Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 172 (2005) 279–295
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

The determination of pair distance distributions by pulsed ESR using Tikhonov regularization

Yun-Wei Chiang, Peter P. Borbat, Jack H. Freed*

Baker Laboratory of Chemistry and Chemical Biology, and National Biomedical ACERT Center for Advanced ESR Technology, Cornell University, Ithaca, NY 14853-1301, USA

Received 3 August 2004; revised 22 October 2004
Available online 2 December 2004

Abstract

Pulsed ESR techniques with the aid of site-directed spin labeling have proven useful in providing unique structural information about proteins. The determination of distance distributions in electron spin pairs directly from the dipolar time evolution of the pulsed ESR signals by means of the Tikhonov regularization method is reported. The difficulties connected with numerically inverting this ill-posed mathematical problem are clearly illustrated. The Tikhonov regularization with the regularization parameter determined by the L-curve criterion is then described and tested to confirm its accuracy and reliability. The method is applied to recent experimental results on doubly labeled proteins that have been studied using two pulsed ESR techniques, double quantum coherence (DQC) ESR and double electron–electron resonance (DEER). The extracted distance distributions are able to provide valuable information about the conformational constraints in various partially folded states of proteins. This study supplies a mathematically reliable method for extracting pair distributions from pulsed ESR experimental data and has extended the use of pulsed ESR to provide results of greater value for structural biology.

Keywords: Pulsed ESR; Dipolar spectrum; Double quantum coherence ESR; Double electron–electron resonance; Tikhonov regularization; Spin-labeled proteins

1. Introduction

The measurement of long distances ranging from those in a single biomolecule to the domain size of protein complexes by ESR methods is a relatively new method in structural biology. Based on cw (continuous wave) [1,2] and pulsed ESR techniques [1], the distances available from ESR span the range of distances between electron spins from 0.8 to almost 8 nm. As compared to mature fluorescence resonance energy transfer (FRET) techniques [3,4], ESR provides several important advantages in terms of smaller molecular size of the probes for a less perturbing original structure, and simpler labeling protocols; (two different probes are required for FRET). ESR distance measurement techniques have been successful in various aspects for studying biological systems. In combination with site-specific double-spin-labeling techniques [5] it has become a valuable tool for determining long range distance constraints of the protein backbone and obtaining information on conformational mobility in specific loop regions [1,6–8].

Double-quantum coherence (DQC)-ESR [9–11] has been shown to be a promising approach to determine intramolecular distances in proteins [12]. Its capability for large distance measurement in biological systems has recently been expanded from 5 nm to about 7 nm by the double-quantum filtered refocused (DQFR)-electron spin-echo (ESE) technique, which is based on the partial suppression of the effects of nuclear spin diffusion

* Corresponding author. Fax: +1 607 255 6969.
E-mail address: jhf@ccmr.cornell.edu (J.H. Freed).

1090-7807/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2004.10.012

on the electron spin phase memory time T_m [13]. The notable features of DQC-ESR are: (i) with appropriate phase cycling it is able to isolate just the double quantum coherence signal, possible in a system of two electron spins such as a double-labeled molecule, from which the dipolar interaction is directly extracted; (ii) the standard form of DQC [12] based on strong microwave pulses yields time-domain signals that are distinct from relaxation decay; (iii) dilute systems, such as solutions of large proteins and their complexes, can yield signals an order of magnitude greater than that of techniques that rely on weaker pulses; (iv) thus DQC is desirable when working with small amounts of bilabeled biomolecule and/or when greater sensitivity is needed for measuring longer distances.

Among the other methods, the older technique of PELDOR [14–16], which is also known as DEER (double electron–electron resonance) [17,18] is the most familiar, and it has been applied more often to investigate the structure of biomolecules. Both, DQC-ESR and DEER are based on detecting the spin-echo, but the effect of the dipolar coupling enters differently. DQC-ESR is based on the development of electron–electron coherence in the system of coupled electron spins, e.g., on anti-phase coherence $(2S_{1+}S_{2z} + 2S_{2+}S_{1z}) \times \sin(at/2)$, which can be converted to double-quantum coherence $S_{1+}S_{2+} - S_{1-}S_{2-}$ and back, with the proper pulse sequence [9]. Here, a is the value of the dipolar coupling. Both spins must be manipulated with coherent pulses in a concerted way for the method to work. DEER, on the other hand, records the evolution of the in-phase single-quantum coherence $S_{1+}\cos(at/2)$ that is affected by the interaction with the 2nd spin, flipped independently by a single pulse at its resonant frequency, which must be sufficiently different from that of the first spin. No coherence development between the two spins is necessary, regardless of the fact that DEER can be described in terms of anti-phase coherence, which accounts for the "missing" part of the detected echo. Thus, even though the two methods detect the evolution of the single-electron coherence caused by interaction with another spin, they do it quite differently. As a result the two methods can be expected to complement each other. Due to its simpler (compared to DQC) requirements on pulse strength and manipulations, DEER has been used by several groups in studies of intramolecular organization in biological systems [19–24].

In the past the average distance in a pair of electron spins has been estimated by Fourier transforming the time-domain experimental data to obtain the singular frequencies of the Pake pattern function [25]. An estimate of the pair distribution can be obtained by several indirect methods. One way is to fit a reasonable model to the experimental data. In early studies with PELDOR, a simple model for the distance distribution was presumed to fit the experimental time evolution data [15,16]. A current advantage of modeling is that one can parametrize and attempt to fit subtle aspects of the method [20]. Borbat et al. [12] obtained more detailed pair distributions in a protein system using DQC-ESR by considering refined geometrical models based on Monte Carlo simulations of the conformations of the methanethiosulfonate spin label (MTSSL), which was attached at two sites of the protein. Fig. 1 shows the scheme for how the average distance between a pair of spins in proteins can be obtained (by Fourier transform) and how the pair distribution can be extracted (by the method presented in the present report) from the dipolar time evolution signal produced with pulsed ESR techniques. MTSSL is shown in (Fig. 1F) and was used as the spin label in the present study.

As opposed to the use of a parameterized model for fitting the experimental data, here we concentrated on solving directly the inverse problem of reconstruction of the distance distribution from experimental data. The equation connecting the pair distribution ($P(r)$ in Fig. 1E) and experimental time evolution data ($V(t)$ in Fig. 1B) is the Fredholm equation of the first kind, (a detailed account of which is given in the next section), which has been recognized as representing an ill-posed problem [26]. This means that the solution of this equation has been proven to be mathematically unstable, i.e., a number of possible solutions, which may be totally unreasonable from, e.g., the structural perspective, may approximate the experimental data well within the error bounds. A large variety of physical problems belong to this class, a close example is the problem of "de-Paking" in solid-state NMR. This is the reason why a physically reasonable model was utilized for fitting the experimental time evolution data provided by DQC ESR. Recently, a method for direct conversion of ESR dipolar time evolution data to distance distributions was proposed by Jeschke et al. [27]. They found that with judicious discretization of the equation it was possible to convert the ill-posed problem to a reasonably well-posed problem. Very recently, the data analysis procedures for this ill-posed problem (more specifically, the dipolar evolution function for PELDOR measurements) has been aided by employing several methods, including Tikhonov regularization, coordinate transformation, and polynomial interpolation [28,29].[1]

In this work, the Tikhonov regularization method is applied to reconstruct distance distributions, and its usefulness for this task is investigated. Starting with the analysis based on singular value decomposition (SVD), we illustrate the difficulties connected with the recovery of the distance distribution from the spectroscopic data. The quality of the solution, $P(r)$ obtained by Tikhonov

[1] We thank the reviewers for calling our attention to the very recent references [28,29].

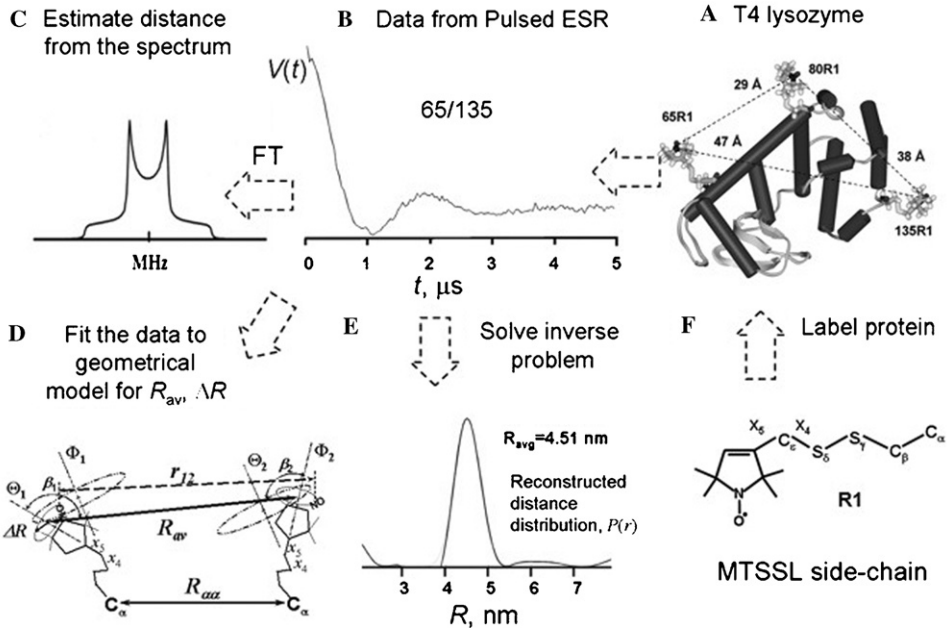

Fig. 1. Distance measurements by pulsed ESR. (A) A cartoon structure of T4 lysozyme with MTSSL (F) side-chains shown. (B) Time-domain data $V(t)$ from 17.3 GHz DQC for 65/135 labeling. (C) The dipolar spectrum of (B). (D) The geometrical model used to fit the data for R_{av} and ΔR. The model is based on $X4/X5$ [12] approximation of side-chain $R1$ (F) conformations. (E) The pair distance distribution reconstructed by solving the corresponding inverse problem (cf. Eq. (1)) using Tikhonov regularization method.

regularization strongly depends on finding the optimal regularization parameter, λ. Whereas there are many methods for determining the optimal value [30], the L-curve criterion was selected, in particular because it leads to a convenient graphical tool for displaying the effect of λ on the regularized solution. We have tested it with two representative distributions of distinct shape, a bimodal one based on two Gaussians and a box-like one with sharp edges.

We apply this method to biomolecules, by analyzing both DQC-ESR and DEER experiments that have been performed on the proteins T4 lysozyme and cytochrome c and its various intermediate conformations between native and unfolded states. The distance distributions, estimated by using the Tikhonov regularization on experimental data, are in good agreement with the conformations of these proteins resolved by other methods. Our analysis does provide pulsed ESR with an additional feature to approach the subject of structural biology.

