Anna G. Matveeva, Yulia V. Yushkova, Sergei V. Morozov, Igor A. Grygor'ev and Sergei A. Dzuba*

Multi-Gaussian Monte Carlo Analysis of PELDOR Data in the Frequency Domain

DOI 10.1515/zpch-2016-0830
Received June 17, 2016; accepted October 6, 2016

Abstract: Pulsed double electron–electron resonance technique (PELDOR or DEER) is often applied to study conformations and aggregation of spin-labelled macromolecules. Because of the ill-posed nature of the integral equation determining the distance distribution function, a regularization procedure is required to restrict the smoothness of the solution. In this work, we performed PELDOR measurements for new flexible nitroxide biradicals based on trolox, which is the synthetic analogue of α-tocopherol; spin-labelled trolox derivatives are investigated as potential anti-cancer drugs. We use regularization by an approximation of the solution with a sum of limited number of Gaussians, by varying their positions, widths and amplitudes. Their best-fitted values were found by a completely random Monte Carlo process. The use of the frequency-domain PELDOR data allowed diminution of the artifacts induced by spin–spin electron–nuclear and intermolecular electron–electron interactions. It was found that for the all biradicals studied, the use of three Gaussians was enough for good agreement with the experiments. The number of trials for obtaining satisfactory result was found to be quite reasonable, which is explained by presence of the singularity in the core of integral equation. The maxima of inter-spin distance distribution for different biradicals were found to vary between 1.5 and 2.3 nm, depending on the linkers between the Trolox core and nitroxides. The distance distributions around these positions reflect flexibility of the biradicals.

***Corresponding author: Sergei A. Dzuba,** Voevodsky Institute of Chemical Kinetics and Combustion, Novosibirsk 630090, Russia; and Novosibirsk State University, Novosibirsk 630090, Russia, e-mail: dzuba@kinetics.nsc.ru
Anna G. Matveeva: Voevodsky Institute of Chemical Kinetics and Combustion, Novosibirsk 630090, Russia; and Novosibirsk State University, Novosibirsk 630090, Russia
Yulia V. Yushkova and Igor A. Grygor'ev: Vorozhtsov Novosibirsk Institute of Organic Chemistry, Novosibirsk 630090, Russia
Sergei V. Morozov: Novosibirsk State University, Novosibirsk 630090, Russia; and Vorozhtsov Novosibirsk Institute of Organic Chemistry, Novosibirsk 630090, Russia

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

Keywords: biradicals; DEER; DNA; Fredholm equation regularization; Monte Carlo fitting; peptide antibiotics; RNA; spin labels; Trolox.

Dedicated to: Kev Salikhov on the occasion of his 80th birthday.

1 Introduction

Pulsed double electron–electron resonance technique (DEER, also known as PELDOR) [1, 2] allows measurement of the distance distribution function, $P(r)$, for distances between two paramagnetic centers. For doubly spin-labelled molecules (biradicals), PELDOR data provides information on their conformational flexibility. Some recent applications of the technique are summarized in reviews [3–10].

The $P(r)$ function can be obtained in the range between 1.5 and 8 nm, as a solution of a first-kind Fredholm integral equation [11–20]. Solving of this equation, however, is an ill-posed problem: slight variation of the input experimental data may result in a large variation of the solution. Therefore, some *a priori* properties of the solution must be postulated, which is called regularization of the solution; the most commonly employed approach is the Tikhonov regularization. However, in some cases the existing approaches may produce artifacts – e.g. when narrow and broad distributions are superimposed [16, 17]. Regularization can also be done in other ways – by approximation of the solution as a sum of several Gaussian distributions (multi-Gaussian fit) [21–23], or by increasing the discretization length up to the threshold where numerical calculation of the integral still provides agreement with the experimental data [24, 25].

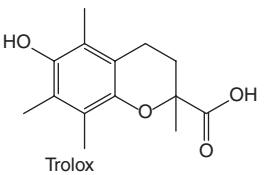
In this work, we show that a multi-Gaussian fit [21–23] can be satisfactorily applied to study flexible synthetic biradicals. We use here exclusively a random Monte Carlo procedure for searching for the best-fitted solution, without the downhill simplex procedure employed in [21–23], so making the searching scheme extremely simple and easily programmable. Also, because the consequent trials are uncorrelated, trapping in the local minima is automatically avoided. Another point of the approach employed here is the use of the frequency-domain experimental PELDOR data, which allows evaluation of the initial fitting parameters and diminishes the artifacts induced by spin–spin electron–nuclear and intermolecular electron–electron interactions.

At first glance, the number of Monte Carlo trials must be extremely large, because for N Gaussians $3N$ independent parameters must be found – their positions, widths, and amplitudes ($3N-1$ after normalization). However, the contribution of different distances when solving the integral equation is essentially

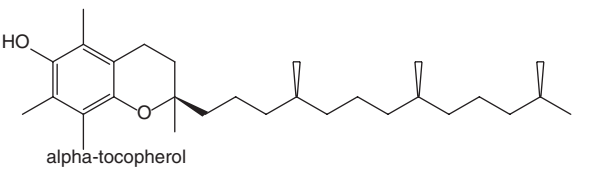
Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

non-equivalent because of the singularity of the kernel of the frequency-domain integral equation. Therefore the number of trials may be reasonable.

The biradicals under study are based on Trolox – the synthetic analogue of α-tocopherol (see the scheme below). α-Tocopherol has many important functions; in particular it serves as a universal protector of cell membranes from oxidative damage. Therefore, its derivatives may also have physiological activity. Spin-labelled Trolox derivatives are nowadays investigated as potential anti-cancer drugs [26]. These studies are focused on the possible ability of nitroxide-containing Trolox derivatives to protect bacterial cells from spontaneous and peroxide-induced mutagenesis and on their cytotoxicity against different tumors [26].

