

OPTESIM, a versatile toolbox for numerical simulation of electron spin echo envelope modulation (ESEEM) that features hybrid optimization and statistical assessment of parameters

Li Sun, Jessica Hernandez-Guzman, Kurt Warncke *

Department of Physics, Emory University, Atlanta, GA 30322, USA

ARTICLE INFO

Article history:

Received 31 December 2008

Revised 25 April 2009

Available online 29 May 2009

Keywords:

ESEEM

EPR

Optimization

Simulation

ABSTRACT

Electron spin echo envelope modulation (ESEEM) is a technique of pulsed-electron paramagnetic resonance (EPR) spectroscopy. The analysis of ESEEM data to extract information about the nuclear and electronic structure of a disordered (powder) paramagnetic system requires accurate and efficient numerical simulations. A single coupled nucleus of known nuclear g value (g_N) and spin $I = 1$ can have up to eight adjustable parameters in the nuclear part of the spin Hamiltonian. We have developed OPTESIM, an ESEEM simulation toolbox, for automated numerical simulation of powder two- and three-pulse one-dimensional ESEEM for arbitrary number (N) and type (I, g_N) of coupled nuclei, and arbitrary mutual orientations of the hyperfine tensor principal axis systems for $N > 1$. OPTESIM is based in the Matlab environment, and includes the following features: (1) a fast algorithm for translation of the spin Hamiltonian into simulated ESEEM, (2) different optimization methods that can be hybridized to achieve an efficient coarse-to-fine grained search of the parameter space and convergence to a global minimum, (3) statistical analysis of the simulation parameters, which allows the identification of simultaneous confidence regions at specific confidence levels. OPTESIM also includes a geometry-preserving spherical averaging algorithm as default for $N > 1$, and global optimization over multiple experimental conditions, such as the dephasing time (τ) for three-pulse ESEEM, and external magnetic field values. Application examples for simulation of ^{14}N coupling ($N = 1, N = 2$) in biological and chemical model paramagnets are included. Automated, optimized simulations by using OPTESIM lead to a convergence on dramatically shorter time scales, relative to manual simulations.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Electron spin echo envelope modulation (ESEEM) is a technique of pulsed-electron paramagnetic resonance (EPR) spectroscopy that reveals the energies of electron-nuclear hyperfine (hf) interactions, and, for nuclei with $I \geq 1$, nuclear quadrupole interaction (nqi) energies [1,2]. Nuclear state mixing and a bandwidth of the microwave pulse magnetic field (B_1) larger than the hf splitting afford favorable conditions for the detection of ESEEM from single or small numbers of electron-nuclear couplings. The modulation of the echo envelope is collected in the pulse time domain, and is cosine Fourier-transformed to create an ESEEM frequency spectrum. Schematic pulse timing diagrams for the standard one-dimensional two-pulse and three-pulse ESEEM experiments are shown in Fig. 1. In certain cases, qualitative information about the identity, number, and magnitude of the coupling of the coupled nuclear

spin(s) can be gained from inspection of the modulation periodicities in the waveform or the line positions in the Fourier transform (FT). However, simulation of the ESEEM is required to extract the coupling parameters that give the desired detailed information about the nuclear and electronic structure of the molecular paramagnet, such as the hf tensor and electric field gradient (efg) tensor (nqi parameters). Different programs for the numerical simulation of ESEEM have been reported and used to interpret one-dimensional and two-dimensional experiments [3–9]. Although the simulations yield, in many cases, a high level of agreement with experiment, the process of matching simulation and experiment remains both subjective and time-consuming, and is the bottleneck in the interpretation of ESEEM experiments. Here, we describe an ESEEM simulation software toolbox that incorporates automated fitting with an efficient hybrid optimization approach, and provides an assessment of the statistical significance of the simulation parameters.

The overall design goal of the tool box (denoted by OPTESIM, for optimization of ESEEM simulation) is to provide a modular, flexible set of routines for an automated, optimization-based simulation of diverse paramagnetic systems (arbitrary number of hf coupled

* Corresponding author. Address: Department of Physics, N201 Mathematics and Science Center, 400 Dowman Drive, Emory University, Atlanta, GA 30322-2430, USA. Fax: +1 404 727 0873.

E-mail address: kwarncke@physics.emory.edu (K. Warncke).

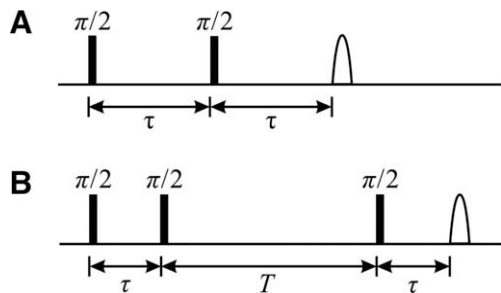


Fig. 1. Pulse timing diagrams for the two- and three-pulse ESEEM pulse sequences. (A) Two-pulse sequence: $\pi/2 - \tau - \pi/2 - \tau - \text{detection}$. (B) Three-pulse sequence: $\pi/2 - \tau - \pi/2 - T - \pi/2 - \tau - \text{detection}$.

nuclei with arbitrary nuclear spin; different classes of electron spins). For this purpose, a collection of functions for scripting is most appropriate. The Matlab (Mathworks, Natick, MA, USA) programming environment is chosen for the development of OPTESIM, because of the following features: (1) Matlab is generally available for academic use, (2) Matlab is easier for further development, compared to a C++ or FORTRAN environment, and (3) a number of literature numerical and optimization routines are available in Matlab. In addition, a practical consideration is that experiment control and data collection software can be developed in Matlab, which allows a seamless flow from ESEEM data acquisition to end analysis. The core calculations of the nuclear eigenvalues and eigenvectors and the time evolution of the density matrix use direct diagonalization methods [10,11]. The optimization is achieved by using the minimization of the least-squares residuals of the match of the simulated to the experimental ESEEM, and is generally carried out in the time-domain, where distortions of the FT owing to waveform truncation are absent. Algorithms for the search of the hf, nqi, and system-specific Euler angle parameter space are as follows: genetic algorithm [12], simulated annealing [13], Nelder–Mead simplex [14], Gauss–Newton method, and Levenberg–Marquardt method [15]. A statistical assessment of the uncertainty in a parameter, with respect to other parameters, is given by the calculated simultaneous confidence region for each parameter.

We apply OPTESIM to the simulation of a system, which had been previously simulated by “manual” simulation, and to one which is too complex to efficiently simulate by “manual” simulation. In each case, a two-tier approach of genetic algorithm (coarse grained search for basin around the global minimum) followed by simplex (fine grained search for the global minimum within the basin) optimization are used. The automated, optimization simulations converge on dramatically shorter timescales, relative to manual simulations.

