

SDSL-ESR-based protein structure characterization

Janez Štrancar · Aleh Kavalenka · Iztok Urbančič ·
Ajasja Ljubetič · Marcus A. Hemminga

Received: 15 May 2009 / Accepted: 23 June 2009 / Published online: 11 August 2009
© European Biophysical Societies' Association 2009

Abstract As proteins are key molecules in living cells, knowledge about their structure can provide important insights and applications in science, biotechnology, and medicine. However, many protein structures are still a big challenge for existing high-resolution structure-determination methods, as can be seen in the number of protein structures published in the Protein Data Bank. This is especially the case for less-ordered, more hydrophobic and more flexible protein systems. The lack of efficient methods for structure determination calls for urgent development of a new class of biophysical techniques. This work attempts to address this problem with a novel combination of site-directed spin labelling electron spin resonance spectroscopy (SDSL-ESR) and protein structure modelling, which is coupled by restriction of the conformational spaces of the amino acid side chains. Comparison of the application to four different protein systems enables us to generalize the new method and to establish a general procedure for determination of protein structure.

Keywords Rotational conformational space modelling (CSM) · ESR/EPR spectral simulation and optimization · GHOST condensation · Protein structure optimization · Site-directed spin labelling (SDSL)

Abbreviations

NMR	Nuclear magnetic resonance
SDSL	Site-directed spin-labelling
ESR	Electron spin resonance
EPR	Electron paramagnetic resonance
GHOST	Condensation algorithm that filters and groups the solutions found in optimization runs
NTAIL	C-terminal domain of nucleoprotein of the measles virus

Introduction

Proteins are key molecules in cells of living organisms, including human beings. Knowledge about protein structure and function provides important insights and practical applications in medicine, agriculture, nutrition, and industry (Lehninger et al. 2005). The most powerful techniques of protein structure determination are X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Nevertheless, for proteins that are difficult to crystallize, or to concentrate, very limited structural information is available. Therefore, it is not surprising that determination of the structures of membrane proteins is one of the most challenging fields of structural biology and structural proteomics (Lacapere et al. 2007; Torres et al. 2003). Because of the very heterogeneous environment in which they are found, classical methods have difficulties with

The more you see: Spectroscopy in molecular biophysics.

J. Štrancar (✉) · A. Kavalenka · I. Urbančič · A. Ljubetič
Laboratory of Biophysics, Solid State Physics Department,
“Jožef Stefan” Institute, Jamova 39, 1000 Ljubljana, Slovenia
e-mail: Janez.Strancar@ijs.si

A. Kavalenka · M. A. Hemminga
Laboratory of Biophysics, Wageningen University,
Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

determination of the structure of membrane proteins. In fact, less than 1% of known protein structures correspond to membrane proteins, although one-third of all proteins are membrane proteins (White 2009).

X-ray crystallography and NMR spectroscopy are also less successful in determination of the structure of so-called intrinsically unstructured or intrinsically disordered proteins (Dyson and Wright 2005). Intrinsically disordered proteins (IDP) consist of dynamic ensembles of inter-converting conformers and they exert their biological function by recognizing their binding partners through their disordered regions (Bourhis et al. 2007; Dunker et al. 2001, 2005, 2008; Dyson and Wright 2005; Ferron et al. 2006; Fink 2005; Receveur-Bréchet et al. 2006; Tompa 2002; Uversky 2002; Wright and Dyson 2009). Because of their inherent flexibility, intrinsically disordered proteins generally fail to crystallize in the absence of their partner(s). In the rare cases where crystallization of the free form is successful, it only leads to a snapshot of a single conformation that is not representative of the whole conformational ensemble (Timsit et al. 2006).

Difficulties in the application of standard high-resolution methods for characterization of the three-dimensional structure of intrinsically disordered and membrane proteins therefore call for the development of alternative approaches. Low-resolution structural data can be obtained with small-angle X-ray scattering (SAXS) (Petoukhov and Svergun 2005; Svergun and Koch 2003), circular dichroism (CD) (Fasman 1996; Kelly and Price 2000; Uversky 2002), and atomic force microscopy (AFM) (Muller and Engel 2008; Pebay-Peyroula 2008). Molecular dynamics simulations and other computational techniques reinforce alternative experimental methods, for example NMR spectroscopy (Arora and Tamm 2001; Castellani et al. 2002; Dominguez et al. 2003; Watts et al. 2004; Wüthrich 1986) and electron microscopy (EM) (Fleishman et al. 2006; Henderson 2004). One alternative technique is site-directed spin labelling (SDSL) electron spin resonance (ESR). This technique enables both structural and dynamic characterization of the local conformations of a membrane protein (or any other protein) in its native environment (Alexander et al. 2008; Fanucci and Cafiso 2006; Hemminga 2007; Stopar et al. 2005, 2006). Site-directed spin labelling at multiple sites of proteins has been widely used for protein structure characterization (Huang and Cafiso 2008; Hubbell et al. 1998; Jao et al. 2008; Li and Fung 2009; Pistolesi et al. 2006).

Recently, we developed a novel approach for protein structure characterization (Fig. 1), based on modelling of the conformational space of the side chains of the amino acid residues. This methodology makes use of structural constraints extracted from SDSL-ESR spectroscopic data at

multiple protein sites. In this paper we provide an overview of our latest progress in this field.

Structure characterization based on side chain conformational space restrictions

Initially, our method of protein structure determination was applied to membrane-embedded M13 major coat protein (Štrancar et al. 2009) by using a combination of high-throughput SDSL-ESR experiments (Stopar et al. 2006) and protein modelling. The basic unit of the methodology is the restricted conformational space. Comparison of the simulated restrictions of the rotational conformational space of spin label side chains and the experimental conformational space determined from SDSL-ESR spectra is used to optimize the dihedral angles of the protein backbone, and its relative position and orientation (Fig. 1b). This comparison leads to a family of favourable three-dimensional structures of a protein in a protein–lipid system (Kavalenka et al. 2009a). At a later stage, this approach was employed for characterization of the structure of intrinsically disordered NTAIL protein in a complex with a partner protein XD (both are measles virus proteins) under different experimental conditions (Belle et al. 2008; Kavalenka et al. 2009b). Recently, several improvements were introduced in the modelling of the conformational space and the calculation of the restrictions of the conformational space. As will be described in this paper, these advances in methodology were checked by application of the method to two other proteins: human pancreatic lipase protein (Belle et al. 2007) and equinatoxin II (Malovrh et al. 2003).

SDSL-ESR-detected local restrictions

In the experimental part of the methodology, site-directed mutagenesis is used to replace a strategically chosen amino acid residue with a cysteine. This cysteine is then targeted by a spin label (Hubbell et al. 1998). After protein purification, concentration, and—if needed—reconstitution into the membrane, the spin-labelled protein sample is prepared for ESR experiments (Fig. 1a). Taking into account that the spectral lineshape is highly sensitive to the motional properties of the spin label, the temperature has to be chosen in such a way that these motional properties will depend primarily on the protein structure, and not on the internal label properties. As a result of this condition the conformational space of a spin label attached to a protein, to be fully exploited, should be restricted by structural elements such as the protein backbone, rotational space of neighbouring amino acids, and lipids (for a membrane protein).

