Decomposition of Protein Tryptophan Fluorescence Spectra into Log-Normal Components. I. Decomposition Algorithms

Edward A. Burstein, Sergei M. Abornev, and Yana K. Reshetnyak
Institute of Theoretical and Experimental Biophysics, Russia Academy of Sciences, Pushchino, Moscow region, Russia 142290

ABSTRACT Two algorithms of decomposition of composite protein tryptophan fluorescence spectra were developed based on the possibility that the shape of elementary spectral component could be accurately described by a uniparametric log-normal function. The need for several mathematically different algorithms is dictated by the fact that decomposition of spectra into widely overlapping smooth components is a typical incorrect problem. Only the coincidence of components obtained with various algorithms can guarantee correctness and reliability of results. In this paper we propose the following algorithms of decomposition: (1) the **SI**mple fitting procedure using the root-**M**ean-**S**quare criterion (SIMS) operating with either individual emission spectra or sets of spectra measured with various quencher concentrations; and (2) the pseudo-graphic analytical procedure using a **PH**ase plane in coordinates of normalized emission intensities at various wavelengths (wavenumbers) and **RE**solving sets of spectra measured with various **Q**uencher concentrations (PHREQ). The actual experimental noise precludes decomposition of protein spectra into more than three components.

INTRODUCTION

The fluorescence parameters of tryptophan residues are sensitive to the microenvironment of fluorophore in protein structure. For this reason, fluorescence characteristics are widely used to study physico-chemical and dynamic properties of tryptophan microenvironment in proteins and the structural transitions and behavior of protein molecules as a whole (Lakowicz, 1983, 1999; Demchenko, 1986). The overwhelming majority of proteins exhibit smooth, nonstructured spectra of tryptophan fluorescence, which often contain more than one component. The multicomponent nature of protein spectra makes the unequivocal interpretation of them difficult and poses a task of development of methods for the decomposition of tryptophan fluorescence spectra into elementary components (Burstein et al., 1973; Burstein, 1977).

The problem of decomposition of multicomponent spectra belongs to the class of typical, so-called reverse problems, because one must determine the parameters of spectral components from the overall experimental spectrum, where the components are indirectly manifested. Solutions of such problems are, as a rule, unstable against slight variations in the input data (noise). Because the real input data are known approximately (i.e., with some experimental error), this instability results in an inevitable ambiguity of the solution within the given accuracy. In this respect Tikhonov and Arsenin (1986) classified such a problem as an incorrect one. To obtain a sufficiently stable solution, it is necessary to formulate a principle of choosing among the possible solutions, based on an additional information about the system under study and the solution quality. The application of additional information forms a basis for regularizing the solving (Tikhonov and Arsenin, 1986). The regularizing factors (functions, algorithms, or logical premises) allow one to develop practical ways of solving incorrect problems. To decompose the multicomponent protein tryptophan spectra, we used the following regularizing factors (Abornev and Burstein, 1992; Abornev, 1993):

1. *The spectrum of an elementary component on the frequency (wave number) scale is described by a biparametric (maximal amplitude and maximum position) log-normal function* (Burstein and Emelyanenko, 1996). The quadriparametric (maximal amplitude, I_m, spectral maximum position, ν_m, and positions of half-maximal amplitudes, ν_- and ν_+; see Fig. 1) log-normal function has been proposed by Siano and Metzler (1969) for describing the absorption spectra of complex molecules and was later successfully used to resolve multicomponent absorption spectra, including those of biological systems (Metzler et al., 1972, 1985, 1991; Morozov and Bazhulina, 1989). The log-normal function used in its mirror-symmetric form has been shown to accurately describe fluorescence spectra as well (Burstein, 1976; Burstein and Emelyanenko, 1996). The straight linear relationships between positions of maximal (ν_m) and two half-maximal amplitudes (ν_- and ν_+) have been revealed for a large series of monocomponent spectra of small tryptophan derivatives in various solvents and allowed to reduce the number of unknown parameters from four to two (Burstein and Emelyanenko, 1996). Such a reduction of number of parameters sought is known to make a decomposition much more unambiguous (Antipova-Korotaeva and Kazanova, 1971). As a result, the biparametric log-normal function (uniparametric one for the spec-

Received for publication 12 January 2001 and in final form 7 June 2001.
Address reprint requests to Dr. Edward A. Buzstein, Institute of Theoretical and Experimental Biophysics, Russia Academy of Sciences, Pushchino, Moscow region, Russia 142290. Tel.: 0967-733319; Fax: 0967-790553; E-mail: burstein@fluoz.iteb.sezpukhov.su.

0006-3495/01/09/1699/11 $2.00

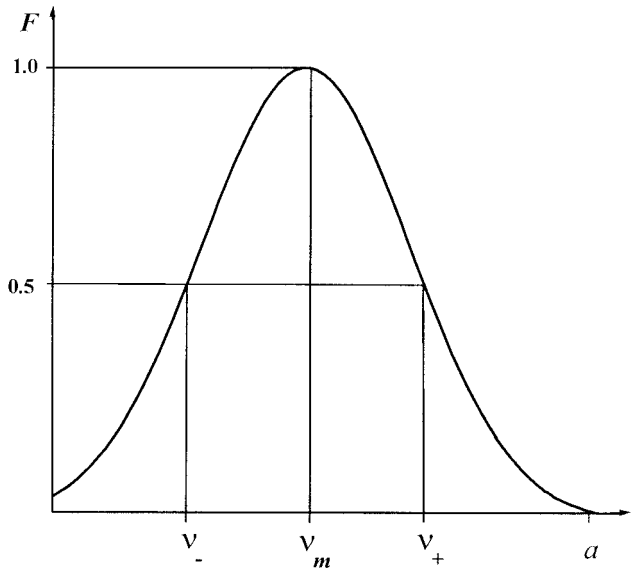

FIGURE 1 The quadriparametric (maximal amplitude, I_m, spectral maximum position, ν_m, and positions of half-maximal amplitudes, ν_- and ν_+) log-normal function.

tral shape) for fluorescence spectra of tryptophan and its residues in proteins appears as follows:

$$\begin{cases} I(\nu) = I_m \cdot \exp\left\{ -\dfrac{\ln 2}{\ln^2 \rho} \cdot \ln^2\left(\dfrac{a-\nu}{a-\nu_m}\right)\right\} & (\text{at } \nu < a) \\ I(\nu) = 0 & (\text{at } \nu \geq a) \end{cases} \quad (1)$$

where I_m is the maximal intensity; ν is the current wavenumber; ρ is the band asymmetry parameter $\rho = (\nu_m - \nu_-)/(\nu_+ - \nu_m)$; a is the function-limiting point position $a = \nu_m + (\rho \cdot (\nu_+ - \nu_-))/(\rho^2 - 1)$. Therewith, ν_m, ν_+, and ν_- are related as (Burstein and Emelyanenko, 1996):

$$\nu_+ = 0.831 \cdot \nu_m + 7070 \ (\text{cm}^{-1})$$

$$\nu_- = 1.177 \cdot \nu_m - 7780 \ (\text{cm}^{-1}) \quad (2)$$

2. *The shape and position of tryptophan emission spectra remain unchanged at fluorescence quenching with water-soluble small quenchers* (Burstein, 1968, 1977; Lehrer, 1971; Lehrer and Leavis, 1978). A series of spectra measured at various quencher concentrations represents a set of data wherein the position and shape parameters of components are constant and only their relative contributions are variable. Such an expansion of the statistical mass of data, compared with that of an individual spectrum, also aids the rise of solution reliability.

