Donor−Acceptor Distance Distributions in a Double-Labeled Fluorescent Oligonucleotide Both as a Single Strand and in Duplexes†

Kay M. Parkhurst and Lawrence J. Parkhurst*

Department of Chemistry, University of Nebraska−Lincoln, Lincoln, Nebraska 68588-0304

Received July 19, 1994; Revised Manuscript Received October 14, 1994®

ABSTRACT: A 16-mer deoxyribonucleotide was labeled at the 5′-end with x-rhodamine and at the 3′-end with fluorescein. Förster resonance energy transfer was used to determine the distribution, $P(R)$, of donor−acceptor distances for the oligomer in three duplex structures (hybridized to complementary strands having 10, 16, and 24 bases) and as a single strand; measurements were made in 0.18 M NaCl and 1 M KCl solutions. These distributions were derived from lifetime measurements made in the frequency domain using a multiharmonic frequency phase fluorometer. The fluorescein fluorescence decay for each duplex structure was fit very well with $P(R)$ modeled as a shifted Gaussian. On the basis of the mean donor−acceptor distance, these structures for the 16-mer and the 24-mer were consistent with that of B-DNA. For the 16-mer duplex in 0.18 M NaCl, the distribution was centered at 68.4 Å with $\sigma = 6.4$ Å. For the 10-mer duplex, the mean donor−acceptor distance, 61.8 Å, is much larger than that for a structure with the fluorophores extended perpendicular to the helix axis. For the single-strand data, various high-quality fits yielded physically unreasonable distributions or could not accurately account for the acceptor response. Considered analyses suggested that the single-strand distribution was best represented by a shifted Gaussian, with $\bar{R} = 51.5$ Å and $\sigma = 10.0$ Å in 0.18 M NaCl. In all cases, σ increased and $\bar{R}$ decreased in 1 M KCl relative to their values in 0.18 M NaCl, consistent with the increased flexibility of the polymer.

Over the past several years, we have been carrying out solution studies with a fluorescently double-labeled oligodeoxyribonucleotide (Parkhurst & Parkhurst, 1992, 1993, 1994). In the preceding paper in this issue (Parkhurst & Parkhurst, 1995), we have reported the results of our work that involved using the labeled oligomer to signal hybridization in solution and to determine the kinetics of hybridization, exploiting the pronounced differences in the steady-state fluorescence emission for the oligomer as a single strand and in a duplex structure. The oligomer has 16 bases, with x-rhodamine covalently bound to the 5′-end and fluorescein bound to the 3′-end: x-rhodamine•5′-GTAAAACTGACGGCCAG-3′•fluorescein (Figure 1). The fluorescein and x-rhodamine act as a donor/acceptor pair, respectively, for Förster resonance energy transfer, a coupled dipole interaction that results in the direct transfer of excitation energy from the donor to the acceptor chromophore. Because the efficiency of transfer is dependent on the inverse sixth power of the distance between the two fluorophores, the integrated fluorescence emission intensities sensitively reflect changes in that distance. Resonance energy transfer has been used extensively to obtain the mean distance between two fluorophores that are bound to a relatively rigid structure (Latt et al., 1965; Stryer & Haugland, 1967; Steinberg, 1971; Yang & Söll, 1974; Stryer, 1978; Cheung, 1991a). More recently, energy transfer has been used to obtain the distribution of distances, $P(R)$, between two fluorophores bound to a nonrigid structure (Grinvald et al., 1972; Lakowicz et al., 1987, 1988; Cheung, 1991b; Haran et al., 1992, Eis & Lakowicz, 1993).

We have extended our studies with the labeled oligomer, analyzing fluorescence lifetime measurements made in the frequency domain to obtain the distribution of donor−acceptor distances for the oligomer as a single strand and in three duplex structures, at low and high salt concentrations. We explored the effect on the fit of diverse distributions, particularly for the data obtained from the single-stranded oligomer, and arrived at a model-free method of fitting. The single-strand data were well fit by significantly different models; the plausibility of the associated distribution and the predicted response of the acceptor, x-rhodamine, were used to determine the $P(R)$ that best represented the data.

EXPERIMENTAL PROCEDURES

Tris-HCl and quinine, as the hemisulfate salt, were purchased from Sigma (St. Louis, MO). The 16-mer deoxyribonucleotide with covalently bound 5′-x-rhodamine and 3′-fluorescein (^{R}*oligo*^{F})[1] was synthesized and HPLC-purified by Research Genetics, Inc. (Huntsville, AL), as was the single-labeled oligomer with only 3′-fluorescein (oligo*^{F}). The 16-mer complement (oligo$_{16c}$), the 10-mer complement to the middle 10 bases (oligo$_{10c}$), and a 24-mer complement (oligo$_{24c}$, the 16-mer complement with an additional 5′-TCAC

† This work was supported by National Institutes of Health Research Grant DK 36288 and by grants from the Center for Biotechnology, University of Nebraska.

* Author to whom correspondence should be addressed. Telephone: (402) 472-3316. e-mail: lparkhurst@unl.edu. FAX: (402) 472-9402.

® Abstract published in *Advance ACS Abstracts,* December 1, 1994.

[1] Abbreviations: all oligonucleotides are oligodeoxyribonucleotides; ^{R}*oligo*^{F}, the 16-base oligomer 5′-GTAAAACTGACGGCCAG-3′ with x-rhodamine (structure shown in Figure 1) covalently bound at the 5′-end and fluorescein at the 3′-end; oligo*^{F}, the 16-base oligomer with only 3′-fluorescein; oligo$_{10c}$, the 10-mer complementary to the middle 10 bases of the labeled oligomer; oligo$_{16c}$, the 16-mer complement to the labeled oligomer; oligo$_{24c}$, a 24-mer identical to oligo$_{16c}$ but with additional 5′-TCAC and 3′-TACA.