At low temperature the conformational space degenerates into a small number of low-energy conformational

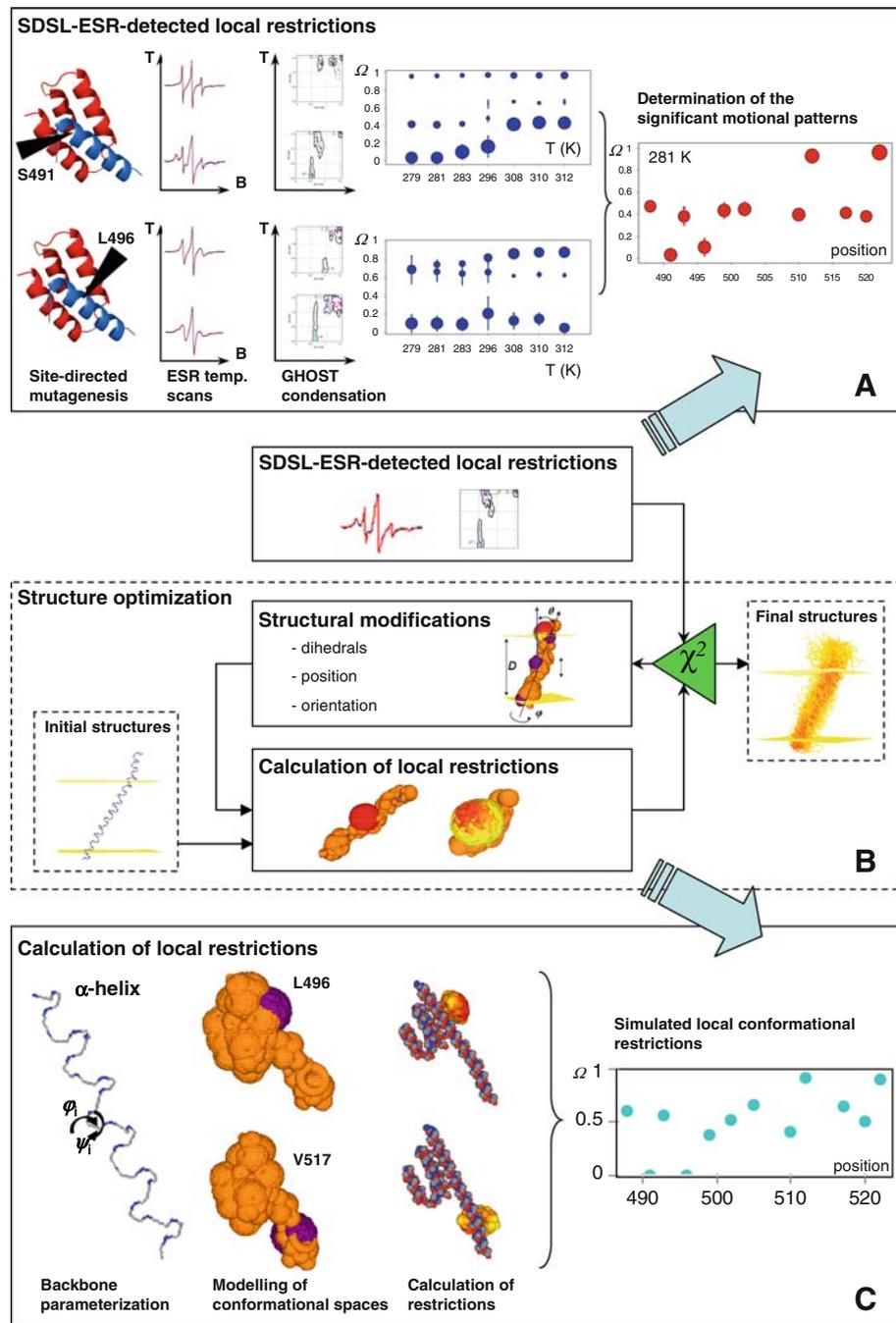


Fig. 1 Overview of the SDSL-ESR approach for protein structure determination. **a** Detection of the local restrictions from SDSL-ESR spectroscopic data. The method is illustrated for the NTAIL–XD protein complex spin labelled at two typical mutant positions S491 and L496. The method is based on site-directed mutagenesis, measurements of ESR spectra at different temperatures (the experimental spectra are shown in blue, the simulated spectra are in red), GHOST condensation, and determination of the significant motional patterns for each mutant position. **b** The structure determination approach is based on modelling of the conformational space of the amino acid side chains. This modelling is coupled to SDSL-ESR

spectroscopy, and enhanced by structure optimization. Optimization starts with an initial structure, and finally provides a family of favourite structures. **c** The determination of local restrictions is based on modelling of the conformational space. This method includes modelling of the protein structure (parameterized by pairs of backbone dihedral angles φ and ψ), modelling of conformational spaces of the side chains, and calculation of the conformational space restrictions. The method is illustrated for the NTAIL–XD protein complex spin labelled at two typical mutant positions L496 (at the interacting part of the NTAIL–XD complex) and V517 (at the disordered part of NTAIL) (Kavalenka et al. 2009b)

states (rotamers), which depend mostly on minimization of the internal energy in accordance with the angular potentials of the rotamer. However, when the temperature increases, the side chains start to exploit their full rotational spaces, which at the same time become restricted because of steric overlap with the more rigid backbone and because of sharing physical space with the side chains of the neighbouring amino acids. If the temperature is increased even further, the backbone can lose its stable conformation. This state of the backbone will lead to an undesired situation for our methodology and should be avoided. Therefore, the lifetime of the protein backbone should be long in comparison with the nanosecond time-scale of the ESR experiment. Under this condition, slow backbone motions will not be reflected in the ESR spectral lineshape, which is then mainly determined by the fast motions of the spin label side chain. If the lifetime of a protein conformation is shorter than the ESR time window, both backbone and side chains will contribute to the ESR spectra, preventing the extraction of useful structural information from the lineshape. In such a case, the protein backbone dynamics might be slowed down by increasing the viscosity of the environment, e.g., by addition of sucrose, or by lowering the temperature.

In any case, it is advantageous to measure ESR spectra at several temperatures. After acquiring a temperature series of ESR spectra at each mutant position, spectral simulations and optimizations are used to extract the appropriate motional patterns (Fig. 1a). Because the approach of ESR spectra simulation has been developed previously (Schindler and Seelig 1973; Štrancar et al. 2000), only the main issues related to the physical background of the spectral parameters are discussed here. First, two parameters are used to parameterize the partial averaging of the rotational motion within a cone model, i.e., by defining the anisotropy of the cone with the opening angle ϑ_0 and an asymmetry angle φ_0 . Second, the traces of the interaction tensors \mathbf{g} and \mathbf{A} are linearly corrected with the parameters p_A (Marsh 1981) and $Prot$ (Steinhoff et al. 2000) that take into account the effects of polarity and proticity, respectively. Third, when calculating the convolution of the magnetic field distribution and basic lineshape, two linewidth parameters are also used: a single (effective) rotational correlation time, τ_c , and an additional broadening constant W . The rotational correlation time defines a Lorentzian-type lineshape in the motional narrowing approximation (Nordio 1976), while the additional broadening constant arises primarily from unresolved hydrogen superhyperfine interactions and contributions from paramagnetic impurities (e.g., oxygen), external magnetic field inhomogeneities, field modulation effects, and spin–spin interaction.

To take into account a superposition of motional patterns, which often arise because of multiple global

conformations or because of heterogeneity at local sites, the basic set of parameters ϑ_0 , φ_0 , τ_c , W , p_A , and $Prot$ is expanded for each of the N_c spectral components. In addition, there are $N_c - 1$ weights, d , of these spectral components. Thus altogether, there are $7N_c - 1$ spectral parameters, which have to be resolved by the spectral optimization routine. Taking into account the resolution limit of spin label ESR to be around 30 parameters, this allows the use of at most four spectral components (Stopar et al. 2006; Štrancar 2007; Štrancar et al. 2005). A hybrid evolutionary optimization (HEO), a combination of a genetic algorithm and a downhill-simplex local search (Filipič and Štrancar 2001; Štrancar et al. 2005) is used to find the set of spectral parameters that produces the best fit to the experimental ESR spectrum. To guide the optimization, which solves an inverse problem, a common fitness function is introduced. Typically the fitness function is the reduced χ^2 , calculated from the sum of the squared residuals between the experimental and simulated spectral points divided by the squared standard deviation of the experimental points and by the number of points in the experimental spectrum (in our case 1,024).