2. Mathematical formulation of the dipole–dipole interaction

A complex spin system is characterized by dipole–dipole interactions amongst all its spins. In the ideal case of a very dilute distribution of isolated molecules bearing two electron spins, the spin system can be viewed as an ensemble of spin pairs, each coupled via their intramolecular dipolar coupling but also weakly coupled to all other spins via intermolecular dipolar couplings. The signals obtained from either DQC or DEER reflect the effects of these two types of interactions where, of course, only the intramolecular couplings are of interest to us from the viewpoint of a structural study. It has been recognized that the intermolecular interactions produce a significant effect on the signal, causing single-exponential signal decay by the mechanism of instantaneous diffusion [31,32]. It was shown to appear as a baseline effect in the DQC time-domain data from a system of dilute bilabeled molecules [12]. In turn, the intramolecular couplings lead to the finite-depth modulation of the signal. The baseline function with respect to time was found to be approximately linear in the DQC signal envelope, and the logarithmic plot of the DEER signal is linear as well. This baseline function can be removed, as has been shown elsewhere [12,16].

The intramolecular dipole–dipole interactions can be expressed in the following form given by equations:

$$\int_{R_{\min}}^{R_{\max}} \kappa(r,t)P(r)\,\mathrm{d}r = V(t), \tag{1}$$

$$\text{with } \kappa(r,t) = \int_0^1 \cos[(1-3x^2)\omega_d t]\,\mathrm{d}x,$$

$$\omega_d = \frac{\gamma_e^2\hbar}{r^3}, \quad \text{and } x = \cos\theta, \tag{2}$$

where $V(t)$ is the intramolecular signal obtained from dipolar spectroscopy experiments, $\kappa(r,t)$ represents the ensemble average of the dipolar coupling over all possi-

ble molecular orientations for given r. The angle θ is the angle between the vector $\mathbf{r}$ connecting the two spins and the direction of the static magnetic field, B_0. $P(r)$ is the distance distribution in spin pairs defined on the interval $[R_{\min}, R_{\max}]$. Eq. (1) is in the form of a Fredholm equation of the first whose kernel $\kappa(r,t)$ is given by Eq. (2). The inversion of this equation for obtaining $P(r)$ is thus an ill-posed problem, which means that many different functions, $P(r)$ may yield a predicted $V(t)$ that agrees with the experimental $V(t)$ within the error bounds of the experiment.

Strictly speaking, the kernel $\kappa(r,t)$ is suitable for an ideal case, since for example, it neglects any intermolecular dipolar interactions, (or else they have been effectively removed as discussed above). In addition, a more rigorous kernel, $\kappa(\mathbf{r},t)$, would also account for correlation of the orientation of the radial vector between the two spin labels in the molecule and the orientations of the nitroxide magnetic tensors, which would lead to orientation selection when finite pulses are used that provide incomplete spectral excitation. Whereas it is possible to account for these and other possible effects [9,23], it would greatly complicate $\kappa(\mathbf{r},t)$ by adding such ill-defined features into the analysis. On the contrary, the form of the intramolecular kernel given by Eqs. (1) and (2) is believed to constitute a reasonable approximation, since typically spin-labels used in protein labeling have enough flexibility to significantly reduce complicating effects from such orientational correlations [12]. Clearly, in all these experiments, the data requires some form of largely empirical conditioning to account for known effects, such as intermolecular couplings or relaxation, before they can be used to recover distributions in distances.

3. Shortcomings in the application of the linear regression method

To invert this Fredholm equation of the first kind numerically, we should first convert it to a discrete form. There are many ways to discretize integral equations, and we do not attempt to cover this topic extensively here. The quadrature method is used below to approximate the integral equation in Eq. (1) with a weighted sum as

$$V(t) = \int_{R_{\min}}^{R_{\max}} \kappa(r,t)P(r)\,\mathrm{d}r \approx \sum_{j=1}^{N} w_j \kappa(t_i, r_j)P(r_j), \tag{3}$$

where $w_j = (R_{\max} - R_{\min})/N$, $r_j = (j - 1/2)w_j + R_{\min}$ for $j = 1,2,\ldots,N$, and $t_i = (i - 1/2)(T_{\max}/M)$ for $i = 1,2,\ldots,M$. N and M represent the numbers of solution points and time-domain data points, respectively. Hence, we convert the equation to a system of linear algebraic equation represented by,

$$S = KP \tag{4a}$$

with the elements given by

$$k_{ij} = w_j \kappa(t_i, r_j) \text{ and } s_i = V(t_i)$$
$$\text{for } i = 1,2,\ldots,M,\ \ j = 1,2,\ldots,N, \tag{4b}$$

where K denotes an operator which maps the function P into the experimental data vector S. And as noted, K is an $M \times N$ matrix with $M \geqslant N$.

According to the singular value decomposition (SVD) method, K can be decomposed as

$$K = U\Sigma V^{\mathrm{T}} = \begin{pmatrix} u_{11} & \cdot & \cdot & u_{1N} \\ \cdot & & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot \\ u_{M1} & \cdot & \cdot & u_{MN} \end{pmatrix} \cdot \begin{pmatrix} \sigma_{11} & & & \\ & \cdot & & 0 \\ & & \cdot & \\ & 0 & & \sigma_{NN} \end{pmatrix} \cdot \begin{pmatrix} v_{11} & \cdot & \cdot & v_{1N} \\ \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ v_{N1} & \cdot & \cdot & v_{NN} \end{pmatrix}^{\mathrm{T}} . \tag{5}$$

Thus, K can be written as the product of an $M \times N$ column-orthogonal matrix U, whose columns are composed of the M-dimensional vectors u_i, an $N \times N$ diagonal matrix Σ with nonnegative elements σ (the singular values), and the transpose of an $N \times N$ column-orthogonal matrix V, whose columns are composed of the N-dimensional vectors v_i, such that $U^{\mathrm{T}}U = I$ and $V^{\mathrm{T}}V = I$ [33]. What is common for all discrete ill-posed problems is that the matrix Σ is a diagonal matrix which has nonnegative diagonal elements appearing in decreasing order such that $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_N \geqslant 0$, and an increase of the dimensions of K will increase the number of small singular values.

It can be readily shown that the solution to Eq. (4a) can be written as

$$P = \underset{N\times N}{(V)} \cdot \underset{N\times N}{(\mathrm{diag}(1/\sigma_i))} \cdot \underset{N\times M}{\left(U^{\mathrm{T}}\right)} \cdot \underset{M\times 1}{(S)} = \sum_{i=1}^{N} \frac{u_i^{\mathrm{T}} S}{\sigma_i} v_i. \tag{6}$$

The solution vector P only exists if the right-hand side of Eq. (6) indeed converges. From Eq. (6), we see that P is expressed in terms of the singular vectors v_i and their corresponding expansion coefficients $u_i^{\mathrm{T}}S/\sigma_i$. One can therefore completely characterize the solution P by an analysis of the coefficients and the singular vectors.

Let us assume that the errors in the given problem are restricted to the right-hand side of Eq. (4a), such that the given data vector S can be written as

$$S = S^* + \varepsilon, \quad S^* = KP^*, \tag{7}$$

where S^* represents the exact unperturbed data, and the vector ε represents the noise in the data. At this stage, it

is necessary to introduce the following assumption, which is referred to as the Picard condition [34], in order to characterize the solution P. In order for Eq. (4a) to have a solution, it is necessary and sufficient that

$$\sum_{i=1}^{M} \frac{1}{\sigma_i^2} |\langle S, u_i \rangle|^2 < \infty,$$

where the brackets denote the inner product. For the discrete algorithm used herein it means that the best approximate solution of $KP = S$ exists only if the SVD coefficients $|u_i^T S|$ decay sufficiently fast with respect to singular values σ_i with increasing i.

Here, we consider the bimodal distribution, shown in Fig. 2A, whose time-domain signal vectors S and S^* are shown in Fig. 2B, as an illustration of how the ill-posed problem can be examined using a Picard plot. Fig. 3 shows the Picard plot of the first 100 singular values σ_i, coefficients $|u_i^T S|$, and coefficients $|u_i^T S/\sigma_i|$ of the solution for the case of $S = S^*$, which means that no noise is present on the right-hand side of Eq. (4a). The coefficients $|u_i^T S|$ decay with i faster than σ_i does before both coefficients level off at the rounding-off error ($\sim 10^{-17}$) by ca. the first 50 singular values. This indicates that a solution P cannot be recovered by SVD once the number of desired singular values in the solution, i is greater than ca. 50, because the coefficients $|u_i^T S|$ fail to meet the requirement of the Picard criterion. To illustrate how dramatically a few small singular values can spoil the estimated solution, two solutions of this case, (i.e., noise-free) were obtained by truncating the respective singular values by the 53rd and 54th and are shown in Figs. 3B and C. The estimated solution obtained with the first 53 singular values is acceptable; whereas, the solution obtained with the first 54 singular values is entirely garbled due to just a few very small singular values. This indicates that after ca. σ_{50} the decreasing singular values become dominant in the solution, so that by σ_{54} the solution has been greatly distorted just by computer round-off error. The Picard condition is thus found useful in evaluating a system of linear equations. Fig. 3D shows the Picard plot for the case where the SNR of the S vector is set to 100. The coefficients $|u_i^T S|$ fail throughout to meet the requirement of the Picard condition, because the coefficients $|u_i^T S|$ do not decay fast enough relative to the singular values σ_i. This indicates that for the kernel of interest (cf. Eq. (2)) the problem is severely ill-posed at a noise level of SNR ~ 100 as does the system of corresponding linear equations (cf. Eq. (4a)). If one attempts to reconstruct an approximate kernel with proper discretization, he may be misled to think he is closer to the correct solution than is really the case just because some of singular values have contributed significantly to the variations in the solution. Thus, for this case of SNR ~ 100 it appears impossible to recover the solution P by SVD.