0006-2960/95/0434-293$09.00/0

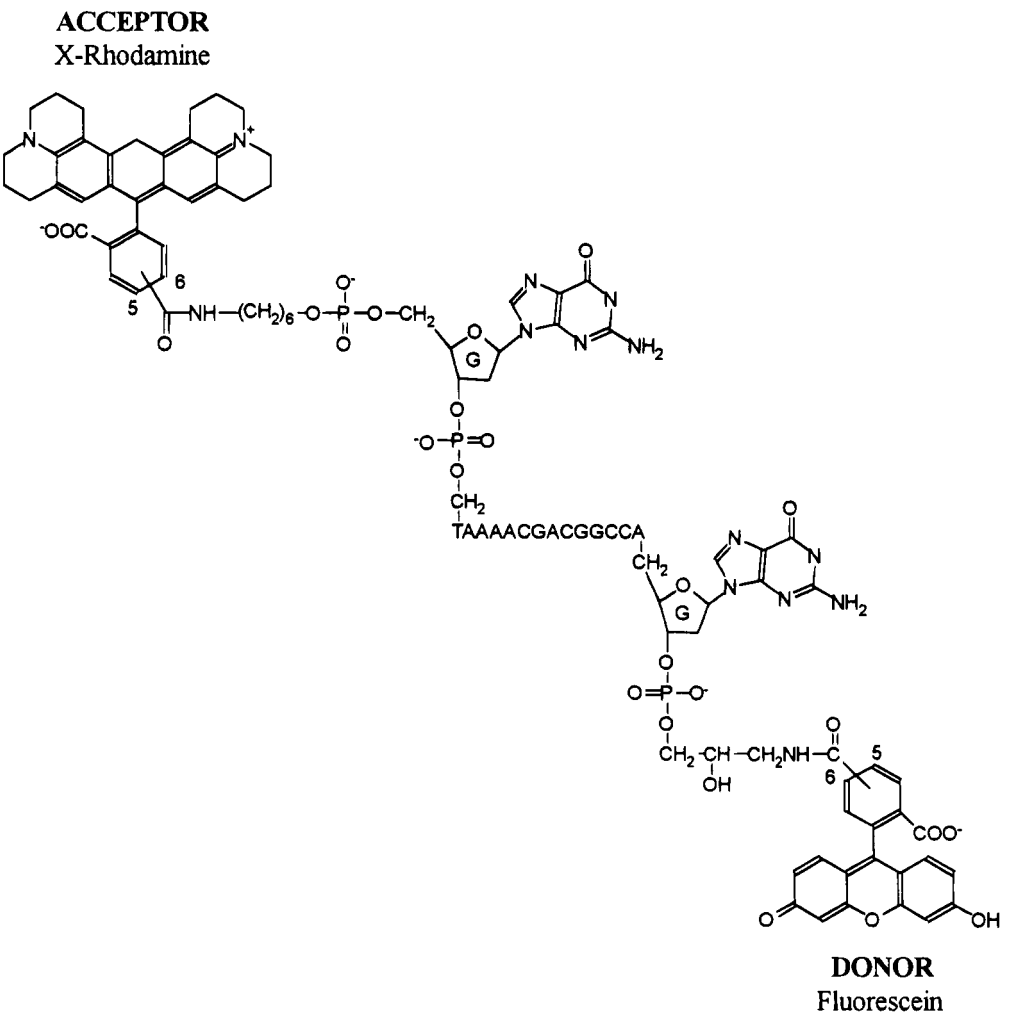

FIGURE 1: Fluorescently double-labeled 16-mer oligodeoxyribonucleotide. x-rhodamine is bound at the 5′-end with a 6-carbon linker arm, and fluorescein is bound at the 3′-end with a 3-carbon linker.

and 3′-TACA) were prepared in our departmental oligonucleotide synthesis facility and shown by gel electrophoresis to be >95% pure. The buffer was 0.01 M Tris (pH 8) and 0.001 M EDTA, with either 0.18 M NaCl or 1 M KCl, as noted in the text. The concentrations of oligo*^{F} and ^{R}*oligo*^{F} were 140 and 250 nM, respectively, either alone or with a 40% excess of complement. Duplex structures were prepared at least 45 min before the measurements were made; steady-state work had shown that, with these solution conditions, hybridization would be essentially complete in that time. All measurements were made at ambient temperature (20−22 °C).

Instrumentation, Theory, and Data Analysis. In the absence of transfer, the fluorescence decay of the donor in the time domain is described by

$$I_d(t) = I_d^0 \sum \alpha_{di} \exp[-(k_F + k_I)_{di} t] \quad (1)$$

where I_d^0 is the initial fluorescence intensity, and the rate constants k_F and k_I pertain to fluorescence decay and nonradiative decay (including intersystem crossing), respectively. The inverse of the donor lifetime, τ_d^{-1}, is the sum of these intrinsic decay processes, $k_F + k_I$. Donor decay can be mono- or multiexponential, the latter a consequence of variations in k_I.

When an acceptor is present at a fixed distance R from the donor, a second term is needed to account for decay due to transfer:

$$I_{da}(t) = I_d^0 \sum \alpha_{di} \exp[-[(k_F + k_I)_{di} + k_t]t] \quad (2)$$

where k_t is the rate constant for transfer ($=1/\tau_{D^*}(R_o/R)^6$), and the parameters α_{di} and $(k_F + k_I)_{di}$ are assumed to be unchanged by the presence of acceptor. R_o is the Förster distance, the distance at which the efficiency of transfer is 0.5, that is, $k_t = (k_F + k_I) = \tau_{D^*}^{-1}$. τ_{D^*} in the additional term $(1/\tau_{D^*}(R_o/R)^6$, eq 3) therefore is the donor lifetime uniquely associated with a particular value of R_o and remains constant so long as acceptor absorption and donor emission spectra remain unchanged. When the donor and acceptor are bound to a nonrigid moiety, their distance apart varies over some probability distribution $P(R)$ (Cantor & Pechukas, 1971), and eq 2 becomes

$$I_{da}(t) = I_d^0 \int_0^\infty P(R) \sum \alpha_{di} \exp\left[-\left(\frac{t}{\tau_{di}} + \frac{t}{\tau_{D^*}}\left(\frac{R_o}{R}\right)^6\right)\right] dR \quad (3)$$

The fluorescence decay of the donor can also be obtained in the frequency domain, by measuring the phase and amplitude of the response to modulated excitation. The sine and cosine Fourier transforms (N and D, respectively) of eq 3 yield the equivalent frequency response functions:

$$N(\omega) = \left(\frac{1}{J}\right)\int_0^\infty \sum_i \frac{P(R)\alpha_{di}\omega\tau_{dai}^2}{1+\omega^2\tau_{dai}^2}\,dR \quad (4)$$

$$D(\omega) = \left(\frac{1}{J}\right)\int_0^\infty \sum_i \frac{P(R)\alpha_{di}\tau_{dai}}{1+\omega^2\tau_{dai}^2}\,dR \quad (5)$$

where ω is the frequency (in radians second^{-1}), $1/\tau_{dai} = 1/\tau_{di} + 1/\tau_{D^*}(R_o/R)^6$, and the normalization term $J = \sum\alpha_{dai}\tau_{dai}$.