3. *The change of amplitudes of individual components with quenching obeys the Stern-Volmer law* (Burstein, 1977; Lehrer, 1971; Lehrer and Leavis, 1978).

4. An additional important factor, which results in rise of reliability of the decomposition procedure, is the use of *statistically redundant information*: practically all experimentally determined points in the input spectra are used for the decomposition, i.e., the number of experimental points under analysis far exceeds the number of parameters sought. This approach attenuates the effect of occasional noise and improves the accuracy of results (Akseenko et al., 1989).

The above-mentioned factors allowed us to develop several methods for sufficiently stable decomposition of composite tryptophan fluorescence spectra of proteins without exceeding the experimental error. The practical need in more than one mathematically different algorithm is a consequence of the fact that the individual component spectra are wide (25–61 nm) compared with the spectral interval within which their maxima may be positioned (307–355 nm). Thus, the coincidence of parameters of components revealed with diverse methods can guarantee the reliability of decomposition results for a given protein. Here we describe two methods, first of which is based on the fitting of experimental spectra by a sum of log-normal components using the root-mean-square criterion of fitting quality, and the second algorithm uses an analytical pseudo-graphic solving the task.

MATERIALS AND METHODS

Materials

The annexin VI was supplied by Dr. Andrzej Sobota, N. Nencki Institute of Experimental Biology, Warsaw, Poland (Bandorowicz et al., 1992; Sobota et al., 1993). KI, KC, and $Na_2S_2O_3$ were commercial preparations of ultra-high purity of Russian industry production.

Fluorescence spectra

The fluorescence spectra were measured using the lab-made spectrofluorimeter with registration from the front surface of the cell (Bukolova-Orlova et al., 1974). The 296.7-nm mercury line from the ultra-high-pressure mercury lamp SVD-120A (The Moscow Electrolamp Factory, Moscow) was used for excitation. The slit widths of the excitation and output monochromators did not exceed 2 nm. After correction for instrument spectral sensitivity, the intensities were proportional to the number of photons emitted in the unit wavelength interval. Under the measurements of fluorescence spectra with different concentrations of KI the total ionic strength was kept constant (0.4 M) by addition of KCI. The stock solution of KI contained $Na_2S_2O_3$ to prevent oxidation of I^-. The decomposition algorithms were used as Visual Basic programs in personal computers.

THE FITTING ALGORITHM WITH MINIMIZATION OF ROOT-MEAN-SQUARE RESIDUES (SIMS)

Algorithm description

The below-described algorithm we called SIMS (**SI**mple fitting procedure using the root-**M**ean-**S**quare criterion) (Abornev and Burstein, 1992; Abornev, 1993).

Because, under the different concentrations of external fluorescence quenchers, the position and shape of spectral components remain unchanged while the relative contributions of components to the overall spectrum are changed, then the experimental spectra can generally be described as follows:

$$F(i, j) = \sum_{k=1}^{L} I(k, i) \cdot \varphi(k, j) \tag{3}$$

where $i = 1, \ldots, N$ is the number of spectra corresponding to the ith quencher concentration, $c(i)$; $j = 1, \ldots, M$ is the number of current frequency (wave number), $\nu(j)$; $k = 1, \ldots, L$ is the number of components determined by the position of its spectral maximum, $\nu_m(k)$; $F(i, j)$ is the experimental intensity of fluorescence on the wave number scale in the ith spectrum at the jth current frequency $\nu(j)$; $\varphi(k, j)$ is the value of the log-normal function with a maximum at $\nu_m(k)$ at current frequency $\nu(j)$ with unite maximal amplitude (at given k and j, this value is the same for any of the N spectra); and $I(k, i)$ is the maximal amplitude of the kth component in the ith spectrum.

The problem is to find the positions $\nu_m(k)$ and the maximal intensities $I(k, i)$ of log-normal spectral components from the set of experimental spectra $F(i, j)$; Eq. 3, however, cannot be solved analytically because the log-normal function $\varphi(k, j)$ is essentially transcendental with respect to the unknown $\nu_m(k)$ (see Eq. 1). Hence, the solution can be sought for by approximation, e.g., using the fitting of $\nu_m(k)$ values by minimization of residuals. At each step of this process, the transcendental $\varphi(k, j)$ terms are computed for a given $\nu_m(k)$ value, and then the corresponding $I(k, i)$ values are easily determined analytically, solving the set of linear equations.

The fact that individual components in protein fluorescence spectra are very broad and mutually overlapped poses severe limitations on the procedure of searching for a functional minimum. Attempts to use modern fast fitting methods revealed a strong dependence of solutions on the initial conditions. Only the exhaustive enumeration of ν_m values (with successively diminishing steps from ~8 nm down to 0.1 nm) avoided "trapping" in the local minima of the functional (rms residuals) and, thus, to find its global minimum. It is essentially important in the presence of experimental noise. Moreover, it obviates the need to set any arbitrary initial conditions, which often leads to the erroneous result of solving an incorrect reverse problem (Tikhonov and Arsenin, 1986). This notwithstanding, the results of decomposition of experimental and simulated multicomponent spectra showed that the typical experimental noise of ~0.5–1.5% does not permit a sufficiently reliable decomposition for more than three spectral components. Therefore, we shall consider this limiting case with $L \leq 3$ in describing the algorithm. Uni-, bi-, and tri-component solutions are searched independently by turn for the set of experimental spectra. However, we shall consider below the tri-component solution as a more general case.

With fixed $\nu_m(1)$, $\nu_m(2)$, and $\nu_m(3)$ values at current wave number j at each fitting step, the solution can be found on the basis of the minimal least-square formalism, i.e., when the S would be minimal:

$$S = \sum_{i=1}^{N}\sum_{j=1}^{M}\left[\sum_{k=1}^{3} I(k, i) \cdot \varphi(k, j) - F(i, j)\right]^2$$

The unknowns are $I(k, i)$. The $\varphi(k, j)$ values are calculated from Eq. 1 at given $\nu_m(k)$ and $\nu(j)$ values. The criterion S is minimal when $dS/dI(k, i) = 0$. These conditions allow construction of N sets of three equations in each:

$$\frac{dS}{dI(k, i)} = \sum_{j=1}^{M}\left[\sum_{k=1}^{3} I(k, i) \cdot \varphi(k, j) - F(i, j)\right] \cdot \varphi(k, j) = 0 \tag{4}$$

or, after opening the brackets,

$$\sum_{j=1}^{M}\varphi(k, j) \cdot \sum_{k=1}^{3} I(k, i) \cdot \varphi(k, j) = \sum_{j=1}^{M} F(i, j) \cdot \varphi(k, j)$$

Transposing the summation over k and j in the left part, we can write down the whole set of Eq. 4:

$$\begin{cases} \sum_{k=1}^{3}\left[I(k, i) \cdot \sum_{j=1}^{M}\varphi(k, j) \cdot \varphi(1, j)\right] = \sum_{j=1}^{M} F(i, j) \cdot \varphi(1, j) \\ \sum_{k=1}^{3}\left[I(k, i) \cdot \sum_{j=1}^{M}\varphi(k, j) \cdot \varphi(2, j)\right] = \sum_{j=1}^{M} F(i, j) \cdot \varphi(2, j) \\ \sum_{k=1}^{3}\left[I(k, i) \cdot \sum_{j=1}^{M}\varphi(k, j) \cdot \varphi(3, j)\right] = \sum_{j=1}^{M} F(i, j) \cdot \varphi(3, j) \end{cases} \tag{4a}$$

Because all the sums over j are known, we obtain N nonuniform sets ($i = 1, \ldots, N$) of linear equations, where each set contains three equations and can be solved independently, the main determinant being the same for all N sets. These canonical sets of linear equations are then solved (i.e., the $I(k, i)$ amplitudes are evaluated) using the routine Gauss method. An analogous algorithm was developed for the decomposition of an individual emission spectrum (the program SIMS-MONO). In this case, the set of equations is constructed using an expanded set of points of a single spectrum.

