Chapter 10

Global Analysis of Time-Resolved Fluorescence Data

Anatoli V. Digris, Eugene G. Novikov, Victor V. Skakun, and Vladimir V. Apanasovich

Abstract

In this chapter, we describe the global analysis approach for processing time-resolved fluorescence spectroscopy data of molecules in the condensed phase. Combining simultaneous analysis of data measured under different experimental conditions (spatial coordinates, temperature, concentration, emission wavelength, excitation intensity, etc.) with the fitting strategy, enabling parameter linkage and thus decreasing the total amount of estimated variables, makes global analysis more robust and more consistent compared to a sequential fit of single experimental data. We consider the main stages of the global analysis approach and provide some details that are important for its practical implementation. The application of the global approach to the analysis of time-resolved fluorescence anisotropy is demonstrated on experimental data of (enhanced) green fluorescent protein in aqueous solution.

Key words Global analysis, TCSPC, Time domain, Frequency domain, Deconvolution

1 Introduction

Structural and dynamical properties of complex, fluorescent biological and chemical systems can be successfully studied using time- and frequency-domain methods of time-resolved fluorescence spectroscopy [1]. The most commonly used time-domain experimental method is time-correlated single-photon counting (TCSPC). In this method the fluorescent sample is excited by a series of short laser pulses of fixed repetition rate and the fluorescence decay is measured and stored in computer memory [1]. Quantitative information about the explored molecular system is obtained by fitting the model-generated decay curve to the experimental sample fluorescence decay. This gives the values of model parameters that characterize the sample under study. In the time-resolved frequency-domain approach [1] the sample is excited by intensity-modulated light at different modulation frequencies. The emission signal of the molecular system is then demodulated and

Yves Engelborghs and Antonie J.W.G. Visser (eds.), *Fluorescence Spectroscopy and Microscopy: Methods and Protocols*, Methods in Molecular Biology, vol. 1076, DOI 10.1007/978-1-62703-649-8_10,

phase-shifted with respect to the modulated excitation light. The dependencies of degree of modulation and phase shift on the modulation frequency are measured and analyzed by the appropriate model to estimate the values of fit parameters characterizing the sample under study.

The experimental methods of time-resolved fluorescence spectroscopy are frequently applied to measure fluorescent molecules in the condensed phase under different experimental conditions such as different temperatures, emission wavelengths, and concentrations. Since some parameters characterizing the investigated molecular system may be independent on the experimental conditions, these sets of data can be simultaneously analyzed using a global analysis approach [2, 3]. The fit parameters that are the same in different experiments are linked and thus kept equal for the corresponding measured curves participating in the fit. The other parameters that change with experimental conditions are fitted independently. As the overall number of fitted parameters decreases, whereas the amount of experimental data remains the same, the global analysis approach has proved to be more robust and more consistent compared to the separate analysis of single curves.

The abilities of the time-resolved fluorescence spectroscopy are significantly extended with introduction of the fluorescence lifetime imaging microscopy (FLIM) technique [1, 4, 5]. This technique uses a wide-field or a confocal microscope to perform the time- or frequency-domain measurements in various spatial points of the sample, thus resolving the immediate molecular environment and the state of a fluorescent molecule in a microscopic image. Another example is the use of genetically encoded fluorescent proteins such as green fluorescent protein (GFP), which can be attached to the protein of interest and expressed in living cells [6]. When two different proteins are labeled with two different variants of GFP, for instance cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP) in living cells, the distance between these proteins when interacting can be mapped in a living cell by using FLIM combined with Förster resonance energy transfer (FRET) [7]. In the following we use the term donor for a protein that is tagged with CFP and the term acceptor for the other protein that is tagged with YFP. In other biosensors donor and acceptor are both linked to a common protein that changes conformation when a ligand is bound and thus alters the donor–acceptor distance. When FRET occurs (i.e., when donor and acceptor are a few nanometers apart) the reaction between excited donor and acceptor in the ground state will decrease the fluorescence lifetime of the donor in various points of the image. The appearance of double-exponential fluorescence intensity decays is frequently encountered because of the presence of a mixture of FRET-active and FRET-inactive donor molecules [8, 9]. To detect the spatially

dependent relative contribution of donors participating in FRET global analysis may be applied. The fluorescence intensity decays measured in different points of the image are analyzed globally using double-exponential models. The corresponding fluorescence lifetimes for all space points are linked whereas the pre-exponential factors for each exponent are fitted independently (for an example *see* ref. 10).

In this chapter we present the main stages required to perform the global analysis of time-resolved fluorescence spectroscopy data. We focus on some practical details important for its implementation. As illustration we show the application of global analysis for fitting the time-resolved fluorescence anisotropy of enhanced GFP from *Aequorea victoria* (abbreviated as EGFP) in aqueous, buffered solution.

2 Materials

2.1 EGFP

The purification of EGFP has been described previously [11, 12]. Protein solutions were prepared in 20 mM Tris buffer (pH 7.0).

2.2 Time-Resolved Polarized Fluorescence Measurements

Time-resolved fluorescence measurements (20 °C) were carried out using a mode-locked laser pumped by a continuous-wave laser for excitation and TCSPC as the detection technique, as extensively described previously [8, 11, 13]. The samples were excited with plane-polarized light pulses (0.2 ps FWHM) at an excitation frequency of 3.8 MHz and both parallel- and perpendicular-polarized fluorescence intensities were detected.

Experiments with EGFP were conducted at 470-nm excitation and fluorescence detection via the combination of 515-nm cutoff and 512-nm interference filters.

The reference compound for time-resolved fluorescence decay experiments with 470-nm excitation and 512-nm detection was erythrosine B in aqueous solution having a fluorescence lifetime of 85 ps. In all experiments the binning time was 5 ps and 4,096 time channels were used.

3 Methods

3.1 General Global Analysis Algorithm

The global analysis method is based on the simultaneous processing of more than one separately measured data set of a fluorescent species in the condensed phase (*see* Fig. 1). Such data sets are usually obtained by measuring samples under different experimental conditions (temperature, emission wavelength, excitation intensity, etc.). The application of global analysis becomes meaningful if at least one of the unknown parameters can be assumed to be the same for different measurements. Therefore these parameters, even

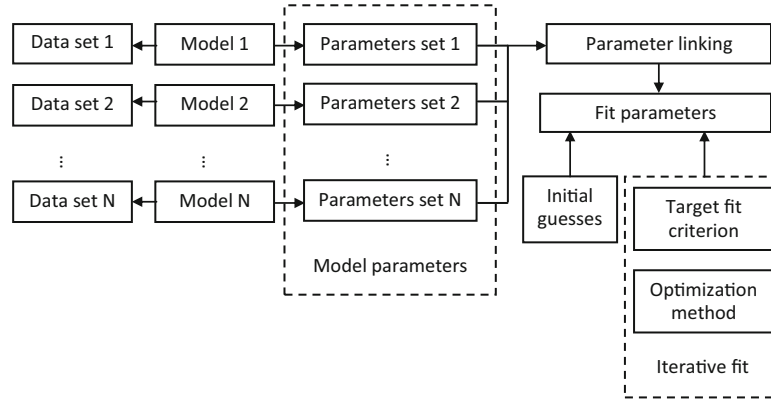

Fig. 1 General scheme of global analysis

though being assigned to different experimental data sets, are united into one parameter in the fitting procedure. The global analysis is more robust and consistent than the separate analysis of individual data sets, because it exploits measured information available from several data sets in a more efficient way by reducing the total number of fit parameters.

