ESR experimental data analysis using simulation-based fitting approach

Katerina V. Makarova [1)], Petr V. Nazarov [1)], Vladimir V. Apanasovich [1)], Marcus A. Hemminga [2)]

1) Department of Systems Analysis, Belarusian State University, Skoryna Ave 4, 220028, Minsk, Belarus, katena@tut.by, www.sstcenter.com/dsa/

2) Laboratory of Biophysics, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands, www.mf.wau.nl/hemminga/

Abstract: *A mathematical model of ESR in isotropic case together with algorithms of analysis of experimental ESR spectra is proposed. The developed methods allow extracting physical parameters of multi component heterogeneous systems; and therefore they overcome the limitations of standard direct spectra analysis*

Keywords**:** *ESR, simulation, membrane proteins.*

1. INTRODUCTION

Experimental data analysis is one of the most important stages of different spectroscopic studies aimed on complex biomolecular systems. This statement especially holds for electron spin resonance (ESR) spectroscopy, which is one of the successful tools for studying such complex biomolecular systems as membrane proteins. The data analysis in ESR spectroscopy is rather complex because of the following reasons: there could be a number of unknown parameters; almost all dependencies between them and the processes occur are non-linear; and experimental data are distorted by noises and inaccuracies of a registration system. These facts impel to analyse experimental data *via* the multi-parametric optimisation approach (fitting). The idea of fitting is the approximation of experimental data by synthetic data obtained by simulation or mathematical modelling.

The current work is aimed to study small peptides by ESR spectroscopy. The adequate mathematical model of ESR in the case of isotropic motion of spin labels was built and then applied for analysis of experimental data. The performed study provides significant information about the influence of the presence of short peptides in a lipid bilayer.

2. ESR SPECTROSCOPY

The ESR spectroscopy is based on Zeeman effect – splitting of electron energy states in an external magnetic field. Large molecules, such as membrane proteins, are labelled using rather small spin labels containing nitroxide group. Being placed into strong magnetic field, electron energy states are split and therefore can be studied using standard absorption spectroscopy technique. The resulted absorption ESR spectrum contains precise information about environment of spin labels, their mobility, interaction between them, etc.

Therefore this technique provides distance information within the range of 5-20 Å together with information about protein mobility, which is sufficient to study the geometry of small membrane proteins and their complexes. In our days ESR has matured into a powerful, versatile and non-destructive analytical method. Unlike many other techniques, ESR can produce meaningful structural and dynamical information, even from ongoing chemical and/or physical processes without influencing the process itself. Therefore, ESR is an ideal technique in the area of structural proteomics as a tool to gain information about mobility, geometry and functions of membrane proteins.

However the complexity of protein-lipid systems hampers and limits the analytical interpretation of ESR spectra, especially in the case of non-homogenous environment of spin labels in the sample (multicomponent systems) [6]. Standard methods of spectra analysis includes direct extraction of line shape parameters: measurement of peak heights, splitting, and overall broadening. These parameters of a line shape do not give precise information about physical processes inside experimental system.

In heterogeneous systems, when experimental ESR spectrum is a superposition of a number of homogenous spectra from spin labels being in different conditions it is possible to measure the degree of partitioning from the line width ratios in the high-field region of the spectrum [7]. However, there are a number of broadening effects in experiments and therefore lines are not well resolved. Two spectra are overlapped and the approach of Bales [7] cannot be used.

3. GENERAL APPROACH TO PARAMETER DETERMINATION

To analyze multicomponent ESR spectra mathematical or simulation modeling has to be used. The idea of the analysis applied in current work is the following. After developing of the mathematical model, experimental data are approximated by simulated ones. The Simulation-based fitting (SBF) includes tuning of real physical parameters of the model to minimize deviation between experimental and simulated data. As a result one can get precise information about physical processes occurring in the system and their parameters [8].

A complex biophysical system Θ can be characterized by a vector of input parameters $\mathbf{P} = (p_1, p_2, p_3, \ldots)$ of the system Θ. After a number of experimental studies on the system Θ are carried out with different input parameters, the vector of output values **F** can be obtained. In this case,

the system can be considered as an operator performing the following operation:

$$\Theta\ (\ p_1\ ,\ p_2\ ,\ p_3\ ,...)\ =\ \Theta\ (\mathbf{P}\ )\ =\ \mathbf{F} \qquad (1)$$

Usually, some input parameters are known. Let us denote them $\mathbf{P}_0$, for example let $\mathbf{P}_0 = (p_1, p_2)$. Other parameters, which should be extracted, are denoted as $\mathbf{P}_X$, suppose $\mathbf{P}_X = (p_3,\ p_4\ldots)$. The vector of input parameters therefore includes a combination of known and unknown parameters $\mathbf{P} = (p_1, p_2, p_3\ \ldots) = (\mathbf{P}_0, \mathbf{P}_X)$. The extraction of $\mathbf{P}_X$ is the aim of the analysis.