Fluorescence lifetime measurements were made using an SLM-Aminco Model 4850 multiharmonic frequency phase fluorometer. The fluorescein was excited with the 488 nm line from an Ar^+ laser (Lexel, Model 75), which was modulated simultaneously at 28 frequencies from 5 to 140 MHz with a base frequency of 5 MHz; a 520 nm interference filter (Oriel Corp., Stratford, CT) was positioned between the sample and the photomultiplier. The SLM software was used to collect and least-squares fit the emission data to obtain the phase shift and modulation that best described the fluorescein emission response at each frequency and, subsequently, to determine the optimal values for τ. These relationships and methods of analysis are discussed elsewhere (Lakowicz et al., 1984). In all cases, the data were very well fit by a biexponential decay model (χ^2 values of 0.2−1.6). This means only that the data were very well represented by a biexponential model that approximates a distribution of lifetimes and does not necessarily imply decay from two components. The weighting factor a in the SLM software is the fractional contribution of each component to the total integrated intensity; these a's were converted to preexponential α's by solving the expression $a_i = (\alpha_i\tau_i/\sum\alpha_i\tau_i)$ for each α_i, with normalization $\sum\alpha_i = 1$. These measurements and analyses were made on donor only (oligo*^{F}) and donor/acceptor (^{R}*oligo*^{F}) molecules, each as a single strand and in three duplex structures (hybridized to $oligo_{16c}$, $oligo_{10c}$, and $oligo_{24c}$). Data for each of these four donor−donor/acceptor pairs were collected in buffered solutions containing either 0.18 M NaCl or 1 M KCl. For each sample, 5 replicate measurements were made and averaged to give a single data set; 6−10 such data sets were collected during each period of data acquisition (on any given day), and the optimal data set, based on the χ^2 values, was chosen. Phase ($\phi(\omega)$) and modulation ($m(\omega)$) vectors for each case were composite files averaged over 2−5 such optimal sets. The error vectors for phase and modulation were the standard deviations at each frequency for the initial five replicate measurements; a composite error vector was then made for both phase and modulation.

For each pair, the data were analyzed for $P(R)$ as follows: The observed donor/acceptor response was obtained by constructing, at each frequency, the sine and cosine Fourier transforms from the phase shift and modulation vectors, using the following:

$$D(\omega)_{obs} = \frac{m(\omega)}{[1+\tan^2\phi(\omega)]^{1/2}} \quad (6)$$

and

$$N(\omega)_{obs} = [m(\omega)^2 - D(\omega)^2]^{1/2} \quad (7)$$

(Lakowicz et al., 1984). These observed values for $N(\omega)$ and $D(\omega)$ were used on the left-hand sides of eqs 4 and 5. Weighting vectors for the sine and cosine transforms were generated from the composite raw phase and modulation vectors and their associated error vectors using standard error propagation. On the right-hand sides of eqs 4 and 5, the α's pertain to the donor-only case and $1/\tau_{dai} = 1/\tau_{di} + 1/\tau_{D^*} \times (R_o/R)^6$. Various expressions were used for $P(R)$, and a Simplex algorithm (Noggle, 1985) was used to minimize the weighted sum of squared residuals in eqs 4 and 5 as the parameters within the distribution function were varied. Within the Simplex, the numerical integration over $P(R)$ was carried out by a variable-step Simpson's rule integration over the interval 3.5−120 Å. The minimum distance between the two fluorophores that is physically reasonable is the sum of their van der Waals radii, ~3.5 Å, and the maximum distance, the length of the fully extended oligomer plus the two linker arms, we calculated to be 120 Å (see Figure 1). The intervals were subdivided until the absolute value of $(T - M)/S$ was $< 1E-5$, where T, M, and S are trapezoidal, midpoint, and Simpson's rule integrations, respectively, with $S = (T + 2M)/3$. The integrals in eqs 3−5 were normalized to the distribution function, $P(R)$, by dividing by $\int P(R)\,dR$ over the same range of integration.

The Förster critical distance, R_o, was calculated using

$$R_o\,(\text{Å}) = 9.79 \times 10^3\,(\kappa^2 n^{-4}\phi_d\int_{da})^{1/6} \quad (8)$$

(Lakowicz, 1983), where κ is the angular portion of the dipole−dipole interaction for the transition dipoles of the donor and acceptor, n is the refractive index of the medium, and ϕ_d is the quantum yield of the donor (Förster, 1948). The overlap integral of the donor and acceptor is defined as

$$\int_{da} = \int_0^\infty F_d(\lambda)\,\epsilon_a(\lambda)\,\lambda^4\,d\lambda \quad (9)$$

where $F_d(\lambda)$ is the normalized fluorescence emission of the donor at each wavelength, $\epsilon_a(\lambda)$ is the extinction coefficient (in M^{-1} cm^{-1}) of the acceptor at each wavelength, and λ is wavelength in centimeters. The value of n was fixed at 1.343 (Washburn, 1930) and ϕ_d at 0.93 (Martin & Lindqvist, 1975). Initially, κ^2 was assigned the value of $^2/_3$, which assumes complete rotational averaging. This assumption is in accord with the work of Hochstrasser et al. (1992), who concluded from fluorescence anisotropy measurements that nearly complete rotational averaging was occurring for fluorophores with linker arms similar to those on our dyes. The response function of the photomultiplier was calibrated using quinine sulfate (Melhuish, 1972), and the fluorescein emission spectrum was corrected accordingly.

P(R). Our *a priori* model for the duplex structures was that of a rigid central region with random orientation of the dyes on either end. In a simple Monte-Carlo simulation of the end-to-end distribution based on this model, the dyes were allowed to occupy any point within the sphere defined by the maximum length of their tethers (except those points within the helix cylinder). Using 10 000 random positions for the donor and acceptor, we obtained a distribution that very closely approximated a shifted Gaussian. Further, the vast majority of the positions occupied by the dyes were within the thin surface shell of the sphere, so that σ and $\bar{R}$ for the distribution largely reflected the arcs swept out by the angular motion of the dyes. The lifetime data for the three duplex structures, in 0.18 M NaCl and in 1 M KCl,

were subsequently fit to a shifted Gaussian distribution:

$$P(R) = A \exp[-(R - \bar{R})^2/2\sigma^2] \quad (10)$$

where A is a normalization constant.

Our initial model for the single strand was that of a random coil with freely jointed segments, in which case, for $R \ll R_{maximum}$ (~120 Å), the distribution would approximate the form (Tanford, 1961)

$$P(R) = AR^2 \exp(-BR^2) \quad (11)$$

which is a skewed Gaussian with $B = (^3/_2)(1/\overline{R^2})$, $A = 4\pi \times (B/\pi)^{3/2}$, and only one fitted parameter, $\overline{R^2}$. The quantity $\bar{R} = \int R\, P(R)\, dR$, and $\sigma^2 = \overline{R^2} - (\bar{R})^2$. The very poor fit to this model led to extensive analyses of the single-strand data. We postulated that the average value of κ^2 might not be $^2/_3$, and κ^2 became a fitted parameter. This resulted in no improvement in the fit. In fact, because κ, σ, and $\bar{R}$ are correlated, for $0 < \kappa^2 < 4$, the values for σ and $\bar{R}$ adjust in a three-parameter fit to give essentially the same quality of fit for all values of κ^2. For $0.45 < \kappa^2 < 0.90$, the values of σ and $\bar{R}$ changed by <7%. We then assumed that κ^2 was not dynamically averaged in the single-stranded oligomer, and we introduced a second integral into eqs 4 and 5, with the κ^2 distribution modeled in various ways. The details of the κ^2 fitting have been reported elsewhere (Parkhurst & Parkhurst, 1994). We were unable to significantly improve the fit in these analyses and concluded that the distributions of R and κ^2 might not be separable.

