Global Analysis of Steady-State Polarized Fluorescence Spectra Using Trilinear Curve Resolution

G. R. Phillips and S. Georghiou

Molecular Biophysics Laboratory, Department of Physics, University of Tennessee, Knoxville, Tennessee 37996-1200 USA

ABSTRACT Global analysis using trilinear curve resolution is described and shown to be a powerful method for the resolution of polarized fluorescence data arrays, in which the measured fluorescence intensity is a separable function of polarization orientation, excitation wavelength, and emission wavelength. This methodology is applicable to mixtures the components of which have linearly independent excitation and emission spectra and distinct anisotropies. Normalized excitation and emission spectra of individual components can be uniquely determined without prior assumptions concerning spectral shapes (e.g., sum of Gaussians) and without the uncertainties inherent in bilinear techniques such as principal component analysis or factor analysis. The normalized excitation and emission vectors are combined with the total absorption spectrum of the multicomponent mixture to compute absolute absorption and emission spectra. The precision of this methodology is evaluated as a function of noise, overlap, relative intensity, and anisotropy difference between components using simulated mixtures of the DNA bases. The ability of this method to extract individual spectra from steady-state fluorescence data arrays is illustrated for mixtures containing two and three components.

INTRODUCTION

Due to its inherent sensitivity, fluorescence spectroscopy has been very useful in the investigation of biophysical systems. The enhanced information content that results from measuring fluorescence intensity as a function of both excitation wavelength and emission wavelength has long been recognized (1). Two-dimensional factor analysis (2) is the standard method used to extract the spectra of individual species from the excitation-emission matrix (EEM) of a mixture without prior knowledge of, or assumptions about, the individual spectra. Eigenvector analysis is first used to determine the abstract factors, also known as principal components or singular vectors, of the data matrix. These abstract factors are linear combinations of the absorption and emission spectra. Then non-negativity and normalization constraints are used to determine those linear combinations of the abstract factors that are suitable candidates for the spectra. Except for the trivial case of little or no overlap, however, these methods cannot produce unique resolutions of overlapped component spectra. Warner et al. (3,4) provide a detailed discussion of the ambiguities in the component spectra extracted from binary mixtures.

Global analysis is another method for analyzing two-dimensional fluorescence data sets. In contrast to methods based on factor analysis, global analysis requires a knowledge of the functional dependence of the fluorescence intensity on one dimension. For example, the time-resolved fluorescence decay of a mixture measured at several emission wavelengths can be analyzed for component decay times and emission spectra by modeling the time-dependent behavior as a sum of exponential terms (5,6). But this approach is not applicable to steady-state measurements, where assumptions cannot be made about spectral shape. A comprehensive account of the theory and algorithms of global analysis has been published by Beechem et al. (7).

Incorporation of a third independent dimension uniquely determines the component spectra to within a scalar constant. In addition to excitation and emission wavelengths, chromatographic elution time (8), the time following pulsed excitation (9–11), modulation frequency (12, 13), steady-state fluorescence polarization (14), chemical treatment (15), and fluorescence quenching (16) have been used as the third dimension. Chromatographic methods are quite useful in analytical applications but cannot be used for fluorophores that are integral parts of biomolecules, such as nucleic acids or proteins. Chemical treatment by Mg^{2+} has been used to modify the fluorescence of chloroplasts (Ref. 15 and references therein) but is not generally applicable to biomolecules. Fluorescence quenching (16) is an interesting method, albeit one that does not enable a straightforward interpretation of the data because of uncertainties regarding (a) the quenching mechanism, (b) steric hinderance effects, and (c) electrostatic effects in the case of charged quenchers. (See Ref. 17 for a recent comprehensive review of fluorescence quenching.) For technical reasons, both pulsed and frequency-modulated fluorescence measurements are very difficult to make for systems with decay components in the picosecond or femtosecond time regimes. By contrast, the measurement of steady-state EEM at two different polarization orientations

is more easily and accurately obtained. In addition, the arc lamps used in steady-state fluorescence measurements have a stability of about $\pm 1\%$ over a period of several hours (Georghiou and Phillips, unpublished observations), which cannot be currently matched by pulsed lasers used in time-resolved fluorescence spectroscopy.

Previous work on polarized excitation-emission matrices by Weidner and Georghiou (14) was limited to square matrices where the number of emitting species in the mixture equals the number of excitation and emission wavelengths. One step in that technique is the calculation of the inverse matrix of each polarized EEM—thus the limitation to square matrices of full rank. Random noise and the desired spectral signal are both included in all calculations, resulting in error propagation and making the methodology rather sensitive to noise. This sensitivity to noise is particularly undesirable for multicomponent mixtures with severely overlapped component spectra.

This paper outlines an alternative approach to the analysis of polarized fluorescence data matrices that can uniquely determine anisotropies and component spectra without artificial limits on the number of wavelengths and with much less sensitivity to noise. (Because the data sets examined here cannot be analyzed by the technique presented in Ref. 14, it is not possible to quantify the difference in sensitivity between the two methods.) This method is applicable to multicomponent systems with linearly independent component excitation and emission spectra, and two or more components with distinct anisotropies. Normalized excitation and emission spectra are first computed using the robust singular value decomposition and employing empirical methods to determinate the number of independent species, thereby eliminating the need for prior knowledge of the number of components. Next the functional dependence of the polarized fluorescence is used to determine absolute spectra and anisotropies.

THEORY

The methodology presented here is developed for polarized fluorescence data matrices measured with a single polarizer but can easily be adapted to data sets measured using two polarizers. (The use of only one polarizer significantly reduces noise and has been found to be necessary for fluorophores with very low (e.g., $\sim 10^{-4}$) fluorescence quantum yields (18,19)). The theoretical development also assumes that corrections have already been made for the variation in excitation light intensity with wavelength, the wavelength dependence of the sensitivity of the detection system, and the background signal from solvent fluorescence or from scattered light. Throughout this article, bold uppercase letters denote matrices, bold lowercase letters signify column vectors, and the superscript T indicates the transpose of a matrix or vector.