Now it is clear that the singular values continue to decrease as the number of SVD components increase. Because the inverse square of the σ_i contribute to the solution P, a small amount of noise in S may produce extremely large variations in P, if some singular values

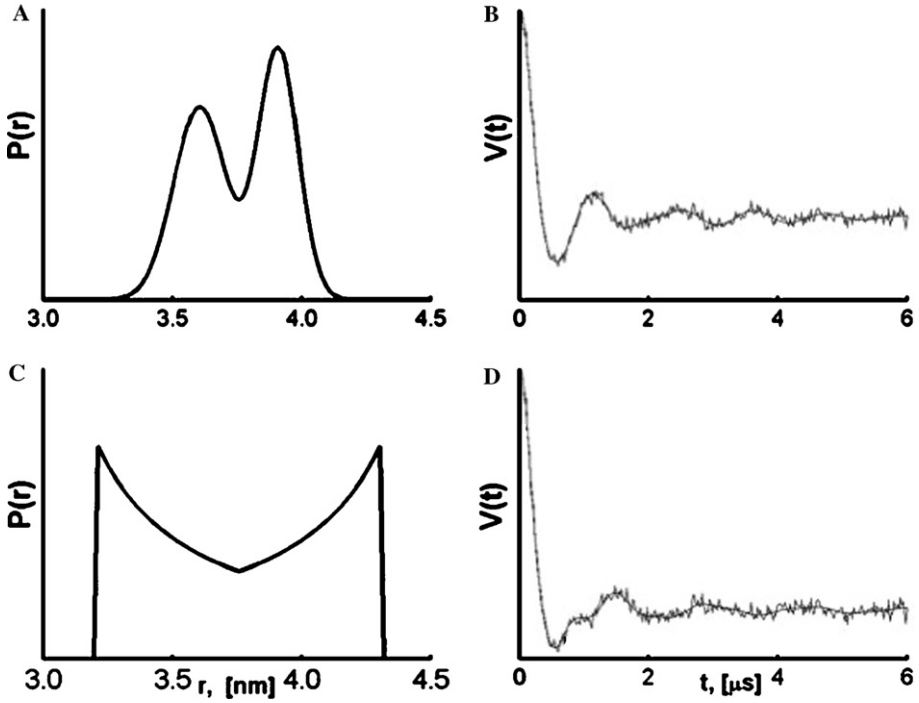

Fig. 2. Two models of simulated distance distribution, $P(r)$ are illustrated in (A), a bimodal distribution taken as a sum of two Gaussians, and (C), a box-like distribution. The respective dipolar signals $V(t)$ simulated by Eq. (1) are shown in (B) and (D), with the noise that was added to yield SNR's of ca. 50.

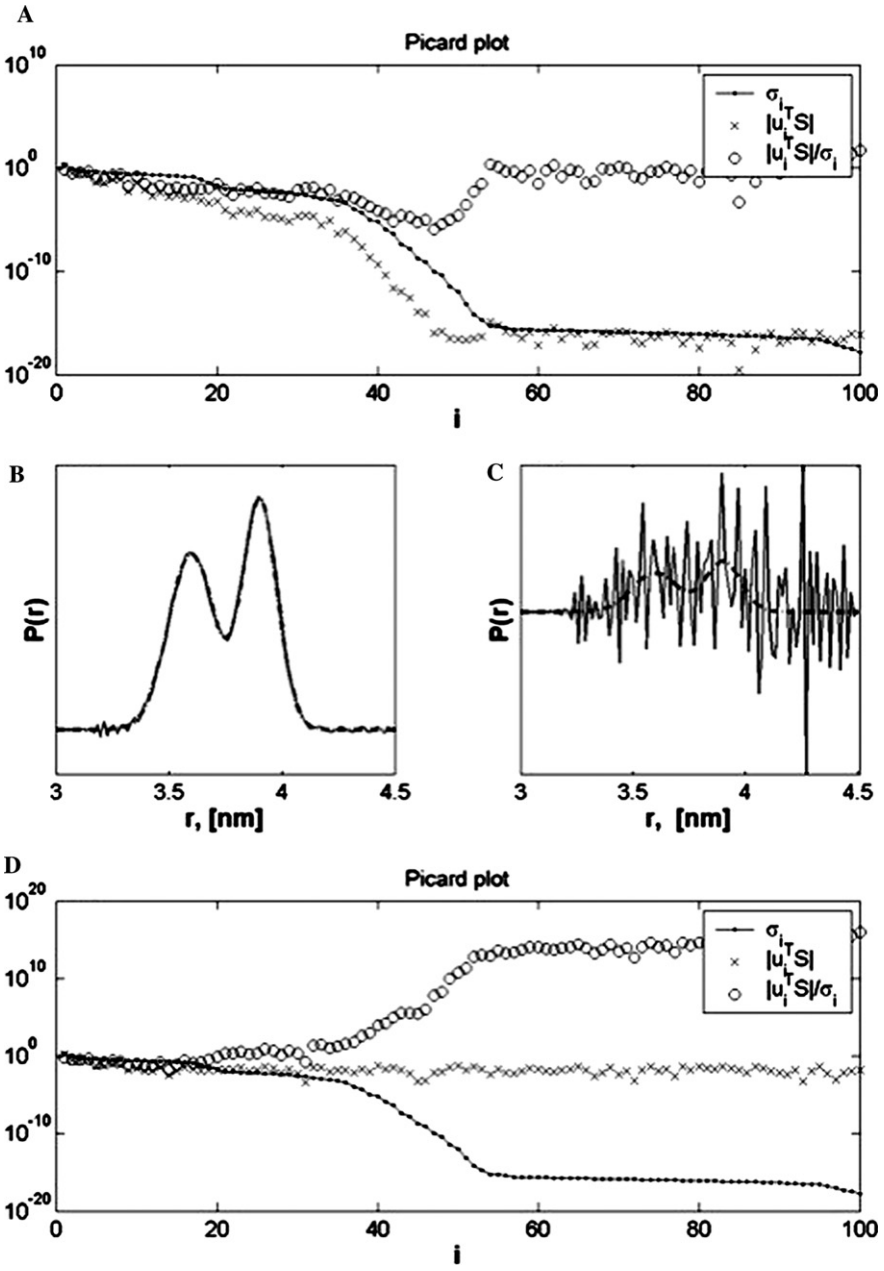

Fig. 3. (A) The Picard plot of the linear system, (cf. Eq. (4a)) for the bimodal distribution. Noise is set to zero in the vector S. The coefficients $|u_i^T S|$ decay faster than the singular values, σ_i for ca. $i \geqslant 50$. The solution, P becomes dominated and distorted by the decreasing σ_i as i is increased to greater than ca. 50. The estimated distributions obtained using the first 53 and 54 σ_i are shown in (B) and (C), where the original distribution is plotted by a dashed line. (D) The Picard plot for the case of (A) with the SNR increased to ca. 100. The system is ill-posed because the coefficients $|u_i^T S/\sigma_i|$ diverge quickly and dominate the solution, P starting from even the first few singular values.

become too small. This corresponds to large errors. The solution P obtained from some data S may deviate strongly from true values. Hence, even one small singular value is very undesirable. It means that somehow information is lost and cannot be restored by inversion. The Picard condition provides a valuable tool to determine the severity of the ill-posed inverse problem for the system given by Eq. (1) with the kernel of Eq. (2). A method which can extract a numerically stable $P(r)$ from experimental data is necessary in applications of pulsed ESR to distance measurements. In other words, the problem should be regularized by transforming the Fredholm equation of the first kind into a form that is not ill-posed.

4. Tikhonov regularization

The Tikhonov regularization method is based on a modification of the Fredholm integral equation of the first kind that is intended to stabilize its solution.

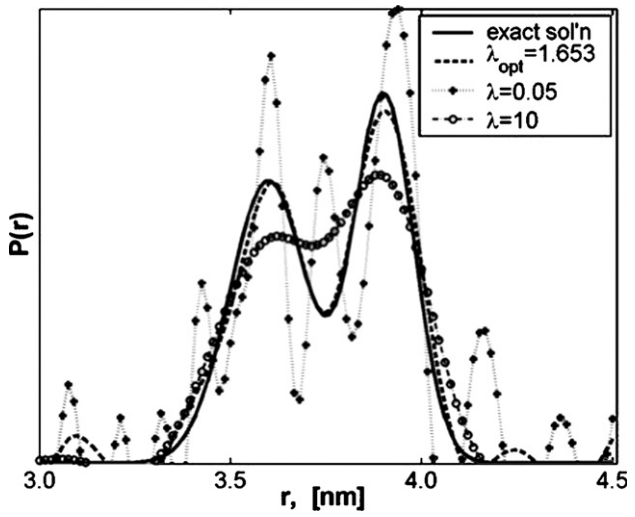

Fig. 4. The influence of the regularization parameter on the regularized solution is illustrated by varying the regularization parameter λ. As $\lambda \ll \lambda_{\text{opt}}$ the regularized solution has many peaks because of the inherent instability of the residual norm $\|KP - S\|$. The second term of Eq. (8) stabilizes the solution. As λ increases to $\lambda \gg \lambda_{\text{opt}}$ it results in an over-smoothed solution. The exact solution and the solution with optimal regularization parameter are also shown in the figure for comparison.

This is achieved first by constructing the following functional:

$$\Phi[P] = \|KP - S\|^2 + \lambda^2 \|LP\|^2 \tag{8}$$

with λ being the so-called regularization parameter, L is an operator, usually the identity ($LP = P$) or second derivative ($LP = P''$). Then the regularized solution P_λ is formally obtained by minimizing the functional in Eq. (8), yielding: [35,36]

$$P_\lambda = \left(K^{\mathrm{T}}K + \lambda^2 L^{\mathrm{T}}L\right)^{-1} K^{\mathrm{T}}S. \tag{9}$$

For an appropriate value of λ, the first term on the right hand side of the Eq. (8) forces the result to become compatible with the data. The second term leads to a smoothed estimate of the solution. The quality of the result depends strongly on the regularization parameter λ. If λ is too small, the result will show artificial peaks. If λ is too large, the result will be over-smoothed. Fig. 4 shows a comparison between the exact solution of the bimodal distribution and the solutions with Tikhonov regularization for $\lambda \gg \lambda_{\text{optimal}}$, $\lambda = \lambda_{\text{optimal}}$, and $\lambda \ll \lambda_{\text{optimal}}$. Thus, for this reason, a reliable method for determination of the regularization parameter is crucial to solving our inverse problem.

5. Determination of the regularization parameter

The main problem when using the Tikhonov regularization method is the determination of an appropriate regularization parameter λ. Several numerical methods for determining the regularization parameters were critically discussed in the literature [30]. Other methods are based on tracking the changes in the size of the regularized solution versus the size of the corresponding residual to determine an optimal λ. The procedure represented in graphical form is referred to as the L-curve criterion. The use of such a criterion in connection with ill-posed least squares problems goes back to Miller [37] and Lawson and Hanson [38]. The L-curve criterion is clearly illustrated and extensively applied to the analysis of discrete ill-posed problems by Hansen [39]. More properties of the L-curve are derived by Hansen and O'Leary [40] where it was also shown that the characteristic L-shaped corner is better pronounced in a double logarithmic plot. Many properties of the L-curve for Tikhonov regularization have been investigated [40,41]. Two limitations [42] of the L-curve criterion should be mentioned and checked in the computation: (i) if the corresponding SVD coefficients $|v_i^{\mathrm{T}}P^*|$ decay very fast to zero such that the solution P^* (exact solution) is dominated by the first few SVD components, the optimal regularization parameter may not be located at the corner of the L-curve; (ii) as pointed out by Vogel [43], the λ that is based on the L-curve criterion may deviate from the optimal λ when N, the vector rank of the regularized solution, is too large. Note that it depends on the specifics of the kernel for the problem regarding how large a value of N would cause a failure of the L-curve approach to determine λ. A detailed analysis of the problem for the kernel of interest and the destabilizing effects caused by the experimental noise is given in the later sections.