The hybrid evolutionary optimization routine starts with a random initialization of a population of 400 solutions, and continues with the tournament selection and application of genetic operators (i.e., three-point crossover, uniform mutation, and local improvements performed with downhill-simplex) for 100 generations (Filipič and Štrancar 2001; Štrancar et al. 2005). The elite set is used to keep track of the best individuals. The implementation of a shaking operator guarantees diversity even within a single hybrid evolutionary optimization run (Kavalenka et al. 2005), reducing the number of optimization runs down to 20. Such a number is sufficient to accumulate a final set of 200 sets of parameters of spectral components, which are then filtered, grouped and graphically presented with a so-called GHOST condensation algorithm. The efficiency of the spectral simulations and optimization of the spectral parameters is checked by inspecting the values of the fitness function χ^2 , which should be below 10 at a signal-to-noise ratio of about 200. Another important indicator of successful spectral optimization is an equal contribution of different runs into the final set of solutions, which is measured in terms of the run flatness parameter. This value should be above 70% (Kavalenka et al. 2005; Štrancar et al. 2005).

In the GHOST condensation algorithm, solution density filtering eliminates isolated (less frequent) solutions, while filtering against the goodness of fit only maintains the solutions which successfully describe the spectrum (Štrancar et al. 2005). The filtered solutions are grouped into domains (motional patterns) and presented in terms of two-dimensional cross-sections ($\vartheta_0 - \varphi_0$, $\vartheta_0 - \tau_c$, etc., as in Fig. 1a) for visual checking of the solutions. There is no

need to predefine the complexity (i.e., the number of spectral components when simulating the experimental spectra) at any spin label position in advance, because it is obtained automatically from the GHOST condensation and domain recognition procedure. Finally, the solution domains are parameterized by a centre-of-mass and second moment of each type of spectral parameter. Some parameters are then combined into more physically relevant quantities, such as the free rotational space Ω and normalized rotational diffusion constant D :

$$\Omega = \frac{\vartheta_0 \varphi_0}{(\pi/2)^2}, \quad (1)$$

$$D = \frac{\vartheta_0 \varphi_0}{4\tau_c}. \quad (2)$$

Both Ω and D describe the motional patterns as detected by SDSL-ESR at each mutant position. The free rotational space Ω describes the anisotropy of the rotational diffusion of a spin-labelled side chain and represents the local restrictions imposed by all surrounding structural elements. The normalized rotational diffusion constant D describes the rate of rotational diffusion of the spin label and is defined in such a way that it eliminates the influence of the anisotropy of the rotational motion on the rotational correlation time. Therefore, it reports about the environment in which the spin label is wobbling. Typically, the normalized rotational diffusion can show a transition from amino acid side chains positioned in the aqueous solution to positions within a membrane. Additionally, it superimposes also the effect of backbone mobility, i.e., if backbone movement becomes fast then the normalized rotational diffusion constant D also will increase.

Spectral simulation and automatic optimization of the spectral parameters help to characterize the site-specific motional properties in a well-defined and high-throughput manner. However, the detected motional patterns should be checked for reliability before interpretation, or before further usage for protein modelling and structural optimization. As ESR spectra are always noisy, it is impossible to analyse a single spectrum precisely. Thus, to increase the reliability of the analysis, ideally a suitable series of ESR spectra has to be measured, analyzed and interpreted jointly. Depending on the subject of the research, an experimental series of different spin labels, various environments and/or chemical concentrations can be applied. However, performing measurements and comparing results at different temperatures is the most straightforward approach to removing artefacts from the spectral analysis. This cleaning can be easily done in terms of a so-called “bubble diagram” (Figs. 1a, 2). In this diagram the average characteristic values of the chosen parameter (such as the free rotational space Ω , rotational diffusion D , etc.) are

plotted against the property of the series (e.g., temperature, or mutant position). The bubble size is related to the spectral weight of the motional pattern, and the vertical bar at each bubble represents the second moment of a distribution of that particular motional pattern in the phase space. Such a representation allows the determination of general trends of significant patterns in the data series, and recognition of numerical and computational artefacts by applying the following criteria. Reliable domains should appear regularly and consistently in the series of the external variable, e.g., by having a locally monotonous temperature dependence of Ω . Further, isolated solutions and solutions with a spectral weight below a certain threshold of a few percent are discarded. In addition, spectrally irrelevant solutions (e.g., that describe none of the spectral features) are deleted. In the final presentation only the patterns that meet *all* mentioned conditions are kept.

To illustrate this “cleaning” strategy, a check of a typical temperature-dependent series of ESR measurements at a chosen mutant position, resulting in a series of motional patterns, is shown in Fig. 2. Irregular temperature behaviour (at 279, 281, and 283 K) indicates a deviation due to inappropriate filtering of the optimization results. It is expected that the temperature dependence of any parameter of our system is monotonous, unless it feels a major structural rearrangement, for example a phase transition. It is therefore expected that motional patterns evolve smoothly in some small temperature range. Therefore, deviating solutions are very likely to be caused by numerical artefacts. For the same reason, isolated solutions originating from either insignificant motional patterns with a small spectral weight, or inappropriate spectral components fitting noisy spectral details (e.g. small-weight patterns at temperatures 308, 310, 312 K), are also irrelevant (Fig. 2a). To increase the accuracy of the structure determination, all these inappropriate solutions should be systematically removed.

Temperature dependencies can also be used to verify the main assumption in modelling the conformational space, i.e., that the backbone motion should be slow on the ESR time scale, whereas the motion of the side chains should be fast. This effect can be diagnosed by a sudden transition in the free rotational space as a function of temperature. In such a case, ESR experiments at several temperatures are needed to identify whether the protein is in a permanent disordered state, or if the lifetime of the backbone conformation is too short.

Finally, the complexity of the motional patterns should also be taken into account. More than one reliable motional pattern at a site obviously means that the spin label feels different restrictions to its wobbling space. For example, a completely unrestricted motional pattern could indicate