2. Theory

2.1. Spin Hamiltonian

The stationary-state Hamiltonian for treatment of the interaction of an electron spin ($S = 1/2$) with one nuclear spin (I) is formulated with a hf coupling term, a nuclear Zeeman term, and a nqi term (for $I \geq 1$), as follows:

$$\mathbf{H} = \hbar \bar{S} \cdot \mathbf{A} \cdot \bar{I} - g_N \beta_N \bar{B}_0 \cdot \bar{I} + \bar{I}' \cdot \mathbf{Q} \cdot \bar{I}' \quad (1)$$

where β_N , \bar{S} , and \bar{I} are the nuclear magneton, electron spin operator, and nuclear spin operator, respectively, g_N is the nuclear g -value, \mathbf{A} is the hf coupling tensor, and \mathbf{Q} is the nqi tensor [16]. The I

and I' indicate that the eigenfunctions of \mathbf{A} and \mathbf{Q} are, in general, different. The hf tensor has the principal components, $\mathbf{A}_p = [A_{xx} \ A_{yy} \ A_{zz}]$, and is composed of an isotropic part, A_{iso} , and a dipolar part. In the point dipole approximation, the dipolar part is given by an axially symmetric dipolar tensor, $[-A_{dip} \ -A_{dip} \ 2A_{dip}]$, where $A_{dip} = (\mu_0/2h)g_e\beta_e g_N \beta_N r_{eff}^{-3}$, and g_e , β_e , and r_{eff} are the electron g -value, Bohr magneton, and an effective distance between the electron and the nuclear spins, respectively [16]. The hf tensor \mathbf{A}_p is rotated to the molecular frame, which is defined by a reference principal axis system (PAS) that is provisionally related to the electron g tensor PAS, by the following operation:

$$\mathbf{A} = \mathbf{R}_A^T \mathbf{A}_p \mathbf{R}_A \quad (2)$$

where \mathbf{R}_A is a rotation matrix, which is defined by the Euler angles, $[\alpha_A \ \beta_A \ \gamma_A]$, to rotate from the electron g tensor PAS to the hf coupling tensor PAS.

The nqi tensor, \mathbf{Q} , is defined by the nuclear quadrupole coupling constant, e^2qQ/h , and the efg asymmetry parameter, η , where e , q , and Q are the elementary charge, the magnitude of the principal component of the efg tensor, and the nuclear quadrupole moment, respectively. In its PAS, the traceless nqi tensor $\mathbf{Q}_p = [Q_{xx} \ Q_{yy} \ Q_{zz}]$ is related to e^2qQ/h and η by the following expressions:

$$Q_{zz} = e^2qQ/(2I(2I-1)\hbar) \quad (3)$$

$$\eta = (Q_{xx} - Q_{yy})/Q_{zz} \quad (4)$$

where $|Q_{zz}| \geq |Q_{yy}| \geq |Q_{xx}|$. In the toolbox, the orientation between the nqi tensor PAS and the hf tensor PAS is defined by the Euler angles $[\alpha_Q \ \beta_Q \ \gamma_Q]$, which consequently define a rotational matrix, \mathbf{R}_Q , and nqi tensor \mathbf{Q} in the molecular PAS (g tensor PAS) to be expressed as follows:

$$\mathbf{Q} = \mathbf{R}_A^T \mathbf{R}_Q^T \mathbf{Q}_p \mathbf{R}_Q \mathbf{R}_A \quad (5)$$

This definition of \mathbf{Q} by using a two-stage rotation allows an additional constraint on the mutual orientations of \mathbf{A} and \mathbf{Q} during numerical optimization, which adds to the flexibility in the specification of the geometry model for multiple electron-nuclear interactions.

To summarize, the general coupled electron – single nucleus system for $I \geq 1$ is parameterized by using the following 13 parameters: g_N , I , A_{xx} , A_{yy} , A_{zz} , α_A , β_A , γ_A , e^2qQ/h , η , α_Q , β_Q , and γ_Q (for $I = 1/2$, the five nqi-associated parameters are dropped). In practice, the coupled nucleus is assigned based on characteristic spectral features or a priori knowledge of the system, so that g_N and I are fixed, which entails eleven adjustable parameters. In disordered, or powder, systems (no preferred orientation of the molecular, or g , PAS relative to B_0), α_A , β_A , and γ_A are not considered, and the number of adjustable parameters is reduced to eight. The point dipole assumption for electron-nuclear dipolar coupling reduces the number of adjustable parameters to seven. Additional nuclear couplings introduce three Euler angle parameters that describe the mutual orientations of the hf PAS, in addition to the individual nuclear coupling parameters, described above. In OPTESIM, all of the parameters can be subjected to the numerical optimization.

2.2. ESEEM calculation

The time-domain ESEEM calculation incorporates the full density matrix diagonalization from the Mims density matrix treatment of ESEEM [10,11]. Although the time evolution simulation of a pulsed magnetic resonance experiment is straightforward [7,8], the explicit evaluation of the transition frequencies and intensities is computationally advantageous for optimization purposes, relative to methods that calculate the time evolution of the density matrix. Specifically, the following adapted numerical

procedure [1,2,17], which we have used previously in ESEEM simulations [9], is found to be most efficient for constructing the two-pulse and three-pulse modulation in OPTESIM.

The Hamiltonian is first partitioned into sub-matrix representations \mathbf{H}_α and \mathbf{H}_β corresponding to α ($m_s = 1/2$) and β ($m_s = -1/2$) electron spin manifolds. \mathbf{H}_α and \mathbf{H}_β are diagonalized separately with unitary matrices \mathbf{M}_α and \mathbf{M}_β to obtain the eigenfrequencies, ν_α and ν_β , belonging to the α and β manifolds, respectively. The Hamiltonian, \mathbf{H} , is transformed into \mathbf{H}' , which corresponds to the coupled representation, as follows:

$$\mathbf{H}' = \begin{bmatrix} 0 & \mathbf{M} \\ \mathbf{M}^\dagger & 0 \end{bmatrix} \quad (6)$$

where $\mathbf{M} = \mathbf{M}_\alpha^\dagger \mathbf{M}_\beta$.

The modulation frequencies and amplitudes are then calculated for specific types of experiments. For example, for the two-pulse echo envelope modulation (pulse sequence, Fig. 1), the frequencies ω and amplitudes A are as follows:

$$\begin{aligned} \omega_{lmjk} &= 2\pi(\nu_{\beta l} + \nu_{\alpha m} - \nu_{\beta j} - \nu_{\alpha k}) \\ A_{lmjk} &= \mathbf{M}_{jk}^\dagger \mathbf{M}_{kl} \mathbf{M}_{lm}^\dagger \mathbf{M}_{mj} \end{aligned} \quad (7)$$

The indices i, m, j , and k allow permutation through all elements of \mathbf{M} and the eigenvalues.