The L-curve is a parametric plot of type $(\rho^2(\lambda), \eta^2(\lambda))$ [40], where $\rho(\lambda)$ and $\eta(\lambda)$ represent the norms of the regularized solution $\|LP\|$ and the residual $\|KP - S\|$, respectively. The underlying idea of the L-curve method is that a good method for choosing the regularization parameter for discrete ill-posed problems must incorporate information about the solution norm $\|LP\|$ in addition to using the information about the residual norm $\|KP - S\|$. This is indeed very consistent with the spirit of the Tikhonov regularization, which is poised to seek a reasonable balance in keeping both of these values small. Minimization of the Tikhonov functional of Eq. (8) is a compromise between minimizing the residual norm $\|KP - S\|^2$, and keeping the solution norm $\|LP\|^2$ small, i.e., enforcing stability of the solution. The L-curve has a distinct L-shaped corner located exactly where the solution P_λ changes in nature from being dominated by regularization errors to being dominated by noise in the S vector. It has been proven that the L-curve always has an L-shaped appearance [40]. Hence the corner of the L-curve corresponds to a desirable balance between minimized norms, and the corresponding λ is a good one. The idea for choosing the optimal λ is to choose the point on the L-curve that is at the corner

of the curve corresponding to its maximum curvature. As noted above, the rationale behind this criterion is that the corner corresponds to a solution in which there is a reasonable balance between the regularization and perturbation errors. This interpretation is of particular importance for Tikhonov regularization due to its corner-stone variation theorem [26] that P_λ is the unique minimizer of the Tikhonov functional.

The implementation of the L-curve method is based on locating the unique point of maximum curvature in the plot of $(\rho^2(\lambda), \eta^2(\lambda))$. It is obtained by consecutive substitutions of λ into Eq. (9) to yield P_λ from which the corresponding values of $\rho(\lambda)$ and $\eta(\lambda)$ (defined above) are obtained. One then generates an L-curve parametric plot whose x and y axes are $\rho^2(\lambda)$ and $\eta^2(\lambda)$, respectively. The point of maximum curvature is then obtained directly by finding the maximum of $d(\rho^2)/d(\eta^2)$ [40,42]. In addition, the L-curve plot provides a useful graphic illustration for the basic mathematical idea of the Tikhonov regularization method, (i.e., the corner represents a balance in keeping both norms small, as mentioned above). Note that the L-curve plots shown in the sequel are displayed as double logarithmic plots (i.e., plots of log ρ vs. logη) to best maintain the L-shaped curve. It has been found that the behavior of the L-curve is more easily seen and better pronounced if plotted in this way [40].

6. Results and discussion

6.1. Application to simulated data

To simulate experimental data, two different models of the distance distribution $P(r)$ were chosen and shown in Fig. 2A (a bimodal Gaussian distribution with two peaks, whose respective parameters (mean, standard deviation) are 3.6, 1 nm and 3.9, 0.77 nm) and Fig. 2C (a box-like distribution with two maxima located at 3.2 and 4.3 nm and the curve between the peaks decreases as r^{-1} from each maximum such that it is symmetric about the center at 3.75 nm). Their respective time evolution signals (i.e., the vector S) were calculated according to Eq. (1) and are shown in Figs. 2B and D. The number of solution points (i.e., N; cf. Eq. (4b)) was set to be the same as the number elements of the vector S (i.e., M; cf. Eq. (4b)) for all cases studied, although in some high SNR cases we found reducing N not only can shorten the computation time but also does not affect the estimated distribution. In other words, K was usually chosen to be a square matrix, which provides the maximum possible number of points in the solution.

The estimated distance distributions $P(r)$ extracted for three different noise levels in dipolar signals are shown on the right-hand side of Fig. 5 by thick solid lines where the original distributions are plotted by thin solid lines. The respective corresponding L-curves are shown on the left-hand side. The optimal regularization parameter, which minimizes the Tikhonov functional in Eq. (8), for the respective case is marked on the corresponding L-curve plot. All results were obtained by using the "Regularization Toolbox" of the Matlab software package, which was originally created by Hansen's group [44] and adapted for present purposes. The reconstruction of the distance distribution with the Regularization Tools package is remarkably good, especially for cases with high SNR in the dipolar signal. The original bimodal distribution is well recovered in the cases of SNR $\sim$ 500 and 100. The optimal λ, which represents the maximum curvature of the L-curve, is properly located on the corner for these two cases. As the SNR is decreased to 30, the maximum curvature of the L-curve fails to locate at the *visual* corner of the L-curve in log-log scale. This indicates the L-curve criterion may break down under SNR $\sim$ 30, therefore the distribution is poorly recovered. In other words, the result shows that the estimated solution is improved as the SNR in S is increased, as one would expect. The increased noise in the vector S is the factor that causes the corner of the L-curve to become less sharp, and therefore contributes to the failure to obtain an optimal regularization parameter. As shown in Fig. 5, the estimated distribution thus first becomes oscillatory about zero in the peripheral regions of the distribution that is properly recovered from a high SNR of ca. 500 and then becomes distorted globally as the SNR decreases to ca. 30.

The estimated solutions for the box-like distribution are shown in Fig. 6. The SNR's for the respective cases are shown in the insets. It is obvious that higher SNR is required, compared to the bimodal distribution, to reconstruct this type of distribution. The same oscillations (which are caused by the noise) in the peripheral regions of the distribution were observed. For the case of SNR $\sim$ 30, the optimal λ considerably deviates from the corner. This indicates that the original distribution may not be realistically recovered at this noise level as shown on the right-hand side. In the estimated distribution the two sharp peaks in the original distribution are barely seen, whereas, the average distribution remains close to the original. An SNR $\sim$ 30 is, in general, smaller than in a good DQC or DEER experiment. The box-like distribution is a challenge for any of the regularization methods. Thus, it is fair to say that Tikhonov regularization with the L-curve works quite well for the kernel of Eq. (2).

It should be noted that the estimated solution $P(r)$ was found to be oscillating about zero in the peripheral regions of the distribution for all of the cases studied. These oscillations were found to exist even in the case of SNR $\sim$ 2000, which is much greater than in typical experiments. Although negative values of $P(r)$ are

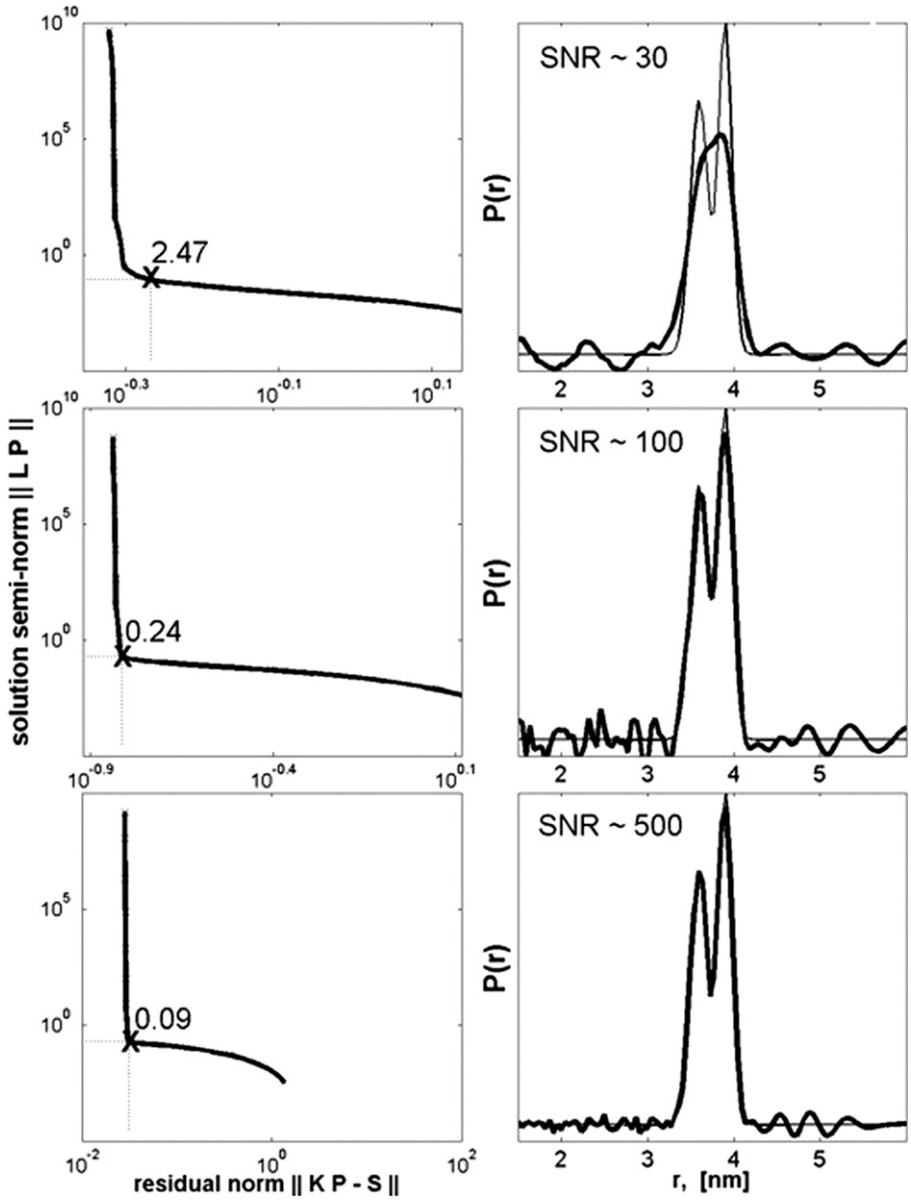

Fig. 5. The bimodal Gaussian-like distribution is recovered from three different noise levels (SNR $\sim$ 30, 100, and 500 corresponding to top, middle, and bottom graphs, respectively) in the simulated time evolution data using Tikhonov regularization. The recovered distribution is plotted by a thick solid line on the right-hand side, while the original distribution is shown by a thin line. The respective corresponding L-curves are shown on the left-hand side, which themselves illustrate to some extent how ill-posed is the case with the applied noise. The optimal λ for each case is marked on the L-curve plot.

physically unreasonable, it was not necessary to modify the current algorithm to force the estimated solution be positive since (i) as shown above, this oscillation does not affect the average distance even for the case of SNR $\sim$ 30; (ii) the oscillation in the peripheral regions never becomes dominant in intensity as compared to the major distribution in all of the cases studied; (iii) although the peripheral distribution could become dominant in cases of very low SNR, and thus be mistaken as the main distribution in such low SNR cases, the optimal λ will have deviated significantly from the corner (see below). Thus, it is important to observe the location of the optimal λ, so that one can know how well the distribution is recovered. As far as choosing the operator L in Eq. (8), we found from the model tests that for the case of high SNR (such as SNR $\sim$ 500) the solutions obtained with either L equal the identity or the second derivative operator are almost the same; whereas, for the case of SNR $\sim$ 30 the solutions were better recovered with the identity operator; (e.g., the estimated solution of the bimodal distribution of SNR $\sim$ 30 was recovered to be a simple Gaussian distribution, if the second derivative operator is used to smooth the distribution). Thus, we selected the identity operator