Then, at each step we determine the sum of absolute values (modules) of residuals S (differences between calculated and experimental intensities):

$$S = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{L}\left|I(k, i) \cdot \varphi(k, j) - F(i, j)\right| \tag{5}$$

and the parameter D, which characterizes the quality of accordance of spectral components quenching with the Stern-Volmer law, i.e.,

$$D = \frac{1}{3} \cdot \sum_{k=1}^{L} R_{sd}(k) \tag{6}$$

$$R_{sd} = \frac{1}{Y(k, N)} \cdot \left[\sum_{i=1}^{N} \frac{(Y(k, i) - X(k, i))^2}{N} \right]^{1/2} \tag{6a}$$

Here R_{sd} is the relative root-mean-square residual between the values $X(k, i) = I(k, 1)/I(k, i)$ determined by solving Eq. 4a and the $Y(k, i)$ values calculated from the linear equation of the Stern-Volmer law:

$$Y(k, i) = K_{s\nu}(k) \cdot a(i) + B(k) \tag{7}$$

where $a(i)$ is an activity of ionic quenchers that is calculated using values of concentrations $c(i)$ (Hodgman et al., 1955). In the programs, the relations between $c(i)$ and $a(i)$ values were included analytically in the polynom forms.

The resulting combined minimization criterion (functional) is used in the form:

$$S_1 = S \cdot (1 + D)$$

A set of components, i.e., the values of spectral maxima positions ν_m and maximal amplitudes I_m, corresponding to the global minimum of S_1 is considered as the solution of Eq. 4a. The above-described algorithm of three-component decomposition can be, in principle, expanded over an arbitrary number of components.

The procedure of searching for a sufficient number of components describing a series of experimental spectra of a protein is carried out as follows. The experimental series of spectra is consecutively decomposed into one, two, and three components. Because the experimental spectra are measured with a constant wavelength increment, each value set (i.e., $\nu_m(1)$, $\nu_m(2)$, and $\nu_m(3)$) on the frequency scale ($\nu_m(cm^{-1}) = 10^7/\lambda_m(nm)$) is determined by exhaustion in the wavelength range from 300 to 370 nm with consecutive three-times shortening steps from 8.1 to 0.1 nm. The intensities on the frequency (F_ν) and the wavelength (F_λ) scales are related as $F_\nu = F_\lambda \cdot \lambda^2$; hence the spectra of individual components are converted onto the wavelength scale and the maximum position of components ($\nu_m(k)$) are presented on the wavelength scale ($\lambda_m(k)$).

To estimate the quality of decomposition, the relative rms residual of theoretical and experimental spectra is expressed as a percentage of maximal amplitude F_m of the spectrum, measured in the absence of quencher:

$$Ts = T \cdot (1 + D) \tag{8}$$

Where D is determined with Eq. 6 and

$$T = \frac{1}{M} \cdot \sum_{j=1}^{M} S^2(j) \tag{9}$$

$$S(j) = \frac{1}{N} \cdot \sum_{i=1}^{N} s(i, j) \tag{10}$$

$$s(i, j) = \frac{\sum_{k=1}^{L} I(k, i) \cdot \varphi(k, j) - F(i, j)}{F_m(c = 0)} \tag{11}$$

To choose among the one-, two-, or three-component solutions as being more reliable, we used the discriminant D_s values equal to the product of functional T_s by the number of components searched for, i.e., the number of parameters (L) varied under fitting:

$$D_s = L \cdot T_s \tag{12}$$

Before determining the final results, the same procedure is used for the smoothing experimental spectra. As a rule, the spectra contain some points distorted by Raman line, scattered mercury lines from the light source, and/or by a random noise. After each decomposition cycle, the intensity values differing by >2% from the theoretical ones were changed to be equal to the latter. The smoothing cycles are repeated while such differences disappear, but their number is not to exceed 10 to avoid an eventual distortion of the spectral shape.

Properties of the algorithm

To test how various factors affect the accuracy of solution, a series of decompositions were carried out for simulated spectra. The latter were sums of two or three log-normal curves varying in the positions of maxima and relative amplitudes. Because, above all, we were interested in the quality of spectral resolution of components, the precision was evaluated as a value of $\Delta\lambda$, which is a mean absolute difference between the positions of component maxima (in nanometers) that were preset in simulation and those obtained after the decomposition. Fig. 2 shows the $\Delta\lambda$ value dependencies for two- and three-component decompositions, crosses and squares, respectively, of the amplitude of randomly introduced noise, S% (panel *A*); the number of spectra (N) with various "quencher concentrations" used (panel *B*); the distance between the component spectral maxima, $\Delta\lambda_{max}$, nm (panel *C*); the ratio of Stern-Volmer constants of quenching for the components, $K2/K1$ (panel *D*); the number of registered points in each spectrum, M (panel *E*); and the contribution of one component in the total spectrum, as an intensity ratio $I1/(I1 + I2)$ (panel *F*). The noise was introduced as random equally probable positive and negative deviations with amplitude from 0 to S%