The excitation-emission-polarization array (EEPA), denoted $\mathbf{G}$, of a dilute solution of K noninteracting components

is defined as

$$\mathbf{G} = \sum_{k=1}^{K} \mathbf{a}_k \otimes \mathbf{f}_k \otimes \mathbf{o}_k = \sum_{k=1}^{K} \mathbf{x}_k \otimes \mathbf{y}_k \otimes \mathbf{z}_k \qquad (1)$$

where the vector $\mathbf{a}_k$ contains the absorption spectra, $\mathbf{f}_k$ contains the absolute emission spectra, $\mathbf{o}_k = [(2 + r_k)/3, (1 - r_k)2/3]^T$, $\mathbf{x}_k$ contains the (normalized) excitation spectra, $\mathbf{y}_k$ contains the (normalized) emission spectra, and $\mathbf{z}_k = (\|\mathbf{a}_k\| \|\mathbf{f}_k\|)\mathbf{o}_k$ for the kth component. Additionally, r_k is the anisotropy of the kth component, $\| \|$ denotes the Euclidean norm, and $\otimes$ denotes the tensor product. (For an EEPA measured using two polarizers, $\mathbf{o}_k = [(1 + 2r_k)/3, (1 - r_k)/3]^T)$. The vectors $\mathbf{x}_k$, $\mathbf{y}_k$, and $\mathbf{z}_k$ extracted from $\mathbf{G}$ using trilinear curve resolution are combined with the total absorption spectra of the multicomponent mixture to compute $\mathbf{a}_k$ and $\mathbf{f}_k$.

Since the EEPA has only two slices in the third dimension, the resolution of the three-way array $\mathbf{G}$ can be viewed as the simultaneous analysis of two polarized excitation-emission matrices (PEEMs). The PEEMs for a mixture of K fluorophores can be expressed by an $I \times J$ matrix, $\mathbf{I}_v$ or $\mathbf{I}_h$, where the ith row contains the emission spectra corresponding to excitation at wavelength λ_i and the jth column contains the excitation spectrum monitored at emission wavelength λ_j. Here, I is the number of excitation wavelengths, J is the number of emission wavelengths, and the subscripts v and h refer to vertically and horizontally polarized fluorescence. The noise-free PEEMs for a dilute solution of noninteracting fluorophores are

$$\mathbf{I}_v = \mathbf{X} \mathbf{D}_v \mathbf{Y}^T \qquad (2a)$$

$$\mathbf{I}_h = \mathbf{X} \mathbf{D}_h \mathbf{Y}^T \qquad (2b)$$

where the columns of the $I \times K$ matrix $\mathbf{X}$ and the $J \times K$ matrix $\mathbf{Y}$ contain the normalized excitation and emission spectra of the individual chemical species, $\mathbf{x}_i$ and $\mathbf{y}_i$, respectively. The $K \times K$ matrices $\mathbf{D}_v$ and $\mathbf{D}_h$ have diagonal elements $[\mathbf{D}_v]_{k,k} = z_{k,1}$ and $[\mathbf{D}_h]_{k,k} = z_{k,2}$ and off-diagonal elements equal to zero.

For noise-free data, $\mathbf{I}_v$ and $\mathbf{I}_h$ share a common row space, $\mathbf{Y}$, and a common column space, $\mathbf{X}$. The column vectors $\mathbf{x}_i$ and $\mathbf{y}_i$, for $i = 1, \ldots, K$, contain the excitation and emission spectra of the components and constitute a set of basis vectors for the column and row spaces, respectively. Since basis vectors are not uniquely defined, a set of orthonormal basis vectors for the common column and row spaces of both noise-free polarized fluorescence matrices is first determined and then transformed into the physical spectra.

Orthonormal basis vectors can be computed for noise-free data using the singular value decomposition (SVD) of either $\mathbf{I}_v$ or $\mathbf{I}_h$. For example, $\mathbf{I}_v = \mathbf{U}\mathbf{S}_v\mathbf{V}^T$, where $\mathbf{U}$ is an $I \times J$ orthonormal matrix (i.e., $\mathbf{U}^T\mathbf{U} = \mathbf{I}$), $\mathbf{S}_v$ is a $J \times J$ diagonal matrix containing the singular values in descending order, $\mathbf{V}$ is a $J \times J$ unitary matrix (i.e., $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I}$) (20). The rank of $\mathbf{I}_v$ equals the number of positive singular

values. The left and right singular vectors corresponding to positive singular values constitute a set of orthonormal basis vectors.

Because experimental PEEMs contain noise, the SVD of I_v and I_h will not produce the same singular vectors. Orthonormal basis vectors for the column and row spaces can be computed from the SVD of the augmented matrices:

$$W_1 = (I_v, I_h) = U_1 S_1 V_1^T \qquad (3a)$$

$$W_2 = \begin{pmatrix} I_v \\ I_h \end{pmatrix} = U_2 S_2 V_2^T \qquad (3b)$$

where W_1 contains I rows and $2J$ columns, and W_2 contains $2I$ rows and J columns. U_1 and V_2 can also be computed by alternating least squares (9); however, the singular value decomposition used here is more robust (20). Since all singular values will be positive, the number of independent species, denoted by K, is determined from U_1 and V_2, if it is not already known. Only the first K singular vectors contain spectral information; the other J-K columns predominantly contain random error and can be eliminated without significant distortion of the spectra.

Let $\bar{U}$ and $\bar{V}$ denote the matrices containing the K significant singular vectors in U_1 and V_2. High frequency random noise is removed from the truncated matrices and, permitting the noise-free data to be approximated,

$$P_U I_v P_V = \bar{U} C_v \bar{V}^T \cong X D_v Y^T \qquad (4a)$$

$$P_U I_h P_V = \bar{U} C_h \bar{V}^T \cong X D_h Y^T \qquad (4b)$$

where $P_U = \bar{U}\bar{U}^T$ and $P_V = \bar{V}\bar{V}^T$ are projection matrices onto the common column and row spaces and C_v and C_h are defined in terms of the noisy PEEMs

$$C_v = (\bar{U}^T\bar{U})^{-1}\bar{U}^T I_v \bar{V}(\bar{V}\bar{V}^T)^{-1} \qquad (5a)$$

$$C_h = (\bar{U}^T\bar{U})^{-1}\bar{U}^T I_h \bar{V}(\bar{V}\bar{V}^T)^{-1} \qquad (5b)$$

Since the columns of both $\bar{U}$ and X span the same space, and the columns of both $\bar{V}$ and Y span another space, there exist $K \times K$ transformation matrices P and Q such that $X = \bar{U}P$ and $Y = \bar{V}Q$. Substituting and rearranging yields

$$PD_v Q^T = C_v \qquad (6a)$$

$$PD_h Q^T = C_h \qquad (6b)$$

The transformation matrices are calculated by solving the eigenvalue equations

$$C_v C_h^{-1} = P(D_v D_h^{-1})P^{-1} \qquad (7a)$$

$$C_v^{-1}C_h = (Q^T)^{-1}(D_v^{-1}D_h)Q^T \qquad (7b)$$

It should be noted that P and Q are easily calculated without the need for any assumptions about spectral shapes (e.g., non-negativity). Additionally, the use of projection matrices in Eq. 4 reduces the influence of random noise and enables the calculation of the inverse of C_v and C_h whether the PEEMs are square matrices or not.