Diverse distributions were used in fitting the single-strand data. These included a shifted Gaussian, a sum of shifted Gaussians, and a shifted Gaussian with a preexponential variable term, C, of the form

$$P(R) = A(R - C)^n \exp[-(R - \bar{R})^2/2\sigma^2] \quad (12)$$

with $n = 1$ and 2. Another model assumed that there might be both random coil, described by a random walk model, and structure, characterized by a shifted Gaussian:

$$P(R) = C_1A_1 \exp[-(R - \bar{R})^2/2\sigma^2] + (1 - C_1)A_2R^2 \exp(-BR^2) \quad (13)$$

For each of these distributions, κ^2 was both fixed at $^2/_3$ and allowed to vary. For $P(R)$ as a shifted Gaussian, a sum of $P(R)$ and a skewed Gaussian was also examined:

$$[A + (1 - A)(1 - e^{-\theta R})^n]P(R) \quad (14)$$

with additional fitted parameters A, θ, and n. Equations 4 and 5 were also modified to include a constant contribution from free fluorescein or from a contaminant of oligo*^{F} in the ^{R}*oligo*^{F} solution, and data were fit using various of these forms for $P(R)$.

The general problem of fitting for $P(R)$ is the form of eq 15 in the frequency domain or eq 16 in the time domain:

$$N(\omega) = \int_0^\infty K(\omega,R)\, P(R)\, dR \quad (15)$$

$$I(t) = \int_0^\infty K'(t,R)\, P(R)\, dR \quad (16)$$

These are examples of Fredholm integral equations of the first kind, with known kernels K and K', and can be solved for $P(R)$ by expanding $P(R)$ in an orthonormal basis set. A reasonable basis set for this problem would be the Hermite-polynomials multiplied by the corresponding Gaussian weighting factor, since they are orthonormal over the interval $[0,\infty]$ and contain a Gaussian function that alone should approximate many random end-to-end distributions reasonably well. Through order n, the function is

$$P_n(R) = \sum_{v=0}^{n} C_vN_vH_v[\alpha^{1/2}(R - \bar{R})] \exp[-\alpha(R - \bar{R})^2/2] \quad (17)$$

where the C_v's are fitting coefficients, N_v is a normalization constant $(=(2^v v!)^{-1/2}(\alpha/\pi)^{1/4})$, and H_v is the Hermite polynomial of order v [$H_0 = 1$, $H_1 = 2\alpha^{1/2}(R - \bar{r})$, etc.; recurrence relation: $H_{v+2}(x) = 2xH_{v+1}(x) - 2(v + 1)H_v(x)$]. For eq 17, the fitting parameters are thus α, R, and the C_v's.

Most of the distributions described here gave fits with unacceptably large sums of squared residuals when the phase and modulation vectors for frequencies from 5 to 140 MHz were used. Those distributions that gave good fits included regions of probability that were physically unreasonable. These analyses all relied upon the donor decay changing in the presence of acceptor due *only* to energy transfer, that is, α_{di} and the τ_{di} part of τ_{dai} in eqs 4 and 5 must be the same as α_{di} and τ_{di} for oligo*^{F}. In subsequent analyses of the lifetime data, this constraint was removed. The α and τ_{di} terms (within τ_{dai}) in eqs 4 and 5 that pertain to the donor only became fitted parameters, along with σ and $\bar{R}$ in the shifted Gaussian distribution (eq 14). Because the α's were normalized to $\sum\alpha_i = 1$, α_2 was defined as $(1 - \alpha_1)$, and the five fitted parameters were σ, R, α_1, and τ_{d1} and τ_{d2} (within the τ_{dai} terms). The data for the three duplex structures and the single strand, in 0.18 M NaCl and in 1 M KCl, were all fit in this way. The quality of the fits for duplex and single-strand structures was also studied as the phase and modulation data sets were progressively truncated from the high-frequency end. The steady-state acceptor emission was also calculated by integrating the efficiency of transfer over the distribution, $P(R)$, to provide a constraint on the donor parameters.

RESULTS

The critical distance R_o for the donor/acceptor pair was calculated using eq 8 to be 59.6 Å. This distance is for fluorescein/x-rhodamine, and for fluorescein, $\tau = 3.96$ ns. If k_I changes, R_o also changes, but k_t must remain invariant to such changes. We therefore write eq 3 with the τ_D fixed at 3.96 ns and denote *that* τ_D as τ_{D*}, linked to $R_o = 59.6$ Å. The distance between the 3′- and 5′-ribosyl oxygens in a 16-mer B-DNA duplex structure is 54.7 Å based on molecular modeling, making the fluorescein/x-rhodamine pair well suited to the study of this structure.

The phase shift and modulation responses for the three duplex structures, in 0.18 M NaCl and in 1 M KCl, were fit very well to a biexponential decay model, and the α's and τ's are given in Table 1. Excellent fits to eqs 4 and 5 were obtained with these data, with $P(R)$ modeled as a shifted Gaussian. The optimal values obtained for σ and $\bar{R}$ are shown in Table 2. In 0.18 M NaCl, σ for each of the three duplexes is ~6.5 Å. The mean distance between the two fluorophores was nearly the same for the 16-mer and 24-

Table 1: Optimal Values for the α's and τ's That Characterized the Biexponential Decay of the Fluorescein Emission in the Oligo*F (D) and R*Oligo*F (D/A) Molecules as a Single Strand and in the Duplex Structures (in 0.18 M NaCl and 1 M KCl)

	0.18 M NaCl				1 M KCl			
	α_1	τ_1 (ns)	α_2	τ_2 (ns)	α_1	τ_1 (ns)	α_2	τ_2 (ns)
single strand								
D	0.75	3.94	0.25	0.33	0.77	4.11	0.23	1.09
D/A	0.19	2.77	0.81	0.58	0.15	3.285	0.85	0.47
10-mer duplex								
D	0.81	3.97	0.19	0.67	0.79	3.92	0.21	0.79
D/A	0.62	2.42	0.38	0.84	0.53	2.31	0.47	0.73
16-mer duplex								
D	0.78	3.85	0.22	0.53	0.68	3.76	0.32	0.52
D/A	0.71	2.75	0.29	0.63	0.62	2.54	0.38	0.46
24-mer duplex								
D	0.87	4.09	0.13	0.55	0.84	4.19	0.16	0.98
D/A	0.82	2.99	0.18	0.80	0.76	3.04	0.24	0.90

Table 2: Values for σ and $\bar{R}$ That Best Describe the 3′-Donor−5′-Acceptor Distance Distribution for R*Oligo*F as a Single Strand and in Three Duplex Structures (in 0.18 M NaCl and 1 M KCl)