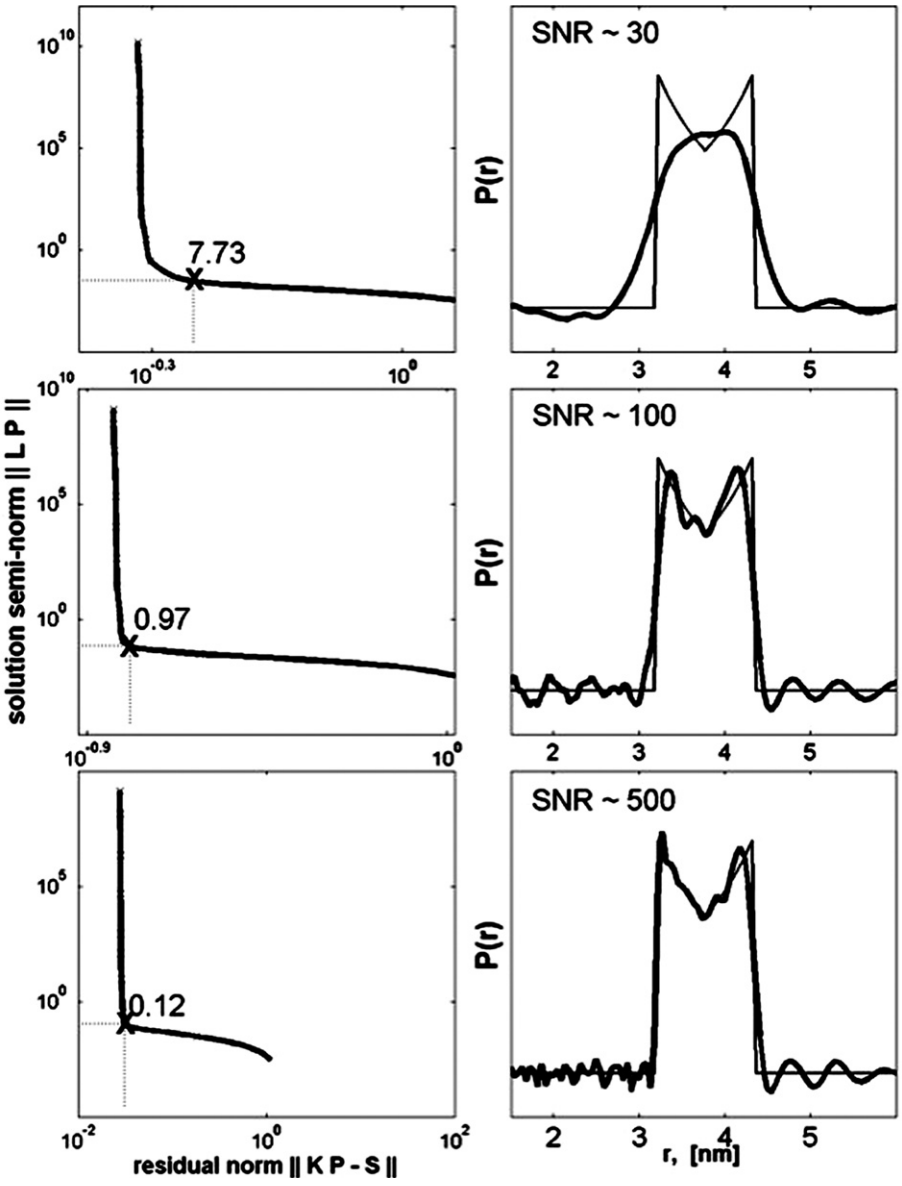

Fig. 6. The box-like distribution is recovered from three different noise levels (SNR $\sim$ 30, 100, and 500 corresponding to top, middle, and bottom graphs, respectively) in the simulated time evolution data using Tikhonov regularization. The recovered distribution is plotted by a thick solid line on the right-hand side while the original distribution is shown by a thin line. The respective corresponding L-curves are shown on the left-hand side of the figure. The optimal λ for each case is marked on the L-curve plot. Although the distribution with SNR $\sim$ 30 is poorly recovered, the average distance is close to a correct value.

to recover the distributions from the actual experimental data, which will be shown in the next section.

In summary, the optimal λ is determined by the maximum curvature of the L-curve, and it is found to be located at the visual corner of the L-curve in a log–log scale plot in most cases. In other words, the location of the optimal λ in log-log scale plot of the L-curve indicates how well the original distribution can be recovered under the given noise level. As shown above, the increased noise in the vector S causes the optimal λ to deviate from the L-curve in the log–log scale plot resulting in a poorly reconstructed distribution (the bimodal and box-like distributions with SNR $\sim$ 30).

6.2. Experimental methods

The DQC experiments at 17.3 GHz were performed as described previously [12], but with the use of a dielectric resonator in order to improve B_1 and the sensitivity. For the same frequency, the DEER option was provided as an extension of the existing pulse spectrometer by adding a pumping arm. It consists of a microwave

source, pulse channel, and Ku-band TWT amplifier. The pumping pulse was injected into the waveguide through the 10 dB directional coupler. For both types of experiments a continuous flow cryostat CF935 (Oxford Instruments), housing the dielectric resonator, was employed to stabilize the sample temperature. In the DQC experiments the 6-pulse sequence ($\pi/2$–π–$\pi/2$–π–$\pi/2$–π) with $\pi/2$ pulses of 3.2 ns and π pulses of 6.2 ns was used. In 4-pulse DEER [14], the center maximum of the ESR spectrum was used for pumping and the low-field edge for detection. The frequency difference was set to 65–70 MHz. For detection 16 ns $\pi/2$ and 32 ns π pulses were used; a 25 ns π-pulse was used for pumping.

Two proteins were studied, T4-lysozyme (T4L) and iso-cytochrome *c* (iso-1-cyt *c*). The residues of interest were substituted by cysteines and labeled with the nitroxide spin label, (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl) methanethiosulfonate (MTSSL). The respective nitroxide side-chain R1 is shown in Fig. 1. T4L was prepared at concentrations of 50–100 μM in 30 wt% glycerol-d_8/D_2O and studied by DQC. Iso-1-cyt *c* at a concentration of 500 μM in 30 wt% glycerol–water was studied by DEER. The Heme iron in iso-1-cyt was reduced under anaerobic conditions to lengthen the T_2's of nitroxide spin-labels. For the experiments with partially unfolded protein, guanidinium hydrochloride (GdnHCl) was added to iso-1-cyt solutions in concentrations of 0.7 and 1.5 M to achieve different degrees of unfolding. For all experiments the samples were vitrified by shock-freezing in liquid nitrogen and transferred to the cryostat for measurements conducted at 70–77 K. Data collection times were in the range of 0.5 to 4 h.

6.3. Distance distributions in T4L by DQC

Figs. 7A and 8A show the DQC experimental time evolution signals of T4L 65/135 and 61/80 mutants from the six-pulse DQC sequence after correction for the decay, caused by phase relaxation due to nuclear spin diffusion.

Fig. 7B shows the result for the distance distribution of spin-pairs (solid line) recovered by Tikhonov regularization, with the L-curve shown as an inset to Fig. 7A. The shape of the L-curve is well maintained and the optimal λ does not deviate from the sharp corner. As pointed out above, this indicates that the distribution function $P(r)$ can be reasonably recovered using Tikhonov regularization. The corresponding average distance for this 65/135 doubly labeled mutant of T4L is 4.52 nm, which is consistent with the known crystal structure of T4L and our previous study wherein the dipolar spectra were fitted with a geometrical model based on the expected side-chain conformation of MTSSL [9,10]. There are some perturbations and oscillations about zero in the estimated distribution in the peripheral regions.

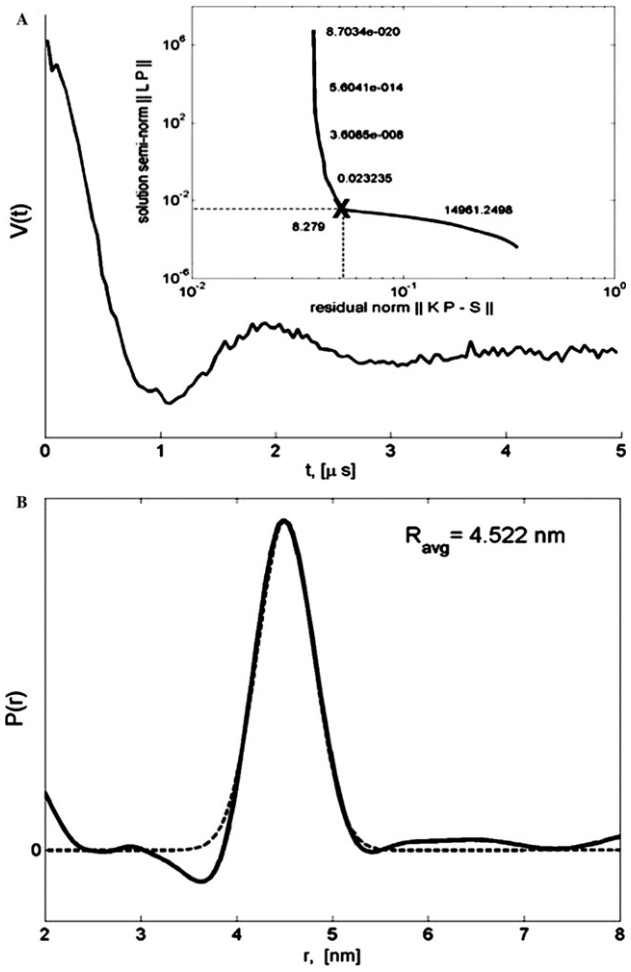

Fig. 7. (A) The dipolar time evolution spectrum of the T4L 65/135 mutant, with 40 ns sampling spacing, obtained from six-pulse DQC-ESR at 17.3 GHz. The L-curve for selecting the optimal regularization parameter, 8.279 (marked by a cross) is shown as an inset. (B) The distribution obtained is plotted by a solid line whose average distance is found to be 4.52 nm. This distribution is found to be similar to a Gaussian distribution, which is plotted by the dashed line with an average and standard deviation of 4.51 and 0.29 nm, respectively. In addition to being consistent with the previous study (~ 4.63 nm),[12] the Gaussian-like distribution indicates the side-chains are likely to exhibit less restricted conformations than assumed in the *X*4–*X*5 model.[12]

The oscillations, as shown in the model studies using simulated data, are due to perturbations (noise and distortions) in the time-domain signal, whereas it affects the average distance of the recovered distribution insignificantly even for the case of SNR ~ 30. As discussed above, only positive values of the distribution are physically relevant; small negative excursions fall outside the region where the relevant part of the distribution is located and can be ignored. The estimated distribution, shown in Fig. 7B, appears to be a Gaussian-like distribution. For better comparison, a Gaussian distribution, $\hat{S} = 0.29$ nm, $R_{avg} = 4.51$ nm, (where $\hat{S}$ represents the standard deviation of Gaussian distribution) is also plotted in Fig. 7B (dashed line). This is in good agreement with our previous result [12] (for this 65/135