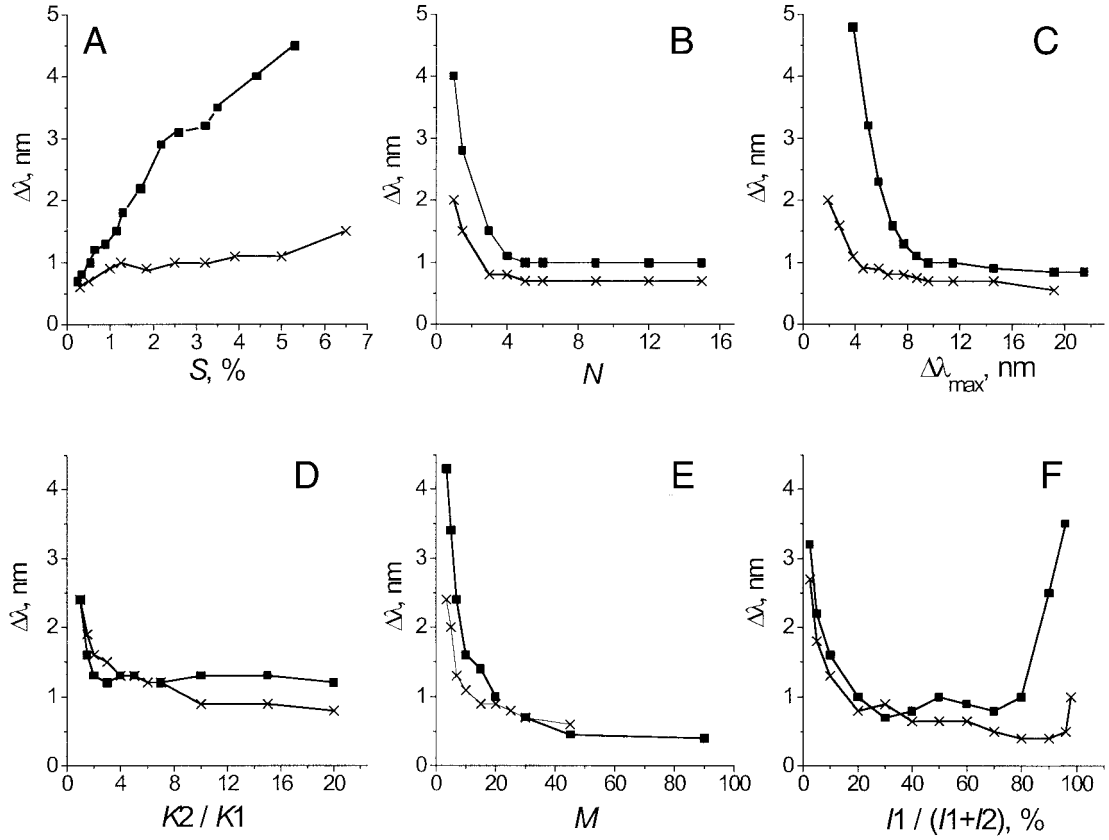

FIGURE 2 The properties of the SIMS algorithm. The dependencies for two- and three- component decompositions, crosses and squares, respectively, between $\Delta\lambda$ values (mean absolute difference between positions of component maxima (in nanometers) that were preset in simulation and those obtained after the decomposition) and the amplitude of randomly introduced noise, $S\%$ (*A*); the number of spectra (N) with various "quencher concentrations" used (*B*); the distance between the component spectral maxima, $\Delta\lambda_{max}$, nm (*C*); the ratio of Stern-Volmer constants of quenching for the components, $K2/K1$ (*D*); the number of registered points in each spectrum, M (*E*); and the contribution of one component in the total spectrum, as an intensity ratio $I1/(I1 + I2)$ (*F*).

of the theoretical amplitude. The maximal "quencher concentrations" were such that would reduce the amplitude of a total spectrum approximately by half. Varying one parameter, the others were held constant at the following standard values: $S\% = 0.6\%$; $N = 5$; $M = 15$; $\Delta\lambda_{max} = 10$ nm; $K1 = 0.1\ M^{-1}$; $K2 = 3.0\ M^{-1}$; $I1 = I2$. For three-component decomposition the standard values of $S\%$, N, M, and $\Delta\lambda_{max}$ are the same as for two-component ones, but $K1 = 1\ M^{-1}$, $K2 = 5\ M^{-1}$, $K3 = 0\ M^{-1}$, and $I1 = I2 = I3$. Only $K1$ and $K2$ change in panel *D*, and $I1$ and $I2$ in panel *F*. As can be seen, the method provides an acceptable accuracy ($\Delta\lambda < 1$ nm in the two-component and <1.5 nm in the three-component decomposition) in determining the true positions of component maxima at values of initial parameters usually existing in practice: $S\% = 0.5$–1.5; $N = 3$–10; $M = 10$–20; $\Delta\lambda_{max} > 7$ nm; contribution of an individual component of 10–90%. The accuracy of decomposition into three components is somewhat worse than that into two components; however, it is also quite satisfactory, taking into account the large overlapping of the component spectra.

THE ANALYTICAL ALGORITHM BASED ON THE PSEUDO-PHASE REPRESENTATION (PHREQ)

The algorithm PHREQ (the **PH**ase-plot-based **RE**solution using **Q**uenchers) is an analytical realization of a graphical way of two-component decomposition of a set of protein-tryptophan fluorescence spectra measured at various concentrations of quencher (Abornev, 1993). The method uses quasi-phase representation of parameters characterizing the shape of fluorescence spectra (Burstein, 1976), which was already successfully applied for analysis of protein structural transitions (Kaplanas et al., 1975; Permyakov et al., 1980a,b).

In the most general case, to use any physical parameter P either to characterize the shape of transition curve $A \rightarrow B$ or to estimate the proportion of components in the mixture of A and B components, this parameter (P) should be linearly related to either extent of transition completing (α) or contribution of a mixture component into total concentration.

$$\alpha = \frac{c_B}{c_A + c_B}$$

where c_A and c_B are concentrations of A and B. Then, the composite parameter P for the system can be expressed as:

$$\begin{aligned} P &= \beta \cdot P_A + \alpha \cdot P_B \\ &= (1 - \alpha) \cdot P_A + \alpha \cdot P_B \\ &= P_A + \alpha \cdot (P_B - P_A) \end{aligned} \quad (13)$$

Therefore, P is the weighted mean of values P_A and P_B, which characterize the pure A and B components, respectively. The weight factors are $\beta = 1 - \alpha$ and α, respectively.

A two-state transition or a mixture can be characterized by two mutually independent physical parameters P_1 and P_2 linearly related to α:

$$P_1 = \beta \cdot P_{A1} + \alpha \cdot P_{B1}$$

$$P_2 = \beta \cdot P_{A2} + \alpha \cdot P_{B2}$$

It can be simply shown that parameters P_1 and P_2 are linear functions of one another at the same α value, i.e.,

$$P_1 = k \cdot P_2 + m \quad (14)$$

Where k and m are factors expressed through the values of P_{A1}, P_{A2}, P_{B1}, and P_{B2}, and the plane with coordinates (P_1, P_2) possesses a property of phase-plane. In such a plane the states A and B are represented by points (P_{A1}, P_{A2}) and (P_{B1}, P_{B2}), respectively. The transition between A and B is reflected by the totality of phase-points, which should be onto the straight line connecting points A and B. At any intermediate point Z with coordinates (P_1, P_2) located on this line ($0 < \alpha < 1$), α is proportional to the ratio of lengths of segment between the points Z and A and segment between points A and B, i.e., α = AZ/AB. These simple relationships allow graphic determination of contributions of components A and B at any step of transition or in two-component mixtures.

By the way, in the case of transitions measured by kinetics or equilibrium shift, the deviation from linearity of trajectory AB suggests the existence of one or more intermediate states in the process. In the simplest cases it could be evaluated P_1 and P_2 for the intermediate state by extrapolating the initial and last linear parts of the trajectory to point of their intersection, assumedly representing the intermediate in the plane.

In case of decomposition of fluorescence spectra, the "physical" state is a position and shape of a spectral component, which remains unchanged under different concentrations of fluorescence quenchers. However, added quenchers perturb the "spectral" state, i.e., change the ratios of component contributions. The role of parameters P_1 and P_2 for two-component fluorescence spectra plays the emission intensities $F(\nu_1)$ and $F(\nu_2)$ measured at different wavenumbers ν_1 and ν_2. Such an approach demands the equality of the number of photons absorbed by both components in the unit time interval. The parameters $F(\nu_1)$ and $F(\nu_2)$ reflect the contributions of α and β of two components into the total emission spectrum. To obtain the linear trajectory on the quasi-phase plane [$F(\nu_1)$, $F(\nu_2)$] it can change the α value by measuring spectra at various concentrations of fluorescence quenchers (Cs^+, I^-, acrylamide, etc.), which change intensities of two components in different degrees depending on solvent accessibility of fluorophore(s) from which a component originates, but do not affect the shape and maximum positions of spectral components (Burstein, 1968, 1976, 1977, Lehrer, 1971; Lehrer and Leavis, 1978).

