable.

As discussed before, our program chooses τ_i to ensure that $R_{i+1} - R_i$ is constant. Under such circumstances, a_i is directly proportional to $P(R_i)$. Using the data of Figure 2a, the recovered $P(R)$ is obtained and presented in Figure 3.

Also shown in Figure 3 is the distance distribution function $P(R)$ which was inputted to generate the decay curve of Figure 1. The agreement between the inputted and the recovered $P(R)$ is excellent.

The superb agreement between the inputted and the recovered $P(R)$ is rewarding. The recovered $P(R)$ should be the same as that inputted and be unique. However, this is true only if the precision and time resolution of fluorescence decay data used for analysis are both extremely high. When such data are plagued with statistical errors and suffer from limited time resolution, many statistically acceptable $P(R)$ may exist and the one which is recovered may not be the one which is inputted.

Further examination of the decay curve of Figure 1 shows that counts at its peak channel is 3.55×10^6. This number is large but achievable experimentally by making use of a laser as the excitation source. As a matter of fact, the accurate recovery of the present $P(R)$ requires much fewer

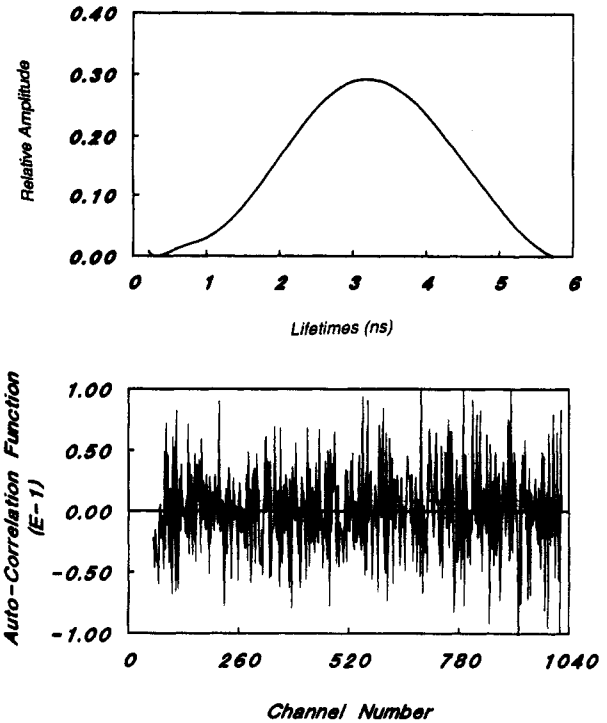

Figure 2. (a, Top) Lifetime distribution recovered from fitting the decay data shown in Figure 1 using our ESM program. The final χ^2 is 1.034. (b, Bottom) Autocorrelation function of the weighted residues of fitting.

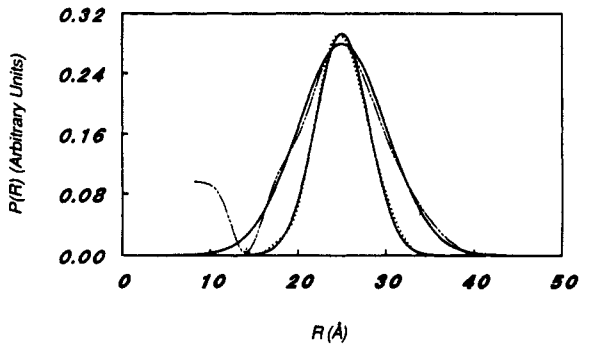

Figure 3. Comparison of inputted and recovered distance distribution functions. Curves (—) are distance distribution functions calculated from eq 24 using R_a = 25 Å and σ = 3 Å and σ = 5 Å, respectively. Curves (···), σ = 3 Å, and (- - -), σ = 5 Å, are those recovered from analyzing simulated fluorescence decay curves.

counts than is shown here, due to the narrowness of the $P(R)$ distribution. The effect of data precision or signal to noise ratio, which increases with the square root of peak channel counts, on the correct recovery of $P(R)$ will be discussed later.