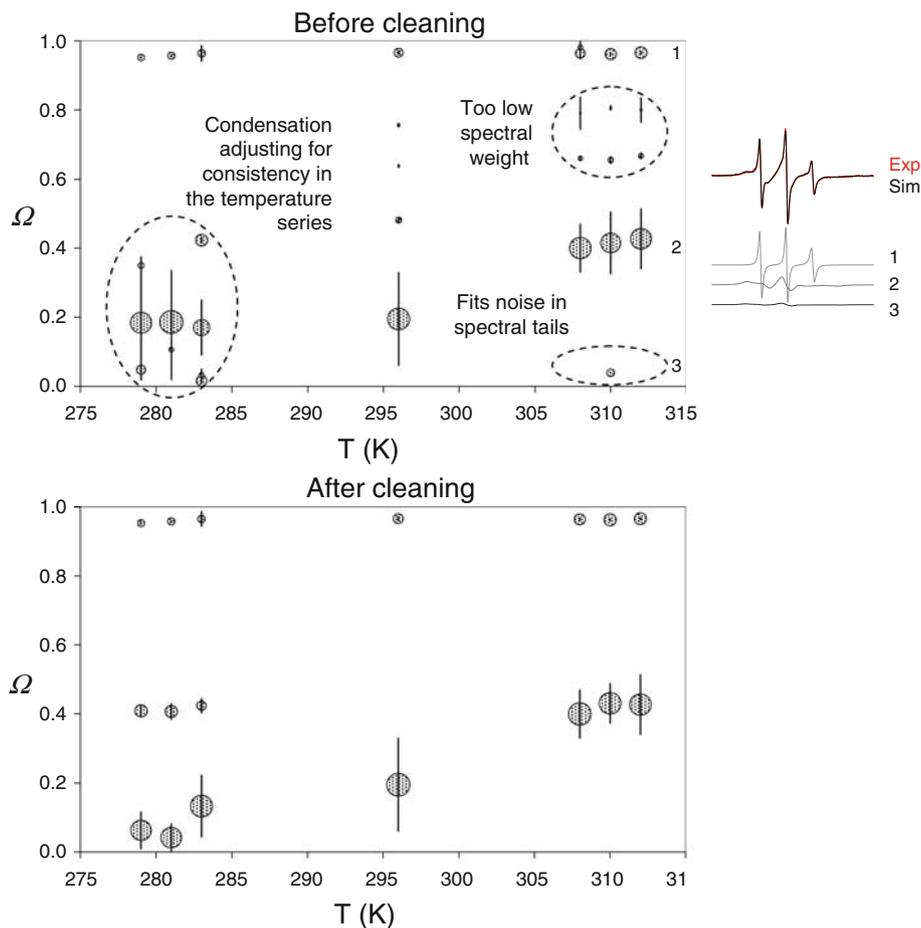


Fig. 2 An example of motional pattern cleaning. A temperature-dependent series of bubble diagrams at each spin-labelled protein site (an example of a bubble diagram is shown for the NTAIL–XD protein complex spin labelled at position S491; Kavalenka et al. 2009b) is used to detect insignificant and/or false solutions. High values of Ω (between 0.7 and 1) correspond to (nearly) unrestricted motional patterns of the spin label, whereas low values (between 0 and 0.25) imply very high restrictions. Adjustment of the condensation procedure enables achievement of consistency of motional patterns in the

temperature series (see the motional patterns at 279, 281, 283 K marked with a *dashed oval*). Removing spectral components of low intensity, or components that fit noise in the tails in the ESR spectrum (see motional patterns at 308, 310, 312 K marked with *dashed ovals*) allows focussing on the most important meaningful motional patterns. For illustration, the motional patterns at 308, 310, 312 K are numbered 1–3 and the corresponding ESR spectral components of the simulated 310 K spectrum are presented on the *right*

nonspecific labelling at unrestricted positions in the protein (i.e., the terminal ends). However, any situation where more than one motional pattern is revealed with a significant contribution means that there are coexisting local protein conformations. In modelling of the protein any of these patterns can be used, and appropriate structures can be derived accordingly.

Local restrictions determined by modelling of the conformational space

In SDSL-ESR spectroscopy a protein is labelled at a specific site with a spin label of a size slightly larger than the size of the largest amino acid residues. Together with high-temperature conditions, this guarantees that the fully

exploited conformational space of the amino acid side chain becomes restricted by steric overlap with the local backbone conformation and by the conformational spaces of the neighbouring amino acids. In addition, the surrounding phospholipids (for membrane proteins) affect the conformational space of the spin label. To employ these restrictions for protein structure determination, the conformational space has to be measured experimentally and simulated at the same time and then compared. In this respect, it is important to note that the ESR experiment is insensitive to the exact atomic coordinates, but very sensitive to the motional anisotropy of the nitroxide group. Therefore there is no need for precise calculation of a side chain conformation. Instead, the relative probability of side chain conformations in the conformational space of the

spin label has to be determined. A calculation should also take into account the average space-sharing effects of all surrounding wobbling chains from the neighbouring amino acid side chains of the protein(s) and alkyl chains of the lipids, when present.

The protein backbone structure is parameterized by the dihedral angles φ_i and ψ_i at each i -th amino acid residue, following the attachment of the amino acid side chains to the backbone (Fig. 1c). It is assumed that atomic structures of the spin label and amino acid residues can be constructed using an approximation of fixed bond lengths and bond angles (Xiang and Honig 2001), based on previously reported values (Engl and Huber 1991; MacKerell et al. 1998; Word et al. 1999). In the modelling, the unrestricted conformational spaces of all amino acid side chains are attached to their respective backbone C^α atoms, which are assumed to be fixed in space. The unrestricted conformational space is a result of the rotations of the side chain of an amino acid, or of a spin label around single bonds. The side chain is rotated around its single bonds in different steps in accordance with the type of amino acid residue (or spin label). The steps in these rotations are derived by taking into account that the most computationally demanding step in conformational space analysis is calculation of the restrictions of the conformational space. This calculation involves checking of the overlap between side chains, which quadratically depends on the number of rotamers in the conformational space. Therefore, it is clear that the number of rotamers has to be reduced as much as possible, however, while maintaining a certain degree of accuracy.

Calculation of the restrictions of the conformational space of the spin label (Štrancar et al. 2009) starts with the determination of the effect of backbone overlap. This calculation involves checking all the individual spin label conformations for overlap with the backbone atoms. While constructing the unrestricted conformational space, conformations that have internal overlap are discarded. The effective van der Waals radii, which are used for internal steric clash checking, are the original Van der Waals radii reduced for allowing a soft overlap according to Grigoryan et al. (2007), Ho et al. (2003), Shetty et al. (2003), and Tomblato et al. (2006). Because the backbone motion is assumed to be fixed, the statistical weight of a conformation that overlaps with the backbone is set to zero, i.e., all such conformations are forbidden. All conformations that are not forbidden are then checked for overlap with neighbouring amino acid side chains. This overlap is determined by a reduction of the statistical weight of the i -th conformation P_i^{initial} proportional to the number of conformations of the k -th neighbouring residue N_k^{overlaps} , with which it shares space:

$$F_k^i = \frac{N_k^{\text{all}} - N_k^{\text{overlaps}}}{N_k^{\text{all}}}, \quad (3)$$

where N_k^{all} is the total number of the k -th neighbouring residue. If there is more than one overlapping neighbouring side chain, the probability for each of the overlapping pairs of conformations is factorized:

$$P_i' = P_i^{\text{initial}} \prod_k^n F_k^i. \quad (4)$$

If the spin-labelled protein site is in a transmembrane region, the conformational space is further reduced because of the restrictive effect of the fluctuating alkyl chains of the phospholipids and the restrictive effect of lipid head groups (Marsh 2008). In contrast with the restriction calculations that arise from the backbone and side chains, in the case of lipid effect the steric overlapping effect cannot be derived explicitly by calculating the overlap of atoms and groups. This arises because the position of the atoms of the lipids is not precisely known. Therefore, it is clear that the lipid effect has to be introduced in the calculations in a more phenomenological way.

In the simplest approximation the effect of the alkyl chains of the phospholipids should take into account following assumptions (Štrancar et al. 2009):

- 1 side chain conformations, which stretch out from the main body of the protein perpendicular to the lipids, should be restricted to the greatest extent;
- 2 there are minimal restrictions in cases of parallel alignment to the membrane normal;
- 3 lipid ordering is effective as soon as there is any non-zero angle between the side chain of a spin label and a lipid alkyl chain, meaning that the derivative of the lipid effect should be linear when ϑ angle approaches zero;
- 4 perpendicular and near-perpendicular conformations should be restricted by approximately a similar extent, meaning that the derivative of the lipid effect should be zero when the angle approaches $\pi/2$; and
- 5 the amplitude of the lipid effect on the conformational space of the side chains can be deduced from the effect of electron density profile (static restriction, highest at the membrane surface; Wiener and White 1992) and from the effect of lipid chain rotational conformational space (dynamic restrictions, increasing towards the centre of membrane; Vermeer et al. 2007).