The global analysis method can be applied to fit the time-resolved fluorescence data obtained from both time-domain and frequency-domain measurements and includes several consecutive steps.

1. Collect the data to be analyzed and prepare the data sets. The term "data set" refers to the set of information that can be analyzed by a single mathematical model with a well-defined set of parameters. The data set contains the measured, possibly multidimensional, data from the sample (e.g., fluorescence intensity decay, phase shift, demodulation). In addition, either prior known or experimentally obtained information required for the analysis can also be included.

   For time-domain measurements, each data set contains the temporal information (either as time points or as time step and the number of time points), sample fluorescence intensity decay, and either the instrumental response function (IRF, obtained with a scattering solution) or the fluorescence intensity decay of a reference compound with known fluorescence lifetime. The latter is used for deconvolution of measured data (*see* **step 2**). A background sample such as solvent alone, obtained at the same conditions as the investigated sample, accounting for impurities or Raman scattering in the sample may also be added to the data set.

   For frequency-domain measurements the data set includes the sequence of modulation frequencies and values for phase shift and degree of modulation obtained for each frequency.

Optionally, the data set can include the list of external parameters reflecting the conditions of the measurement (e.g., temperature, excitation/emission wavelength, polarization angle). As these parameters are set by the user during the measurement and are typically known precisely, they can be used by some models as preset (fixed) values to calculate the theoretical curve.

2. Select the model for each data set created in **step 1**. Depending on the model, an algorithm should be developed and implemented generating the theoretical curve that is fitted to the measured one.

   The analytical models that are applied for fitting the fluorescence intensity decays are based on the following general equations.

   If an IRF is used for deconvolution, the basic model for calculating the theoretical decay $f^{\mathrm{th}}(t)$ takes the form

$$f^{\mathrm{th}}(t) = (g(t+\delta) - b) \otimes I(t, \vec{A}) + \gamma B(t) + c + n \times (g(t+\delta) - b), \quad (1)$$

   where $g(t+\delta)$ denotes the instrument response function with a time shift of δ relative to the emission response; $B(t)$ is the measured intensity of time-dependent background emission contributing with a relative weight of γ; b and c are constant values for dark noise and time-uncorrelated background photons in g and f, respectively; n is the scattered light coefficient and $I(t, \vec{A})$ represents a model function with the vector of fit parameters $\vec{A}$.

   For deconvolution using a single-lifetime reference compound the basic model takes the form

$$f^{\mathrm{th}}(t) = (f_{\mathrm{ref}}(t) - b) \otimes \left(\delta(t) I(0, \vec{A}) + \frac{I(t, \vec{A})}{\tau_{\mathrm{ref}}} + \frac{\mathrm{d}I(t, \vec{A})}{\mathrm{d}t} \right) + \gamma B(t) + c, \quad (2)$$

   where $f_{\mathrm{ref}}(t)$ denotes the measured one-exponential reference compound fluorescence intensity decay with decay time τ_{ref} and $\delta(t)$—Dirac delta-function.

   To complete the model, the function $I(t, \vec{A})$ should be defined. It reflects typical physical processes taking place in the sample under study. A number of models for $I(t, \vec{A})$ have been published [1, 14, 15]. These include simple models, such as a sum of exponentials or stretched exponential, as well as more complicated ones used in the compartmental formalism. One of the most widely used equations for $I(t, \vec{A})$ is the sum of exponential terms:

$$I(t, \vec{A}) = \sum\nolimits_{j=1}^{M} p_j \mathrm{e}^{-t/\tau_j}, \quad (3)$$

where p_j and $\tau_j (j = 1, \ldots, M)$ are, respectively, the amplitudes and decay times of corresponding exponents and M is the number of exponents in the sum. The multi-exponential model can be used to fit the fluorescence intensity decays obtained from the mixture of noninteracting species, donor molecules when resonance energy transfer process takes place, etc. [1].

The analytical models for fitting the data obtained from frequency-domain measurements are based on the sine $N(\omega, \vec{A})$ and cosine $D(\omega, \vec{A})$ Fourier transforms of the function $I(t, \vec{A})$:

$$N(\omega, \vec{A}) = \frac{\int_0^\infty I(t, \vec{A}) \sin(\omega t) \mathrm{d}t}{\int_0^\infty I(t, \vec{A}) \mathrm{d}t},$$

$$D(\omega, \vec{A}) = \frac{\int_0^\infty I(t, \vec{A}) \cos(\omega t) \mathrm{d}t}{\int_0^\infty I(t, \vec{A}) \mathrm{d}t},$$

where ω is the angular modulation frequency of the excitation light. The general model equations for phase shift $\varphi(\omega, \vec{A})$ and degree of modulation $m(\omega, \vec{A})$ are

$$\begin{aligned} \varphi(\omega, \vec{A}) &= \mathrm{arctg}\Big(N(\omega, \vec{A}) / D(\omega, \vec{A})\Big), \\ m(\omega, \vec{A}) &= \sqrt{N^2(\omega, \vec{A}) + D^2(\omega, \vec{A})}. \end{aligned} \tag{4}$$

If $I(t, \vec{A})$ is a sum of exponents and defined by Eq. 3, the models for $\varphi(\omega, \vec{A})$ and $m(\omega, \vec{A})$ become

$$\varphi(\omega, \vec{A}) = \mathrm{arctg}\left(\frac{\sum_{j=1}^{M} \frac{\omega p_j \tau_j^2}{1+\omega^2 \tau_j^2}}{\sum_{j=1}^{M} \frac{p_j \tau_j}{1+\omega^2 \tau_j^2}}\right), \tag{5}$$

$$m(\omega, \vec{A}) = \frac{1}{\sum_{j=1}^{M} p_j \tau_j} \sqrt{\left(\sum_{j=1}^{M} \frac{\omega p_j \tau_j^2}{1+\omega^2 \tau_j^2}\right)^2 + \left(\sum_{j=1}^{M} \frac{p_j \tau_j}{1+\omega^2 \tau_j^2}\right)^2}. \tag{6}$$

Note that parameters $p_j (j = 1, \ldots, M)$ in Eqs. 5 and 6 define the contribution of corresponding components in the sum, whereas amplitudes in Eq. 3 are also responsible for fitting the amplitude of the fluorescence decay.