In contrast to two-dimensional data analysis, the transformation matrices P and Q are uniquely determined. The columns of X and Y are uniquely defined except for sign; that is, if two of the three vectors x_1, y_1, and z_1 are multiplied by -1, the resulting vectors are also solutions to Eqs. 1 and 2. This indeterminacy is removed by requiring the column vectors in X and Y to have unit length and non-negative area. Note that this differs from the non-negativity constraints used in two-dimensional methods that require each element in X and Y to be non-negative.

The anisotropies are combined with normalization constants and wavelength-independent factors in the $K \times K$ diagonal matrices D_v and D_h. The values of the diagonal elements of D_v and D_h are obtained from linear regression

$$C_v = \sum_{k=1,K} p_k q_k^T [D_v]_{k,k} \qquad (8a)$$

$$C_h = \sum_{k=1,K} p_k q_k^T [D_h]_{k,k} \qquad (8b)$$

where p_k and q_k denote the kth columns of P and Q.

Combination of normalized excitation spectra, x_k, with the steady-state total absorption spectrum measured at the same excitation wavelengths allows absolute absorption spectra to be calculated. The total absorption spectrum, $a_{tot} = \sum a_k$, corrected for path length differences between absorption and fluorescence measurements, is a linear combination of the excitation spectra: $a_{tot} = X\ b$. The vector b contains the Euclidean norm of the absorption spectra of each component (i.e., $b_k = \| a_k \|$) and can be calculated from linear regression. The absorption spectra of the kth species is the product of the kth column of X times the kth element of b. This approach should be practical for absorption spectra with small overlaps; the value of this approach for strongly overlapped spectra, which are common in biophysical problems, is evaluated in this paper using mixtures of the DNA bases.

The only assumption used so far is that the EEPA is a separable function of the excitation wavelength, emission wavelength, and polarization orientation (see Eq. 1). The functional dependence of D_v and D_h upon the anisotropy is the only additional assumption required for computation of the anisotropy and absolute emission spectra. The anisotropy of the kth component, r_k, is computed from the eigenvalues of Eq. 7a, which equal $\frac{1}{2}(2 + r_k)/(1 - r_k)$, for $k = 1, \ldots, K$. Combination of r_k and b_k with the diagonal elements of D_v allows the normalization constant for the absolute emission spectra to be calculated: $\| f_k \| = 3[D_v]_{kk}/(b_k(2 + r_k)) = 3[D_h]_{kk}/(b_k(1 - r_k))$. The absolute emission spectra, f_k, are estimated as $\| f_k \|$ multiplied by the normalized emission spectra, y_k. Error propagation from earlier computations influences the precision of absolute spectra and is examined below.

EXPERIMENTAL SECTION

Excitation and emission spectra, anisotropies, and fluorescence quantum yields in neutral aqueous solutions of adenine, thymine, cytosine, and guanine were obtained from the

literature (19). Although the anisotropy of guanine exhibits a dependence on the excitation wavelength, it has been included to test the performance of the curve resolution method under different conditions. A wavelength-independent anisotropy of 0.25 was chosen (see Table 1 in Ref. 19). The spectra were digitized in 2-nm increments and scaled with literature values of the quantum yields (19) and the extinction coefficients (21) since the spectrum of each base in Ref. 19 was normalized to unit amplitude.

Simulated EEPAs were generated by adding Gaussian random noise with mean zero and SD equal to $(Y + b)^{\frac{1}{2}}$ to the noise-free fluorescence signal, where Y is the noise-free fluorescence intensity and b approximates the average background signal. The parameter b was chosen to be equal to 50% of the mean noise-free fluorescence data matrix.

Data analysis programs were initially developed in the MATLAB programming language (The Mathworks, South Natick, MA) on a MacIntosh IIx computer. For simulation studies, the programs were written in the Microsoft FORTRAN and C languages and executed on a 80386 computer with a 80837 numeric coprocessor. Analysis of a 36 × 56 × 2 EEPA required less than 1 min.

RESULTS AND DISCUSSION

The ability to resolve EEPAs into component spectra is a complex function of spectral shape, spectral overlap, relative intensity, noise level, and anisotropy differences. Changes in one parameter affect other parameters, making a systematic study difficult. The present work has employed mixtures of adenine, thymine, cytosine, and guanine in neutral aqueous solution. The resolution of the spectra of DNA into those of its components is a very challenging problem in part because of the large overlap in the component spectra (21,22).

Fig. 1 shows the absorption and fluorescence spectra of the four bases. The fluorescence spectra have been normalized to unit area over all wavelengths with a nonzero emission intensity. Physical properties are listed in Table 1. Unless specifically stated otherwise, EEPAs were simulated for excitation wavelengths between 230 and 300 nm in 2-nm increments and emission wavelengths between 290 and 400 nm in 2-nm increments. Simulation studies have been very effective in evaluating the performance of multivariate techniques. Various factors that affect the precision of the results can be varied, and new data sets can be easily acquired.

Binary mixtures

Although this methodology is not limited to two-component EEPAs, its capabilities will first be evaluated using equimolar binary mixtures. Table 2 lists the six mixtures simulated, along with their spectral overlaps and the relative intensities of the components in the EEPAs. The relative intensity shown in Table 2 is the ratio of the area under the (unpolarized) EEM of the first component to the area under the EEM of the second component. The overlap between two
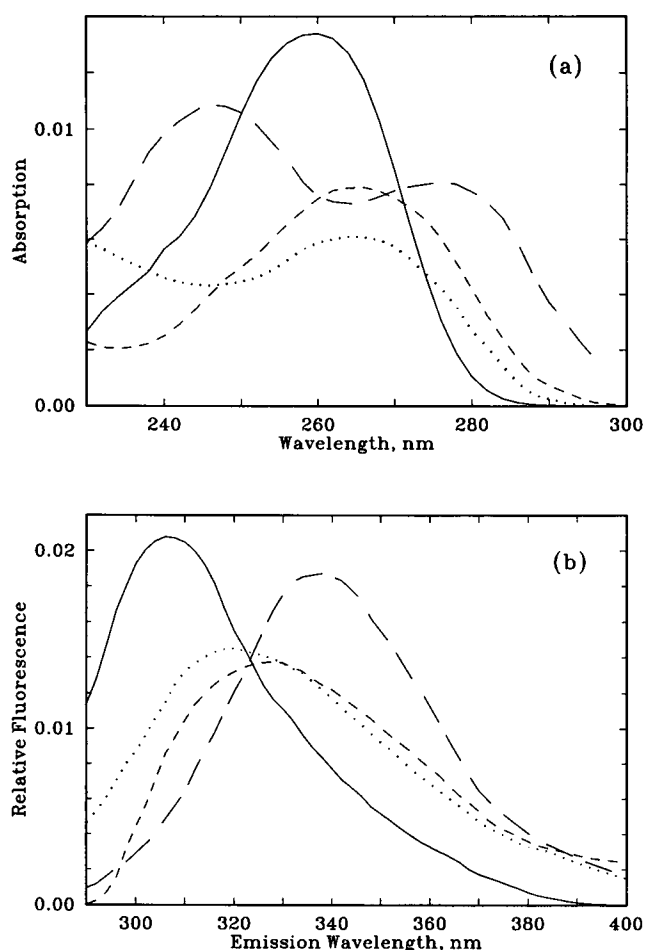