Let us assume that for system Θ it is possible to build an adequate model (mathematical or simulation one), which performs operation (2) with the same physical parameters **P**.

$$\Xi\ (\mathbf{P}\ )\ \equiv\ \Xi\ (\mathbf{P}_0\ ,\mathbf{P}_X\ )\ =\ \mathbf{F}^*\ , \qquad (2)$$

where $\mathbf{F}^*$ contains the simulated output values, which should approximate the experimental ones.

The determination of the unknown parameters $\mathbf{P}_X$ is carried out in the form of SBF. The flow diagram of this method is shown in Fig. 1 [8].

The following steps can be identified in SBF:

1. Output values **F** are obtained experimentally (see blocks 1-3).
2. An adequate model Ξ of system Θ, which performs operation (2), is created (block 5).
3. An initial estimation $\mathbf{P}_X{}^*$ is made for $\mathbf{P}_X$ (block 4).
4. An optimization algorithm, using a variation of parameters $\mathbf{P}_X{}^*$ minimizes the discrepancy function $\|\mathbf{F}^*-\mathbf{F}\|$ (blocks 6-8 and 5 again).
5. Finally, the fitted parameters $\mathbf{P}_X{}^*$ which should estimate the experimental parameters $\mathbf{P}_X$, are obtained (block 9).

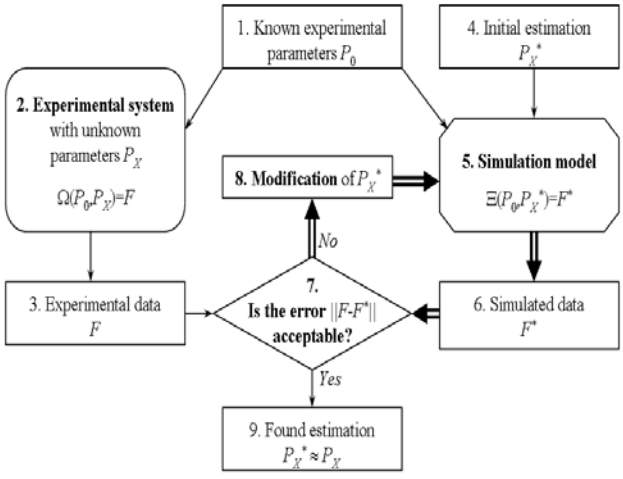


Figure 1. Flow diagram of simulation-based fitting

4. MATHEMATICAL MODEL OF ESR SPECTRA

To analyze ESR spectra from heterogeneous systems, the mathematical model of ESR has been built. Present work is focused on the situation of fast isotropic motion of nitroxide spin labels.

In the ideal case of fast isotropic motion (such as a liquid) ESR spectrum for a nitroxide label consists of three Lorenzian lines [9]. However, due to the interactions of spin labels with other atoms (oxygen, hydrogen) energy levels are split and broadened. The resulted spectrum is a superposition of many slightly shifted Lorenzian peaks – a statistical distribution on unresolved hyperfine interactions. This gives a Gaussian function [9] or, in vast majority of cases, a mixture of Gaussian and Lorenzian functions or Voight function (convolution of Gaussian and Lorenzian).

The spectra are unresolved, either intrinsically by nature of the spin label and the experiment, or artificially, because of broadening caused by oxygen or other paramagnetic species.

The line broadening effects due to ^{13}C, ^{1}H (they are presented in the spin labels), spin-spin exchange interaction and presence of the oxygen in the system were taken into account.

5. OPTIMIZATION ROUTINE

In current work the downhill simplex method, introduced by Nelder and Mead [10], was used for parameters tuning. The method requires only function evaluations, not derivatives. It has better convergence and higher noise stability compared to those for the gradient methods.

To navigate the optimization, an objective fitting function ξ was introduced. It measures the goodness of fit of the simulated spectrum $\mathbf{Y}^{sim}$ to the experimental one ($\mathbf{Y}^{exp}$).

The following objective function was introduced:

$$\xi = \sum_{i=1}^{N} w_i \sqrt{\left| Y_i^{\exp} - Y_i^{sim} \right|}\ , \qquad (3)$$

where w_i – is the weight (value of importance) of the *i*-th point in experimental data. This objective function takes into account the importance of parts of ESR spectra.

The importance of different parts in ESR spectrum varies. For instance, the background is of zero importance, and the right peak ($M_I = -1$) can give more precise information about partitioning than peak with $M_I = 0$. At the same time it is not possible to fit a single peak because other peaks contains important information about correlation time and Gaussian broadening. To get reliable information three peaks in experimental spectrum have to be fitted with different weights.

A number of numerical experiments were performed to determine the best levels of weight function for each of three peaks.