	0.18 M NaCl		1 M KCl	
R*oligo*F	σ (Å)	$\bar{R}$ (Å)	σ (Å)	$\bar{R}$ (Å)
single strand	10.01	51.5	15.3	45.3
10-mer duplex	6.8	61.8	8.4	58.7
16-mer duplex	6.4	68.5	8.8	66.5
24-mer duplex	6.5	71.4	10.5	70.9

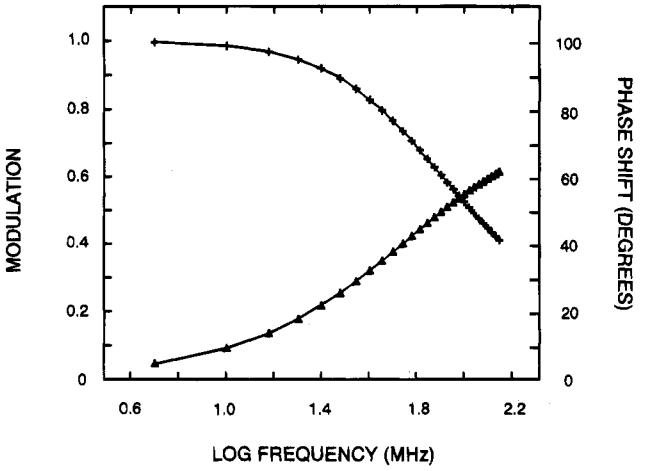

FIGURE 2: Theoretical values (lines) and observed phase shift (+) and modulation (▲) as a function of frequency for the duplex composed of R*oligo*F and oligo$_{16c}$ (in 0.18 M NaCl), with $P(R)$ modeled as a shifted Gaussian. The theoretical values were calculated from the sine and cosine Fourier transforms obtained at each frequency using the optimal values for σ and $\bar{R}$.

mer duplexes, 68.5 and 71.4 Å, respectively, but was significantly less, 61.8 Å, for the 10-mer duplex. In 1 M KCl, the value of σ increased by 2 Å for the 10- and 16-mer duplexes and by 4 Å for the 24-mer duplex; the mean distance between the fluorophores decreased slightly (0.5−2 Å) for each duplex. The sine and cosine Fourier transforms of the donor/acceptor response, to which the data were fit, were recovered in all cases to within experimental error. The theoretical values for the phase shift and modulation at each frequency were determined from the sine and cosine transforms using eqs 6 and 7 and are shown in Figure 2, together with the observed values, for the oligo$_{16c}$ duplex structure in 0.18 M NaCl. This plot is typical of that for

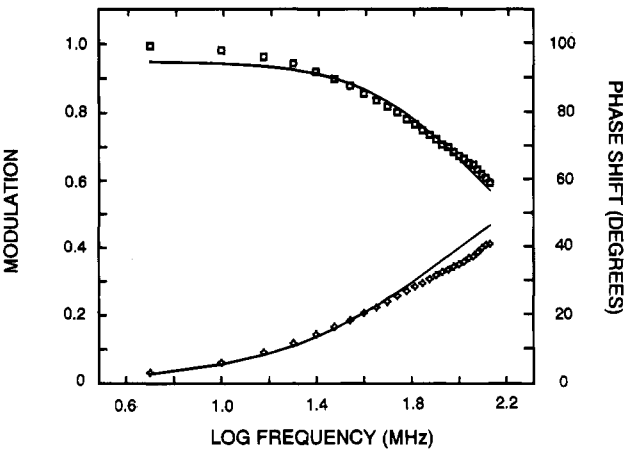

FIGURE 3: Phase shift and modulation plots, showing the lack of fit of the data from the single-stranded oligomer (in 0.18 M NaCl) to a shifted Gaussian distribution, using the donor properties of oligo*F. Observed phase shift (□) and modulation (◇) are shown with the theoretical values (lines) obtained at the least-squares minimum for σ (10.01 Å) and $\bar{R}$ (51.5 Å). The average residual for this fit is ~5× larger than the experimental error.

each of the duplex cases. Notably, essentially the same values for σ and $\bar{R}$ were obtained using data sets truncated (from the high-frequency end) to include only seven frequencies from 5 to 35 MHz, with little change in the quality of the fit (σ_{fit}).

The distribution for the 16-mer duplex in 0.18 M NaCl is shown in Figure 6. For the Monte-Carlo simulations, the maximum lengths of the tethers from the ribosyl oxygens to the centers of the dyes were 16.3 Å for the fluorescein and 20.3 Å for x-rhodamine. The helix was considered to be rigid. Letting all points within a sphere (with radius equal to the maximum length of the tether) centered on the ribosyl oxygens, but excluding the DNA helix, be available for the dyes, we obtained σ = 9.7 Å and $\bar{R}$ = 60.1 Å. With one adjustable parameter, the excluded distance from the cylinder, which accounts for van der Waals contacts between the tether and the helix as well as the excluded volume of the tether, we obtained σ = 8.4 Å and $\bar{R}$ = 66.9 Å for an excluded distance of 7.25 Å. These σ and $\bar{R}$ values compare very well with the experimental values for 1 M KCl (σ = 8.8 Å and $\bar{R}$ = 66.5 Å), where charge interactions between the dyes and the helix are greatly reduced.

Obtaining an acceptable fit to the single-strand data was considerably more complex. The best fit to the random walk model (eq 11) gave an rms residual ~20× greater than the experimental error, with an rms end-to-end distance, $(\bar{R^2})^{1/2}$, of 54.6 Å. For a sufficiently long random coil, this should be the proper $P(R)$, but as noted by Yevich and Olson (1979), for short single-strand oligonucleotides, $P(R)$ distributions are not expressible in simple mathematical forms. Figure 3 shows the best fit to the experimental phase shift and modulation data for the single strand in 0.18 M NaCl, with $P(R)$ modeled as a shifted Gaussian. The rms weighted residual for this fit was 5 times the experimental error. The optimal values for σ and $\bar{R}$ were 10.0 and 51.5 Å, respectively, in 0.18 M NaCl. This distribution is shown in Figure 6. As the phase and modulation vectors were truncated, the fit to the symmetric Gaussian improved dramatically and was within experimental error for frequencies ≤40 MHz, while the optimal parameter values changed by <5%. When a shifted Gaussian was used with eq 14,

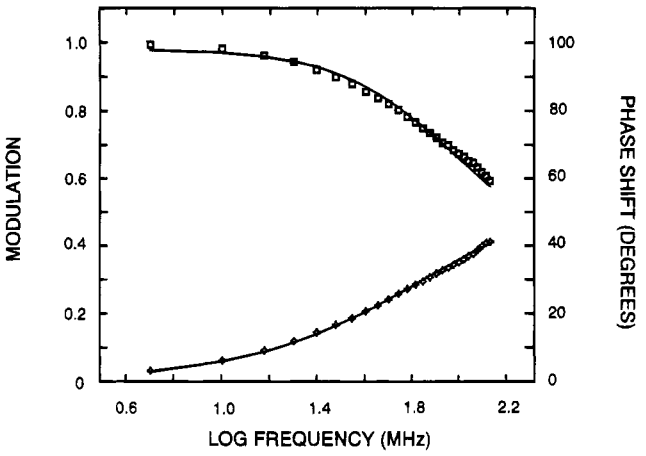

FIGURE 4: Theoretical (lines) and observed phase shift (□) and modulation (◇) vs frequency for the single-strand oligomer (in 0.18 M NaCl), obtained using Hermite polynomials (with $n = 3$) to model $P(R)$.