Figure 3 also contains the inputted and recovered distance distribution functions for R_a = 25 Å and σ = 5 Å. The agreement between the two in the higher end of R is excellent. The agreement at the lower end is poor. Increasing the precision of the generated fluorescence decay data did not improve the shape of recovered $P(R)$. The poor agreement at the lower end of $P(R)$ can be attributed to the limited time resolution of the decay data. The time increment, Δt, of the decay data is 0.040 ns. Any lifetime below this value is unlikely to be resolved by the program. In fact, Livesey and Brochon[16] have suggested that amplitude coefficients of lifetimes below $5\Delta t$ could not be properly recovered by an MEM program. Supposing that the coefficients for lifetimes which are larger than $5\Delta t$ are acceptable, one can conclude that $P(R)$ values with R larger 14.0 Å or $0.56R_0$ should agree with the inputted $P(R)$ well, where R = 14.0 Å is calculated using eq 9 and $\tau_i = 5\Delta t$, i.e., 0.20 ns. It can be seen from Figure 3 that after 14.0 Å, the agreement between inputted $P(R_i)$ and recovered $P(R_i)$ improves quickly.

Upper and Lower Effective Limits of a Spectroscopic Ruler. The correctly-recoverable $P(R_i)$ values are those with R_i larger than some critical value denoted by R_1*. The magnitude of R_1*, as has been discussed, is governed by the shortest lifetime τ_1* with a correctly-recoverable amplitude coefficient. The value of τ_1* can range from Δt to $5\Delta t$, depending on the distribution function $P(R)$. As will be seen later, τ_1* can indeed go to as low as Δt under favorable circumstances. Suppose that $\tau_1^* = \Delta t$, i.e., 0.04 ns; the R_1* value calculated for this case is $0.43R_0$. Using $\tau_1^* = 5\Delta t$, $R_1^* = 0.56R_0$ is calculated. Thus, it is concluded that the effective lower range of a spectroscopic ruler is $\sim (1/2)R_0$.

The upper limit of resolvable τ is more difficult to set. Our experience suggests that the upper effective range of a spectroscopic ruler is around $1.5R_0$. The upper effective range depends on the precision of decay data and other factors. The reason for the existence of such a upper effective range can be understood as follows. Suppose that there is a distance distribution function that spans the distance range from zero to $5R_0$. If the whole $5R_0$ range is sliced into 100 equal intervals, the increment in R is $0.05R_0$. Assuming τ_0 = 6.65 ns, the values of τ_i calculated from eq 9b for i = 99 and i = 100 are 6.649534 and 6.649561 ns, respectively. Due to the small difference between the two lifetimes and the relatively large time Δt, i.e., 0.040 ns in the present case, it is very unlikely that the program would resolve a_i values for these two components.

Conformations of Polymer Chains. The end-to-end distance distribution function of polymer chains is an important characteristic function for polymer conformations. The attachment of a donor group to one end of a chain and the attachment of an acceptor group to the other has previously allowed the determination of ensemble-averaged end-to-end distances.[5] The use of the present technique will allow the direct determination of the end-to-end distance distribution function of the connecting chain, a task which has never been achieved before.

In a Θ solvent, the end-to-end distance distribution function of a polymer chain is given by[22]

$$P(R) = 4\pi R^2\left(\frac{3}{2\pi R_n^2}\right)^{3/2} \exp[-(3/2)(R/R_n)^2] \quad (25)$$

where R_n is the root-mean-square end-to-end distance. Suppose that fluorene and 1,8-diphenyloctatetraene can be attached to the opposite ends of a polymer chain. The critical energy transfer distance between fluorene and 1,8-diphenyloctatetraene should be close to 40 Å. The molecular weight of the polymer can be so controlled that R_n = 40 Å. Using $R_0 = R_n$ = 40 Å, τ_0 = 6.65 ns, and a lamp profile with its shape the same as that shown in Figure 1 but with Δt = 0.0435 ns, a decay curve for fluorene has been generated for such a system. The counts at the peak and final channel were found to be 2.7×10^6 and 750, respectively.