3. Link parameters: Every model selected in **step 2** contains the set of fit parameters. The term "fit parameter" refers to the value of interest that should be estimated during the analysis. For each fit parameter the range of admissible values can be set by defining the minimum and maximum constraints. The number of fit parameters and their physical interpretation depend

on the selected model. The models based on Eqs. 1 and 2 have experiment-specific fit parameters (δ, b, γ, c n for Eq. 1 and b, γ, c, τ_{ref} for Eq. 2) and sample-specific parameters defined by $I(t, \vec{A})$. For the multi-exponential model defined by Eq. 3 for time-domain measurements and by Eqs. 5 and 6 for frequency-domain measurements the model-specific parameters are p_j and $\tau_j (j = 1, \ldots, M)$. The fit parameters from all models make up the common list of fit parameters. If for certain parameters the value is known or has been already determined from other, independent experiments, its value can be fixed so that it will remain unchanged during the fit.

The key feature of the global analysis is the parameter linkage. The linking procedure can be introduced as creating a group of fit parameters that must be kept equal during the fit. The settings (value, minimum, maximum, etc.) of fit parameters included in such group always match the appropriate settings of the group itself. In most cases one parameter group contains the fit parameters that are supposed to have the same value. The linked fit parameters are excluded from the common fit parameter list, whereas the created fit parameter group is added to this common list as a new fit parameter. Since the link procedure replaces several fit parameters in the common list by a group, the total number of parameters to be estimated is reduced. For the analysis to be global, at least one fit parameter group should be created. Generally the number of such global groups is more than one.

4. Initial guess generation: Initial guesses (IG) are required to make the forthcoming iterative fitting procedure fast and reliable. Therefore the IG should be reasonably close to the optimal values of fit parameters that correspond to the best fit of experimental data. If IG are chosen randomly and are located relatively far from the optimal values of fit parameters, the iterative algorithm may converge rather slowly and in some cases may stop untimely, thus leading to biased parameter estimates. To automatically generate the initial guesses a fast non-iterative algorithm should be applied. Such algorithms always take into account the particularity of the models selected for the fit and therefore are not available for any combination of models within global analysis. If such algorithms exist, it is always recommended to use them in combination with the iterative fit routine.

   In many practical applications such as FLIM the multi-exponential model can be applied for the global analysis of fluorescence intensity decays [2, 3]. The model is defined by Eqs. 1 and 3, for which the corresponding decay times are assumed to be equal for different data sets and therefore they

should be linked, whereas the amplitudes are allowed to be different. If instrumental distortions in Eq. 1 are negligibly small (parameters δ, b, γ, c n equal to 0), the phase plane method can be used to generate initial guesses [16, 17]. It is based on the equivalence of the model given by Eq. 1 with $I(t, \vec{A})$ defined by Eq. 3 to the following integral equation:

$$\sum_{j=0}^{M} c_j F_{ij}(t) = \sum_{j=1}^{M} \rho_{ij} G_{ij}(t), \tag{7}$$

where

$$\begin{aligned}
F_{i0}(t) &= f_i(t), \\
F_{ij}(t) &= f_i(t) \otimes \varphi_j(t) = \int_0^t \frac{f_i(x)(t-x)^{j-1}}{(j-1)!}\,\mathrm{d}x, \\
G_{ij}(x) &= g_i(t) \otimes \varphi_j(t) = \int_0^t \frac{g_i(x)(t-x)^{j-1}}{(j-1)!}\,\mathrm{d}x, \\
\varphi_j(t) &= \frac{t^{j-1}}{(j-1)!},
\end{aligned}$$

$i = 1, \ldots, m$, where m is the number of data sets. The coefficients c_k are functions of the decay times only. As the decay times are linked, c_j do not depend on the data set index i. Coefficients ρ_{ij} depend on both decay times and amplitudes. These coefficients are different for each data set because exponential amplitudes are not linked. Taking into account Eq. 7 the algorithm of the phase plane method can be described as follows.

(a) Calculate integral functions $F_{ij}(t)$ and $G_{ij}(t)$. Since in practice $f_i(t)$ and $g_i(t)$ are obtained on discrete set of time points $t_k (k = 1, \ldots, N)$ the integration is replaced by summation and discrete convolution of $f_i(t_k)$ and $g_i(t_k)$ with $\varphi_j(t_k)$ is calculated giving $F_{ij}(t_k)$ and $G_{ij}(t_k)$.

(b) For each data set construct the following matrices:

$$\overline{F_i} = \begin{pmatrix} F_{i0}(t_1) & \cdots & F_{i0}(t_N) \\ F_{i1}(t_1) & \cdots & F_{i1}(t_N) \\ \vdots & \ddots & \vdots \\ F_{iM}(t_1) & \ldots & F_{iM}(t_N) \end{pmatrix}; \quad \overline{G_i} = \begin{pmatrix} G_{i1}(t_1) & \cdots & G_{i1}(t_N) \\ \vdots & \ddots & \vdots \\ G_{iM}(t_1) & \ldots & G_{iM}(t_N) \end{pmatrix};$$

$$\overline{W_i} = \begin{pmatrix} 1/f_i(t_1) & 0 & \cdots & 0 \\ 0 & 1/f_i(t_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/f_i(t_N) \end{pmatrix}.$$

(c) Obtain the vector of coefficients c_j by the linear least-squares method. To do this prepare for each data set the matrices:

$$\overline{R_i} = \overline{F_i}\,\overline{W_i}\,\overline{G_i}'\left(\overline{G}\;\overline{W}\;\overline{G_i}'\right)^{-1}; \quad \overline{D_i} = \overline{F_i} - \overline{R_i}\;\overline{G_i}; \qquad \bar{S} = \sum\nolimits_{i=0}^{m} \overline{D_i}\;\overline{D_i}'. \tag{8}$$

Find the coefficients c_j by solving the system of linear equations. This solution in matrix form can be represented as follows:

$$\bar{C} = \bar{A}/\bar{B}, \tag{9}$$

where $\bar{C}$ is a vector of coefficients $c_j (j = 1, \ldots, M)$; matrices $\bar{A}$ and $\bar{B}$ are obtained from matrix $\bar{S}$:

$$\bar{A} = \begin{pmatrix} S_{22} \cdots \\ \vdots \ddots \\ S_{(M+1)2} \cdots \end{pmatrix}; \quad \bar{B} = -\left(S_{21} \vdots S_{(M+1)1}\right).$$

(d) Calculate estimations for decay times as roots of the following polynomial:

$$\sum_{j=0}^{M} (-1)^j c_j \tau^{j-M} = 0, \tag{10}$$

where $c_0 = 1$; $c_j (j = 1, \ldots, M)$ are taken from vector $\bar{C}$ obtained above from Eq. 9. In most cases the roots of polynomial in Eq. 10 are calculated numerically [18].