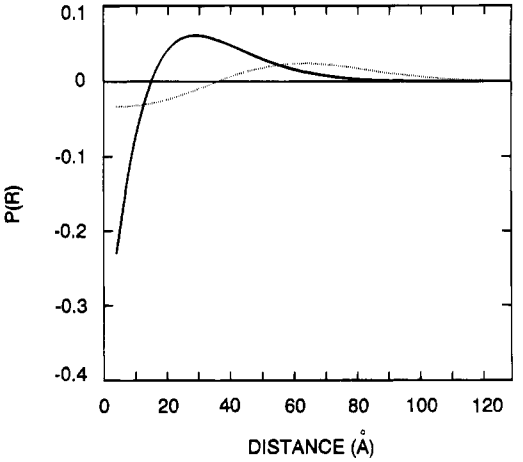

FIGURE 5: 3′−5′ distance distributions obtained for the single-stranded oligomer for $P(R)$ obtained from Hermite polynomials. The distribution obtained fitting the complete data set is shown as a solid line; numerous of the $P(R)$ forms yielded distributions with this general shape. As the data set was truncated from the high-frequency end, the shape of the distribution approached that of a shifted Gaussian centered at ~60 Å; the distribution shown (•••) was obtained by fitting the data from 5 to 40 MHz. Both distributions are normalized over the interval 3.5−120 Å.

the coefficient $(1 - A)$ weighting the additional term approached zero, with essentially no change in the fit. Modification of various of the $P(R)$ forms to include a constant contribution from the fluorescein-only labeled oligomer also resulted in the coefficient for the additional term approaching zero. The introduction of additional parameters that allowed the overall shape of the distribution to change, as with eqs 15 and 16, resulted in a significantly improved fit, with the average residual ~2× the inherent error. However, the fitted distributions had a large region of negative amplitude as R decreased, which was physically unreasonable. When the data were fit using Hermite polynomials (eq 17), the average residual was 1.5× the experimental error, with no improvement in the fit from $n = 2$ to $n = 4$. The theoretical phase shift and modulation derived from this fit are shown in Figure 4, together with the experimentally observed values. The distribution obtained from this fit was very similar to that for eqs 15 and 16, with a large region of negative $P(R)$, and is shown in Figure 5. As the data vector was truncated from 140 to 40 MHz, the fit improved to well within the experimental error; the shape of the distribution changed, approaching that of a shifted Gaussian centered at ~60 Å (Figure 5), bit still with an area of negative probability.

The above analyses rely on the assumption that the donor decay parameters in oligo*^{F} are unchanged in ^{R}*oligo*^{F}, except for addition of k_t for the latter case. It is this assumption that allows one to use the α_{di}'s and τ_{di}'s measured for the donor-only decay on the right-hand sides of eqs 4 and 5, thus representing the donor decay in the presence of acceptor in terms of the decay of the donor-only; the latter is modified by the term $1/\tau_{D^*}((R_o/R)^6$ to account for energy transfer. This assumption was called into question by fitting eqs 4 and 5 for the α's and τ's that described the donor-only decay, in addition to the distribution parameters. This procedure was used for the three duplex structures, with $P(R)$ modeled as a shifted Gaussian. The values obtained for α_1, τ_1, α_2, and τ_2 were very close to those actually measured for oligo*^{F}, in both low and high salt. For each case, the Simplex algorithm was run repeatedly using widely different starting conditions; without exception, the optimal parameters represented a clear minimum. That is, there was sufficient information in the phase shift and modulation data to allow the hypothetical donor-only properties to be extracted. The values obtained for the 16-mer duplex structure in 0.18 M NaCl were typical: the *fitted* values obtained for the donor-only decay in eqs 4 and 5 were $\alpha_1 = 0.81$, $\tau_1 = 4.2$ ns, $\alpha_2 = 0.19$, and $\tau_2 = 0.48$ ns, which are very similar to those measured for oligo*^{F} in a duplex (see Table 1). The values obtained for σ and $\bar{R}$ were 6.9 and 66 Å, respectively, also very similar to the values obtained in the 2-parameter fit. The sum of squared residuals was the same as that for the 2-parameter fit. For the single strand, the parameters describing donor decay obtained from the 5-parameter fit significantly differed from those actually measured for oligo*^{F}. With $P(R)$ modeled as a shifted Gaussian, the fitted values (for 0.18 M NaCl) were $\alpha_1 = 0.21$, $\tau_1 = 3.9$ ns, $\alpha_2 = 0.79$, $\tau_2 = 0.64$ ns, $\sigma = 9.6$ Å, and $\bar{R} = 68.8$ Å, and the data were fit to within experimental error. An equally good fit was obtained with essentially identical values for the α's and τ's, but with $\sigma = 12.7$ Å and $\bar{R} = 72$ Å; the correlation of these parameters led to a double minimum during the minimization.

The fitting algorithm was run repeatedly with diverse initial conditions, with convergence invariably occurring at one of these two minima. The two lifetimes were similar to those measured for oligo*^{F}; however, the values for the α's, $\alpha_1 = 0.21$ and $\alpha_2 = 0.79$, were very different from those measured, where $\alpha_1 = 0.75$ and $\alpha_2 = 0.25$ (with α_1 always assigned to the larger τ). The distribution for the single strand, with $\sigma = 9.6$ Å and $\bar{R} = 68.8$ Å, is shown in Figure 6. The 5-parameter fit for the single strand in 1 M KCl gave similar results. When the value of τ_1 was constrained in the fitting routine to be ≤4 ns, the optimal parameters were $\alpha_1 = 0.17$, $\tau_1 = 4$ ns, $\alpha_2 = 0.83$, $\tau_2 = 0.51$ ns, $\sigma = 18$ Å, and $\bar{R} = 78$ Å. When no constraints were imposed, $\alpha_1 = 0.17$, $\tau_1 = 4.2$ ns, $\alpha_2 = 0.83$, $\sigma = 17$ Å, and $\bar{R} = 74$ Å. The fit to these two cases was identical and within experimental error. The essential result obtained from the 5-parameter fit was that σ for the single-strand distribution was significantly larger than that for the duplexes, but the mean end-to-end distance was nearly unchanged from those of the 16- and 24-mer duplex structures.