In the ideal imaginary case, when the quenching does not change summary emission quantum yield, on the quasi-phase plane [$F(\nu_1)$, $F(\nu_2)$] the points obtained at several quencher concentrations lie on a straight line connecting the points of "pure" spectral forms ("ν_{m1}" and "ν_{m2}" in Fig. 3), of which the total spectra consist ($\alpha = 0$ and $\alpha = 1$). In reality, to exclude the perverted effects of quenchers on total quantum yield it is necessary to normalize $F(\nu_1)$ and $F(\nu_2)$ values by either total surface area under spectrum or by emission intensity at any third, constant wave number ν_n, $F(\nu_n)$. Because the precise measurement of the area under an experimental spectrum is almost impossible, we used normalizing by $F(\nu_n)$:

$$P_{n1} = \frac{F(\nu_1)}{F(\nu_n)}$$

$$P_{n2} = \frac{F(\nu_2)}{F(\nu_n)} \quad (15)$$

The $F(\nu_1)$, $F(\nu_2)$, and $F(\nu_n)$ can be represented as the combinations of normalized log-normal functions ($\varphi(\nu_{mi}, \nu_j)$) at wavenumbers ν_1, ν_2, and ν_n with maxima at ν_{m1} and ν_{m2} according to Eq. 3, i.e.,

$$P_{n1} = \frac{I_1 \cdot \varphi(\nu_{m1}, \nu_1) + I_2 \cdot \varphi(\nu_{m2}, \nu_1)}{I_1 \cdot \varphi(\nu_{m1}, \nu_n) + I_2 \cdot \varphi(\nu_{m2}, \nu_n)}$$

$$P_{n2} = \frac{I_1 \cdot \varphi(\nu_{m1}, \nu_2) + I_2 \cdot \varphi(\nu_{m2}, \nu_2)}{I_1 \cdot \varphi(\nu_{m1}, \nu_n) + I_2 \cdot \varphi(\nu_{m2}, \nu_n)}$$

Thus, P_{n1} and P_{n2} could be presented as:

$$\begin{aligned} P_{n1} &= \varphi_r(\nu_{m1}, \nu_1) + \alpha \cdot [\varphi_r(\nu_{m2}, \nu_1) - \varphi_r(\nu_{m1}, \nu_1)] \\ P_{n2} &= \varphi_r(\nu_{m1}, \nu_2) + \alpha \cdot [\varphi_r(\nu_{m2}, \nu_2) - \varphi_r(\nu_{m1}, \nu_2)] \end{aligned} \quad (16)$$

where $\varphi_r(\nu_{mi}, \nu_j) = \varphi(\nu_{mi}, \nu_j)/\varphi(\nu_{mi}, \nu_n)$ and $\varphi(\nu_{mi}, \nu_j)$ are values of log-normal functions with maximal amplitudes equal 1 (see Eqs. 1 and 2) and maximum positions ν_{mi} at current wavenumber ν_j. In such a representation, α means the contribution of the component with the maximum position at $\nu_{m2}(f(2))$ in the normalizing fluorescence intensity at ν_n, $F(\nu_n)$.

$$\alpha \equiv f(2) = \frac{I_2 \cdot \varphi(\nu_{m2}, \nu_n)}{I_1 \cdot \varphi(\nu_{m1}, \nu_n) + I_2 \cdot \varphi(\nu_{m2}, \nu_n)} \quad (17)$$

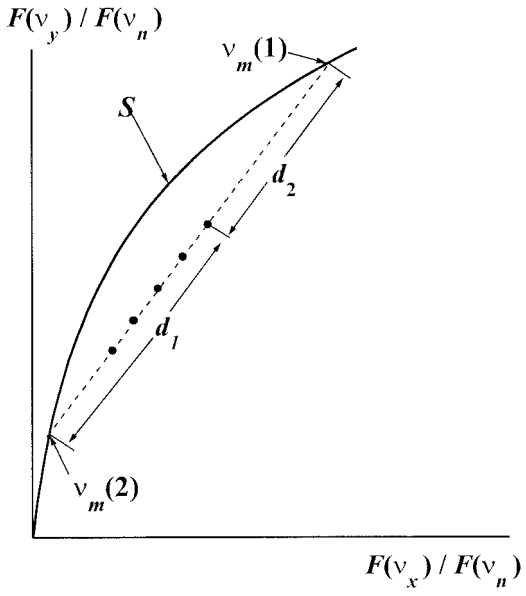

FIGURE 3 The representation of fluorescence spectra measured at different concentration of quenchers as points on the quasi-phase plane. The curve *S* corresponds to the totality of all possible elementary log-normal functions.

As well as in the case of $P_1 = F(\nu_1)$ and $P_2 = F(\nu_2)$, P_{n1} and P_{n2} are mutually linearly related (the equation that is analogous to Eq. 14):

$$P_{n1} = k \cdot P_{n2} + m$$

where:

$$k = \frac{\varphi_r(\nu_{m2},\nu_1) - \varphi_r(\nu_{m1},\nu_1)}{\varphi_r(\nu_{m2},\nu_2) - \varphi_r(\nu_{m1},\nu_2)}$$

$$m = \frac{\varphi_r(\nu_{m1},\nu_1) \cdot \varphi_r(\nu_{m2},\nu_2) - \varphi_r(\nu_{m2},\nu_1) \cdot \varphi_r(\nu_{m1},\nu_2)}{\varphi_r(\nu_{m2},\nu_2) - \varphi_r(\nu_{m1},\nu_2)}$$

Therefore, the points obtained at various quencher concentrations form the linear track on the phase-plane (P_{n1}, P_{n2}) (see Fig. 3). Thus, the phase-plot in coordinates (P_{n1}, P_{n2}) (Fig. 3) can be used for estimating the main parameters of the two-component spectrum, i.e., ν_{m1} and ν_{m2} and their relative contributions α and $(1 - \alpha)$. To estimate the components' maximal positions ν_{m1} and ν_{m2} we used the extrapolation of the linear track through the experimental points, obtained with various quencher concentrations, up to