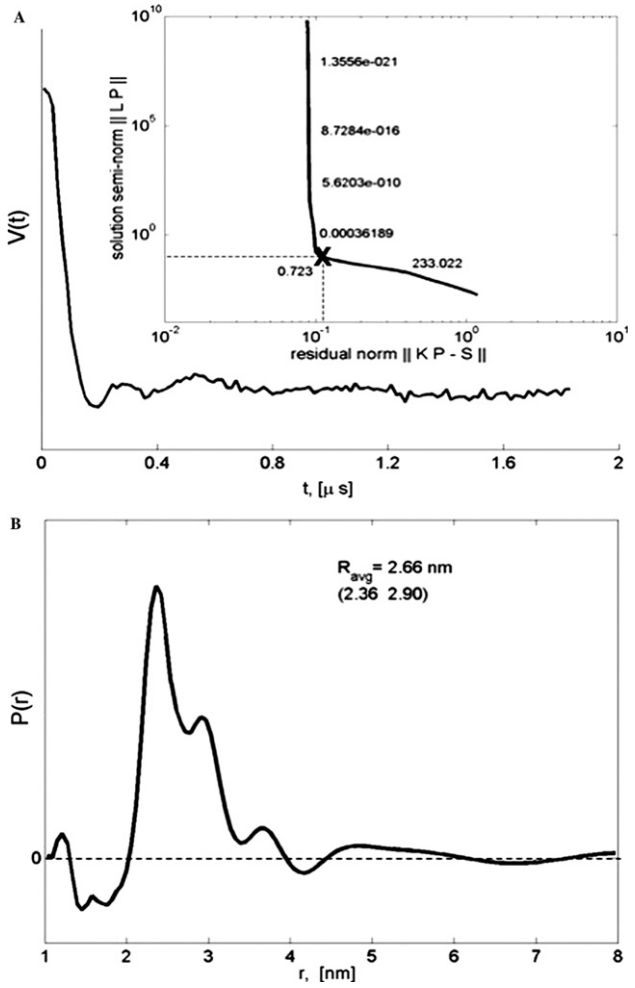

Fig. 8. (A) The dipolar time evolution spectrum of the T4L 61/80 mutant, with 16 ns sampling spacing, obtained from six-pulse DQC-ESR at 17.3 GHz. The L-curve for selecting the optimal regularization parameter, 0.723 (marked by a cross) is shown as an inset. (B) The estimated distribution is plotted whose average distance the main distribution is found to be 2.66 nm. It shows a bimodal distribution with a relative weight of the components of 2:1 located at 2.36 and 2.90 nm, respectively. The previous study [12] also showed a similar bimodal distribution.

mutant, $\hat{S} \sim 0.22$ nm, $R_{\text{avg}} = 4.63$ nm). Two matters should be emphasized: (i) the real pair distribution of the 65/135 does not have to be Gaussian-like; (ii) we have no intent, at this stage, to uniquely determine the structural details of R1 side-chains that lead to the distributions, since different conformations of the first three bonds would likely result in other reasonable structures that agree with the experiment. Our results should just be taken as a demonstration that the values of the average distance obtained here are consistent with our knowledge about the T4L crystal structure.

The DQC time evolution data for the 61/80 mutant is shown in Fig. 8A and the respective L-curve plot is shown in the inset. The L-curve has a good shape. It indicates that the distribution should be satisfactorily recovered from the experimental data using Tikhonov regularization. Fig. 8B shows the estimation for the distance distribution of pairs for the 61/80 mutant. The estimated solution is a bimodal distribution with the respective components in the ratio of 2 to 1 being centered at 2.36 and 2.9 nm, respectively. The average distance for the main distribution of 61/80 is 2.66 nm. As we have determined from fitting to a model [12], distances between the mutants 61/80 showed a bimodal distribution, with peaks at 2.9 and 3.4 nm and relative weights of components are 2 and 1, respectively; the average distance was ca. 3.06 nm. Both methods agree on the notion that the mutant is likely to have a bimodal distribution.

6.4. Iso-1-cytochrome c folding conformation study by DEER

One of the major topics in structural biology is to understand the dynamic conformations of biomolecules as they fold into or unfold from their tertiary and secondary structures. A complete study of dynamic protein conformation consists of two types of information: the kinetics of folding/unfolding and the conformations of possible intermediates. The former has been studied by Scholes et al. using the rapid-mix flow ESR technique [45–47] on the heme-containing protein iso-1-cyt *c*. In the present study we used iso-1-cyt *c* double labeled at S47C and K79C to study the pair distributions in the folded protein and in partially unfolded states induced by variation in the GdnHCl concentration.

The time evolution data obtained from DEER experiments at 17.3 GHz for iso-1cyt *c* in its completely folded state is shown in Fig. 9A, with the L-curve plotted in the inset. The optimal λ is located well in the corner of the L-curve, although the corner is not as sharp as those in the model studies. The estimated distribution is shown in Fig. 9B, a narrow pair distribution (as compared with the distribution from the unfolded iso-1-cyt *c* shown below) with an average distance of 1.96 nm for the main distribution. An average distance ca. 1.4 nm was estimated with the cw ESR method using the spectral convolution scheme [48], which assumes the additional spectral broadening in a double-labeled protein can be approximated as a Lorentzian r^{-6} broadening, but no information about the pair distribution is available from the cw ESR.

Two cases of the spectra from the mutant iso-1-cyt *c* in different partially folded states are shown in Figs. 10 and 11, which represent the proteins being unfolded in [GdnHCl] of 0.7 and 1.5 M, respectively. Note that the mutant iso-1-cyt *c* was found to be completely unfolded in 2.0 M GdnHCl. For the case of the partially folded state in 0.7 M GdnHCl, the optimal λ locates properly around the corner of the L-curve. The estimated distribution is plotted in Fig. 10B, whose average distance

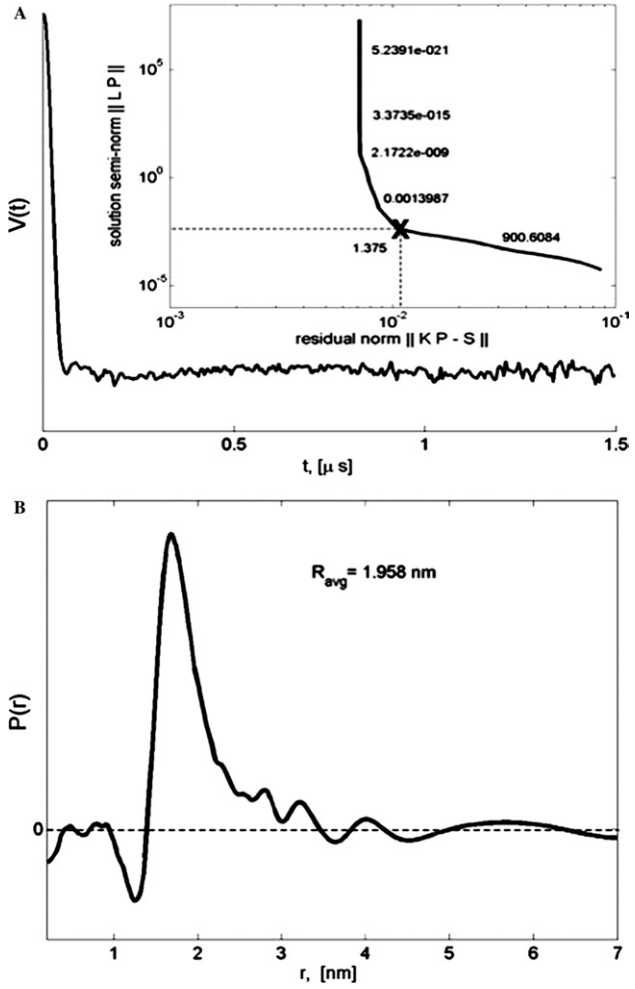

Fig. 9. (A) The dipolar time evolution spectrum of the iso-1-cyt *c* S47C/K79C mutant, with 5 ns sampling spacing, obtained from four-pulse DEER experiments. The proteins are in a completely folded state. The L-curve for selecting the optimal regularization parameter, 1.375 (marked by a cross) is shown as an inset. (B) The estimated distribution with an average distance ca. 1.96 nm for the main distribution. The distribution is narrower as compared to other cases in the present study.

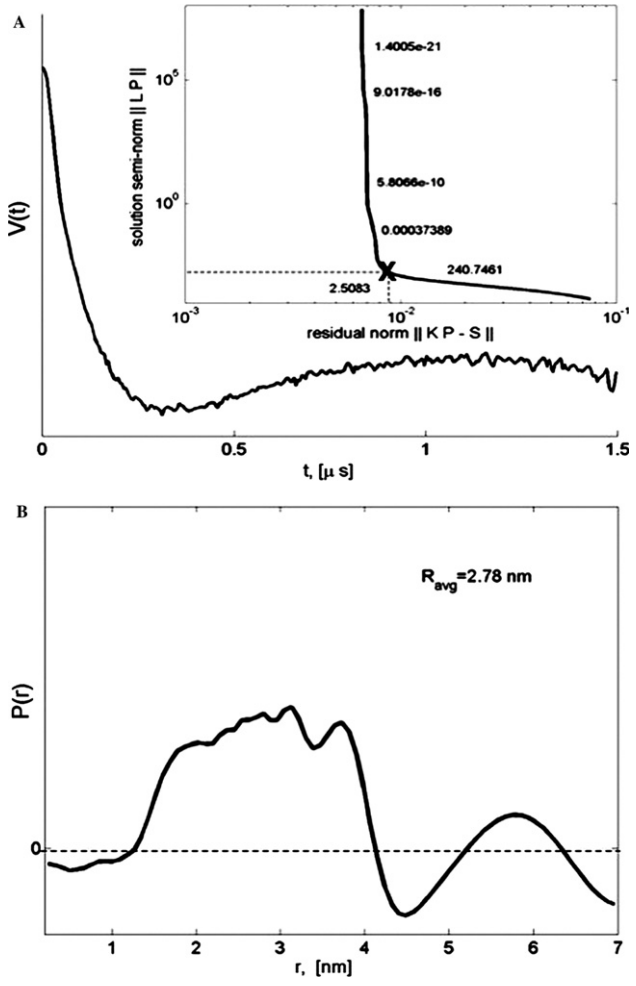

Fig. 10. (A) The dipolar time evolution spectrum of the iso-1-cyt *c* S47C/K79C mutant, with 5 ns sampling spacing, obtained from four-pulse DEER experiments at 17.3 GHz. The proteins are in a partially folded state induced with [GdnHCl] = 0.7 M. The L-curve for selecting the optimal regularization parameter, 2.508 (marked by a cross) is shown as an inset, where the optimal regularization parameter locates properly around the corner. (B) The estimated distribution with an average distance 2.78 nm from the main distribution. The distribution is broadened as compared to the case of the completely folded state.

is ca. 2.78 nm for the main distribution. The pair distribution is much broader than in the completely folded state, and its average distance is found be larger due to partial folding. This information, which indicates less restricted motion of the R1 side-chain and larger distances between the probes, is in good agreement with general expectations on the partially folded state of the protein. This broadened distribution has also been suggested by a cytochrome *c* folding landscape mapped from a FRET study [49].