All these approximations can be merged into a probability function for the lipid effect, given by:

$$P_i = P_i'(1 - \sin \vartheta_i), \quad (5)$$

where P_i' is the statistical weight of the i -th conformation after the restrictive effects of protein backbone and

neighbouring residue side chains have been applied. This description of the lipid effect is in agreement with results of recent molecular dynamics simulations studies, which show that aromatic, polar, and charged amino acid side chains tend to orient along the membrane normal (Johansson and Lindahl 2006; MacCallum et al. 2008). As can be seen from Eq. 5, we assume that the lipid effect is depth-independent. This is reasonable, as the two restrictive effects from virtually fixed headgroups and more flexible tails sum almost to a constant effect at different membrane depths (Vermeer et al. 2007; Wiener and White 1992). Finally, the probability of each conformation can be derived by combining all restrictive effects (Štrancar et al. 2009):

$$P_i = P_i^{\text{initial}} \begin{cases} 0, \text{backbone overlap} \\ 1, \text{no backbone overlap} \end{cases} \prod_k^n \left(1 - \frac{N_k^{\text{overlaps}}}{N_k^{\text{all}}} \right) (1 - \sin \vartheta_i). \quad (6)$$

Because ESR spectroscopy is sensitive to the orientation of the spin label nitroxide group relative to the external magnetic field, the distribution of conformational orientations is reflected in the measured ESR spectrum. From all restricted rotamers of the modelled conformational space of the spin label, the distribution of nitroxide NO vectors can be characterized and compared with the related characteristics extracted from the ESR spectra. The cone model of spin label motion that is used in the analysis of experimental ESR spectra (Kavalenka et al. 2005; Stopar et al. 2006; Štrancar et al. 2005), is parameterized with the angles ϑ_0 and φ_0 , both defined in the range $(0; \pi/2)$. These angles describe the amplitude and anisotropy of the spin label rotational motion within a cone, respectively. For example, a large ϑ_0 indicates a larger (more open) cone (fewer restrictions of the conformational space from the top), while a large φ_0 indicates less anisotropy (more symmetry), i.e., fewer restrictions of the conformational space from the sides. Both angles ϑ_0 and φ_0 are combined into a simulated normalized free rotational space Ω_{sim} (Stopar et al. 2006), which is defined similar to Eq. 1 (Kavalenka et al. 2009a).

The best way to check the accuracy of the modelling of the conformational space is to compare restrictions calculated from known structures at particular sites of a protein with the restrictions detected by SDSL-ESR, or with those derived from molecular dynamics simulations. This was done while exploring the equinatoxin II system (Malovrh et al. 2003) (position 18) and human pancreatic lipase (Belle et al. 2007) (position 249). The structures of both proteins have been determined by X-ray crystallography (Athanasidis et al. 2001; Winkler et al. 1990). For equinatoxin II, NMR structures also were available (Hinds et al. 2002). These protein structures were then used to

calculate the restrictions of the conformational space and to compare them with those detected by SDSL-ESR. As can be seen in Fig. 3, the measured and simulated free rotational spaces agree well. The only exception is the result from molecular dynamics simulations, which shows a too restricted conformational space. This discrepancy can be explained by the relatively short simulation time of the trajectory (20 ns). As the length of the trajectory is only a few times longer than the ESR time-window, only a few conformations of the full conformational space will be sampled, resulting in a small (restrained) conformational space. Based on these analyses, we conclude that relatively simple modelling of the conformational restrictions of the amino acid and spin label side chains can be used to simulate the SDSL-ESR data at local protein sites, without considering the details of the local dynamics of the side chains.

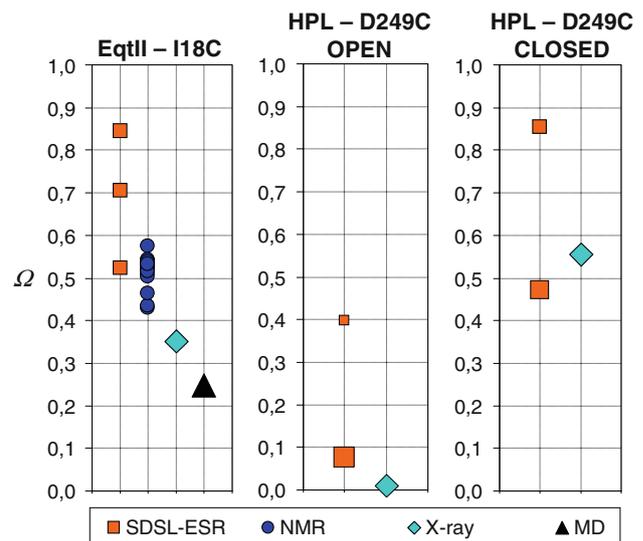


Fig. 3 Comparison of the measured and calculated restrictions of the conformational space (Ω). Calculated restrictions are derived using modelling of the conformational space from known protein structures, or by using molecular dynamics simulations. Measured restrictions are derived from analysis of the experimental SDSL-ESR spectra using GHOST condensation. Values for the free rotational space are presented for equinatoxin II (*Eqt II*) at mutant position 18 and human pancreatic lipase (*HPL*), mutant position 249 in the open and closed conformations. *Squares* indicate experimental SDSL-ESR results, *circles* result from modelling of the conformational space from the NMR data (Hinds et al. 2002), *diamonds* result from modelling of the conformational space from the X-ray data (Athanasidis et al. 2001; Winkler et al. 1990), *triangles* result from modelling of the conformational space from the molecular dynamics data. The molecular dynamics simulation was performed using the CHARMM 27 force field (Brooks et al. 1983), the spin label parameters were obtained from Fajer et al. (2007), and the restrictions were calculated from a 20-ns trajectory of EqtII in explicit water. The size of the symbols is proportional to the weight of the motional pattern detected with the GHOST condensation method

Structure optimization

The comparison of the simulated values for Ω_{sim} with the normalized free rotational space Ω_{exp} extracted from SDSL-ESR experimental data (Stopar et al. 2006; Štrancar et al. 2009) is used to govern an optimization algorithm, which tunes the secondary structure of the protein and the parameters of its relative orientation and position (Fig. 1b). The optimization module is based on a stochastic algorithm of the Metropolis Monte Carlo family (Kirkpatrick et al. 1983) with several elements of the evolutionary optimization (mutation, crossover, elite operators) (Eiben and Smith 2003; Fogel et al. 2000). However, unlike the conventional evolutionary algorithm, each optimization run optimizes just a single structure (and not a population of structures). A single run of optimization (Fig. 4), which counts for 200 generations, starts with initialization of a protein structure (setting the pairs of backbone dihedral angles $\{\varphi_i, \psi_i\}$ and the relative orientation and position of the protein in the system) (Table 1; Fig. 5), and initialization of the optimization parameters and constants (for example selection, mutation, crossover, elite, and shaking).

In each generation, the optimization parameters and the dihedral angles of the protein backbone are updated first (these parameters usually change with run generation number N_{gen}). Then the current structure of the protein system is modified by *internal operators* (modification of

the secondary structure of a protein via a mutation and crossover operators applied to the backbone dihedral angles) and *external operators* (modification of the position and orientation of the protein towards the membrane, or towards the other protein). The external operators also include a rotation of the protein around its long axis, given by the angle φ (relevant for helical chains; Fig. 5). Table 1 gives the parameters of the protein–lipid model, or protein complex model that can be optimized. At several stages in the optimization procedure, the protein structure in the system is checked for steric clashes. In cases of internal steric overlap, the algorithm returns and makes another try with the current operator. A maximum number of clashed structures is allowed in one generation. If this number is achieved the current problematic structure is replaced with the initial structure to protect the algorithm from going into a dead end. After the new structure is generated, the local restrictions at the mutant positions are calculated, and the obtained restriction profile Ω_{sim} is compared with the experimental restriction profile Ω_{exp} by a goodness of fit as follows:

$$\chi^2 = \frac{1}{N} \sum_i \left(\frac{\Omega_{exp,i} - \Omega_{sim,i}}{\Delta\Omega_{exp,i}} \right)^2, \tag{7}$$

where N is the number of spin-labelled mutants, $\Omega_{exp,i}$ and $\Omega_{sim,i}$ correspond to experimentally derived and simulated free rotational space values at the i -th mutant position, and

Fig. 4 Scheme of a single run of the algorithm for protein structure optimization. The algorithm is split into several functional parts: internal structure optimization operators (red box), external structure optimization operators (yellow box), restrictions calculation and data fitting (orange box), decision-making part (green box). The algorithm includes modelling of the lipid effect, as is needed for membrane proteins (Kavalenka et al. 2009a). The run generation number is N_{gen} . N_{max} is the maximum number of generations, typically 100