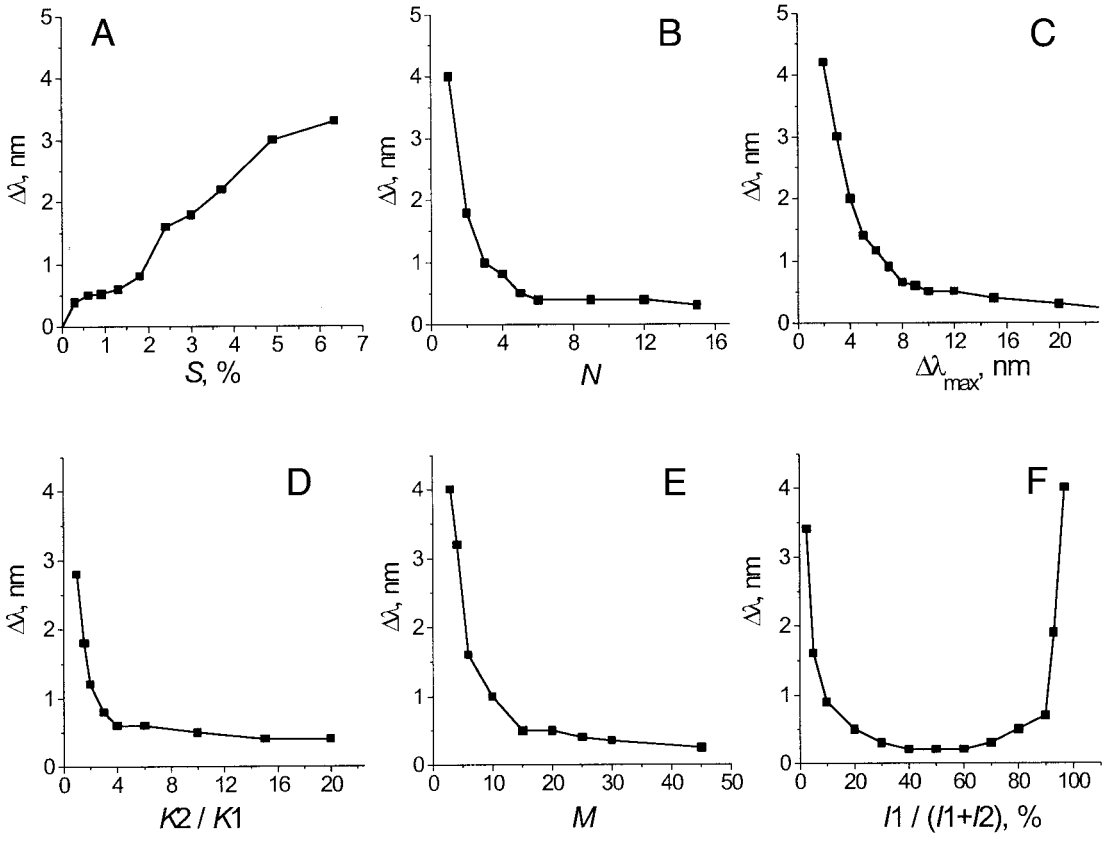

FIGURE 4 The properties of the PHREQ algorithm. The dependencies for two-component decompositions between Δλ values (mean absolute difference between positions of component maxima (in nanometers) that were preset in simulation and those obtained after the decomposition) and the amplitude of randomly introduced noise, *S*% (*A*); the number of spectra (*N*) with various "quencher concentrations" used (*B*); the distance between the component spectral maxima, $\Delta\lambda_{max}$, nm (*C*); the ratio of Stern-Volmer constants of quenching for the components, *K*2/*K*1 (*D*); the number of registered points in each spectrum, *M* (*E*); and the contribution of one component in the total spectrum, as an intensity ratio *I*1/(*I*1 + *I*2) (*F*).

TABLE 1 The results of decomposition of tryptophan fluorescence spectra of annexin VI (50 mM cacodylate, pH 7.0) measured at three different concentrations of KI (0.0, 0.2, and 0.4 M) into log-normal components by the SIMS and PHREQ methods

N	λ_m (nm)	ν_m (cm^{-1})	I_i	A_i	S_i (%)	$K_{s\nu}$ (M^{-1})	$K_{s\nu}^{rel}$ (%)	R	B	R_{sd}
SIMS, 1-component solution (D_s = 3.222)										
1	331.3 ± 0.5	29949 ± 45	10225	494795	100	1.05 ± 0.17	7.2 ± 1.2	0.986	1.01	0.015
			8571	414728	100					
			7939	384178	100					
SIMS, 2-component solution (D_s = 0.757)										
1	326.0 ± 0.5	30460 ± 46	8200	441495	74.0	0.82 ± 0.08	5.6 ± 0.5	0.995	1.01	0.008
			7203	387793	77.6					
			6692	360302	77.9					
2	347.7 ± 0.5	28458 ± 40	2875	154778	26.0	1.88 ± 0.48	12.9 ± 3.3	0.969	1.04	0.035
			2075	111720	22.4					
			1897	102159	22.1					
SIMS, 3-component solution (D_s = 1.045)										
1	317.2 ± 0.5	313.48 ± 49	1144	64681	10.3	0.06 ± 0.39	0.4 ± 2.7	0.159	1.31	0.042
			1040	58768	11.2					
			1128	63779	13.9					
2	329.4 ± 0.5	30139 ± 45	8422	478038	75.6	1.05 ± 0.07	7.2 ± 0.5	0.998	1.01	0.006
			7210	407500	77.5					
			6528	369004	75.4					
3	356.3 ± 0.5	27724 ± 38	1575	88995	14.1	2.09 ± 0.79	14.3 ± 5.4	0.936	1.06	0.054
			1057	59734	11.4					
			1001	56595	11.6					
PHREQ (D_s = 0.527)										
1	325.5 ± 2.9	30514 ± 270	7660	363370	67.7	0.75 ± 0.02	5.1 ± 0.1	1.0	1.00	
			6875	326121	72.8					
			6346	301067	72.7					
2	346.5 ± 3.1	28564 ± 253	3586	173357	32.3	1.94 ± 0.60	13.3 ± 4.1	0.956	1.04	
			2522	121928	27.2					
			2343	113292	27.3					

N = the number of log-normal components;
λ_m = the maximum position of log-normal components in wavelength scale;
ν_m = the maximum position of log-normal components in frequency scale;
I_i = the maximal intensity of log-normal components at ith concentration of KI, i.e., at 0, 0.2, and 0.4 M;
A_i = the area of log-normal component under the total spectrum at ith concentration of KI;
S_i = the contribution of log-normal component (in percent) into the area under the total spectrum at ith concentration of KI;
$K_{s\nu}$ = the Stern-Volmer quenching constant for each log-normal component;
$K_{s\nu}^{rel}$ = the relative Stern-Volmer quenching constant, i.e., the ratio of the $K_{s\nu}$ for each log-normal component to the $K_{s\nu}$ for free aqueous tryptophan emission quenching with KI, which was taken as 14.6 M^{-1} (Burstein, 1977);
R = the coefficients of linear correlation of the parameters $X(k, i) = I(k, 1)/I(k, i)$ and $c(i)$ on the plots in Stern-Volmer coordinates;
B = the free parameter in Eq. 7;
R_{sd} = the relative root-mean-square residuals, see Eq. 6a.

its intersection with the curve S, which corresponds to the totality of all possible elementary log-normal functions normalized by the same $F(\nu_n)$ value as the coordinate values P_{n1} and P_{n2}. From the distances between the experimental point and the points ν_{m1} and ν_{m2}, d_2 and d_1, respectively, the contributions of components can be calculated:

$$\alpha \equiv f(2) = \frac{d_2}{d_1 + d_2}$$

and

$$1 - \alpha \equiv f(1) = \frac{d_1}{d_1 + d_2}$$

Using the PHREQ algorithm all points of experimental spectra are analyzed and the solutions obtained at any different ν_1, ν_2, and ν_n are averaged.

Properties of the algorithm

Analogously to the testing of the various factors affecting the accuracy of solution obtained by the SIMS algorithm, we carried out a series of decompositions for simulated spectra (the sums of two log-normal curves with various positions of maxima and relative amplitudes) by the PHREQ method. Fig. 4 shows the dependencies for two-component decomposition $\Delta\lambda$ values (mean absolute differ-