The other case for a more unfolded state of the mutant iso-1-cyt *c* is shown in Fig. 11. The experimental time evolution data (cf. Fig. 11A) shows that there is significant noise, which is greater than in the other cases studied in this work. The optimal λ seems to fail to be located properly at the corner of the L-curve (cf. the inset of Fig. 11A). This apparently is caused by the increased noise in the time evolution data as shown in Fig. 11A, as compared to the DQC dipolar spectra shown in Figs. 7A and 8A. At present we cannot adequately assess how reliably the pair distribution is reconstructed, since the deviation of the optimal λ from the corner indicates the distribution may have been poorly recovered. Based on the model studies shown in Figs. 5 and 6 which demonstrate that the distribution may not be faithfully recovered under severe noise condition but the average distance is still reasonable, we believe that the estimated distribution (Fig. 11B) still provides some important information: (i) the obtained broad pair

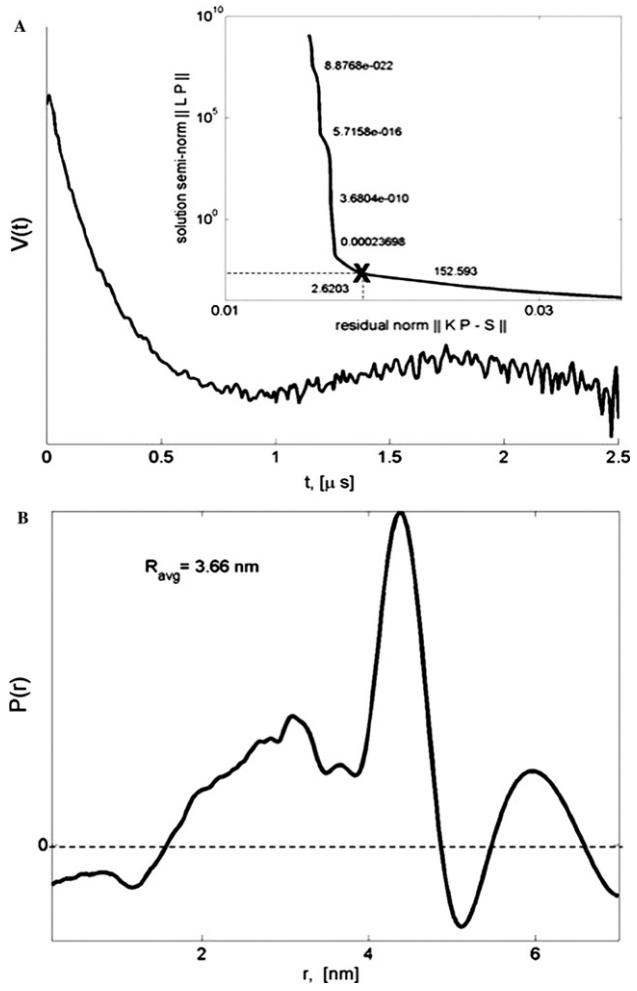

Fig. 11. (A) The dipolar time evolution spectrum of the iso-1-cyt *c* S47C/K79C mutant obtained with 10 ns sampling spacing from four-pulse DEER experiments at 17.3 GHz. The proteins are in a partially folded state induced with [GdnHCl] = 1.5 M. (The proteins are fully unfolded in 2.0 M GdnHCl). The L-curve for selecting the optimal regularization parameter, 2.62 (marked by a cross) is shown as an inset, where the optimal regularization parameter is found to slightly deviate from the corner due to great noise in the spectrum. (B) The estimated distribution with an average distance 3.66 nm for the main distribution. The distribution is broadened and of a different type from Fig. 10B.

distribution (as compared to the folded state) supports the fact that the protein is far from being in its folded state; (ii) the average intramolecular distance between two probes is ca. 3.66 nm for the main distribution, which indicates the probes at 47C and 79C are indeed more distant than the case of [GdnHCl] = 0.7 M. In general, the SNR in the dipolar spectra obtained from four-pulse DEER experiments was not as large as that found for T4 lysozyme in six-pulse DQC-ESR, as can be seen in Figs. 7–11. This may be due to the fact that in this case the distributions were much broader making the separation of intramolecular and intermolecular parts of signals more problematic.

6.5. Comments on solving the inverse problem

Inverse problems are encountered in various fields of science, such as problems arising in image reconstruction, determination of thermal parameters of the material from temperature measurements, and X-ray tomography in medical applications [26]. It is quite common for linear inverse problems to lead to an integral equation of the first kind, which is usually found to be ill-posed in dealing with experimental data. In a strict sense, the solution to the ill-posed problem might not be unique and/or might not depend continuously on the measured data. While the study of inverse problems usually involves questions about how to enforce uniqueness by additional information or assumptions, an appropriate regularization method is necessary to treat the instability of the problem and obtain reliable solution reconstruction, [50].

In addition to the inverse problem presented in the present study, there are several spectroscopic techniques requiring solving the inverse problem of the extraction important information from experimental data, such as the dePake-ing [51] approach for deconvoluting NMR spectra, and measuring the distance between two different chromophores in FRET.

DePake-ing of NMR spectra is a procedure for decomposition of the deuterium NMR spectrum of the Pake doublets from randomly oriented domains and/or from nuclei with different quadrupolar couplings into the equivalent spectra from an oriented system. The general dePake-ing problem is closely related to the ESR problem of recovering the distance distribution, so it can be solved using either an iterative algorithm [52] or regularization methods such as Tikhonov regularization [53–55] whose regularization parameter was usually determined using a self-consistency (SC) method [56,57]. The SC method for determining the regularization parameter has by far been found useful in various fields, including NMR and photo-correlation spectroscopy [58]. In the initial period of this work the SC method was used for the narrow distributions of T4 Lysozyme. However, due to the absence of a graphical interface in the SC method, we found it more convenient to use the L-curve in the present study. Furthermore, we have found that the L-curve gives reliable results even for broad distributions.[2]

[2] *Added in revision:* We have considered a model distribution function utilized in [28] that leads to a broad distribution; i.e., a sum of three equally weighted Gaussian peaks at 3, 4, and 5 nm with respective standard deviations of 0.4, 0.5, and 0.2 nm and a noise level ca. 0.025 (the lowest SNR tested in [28]). We found that the use of the L-curve criterion provided a very good reconstruction of $P(r)$ for this case, which was found to be troublesome with the SC method in [28] using the form of the kernel of Eq. (2).

The FRET technique is a popular technique for estimating pair distance distributions from the measured spectral data. However, the kernel of the equation for FRET is in the simple form of an exponential function, which is significantly different from the kernel given by Eq. (2). Such an exponential kernel can be solved with various exponential analysis methods [59]. The distribution of distances between donor and acceptor-labeled residues in a polypeptide thus can be extracted from an analysis of FRET kinetics using least-squares fitting for general cases [49]. The problem becomes ill-posed in the presence of noise. A method [60], which is a combination of the Maximum entropy method and the Tikhonov regularization algorithm, has also been used in a complementary fashion to adjust the least-squares fitting results in case of the significant noise.

7. Summary and conclusions

This study, using the Tikhonov regularization method for extracting distance distributions in pairs from dipolar time evolution spectra, has demonstrated an enhanced potential of pulsed ESR in its applications to protein structure. The kernel of the equation, which connects the dipolar time evolution spectra to the distance distributions in pairs, has been shown to result in an ill-posed problem of solution reconstruction. Therefore, a mathematically stable regularization method was used to solve the inverse problems encountered in this study. Its application to simulated data, for both a bimodal Gaussian-like distribution and a box-like distribution, has shown that Tikhonov regularization could yield reliable and rather good solutions for reasonable SNRs. Furthermore, experimental data from DQC-ESR and DEER techniques have been analyzed by this method. The pair distance distributions for the doubly labeled 65/135 and 61/80 T4 Lysozyme mutants estimated using Tikhonov regularization are consistent with the known crystal structure of T4L and our previous study. In the analysis of pulsed ESR experiments on mutants of cytochrome *c*, the pair distributions depend on the degree of folding, which is of value in studying protein folding.

The unique minimizer obtained by Tikhonov regularization guarantees the optimal solution to the ill-posed problem of interest for the level of noise in the experimental data. Given continuing improvement in the quality of the dipolar data provided by DQC-ESR and DEER, these methodologies should greatly benefit protein structural studies based on using distance constraints from them. With further improvement in decreasing noise, distance distributions extracted by Tikhonov regularization should be able to provide a more detailed picture of the structure of bilabeled biomolecules.

Acknowledgments

YWC thanks Professor Per Christian Hansen at Technical University of Denmark for many useful discussions about the L-curve method during the course of this work. The samples of T4L and iso-1-cytochrome *c* were kindly supplied by H.S. Mchaourab and C.P. Scholes. This work was supported by grants from NIH/GM, NIH/NCRR and NSF/Chemistry. A preliminary account of this work was presented at the 45th Rocky Mountain Conference on EPR (Denver, CO, July 2003) and the 5th Meeting of the European Federation of EPR Groups, (Lisbon, Portugal, Sept, 2003).

