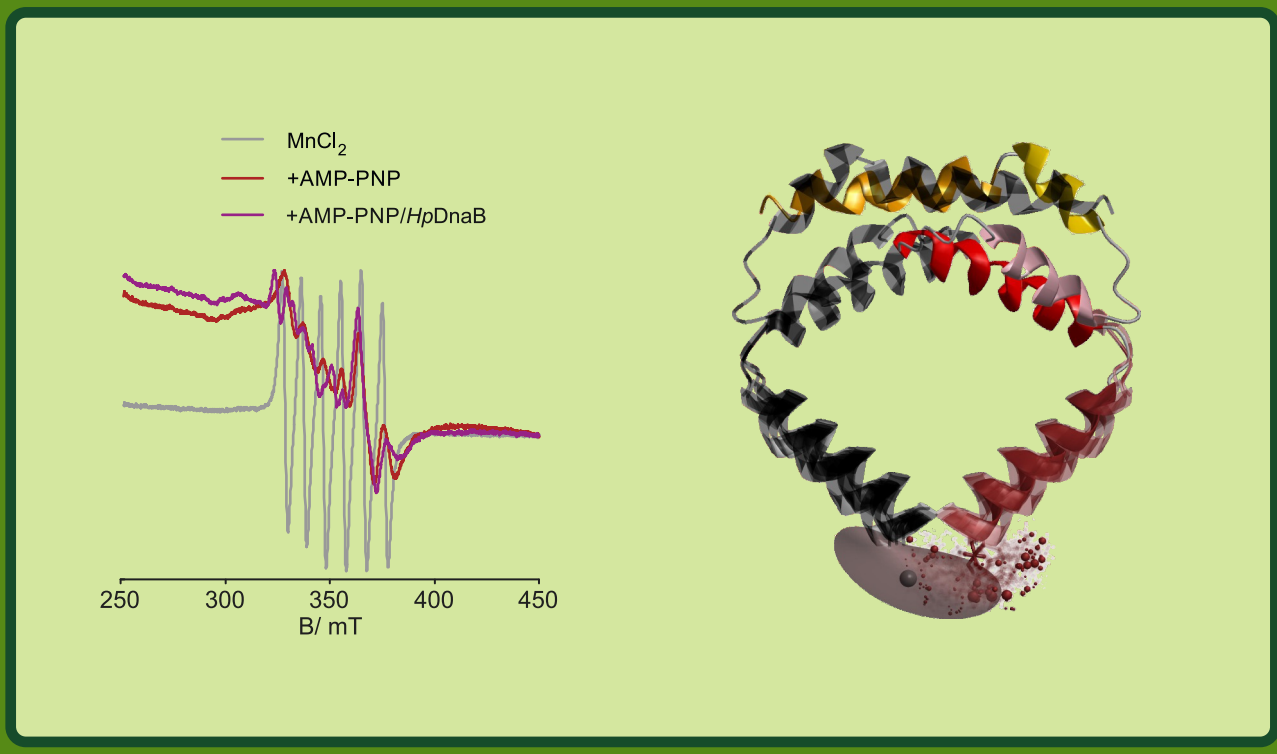
Lecture Notes

Molecular and Structural Biology V

Studying Macromolecules by EPR

Gunnar Jeschke



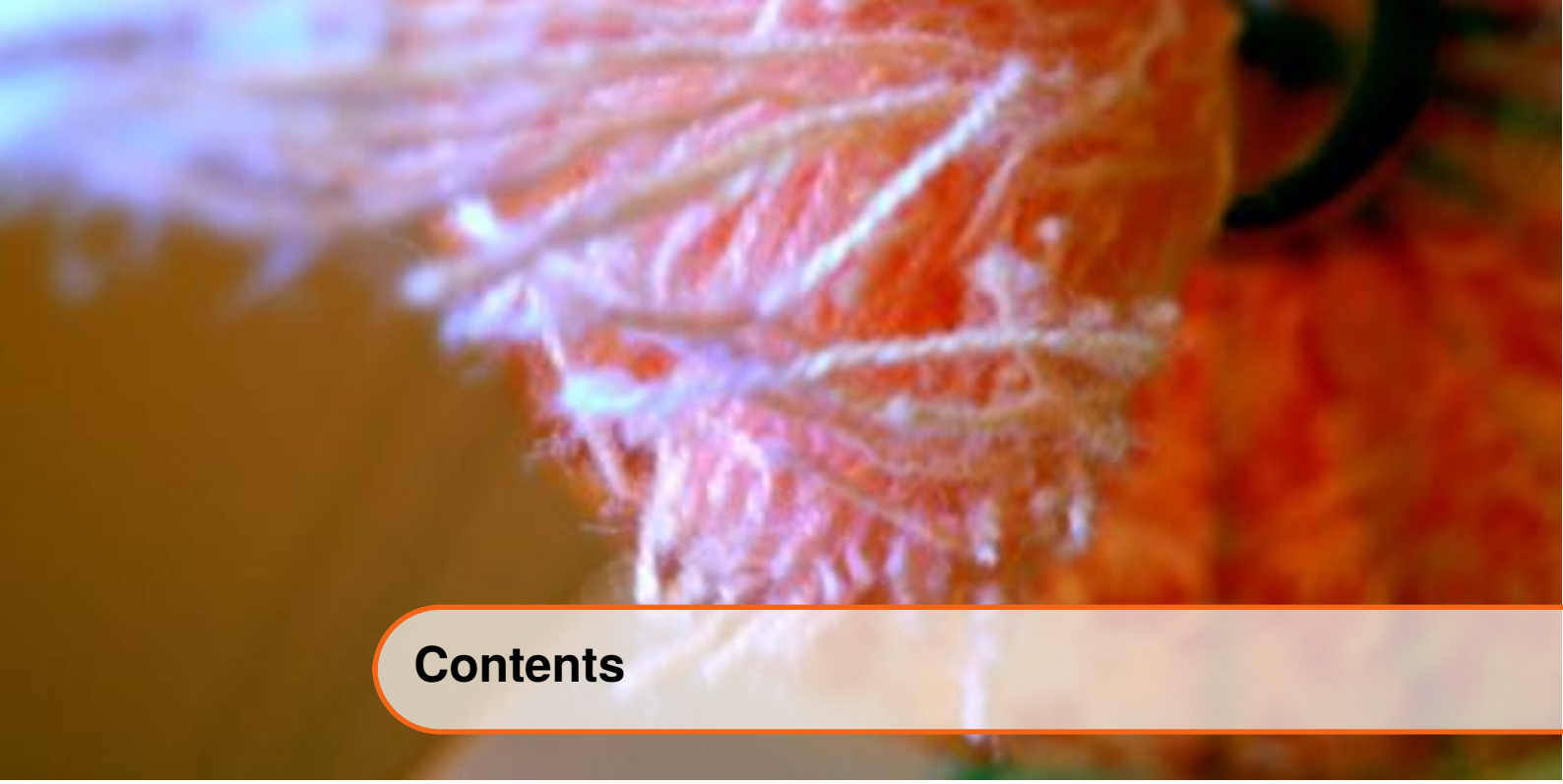
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Title image: Left - Binding of Mn(II) to a nucleotide binding domain [Wie+17b]
Right - Localization of the end of α-helix 6 in the active form of the pro-apoptotic protein Bax [Ble+14b]
Chapter 2 Singly occupied molecular orbital of a tyrosyl radical (own work)
Chapter 3 Resonator for EPR and Overhauser dynamic nuclear polarization of aqueous samples (own work)
Chapter 4 [4Fe-4S] iron-sulfur cluster in *Ectothiorhodospira halophila* HiPIP (PDB #1HIP, visualized by MMM)
Chapter 5 Visualization of MTS spin labels attached to residues 72 and 131 of T4 Lysozyme (own work)
Chapter 6 Ensemble model generated from simulated restraints for p27KID (own work)
Chapter 7 Screen shot of MMM during solving one of the tutorial problems (own work)

Published by Gunnar Jeschke

http://www.epr.ethz.ch

First printing, May 2017



Contents

1 — Preface

1.1 General Remarks

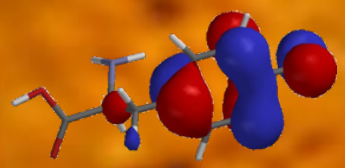
Electron Paramagnetic Resonance (EPR) spectroscopy relies on the presence of unpaired electrons, which are rare since chemical bonding is electron pairing. On the one hand, this restricts direct application of EPR spectroscopy to the small subset of biological macromolecules that feature intrinsic paramagnetic centers or can be put in paramagnetic states. On the other hand, it opens up the possibility to use EPR spectroscopy as a probe technique by combining it with site-directed spin labeling (SDSL). Since electron spins couple only weakly to their environment, EPR can be used as a probe technique under a wide range of conditions and in a wide range of biological systems. In this lecture course we consider both the characterization of intrinsic paramagnetic centers in biological systems and EPR spectroscopy as an information source for hybrid structure determination when applied in conjunction with SDSL. The focus of this course is not on the inner working of EPR experiments, but rather on the proper choice of an experiment for a given problem, on sample preparation, and on analysis of the experimental data including their use in modeling of structures.