The echo modulation amplitude $E(\tau)$ is formulated as follows [1,2,17]:

$$E(\tau) = \sum_{l,m,j,k} A_{lmjk} e^{i\omega_{lmjk}\tau} \quad (8)$$

For the three-pulse echo modulation (pulse sequence, Fig. 1), the frequencies ω and amplitudes A are separated into terms representing the α and β manifolds, as follows:

$$\begin{aligned} \omega_{\alpha jk} &= 2\pi(\nu_{\alpha j} - \nu_{\alpha k}) \\ \omega_{\beta ml} &= 2\pi(\nu_{\beta m} - \nu_{\beta l}) \\ A_{\alpha jk} &= \sum_{m,l} \mathbf{M}_{jm} \mathbf{M}_{mk}^\dagger \mathbf{M}_{kl} \mathbf{M}_{lj}^\dagger e^{i(\omega_{\alpha jk} + \omega_{\beta ml})\tau} \\ A_{\beta ml} &= \sum_{k,j} \mathbf{M}_{lj} \mathbf{M}_{kl}^\dagger \mathbf{M}_{mk} \mathbf{M}_{jm}^\dagger e^{i(\omega_{\alpha jk} + \omega_{\beta ml})\tau} \end{aligned} \quad (9)$$

The simulated echo modulation amplitude $E(\tau, T)$ is formulated as follows [1,2,17]:

$$E(\tau, T) = \sum_{j,k} (A_{\alpha jk} e^{i\omega_{\alpha jk}T} + A_{\beta ml} e^{i\omega_{\beta ml}T}) \quad (10)$$

For flexibility, users can easily substitute their own time-domain ESEEM calculations into OPTESIM as Matlab functions. Therefore, experimental systems and pulse sequences other than those considered here [1,2,17] can be addressed by using the optimization and statistical assessment features of OPTESIM.

2.3. Powder average for disordered systems: general and geometry-preserving treatments

For a disordered, or powder, system, the ESEEM represents the average over all possible orientations of the external magnetic field vector B_0 , with respect to the molecular PAS [1,2,16,17]. The number of orientations of B_0 relative to the molecular PAS that can be evaluated is practically limited by the available computational power (time scale of the calculations) and the required accuracy. The orientation sampling vectors can be chosen as either a set of random unit vectors, or a set of evenly-distributed points on a unit sphere, or octant, if symmetry permits. Several methods have been used in the latter cases [18,19]. In OPTESIM, a geodesic-sphere sequence generator evenly distributes vectors on a unit sphere. The

generator subdivides each of the triangular faces of an icosahedron into four smaller triangles on a unit sphere in an iterative manner, yielding geodesic-sphere sequences at different levels of angular resolution. Users of OPTESIM also have the option to specify their own orientation vectors, as well as the weights of particular subsets of vectors. This approach is used for simulating orientation-selection experiments, in cases where the assumption of an isotropic g -tensor cannot be made [9,20].

The overall envelope modulation when multiple nuclei are coupled to the electron is combined according to the product rule. Specifically, for the two-pulse sequence, the modulation is as follows [21],

$$E_{\text{tot}}(\tau) = \prod_i E_i(\tau) \quad (11)$$

and for the three-pulse sequence, the product is taken separately over the α and β manifolds, and the overall envelope modulation is combined, as follows [22],

$$E_{\text{tot}}(\tau, T) = \prod_i E_{\alpha i}(\tau, T) + \prod_i E_{\beta i}(\tau, T) \quad (12)$$

In the case of electron spin coupling with multiple nuclei, it is computationally expedient to perform the powder, or spherical, average of individual electron–nuclear interactions prior to combination of the individual $E_i(\tau)$, or $E_{\alpha, i}(\tau, T)$ and $E_{\beta, i}(\tau, T)$, modulation terms. Information about the mutual orientation of the nuclei, which is contained in the order ≥ 2 modulation product terms, is lost with this treatment. In OPTESIM, the combination rule is specified at the level of a physical model of the geometry of electron–nuclear interactions, as defined by Euler angles that specify the mutual orientation(s) of the different hf PAS. Therefore, the ESEEM is calculated and combined at each powder average orientation, and subtle structural information arising from the geometry of the multiple hf PAS can be retrieved from the overall averaged spectrum.

2.4. Optimization and statistical analysis of experimental parameters

Optimization is achieved by fitting the simulated spectrum (\tilde{Y}) to the experimental one (Y) by using the minimization of the least-squares residuals, which are a function of the parameters ($\vartheta = [\theta_1, \theta_2, \dots, \theta_L]$) of the Hamiltonian. The goodness of least-squares fitting for one spectrum is defined as $Q^2 = \sum_i (Y_i - \tilde{Y}_i)^2$. For a global fitting of multiple spectra with different experimental parameters, Q^2 for the individual spectrum is weighted by its own noise level. The global goodness of fitting is defined as:

$$Q^2 = \sum_k \sum_{i=1}^{n_k} (Y_{k,i} - \tilde{Y}_{k,i})^2 / \sigma_k^2 \quad (13)$$

where σ_k^2 and n_k are the noise variance and number of points in spectrum k , respectively. We assume that the experimental envelope modulation is free of artifacts (for example microwave phase drift or “glitches” above the rms noise level), which is the case in our experiments [23–25], and that the noise variance of each point in the ESEEM waveform is identical (dead time points removed). In this normalized representation and with the two assumptions, the variance of the simulation error of all data points ($Y_{k,i}$) is estimated as follows:

$$S^2 = Q_{\text{min}}^2 / (N - L) \quad (14)$$

where Q_{min}^2 is the optimized Q^2 minimum at the global minimum, N is the total number of experimental points, as given by $N = \sum_k n_k$, and L is the number of independent parameters for optimization. $Q^2(\vartheta)$ can be expanded at $\hat{\vartheta}$, as:

$$Q^2(\vartheta) = Q_{\min}^2 + \frac{1}{2} \sum_{ij} \left(\frac{\partial^2 Q^2}{\partial \vartheta_i \partial \vartheta_j} \right)_{\vartheta=\hat{\vartheta}} (\vartheta_i - \hat{\vartheta}_i)(\vartheta_j - \hat{\vartheta}_j) + \dots \quad (15)$$

where $\hat{\vartheta}$ are the optimal parameters when $Q^2(\hat{\vartheta}) = Q_{\min}^2$. Assuming that the terms with order >2 are negligible, and that the variances of the measurement errors follow a normal distribution, then the term:

$$a = \frac{1}{2S^2} \sum_{ij} \left(\frac{\partial^2 Q^2}{\partial \vartheta_i \partial \vartheta_j} \right)_{\vartheta=\hat{\vartheta}} \quad (16)$$

follows a χ^2 distribution with a freedom of L [26]. The simultaneous confidence region of parameters ϑ can be determined by the following expression:

$$Q^2(\vartheta) = Q_{\min}^2 + aS^2 \quad (17)$$

The simultaneous confidence region gives an estimation of the uncertainty of a selected optimization parameter with respect to the other parameters. The probability of the simultaneous confidence region containing the true values of ϑ_i is determined by $\chi^2(L)$ and a , as a cumulative distribution of $\chi^2(L)$ from $-\infty$ to a . For example, when $L = 2$ and $a = 1^2, 2^2$, or 3^2 , the simultaneous confidence probability is 39.3%, 86.5%, or 98.9%.

3. Simulation of ESEEM

The OPTESIM software consists of 28 Matlab functions and eight Java classes that can be divided into the following four categories: experimental data filtering, numerical simulation, simulation parameter optimization, and distributed computation framework. These categories are incorporated into four stand-alone modules that are integrated in OPTESIM, as described below. The modules may be individually substituted by the routines of other users, if desired. Table 1 shows a code fragment example of OPTESIM.