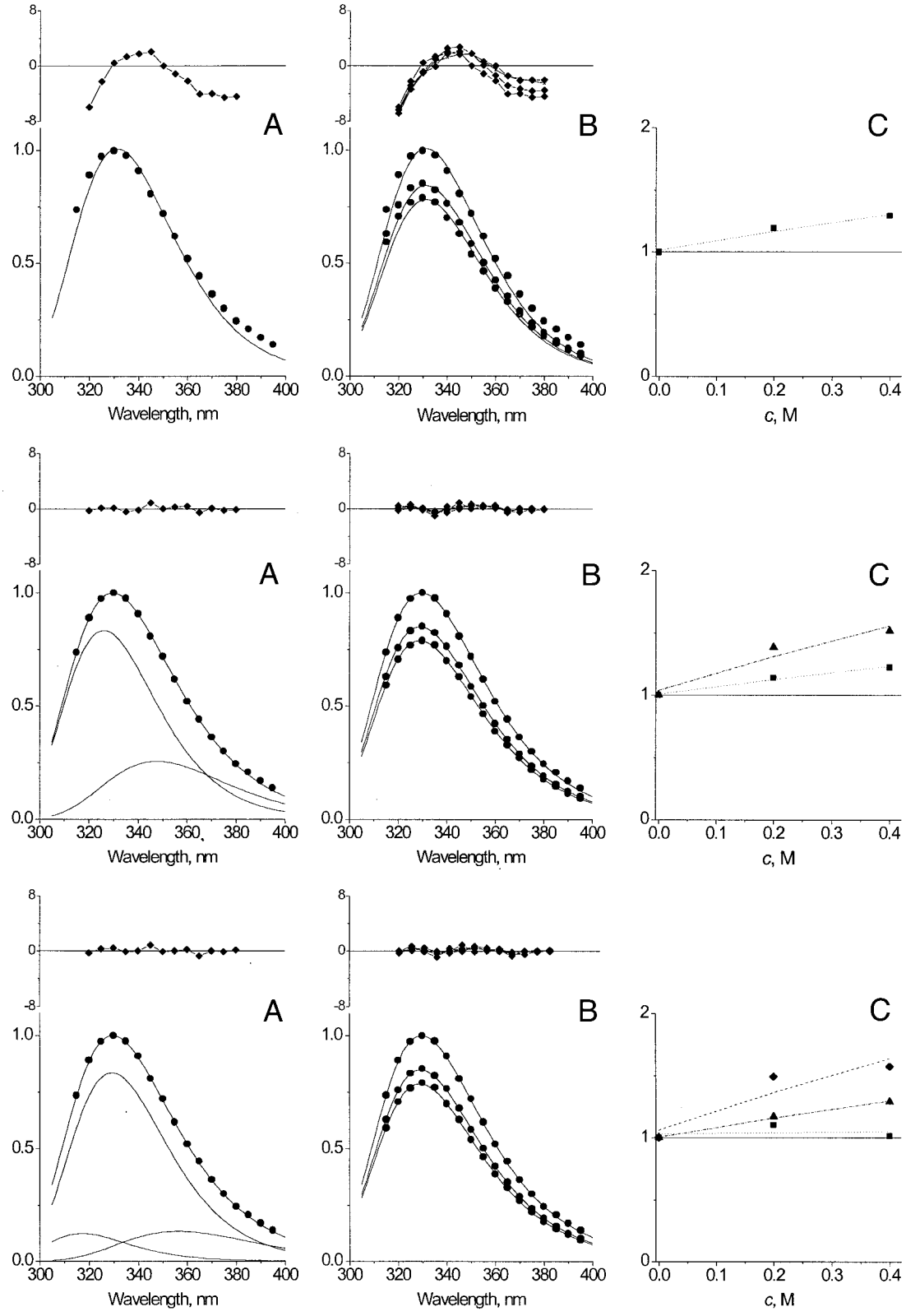

FIGURE 5 Decomposition of tryptophan fluorescence spectra of annexin VI measured at different concentrations of KI (0, 0.2, and 0.4 M) by one-, two- and three-components by the SIMS algorithm. Panels *A* represent the experimental spectra measured without fluorescence quencher (*points*) and calculated theoretical spectra, which are the sum of log-normal components (*curves*), and root-mean-square deviations between two kinds of spectra. Panels *B* represent the experimental spectra at different concentrations of KI (*points*) and calculated theoretical spectra (*curves*), and root-mean-square deviations between two kinds of spectra. Panels *C* represent the Stern-Volmer plots for each calculated component.

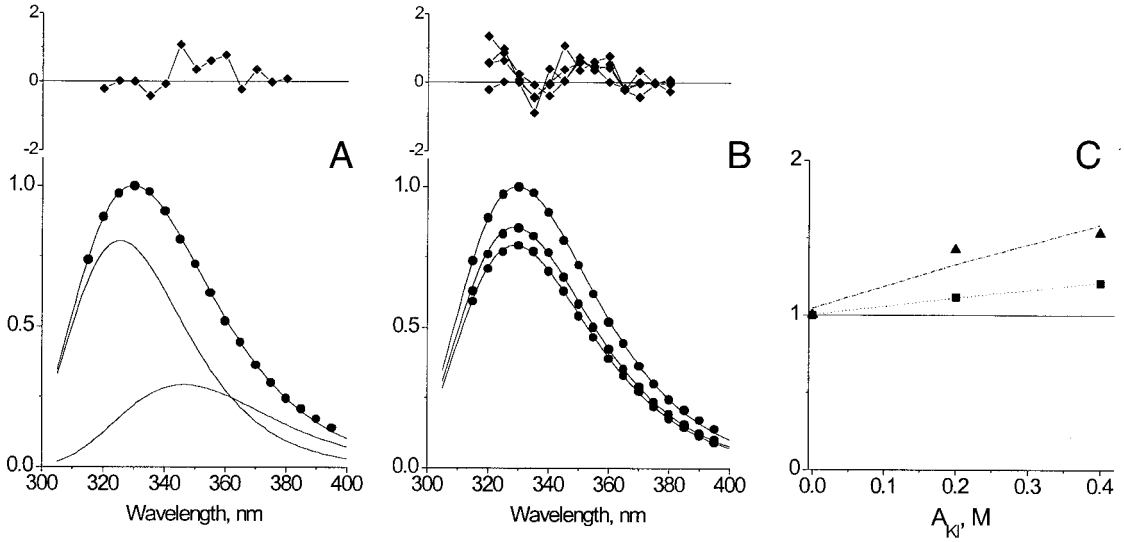

FIGURE 6 Decomposition of tryptophan fluorescence spectra of annexin VI measured at different concentrations of KI (0, 0.2, and 0.4 M) by two components by the PHREQ algorithm. Panel *A* represents the experimental spectra without quencher (*points*) and calculated theoretical spectra, which are the sum of two log-normal components (*curves*), and root-mean-square deviations between two kinds of spectra. Panel *B* represents the experimental spectra at different concentrations of KI (points) and calculated theoretical spectra (*curves*), and root-mean-square deviations between two kinds of spectra. Panel *C* represents the Stern-Volmer plots for two calculated components.

ence between positions of component maxima (in nanometers) that were preset in simulation and those obtained after the decomposition) and the amplitude of randomly introduced noise, *S*% (panel *A*); the number of spectra (*N*) with various "quencher concentrations" used (panel *B*); the distance between the component spectral maxima, $\Delta\lambda_{max}$, nm (panel *C*); the ratio of Stern-Volmer constants of quenching for the components, *K*2/*K*1 (panel *D*); the number of registered points in each spectrum, *M* (panel *E*); and the relative contribution of one component in the total spectrum, as an maximal intensity ratio *I*1/(*I*1 + *I*2) (panel *F*). The noise was introduced as random equally probable positive and negative deviations with amplitude from 0 to preset *S*% (in percent to the maximal spectrum amplitude). The maximal "quencher concentration" was chosen so that it would reduce the amplitude of a total spectrum approximately by half. Varying one parameter, the others were held constant at the following standard values: $S\% = 0.6\%$; $N = 5$; $M = 15$; $\Delta\lambda_{max} = 10$ nm; $K1 = 0.1\ M^{-1}$; $K2 = 3.0\ M^{-1}$; $I1 = I2$. As well as the SIMS algorithm, the PHREQ provides an acceptable accuracy ($\Delta\lambda < 1$ nm) of decomposition at real conditions.