FIGURE 1   (a) Absorption and (b) fluorescence spectra of 1 μM solutions of adenine (——), thymine (– – – –), cytosine (·····), and guanine (— —). The fluorescence curves have been normalized to unit area and do not include the quantum yields. (The data have been obtained from Refs. 19 and 21.)

TABLE 1 Physical properties of DNA bases

Base	Quantum yield*	Molar extinction coefficient[‡]	Anisotropy*
Adenine	2.6×10^{-4}	13.4×10^3	0.24
Thymine	1.0×10^{-4}	7.9×10^3	0.34
Cytosine	0.8×10^{-4}	6.1×10^3	0.30
Guanine	3.0×10^{-4}	10.7×10^3	0.25

*Ref. 19.

[‡]Maximum molar extinction coefficient in units of M^{-1} cm^{-1} (Ref. 21).

spectra, say x_1 and x_2, is defined (3, 4) as

$$\zeta = x_1^T x_2 / (\| x_1 \| \| x_2 \|) \tag{9}$$

and has also been known as spectral similarity (23). Superimposable spectra have an overlap of $\zeta = 1$; spectra with no points in common have a similarity or overlap of $\zeta = 0$. Besides excitation and emission spectra, the third dimension (i.e., polarization orientation) can also be characterized by an overlap. The total overlap equals the product of the overlap

TABLE 2 Spectral characteristics of binary mixtures

Data set	Bases*	Spectral overlap‡				Intensity§
		ζ_x	ζ_y	ζ_z	Total	
1	A, T	0.929	0.784	0.996	0.725	4.0:1
2	A, C	0.913	0.895	0.998	0.816	4.9:1
3	T, C	0.935	0.976	0.999	0.912	1.2:1
4	G, T	0.894	0.650	0.997	0.859	6.3:1
5	G, C	0.938	0.912	0.999	0.855	7.7:1
6	G, A	0.854	0.650	1.000	0.555	1.6:1

*Abbreviations: A, adenine; T, thymine; C, cytosine; G, guanine.

‡Total overlap = $\zeta_x \zeta_y \zeta_z$, where ζ_x is the overlap between x_k, ζ_y is the overlap between y_k, and ζ_z is the overlap between the o_k.

§The relative intensity is defined as the ratio of the area under the EEM of the first component to the area under the EEM of the second component.

in each of the three dimensions (see Eq. 1). Theoretically the resolution of the three-way data array **G** can be enhanced by differences in anisotropy between two components. If the vectors z_1 and z_2 for the components of a binary mixture are orthogonal (i.e., have zero overlap), then the individual spectra can be uniquely resolved no matter how great the overlap in the excitation and emission spectra. This case does not occur for EEPAs, however. The overlap in the third dimension will be large (≥ 0.996 for the six mixtures in Table 2) and at first glance does not appear to significantly enhance the resolution of individual components.

Previous work by Millican and McGown (13) on the resolution of the phase-resolved fluorescence of binary mixtures does not transfer to EEPAs. That work studied excitation-emission-frequency arrays where the third (i.e., frequency) dimension consisted of three to six different modulation frequencies chosen to enhance individual components. For example, a mixture of two fluoranthenes and three modulation frequencies has an overlap in the frequency dimension equal to 0.82, thereby decreasing the total overlap and increasing the resolution of the two components. This additional resolution is not available in EEPAs since (a) the third dimension (i.e., polarization orientation) is limited to only two levels and (b) these levels are fixed and cannot be experimentally optimized. Consequently, it is important to evaluate what limits, if any, this places on the resolution of EEPAs.

For each mixture in Table 2, the signal-to-noise ratio (S/N) was varied by changing the maximum intensity in I_v. Variation in S/N may correspond to different integration times, number of scans added together, excitation light intensity, concentrations, or fluorescence quantum yields. Because there is no generally accepted definition for the S/N of a multiway data array with nonconstant noise, it is convenient for this study to define the relative signal-to-noise ratio (RSN) as the signal-to-noise ratio at the peak of I_v. Consequently, the S/N is much lower at all other wavelengths and throughout I_h. These values of RSN are not comparable to the S/N values calculated using other definitions, but span the amount of noise typically found in the EEPAs of biophysical systems. Twenty simulated data matrices were generated for

each set of conditions, and the variation of results between individual analyses was used to estimate the precision of the analysis.

Fig. 2 shows a simulated EEPA for mixture 1 with RSN = 50. The extracted excitation and emission spectra are shown in Fig. 3a along with the spectra used to simulate the EEPA. Most of the variation is concentrated in the spectra for thymine. This is a result of the difference in intensities between the two components (see Table 2) and is not immediately apparent since these are normalized spectra. Absolute absorption and absolute emission spectra are shown in Fig. 3b. The factors influencing the precision of these spectra are discussed below. A visual inspection of Fig. 2 suggests that only one rather than two components are present, making the results of the analysis in Fig. 3 even more impressive.

Normalized excitation and emission spectra

The relative standard error (RSE) for spectral curves is computed as the SD between the synthetic and extracted spectra averaged over 20 simulations and divided by the Euclidean norm of the synthetic spectra. The RSEs for the excitation and emission spectra are shown in Figs. 4 and 5 as a function of noise for the mixtures in Table 2. For clarity, not all 12 curves are shown. In discussing the trend in the RSE, the following convention will be used: $RSE(x_A, 1)$ symbolizes the RSE in x_A in mixture 1, for example. The letter is an abbreviation for the base (i.e., A = adenine, T = thymine, C = cytosine, and G = guanine), and the number corresponds to the binary mixture in Table 2.