Table 1
A code Fragment example of OPTESIM.

```
%Load experimental spectrum
spectrum = filterRaw(filename, filter_options);

%Define a model of one deuterium and one proton with a
%product rule
optmodel = eseOptModel({'H2','H1'}, '1*2');

%Set the nuclear parameters for Nucleus 1
optmodel = changeOptModelValue(optmodel,...
    1'eeqQ', 1, 0.2, 'eta', 1, 0.1);

%Set the nuclear parameters for Nucleus 2
optmodel = changeOptModelValue(optmodel,...
    'eeqQ', 2, 0.2, 'eta', 2, 0.1);

%Register variables for Optimization
%Parameter, nucleus number, index, min value, max value
optmodel = registerOptModelVariable(optmodel,...
    'Aiso', 1, 1, -10, 10, ...
    'r', 1, 2, 0, 10, ...
    'Aiso', 2, 3, -10, 10, ...
    'r', 2, 4, 0, 10);

%Set initial values for
%Aiso1, r1, Aiso2, r2, respectively
x = [6.5,2.5,1.5,2.5];

%Use simplex methods for optimization
options = optimset('OutputFcn', @esePlot, 'Display', 'iter');
x_opted = fminsearch(@eseOptObjective, x, options);
```

3.1. Experimental data filtering

Experimental data filtering allows extraction of experimental parameters from saved experimental data files and processing of the modulation amplitude waveform for simulation analysis. For two-pulse ESEEM, the extracted experimental parameters are experiment type, τ value, and B_0 field. For three-pulse ESEEM, the additional experimental parameter, T , is extracted. An example data filtering routine that is compatible with our experimental control software is provided with the package. The ESEEM waveform is then processed, as follows. If necessary, baseline decay of the echo envelope, owing to spin–lattice (T_1) or spin–spin (T_2) relaxation processes, is removed by using an empirical decay function. This is accomplished by fitting the waveform to a user specified monoexponential, biexponential, or stretched exponential function, and then subtracting the decay function from the experimental waveform. Any points recorded within the spectrometer deadtime are removed. The deadtime portion of the waveform is reconstructed as part of the optimization during the simulation. The average noise variance (σ^2) of the processed waveform (minus the deadtime fill points) is calculated.

3.2. Numerical simulation

Numerical simulations of ESEEM are performed for one or more electron–nuclear couplings with the identity of each nucleus supplied by the user, corresponding to the parameters: g_N and I . This information can often be gleaned from a qualitative interpretation of the spectrum, or with the aid of multiple microwave frequency/ B_0 experiments. In the case of unidentified nuclei or numbers of nuclei, the most likely models can be assessed from the statistical criteria in OPTESIM for separate simulations that incorporate different types and numbers of coupled nuclei. A model for the combination of the ESEEM from multiple nuclei is also specified. In the most simple case of all nuclei coupled to the same electron spin, and a single nuclear isotope for each coupling, the envelope modulation for two-pulse and three-pulse experiments is combined according to the product rule expressions in Eqs. (11) and (12), respectively. When the experimental ESEEM is known to arise from sub-populations of electron spins with different nuclear coupling(s), the “sum rule” [21,27] for combination of the individual ESEEM is used, with a weighting factor for each population, to generate the total ESEEM. Practical cases where the sum rule and weighting are necessary are incomplete or heterogeneous isotopic substitution of a nuclear site, conformational isomers of the same paramagnetic molecule, and different paramagnetic molecules. OPTESIM allows users to specify the particular combination rule by using a numerical expression. For example, in the case of two-pulse ESEEM from a molecule with uniform isotope content at coupled nuclear site 1, and heterogenous isotopic content at coupled nuclear site 2, the combination of the individual modulation is specified by “ $E_1[C_{2,a}E_{2,a} + C_{2,b}E_{2,b}]$ ”, where the E_i are the envelope modulation and the $C_{i,a/b}$ are the normalized isotope substitution factors at site 2. The weighting can be included in the simulation as an adjustable parameter. In cases of three-pulse experiments, this combination is done by separately combining the modulation from the α and β manifolds (Eq. (12)).

The calculation of the powder average ESEEM is achieved by using the following steps: (1) the user defines the nuclear parameters for each coupled nucleus and the combination rule. (2) The Hamiltonians for each electron–nuclear coupling are calculated for one orientation of the external magnetic field. (3) The modulation envelopes for each nuclear coupling are calculated based on the Hamiltonians calculated in step 2. (4) The modulation envelopes are combined to attain the overall spectrum for the one

orientation. (5) Steps 2–4 are repeated until the spectra include averaging over all orientations of B_0 relative to the reference PAS.

3.3. Optimization

The simulated ESEEM is matched to the filtered experimental spectrum after the following transformation,

$$\tilde{Y} = c_1(\hat{Y} + c_2)e^{c_3t} + c_4 \quad (18)$$

where \tilde{Y} and \hat{Y} are the transformed and original simulation envelope modulation, respectively, and t represents τ or $\tau + T$, for the two- or three-pulse experiments respectively. The four real constants c_{1-4} are calculated by minimizing the squared residuals of the match between the transformed amplitude and the filtered experimental data. The c_i are defined, as follows: c_1 is the scaling factor for the simulated waveform, c_3 compensates for modulation amplitude decay, and c_2 and c_4 are offset terms for adjusting the waveform prior to, or subsequent to, the decay operation. The scaling constants, c_1 , c_2 , and c_4 , in Eq. 18 are used when the target ESEEM originates from a sample in which the contributions to the ESEEM from nuclear couplings, other than the target couplings of the simulation, are not known. The minimum squared residual is set as the objective function for optimization over the adjustable parameters. For one model of number and type of nuclear couplings and method of combination, users can specify multiple experimental spectra under different experimental conditions (for example, τ value, B_0 value for the same g_e value) for a global optimization, with the fitting objective function defined as in Eq. (13). A standardized definition of the objective function in OPTESIM allows users to take full advantage of available Matlab optimization routines.

In our experience, owing to the large dimension of the parameter space and the rugged landscape of the objective function, gradient-based local optimization methods often fail to find the set of global optimal parameters because of the existence of many local minima, and non-gradient-based optimization methods require extensive computation power. Therefore, the computation speed becomes the bottleneck for the scale and accuracy of optimizations. One solution is to use faster C/C++ code for the core calculations that are crucial to computational speed. In our test, using a compiled C++ executable (MEX-files) for the modulation envelope calculation in Matlab resulted in a decrease of up to 1.5-fold in the time of the overall simulation. Another solution is to distribute the computation of the powder averaging to the available computational resources on a local network.

3.4. Distributed computing framework

A Java-RMI based distributed computational framework is included in OPTESIM. This framework allows users to build a distributed system of ESEEM simulation on their own PC computer hardware resources. The orientation sampling vectors for the powder average are first divided on the local computer and sent to different remote PCs according to their computational speed. The results are then collected and averaged on the local PC. A linear increase in computation speed with respect to the number of PCs can be achieved with this approach.