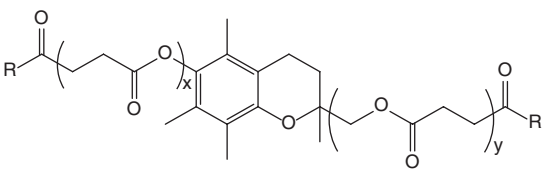


Trolox



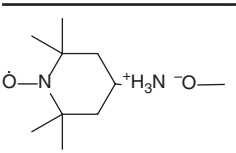
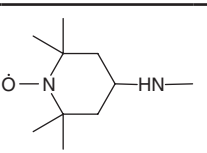
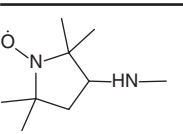
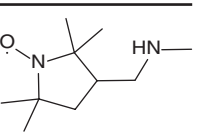
alpha-tocopherol

The structures of biradicals studied in this work are



where R is a spin label, and x (= 0 or 1) and y (= 0 or 1) indicate the presence of additional inserts. The substances studied in this work are listed in Table 1. For convenience, all the substances are grouped by the type of spin label (indicated by a Roman number). The presence of additional inserts is marked as two Arabic numbers.

Tab. 1: Structures of biradicals studied in this work.

			
I	**II**	**III**	**IV**
x = 1; y = 0	x = 1; y = 0	x = 1; y = 0	x = 1; y = 0
I-10	**II-10**	**III-10**	**IV-10**
x = 0; y = 1	x = 1; y = 1	x = 1; y = 1	x = 1; y = 1
I-01	**II-11**	**III-11**	**IV-11**

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

Some of the studied compounds (**I-10** and **I-01**) were synthesized in [27]; others were originally synthesized. Note that **I-10** and **I-01** are salt hybrids, and the six others are covalently bonded true biradicals.

2 PELDOR data analysis

PELDOR technique is based on electron spin echo (ESE) spectroscopy – a pulsed version of electron paramagnetic resonance (EPR); PELDOR signal is the ESE signal diminishing under application of an additional microwave pumping pulse, with a scanning delay T [1–10]. For doubly spin-labelled molecules, the PELDOR time trace depends on two contributions: the intramolecular one, arising from interactions between two spin labels in the molecule, and the intermolecular one, arising from interactions between labels in different molecules. These two contributions can be assumed to be independent so that the PELDOR time trace is presented as a product:

$$V(T) = V_{\text{INTRA}}(T)V_{\text{INTER}}(T) \tag{1}$$

$V_{\text{INTER}}(T)$ often obeys a simple exponential dependence,

$$V_{\text{INTER}}(T) = V_0 \exp(-\text{const} * T) \tag{2}$$

For $V_{\text{INTRA}}(T)$, a normalized form,

$$V_N(T) = \frac{V_{\text{INTRA}}(T) - V_{\text{INTRA}}(\infty)}{V_{\text{INTRA}}(0) - V_{\text{INTRA}}(\infty)}, \tag{3}$$

is obtained. It is assumed here that $V_{\text{INTRA}}(T)$ attains a constant value $V_{\text{INTRA}}(\infty)$ at large T. Note that ratio $(V_{\text{INTRA}}(0) - V_{\text{INTRA}}(\infty))/V_{\text{INTRA}}(0)$ is the efficiency of the pumping pulse action, the dimensionless value p_B [2, 10].

The theory [2–10] predicts that cosine Fourier transform of Eq. (3), $F(\nu)$, may be presented as

$$F(\nu) = \int_0^\infty K(\nu, r)P(r)dr, \tag{4}$$

where $P(r)$ is the distance distribution function between the two spin labels in the molecule, which is assumed to be normalized, $\int_0^\infty P(r)dr = 1$, and

$$K(\nu, r) = 2\int_0^\infty \cos(2\pi\nu T)dT \int_0^{\pi/2} \sin\theta d\theta \cos\left(2\pi \frac{A}{r^3}(1 - 3\cos^2\theta)T\right), \tag{5}$$

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

where A is a constant which for nitroxide spin labels is close to 52.4 MHz $*$ nm^3. $F(\nu)$ is a frequency-domain PELDOR lineshape, which is called a dipolar Pake (resonance pattern) spectrum. Note that $F(\nu)$ is also normalized, $\int_{-\infty}^{\infty} F(\nu)d\nu = 1$.

In most applications, a PELDOR experiment is analysed in the time domain, for data presented by Eq. (3) [2–10]. However, the experimentally obtained Pake spectrum given by Eq. (4) is also eligible for the analysis [11, 13, 15]. The frequency-domain data may have the advantages that [24, 25] (i) the direct assessment is possible of the distance interval, from r_{min} to r_{max}, in which $P(r)$ has significant values; (ii) artifacts appearing in ESE signal formation because of the intervention of electron–nuclear envelope modulation (ESEEM) can be avoided, because these artifacts result in peaks at known frequencies (near ± 14 MHz for proton-induced ESEEM); and (iii) imperfect background correction can be clearly seen as an artifact at zero frequency so it also can be disregarded in the fitting process.

Note that r_{min} and r_{max} can be assessed from the relations $r_{min(max)}/\text{nm} = (52\ \text{MHz}/\nu_{max(min)})^{1/3}$, where ν_{max} is the high-frequency boundary of the Pake spectrum, and ν_{min} is taken slightly below the low-frequency maximum of the spectrum [25], with the exact $r_{min(max)}$ values corrected *a posteriori* by comparing the experimental and simulated Pake spectra.

In order to find the distance distribution function $P(r)$, we used approximation of $P(r)$ by several Gaussian distributions [21–23]. Each Gaussian component is determined by the position of the maximum, the width and the relative weight, so the trial distance distribution function is constructed as

$$P_{trial}(r) = \sum_{i=1}^{N} a_i \frac{1}{\sqrt{2\pi}\delta_i} \exp\left(-\frac{(r-r_i)^2}{2\delta_i^2}\right), \tag{6}$$

with $\sum_{i=1}^{N} a_i = 1$, where N is the number of Gaussians. The sets of a_i, δ_i, r_i are random values. The a_i values vary between 0 and 1, the intervals for variation of δ_i and r_i can be estimated from the frequency-domain PELDOR spectra [25]. It is assumed in Eq. (6) that $\delta_i << r_i$, which is normally fulfilled in PELDOR applications.