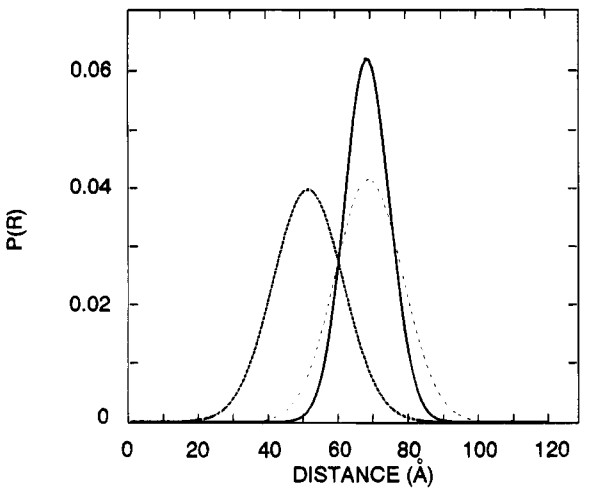

FIGURE 6: Optimal distributions (in 0.18 M NaCl) with $P(R)$ modeled as a shifted Gaussian. For the 16-mer duplex structure (—), σ = 6.4 Å and $\bar{R}$ = 68.5 Å; for the single strand (- - -), σ = 10.01 Å and $\bar{R}$ = 51.5 Å. The light line (— —) shows the distribution for the single strand derived from the 5-parameter fit (σ = 9.6 Å and $\bar{R}$ = 68.6 Å), which removed constraints on the donor-only decay; although the fit to the fluorescein decay was excellent, the results could not accommodate the emission response of the x-rhodamine.

To more clearly evaluate these results, the change upon hybridization in the integrated emission intensity of the acceptor, x-rhodamine, was considered. The integrated intensity of x-rhodamine for any given R is $(\tau_{\text{x-rhodamine}})(E_t)$. The efficiency of transfer, E_t, is $k_t/(k_F + k_I + k_t)$, and $\tau_{\text{x-rhodamine}}$ was found to have the same value in the single strand as in the duplex. To accommodate a distribution of distances, the average E_t can be evaluated from $\int P(R)\, E_t\, dR$, and this was done for each of the two components of the donor decay, α_1 and α_2. The integrated intensity was thus calculated for x-rhodamine in the duplex and in the single strand, with $P(R)$ modeled as a shifted Gaussian, and a ratio of these values was taken. When the parameters used in the integral for the single strand were those from the 2-parameter fit, the ratio of the x-rhodamine integrated intensity in the duplex to that in the single strand was 0.47, which is in good accord with the 0.5 value observed in the steady state. When the single-strand parameters were those obtained from the 5-parameter fit, however, the ratio was 1.86. Thus, although the fluorescein data were fit extremely well when the constraints on the donor-only decay were removed, the results did not account for the x-rhodamine response observed in the steady state. When the distribution for the single strand was described by a Hermite polynomial (with n = 2), the ratio was 0.49, essentially the same as that observed in the steady state.

DISCUSSION

Diverse expressions can be used to model $P(R)$ in eqs 4 and 5. In cases where the appropriate model is not known or cannot be readily determined, $P(R)$ can be extracted from the fluorescence decay data as a sum of weighted Hermite polynomials, a procedure that is model independent, provided the kernel K or K' is known. This approach was useful in analyzing the single-strand data, for which we encountered a consistent lack of fit with $P(R)$ modeled in various ways. Hermite polynomials have been used previously (Flory, 1969) in expansions of the distribution function for the end-to-end vector.

The 16- and 24-mer duplexes have $P(R)$ parameters that are consistent with the B-DNA structure. These duplex structures are considered to be sufficiently rigid that the value of σ must derive primarily from the random motion of the dye linker arms. These results are in accord with those of Hochstrasser et al. (1992), who labeled the 5′-end of each strand of a 9-mer duplex and obtained the end-to-end distance distribution from measurements made in the time domain. Sixteen base pairs compose 1.5 helical turns, causing the 3′- and 5′-ends to be positioned on opposite sides of the helical cylinder. Base pair 11 lies directly above base pair 1 when viewed down the helix axis. Consider the duplex formed when the 16-mer ^{R}*oligo*^{F} hybridizes with oligo$_{10c}$, forming 10 base pairs with 3 bases left unpaired at each end. Base pairs 1 and 10 are separated by 30.6 Å parallel to the helix axis and by 36° in a plane perpendicular to the helix axis. Were the three terminal nucleotides and dyes maximally extended and perpendicular to the helix axis, the separation between dye centers would be 40 Å (assuming the extended length of the 16-mer oligonucleotide plus dyes to be 120 Å). For the distribution parameters obtained in our fitting to be consistent with the B structure of DNA, the $\bar{R}$ value of 61.8 Å must correspond to an average structure having the three nonpaired bases in an extended conformation, as if continuing the helix. The value of σ, 6.8 Å, is scarcely different from that in the 16-mer duplex, also suggesting significant structure in the nonpaired ends of the 16-mer oligo when hybridized to the 10-mer complement. Clegg et al. (1993) considered the helical geometry in interpreting their steady-state energy transfer measurements on typical duplex structures with all bases paired.

In all cases, for both the single strands and the duplex structures, the width of the distribution increased and the mean end-to-end distance decreased in 1 M KCl relative to those in 0.18 M NaCl, which is consistent with increased flexibility due to further neutralization of negative charges along the phosphate backbone.

Interpretation of the results from the single-strand data presents a dilemma. If the full data set (5−140 MHz) is analyzed and the donor α's and τ's in eqs 4 and 5 are assumed to be those of the fluorescein in oligo*^{F}, then a model independent $P(R)$ is obtained (Figure 5) in a very good fit, but with a region of negative amplitude. Yevich and Olson (1979), in their study of random-coiling polynucleotides, found regions of negative probability around the origin in Hermite series expansions of $\mathbf{W_p}(\boldsymbol{\varrho})$ (the 3-dimensional spatial distribution of all chain vectors, expressed in terms of the displacement vectors ($\boldsymbol{\varrho}$) of the chain termini relative to the persistence vector ($\mathbf{p}$)), which were attributed to truncation of the Hermite series. In our study, the negative region for $P(R)$ is very sensitive to phase and modulation data; truncation of the high-frequency data changes the shape of the distribution to approach that of a shifted Gaussian centered at ~60 Å, although an area of negative probability is retained (Figure 5). This fit predicts a change upon hybridization for the x-rhodamine emission that is essentially identical to the change measured in the steady state. If the α_{di}'s and τ_{di}'s are allowed to vary in a 5-parameter fit with $P(R)$ as a shifted Gaussian, convergence is obtained, the fit is within experimental error (Figure 4), but the change in the integrated emission intensity of the x-rhodamine is predicted to increase by ~2-fold upon hybridization, whereas the observed value is a factor of 2 *decrease*. If, in the