4. Results

4.1. Example 1: Global simulation of ESEEM from a single ^{14}N nucleus in an enzyme radical reaction intermediate

In three-pulse ESEEM studies of the cryotrapped 2-aminopropanol-generated substrate radical catalytic intermediate in the active

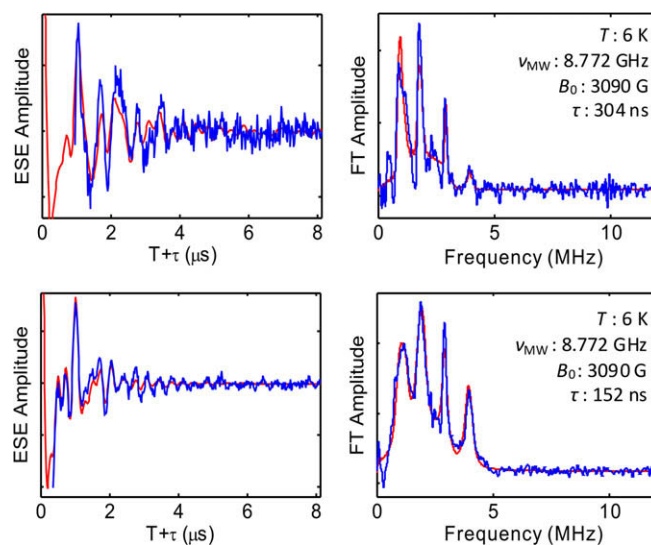


Fig. 2. ^{14}N ESEEM collected from the substrate radical intermediate in ethanolamine ammonia-lyase, and Fourier transform, and simulation (red line) representing a global optimization. (top) $\tau = 304$ ns. (bottom) $\tau = 152$ ns. Simulation parameters are listed in Table 2. Conditions [24]: microwave frequency, 8.772 GHz; magnetic field, 3090 Gauss; temperature, 6 K; microwave pulse width, 20 ns. (For interpretation of color mentioned in this figure legend the reader is referred to the web version of the article.)

Table 2

Simulation parameters for ESEEM from the coupled ^{14}N in the substrate radical intermediate in ethanolamine ammonia-lyase. Euler angles correspond to the relative orientation of the hf and nqi PAS.

Nuclear parameters	Values	Simultaneous confidence (0.99) region
A_{iso} (MHz)	0.933	0.910–0.958
r_{eff} (Å)	3.90	3.63–4.34
e^2qQ/h (MHz)	3.075	3.068–3.084
η	0.579	0.568–0.590
α (°)	30.1	27.4–36.7
β (°)	60.1	55.5–64.9
γ (°)	14.7	6.9–39.0

site of coenzyme B_{12} -dependent ethanolamine ammonia-lyase (EAL) from *Salmonella typhimurium*, a ^{14}N nucleus was found to be coupled to the substrate radical [28,29]. The narrow features in Fourier transforms of the three-pulse ^{14}N ESEEM at 0.8–0.9, 1.7–2.0, and 2.9–3.0 MHz correspond approximately to the ν_0 , ν_- and ν_+ nqi frequencies of ^{14}N (α electron spin manifold for $A_{\text{iso}} > 0$), and the broad feature positioned around 4 MHz is the $\Delta m_{1,\beta} = \pm 2$, or “double quantum”, feature. This spectral pattern is characteristic of the “exact cancellation” condition, in which the hf coupling and nuclear Zeeman contributions approximately cancel in one electron spin manifold [30,31]. The coupled ^{14}N has been assigned to a nitrogen belonging to the guanidinium group of Arg160 in the EutB subunit of EAL [24,32]. Arg160 makes hydrogen bond contact with the substrate in the active site [24]. Multi-frequency three-pulse ESEEM spectra obtained at $g = 2.02$ for the substrate radical were analyzed by using OPTESIM for a global optimization over the ^{14}N nuclear parameters. Fig. 2 shows the global fitting in both time and frequency domains at two different τ values. Table 1 lists the optimal parameters found for the simulation and their simultaneous trusted regions at the confidence level of 0.99. The parameter values are consistent with previously published results [24,29]. However, the simulation with hybrid optimization converged to the parameters presented in Table 2 in approximately 1.5 h, whereas the trial-and-error approach used in the original simulation [29] required months of intermittent effort.

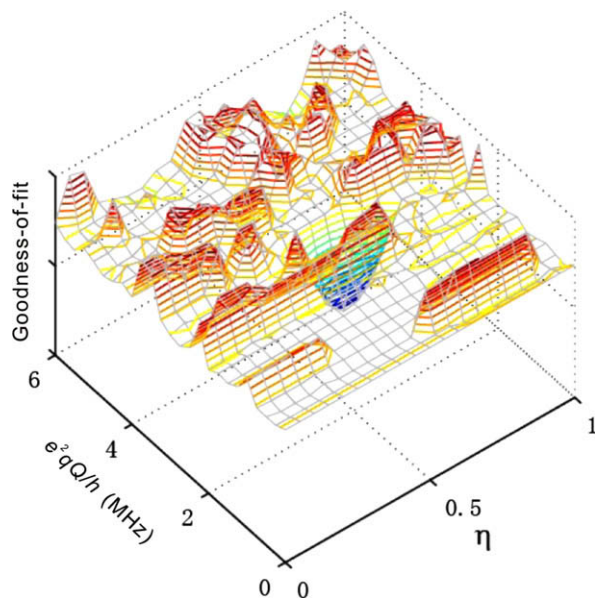


Fig. 3. Goodness-of-fit landscape for the parameters e^2qQ/h and η . Other simulation parameters are fixed to their optimal values, as listed in Table 2.

Fig. 3 shows a goodness-of-fit landscape plot, which is constructed by varying the e^2qQ/h and η values from 0 to 6 MHz and 0 to 1, respectively. This plot represents a simpler optimization problem over e^2qQ/h and η with all other parameters fixed at their optimized values (Table 2). The roughness of the goodness-of-fit landscape prevents local optimization methods, such as conjugated gradient and Nelder–Mead simplex, from finding the global minimum with outlying starting points. Fig. 4A demonstrates the failure of the Nelder–Mead simplex method with random starting points, except for the initial parameter set that is close to the global minimum. In contrast, the non-local genetic algorithm converges to the global minimum within 20 generations of a population size of 50, as shown in Fig. 4B. However, a larger scale genetic algorithm with full parameter optimization requires a larger population size and generation numbers. Our experience shows that an efficient optimization of 10 parameters requires a population size of about 200 for a thorough search of the parameter space. The

algorithm termination condition can be set to a specific number of stale generations [33], and the best individuals will condense at the global minimum in approximately 50 generations. With a hybrid optimization approach, the best individual generated by the genetic algorithm serves as the starting point for further local optimization methods with fast convergence (for example, simplex).