For each $P_{trial}(r)$ function, the trial Pake spectrum $F_{trial}(\nu)$ is calculated by numerical integration of Eq. (4). Then the mean-squared deviations (MSD),

$$\text{MSD} = \frac{1}{J}\sum_{j=1}^{J}(F_{trial}(\nu_j) - F(\nu_j))^2, \tag{7}$$

between the calculated $F_{trial}(\nu_j)$ and the experimental $F(\nu_j)$ values are determined, where J is the number of points sampled in the frequency domain (it was between

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

30 and 60 in this work). Different trials are compared and the distribution $P_{trial}(r)$ providing the smallest MSD value is selected as a solution.

In the computer calculations, the program performed 10^6–10^7 trials which took ~1 min on an ordinary PC.

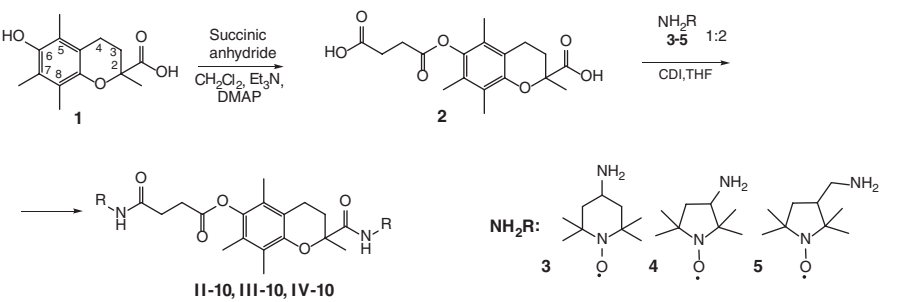
3 Experimental

3.1 Synthesis

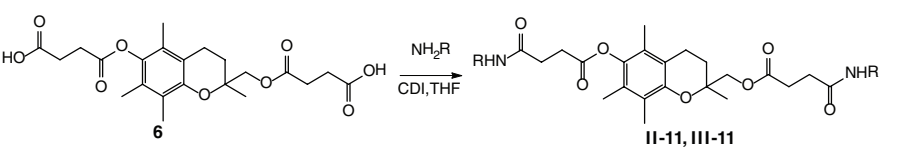
Biradicals **II-10**, **III-10** and **IV-10** were synthesized by reacting succinate Trolox **2** with nitroxyl radicals **3–5** in the ratio 1 : 2 in tetrahydrofuran (THF) in the presence of a dehydrating agent carbonyldiimidazole (CDI) (Scheme 1) [26].

Spin-labelled diamides **II-11, III-11** were obtained by reaction of disuccinate 2-hydroxymethyl derivative Trolox **6** [27] with nitroxyl radicals **3** and **4** taken in the ratio of 1 : 2 in THF in the presence of CDI (Scheme 2).

Compounds **II-11** and **III-11** in Scheme 2 were synthesized in the same way as biradicals **II-10, III-10, IV-10** in Scheme 1. For synthesis of biradical **IV-11**, disuccinate **6** was at first converted to succinamide **7** by reaction with N-hydroxysuccinimide (NHS) in THF in the presence of CDI in the ratio of 1 : 2 [28], then further reacted with radical **5** in CH_2Cl_2 in the ratio 1 : 2, resulting in biradical **IV-11** (Scheme 3).

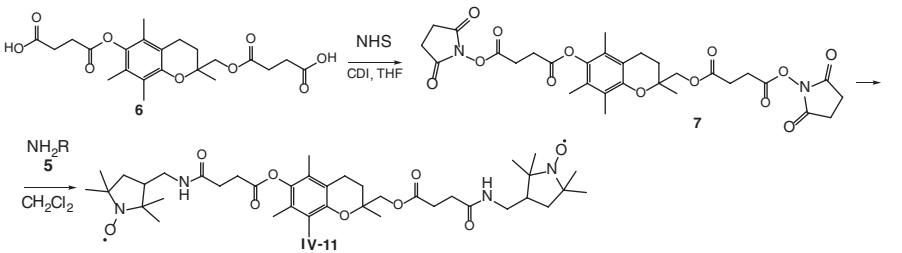


Scheme 1: Synthesis of biradicals **II-10, III-10, IV-10**.



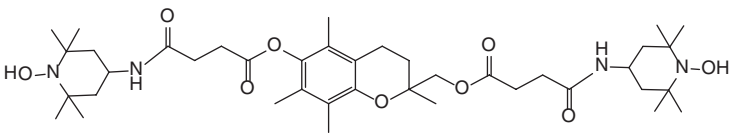
Scheme 2: Synthesis of biradicals **II-11, III-11.**

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM



Scheme 3: Synthesis of biradical **IV-11**.

To confirm the structure of biradical **II-11** by NMR, the corresponding hydroxylamine **II-11a** (see structure below) was obtained by reduction of **II-11** in CD_3OD with zinc in the presence of NH_4Cl [29].



II-11a

In the ^1H NMR spectrum of hydroxylamine **II-11a**, all the characteristic signals of succinyl group, hydroxypiperidine and Trolox fragments were observed. Signals of hydroxypiperidine fragment methyl groups had chemical shifts in ppm: 1.17–1.20. Signals of the methyl group in the phenolic moiety appeared at 1.94–2.04 ppm and a CH_3-2 signal at 1.33 ppm. Signals of CH_2-groups of succinyl fragments appeared in the form of three triplets at 2.49, 2.61 and 2.96 ppm with the constant $J = 6.6$ Hz and at 2.65–2.71 ppm; a CH_2-2 signal appeared at 4.07–4.26 ppm. The purity of compounds **II-11, III-11, IV-11** was checked by thin-layer and high-performance liquid chromatography methods.