Statistical noises were added to the decay curve and the noise-added curve was then fitted using our ESM program to obtain the end-to-end distance distribution function. Illustrated in Figure 4 is the comparison between the recovered and the inputted end-to-end distance distribution function. The recovered end-to-end distance distribution function resembles the inputted one in the effective range of the spectroscopic ruler, i.e., between 20 and 60 Å. Deviations of the recovered $P(R)$ from the

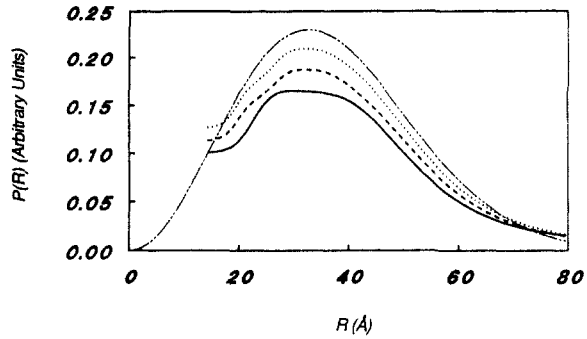

Figure 4. Comparison of inputted (- - -) and recovered end-to-end distance distribution functions of polymer chains. The fluorescence decay curves were generated by assuming $R_0 = R_n = 40$ Å, $\tau_0 = 6.65$ ns, and $\Delta t = 0.0435$ ns. Curve (—) is $P(R)$ recovered from the noise-added decay curve. Curves (- - -) and (···) are $P(R)$ recovered from noise-free decay curves generated from using the laser profile of Figure 1 with the tail included and cut, respectively.

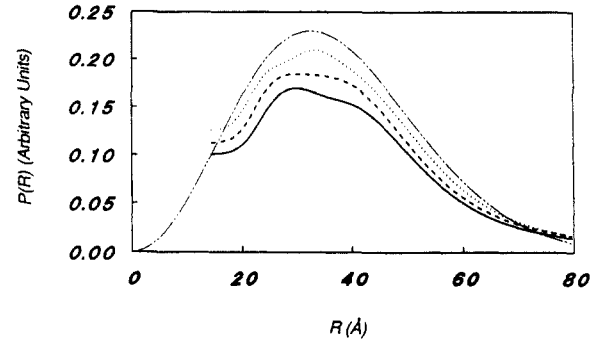

Figure 5. Comparison of $P(R)$ recovered from curves with different peak channel counts. From the bottom up, the peak channel counts of decay curves used for recovering $P(R)$ are 6.74×10^5, 2.7×10^6, and 1.08×10^8, respectively. The last curve (- - -) is the inputted $P(R)$.

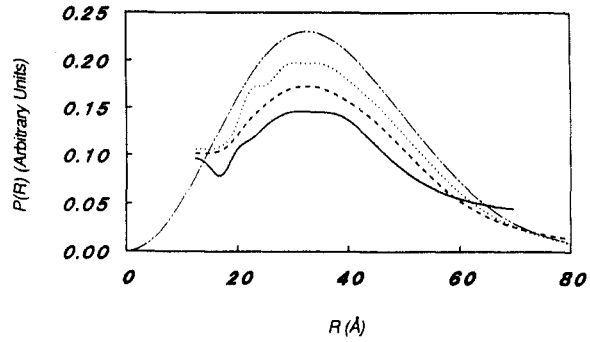

Figure 6. Comparison of $P(R)$ recovered from curves with different R_n/R_0 ratios. All decay curves were generated with R_n fixed at 40 Å. From the bottom up, R_0 values used for decay curve generation were 35, 40, and 45 Å, respectively. The last curve (- - -) is the inputted $P(R)$.

inputted one can be largely attributed to the low signal to noise ratio of the decay curve.