In addition to noise, the resolution of EEPAs is also influenced by overlap, relative intensity, and anisotropy. As
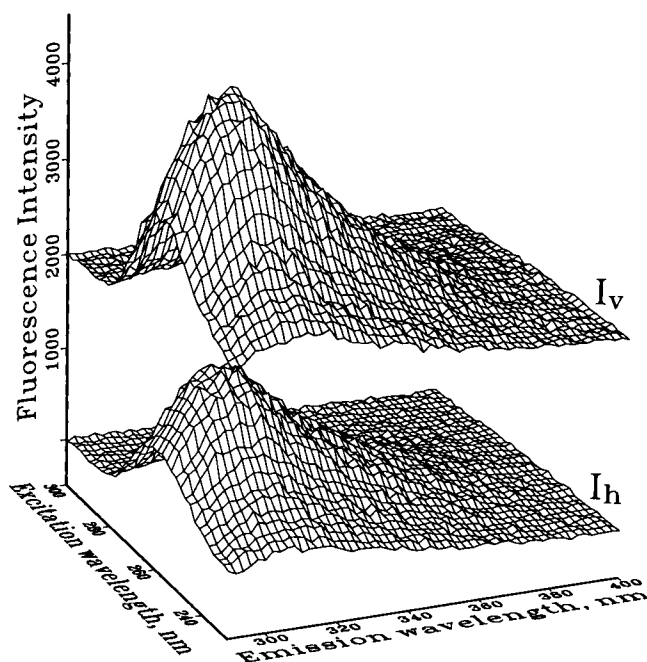

FIGURE 2 Simulated EEPA for mixture 1 (Table 2) corresponding to a RSN of 50 in I_v. I_v is offset by 2000 counts for clarity.
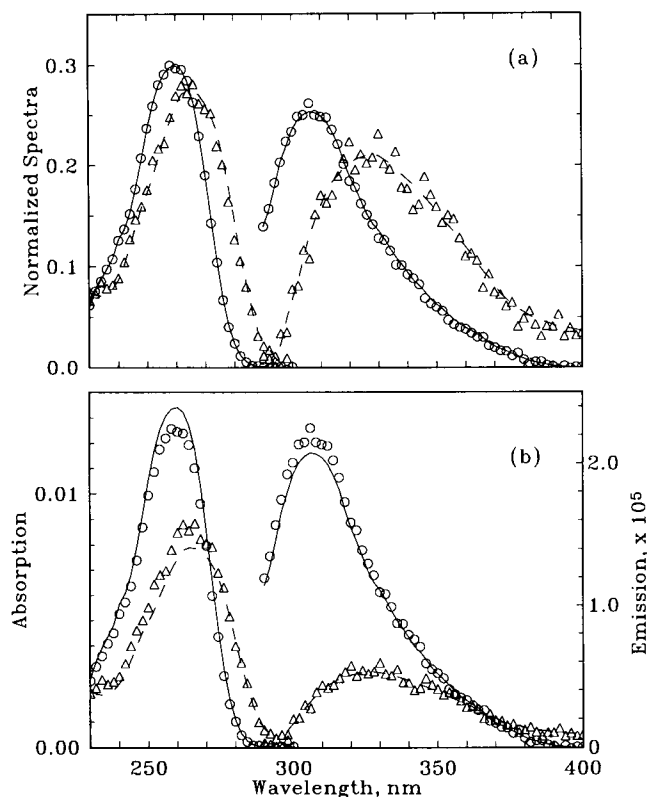
FIGURE 3   (a) Normalized excitation and emission spectra and (b) absolute absorption and emission spectra extracted from the EEPA in Fig. 2. The extracted spectra (○,△) are shown along with the synthetic spectra of adenine (——) and thymine (– – – –).
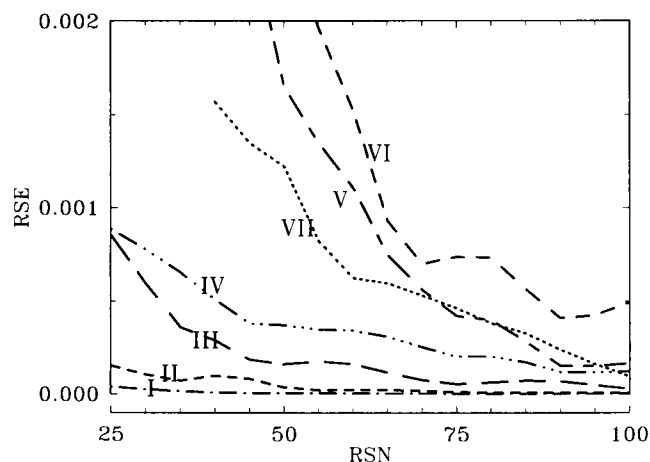

FIGURE 4   RSE for the excitation spectra, x, versus the RSN in binary mixtures: (*I*) RSE (x_A, 1); (*II*) RSE (x_A, 2); (*III*) RSE (x_T, 1); (*IV*) RSE (x_C, 3); (*V*) RSE (x_T, 3); (*VI*) RSE (x_C, 2); and (*VII*) RSE (x_A, 6) and RSE (x_G, 6). For example, III denotes the RSE in x_T in mixture 1.

expected, the RSE decreases (Figs. 4 and 5) as the total overlap between the two components decreases (Table 2). For mixtures with a common component, RSE(x_T, 1) is much less than RSE(x_T, 3), RSE(x_A, 1) is slightly less than RSE(x_A, 2), and RSE(x_G, 4) ≈ RSE(x_G, 5). This is consistent with the decrease in the difference between the total
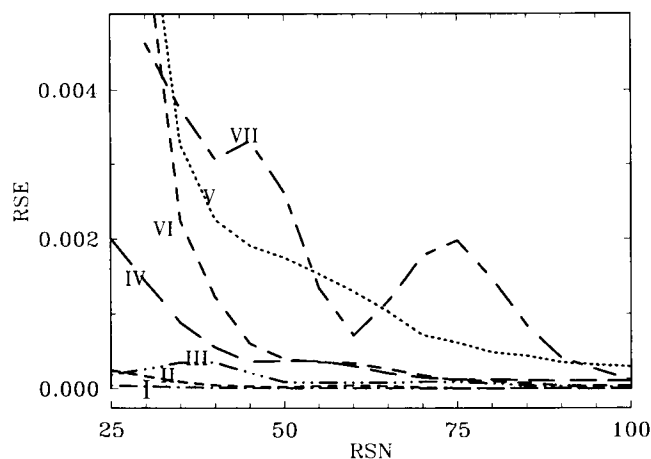