Other details of the synthesis may be found in [26, 27].

3.2 Sample preparation

Salt hybrids and biradicals were dissolved in either methanol or methyltetrahydrofuran (MTF) (both from Ekros-Analytica, St. Petersburg, Russia). Solutions were put in EPR tubes and then frozen by immersion into liquid nitrogen to form a transparent glass. For some experiments, a solution of monoradical **II-H** (compound **3** in Scheme 1) was also prepared.

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

3.3 EPR and PELDOR measurements

An X-band Bruker ELEXSYS E580 EPR spectrometer was used. PELDOR measurements were carried out using a split-ring Bruker ER 4118 X-MS-3 resonator while a Bruker ER 4118X-MD5 dielectric ring resonator was used in continuous wave (CW) EPR spectra measurements. A three-pulse PELDOR setup was employed. The pumping pulse length was 28 ns, and the lengths of the $\pi/2$ and π pulses of the echo-forming detection pulse sequence were 16 and 32 ns, respectively. The time delay between two detection pulses was 800 ns. The pumping pulse was applied in all cases at the frequency ν_B corresponding to the maximum of the echo-detected EPR spectrum and was scanned with a step of 8 ns starting from the negative initial time delay $d_0 = -200$ ns, respective to the first detection pulse. The turning angle of the pumping pulse was set to π in measurements in which the observation frequency ν_A was set equal to ν_B, the pumping pulse was positioned at time delay d_0, and pumping pulse amplitude was varied to adjust the inverted echo signal to its minimum. The starting delay T for the PELDOR time trace analysis ($T = 0$) was determined as described in Ref. [30]. In PELDOR experiments, the difference $\nu_A - \nu_B$ between the detection and pumping frequencies could be varied between 50 and 75 MHz. The in-phase part of the primary echo was integrated with a time gate of 80 ns. The changes in the PELDOR signal $V(T)$ upon passage of the pumping pulse through the detecting pulses were corrected by the method described in Ref. [31].

The resonator was cooled in an Oxford Instruments CF-935 cryostat with flowing cold gaseous nitrogen. The sample temperature was kept near 120 K in CW EPR measurements and near 80 K in PELDOR measurements.

4 Results

4.1 CW EPR spectra

Typical representatives of CW EPR spectra are shown in Figure 1. These spectra are normalized to the same value of the double integral. One can see that the spectra for **I-10** and **II-11** biradicals are slightly broader than the spectrum for monoradical **II-H**. This may be readily ascribed to line broadening due to dipole–dipole and/or exchange interactions in the biradicals. However, as the broadening is small, these interactions are obviously weak in the CW EPR experimental window.

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

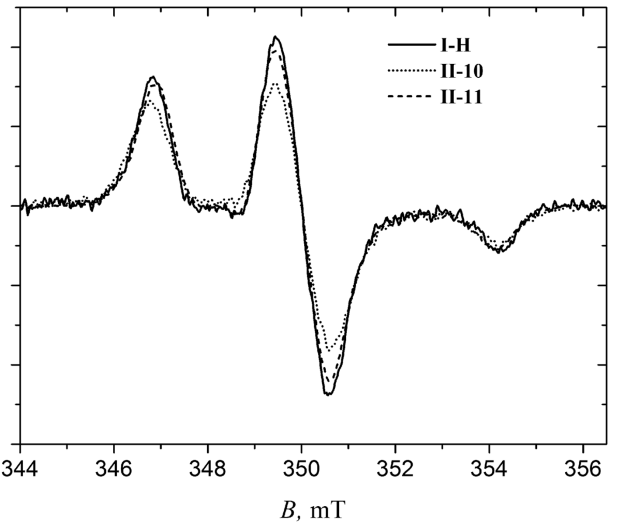


Fig. 1: CW EPR spectra of 1.3×10^{-3} M **I-10** and **II-11** biradicals in methanol at 120 K, normalized to the same integral intensity. The spectrum of monoradical **II-H** recorded at the same conditions is given for comparison.

4.2 PELDOR data for salt hybrids I-10 and I-01

Figure 2 shows on a semilogarithmic scale the PELDOR time traces for salt hybrid **I-10** and **I-01** samples. (The difference $\nu_A - \nu_B$ is 70 MHz.) PELDOR data for covalently-linked biradical **II-11** is also given here. One can see that both salt hybrids show simple exponential decay that is consistent with Eq. (2), so only the intermolecular contribution is present here. The absence of intramolecular contribution unambiguously evidences that the substances perfectly dissociate in the solvent, producing only the corresponding **I** monoradicals. Note that similar results were obtained for both solvents, methanol and MTF.

4.3 PELDOR data for covalently-linked biradicals

It is clearly seen in Figure 2 that for the **II-11** biradical a fast component appears in addition to the slow exponential decay. The PELDOR time trace may be described as a product of intermolecular $V_{\mathrm{INTER}}(T)$ and intramolecular $V_{\mathrm{INTRA}}(T)$ components that is in line with Eq. (1). Analogous results were obtained for all other covalently linked biradicals (data not given). In all cases, the asymptotical $V_{\mathrm{INTER}}(T)$ behavior could be approximated as purely exponential, i.e. fitted by a straight line on a

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

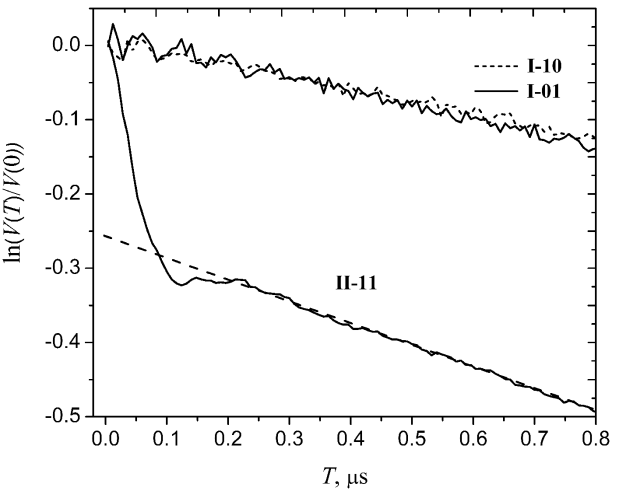


Fig. 2: PELDOR time traces for 10^{-3} M solutions of **I-01**, **I-01** and **II-11** in methanol on a semilogarithmic scale. The difference $\nu_A - \nu_B$ is 70 MHz. The dashed straight line indicates an exponential asymptotic for **II-11**.

semilogarithmic scale (like that shown in Figure 2), so the fast component was extracted for all biradicals from the experimental time traces by subtracting these straight lines.