Effect of Noise Addition to Decay Curves. As discussed in the previous subsection, the $P(R)$ recovered from fitting a noise-added decay curve deviates from that inputted. The deviation has been believed to arise from statistical noise added to the decay data. This point has been proven by fitting the same decay curve but before statistical noise was added to it. $P(R)$ recovered from fitting such a noise-free decay curve is also shown in Figure 4. The agreement between $P(R)$ recovered from the noise-free decay curve and the inputted $P(R)$ improved considerably when compared to the former case.

Effect of Lamp Tailing. Despite the improved agreement between $P(R)$ recovered from fitting the noise-free decay curve and the inputted $P(R)$, there is still some oscillation in the recovered $P(R)$ in the R range of 20–30 Å. The long but low-intensity tail of the laser profile shown in Figure 1, which was used to construct the decay data, was thought to be partially responsible for the oscillation. When such a long tail exists, the tail intensity convolutes with the decay law given by eq 6 via eq 11, which introduces oscillations into the decay curve and affects the correct recovery of $P(R)$. The effect of laser tailing is examined by generating a decay curve from using exactly the same parameters as used before but with the exception that the laser tail was cut; i.e., any data points of the laser profile with counts lower than one-thousandth of that in the peak channel are assumed to have zero intensity. Fitting of such a decay curve without added noise yielded $P(R)$, which is again included in Figure 4. Compared to the $P(R)$ recovered in the former two cases, the current $P(R)$ resembles the inputted $P(R)$ the best.

Effect of Signal to Noise Ratio. End-to-end distance distribution functions $P(R)$ recovered under different conditions shown in Figure 4 illustrate how the addition of noise to a decay curve can affect the shape of $P(R)$ recovered. This subsection discusses how the signal to noise ratio of a decay curve can affect the recovered $P(R)$. Decay curves with different peak channel counts can be easily constructed by scaling the inputted $L(t)$; i.e., increasing the intensity of $L(t)$ increases the intensity of generated decay curves and vice verse. When those decay curves with different peak channel counts are fitted, different $P(R)$ recovered are compared in Figure 5. As is obvious, when the peak channel count increases, the recovered $P(R)$ resembles the inputted $P(R)$ better.

Effect of Different R_0/R_n Ratios. So far, discussion has been concerned with the recovery of $P(R)$ from decay curves which are generated using $R_0 = R_n = 40$ Å or $R_0/R_n = 1$. Experimentally, it is very unlikely that one only deals with samples which possess $R_0/R_n = 1$. To see the effect of varying R_0/R_n on the correct recovery of $P(R)$, fluorescence decay curves were generated using R_0 values from 30 to 50 Å with $\Delta R_0 = 5$ Å as the increment. But the R_n value was fixed at 40 Å. To further simplify complications from effects of noise addition and lamp tailing, the decay curves are generated with the lamp profile of Figure 1 but with the tail cut and no noises are added to decay data.

The $P(R)$ functions recovered from fitting decay data generated using R_0 equal to 35, 40, and 45 Å are shown in Figure 6. When $R_0 = 30$ Å, the recovered $P(R)$ bears little resemblance to the inputted $P(R)$. When $R_0 = 50$ Å, the agreement between recovered $P(R)$ and inputted $P(R)$ is excellent between $R \approx 30$ Å and $R \approx 80$ Å.

Each of the recovered $P(R)$ curves shown in Figure 6 bears close resemblance to the inputted $P(R)$ in the effective range of the individual spectroscopic ruler, i.e., between $(1/2)R_0$ and $(3/2)R_0$. Out of three R_0 and R_n combinations, $R_0 = R_n$ is the best one. In this case, the recovered $P(R)$ is the smoothest. The conclusion of this study is that for the determination of $P(R)$ of polymer chains, the R_0 value should be as close to R_n as possible. If it is not possible, R_0 should be larger than R_n to provide any useful information.