2-parameter fit, $P(R)$ is required to be a shifted Gaussian, the rms residual is about 5 times the experimental error; however, the quality of the fit improves dramatically with truncation of the data set, with *very little* associated change in the optimal values for σ and $\bar{R}$. The changes in integrated emission intensity of the x-rhodamine derived from this fit agree with the observed changes in the steady-state emission. It appears that the latter $P(R)$ is the better choice. If so, the kinetic differences for hybridization between oligo*^{F} and ^{R}*oligo*^{F} reported in the preceding paper (Parkhurst & Parkhurst, 1995) are correlated with altered fractions of statically quenched fluoresceins in the two oligomers—an alteration induced in the 3′-fluorescein by the 5′-x-rhodamine. The possibility cannot be excluded, however, that the donor properties in eqs 4 and 5 *are* altered in ^{R}*oligo*^{F} compared to those of oligo*^{F}, but with no parameter constraints, an erroneous minimum is obtained. We are in the process of collecting time dependent acceptor data so as to fit, simultaneously, donor and acceptor phase and modulation vectors to constrain the fitting parameters.

It may be that a static $P(R)$ is not entirely appropriate for analysis of the data since relative donor/acceptor motion is possible in the single-strand molecule during the donor lifetime. We constructed the following simple 1-dimensional model with one donor site at a distance 60 Å from the acceptor and the other 53.4 Å from the acceptor. For the former case, the distance (R_2) is R_o; for the latter, the transfer rate constant k_t was doubled. The distributions were thus combinations of two Dirac delta functions, a simple comb distribution: $P(R) = 0.5\delta(R - R_1) + 0.5\delta(R - R_2)$. The donor lifetimes for each site in the absence of transfer were set at 4; the donor was allowed to occupy either site (assumed to be equally populated at the instant of excitation) and moved between the two sites with rate constants, k^+ and k^-, assumed to be equal. The eigenvalues for this problem are calculated readily, and the predicted biexponential time course of the donor fluorescence was calculated as a function of $k = k^+ = k^-$. If $k = 0.07$ ns^{-1}, one can recover the distances R_1 and R_2 quite well (54.90 and 61.68 Å). If $k = 0.7$ ns^{-1}, however, one recovers, in a clear misfit, a distance of 56.44 Å for both R_1 and R_2. A good fit is obtained with a shifted Gaussian with $\bar{R} = 56.5$ Å and $\sigma = 0.6$ Å. (Just as in the corresponding NMR experiment involving chemical exchange, the static distribution appears altered by the chemical exchange.) This simple calculation shows that motion of the fluorophore in times comparable to τ_{donor} will not allow the true static distribution to be recovered. In fact, the fit to the true distribution will be worse than to a shifted Gaussian. That Gaussian will give an $\bar{R}$ intermediate between the two static positions and a σ that is much too small to represent the acceptor distances to the two donor sites. A value of 0.7 ns^{-1} for the hopping frequency over 6.6 Å corresponds to a diffusion constant of 0.3×10^{-5} cm^2 s^{-1}, which is approximately what is expected in water at 20 °C for a dye the size of fluorescein plus one base as the diffusing entity.

These considerations suggest that parameters for $P(R)$ for unfolded proteins and single-stranded oligonucleotides may need to be interpreted with some caution and should be used only in a comparative sense. Experiments over a range of viscosities that still allow rotational averaging of κ^2 should further clarify this matter.

REFERENCES

Cantor, C. R., & Pechukas, P. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 2099−2101.

Cheung, H. C. (1991a) in *Topics in Fluorescence Spectroscopy* (Lakowicz, J. R., Ed.) Vol. 2, pp 140−155, Plenum Press, New York.

Cheung, H. C. (1991b) in *Topics in Fluorescence Spectroscopy* (Lakowicz, J. R., Ed.) Vol. 2, pp 158−171, Plenum Press, New York.

Clegg, R. M., Murchie, A. I. H., Zechel, A., & Lilley, D. M. J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 2994−2998.

Eis, P. S., & Lakowicz, J. R. (1993) *Biochemistry 32*, 7981−7993.

Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*, pp 397−398, Wiley-Interscience, New York.

Grinvald, A., Haas, E., & Steinberg, I. Z. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 2273−2277.

Haran, G., Haas, E., Szpikowska, B. K., & Mas, M. T. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11764−11768.

Hochstrasser, R. A., Chen, S.-M., & Millar, D. P. (1992) *Biophys. Chem. 45*, 133−141.

Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*, pp 305−309, Plenum Press, New York.

Lakowicz, J. R., Laczko, G., Cherek, H., Gratton, E., & Limkeman, M. (1984) *Biophys. J. 46*, 463−477.

Lakowicz, J. R., Johnson, M. L., Wiczk, W., Bhat, A., & Steiner, R. F. (1987) *Chem. Phys. Lett. 138*, 587−593.

Lakowicz, J. R., Gryczynski, I., Cheung, H. C., Wang, C.-K., Johnson, M. L., & Joshi, N. (1988) *Biochemistry 27*, 9149−9160.

Latt, S. A., Cheung, H. T., & Blout, E. R. (1965) *J. Am. Chem. Soc. 87*, 995−1003.

Martin, M. M., & Lindqvist, L. (1975) *J. Lumin. 10*, 381−390.

Melhuish, W. H. (1972) *J. Res. Natl. Bureau Standards 76A*, 547−560.

Noggle, J. H. (1985) *Physical Chemistry on a Microcomputer*, pp 145−165, Little, Brown and Company, Boston.

Parkhurst, K. M., & Parkhurst, L. J. (1992) *Abstracts, 11th International Congress on Photobiology*, Kyoto, Japan, p 258, Photobiology Association of Japan, Kyoto, Japan.

Parkhurst, K. M., & Parkhurst, L. J. (1993) *Abstr., Biophys. J. 64*, A266.

Parkhurst, L. J., & Parkhurst, K. M. (1994) Time-Resolved Laser Spectroscopy in Biochemistry (Lakowicz, J. R., Ed.), *Proc. SPIE 2137*, 475−483.

Parkhurst, K. M., & Parkhurst, L. J. (1995) *Biochemistry 34*, 285−292.

Steinberg, I. Z. (1971) *Annu. Rev. Biochem. 40*, 83.

Stryer, L. (1978) *Annu. Rev. Biochem. 47*, 819.

Stryer, L., & Haugland, R. P. (1967) *Biochemistry 6*, 719−726.

Tanford, C. (1961) *Physical Chemistry of Macromolecules*, p 174, J. Wiley and Sons, Inc., New York.

Washburn, E. W., Ed. (1930) *International Critical Tables 7*, pp 12 and 73, McGraw-Hill, New York.

Yang, C. H., & Söll, D. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 2838.

Yevich, R., & Olson, W. K. (1979) *Biopolymers 18*, 113−145.

BI941623Q