After this subtraction, background-corrected $V_{\mathrm{INTRA}}(T)$ time traces were obtained, which were then normalized according to Eq. (3) to produce the $V_N(T)$ time traces. These time traces were Fourier transformed, real (cosine) part was taken and so the Pake spectra were obtained – see Figure 3 (solid lines). In the insert to Figure 3, the input $V_N(T)$ time traces are given on a normal linear scale for the all samples studied (solid lines). One can see that $V_N(T)$ at large T attain asymptotic zero values.

The minor peaks seen in Figure 3 around ± 14 MHz are induced by ESEEM associated with electron–nuclear interaction between the unpaired electron of the spin label and the nearby protons. To avoid undesirable influence of this effect, in Eq. (7) the vicinity of these peaks was omitted. The peaks in the center of the spectrum, seen in the range ± 1 MHz, are induced by imperfect elimination of the background time trace $V_{\mathrm{INTER}}(T)$. These frequencies were also omitted when applying Eq. (7).

4.4 Data analysis with multi-Gaussian fit in a completely random Monte Carlo process

In our calculations, we tried to use the number N of Gaussians equal consequently to 2, 3 and 4. The obtained distance distributions for biradicals **II-10**

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

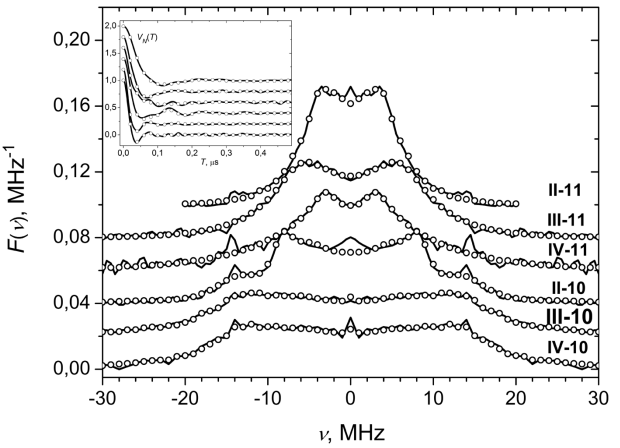


Fig. 3: Pake spectra – cosine Fourier transforms of the time-domain normalised $V_N(T)$ PELDOR data (solid lines), and their simulation (circles) using Eqs. (4–6) for the best-fitted $P_{trial}(r)$ functions found by Monte Carlo multi-Gaussian fitting. Spectra are consequently shifted along the vertical axis by a step of 0.02 MHz^{-1}. The input $V_N(T)$ data are given in the insert (solid lines), with the results of simulations (circles); the data are consequently shifted along the vertical axis by a step of 0.2, in the same order as the Pake spectra. The difference $\nu_A - \nu_B$ is 70 MHz.

are shown in Figure 4, along with results of fitting the Pake spectra (shown in insert). The calculated MSDs were 9.1 10^{-6} MHz^{-2} (for $N=2$), 1.40 10^{-6} MHz^{-2} ($N=3$), and 1.36 10^{-6} MHz^{-2} ($N=4$). For $N=2$ the agreement between experimental and calculated Pake spectra is not satisfactorily: except of large value of the attained MSD, also the peculiarity near the 15 MHz is smeared – see insert to Figure 4. By other words, the experimental data are underfitted in this case. For $N=3$ and $N=4$ a rather good agreement could be achieved, with MSD of the order of the noise level (see below). However, for $N=4$ some instability of the solution was found – different parameters for Gaussians provided similar results in the Pake lineshape (data not given), that is obviously because of the ill-posed nature of the mathematical problem and of the large number of the input parameters. So, 4 Gaussians overfit the experimental spectrum. Note however that the positions of maxima are not influenced by this overfitting (see Figure 4).

Figure 5 show how rapidly the Monte Carlo process converges, for the **II-10** biradical taken as an example. Here, the best fitted distribution functions are shown (the left panel) which were obtained for five different independent realizations, along with the MSD values achieved (the right panel). One can see that 10^6–10^7 runs allow to obtain reproducible results, with the MSD reaching an asymptotic constant value.

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

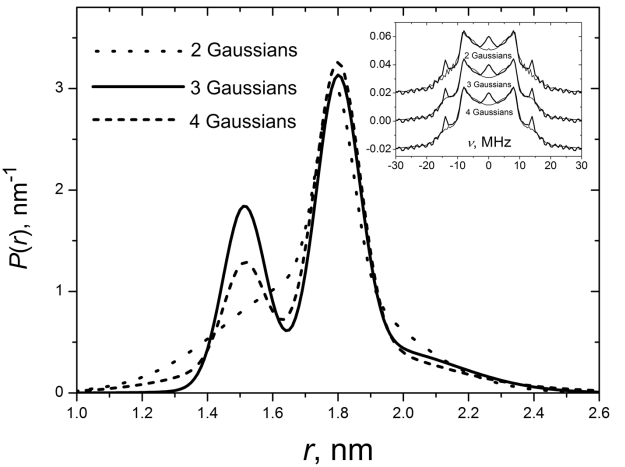


Fig. 4: The best-fitted simulations of the distance distributions obtained for the **II-10** biradical using Monte Carlo multi-Gaussian fitting with indicated different number of Gaussians. The insert shows how these fittings simulate the experimental Pake spectrum. (thin and thick lines, respectively, data are shifted along the vertical axis by a step of 0.02 MHz^{-1}). The number of runs is 10^7 in all cases.