V. Conclusions

A new method is proposed for treating the data of fluorescence quenching of a donor group, attached to one site of a macromolecule, by an acceptor group, attached to the other site, via dipole–dipole energy transfer. Using the suggested data treatment technique, the distance distribution function $P(R)$ between the donor- and ac-

ceptor-labeled sites can be recovered. The technique has been shown to be effective by treating computer-simulated decay curves for two examined situations: (1) the conformational change of proteins upon the addition of denaturant and (2) the determination of end-to-end distance distribution functions of flexible polymer chains. It is expected that the technique can be used in many other systems. Two examples would be its use in the elucidation of conformations of surface-adsorbed polymers[23] and in that of polymers used as drag reduction reagents in a turbulent flowing environment.[24]

Acknowledgment. The author thanks NSERC Canada for financial support of this work and Dr. Alex Siemiarczuk of Photon Technology International for modifying their version of the ESM program to suit the needs of the current research in the early stage of the project. The critical review of the manuscript by Dr. Rodney Roche of the Biological Sciences Department is sincerely acknowledged.

References and Notes

(1) See, for example: Förster, Th. In *Modern Quantum Chemistry*; Sinanöglu, O., Ed.; Academic Press: New York, 1965.
(2) Latt, S. A.; Cheung, H. T.; Blout, E. R. *J. Am. Chem. Soc.* **1965**, *87*, 995.
(3) Styer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 720.
(4) See, for example: Steinberg, I. Z. *Annu. Rev. Biochem.* **1971**, *40*, 83.
(5) Liu, G.; Guillet, J. E.; Al-Takrity, E. T. B.; Jenkins, A. D.; Walton, D. R. M. *Macromolecules* **1990**, *23*, 1393.
(6) Katchalski-Katzir, E.; Haas, E.; Steinberg, I. Z. *Ann. N.Y. Acad. Sci.* **1981**, *366*, 44.
(7) Lakowicz, J. R.; Gryczynski, I.; Wiczk, W.; Laczko, G.; Prendergast, F. C.; Johnson, M. L. *Biophys. Chem.* **1990**, *36*, 99.
(8) See, for example: Dale, R. E.; Eisinger, J.; Blumberg, W. E. *Biophys. J.* **1979**, *26*, 161.
(9) See, for example: Birnbaum, E. R.; Abbott, F.; Gomez, J. E.; Darnall, D. W. *Arch. Biochem. Biophys.* **1977**, *179*, 469.
(10) Berner, V. G.; Darnall, D. W.; Birnbaum, E. R. *Biochem. Biophys. Res. Commun.* **1975**, *66*, 763.
(11) O'Connor, D. V.; Phillips, D. *Time-Correlated Single Photon Counting*; Academic Press: London, 1984.
(12) See, for example: Siemiarczuk, A.; Wagner, B. D.; Ware, W. R. *J. Phys. Chem.* **1990**, *94*, 1661.
(13) James, D. R.; Liu, Y.-S.; Siemiarczuk, A.; Wagner, B. D.; Ware, W. R. *Proc. SPIE Int. Soc. Opt. Eng.* **1987**, *743*, 117.
(14) Wagner, B. D.; Ware, W. R. *J. Phys. Chem.* **1990**, *94*, 3489.
(15) Brochon, J.-C.; Livesey, A. K.; Pouget, J.; Valeur, B. *Chem. Phys. Lett.* **1990**, *174*, 517.
(16) Livesey, A. K.; Brochon, J. C. *Biophys. J.* **1987**, *52*, 693.
(17) Cuthbert, T. R., Jr. *Optimization Using Personal Computers*; Wiley: New York, 1987.
(18) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes*; Cambridge University Press: Cambridge, 1989.
(19) Smith, C. K.; Liu, G. J., manuscript in preparation.
(20) See, for example: Karplus, M.; McCammon, J. A. *Sci. Am.* **1986**, *254*, 42 (April).
(21) Berlman, I. *Energy Transfer Parameters of Aromatic Compounds*; Academic Press: New York, 1973.
(22) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, 1953.
(23) Milner, S. T. *Science* **1991**, *251*, 905.
(24) See, for example: Kulicke, W.-M.; Kötter, M.; Grager, H. *Adv. Polym. Sci.* **1989**, *89*, 1.