The statistical assessment quantifies the importance of individual simulation parameters in reproducing the ESEEM. For discrete systems (single structure), a narrow confidence region indicates that a parameter contributes significantly to the ESEEM. For example, in the case of the ^{14}N ESEEM in Fig. 2, e^2qQ/h and η directly influence the frequency values of the narrow, large-amplitude ν_0 , ν_- , and ν_+ components, and A_{iso} influences the center frequency value of the ν_{dq} component. These parameters therefore have relatively narrow simultaneous confidence regions (Table 1). The r_{eff} parameter influences the width of the ν_{dq} feature, and therefore has a relatively wider confidence region. The Euler angles in Table 1 subtly modify the relative amplitudes of the ν_0 , ν_- , and ν_+ , and ν_{dq} features, and have correspondingly wider confidence regions. The relatively narrow interval for Euler angle, β , in Table 1, and wider confidence regions for α and γ , is consistent with the choice of an axial dipolar (versus rhombic) hf tensor. If the influence of a parameter on the ESEEM is within the experimental noise, then the simultaneous confidence region would approach the full allowed range (for example, -180° to 180° for an angular parameter). The statistical assessment is useful for culling unnecessary parameters, which may reduce computation time.

4.2. Example 2: Mutual orientation of two ^{14}N hf PAS in Cu(II)-bis-histamine complex

The Cu(II)-bis-histamine complex, $\text{Cu}(\text{II})(\text{him})_2(\text{NO}_3)_2$, is a model for bis-*trans* equatorial coordination of Cu(II) by imidazole in proteins. The structure of the complex in Cu(II)-doped $\text{Zn}(\text{II})(\text{him})_2(\text{NO}_3)_2$ crystal has been determined by X-ray crystallography [34], and superhyperfine (shf) coupling parameters for the remote ^{14}N of imidazole have been calculated by using experimentally constrained density functional theory [35]. The Cu(II)(him) $_2$ (NO $_3$) $_2$ complex was synthesized as described [34], and a frozen 5 mM sample was prepared in 1:1 ethylene glycol/water glass in a 0.4 cm outer diameter EPR tube. Three-pulse ESEEM was collected at the B_0 value for maximum ESE amplitude at

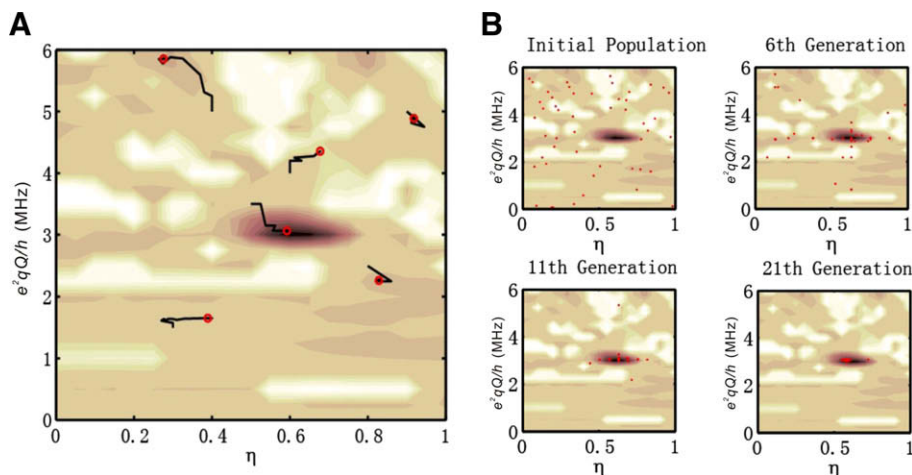


Fig. 4. Trajectories of the goodness-of-fit obtained by using simplex and genetic optimization algorithms. (A) Trajectories of the optimization for the Nelder–Mead simplex method on e^2qQ/h and η . Red circles represent the ending points of the trajectories, which correspond to local minima. The simplex method only succeeds in finding the global minimum when the starting point is close. (B) Optimization trajectory for the genetic algorithm. Red dots represent individuals in a population of 50.

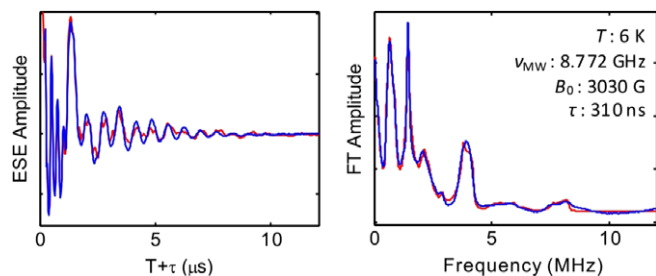


Fig. 5. ESEEM collected from the remote ^{14}N imidazole nuclei in the Cu(II)-bis-histamine complex and Fourier transform, and simulation (red). The simulation parameters are listed in Table 3. Conditions: Microwave frequency, 8.772 GHz; magnetic field, 3030 Gauss; temperature, 6 K; microwave pulse width, 20 ns; $\tau = 310$ ns. (For interpretation of color mentioned in this figure legend the reader is referred to the web version of the article.)

Table 3

Simulation parameters for ESEEM from the remote ^{14}N imidazole nuclei in Cu(II)-bis-histamine complex. Euler angles, α , β , and γ , correspond to the relative orientation of the hf and nqi PAS, and Euler angles α' , β' , and γ' , correspond to the relative orientation of the two ^{14}N hf PAS.

Nuclear parameters	Values	Simultaneous confidence (0.99) region
A_{xx} (MHz)	2.13	2.10–2.15
A_{yy} (MHz)	1.81	1.77–1.84
A_{zz} (MHz)	0.854	0.822–0.886
e^2qQ/h (MHz)	1.51	1.50–1.52
η	0.668	0.653–0.683
α ($^\circ$)	7.28	–1.35 to 16.42
β ($^\circ$)	–9.34	–11.12 to –7.33
γ ($^\circ$)	–9.41	–24.02 to 6.01
α' ($^\circ$)	177.0	168.3 to 200.9
β' ($^\circ$)	32.5	29.6 to 35.5
γ' ($^\circ$)	–6.7	–17.9 to 16.6

$g = 2.069$ in this disordered sample. In the special case of this nominally octahedral Cu(II) complex, all orientations of the g -tensor can be assumed to contribute to the ESE at this g_{\perp} position of the EPR spectrum [36]. Powder simulation of the ESEEM leads to determination of the mutual orientation of the two remote imidazole ^{14}N shf PAS, based on reproduction of subtle spectral features, which exemplifies the efficiency of OPTESIM.

Fig. 5 shows the converged simulation of three-pulse ESEEM collected from the Cu(II)-bis-histamine complex. The simulation parameters are presented in Table 3. The simulation was achieved by using a sequential application of the genetic algorithm (10^4 evaluations) and Nelder–Mead simplex (10^3 iterations) algorithm. The simulation of the experimental ESEEM spectrum involved optimization on eleven shf and nqi parameters. Eight of the parameters, including the three principal components of the remote ^{14}N shf tensor, e^2qQ/h , η , and the three Euler angles that relate the shf and nqi PAS, are assumed to be common to the two ^{14}N nuclei. A significant feature of the FT of the experimental ESEEM, which is enhanced at the selected τ value of 310 ns, is the spectral feature around 8 MHz that arises from combination of the $\Delta m_{l,\beta} = \pm 2$, or “double quantum” modulation from the two coupled ^{14}N . The line shape of this $\Delta m_{l,\beta} = \pm 2$ combination line is sensitively dependent on the mutual orientation of the two remote imidazole ^{14}N shf PAS. This was accounted for in the simulations by varying Euler angles that relate the two remote imidazole ^{14}N shf PAS as part of the optimization. The values of these Euler angles cause extremely subtle changes in the overall ESEEM amplitude, and the parameter space has a large dimension. Therefore, manual fitting and gradient-based optimization methods are exceptionally time-consuming and impractical.