FIGURE 5   RSE for the emission spectra, y, versus the RSN in binary mixtures: (*I*) RSE (y_A, 1); (*II*) RSE (y_A, 2); (*III*) RSE (y_T, 1); (*IV*) RSE (y_C, 3); (*V*) RSE (y_T, 3); (*VI*) RSE (y_C, 2); (*VII*) RSE (y_A, 6) and RSE (y_G, 6). For example, III denotes the RSE in y_T in mixture 1.

overlap between mixtures 1 and 3, 1 and 2, and 4 and 5, respectively. A notable exception to this trend is mixture 6. Here RSE(x_A, 6) ≫ RSE(x_A, 2), rather than much smaller, as the overlap would suggest. This anomaly can be explained by the anisotropy difference of only Δr = 0.01 (Table 1) and by the approximately equal contribution of A and G to the fluorescence in mixture 6. The small anisotropy difference limits the accuracy with which the transformation matrices **P** and **Q** can be determined. The equal contributions of A and G in mixture 6 result in noise being distributed evenly between both A and G instead of being concentrated in the minor component as in mixtures 1 and 2. The similar values of RSE(x_A, 1), where Δr = 0.10, and RSE(x_A, 2), where Δr = 0.06, as well as similar values of the RSE for G in mixture 4 (with Δr = 0.09) and in mixture 5 (with Δr = 0.05) indicate that anisotropy differences of 0.05, and possibly less (preliminary results), are sufficient to adequately resolve component spectra. This, of course, presumes that all other factors are the same in this comparison. A third parameter affecting the precision of extracted spectra is the relative intensity between the two components in the mixture. As noted above, the RSEs of A and G are approximately equal in mixture 6, where both components contribute equally to the total fluorescence. In mixtures where one component dominates, the RSE of the strong component is less than that of the weak component. For example, the RSE of A is much less than that of T in mixture 1, and the RSE of A is much less than that of C in mixture 2. The differences in the fluorescence quantum yield between two components, which have not been included in Fig. 1b and can easily be overlooked, are important factors in the resolution of component spectra.

The dependence of the RSE of the emission spectra on the RSN in Fig. 5 is the same as the RSE of the excitation spectra in Fig. 4, with only small differences. Because of the interaction between overlap, relative intensity, and anisotropy differences, these small differences cannot be reliably ex-

plained. However, it can be safely concluded that the resolution of EEPAs depends on the total overlap, not just on the overlap between excitation or emission spectra. The relationships $RSE(x_G, 4) \approx RSE(x_G, 5)$ and $RSE(y_G, 4) \approx RSE(y_G, 5)$ correlate with the total overlap, for example, but not with the values of ζ_x and ζ_y in Table 2.

Anisotropy differences

The fluorescence intensity of each component in I_v and I_h is proportional to the diagonal elements in D_v and D_h. If all components have the same anisotropy, then measurement of the EEPA is equivalent to measuring the (attenuated) EEM twice and contains no useful information for deriving the rotation matrices P and Q. When the anisotropies of the components are different, one can think of I_v and I_h as two EEMs of a mixture with the same components but with different "effective" concentrations, $\frac{1}{3}(2 + r_k)c_k$ and $\frac{2}{3}(1 - r_k)c_k$, where c_k is the analytical concentration and r_k is the anisotropy. Since the value of the anisotropy is not a variable under the control of the observer, it is important to examine the sensitivity of this analysis to differences in anisotropy.

Table 3 summarizes the relative standard deviation (RSD) of the anisotropy of each base in the mixtures in Table 1. The values in Table 3 have been calculated by pooling the estimated anisotropies for all replicate simulations at levels of the RSN. Except for mixture 6, the same trend in precision will be seen when examining the RSDs at each noise level. The RSD is correlated with the relative contribution of each base to the binary mixture, rather than with the value of Δr, as might first be expected. When the signal of one component dominates, the anisotropy of the component with the smaller relative intensity has a larger RSD (e.g., mixtures 1, 2, 4, 5). Mixture 3 has components with approximately equal contributions and a larger total overlap. In this case, noise is partitioned between both components rather than being concentrated in the minor one.

Mixture 6, $\Delta r = 0.01$, is an exception to the above generalizations. The estimated anisotropy of adenine is very precise, in contrast to that of guanine. It is surprising that 97% of the simulations estimate the anisotropy of guanine with a RSD of <2%. Each estimate of the anisotropy is accurate to

within 4% of the true anisotropy. For the other simulations, the anisotropy of guanine is severely in error. These large errors are detected by unrealistic anisotropy values. Resolution of EEPAs with very small anisotropy differences is enhanced by high S/N.

Absolute absorption and emission spectra

Sanchez and Kowalski (24,25) have proposed a generalized rank annihilation method to compare the two-dimensional EEM of an unknown mixture with one or more calibration mixtures to determine the relative concentrations and normalized spectra of the constituent species. This method requires EEMs of two or more mixtures of different concentrations and cannot be applied to a single unknown mixture or to fluorophores that are integral parts of biomolecules such as nucleic acids or proteins. By comparison with the case of normalized spectra, the determination of absolute spectra has received little attention.

Combination of normalized excitation and emission spectra with the steady-state total absorption spectrum measured at the same excitation wavelengths enables calculation of absolute absorption and emission spectra. This approach should be practical for absorption spectra with small overlaps; however, the value of this approach for strongly overlapped spectra, which are common in biophysical problems, needs to be evaluated. Although the absolute spectra involve the solution of matrix equations via linear least-squares minimization, errors estimated using the variance-covariance matrix greatly underestimate the precision in the absolute absorption spectra. This is a consequence of error propagation in calculations of the normalized excitation spectra, making the theory of errors in linear regression inapplicable.

The RSE of the absolute spectra of each base should have a dependence on noise similar to that of the RSE of the normalized spectra because $a_k \propto x_k$ and $f_k \propto y_k$. If the normalization constants could be exactly determined, the RSE of absolute and normalized spectra would be identical. The RSE of absolute spectra relative to the RSE of normalized spectra indicate how efficiently the normalization constants, and equivalently absolute spectra, can be calculated. The ratios of the RSE of a_k to the RSE of x_k for the binary mixtures considered here are summarized in Table 4; a similiar summary for the absolute and normalized emission spectra is presented in Table 5.