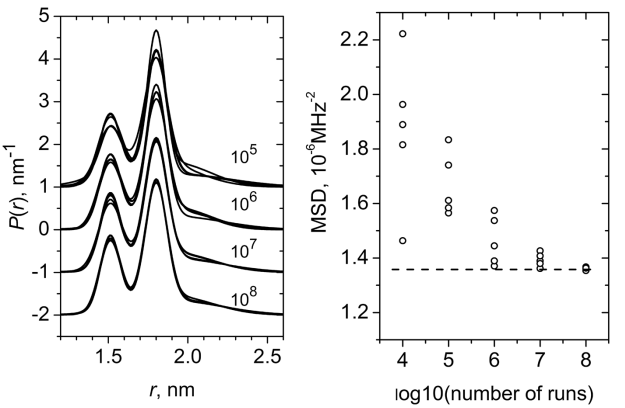


Fig. 5: The illustration of convergence of Monte Carlo process, for the case of biradical **II-10** and 3-Gaussian fitting. Results of five independent realizations are shown for different number of runs indicated, for the distribution function (left panel), and MSD (right panel). Data in left panel are shifted along the vertical axis by a step of 1 nm^{-1}. The dashed line in the right panel shows an asymptotic constant value.

Note that convergence of the Monte Carlo process when fitting the integral equation (4) to a large extent is determined by the singularity in the core $K(\nu, r)$. This singularity makes the contribution of different distances essentially

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

non-equivalent so that each trial "selects" mostly only the Gaussian that is located closer to this singularity, and making the other Gaussians less important. So the number of independent parameters in the Monte Carlo search is effectively reduced.

The asymptotic value in Figure 5 (right) is determined by experimental noise. Note that it is not influenced by the restrictions imposed in the multi-Gaussian approximation – because 3- and 4-Gaussian fits for 10^7 runs provide nearly the same MSD value (see above).

Data in Figure 5 (left) may be employed for assessment of the accuracy how the distribution function can be restored from the real PELDOR experiment. All data here provides the MSD values (see Figure 5 (right)) of the order of experimental noise. The simulations show that the positions of both maxima vary only slightly either for different number of Gaussians employed (3 or 4) – see Figure 4, or for different number of trials – see Figure 5. From data presented in Figures 4 and 5 we estimate the possible variation of the maxima positions as ± 0.005 nm. Meanwhile the data in Figure 4 show that the peak widths may vary by $\pm 10\%$ from their mean values.

To assess the variability of results in various applications, one should compare results of calculations for different number of Gaussians and different number of trials, like it is shown in Figures 4 and 5, having in mind also the MSD level determined by the experimental noise.

To additionally confirm results of the used multi-Gaussian fit approach, we performed also calculations using the traditional Tikhonov regularization method, employing the DeerAnalysis software [18]. In addition, the recently proposed approach [25] which employs increasing the discretization length up to the threshold where numerical calculation of the integral (4) still provides agreement with experiment was also employed. The comparison of these three different approaches for the **II-10** biradical is shown in Figure 6. As in the DeerAnalysis the L-curve for this biradical was found to be smooth (data not given), the optimal regularization parameter was chosen manually by its adjusting for the best coincidence of the experimental and simulated Pake spectra, accompanying with the best coincidence of the obtained $P(r)$ distribution with that found here. The insert to Figure 6 shows comparison of the experimental and simulated Pake spectra in all these three cases. One can see that data in Figure 6 show the similar distance distributions for the all three approaches.

Results of multi-Gaussian fitting for the all series of biradicals studied are shown in Figures 3 and 7. Three-Gaussian fitting ($N=3$) was employed for all cases, with number of trials 10^7. The circles shown in Figure 3 are the approximation of the Pake spectra obtained using the proposed algorithm. (The circles in the insert to Figure 3 shows the analogous time-domain data). One can see a

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

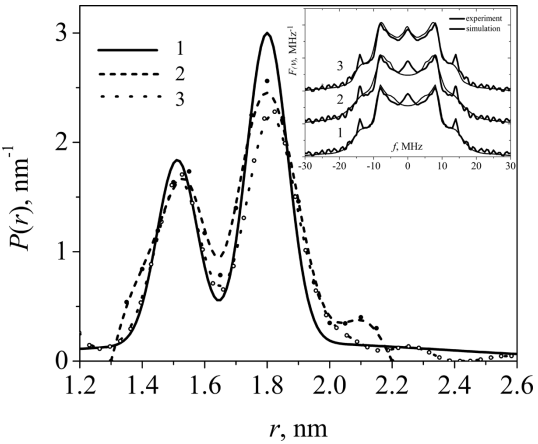


Fig. 6: The best-fitted simulations of the distance distributions obtained for the **II-10** biradical using different approaches: (1) Monte Carlo three-Gaussians fitting as suggested here (the same data as in Figure 4), (2) the increasing of the discretization length in line with [25], (3) DeerAnalysis with Tikhonov regularization algorithm [18]. The insert shows how these approaches simulate (thin lines) the experimental (thick lines) Pake spectrum.

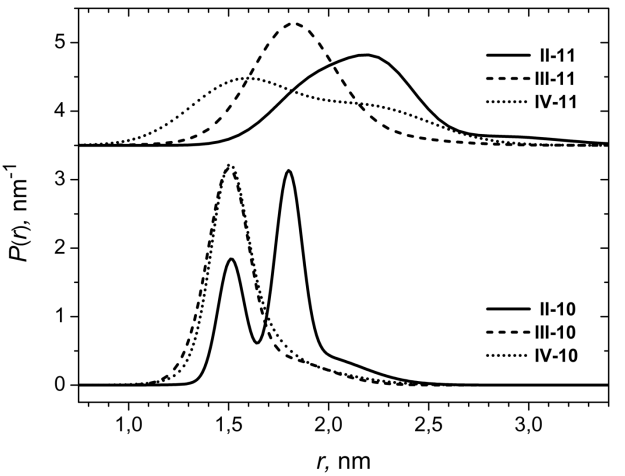


Fig. 7: The best-fitted $P_{trial}(r)$ distributions found in the Monte Carlo search described in the text. The upper curves correspond to 'long' biradicals ($x + y = 2$); the lower curves correspond to 'short' biradicals ($x + y = 1$). The upper curves are shifted along the vertical axis by 3.5 nm^{-1}.