Optimization using the “hybrid” approach with a population size of 500 converged in about 10 h on the distributed network of fourteen Pentium 4 PCs. Table 1 displays the best fit values and simultaneous confidence intervals of the simulation parameters. The shf coupling tensor, nqi parameters, and Euler angles that relate the shf and nqi PAS are in good agreement with values reported for ESEEM-detected imidazole remote ^{14}N coupling in other Cu(II)-imidazole complexes [37,38]. The Euler angles for the relation of the two ^{14}N PAS in Table 3, $\alpha = 177^\circ$, $\beta = 32.5^\circ$, and $\gamma = -6.7^\circ$, indicate that the x - and y -axes are approximately opposed, and that the z -axes are approximately collinear, as depicted in Fig. 6A. The relatively high symmetry of the simulated mutual orientation of the remote ^{14}N shf tensors is consistent with the symmetry in the X-ray crystallographic structure of the Cu(II)(him) $_2$ (NO $_3$) $_2$ complex [34], which is shown in Fig. 6B. A more detailed interpretation of the OPTESIM results for the Cu(II)-bis-histamine complex, and other Cu(II)-imidazole complexes, will be presented in a separate work.

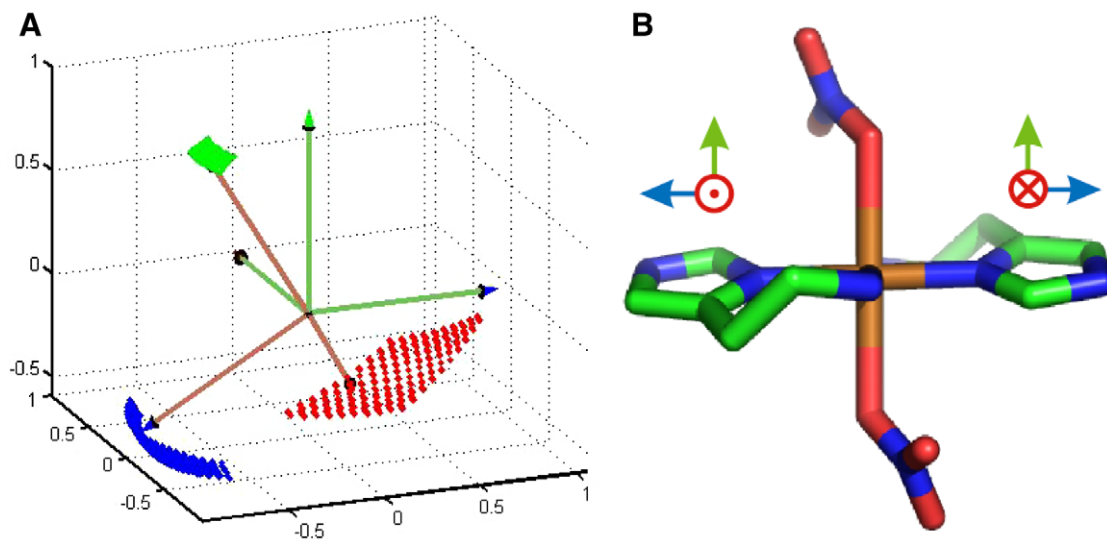


Fig. 6. Physical model for the mutual orientation of the imidazole remote ^{14}N hf PAS in the Cu(II)-bis-histamine complex, and X-ray crystallographic structure of Cu(II)(him) $_2$ (NO $_3$) $_2$ [34]. (A) Best-fit mutual orientation of the hf PAS of the imidazole remote ^{14}N nuclei, obtained by using OPTESIM. The dotted surfaces represent the simultaneous confidence regions of the axes, at level 0.99. (B) X-ray crystallographic structure of Cu(II)(him) $_2$ (NO $_3$) $_2$ [34]. Proposed Cartesian axes that correspond to the directions in (A) are shown.

5. Conclusions

OPTESIM is a Matlab-based toolbox that presents efficient solutions to the analysis of ESEEM by using numerical simulation. Automated, optimized simulations by using the toolbox lead to convergence on dramatically shorter time scales, relative to manual simulations. The efficiency is achieved by using an implementation of the fast algorithm [1,2,17] for translation of the spin Hamiltonian [10,11] into simulated ESEEM, and by the use of a hybrid, coarse-to-fine grain optimization approach. The statistical analysis of the simulation parameters allows the identification of simultaneous confidence regions at specific confident levels, which can be used to cull insignificant parameters to further speed calculations, and which supplies limits on the physical parameters (distances, angular orientations) that are used to construct molecular models from the ESEEM data. The geometry-preserving spherical averaging algorithm for $N > 1$ allows subtle information to be extracted about the relative orientations of the hf PAS, even in powder samples. These features of OPTESIM are exemplified by the applications for simulation of ^{14}N coupling for $N = 1$ and $N = 2$. OPTESIM is available for download from the following url: <http://www.physics.emory.edu/optesim>. The website includes a detailed description of the software, application instructions, and a tutorial.