TABLE 3   Relative standard deviation (RSD) of estimated anisotropy in binary mixtures

Mixture*	Adenine	Thymine	Cytosine	Guanine
1	1.5%	4.3%	—	—
2	1.5%	—	5.9%	—
3	—§	5.4%	3.2%	—
				3.9%
4	—	11%	—	
5	—	—	13%	4.4%
6	1.3%	—	—	<2%‡

*Mixtures as defined in Table 2.

‡Based on 97% of the stimulated EEPAs, the RSD of the anisotropy of guanine is <2%; the remaining simulated EEPAs produce unrealistic estimates for the anisotropy of guanine.

§Blanks indicate that entries are nonapplicable.

TABLE 4   Ratio of RSE between absorption and excitation spectra*

Mixture‡	Adenine	Thymine	Cytosine	Guanine
1	76	6.8	—	—
2	112	—	4.9	—
3	—§	10	29	—
4	—	3.6	—	51
5	—	—	10	59
6	11	—	—	9.0

*RSE of absorption spectra divided by RSE of excitation spectra.

‡Mixtures as defined in Table 2.

§Blanks indicate that entries are nonapplicable.

TABLE 5 Ratio of RSE between absolute and normalized emission spectra*

Mixture[‡]	Adenine	Thymine	Cytosine	Guanine
1	12	1.3	—	—
2	68	—	1.8	—
3	—[§]	7.3	1.3	—
4	—	3.7	—	76
5	—	—	3.0	332
6	29	—	—	6.3

*RSE of absolute emission spectra divided by RSE of normalized emission spectra.
[‡]Mixtures as defined in Table 2.
[§]Blanks indicate that entries are nonapplicable.

As expected, Table 4 shows that absolute spectra are less precise than normalized spectra. The effects of total overlap, relative intensity, and anisotropy differences on calculated spectral shapes has been taken into account by forming the ratio of the RSE of a_k and that of x_k. Since normalization constants are computed using the total absorption spectra, one might expect that mixtures the components of which have large differences in absorption (see Table 1) or the excitation spectra of which lack wavelengths at which only one component absorbs will have greater imprecision. A complicating factor is the correlation between the normalization constants. Specifically, when the absorption of one component is overestimated, the absorption of the second component is underestimated, and vice versa. This makes interpretation of the values in Table 4 difficult, although in general the mixtures the components of which have large differences in the maximum molar extinction coefficient and ζ_x will be less precise. A parallel discussion for Table 5 is more complicated since the estimation of the absolute emission spectra will incorporate errors in the normalized emission spectra as well as errors in the normalization constants for absorption spectra and involves a larger number of wavelengths than the absorption spectra.

Multicomponent mixtures

The algorithm used here is not limited to binary mixtures. This is a significant advantage compared to the nonlinear iterative least-squares algorithm used by Appellof and Davidson (8) and Russell and Gouterman (9,10), which limits the maximum number of resolvable components to the smallest dimension of the three-way data array (i.e., two for an EEPA). Additionally, in the algorithm described here, only two components of a multicomponent mixture must have distinct anisotropies.

This may be illustrated by a mixture containing 1 μM adenine, 3 μM thymine, and 1 μM guanine. The 1:3:1 concentration ratio was chosen to avoid the imprecision in the extracted spectra due to large differences in the relative contributions to the fluorescence. The anisotropy difference between adenine and guanine is only $\Delta r = 0.01$. Although the resolution of multicomponent mixtures only requires that two or more components have different anisotropies, this small value of Δr makes the problem more difficult. The

simulated EEPA (not shown) is a broad surface with the same noise level as the binary mixture of adenine and thymine illustrated in Fig. 2 (i.e., RSN = 50). The extracted spectra, shown in Fig. 6, are less precise for the three-component mixture than for two-component mixtures (compare Figs. 6 and 3). The normalized spectra are more accurately estimated than the absolute spectra, which is not unexpected, since there are no excitation wavelengths at which only one of the bases absorbs (Fig. 1a).

We have observed that simulated three-component mixtures containing either adenine, thymine, and cytosine or adenine, thymine, and guanine with a RSN $\geq$ 50 could successfully resolve four of the six normalized spectra, with the accuracy of the remaining two spectra varying between replications. Increased RSN or the number of wavelengths in the simulated EEPAs allows the components to be estimated more precisely.

The methodology presented here is applicable to biomolecules in which energy transfer between fluorophores is inefficient. Inefficient transfer of energy may be the result of (a) a large distance between the potential donor and the potential acceptor, (b) a low value of the average molar extinction coefficient of the acceptor in the region of overlap between its absorption spectrum and the fluorescence spectrum of the donor, (c) a low value of the fluorescence quantum yield of the donor (in the absence of energy transfer), or (d) an unfavorable orientation between the transition dipole
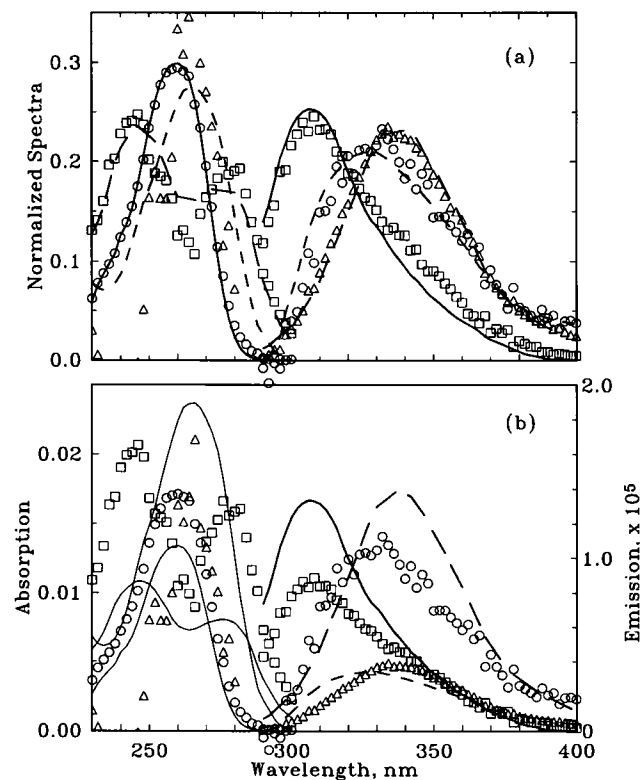

FIGURE 6   (a) Normalized excitation and emission spectra and (b) absolute absorption and emission spectra extracted from an EEPA with RSN = 50. The extracted spectra (O, △, □) are shown along with the synthetic spectra of adenine (——), thymine (– – – –), and guanine (— —).

moments of the donor and the acceptor (26). The incorporation of energy transfer into this analysis is a challenging task, one that has yet to be accomplished by any of the currently available methodologies for analyzing spectral data (8–16). Work in progress in our laboratory is attempting to do that.