(e) For each data set calculate the amplitudes of exponents $p_{ij} (i = 1, \ldots, m;\ j = 1, \ldots, M)$ from the following equation:

$$p_{ij} = \frac{\sum_{k=0}^{M-1} (-1)^k \rho_{i(k+1)} \tau_j^k}{\prod_{m=1, m \neq j}^{M} \left(1 - \frac{\tau_j}{\tau_m}\right)}, \tag{11}$$

where decay times τ_j are obtained by solution of Eq. 10; coefficients ρ_{ij} are elements of vector $\overline{r_i}$ calculated for each data set as follows:

$$\overline{r_i} = \left(1\overline{C}\right) R_i, \tag{12}$$

where $\overline{C}$ is calculated from Eq. 9 and $\overline{R_i}$ obtained from Eq. 8.

5. Iterative fit: The iterative fit procedure is used for getting final estimations of fit parameters that provide the best match of

theoretical curves generated by the particular model and the measured data. To organize the iterative fit the target fit criterion (TFC) and optimization method are required.

The TFC is used to quantify the distance between the theoretical curves generated by the model and the corresponding measured data. The choice of the appropriate TFC is based on the type of statistical noise in the experimental data. For global analysis of time-domain data three TFCs are often used. All these TFCs are based on the maximum likelihood approach [19–22].

When the number of detected photons in each time channel is relatively high, the noise distribution can be well approximated by Gaussian statistics [19] and the TFC is defined by the chi-square equation:

$$\chi^2_G = \frac{1}{n-s-1} \sum_{i=1}^{m} \sum_{k=1}^{N_i} \frac{\left(x_{ik} - F^{\mathrm{th}}_{ik}(\vec{A})\right)^2}{\sigma^2_{ik}}, \tag{13}$$

where $v = n - s - 1$ is the number of degrees of freedom; n is the total number of data points from all data sets that participate in the calculation of sums in Eq. 13; s is the total number of fit parameters (non-linked parameters from all models and all parameter groups) available for optimization; x_{ik}, $F^{\mathrm{th}}_{ik}(\vec{A})$, and σ^2_{ik} are, respectively, experimental value, theoretical value, and variance in the k-th point of the i-th data set. The application of TFC based on Eq. 13 leads to the iterative fit procedure, which is well known as the weighted least-squares method [1, 19, 21]. Some practical remarks related to application of Eq. 13 are given in **Note 1**.

For some time-domain measurements, such as single-molecule fluorescence detection and FLIM, the number of photons registered in each time channel of fluorescence decay is relatively small. At the same time the number of time channels in each decay is large, so the probability to detect a photon within a separate time channel is small. Application of the TFC defined by Eq. 13 for such data leads to getting biased estimates of fit parameters [20]. For such cases the statistical noise is better approximated by the Poisson distribution. Under this assumption the TFC for global analysis takes the form

$$\chi^2_{\mathrm{P}} = \frac{2}{n-s-1} \sum_{i=1}^{m} \sum_{k=1}^{N_i} x_{ik} \ln\left[\frac{x_{ik}}{F^{\mathrm{th}}_{ik}(\vec{A})}\right] - x_{ik} + F^{\mathrm{th}}_{ik}(\vec{A}). \tag{14}$$

The scheme of the single-photon counting technique also allows applying the multinomial distribution for description of statistical noise in the measured fluorescence decays. At each registration cycle of the fluorescence decay measurement no

more than one photon is detected in one of the N channels. Registering the photon in a certain time channel is one of N possible outcomes occurring each time when the photon is detected. Such experimental scheme leads to the multinomial distribution of statistical noise in the measured data. In this case the TFC for global analysis is defined by the following equation:

$$\chi^2_{\mathrm{M}} = \frac{2}{n-s-1} \sum_{i=1}^{m} \sum_{k=1}^{N_i} x_{ik} \ln \left[\frac{x_{ik}}{F^{\mathrm{th}}_{ik}(\vec{A})} \right]. \tag{15}$$

While using Eq. 15 in practice the following condition should be ensured during each iteration of the fit:

$$\sum_{k=1}^{N_i} x_{ik} = \sum_{k=1}^{N_i} F^{\mathrm{th}}_{ik}(\vec{A}), \quad i = 1, \ldots, m. \tag{16}$$

The iterative fit procedure based on either Eq. 14 or Eq. 15 is known as the maximum likelihood estimation method [20–22].

The global analysis of frequency-domain data is performed by least-squared method with the TFC defined by Eq. 13 [3]. The variances σ^2_{ik} in this case are obtained experimentally.

The optimization method finds the estimations of fit parameters that match the minimum of the TFC, thus ensuring the best correspondence between measured and model-generated data. Each optimization method is based on the unique mathematical algorithm that finds the minimum of the TFC by calculating and processing its values for different points in the fit parameter space. One of the most popular classes of the optimization algorithms is based on gradient optimization. It requires calculating the derivatives of the TFC with respect to fit parameters. Gradient optimization methods demonstrate better performance in comparison with non-gradient methods, when the required derivatives are available. Although the derivatives of model functions can often be calculated numerically with sufficient accuracy, analytical derivatives are preferable ensuring higher numerical accuracy and higher processing speed. Gradient methods are not suitable when the theoretical function is obtained by the model that is based on Monte-Carlo simulations. In this case, the model-generated data are obtained by a stochastic simulation algorithm instead of an analytical function and thus contain statistical noise. Therefore, the derivatives required by the gradient method cannot be calculated analytically, whereas the application of numerical algorithms leads to significant errors.

One of the most frequently used gradient algorithms is the Marquardt–Levenberg optimization algorithm [19, 23]. This algorithm consists of the following steps.

(a) Set the initial guesses for fit parameters as a current point in parameter space $a_j^{\text{curr}} = a_j^{\text{in}}, \quad j = 1, \ldots, s$. Initialize the regularization value λ (for example $\lambda = 0.001$).

(b) Calculate the TFC χ^2_{curr} for the current point.

(c) Calculate matrices of the system of linear equations A and B. The elements of these matrices depend on the derivatives of the TFC with respect to the fit parameters and are defined as follows:

$$A_{ij} = -\frac{\partial^2 \chi^2}{\partial a_i \partial a_j}; B_j = -\frac{\partial \chi^2}{\partial a_j} \quad i = 1, \ldots, s; j = 1, \ldots, s. \tag{17}$$

All derivatives are calculated at the current point in the parameter space (at the beginning this is the point defined by initial guesses).