References

- [1] A. Schweiger, G. Jeschke, Principles of Pulse Electron Paramagnetic Resonance, Oxford University Press, Oxford, UK; New York, 2001.
- [2] S.A. Dikanov, Y.D. Tsvetkov, Electron Spin Echo Envelope Modulation (ESEEM) Spectroscopy, CRC Press, Boca Raton, 1992.
- [3] C.P. Keijzers, E.J. Reijerse, P. Stam, M.F. Dumont, M.C.M. Gribnau, Magres – a general program for electron-spin-resonance, Endor and ESEEM, J. Chem. Soc. Faraday T. 1 83 (1987) 3493–3503.
- [4] A. Poppl, M. Gutjahr, M. Hartmann, W. Bohlmann, R. Bottcher, A study of the coordination geometry of the tetrapyrroline copper(II) complex in NaY and MCM-41 by two-dimensional electron spin resonance HYSCORE spectroscopy, J. Phys. Chem. B 102 (1998) 7752–7763.
- [5] A. Poppl, M. Hartmann, W. Bohlmann, R. Bottcher, Coordination geometry of the copper-pyridine complex in frozen solution as studied by proton and deuterium two-dimensional hyperfine sublevel correlation electron spin resonance spectroscopy, J. Phys. Chem. A 102 (1998) 3599–3606.
- [6] A. Poppl, L. Kevan, A practical strategy for determination of proton hyperfine interaction parameters in paramagnetic transition metal ion complexes by two-dimensional HYSCORE electron spin resonance spectroscopy in disordered systems, J. Phys. Chem-US 100 (1996) 3387–3394.
- [7] Z.L. Madi, S. Van Doorslaer, A. Schweiger, Numerical simulation of one- and two-dimensional ESEEM experiments, J. Magn. Reson. 154 (2002) 181–191.
- [8] J.J. Shane, L.P. Liesum, A. Schweiger, Efficient simulation of ESEEM spectra using gamma, J. Magn. Reson. 134 (1998) 72–75.
- [9] J.M. Canfield, K. Warncke, Geometry of reactant centers in the Co-II-substrate radical pair state of coenzyme B-12-dependent ethanolamine deaminase determined by using orientation-selection-ESEEM spectroscopy, J. Phys. Chem. B 106 (2002) 8831–8841.
- [10] W.B. Mims, Amplitudes of superhyperfine frequencies displayed in electron-spin echo envelope, Phys. Rev. B 6 (1972) 3543–3545.
- [11] W.B. Mims, Envelope modulation in spin-echo experiments, Phys. Rev. B 5 (1972) 2409–2419.
- [12] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley Pub. Co., Reading, Mass., 1989.
- [13] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 671–680.
- [14] J.A. Nelder, R. Mead, A simplex-method for function minimization, Computer J. 7 (1965) 308–313.
- [15] K. Levenberg, A method for the solution of certain problems in least squares, Quart. Appl. Math. 2 (1944) 164–168.
- [16] J.E. Wertz, J.R. Bolton, Electron Spin Resonance; Elementary Theory and Practical Applications, McGraw-Hill, New York, 1972.
- [17] S. Stoll, A. Schweiger, Rapid construction of solid-state magnetic resonance powder spectra from frequencies and amplitudes as applied to ESEEM, J. Magn. Reson. 163 (2003) 248–256.
- [18] G.R. Hanson, K.E. Gates, C.J. Noble, M. Griffin, A. Mitchell, S. Benson, XSophe-Sophe-XeprView (R). A computer simulation software suite (v. 1.1.3) for the analysis of continuous wave EPR spectra, J. Inorg. Biochem. 98 (2004) 903–916.
- [19] S. Stoll, A. Schweiger, EasySpin, a comprehensive software package for spectral simulation and analysis in EPR, J. Magn. Reson. 178 (2006) 42–55.
- [20] B.M. Hoffman, J. Martinsen, R.A. Venters, General-theory of polycrystalline Endor patterns – G and hyperfine tensors of arbitrary symmetry and relative orientation, J. Magn. Reson. 59 (1984) 110–123.
- [21] L.G. Rowan, E.L. Hahn, W.B. Mims, Electron-spin-echo envelope modulation, Phys. Rev. 137 (1965) A61–A71.
- [22] S.A. Dikanov, A.A. Shubin, V.N. Parmon, Modulation effects in the electron spin echo resulting from hyperfine interaction with a nucleus of an arbitrary spin, J. Magn. Reson. 42 (1981) 474–487.
- [23] J.M. Canfield, K. Warncke, Active site reactant center geometry in the Co-II-product radical pair state of coenzyme B-12-dependent ethanolamine deaminase determined by using orientation-selection electron spin-echo envelope modulation spectroscopy, J. Phys. Chem. B 109 (2005) 3053–3064.
- [24] L. Sun, O.A. Groover, J.M. Canfield, K. Warncke, Critical role of arginine 160 of the EutB protein subunit for active site structure and radical catalysis in coenzyme B-12-dependent ethanolamine ammonia-lyase, Biochemistry 47 (2008) 5523–5535.
- [25] K. Warncke, Characterization of the product radical structure in the Co-II-product radical pair state of coenzyme B-12-dependent ethanolamine deaminase by using three-pulse H-2 ESEEM spectroscopy, Biochemistry 44 (2005) 3184–3193.
- [26] Y.S. Zhu, Probability and Statistics in Experimental Physics, Science Press, Beijing, 2006.
- [27] K. Warncke, J. McCracken, Analysis of static distributions in hydrogen hyperfine interactions in randomly oriented radicals in the solid state by using 2H electron spin echo envelope modulation spectroscopy: conformational dispersion of b-2H coupling in the model tyrosyl radical, J. Chem. Phys. 103 (1995) 6829–6840.
- [28] S.L. Tan, M.G. Kopczyński, W.W. Bachovchin, W.H. Ormejohnson, B.M. Babior, Electron spin-echo studies of the composition of the paramagnetic intermediate formed during the deamination of propanolamine by ethanolamine ammonia-lyase, and adocbl-dependent enzyme, J. Biol. Chem. 261 (1986) 3483–3485.
- [29] S.C. Ke, K. Warncke, Interactions of substrate and product radicals with Co-II in cobalamin and with the active site in ethanolamine deaminase, characterized by ESE-EPR and N-14 ESEEM spectroscopies, J. Am. Chem. Soc. 121 (1999) 9922–9927.
- [30] W.B. Mims, J. Peisach, Nuclear modulation effect in electron-spin echoes for complexes of Cu2+ and imidazole with N-14 and N-15, J. Chem. Phys. 69 (1978) 4921–4930.
- [31] H.L. Flanagan, D.J. Singel, Analysis of N-14 ESEEM patterns of randomly oriented solids, J. Chem. Phys. 87 (1987) 5606–5616.
- [32] L. Sun, K. Warncke, Comparative model of EutB from coenzyme B-12-dependent ethanolamine ammonia-lyase reveals a beta(8)alpha(8), TIM-barrel fold and radical catalytic site structural features, Proteins 64 (2006) 308–319.
- [33] A. Latif Abuhijleh, C. Woods, I.Y. Ahmed, Synthesis and molecular structure of monomeric copper(II) acetates with 2-methylimidazole and 1,2-dimethylimidazole, Inorg. Chim. Acta 190 (1991) 11–17.
- [34] C. Zhang, C. Duan, Q. Hu, D. Yan, Self-assembly in transition metal complexes: structural characterization of a copper histamine nitrate [Cu(II)(him)2(NO3)2], J. Chem. Crystallogr. 29 (1999) 1153–1155.
- [35] D. Baute, D. Arieli, F. Neese, H. Zimmermann, B.M. Weckhuysen, D. Goldfarb, Carboxylate binding in copper histidine complexes in solution and in zeolite Y: X- and W-band pulsed EPR/ENDOR combined with DFT calculations, J. Am. Chem. Soc. 126 (2004) 11733–11745.
- [36] W. Weltner, Magnetic Atoms and Molecules, Dover Publications, New York, 1989.
- [37] M.J. Colaneri, J. Peisach, An electron spin-echo envelope modulation study of Cu(II)-doped single-crystals of L-histidine hydrochloride monohydrate, J. Am. Chem. Soc. 114 (1992) 5335–5341.
- [38] M.J. Colaneri, J. Peisach, A single-crystal EPR and ESEEM analysis of Cu(II)-doped bis(L-histidinato)cadmium dihydrate, J. Am. Chem. Soc. 117 (1995) 6308–6315.