References

[1] L.J. Berliner, G.R. Eaton, S.S. Eaton, Biological magnetic resonance: distance measurements in biological systems by EPR, Kluwer Academic/Plenum Publishers, New York, 2000.
[2] M.D. Rabenstein, Y.K. Shin, Determination of the distance between two spin labels attached to a macromolecule, Proc. Natl. Acad. Sci. USA 92 (1995) 8239–8243.
[3] P.G. Wu, L. Brand, Resonance energy-transfer—methods and applications, Anal. Biochem. 218 (1994) 1–13.
[4] N.G. Walter, K.J. Hampel, K.M. Brown, J.M. Burke, Tertiary structure formation in the hairpin ribozyme monitored by fluorescence resonance energy transfer, EMBO J. 17 (1998) 2378–2391.
[5] W.L. Hubbell, C. Altenbach, Investigation of structure and dynamics in membrane-proteins using site-directed spin-labeling, Curr. Opin. Struc. Biol. 4 (1994) 566–573.
[6] K. Bhargava, J.B. Feix, Membrane binding, structure, and localization of cecropin–mellitin hybrid peptides: a site-directed spin-labeling study, Biophys. J. 86 (2004) 329–336.
[7] H.J. Steinhoff, B. Suess, Molecular mechanisms of gene regulation studied by site-directed spin labeling, Methods 29 (2003) 188–195.
[8] M. Persson, J.R. Harbridge, P. Hammarstrom, R. Mitri, L.G. Martensson, U. Carlsson, G.R. Eaton, S.S. Eaton, Comparison of electron paramagnetic resonance methods to determine distances between spin labels on human carbonic anhydrase II, Biophys. J. 80 (2001) 2886–2897.
[9] P.P. Borbat, J.H. Freed, in: L.J. Berliner, S.S. Eaton (Eds.), Biological Magnetic Resonance, Kluwer Academic/Plenum Publishers, New York, 2000, pp. 383–459.
[10] P.P. Borbat, J.H. Freed, Multiple-quantum ESR and distance measurements, Chem. Phys. Lett. 313 (1999) 145–154.
[11] S. Saxena, J.H. Freed, Theory of double quantum two-dimensional electron spin resonance with application to distance measurements, J. Chem. Phys. 107 (1997) 1317–1340.
[12] P.P. Borbat, H.S. Mchaourab, J.H. Freed, Protein structure determination using long-distance constraints from double-quantum coherence ESR: study of T4 lysozyme, J. Am. Chem. Soc. 124 (2002) 5304–5314.
[13] P.P. Borbat, J.H. Davis, S.E. Butcher, J.H. Freed, Measurement of large distances in biomolecules using double-quantum filtered refocused electron spin-echoes, J. Am. Chem. Soc. 126 (2004) 7746–7747.
[14] A.D. Milov, A.B. Ponomarev, Y.D. Tsvetkov, Electron electron double-resonance in electron-spin echo—model biradical systems and the sensitized photolysis of decalin, Chem. Phys. Lett. 110 (1984) 67–72.

[15] A.M. Raitsimring, K.M. Salikhov, Electron spin echo method as used to analyze the spatial distribution of paramagnetic centers, Bull. Magn. Reson. 7 (1985) 184–217.
[16] A.D. Milov, A.G. Maryasov, Y.D. Tsvetkov, Pulsed electron double resonance (PELDOR) and its applications in free-radicals research, Appl. Magn. Reson. 15 (1998) 107–143.
[17] V. Pfannebecker, H. Klos, M. Hubrich, T. Volkmer, A. Heuer, U. Wiesner, H.W. Spiess, Determination of end-to-end distances in oligomers by pulsed EPR, J. Phys. Chem. 100 (1996) 13428–13432.
[18] M. Pannier, S. Veit, A. Godt, G. Jeschke, H.W. Spiess, Dead-Time free measurement of dipole–dipole interactions between electron spins, J. Magn. Reson. 142 (2000) 331–340.
[19] C.G. Hoogstraten, C.V. Grant, T.E. Horton, V.J. DeRose, R.D. Britt, Structural analysis of metal ion ligation to nucleotides and nucleic acids using pulsed EPR spectroscopy, J. Am. Chem. Soc. 124 (2002) 834–842.
[20] A.D. Milov, Y.D. Tsvetkov, F. Formaggio, S. Oancea, C. Toniolo, J. Raap, Aggregation of spin labeled trichogin GA IV dimers: distance distribution between spin labels in frozen solutions by PELDOR data, J. Phys. Chem. B 107 (2003) 13719–13727.
[21] O. Schiemann, A. Weber, T.E. Edwards, T.F. Prisner, S.T. Sigurdsson, Nanometer distance measurements on RNA using PELDOR, J. Am. Chem. Soc. 125 (2003) 3434–3435.
[22] M. Bennati, A. Weber, J. Antonic, D.L. Perlstein, J. Robblee, J.A. Stubbe, Pulsed ELDOR spectroscopy measures the distance between the two tyrosyl radicals in the R2 subunit of the *E. coli* ribonucleotide reductase, J. Am. Chem. Soc. 125 (2003) 14988–14989.
[23] A.D. Milov, Y.D. Tsvetkov, F. Formaggio, S. Oancea, C. Toniolo, J. Raap, Solvent effect on the distance distribution between spin labels in aggregated spin labeled trichogin GA IV dimer peptides as studied by pulsed electron–electron double resonance, Phys. Chem. Chem. Phys. 6 (2004) 3596–3603.
[24] G. Jeschke, C. Wegener, M. Nietschke, H. Jung, H.-J. Steinhoff, Interresidual distance determination by four-pulse double electron–electron resonance in an integral membrane protein: the Na+/proline transporter PutP of *Escherichia coli*, Biophys. J. 86 (2004) 2551–2557.
[25] G.E. Pake, Nuclear resonance absorption in hydrated crystals—fine structure of the proton line, J. Chem. Phys. 16 (1948) 327–336.
[26] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer Academic Publishers, Dordrecht, Boston, 1996.
[27] G. Jeschke, A. Koch, U. Jonas, A. Godt, Direct conversion of EPR dipolar time evolution data to distance distributions, J. Magn. Reson. 155 (2002) 72–82.
[28] G. Jeschke, G. Panek, A. Godt, A. Bender, H. Paulsen, Data analysis procedures for pulse ELDOR measurements of broad distance distributions, Appl. Magn. Reson. 26 (2004) 223–244.
[29] M.K. Bowman, A.G. Maryasov, N. Kim, V.J. DeRose, Visualization of distance distribution from pulsed double electron–electron resonance data, Appl. Magn. Reson. 26 (2004) 23–39.
[30] A.N. Tikhonov, Numerical Methods for the Solution of Ill-posed Problems, Kluwer Academic Publishers, Dordrecht, Boston, 1995.
[31] J.R. Klauder, P.W. Anderson, Spectral diffusion decay in spin resonance experiments, Phys. Rev. 125 (1962) 912–932.
[32] A. Abragam, The Principles of Nuclear Magnetism, Clarendon Press, Oxford, 1961.
[33] W.H. Press, Numerical Recipes in Fortran 77 the Art of Scientific Computing, Cambridge University Press, Cambridge, New York, 2000.
[34] C.W. Groetsch, The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Pitman Advanced Pub. Program, Boston, 1984.
[35] S. Twomey, On numerical solution of Fredholm integral equations of first kind by inversion of linear system produced by quadrature, JACM 10 (1963) 97.
[36] A.N. Tikhonov, Solution of incorrectly formulated problems and regularization method, Dokl. Akad. Nauk. Sssr+ 151 (1963) 501.
[37] K. Miller, Least squares methods for ill-posed problems with a prescribed bond, SIAM J. Math. Anal. 1 (1970) 52–74.
[38] C.L. Lawson, R.J. Hanson, Solving Least Squares Problems, Prentice-Hall, Englewood Cliffs, NJ, 1974.
[39] P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, SIAM Rev. 34 (1992) 561–580.
[40] P.C. Hansen, D.P. Oleary, The use of the L-curve in the regularization of discrete ill-posed problems, SIAM J. Sci. Comput. 14 (1993) 1487–1503.
[41] P.C. Hansen:, Rank-deficient and Discrete Ill-posed Problems: Numerical Aspects of Linear Inversion, SIAM, Philadelphia, 1997.
[42] P.C. Hansen, in: P. Johnston (Ed.), Computational Inverse Problems in Electrocardiology, WIT Press, 2000, pp. 119–142.
[43] C.R. Vogel, Non-convergence of the L-curve regularization parameter selection method, Inverse Probl. 12 (1996) 535–547.
[44] P.C. Hansen, Regularization tools Version 3.0 for Matlab 5.2, Numer. Algorithms 20 (1999) 195–196.
[45] V.M. Grigoryants, K.A. DeWeerd, C.P. Scholes, Method of rapid mix EPR applied to the folding of bi-spin-labeled protein as a probe for the dynamic onset of interaction between sequentially distant side chains, J. Phys. Chem. B 108 (2004) 9463–9468.
[46] K.B. Qu, J.L. Vaughn, A. Sienkiewicz, C.P. Scholes, J.S. Fetrow, Kinetics and motional dynamics of spin-labeled yeast iso-1-cytochrome *c*. 1. Stopped-flow electron paramagnetic resonance as a probe for protein folding/unfolding of the C-terminal helix spin-labeled at cysteine 102, Biochemistry 36 (1997) 2884–2897.
[47] K. DeWeerd, V. Grigoryants, Y. Sun, J.S. Fetrow, C.P. Scholes, EPR-detected folding kinetics of externally located cysteine-directed spin-labeled mutants of iso-1-cytochrome c, Biochem. 40 (2001) 15846–15855.
[48] H.S. Mchaourab, K.J. Oh, C.J. Fang, W.L. Hubbell, Conformation of T4 lysozyme in solution. Hinge-bending motion and the substrate-induced conformational transition studied by site-directed spin labeling, Biochemistry 36 (1997) 307–316.
[49] J.G. Lyubovitsky, H.B. Gray, J.R. Winkler, Mapping the cytochrome *c* folding landscape, J. Am. Chem. Soc. 124 (2002) 5481–5485.
[50] V.A. Morozov, M. Stessin, Regularization Methods for Ill-posed Problems, CRC Press, Boca Raton, FL, 1993.
[51] E. Sternin, M. Bloom, A.L. Mackay, De-Pake-Ing of Nmr-Spectra, J. Magn. Reson. 55 (1983) 274–282.
[52] M. Bloom, J.H. Davis, A.L. Mackay, Direct determination of the oriented sample Nmr-spectrum from the powder spectrum for systems with local axial symmetry, Chem. Phys. Lett. 80 (1981) 198–202.
[53] H. Schafer, B. Madler, F. Volke, De-Pake-ing of NMR powder spectra by nonnegative least-squares analysis with Tikhonov regularization, J. Magn. Reson. A 116 (1995) 145–149.
[54] H. Schafer, B. Madler, E. Sternin, Determination of orientational order parameters from H-2 NMR spectra of magnetically partially oriented lipid bilayers, Biophys. J. 74 (1998) 1007–1014.
[55] F.G. Vogt, D.J. Aurentz, K.T. Mueller, Determination of internuclear distances from solid-state nuclear magnetic resonance: dipolar transforms and regularization methods, Mol. Phys. 95 (1998) 907–919.
[56] J. Weese, A reliable and fast method for the solution of Fredholm integral-equations of the 1st kind based on Tikhonov regularization, Comput. Phys. Commun. 69 (1992) 99–111.

[57] J. Weese, A regularization method for nonlinear ill-posed problems, Comput. Phys. Commun. 77 (1993) 429–440.
[58] J. Honerkamp, D. Maier, J. Weese, A nonlinear regularization method for the analysis of photon-correlation spectroscopy data, J. Chem. Phys. 98 (1993) 865–872.
[59] A.A. Istratov, O.F. Vyvenko, Exponential analysis in physical phenomena, Rev. Sci. Instrum. 70 (1999) 1233–1257.
[60] U. Amato, W. Hughes, Maximum-entropy regularization of Fredholm integral-equations of the 1st kind, Inverse Probl. 7 (1991) 793–808.