(d) Modify the diagonal elements of matrix A:

$$A_{ii} = (1 + \lambda) A_{ii}, \quad i = 1, \ldots, s. \tag{18}$$

(e) Solve the system of linear equations:

$$AX = B, \tag{19}$$

where X is a column vector that contains the fit parameter increments Δa_j for the current iteration.

(f) Calculate a new trial point in the fit parameter space:

$$a_j^{\text{tr}} = a_j^{\text{curr}} + \Delta a_j, \quad j = 1, \ldots, s. \tag{20}$$

For the new point calculate the value of the TFC, χ^2_{tr}. For some practical recommendations on this step *see* **Note 2**.

(g) If χ^2_{tr} χ^2_{curr} the new trial point in fit parameter space is accepted and is set as the current point, i.e., $a_j^{\text{curr}} = a_j^{tr}, \quad \chi^2_{\text{curr}} = \chi^2_{\text{tr}}$. The value of the method parameter λ is decreased: $\lambda = 0.1\lambda$. Alternatively, if $\chi^2_{\text{tr}} \geq \chi^2_{\text{curr}}$ the trial point is ignored and the value of method parameter λ is increased: $\lambda = 10\lambda$.

(h) The stop optimization conditions are checked. The first one is defined as follows:

$$\lambda > \lambda_{\max}. \tag{21}$$

This condition is used to stop iterations when the method cannot find the direction to the minimum of the TFC. In practice $\lambda_{\max} = 10^{10}$.

The second stop condition is used when the optimal point in fit parameter space has been successfully found with the required accuracy:

$$\frac{\left|\chi^2_{\text{tr}} - \chi^2_{\text{curr}}\right|}{\chi^2_{\text{curr}}} \leq \epsilon, \tag{22}$$

where ε is the small accuracy value that is set beforehand. In practice it can be equal to 10^{-5}.

When at least one of the stop conditions is satisfied, the iterations are interrupted and the current point defined by $a_j^{\text{curr}}, \quad j = 1, \ldots, s$ is returned as the result. Otherwise the method moves to **step c** and a new iteration starts.

6. Judging the quality of the fit: The final stage of the global analysis procedure consists in evaluation of the quality of the obtained results. For this purpose several fit quality criteria are used. The first is the final value of the TFC calculated by Eqs. 13–15. This value should be close to 1 for a good fit. Besides the TFC two graphical criteria, weighted residuals and autocorrelation function of weighted residuals, are widely used for judging the goodness of the fit. Both these criteria are built individually for each data set and are defined on the same argument range as the measured data.

The weighted residuals are used for visual inspection of the agreement between measured and model-generated data. The weighted residuals correspond to the type of the TFC used for the fit. In the case of chi-square TFC (Eq. 13) the residuals are calculated according to the following equation:

$$R_{ik} = \frac{x_{ik} - F_{ik}^{\text{th}}(\vec{A})}{\sigma_{ik}}, \quad i = 1, \ldots, m; \; k = 1, \ldots, N_i. \tag{23}$$

In the case of Poisson statistics (Eq. 14) the residuals are calculated as

$$R_{ik} = \text{sign}\left(x_{ik} - F_{ik}^{\text{th}}(\vec{A})\right) \times \sqrt{2\left(x_{ik} \ln\left(x_{ik}/F_{ik}^{\text{th}}(\vec{A})\right) + F_{ik}^{\text{th}}(\vec{A}) - x_{ik}\right)}, \tag{24}$$

where $i = 1, \ldots, m; \; k = 1, \ldots, N_i$ and function $\text{sign}(x)$ is equal to -1 in the case of negative x and 1 otherwise. For the multinomial statistics (TFC is defined by Eq. 15) the equation for the weighted residuals takes the form

$$R_{ik} = \text{sign}\left(x_{ik} - F_{ik}^{\text{th}}(\vec{A})\right)\left|2x_{ik} \ln\left(x_{ik}/F_{ik}^{\text{th}}(\vec{A})\right)\right|, \quad i = 1, \ldots, m; \; k = 1, \ldots, N_i. \tag{25}$$

For the successful fits the weighted residuals should be randomly distributed around 0.

The plot of the autocorrelation function of weighted residuals [1, 24] provides another valuable visual check of the quality of the fit. The autocorrelation function of the weighted residuals for i-th data set in the k-th point of argument $C_{ik}(i = 1, \ldots, m; \; k = 1, \ldots, N_i)$ is calculated by the following equation:

$$C_{ik} = \frac{1}{N_i - k + 1} \sum_{j=1}^{N_i - k + 1} R_{ij} R_{i(j+k-1)} \Bigg/ \left(\frac{1}{N_i} \sum_{j=1}^{N_i} R_{ij}^2 \right), \qquad (26)$$

where R_{ij} is the value of weighted residuals calculated for i-th data set in the j-th point of argument and N_i is the number of points for which weighted residuals are calculated. The plots of autocorrelation are usually made across half of the data points $(N_i/2)$. In many cases the visual inspection of the autocorrelation function is more sensitive for the goodness of the fit than a plot of weighted residuals. For successful fits the autocorrelation function of weighted residuals is randomly distributed around zero. Bad fits provide low-frequency periodicity in an autocorrelation plot that can easily be detected by eye. For additional details related to calculation of autocorrelation plots *see* **Note 3**.

Besides fit quality criteria described above additional fit quality criteria (standard deviation of χ^2, Durbin–Watson parameter, etc.) can be considered [24].

As the parameters are estimated from the measured, statistically distorted data, they are known with some error. To quantify the error in the estimated parameters, we use confidence intervals. The confidence interval is the range where the true value of the fit parameter is located with some confidential probability.

The confidence intervals for fit parameters can be found by the exhaustive search method [1, 25]. The advantage of this approach is that it takes into account possible correlations between parameters. The search procedure starts from the optimal value of examined parameter obtained after a successful global fit. The value of the selected fit parameter is shifted from the optimal value to the left or to the right (depending on the bound of the confidence interval that should be found) with the predefined step. After each move the examined parameter is fixed to the shifted value and other parameters are adjusted to find the minimum of the TFC. The new TFC is compared with the χ^2_{lim} level calculated from the statistical F-test as follows:

$$\chi^2_{\text{lim}} = \chi^2_{\text{min}} \left(1 + \frac{s}{n - s - 1} F(s,\, n - s - 1,\, 1 - \beta) \right), \qquad (27)$$

where χ^2_{min} is the minimum value of TFC obtained from the global analysis before calculation of confidence intervals, s is the total number of fit parameters available for optimization, $\nu = n - s - 1$ is the number of degrees of freedom, β is the confidential probability (0.67, 0.95, etc.), and $F(s, n - s - 1, 1 - \beta)$ is the F-statistic. The search procedure stops when the TFC value becomes equal (with some predefined accuracy) to the χ^2_{lim} and the corresponding shifted value of the examined fit parameter is taken as a bound of confidence interval.