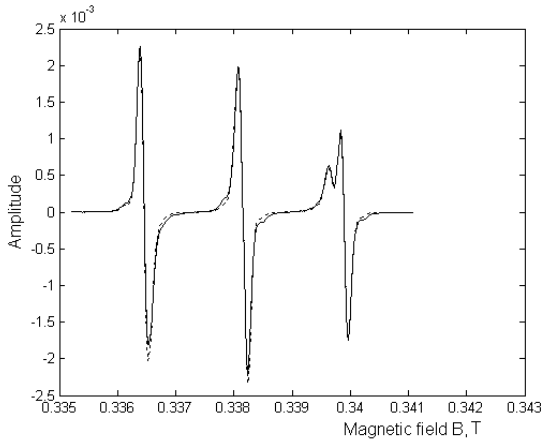
6. RESULT AND DISCUSSION

The two component spectrum is presented in Fig.2 together with the approximation made by the developed model.

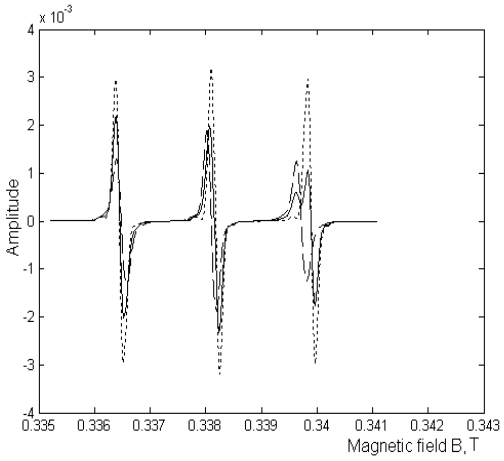
Proposed method of spectra separation allows defining the ratio and parameters for each component of the spectrum. With higher temperature spin labels prefer to

go into the membrane; as a result, the fluidity of the membrane changes.

All algorithms were implemented as functions of MATLAB. They are used in the Laboratory of Biophysics, Wageningen University, The Netherlands, for the analysis of two-component isotropic spectra obtained from membrane-protein systems.



a)



b)

Figure 2. Two component spectrum (a) and its approximation by the model (b)

The developed models and algorithms were applied to analyze the influence of small transmembrane peptides from H+ vacuolar ATPase protein [10] on the spin label partitioning in water and lipid phases together with correlation times and other physical parameters for both phases. The study of those peptides is of great importance for developing new drugs and ways of treatment of osteoporosis disease. The TEMPO spin labels were used. This system gives a two-component ESR spectrum. One component comes from spin labels in lipid phase; the other comes from spin labels in water.

The effect of the presence of peptides and their concentration in the lipid bilayer and temperature dependence were studied.

The experimental data were obtained in the Laboratory of Biophysics, Wageningen University, The Netherlands under supervision of Dr. Marcus A. Hemminga

7. REFERENCES.

[1] Byrne, B**.;** Iwata, S. Membrane protein complexes. *Curr. Opin. Struct. Biol.* 2002, *12*, 239–243.

[2] Arora, A.; Tamm, L. K. Biophysical approaches to membrane protein structure determination. *Curr. Opin. Struct. Biol.* 2001, *11*, 540–547.

[3] dos Remedios, C. G.; Moens, P. D. J. Fluorescence resonance energy transfer spectroscopy is a reliable "ruler" for measuring structural changes inproteins - Dispelling the problem of the unknown orientation factor. *J. Struct. Biol.* 1995, *115*, 175–185.

[4] Marsh, D. Membrane spectroscopy. Electron spin resonance: spin labels. *J. Molecular Biology, Biochemistry and Biophysics* (1981), 31, 51-142.

[5] Nordio. T. Theory of ESR. *Biological Magnetic Resonance, Vol. 1*. Berliner, Lawrence J.; Reuben, Jacques; Editors. USA,1978, 27-176.

[6] Filipič, B. and Štrancar, J., Tuning EPR spectral parameters with a genetic algorithm, *Appl. Soft Comp*. 2001, 1, 83-90.

[7] Bales, B.L. ; Peric M. EPR line shifts and line shape changes due to spin exchange of nitroxide free radicals in liquids. *J.Phys Chem*, 1997, 101, 8707-8716.

[8] Nazarov P.V., Apanasovich V.V., Lutkovski V.M., Yatskou M.M., Koehorst R.B.M., Hemminga M.A. Artificial neural network modification of simulation-based fitting: application to a protein-lipid system, *J. Chem. Inf. Comput. Sci*., 2004, 44, 568-574.

[9] Bales, B.L. ; Peric M. EPR line shifts and line shape changes due to spin exchange of nitroxide free radicals in liquids. *J.Phys Chem*, 1997, 101, 8707-8716.

[10] Nelder, J.A., and Mead, R., *Computer Journal*, 1965, 7, 308–313.

[11] Beutler JA, McKee TC, Novel marine and microbial natural product inhibitors of vacuolar ATPase. *Current Medicinal Chemistry* 2003, 10 (9), 787-796.