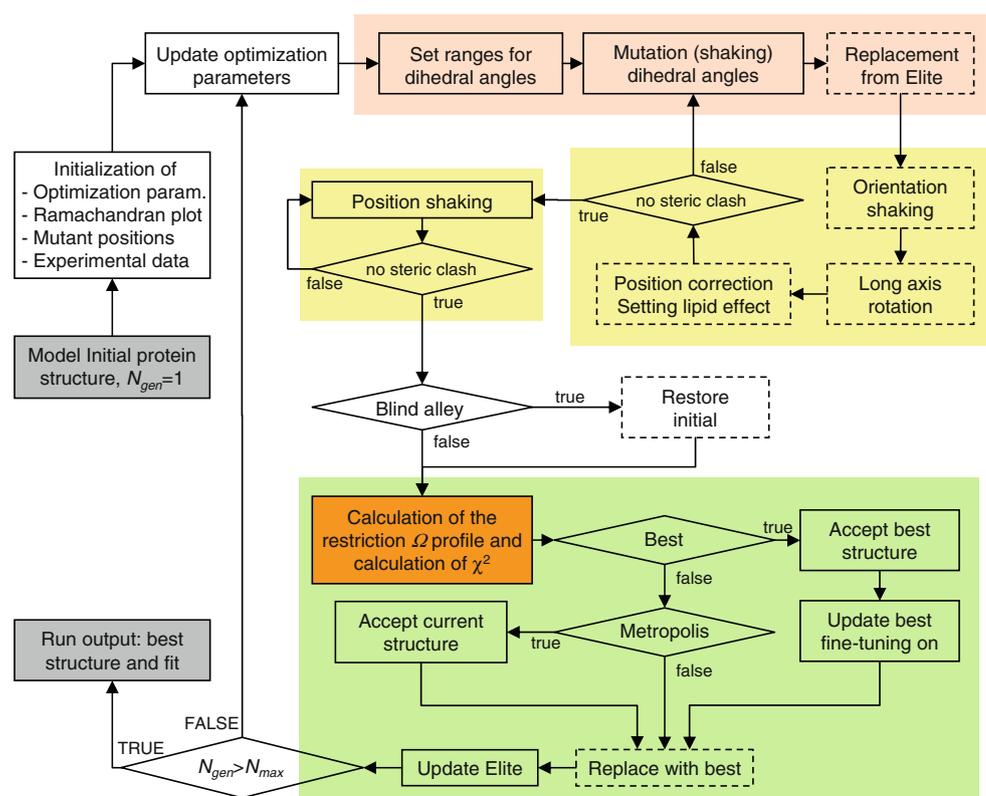


Table 1 Internal and external optimization parameters for different protein systems

Parameter	Unit	Description
Internal parameters		
$\{\varphi_i, \psi_i\}$	°	Pairs of dihedral angles (the first and last angles, φ_1 and ψ_N , are not defined)
φ	°	Rotational angle (rotation of the protein around the long axis)
External parameters—membrane proteins		
tm_{start}		Starting position of the transmembrane region of the protein
tm_{end}		End position of the transmembrane region of the protein
n_{ref}		Reference residue usually in the centre of the protein
D	Å	Steric thickness of the membrane
θ	°	Tilt angle of the protein with respect to the membrane normal
d_{shift}	Å	Shift of the protein in the bilayer along the membrane normal (used for fine-tuning of the transmembrane position of the protein)
External parameters—protein complex		
Δx	Å	Displacement vector of the protein relative to the partner
A	°	Orientation tensor of the protein relative to the partner

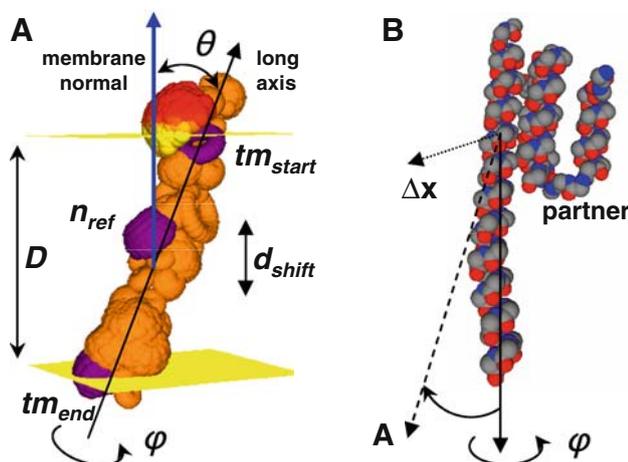


Fig. 5 Parameters for protein structure optimization. **a** Relative position and orientation of membrane-embedded M13 coat protein (Kavalenka et al. 2009a). The protein is shown with the conformational spaces of the amino acid side chains and spin label. The starting, tm_{start} and ending, tm_{end} residues of the transmembrane part of the protein and a reference n_{ref} residue in the centre of the transmembrane domain are highlighted. The yellow planes indicate the restrictive region of the lipid bilayer. **b** NTAIL protein, presented by backbone atoms relative to the partner protein XD (Kavalenka et al. 2009b). Both protein systems are parameterized according to Table 1

$\Delta\Omega_{exp,i}$ represents the second moment of the experimental free rotational space.

The minimum value of χ^2 in a single run is defined as χ_{best}^2 and corresponds to a protein structure which produces restrictions for the conformational space with the best fit to experimental data. If the goodness of fit χ^2 of a current structure is better than χ_{best}^2 , then χ^2 replaces χ_{best}^2 and the current structure becomes the parent for the structure in the next generation and the fine-tuning mutation mode is turned on. Even if χ^2 of the new structure does not outperform χ_{best}^2 , it may still become the parent for the new generation, if the Metropolis criterion (Metropolis et al. 1953) is satisfied. If the new structure is rejected, the parent stays the same as in the previous generation. However, with some probability (depending on the number of consequent unsuccessful generations) the current structure is replaced with the one that corresponds to χ_{best}^2 to let the algorithm produce good-fit structures still within the current generation. Finally, each structure with calculated restrictions is split into short subsequences with the corresponding goodness of fit (fitting the corresponding part of the experimental data). Any successful subsequence updates the elite database, which is used later for elitist crossover. The algorithm repeats this main loop until the maximum number of generations is reached. The outcome of the optimization is always a family of best-fit protein structures, which agree with the experimental SDSL-ESR data. Note that in the case of a membrane-embedded protein the protein structures are found relative to the membrane (Kavalenka et al. 2009a), whereas in the case of a protein–protein complex the family of favourable structures is determined relative to the partner protein (Kavalenka et al. 2009b).

For a protein–protein complex, spin labels can be placed on both proteins and, consequently, both structures can be optimized, including their relative orientation. For larger proteins, or membrane proteins that span the membrane several times, the protein backbone can be split into a set of protein domains, separated by loosely structured loops. The assignment of loosely structured loops and terminal ends could be carried out via the normalized rotational diffusion, being much higher in the unstructured parts compared with the structured parts, for example transmembrane domains (Kavalenka et al. 2009a; Stopar et al. 2005, 2006).