We should note that an alternative, attractive methodology for resolving component spectra may be offered by pulsed (9–11) and frequency-modulated (12,13) fluorescence measurements. In fact, that would represent a multilevel approach in the third dimension instead of the bilevel one (two polarization orientations) presented here for steady-state fluorometric measurements. As was discussed in the Introduction, however, current technical difficulties associated with picosecond and femtosecond measurements as well as with the laser stability have not yet allowed that technique to reach its full potential with regard to resolving component spectra.

CONCLUSIONS

Global analysis using trilinear resolution is shown to be valuable in resolving component spectra from EEPAs. Simulated data sets composed of the severely overlapped spectra of the DNA bases provide a realistic test of this methodology. Despite the existence of only two fixed levels in the third dimension (i.e., polarization orientation) and large overlaps (Table 2), trilinear curve resolution can successfully resolve EEPAs. The results presented here and the availability of computer-controlled fluorometers for data acquisition suggest that this is a promising technique.

REFERENCES

1. Weber, G. 1961. Enumeration of components in complex systems by fluorescence spectrophotometry. Nature (Lond.). 190:27–29.

2. Malinowski, E. R. 1991. Factor Analysis in Chemistry. Ed. 2. John Wiley and Sons, Inc., New York.

3. Warner, I. M., G. D. Christian, E. R. Davidson, and J. B. Callis. 1977. Analysis of multicomponent fluorescence data. Anal. Chem. 49:564–573.

4. Warner, I. M. 1982. The analysis of matrix formatted multicomponent data. In Contemporary Topics in Analytical Chemistry. D. M. Hercules, G. M. Heiftje, L. R. Snyder, and M. A. Evenson, editors. Plenum Publishing Corp., New York. 75–140.

5. Knorr, F. J., and J. M. Harris. 1981. Resolution of multicomponent fluorescence spectra by an emission wavelength-decay time data matrix. Anal. Chem. 53:272–276.

6. Knutson, J. R., J. M. Beechem, and L. Brand. 1983. Simultaneous analysis of multiple fluorescence decay curves: a global approach. Chem. Phys. Lett. 102:501–507.

7. Beechem, J. M., E. Gratton, M. Ameloot, J. R. Knutson, and L. Brand. 1991. The global analysis of fluorescence intensity and anisotropy decay data: second-generation theory and programs. In Topics in Fluorescence Spectroscopy. Vol. 2. J. R. Lakowicz, editor. Plenum Publishing Corp., New York. 241–305.

8. Appellof, C. J., and E. R. Davidson. 1981. Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents. Anal. Chem. 53:2053–2056.

9. Russell, M. D., and M. Gouterman. 1988. Excitation-emission-lifetime analysis of multicomponent systems. I. Principal component factor analysis. Spectrochim. Acta Part A Mol. Spectrosc. 44A:857–861.

10. Russell, M. D., and M. Gouterman. 1988. Excitation-emission-lifetime analysis of multicomponent systems. II. Synthetic model data. Spectrochim. Acta Part A Mol. Spectrosc. 44A:863–872.

11. Russell, M. D., M. Gouterman, and J. A. van Zee. 1988. Excitation-emission-lifetime analysis of multicomponent systems. III. Platinum, palladium and rhodium porphyrins. Spectrochim. Acta Part A Mol. Spectrosc. 44A:873–882.

12. Burdick, D. S., X. M. Tu, L. B. McGown, and D. W. Millican. 1990. Resolution of multicomponent fluorescent mixtures by analysis of the excitation-emission-frequency array. J. Chemometrics. 4:15–28.

13. Millican, D. W., and L. B. McGown. 1990. Fluorescence lifetime resolution of spectra in the frequency domain using multiway analysis. Anal. Chem. 62:2242–2247.

14. Weidner, R., and S. Georghiou. 1990. Global methods for time-resolved and steady-state absorption spectroscopy. Proc. SPIE 1204:717–726.

15. Ross, R. T., C-H. Lee, C. M. Davis, B. M. Ezzeddine, E. A. Fayyad, and S. E. Leurgans. 1991. Resolution of the fluorescence spectra of plant pigment-complexes using trilinear models. Biochim. Biophys. Acta. 1056:317–320.

16. Leurgans, S., and R. T. Ross. 1992. Multilinear models: applications in spectroscopy. Statist. Sci. 7:289–319.

17. Eftink, M. R. 1991. Fluorescence quenching: theory and applications. In Topics in Fluorescence Spectroscopy. Vol. 2. J. R. Lakowicz, editor. Plenum Publishing Corp., New York. 53–126.

18. Ge, G., and S. Georghiou. 1991. Excited-state properties of the alternating polynucleotide poly(dA-dT)·poly(dA-dT). Photochem. Photobiol. 54:301–305.

19. Callis, P. R. 1979. Polarized fluorescence and estimated lifetimes of the DNA bases at room temperature. Chem. Phys. Lett. 61:563–567.

20. Shrager, R. I. 1986. Chemical transitions measured by spectra and resolved using singular value decomposition. Chemometrics Intelligent Lab. Sys. 1:59–70.

21. Voet, D., W. B. Gratzer, R. A. Cox, and P. Doty. 1963. Absorption spectra of nucleotides, polynucleotides, and nucleic acids in the far ultraviolet. Biopolymers. 1:193–208.

22. Georghiou, S., G. R. Phillips, and G. Ge. 1992. Resolution of the electronic absorption spectra of the adenine and thymine residues in poly-(dA)·poly(dT). Biopolymers. 32:1417–1420.

23. Gemperline, P. J., and J. C. Hamilton. 1989. Conditions for detecting overlapped peaks with principal component analysis in hyphenated chromatographic methods. Anal. Chem. 61:2240–2243.

24. Sanchez, E., and B. R. Kowalski. 1986. Generalized rank annihilation factor analysis. Anal. Chem. 58:496–499.

25. Sanchez, E., and B. R. Kowalski. 1990. Tensorial resolution: a direct trilinear decomposition. J. Chemometrics. 4:29–45.

26. Förster, T. 1965. Delocalization excitation and excitation transfer. In Modern Quantum Chemistry, Part III. O. Sinanoglu, editor. Academic Press, New York. 93–137.