3.2 Application of Global Analysis for Fitting Time-Resolved Fluorescence Anisotropy

The global analysis approach described in Subheading 3.1 can be applied for fitting time-resolved anisotropy. The algorithm of global analysis (*see* Subheading 3.1) adopted for this application consists of the following steps.

1. Prepare the data sets. The data sets for anisotropy analysis contain the fluorescence decays measured for different polarization angles Θ. In practice two data sets for parallel ($\Theta = 0$ degrees) and perpendicular ($\Theta = 90$ degrees) polarization components are used. Some practical recommendations related to preparing the data sets with the IRF for anisotropy analysis are given in **Note 4**.
2. Select the models. The model for global analysis of sample polarization components is based on Eqs. 1 or 2 where the function $I(t, \vec{A})$ is defined by the formula

$$I(t,\vec{A}) = H\left(\sum_{j=1}^{M} \alpha_j \mathrm{e}^{-t/\tau_j}\right) \times \left[1 + \left(3\cos^2(\Theta) - 1\right)\left(r_\infty + \sum\nolimits_{K=1}^{L} T_{jk}\,\beta_k \mathrm{e}^{-t/\phi_k}\right)\right], \tag{28}$$

   where H is an amplitude parameter; α_j and τ_j are, respectively, the contributions and decay times of j-th fluorescence exponent $\left(\sum_{j=1}^{M} \alpha_j = 1\right)$; r_∞ is the residual polarization; β_k and ϕ_k are, respectively, the amplitudes and rotational correlation times of corresponding anisotropy exponents; T_{jk} controls the associations $T_{jk} = 1$, if fluorescence and anisotropy exponents are associated, and $T_{jk} = 0$, if not; and M and L are, respectively, the number of fluorescence and anisotropy exponential components.
3. Link the parameters. Since each data set for anisotropy analysis is fitted by the model based on Eq. 16, the set of fit parameters includes H, α_j, τ_j, r_∞, β_k and ϕ_k, and $j = 1, \ldots, M$; $k = 1, \ldots, L$. The value α_1 is not fitted directly because the sum of contributions α_j should be kept equal to 1. Therefore the value α_1 is simply calculated as $\alpha_1 = 1 - \sum_{j=2}^{M} \alpha_j$. The fit parameters with the same names should be linked for all data sets participating in the global fit. The only exception is parameter H, as it takes into account the absolute amplitude of the sample polarization components and therefore it is different for different data sets.
4. Generate initial guesses. Although no special initial guesses algorithm exists for Eq. 28, the start values of some parameters can nevertheless be obtained automatically. For example, if possible ranges for decay times and rotational correlation times are known and a number of fluorescence and anisotropy exponents is selected, the initial guesses for parameters τ_j and ϕ_k can be uniformly distributed in corresponding ranges.

The initial guesses for all contributions α_j can be selected equal to $1/M$. In most cases the initial values for β_k can be set equal to $0.4/L$.

5. Perform the iterative fit. The instructions for the iterative fit are given in Subheading 3.1.
6. Judge the goodness of the fit. To judge the quality of the fit of sample polarization components the fit quality criteria presented in Subheading 3.1 can be applied. Additionally, the measured and model-generated anisotropy decays can be built from polarization components and presented graphically. The anisotropy decays are calculated as follows:

$$r(t) = \frac{I_1(t)/H_1 - I_2(t)/H_2}{I_1(t)(1 - 3\cos^2(\Theta_2))/H_1 - I_2(t)(1 - 3\cos^2(\Theta_1))/H_2}, \quad (29)$$

where $I_1(t)$ and $I_2(t)$ are sample polarization components (measured or model-generated) that correspond to the polarization angles Θ_1 and Θ_2, respectively, and H_1 and H_2 are estimated values of amplitude parameters (Eq. 28). When a shift δ_s exists between measured polarization components $I_1(t)$ and $I_2(t)$, then before calculating the anisotropy decays this shift should be compensated by moving, for example, $I_2(t)$ to $I_1(t)$ over δ_s time units.

When the total IRF is used for fitting polarization data (*see* **Note 4**) the sample G-factor G_s can be estimated using amplitude parameters that correspond to the parallel $H_{||}$ and perpendicular $H_{\perp}$ polarization components:

$$G_s = \frac{H_{||}}{H_{\perp}}. \quad (30)$$

At the same time it is recommended to fix G-factor to the value obtained from the additional measurements, since the relation between parameters $H_{||}$ and $H_{\perp}$ is correlated with anisotropy parameters β_k and ϕ_k. This can be done by fixing $H_{||}$ and $H_{\perp}$ to some values that being substituted to the Eq. 30 give the required value of G_s. In this case the condition $\left(\sum_{j=1}^{M} \alpha_j = 1\right)$ cannot be kept anymore. Therefore all parameters α_j with the same indexes should be linked and the obtained parameter groups should be estimated to fit the amplitudes of the sample decays.

3.3 Example of Global Anisotropy Analysis of Enhanced EGFP

To demonstrate the application of global approach for time-resolved anisotropy analysis the parallel and perpendicular fluorescent components of enhanced green fluorescence protein (EGFP) were measured and fitted globally according to the algorithm described in Subheading 3.2. Global analysis was performed with the TRFA Data Processor Advanced software (SSTC, http://www.sstcenter.com). Since one-exponential reference compound was used for deconvolution, the general model Eqs. 2 and 28 were selected to obtain model-generated curves. The total decays for reference