AN EXAMPLE OF DECOMPOSITION OF EXPERIMENTAL SPECTRA BY SIMS AND PHREQ ALGORITHMS

As an example, we present here results of decomposition of tryptophan fluorescence spectra of annexin VI measured at three different concentrations of quencher KI (0.0, 0.2, and 0.4 M) by SIMS and PHREQ algorithms (Table 1 and Figs. 5 and 6). From the mono-, bi-, and tri-component decompositions obtained by SIMS, the bi-component solution was chosen as the most reliable because it has the least value of discriminant D_s (0.757), i.e., corresponding to the best fit of experimental spectra by theoretical curves. Table 2 summarizes the results of two-component decompositions obtained using both SIMS and PHREQ methods. The fitting algorithm SIMS and the analytical one PHREQ gave very similar results: obtained maximum positions λ_m differ within 1.2 nm, the contributions *S* of components differ within 6.3%, and the relative Stern-Volmer quenching constants $K_{s\nu}^{rel}$ within 0.5%. In the next papers of this series it will be demonstrated that the same values of maximum positions of two spectral components (~325 and 347 nm) will be pre-

TABLE 2 The summary of the decomposition of tryptophan fluorescence spectra of annexin VI into two log-normal components by the SIMS and PHREQ methods

	SIMS			PHREQ		
Comp.	λ_m (nm)	*S* (%)	$K_{s\nu}^{rel}$	λ_m (nm)	*S* (%)	$K_{s\nu}^{rel}$
1	326.0 ± 0.5	74.0 ± 1.0	5.6 ± 0.5	325.5 ± 2.9	67.7 ± 0.5	5.1 ± 0.1
2	347.7 ± 0.5	26.0 ± 2.0	12.9 ± 3.3	346.5 ± 3.1	32.3 ± 1.0	13.3 ± 4.1

dicted based on the analyses of the physical and structural parameters of microenvironments of two tryptophan residues (W192 and W343) in the crystal structure of annexin VI.

The authors are thankful to Dr. Andrzej Sobota for preparation of annexin VI and to Drs. V. I. Emelyanenko and O. A. Andreev for fruitful discussion. Moreover, we are indebted to Drs. D. B. Veprintsev and D. S. Rykunov for their help with computer maintenance.

This work was supported in part by Grants 95-04-12935, 97-04-49449, and 00-04-48127 from the Russia Foundation of Basic Research.

REFERENCES

Abornev, S. M. 1993. Decomposition of composite tryptophan-fluorescence spectra of proteins. Candidate of Sciences (Ph.D.) thesis. The Institute of Theoretical and Experimental Biophysics of the Russia Academy of Science.

Abornev, S. M., and E. A. Burstein. 1992. Resolution of protein tryptophan fluorescence spectra into elementary components *Molec. Biol. (Moscow).* 26:1350–1361 [In Russian; English translation].

Akseenko, V. M., T. N. Shumskaya, V. A. Slapochnikova, and N. V. Shein. 1989. Quantitative analysis of multicomponent systems based on the mathematical treatment of vibrational spectra. *Zhurnal Prikladnoj Spektroskopii (J. Appl. Spectrosc., Minsk).* 51:306–311. [In Russian; English translation].

Antipova-Korotaeva, I. I., and N. N. Kazanova. 1971. Mathematical decomposition of composite spectral contours into components with partially known parameters. *Zhurnal Prikladnoj Spektroskopii (J. Appl. Spectrosc., Minsk).* 14:1093–1096. [In Russian; English translation].

Bandorowicz, J., S. Pikula, and A. Sobota. 1992. Annexins IV (p32) and VI (p68) interact with erythrocyte membrane in a calcium-dependent manner. *Biochim. Biophys. Acta.* 1105:201–206.

Bukolova-Orlova, T. G., E. A. Burstein, and L. Ya. Yukelson. 1974. Fluorescence of neurotoxins from Middle-Asian cobra venom. *Biochim. Biophys. Acta.* 342:275–280.

Burstein, E. A. 1968. Quenching of protein fluorescence. I. Principles of the method. Solutions of tryptophan, tyrosine and denatured proteins. *Biophysics (Moscow).* 13:433–442 [In Russian; English translation].

Burstein, E. A. 1976. Luminescence of protein chromophores. Model studies. *In* Advances in Science and Technology (Itogi Nauki i Tekhniki), ser. Biophysics, Vol. 6. VINITI, Moscow. [In Russian].

Burstein, E. A. 1977. Intrinsic protein luminescence. The nature and application. *In* Advances in Science and Technology (Itogi Nauki i Tekhniki), ser. Biophysics, Vol. 7. VINITI, Moscow. [In Russian].

Burstein, E. A., and V. I. Emelyanenko. 1996. Log-normal description of fluorescence spectra of organic fluorophores. *Photochem. Photobiol.* 64:316–320.

Burstein, E. A., N. S. Vedenkina, and M. N. Ivkova. 1973. Fluorescence and the location of tryptophan residues in protein molecules. *Photochem. Photobiol.* 18:263–279.

Demchenko, A. P. 1986. Ultraviolet Spectroscopy of Proteins. Springer-Verlag. Berlin.

Hodgman, C. D., R. C. Weast, and S. M. Selby, Editors. 1955. Handbook of Chemistry and Physics, Vol. 2, 37th Ed. Chemical Rubber Publ. Co., Cleveland, OH.

Kaplanas, R. I., T. G. Bukolova, and E. A. Burstein. 1975. Fluorescence of β-lactoglobulin AB in various physico-chemical conditions. II. Denaturation by urea and organic solvents. *Mol. Biol. (Moscow).* 9:795–804. [In Russian; English translation].

Lakowicz, J. R. 1983. Principles of Fluorescence Spectroscopy. Plenum Press, New York.

Lakowicz, J. R. 1999. Principles of Fluorescence Spectroscopy, 2nd Ed. Plenum Press, New York.

Lehrer, S. S. 1971. Solute perturbation of protein fluorescence. The quenching of the tryptophan fluorescence of model compounds and of lysozyme by iodide ion. *Biochemistry.* 10:3254–3263.

Lehrer, S. S., and P. C. Leavis. 1978. Solute quenching of protein fluorescence. *Methods Enzymol.* 49G:222–236.

Metzler, C. M., A. E. Cahill, S. Petty, D. E. Metzler, and L. Lang. 1985. The widespread applicability of log-normal curves for the description of absorption spectra. *Appl. Spectrosc.* 39:333–339.

Metzler, D. E., C. Harris, Yang In-Yu, D. Siano, and J. A. Thomson. 1972. Band-shape analysis and display of fine structure in protein spectra: a new approach to perturbation spectroscopy. *Biochim. Biophys. Res. Commun.* 46:1588–1597.

Metzler, C. M., R. Viswanath, and D. E. Metzler. 1991. Equilibria and absorption spectra of tryptophanase. *J. Biol. Chem.* 266:9374–9381.

Morozov, Yu. V., and N. P. Bazhulina. 1989. Electronic structure, spectroscopy, and reactivity of molecules. Nauka, Moscow [In Russian].

Permyakov, E. A., V. V. Yarmolenko, V. I. Emelyanenko, E. A. Burstein, Ch. Gerday, and J. Closset. 1980a. Study of binding of calcium ions to whiting parvalbumin using changes of protein intrinsic fluorescence parameters. *Biophysics (Moscow).* 25:417–422. [In Russian; English translation].

Permyakov, E. A., V. V. Yarmolenko, V. I. Emelyanenko, E. A. Burstein, J. Closset, and Ch. Gerday. 1980b. Fluorescence studies of the calcium binding to whiting (*Gadus merlangus*) parvalbumin. *Eur. J. Biochem.* 109:307–315.

Siano, D. B., and D. E. Metzler. 1969. Band shapes of the electronic spectra of complex molecules. *J. Chem. Phys.* 51:1856–1961.

Sobota, A., J. Bandorowicz, A. Jezierski, and A. F. Sikorski. 1993. The effect of annexins IV and VI on the fluidity of phosphatidylserine/phosphatidylcholine bilayers studied with the use of 5-deoxylstearate spin label. *FEBS Lett.* 315:178–182.

Tikhonov, A. N., and V. Ya. Arsenin. 1986. Methods of solving incorrect problems. Nauka, Moscow. [in Russian].