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

rather good agreement between experiment and simulations. The best-fitted distribution functions $P_{trial}(r)$ are shown in Figure 7.

One more caution with the obtained results may be related with the phenomenon of orientational selectivity in PELDOR experiments [32] – when PELDOR data depend on the mutual orientation of the spin labels in the biradicals. We performed orientation selection experiments in which the difference $\nu_A - \nu_B$ between the detection and pumping frequencies was varied between 50 and 75 MHz – see data in Figure 8 (biradicals **II-10)**. The time-domain data in Figure 8 (left panel) show that for different $\nu_A - \nu_B$ values the obtained results are similar, except of small variation of the depth of the first minimum. However, the Pake spectra (right panel) show that this variation is induced by different contribution of the proton ESEEM around ±14 MHz. Some variations in the central part of the spectra could be attributed to the uncertainty in background correction when refining the pure $V_{INTRA}(T)$ contribution. Anyhow, the shoulders near ±15 MHz are present in all the spectra; this shoulder reflects the maximum of the $P(r)$ distribution seen at 1.5 nm in Figure 7 for this biradical.

Also, the efficiency of the pumping pulse action is almost independent on the $\nu_A - \nu_B$ value – see insert to Figure 8. So we conclude the orientation selection is not critical for our systems. As orientation selection is observed when both nitroxides in a biradicals are strongly spatially restricted [32], this result means that in our systems these nitroxides are spatially distributed.

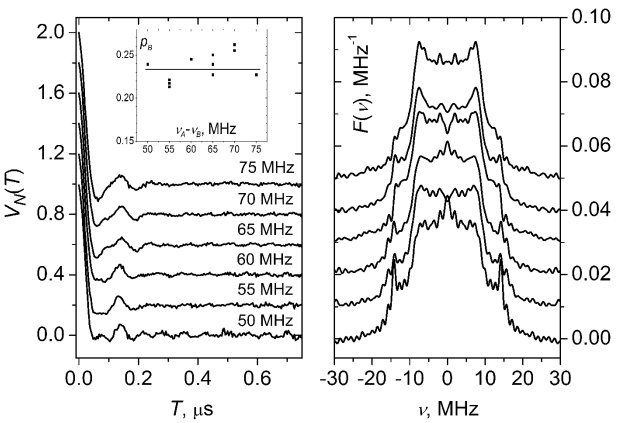


Fig. 8: The time domain (left panel) and the corresponding frequency-domain (right panel) PELDOR data obtained at different difference $\nu_A - \nu_B$ between the detection and pumping frequencies (biradicals **II-10**). Data are consequently shifted along the vertical axis by 0.2 (left) and by 0.01 MHz^{-1} (right). In the insert the efficiency of the pumping pulse action p_B is given for different $\nu_A - \nu_B$ values.

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

5 Discussion

One can see in Figure 7 that for short biradicals ($x + y = 1$) the distribution is narrower than that for the long biradicals ($x + y = 2$). The other point worth mentioning is that a change of the type of spin label may have a dramatic impact on the distribution function, both for short and for long biradicals (except for the cases of **III-10** and **IV-10**).

Note that the fact that the distribution function could be well approximated by three Gaussians does not mean that it contained three separate components. In some cases, such as for biradicals **III-11** and **IV-10**, distance distribution clearly consisted of a single asymmetrical peak that was merely well reproduced on the basis of the three Gaussians.

Figure 7 shows extended background line in the distance distributions $P(r)$ for biradical **II-11**. This line can be ascribed to imperfect elimination of the $V_{INTER}(T)$ contribution in Eq. (1) by the employed simple exponential approximation. However, this effect is small and may be neglected.

Data in Figure 7 show maxima located at the distances of 1.5 nm (biradicals **II-10**, **III-10**, **IV-10**), 1.6 nm (**IV-11**), 1.8 nm (**II-10**, **III-11**), and 2.2–2.3 nm (**II-11**, **IV-11**). The chemical structure of biradicals predicts that 1.5–1.8 nm for short biradicals corresponds to fully extended conformations. For long biradicals, the same is valid for 2.2–2.3 nm. Distance distributions around these positions reflect flexibility of the biradicals. Biradicals **II-10**, **II-11** and **IV-11** obviously possess several conformations. (For **II-10**, presence of two conformations is clearly seen even in Figure 3 from the Pake spectra). Comparing data for short biradicals **II-10** with **III-10** and **IV-10**, one may notice that this flexibility is larger for the 6-member-ring nitroxides then for the 5-member-ring ones. More precise analysis of biradicals conformations is possible applying computational methods of molecular mechanics.

6 Conclusions

In this work, PELDOR measurements were performed for six new flexible nitroxide Trolox-based biradicals. The PELDOR data analysis was based on regularization of an integral equation solution by its restriction to a sum of a limited number of Gaussians. The positions, widths and amplitudes of Gaussians were found by fitting the frequency-domain PELDOR spectra in a completely random Monte Carlo process. In the employed fitting procedure, sticking to local minima is avoided automatically because the consequent Monte Carlo trials are uncorre-

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

lated. The number of Monte Carlo trials was found to be quite reasonable, which is explained by the singularity in the kernel of the frequency-domain integral equation.