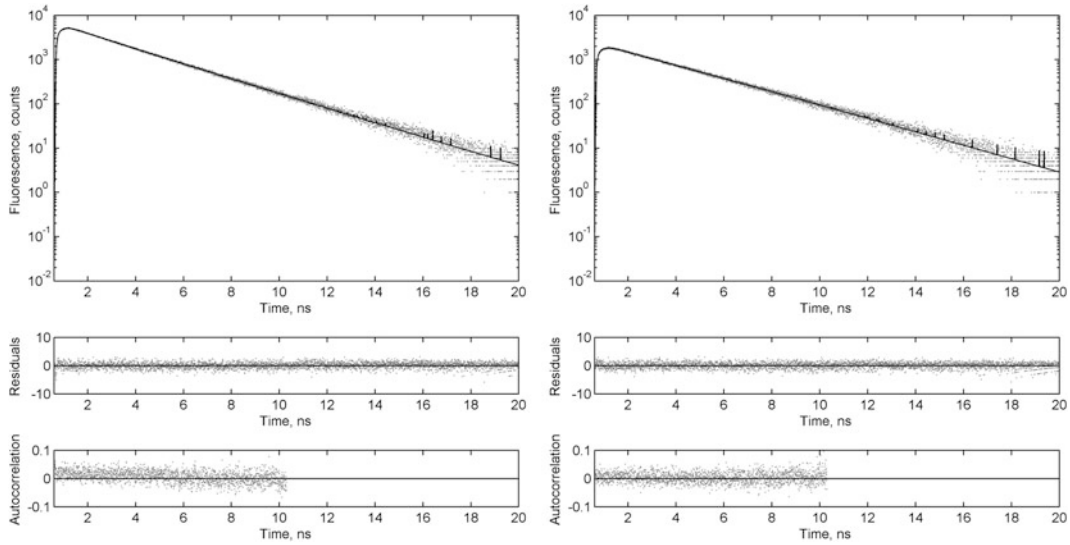

Fig. 2 Experimental (points) and fitted (*solid line*) curves of parallel (*left panel*) and perpendicular (*right panel*) polarization components of EGFP. The results are obtained by the global analysis approach. The quality of the fit can be judged by graphs of weighted residuals and their autocorrelation shown at the *bottom* of both panels. The target fit criterion $\chi^2_G = 1.089$. The estimated values of fluorescence lifetimes are $\tau_1 = 0.771$ ns (with contribution of 8.4 %) and $\tau_2 = 2.735$ ns (with contribution of 91.6 %). The rotational correlation time was estimated at $\phi_1 = 15.6$ ns with initial anisotropy $\beta_1 = 0.383$

compound were prepared from its polarization components according to **Note 4** and added to the corresponding data sets. For both data sets the models defined by Eq. 28 were configured to include two fluorescence exponents ($M = 2$) and one anisotropy exponent ($L = 1$). The values of parameters $H_{||}$ and $H_{\perp}$ were fixed to the same values to keep the $G_s = 1$. Parameters α_1 were linked for both data sets and obtained parameter group was fitted as a separate fit parameter. Reference lifetime parameters τ_{ref} for both data sets were fixed to the prior known value. Since constant background was not observed in the measured data the parameters b and c from Eq. 2 were fixed to 0. The intensity of time-dependent background emission $B(t)$ was available from the measurements for both parallel and perpendicular components with known relative weight equal to 5 (parameters γ for both data sets were fixed to this value). The number of registered photons was high enough to accept the Gaussian approximation of statistical noise in both polarization components of the sample. Therefore the TFC defined by the chi-square Eq. 13 was applied. The Marquardt–Levenberg optimization method was used in the iterative fit. The confidence intervals for fit parameters at 95 % level were estimated by the exhaustive search method.

The results of the global analysis of both EGFP polarization components are presented in Fig. 2. The graphs for anisotropy are shown in Fig. 3. The final value of the TFC obtained after the fit $\chi^2_G = 1.089$. This value as well as the presented graphs for weighted residuals and their autocorrelation functions for both polarization

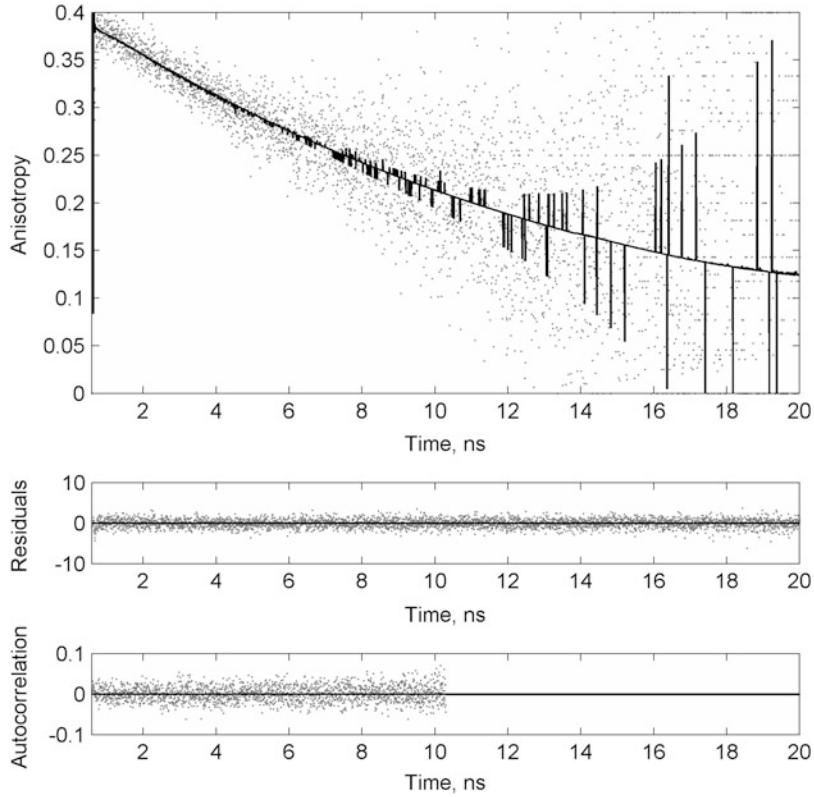

Fig. 3 Experimental (points) and fitted (*solid line*) curves of EGFP anisotropy. The quality of the fit can be judged by presented graphs of weighted residuals and their autocorrelation. Parameter values are given in the legend of Fig. 2

components prove that the fit was successful. The estimated values of EGFP lifetimes are $\tau_1 = 0.771$ [0.714; 0.816] ns (with contribution of 8.4 [7.9; 8.8] %) and $\tau_2 = 2.735$ [2.711; 2.742] ns (with contribution of 91.6 [91.5; 92.3] %). The rotational correlation time was estimated at $\phi_1 = 15.6$ [15.2; 16.0] ns with initial anisotropy $\beta_1 = 0.383$ [0.382; 0.386]. The values shown in square brackets represent the estimated confidence intervals at the 95 % level. The obtained parameter values for EGFP are in good agreement with the results reported previously [11, 26, 27]. This result also demonstrates that the global analysis scheme presented in this section can be successfully applied in practice for fitting the experimental data.

4 Notes

1. If Eq. 13 is applied to fit time-domain data and analytical models are selected for the fit, then the variances in each time point are defined as $\sigma_{ik}^2 = 1/x_{ik}$. The measured fluorescence intensity decays can contain time channels where no photons were detected and therefore $x_{ik} = 0$. Since σ_{ik}^2 for such points

cannot be calculated these points are excluded from the sum in Eq. 13 and total number of points n decreases.