Computational demands

Our method requires solving of an inverse problem, both in the GHOST analysis and in the protein structure optimization. Generally, this means that these procedures are very time-consuming. Therefore, in our work much effort was devoted to speeding up the numeric calculations. Because the computational demand of the modelling of the

conformational space strongly depends on the number of dihedral rotations, we optimized the generation of the conformational space by discretizing the experimentally derived probability histograms of the side chain rotamer angles (Dunbrack 2002; Dunbrack and Cohen 1997; Dunbrack and Karplus 1993; Lovell et al. 2000; Ramya Bhargavi et al. 2003; Vasquez 1996) of each single bond rotation for each type of amino acid side chain. The optimization efficiency was further increased, in particular by introduction of a special operator that keeps track of successful structural segments (that successfully fit the corresponding segment of the SDSL-ESR-based restriction profile) and by introduction of an operator for local structural tuning. Currently, characterization of a 50-amino acid membrane protein with 27 spin label positions (Kavalenka et al. 2009a) takes about five weeks of CPU time on a small 20-core (2.3 GHz) computer cluster. One week was needed to analyse SDSL-ESR data and to extract the corresponding motional GHOST patterns. The GHOST analysis was then used in protein structure optimization, which took another four weeks of CPU time to obtain 1,000 best-fit structures by evolving through approximately 200,000 protein structures. The computational demand is roughly linearly dependent on the protein size. Similarly, the methodology also allows a linear decrease of computational time by increasing the computational power. It should be noted that protein structure optimization is practically inaccessible if molecular dynamics simulations are applied to derive the restrictions of the conformational space, even with a much larger computer cluster size. This impracticability arises because the molecular dynamics simulations should reach an ESR averaging time of a few nanoseconds for each of the structures scanned.

Complementarity to other methods

Our method is comparable with the distance geometry approach employed in two-dimensional solution NMR spectroscopy that also results in a family of structures (Bax 1989; Castellani et al. 2002). However, the number of restraints used in the modelling in our method (free rotational space and rotational diffusion) depends on the number of available spin-labelled protein mutants. Such a data set is smaller than the data set available from NMR spectroscopy (nuclear Overhauser effect, one-bond and three-bond J-coupling, carbon and proton chemical shift, and rotational diffusion anisotropy) (Brunger et al. 1998). The number of restraints could be increased by producing more spin-labelled protein mutants, or by using additional structural restraints from other low-resolution methods (e.g., distance measurements by fluorescence spectroscopy or dual-spin label SDSL-ESR, global conformation constraints by SAXS and CD).

Even though ESR spectroscopy is insensitive to the exact atomic coordinates, sensitivity to the anisotropy of the local conformational space of the spin label and the corresponding modelling of the conformational space allow the determination of the backbone fold with almost atomic resolution. However, there is no structural information about individual side chain conformations. Instead the protein can be represented as a space needed for all the side chains to wobble. The ability to track the protein structure and dynamics in a native environment and at physiological temperature is one of the great advantages of the proposed method. Another advantage is the higher sensitivity of ESR than NMR, which means that much lower concentrations of protein samples are needed to perform the experiment (Hemminga 2007).

Acknowledgments We wish to thank David Stopar and Primož Zihlerl (University of Ljubljana) for many valuable discussions at the very beginning of development of the methodology, Valerie Belle and Bruno Guigliarelli (BIP CNRS Marseille) and Sonia Longhi (AFMB CNRS Marseille) for helping us to develop the algorithm for cleaning the motional pattern profiles on the NTAIL problem. We are also grateful to Jan Premru for important contribution in analysis of the rotamer libraries. This work was carried out with the financial support of the Slovenian Research Agency (program “Experimental biophysics of complex systems”, P1-0060) and COST P15 action support (through a Short Term Scientific Meeting grant).

References

- Alexander N, Al-Mestarihi A, Bortolus M, McHaourab H, Meiler J (2008) De novo high-resolution protein structure determination from sparse spin-labeling EPR data. *Structure* 16:181–195
- Arora A, Tamm LK (2001) Biophysical approaches to membrane protein structure determination. *Curr Opin Struct Biol* 11:540–547
- Athanasiadis A, Anderluh G, Macek P, Turk D (2001) Crystal structure of the soluble form of equinatoxin II, a pore-forming toxin from the sea anemone *Actinia equina*. *Structure* 9:341–346
- Bax A (1989) Two-dimensional NMR and protein structure. *Annu Rev Biochem* 58:223–256
- Belle V, Fournel A, Woudstra M, Ranaldi S, Prieri F, Thome V, Currault J, Verger R, Guigliarelli B, Carriere F (2007) Probing the opening of the pancreatic lipase lid using site-directed spin labeling and EPR spectroscopy. *Biochem* 46:2205–2214
- Belle V, Rouger S, Costanzo S, Liqueire E, Štrancar J, Guigliarelli B, Fournel A, Longhi S (2008) Mapping alpha-helical induced folding within the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by site-directed spin-labeling EPR spectroscopy. *Proteins* 73:973–988
- Bourhis J-M, Canard B, Longhi S (2007) Predicting protein disorder and induced folding: from theoretical principles to practical applications. *Curr Protein Pept Sci* 8:135–149
- Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998)

- Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54:905–921
- Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420:98–102
- Dominguez C, Boelens R, Bonvin AMJJ (2003) HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 125:1731–1737
- Dunbrack JRL (2002) Rotamer libraries in the 21st century. *Curr Opin Struct Biol* 12:431–440
- Dunbrack JRL, Cohen FE (1997) Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci* 6:1661–1681
- Dunbrack JRL, Karplus M (1993) Backbone-dependent rotamer library for proteins application to side-chain prediction. *J Mol Biol* 230:543–574
- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. *J Mol Graph Model* 19:26–59
- Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets: the roles of intrinsic disorder in protein interaction networks. *FEBS J* 272:5129–5148
- Dunker AK, Oldfield C, Meng J, Romero P, Yang J, Chen J, Vacic V, Obradovic Z, Uversky V (2008) The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics* 9:S1
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208
- Eiben AE, Smith JE (2003) Introduction to evolutionary computing. Springer, New York
- Engh RA, Huber R (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr Sect A* 47:392–400
- Fajer M, Fajer PG, Sale KL (2007) Molecular modeling of spin labels. In: Hemminga MA, Berliner L (eds) ESR spectroscopy in membrane biophysics. Springer, New York, pp 253–259
- Fanucci GE, Cafiso DS (2006) Recent advances and applications of site-directed spin labeling. *Curr Opin Struct Biol* 16:644–653
- Fasman GD (1996) Circular dichroism and the conformational analysis of biomolecules. Plenum Press, New York
- Ferron F, Longhi S, Canard B, Karlin D (2006) A practical overview of protein disorder prediction methods. *Proteins* 65:1–14
- Filipič B, Štrancar J (2001) Tuning EPR spectral parameters with a genetic algorithm. *Appl Soft Comput* 1:83–90
- Fink AL (2005) Natively unfolded proteins. *Curr Opin Struct Biol* 15:35–41
- Fleishman SJ, Unger VM, Ben-Tal N (2006) Transmembrane protein structures without X-rays. *Trends Biochem Sci* 31:106–113
- Fogel DB, Bäck T, Michalewicz Z (2000) Evolutionary computation. Philadelphia, Institute of Physics Publishing, Bristol
- Grigoryan G, Ochoa A, Keating AE (2007) Computing van der Waals energies in the context of the rotamer approximation. *Proteins* 68:863–878
- Hemminga MA (2007) Introduction and future of site-directed spin labeling of membrane proteins. In: Hemminga MA, Berliner L (eds) ESR Spectroscopy in Membrane Biophysics. pp, Springer, pp 1–16
- Henderson R (2004) Realizing the potential of electron cryo-microscopy. *Q Rev Biophys* 37:3–13
- Hinds MG, Zhang W, Anderlüh G, Hansen PE, Norton RS (2002) Solution structure of the eukaryotic pore-forming cytolysin equinatoxin II: implications for pore formation. *J Mol Biol* 315:1219–1229
- Ho BK, Thomas A, Brasseur R (2003) Revisiting the Ramachandran plot: hard-sphere repulsion, electrostatics, and H-bonding in the α -helix. *Protein Sci* 12:2508–2522
- Huang H, Cafiso DS (2008) Conformation and membrane position of the region linking the two C2 domains in synaptotagmin 1 by site-directed spin labeling. *Biochem* 47:12380–12388
- Hubbell WL, Gross A, Langen R, Lietzow MA (1998) Recent advances in site-directed spin labeling of proteins. *Curr Opin Struct Biol* 8:649–656
- Jao CC, Hegde BG, Chen J, Haworth IS, Langen R (2008) Structure of membrane-bound α -synuclein from site-directed spin labeling and computational refinement. *Proc Natl Acad Sci USA* 105:19666–19671
- Johansson AC, Lindahl E (2006) Amino-acid solvation structure in transmembrane helices from molecular dynamics simulations. *Biophys J* 91:4450–4463
- Kavalenka AA, Filipič B, Hemminga MA, Štrancar J (2005) Speeding up a genetic algorithm for EPR-based spin label characterization of biosystem complexity. *J Chem Inf Mod* 45:1628–1635
- Kavalenka A, Hemminga MA, Štrancar J (2009a) Optimization of membrane protein structure based on SDSL-ESR constraints and conformational space modeling. *Biophys J* (submitted)
- Kavalenka A, Urbančić I, Belle V, Rouger S, Costanzo S, Kure S, Fournel A, Longhi S, Guigliarelli B, Štrancar J (2009b) Conformational analysis of the partially disordered measles virus NTAIL–XD complex explored by SDSL EPR spectroscopy. *Biophys J* (submitted)
- Kelly SM, Price NC (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr Protein Pept Sci* 1:349–384
- Kirkpatrick S, Gelatt CD Jr, Vecchi MP (1983) Optimization by simulated annealing. *Science* 220:671–680
- Lacapère JJ, Pebay-Peyroula E, Neumann JM, Etchebest C (2007) Determining membrane protein structures: still a challenge! *Trends Biochem Sci* 32:259–270. Epub 2007 May 3. Review.
- Lehninger AL, Nelson DL, Cox MM (2005) Lehninger principles of biochemistry, 4th edn edn. W.H. Freeman, New York
- Li Q, Fung LW-M (2009) Structural and dynamic study of the tetramerization region of non-erythroid α -spectrin: a frayed helix revealed by site-directed spin labeling electron paramagnetic resonance. *Biochem* 48:206–215
- Lovell SC, Word JM, Richardson JS, Richardson DC (2000) The penultimate rotamer library. *Proteins* 40:389–408
- MacCallum JL, Bennett WFD, Tieleman DP (2008) Distribution of amino acids in a lipid bilayer from computer simulations. *Biophys J* 94:3393–3404
- MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
- Malovrh P, Viero G, Serra MD, Podlesek Z, Lakey JH, Macek P, Menestrina G, Anderlüh G (2003) A novel mechanism of pore formation: membrane penetration by the N-terminal amphipathic region of equinatoxin. *J Biol Chem* 278:22678–22685
- Marsh D (1981) Electron Spin Resonance: Spin Labels. In: Grell E (ed) Membrane Spectroscopy. Springer-Verlag, Berlin New York, pp 51–142
- Marsh D (2008) Protein modulation of lipids, and vice versa, in membranes. *Biochim Biophys Acta* 1778:1545–1575
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092