It was found that in all cases the use of only three Gaussians was enough for good agreement with the experiment – for a large variety of different nitroxide biradicals. The suggested approach also allows estimation of the uncertainty of obtained parameters induced by experimental noise. To get the solution, the program performed 10^6–10^7 trials which took ~ 1 min on an ordinary PC.

To assess the variability of results, one should compare calculations for different number of Gaussians and different number of trials, having in mind also the MSD level determined by the experimental noise.

Although the employed multi-Gaussian fit is not a model-free approximation, it nevertheless could be applied for other doubly spin-labelled molecules: proteins, DNA, RNA and peptide antibiotics. The only criterion for checking its validity would be good agreement in each particular case between the experimental and simulated data, attained within the experimental noise. The number of Gaussians must be meanwhile reasonably small – three or four, – otherwise the solution may become unstable because of the ill-posed nature of the mathematical problem.

Acknowledgments: This work was supported by the Russian Science Foundation, project # 15-15-00021.

References

1. A. D. Milov, K. M. Salikhov, M. D. Schirov, Fiz. Tverd. Tela **23** (1981) 975.
2. A. D. Milov, A. G. Maryasov, Yu. D. Tsvetkov, Appl. Magn. Reson. **15** (1998) 107.
3. O. Schiemann, T. F. Prisner, Quart. Rev. Biophys. **40** (2007) 1.
4. A. Savitsky, K. Möbius, Photosynth. Res. **102** (2009) 311.
5. J. P. Klare, H.-J. Steinhoff, Photosynth. Res. **102** (2009) 377.
6. G. W. Reginsson, O. Schiemann, Biochem. J. **434** (2011) 353.
7. G. Jeschke, Ann. Rev. Phys. Chem. **63** (2012) 419.
8. D. Goldfarb, Phys. Chem. Chem. Phys. **16** (2014) 9685.
9. T. F. Prisner, A. Marko, S. Th. Sigurdsson, J. Magn. Reson. **252** (2015) 187.
10. A. D. Milov, Y. D. Tsvetkov, J. Raap, M. De Zotti, F. Formaggio, C. Toniolo, Biopolym. **106** (2016) 6.
11. G. Jeschke, A. Koch, U. Jonas, A. Godt, J. Magn. Reson. **155** (2002) 72.
12. A. D. Milov, Y. D. Tsvetkov, F. Formaggio, S. Oancea, C. Toniolo, J. Raap, Phys. Chem. Chem. Phys. **6** (2004) 3596.
13. M. K. Bowman, A. G. Maryasov, N. Kim, V. J. DeRose, Appl. Magn. Reson. **26** (2004) 23.
14. G. Jeschke, G. Panek, A. Godt, A. Bender, H. Paulsen, Appl. Magn. Reson. **26** (2004) 223.

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM

15. G. Jeschke, A. Bender, H. Paulsen, H. Zimmermann, A. Godt, J. Magn. Reson. **169** (2004) 1.
16. W.-Y. Chiang, P. P. Borbat, J. H. Freed, J. Magn. Reson. **172** (2005) 279.
17. W.-Y. Chiang, P. P. Borbat, J. H. Freed, J. Magn. Reson. **177** (2005) 184.
18. G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, Appl. Magn. Reson. **30** (2006) 473.
19. S. Brandon, A. H. Beth, E. J. Hustedt, J. Magn. Reson. **218** (2012) 93.
20. K. M. Salikhov, I. T. Khairuzhdinov, R. B. Zaripov, Appl. Magn. Reson. **45** (2014) 573.
21. P. G. Fajer, M. Gyimesi, A. Málnási-Csizmadia, C. R. Bagshaw, K. Ilker Sen, L. Song, J. Phys. Condens. Matter **19** (2007) 285208.
22. P. G. Fajer, L. Brown, L. Song, In: ESR Spectroscopy in Membrane Biophysics, (Eds. M. A. Hemminga, L. J. Berliner), Volume 27, Springer, NY (2007), P. 95.
23. K. Ilker Sen, T. M. Logan, P. G. Fajer, Biochem. **46** (2007) 11639.
24. V. N. Syryamina, R. I. Samoilova, Y. D. Tsvetkov, A. V. Ischenko, M. De Zotti, M. Gobbo, C. Toniolo, F. Formaggio, S. A. Dzuba, Appl. Magn. Reson. **47** (2016) 309.
25. S. A. Dzuba, J. Magn. Reson. **269** (2016) 113.
26. O. D. Zakharova, T. S. Frolova, Y. V. Yushkova, E. I. Chernyak, A. G. Pokrovsky, M. A. Pokrovsky, S. V. Morozov, O. I. Sinitsina, I. A. Grigor'ev, G. A. Nevinsky, Eur. J. Med. Chem. **122** (2016) 127.
27. Yu. V. Yushkova, S. V. Morozov, E. I. Chernyak, I. A. Grygor'ev, Chem. Nat. Compd. **52** (2016) 972.
28. H. Fonge, S. K. Chitneni, J. Lixin, K. Vunckx, K. Prinsen, J. Nuyts, L. Mortelmans, G. Bormans, Y. Ni, A. Verbruggen, Bioconjugate Chem. **18** (2007) 1924.
29. I. A. Grygor'ev, L. B. Volodarsky, Zhurnal Organ. Khimii **2** (1974) 118.
30. N. A. Kuznetsov, A. D. Milov, V. V. Koval, R. I. Samoilova, Y. A. Grishin, D. G. Knorre, Y. D. Tsvetkov, O. S. Fedorova, S. A. Dzuba, Phys. Chem. Chem. Phys. **11** (2009) 6826.
31. A. D. Milov, Y. A. Grishin, S. A. Dzuba, Yu. D. Tsvetkov, Appl. Magn. Reson. **41** (2011) 59.
32. C. Abé, D. Klose, F. Dietrich, W. H. Ziegler, Y. Polyhach, G. Jeschke, H.-J. Steinhoff, J. Magn. Reson. **216** (2012) 53.

Brought to you by | Cornell University Library
Authenticated
Download Date | 10/26/16 9:41 AM