2. In most cases the constraints are defined for all fit parameters. The simplest type of the constraint is setting minimum and maximum values but more complex functional constraints are also possible. In the presence of constraints the new values of fit parameters obtained by the optimization method (Eq. 20) should be checked to satisfy them. If the values of some parameters do not satisfy their constraints, the new values of these parameters are replaced by values from the current point in the fit parameter space. Then the new set of fit parameter values will contain the mixture of new values a_j^{tr} obtained by Eq. 20 and current values a_j^{curr}, $j = 1, \ldots, s$. Forming such mixed vector of fit parameter values is more effective than completely decline the new set of values obtained by Eq. 20, because it makes the behavior of the optimization method more flexible in the area close to the boundaries of possible values of fit parameters.
3. As it follows from the Eq. 26, $C_{i1} = 1$. This means that each value of weighted residual R_{ik} is fully correlated to itself. In order to make autocorrelation plots more convenient for visual inspection the value C_{i1} is replaced by 0 on the graphs.
4. For the measured IRF/reference data an additional preprocessing is required before they can be included to the data sets for anisotropy analysis. Let us consider this procedure for the case of an IRF (for a reference compound the same steps should be performed). In order to accurately estimate the value of the sample G-factor after the fit the data sets for anisotropy analysis should include the total IRF that corresponds to the magic angle, although the IRF curves are usually measured on the same polarization angle as sample fluorescence intensity decay. To calculate these total curves the polarization components of IRF $g_{||}(t)$ and $g_{\perp}(t)$ can be used and possible time shift between these components should be taken into account. The time shift δ_{IRF} of perpendicular IRF polarization component in respect to the parallel one is detected by a superposition of their rising edges, thus ensuring better fit of sample decays at the beginning. The IRF for parallel component data set is prepared by the formula

$$g_0(t) = g_{||}(t) + 2 \times G_{\mathrm{IRF}} g_{\perp}(t + \delta_{\mathrm{IRF}}), \tag{31}$$

where G_{IRF} is the G-factor at the wavelength used for IRF measurement. The IRF for perpendicular component data set is calculated as

$$g_{90}(t) = g_{||}(t - \delta_{\mathrm{IRF}}) + 2 \times G_{\mathrm{IRF}} g_{\perp}(t). \tag{32}$$

Application of total IRFs to fit the sample polarization component data potentially allows to restore the sample G-factor after the fit.

References

1. Lakowicz JR (2006) Principles of fluorescence spectroscopy, 3rd edn. Springer, New York
2. Beechem JM, Gratton E, Ameloot M et al (2002) The global analysis of fluorescence intensity and anisotropy decay data: second-generation theory and programs. In: Lakowicz JR (ed) Topics in fluorescence spectroscopy, vol 2. Springer, USA, pp 241–305
3. Verveer PJ, Squire A, Bastiaens PI (2000) Global analysis of fluorescence lifetime imaging microscopy data. Biophys J 78(4):2127–2137
4. Elson D, Requejo-Isidro J, Munro I et al (2004) Time-domain fluorescence lifetime imaging applied to biological tissue. Photochem Photobiol Sci 3(8):795–801
5. Borst JW, Visser AJWG (2010) Fluorescence lifetime imaging microscopy in life sciences. Meas Sci Technol 21:1–21
6. Tsien RY (1998) The green fluorescent protein. Annu Rev Biochem 67:509–544
7. Bastiaens PIH, Squire A (1999) Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell. Trends Cell Biol 9(2):48–52
8. Borst JW, Laptenok SP, Westphal AH et al (2008) Structural changes of yellow Cameleon domains observed by quantitative FRET analysis and polarized fluorescence correlation spectroscopy. Biophys J 95(11):5399–5411
9. Laptenok SP, van Stokkum IH, Borst JW et al (2012) Disentangling picosecond events that complicate the quantitative use of the calcium sensor YC3.60. J Phys Chem B 116 (9):3013–3020
10. Laptenok SP, Borst JW, Mullen KM et al (2010) Global analysis of Förster resonance energy transfer in live cells measured by fluorescence lifetime imaging microscopy exploiting the rise time of acceptor fluorescence. Phys Chem Chem Phys 12(27):7593–7602
11. Visser AJWG, Laptenok SP, Visser NV et al (2010) Time-resolved FRET fluorescence spectroscopy of visible fluorescent protein pairs. Eur Biophys J 39(2):241–253
12. Malikova NP, Visser NV, van Hoek A et al (2011) Green-fluorescent protein from the bioluminescent jellyfish *Clytia gregaria* is an obligate dimer and does not form a stable complex with the Ca(2+)-discharged photoprotein clytin. Biochemistry 50(20):4232–4241
13. Borst JW, Hink MA, van Hoek A et al (2005) Effects of refractive index and viscosity on fluorescence and anisotropy decays of enhanced cyan and yellow fluorescent proteins. J Fluoresc 15(2):153–160
14. Van den Bergh V, Kowalczyk A, Boens N et al (1994) Experimental design in the global compartmental analysis of intermolecular two-state excited-state processes. J Phys Chem 98: 9503–9508
15. Boens N, Kowalczyk A (1996) Identifiability of competitive intermolecular three-state excited-state processes. Chem Phys Lett 260: 326–330
16. Apanasovich VV, Novikov EG (1995) Method for simultaneous reconstruction of parameters for a set of fluorescence intensity decay curves. J Appl Spectrosc 62(6):1069–1074
17. Apanasovich VV, Novikov EG (1996) The method of fluorescence decays simultaneous analysis. Rev Sci Instrum 67(1):48–54
18. Press WH, Teukolsky SA, Vetterling WT et al (2007) Numerical recipes: the art of scientific computing, 3rd edn. Cambridge University Press, New York
19. Bevington PR, Robinson DK (2003) Data reduction and error analysis for the physical sciences, 3rd edn. McGraw-Hill, New York
20. Laurence TA, Chromy BA (2010) Efficient maximum likelihood estimator fitting of histograms. Nat Methods 7(5):338–339
21. Maus M, Cotlet M, Hofkens J et al (2001) An experimental comparison of the maximum likelihood estimation and nonlinear least-squares fluorescence lifetime analysis of single molecules. Anal Chem 73(9):2078–2086
22. Hauschild T, Jentschel M (2001) Comparison of maximum likelihood estimation and chi-square statistics applied to counting experiments. Nucl Instr Meth A 457:384–401
23. Marquardt DW (1963) An algorithm for least-squares estimation of non-linear parameters. J Soc Ind Appl Math 11:431–441
24. Eaton DF (1990) Recommended methods for fluorescence decay analysis. Pure Appl Chem 62(8):1631–1648
25. Johnson ML, Faunt L (1992) M. Parameter estimation by least-squares methods. Methods Enzymol 210:1–37

26. Heikal AA, Hess ST, Webb WW (2001) Multiphoton molecular spectroscopy and excited-state dynamics of enhanced green fluorescent protein (EGFP): acid–base specificity. Chem Phys 274:37–55

27. Hess ST, Sheets ED, Wagenknecht-Wiesner A et al (2003) Quantitative analysis of the fluorescence properties of intrinsically fluorescent proteins in living cells. Biophys J 85 (4):2566–2580