- Muller DJ, Engel A (2008) Strategies to prepare and characterize native membrane proteins and protein membranes by AFM. *Curr Opin Colloid Interface Sci* 13:338–350
- Nordio PL (1976) General magnetic resonance theory. In: Berliner LJ (ed) *Spin labeling: theory and applications*. Academic Press, New York, pp 5–51
- Pebay-Peyroula E (2008) Biophysical analysis of membrane proteins: investigating structure and function. Wiley-VCH, Weinheim
- Petoukhov MV, Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J* 89:1237–1250
- Pistolesi S, Ferro E, Santucci A, Basosi R, Trabalzini L, Pogni R (2006) Molecular motion of spin labeled side chains in the C-terminal domain of RGL2 protein: A SDSL-EPR and MD study. *Biophys Chem* 123:49–57
- Ramya Bhargavi G, Sheik SS, Velmurugan D, Sekar K (2003) Side-chain conformation angles of amino acids: effect of temperature factor cut-off. *J Struct Biol* 143:181–184
- Receveur-Bréchet V, Bourhis JM, Uversky VN, Canard B, Longhi S (2006) Assessing protein disorder and induced folding. *Proteins* 62:24–45
- Schindler H, Seelig J (1973) EPR spectra of spin labels in lipid bilayers. *J Chem Phys* 59:1841–1850
- Shetty RP, De Bakker PI, DePristo MA, Blundell TL (2003) Advantages of fine-grained side chain conformer libraries. *Protein Eng* 16:963–969
- Steinhoff H-J, Savitsky A, Wegener C, Pfeiffer M, Plato M, Mutius K (2000) High-field EPR studies of the structure and conformational changes of site-directed spin labeled bacteriorhodopsin. *Biochim Biophys Acta* 1457:253–262
- Stopar D, Štrancar J, Spruijt RB, Hemminga MA (2005) Exploring the local conformational space of a membrane protein by site-directed spin labeling. *J Chem Inf Mod* 45:1621–1627
- Stopar D, Štrancar J, Spruijt RB, Hemminga MA (2006) Motional restrictions of membrane proteins: a site-directed spin labeling study. *Biophys J* 91:3341–3348
- Štrancar J (2007) Advanced ESR spectroscopy in membrane biophysics. In: Hemminga MA, Berliner L (eds) *ESR spectroscopy in membrane biophysics*. Springer, New York, pp 49–93
- Štrancar J, Sentjurc M, Schara M (2000) Fast and accurate characterization of biological membranes by EPR spectral simulations of nitroxides. *J Magn Reson* 142:254–265
- Štrancar J, Koklic T, Arsov Z, Filipic B, Stopar D, Hemminga MA (2005) Spin label EPR-based characterization of biosystem complexity. *J Chem Inf Mod* 45:394–406
- Štrancar J, Kavalenka A, Zihel P, Stopar D, Hemminga MA (2009) Analysis of side chain rotational restrictions of membrane-embedded proteins by spin-label ESR spectroscopy. *J Magn Reson* 197:245–248
- Svergun DI, Koch MHJ (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep Prog Phys* 66:1735–1782
- Timsit Y, Allemand F, Chiaruttini C, Springer M (2006) Coexistence of two protein folding states in the crystal structure of ribosomal protein L20. *EMBO Rep* 7:1013–1018
- Tombolato F, Ferrarini A, Freed JH (2006) Dynamics of the nitroxide side chain in spin-labeled proteins. *J Phys Chem B* 110:26248–26259
- Tomba P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* 27:527–533
- Torres J, Stevens TJ, Samsó M (2003) Membrane proteins: the ‘Wild West’ of structural biology. *Trends Biochem Sci* 28:137–144. Review. Erratum in: *Trends Biochem Sci* 2003 28:174
- Uversky VN (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* 11:739–756
- Vasquez M (1996) Modeling side-chain conformation. *Curr Opin Struct Biol* 6:217–221
- Vermeer LS, de Groot BL, Reat V, Milon A, Czaplicki J (2007) Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with ²H NMR experiments. *Eur Biophys J* 36:919–931
- Watts A, Straus SK, Grage SL, Kamihira M, Lam YH, Zhao X (2004) Membrane protein structure determination using solid-state NMR. *Methods Mol Biol* 278:403–473
- White S (2009) Membrane proteins of known 3D structure. http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html
- Wiener MC, White SH (1992) Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of X-ray and neutron diffraction data III. Complete structure. *Biophys J* 61:434–447
- Winkler FK, D’Arcy A, Hunziker W (1990) Structure of human pancreatic lipase. *Nature* 343:771–774
- Word JM, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC (1999) Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *J Mol Biol* 285:1711–1733
- Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* 19:31–38
- Wüthrich K (1986) *NMR of proteins and nucleic acids*. Wiley, New York
- Xiang Z, Honig B (2001) Extending the accuracy limits of prediction for side-chain conformations. *J Mol Biol* 311:421–430