Chapter 2 provides a brief overview of the interactions of an electron spin with its environment and of the information on electronic and spatial structurethat can be obtained from them. This Chapter also lists the major experimental techniques that can be used to obtain this information and discusses the choice of EPR frequency. Chapter 3 discusses requirements on sample preparation, in particular concentration and volume, purity with respect to the presence of unwanted paramagnetic centers, optimization of electron spin relaxation time, and the necessity and choice of cryoprotectants. Chapter 4 is devoted to a brief overview of intrinsic paramagnetic centers in biological macromolecules. Chapter 5 discusses SDSL strategies, in particular the choice of labels and of the molecule or site to be labeled and the information that can be obtained in an SDSL study. Finally, Chapter 6 is devoted to the use of EPR information in structure modeling. This comprises the conversion of data from dipolar spectroscopy to distance distribution restraints, the modeling of spin label conformation, and the combination of EPR restraints with restraints from other experimental techniques in hybrid structure determination.

1.2 Suggested Reading & Electronic Resources

There is no introductory level textbook that treats the material of this course. Many of the specific topics are treated on a deeper level in chapters that have been published over the past few decades in the series *Biological Magnetic Resonance*. Where appropriate, references to such chapters or to reviews are given the topical sections. Volume 27 [HE07] is special in that it covers several of the concepts that are treated in this lecture course. A basics introduction into EPR spectroscopy that complements this lecture script has been published in the Oxford Chemistry Primer series by Chechik, Carter, and Murphy [CCM16]. Lectures scripts on EPR spectroscopy in general (German language) can be found on my homepage (`http://www.epr.ethz.ch/education.html`). The EPR script for the Physical Chemistry IV course at ETH (Magnetic Resonance) (`https://www.ethz.ch/content/dam/ethz/special-interest/chab/physical-chemistry/epr-dam/documents/education/EPR_PCIV.pdf`) explains the inner working of the experiments and spin dynamics in detail (English language).

Data analysis and structure modeling require specialized software packages. The most versatile package for spectrum simulations and fitting is EasySpin by Stefan Stoll (`http://www.easyspin.org/`). All worked examples for spectrum simulations in this lecture course are provided as EasySpin scripts that can be run in Matlab®. For intrinsic paramagnetic centers, quantum-chemical computations of spin Hamiltonian parameters may be required and the probably most versatile program for this is the freely available package ORCA (`https://orcaforum.cec.mpg.de/`). The most informative EPR techniques for determining spatial structure of macromolecules are pulsed dipolar spectroscopy techniques such as double electron electron resonance (DEER, also known as pulsed electron electron double resonance, PELDOR). The standard software for analysis of such data is DeerAnalysis (`http://www.epr.ethz.ch/software.html`) that is explained in some more detail in Chapter 6. For developing SDSL strategies and for structure modeling based on SDSL-derived restraints we use the software package MMM (`http://www.epr.ethz.ch/software.html`).



2 — Electron spin as an information source

2.1 Interactions of electron spins

2.1.1 Overview of interactions and associated information

From a physical point of view, spin is angular momentum. It couples to other angular momenta and to magnetic fields. If the spin quantum number is larger than 1/2, it also couples to electric field gradients, but this interaction is not relevant for electron spins. The other angular momenta in molecules to which electron spin may couple are other electron spins (dipole-dipole interaction and exchange coupling), nuclear spins (hyperfine coupling), and the orbital angular momentum of the electrons itself (spin-orbit coupling). Orbital angular momentum is quenched in non-degenerate ground states and therefore, spin-orbit coupling usually arises only through excited states. It is a relativistic effect that increases with the 4^{th} power of the atom mass. Thus, spin-orbit coupling is weak for first- and second-row elements (H, C, N, O), appreciable for transition metal ions, and strong for rare earth ions that do not intrinsically occur in biological systems, but may sometimes be used as spin probes.

Hyperfine coupling

For intrinsic paramagnetic centers, *hyperfine coupling* is the richest source of information since nuclear spins are so abundant. Hyperfine coupling has an isotropic contribution that stems solely from spin density in s orbitals of the nucleus under consideration. This *Fermi contact interaction* informs purely on electronic structure, namely the s spin density on this nucleus, and provides no information on the spatial position of the nucleus. If a sufficient number of isotropic hyperfine couplings can be assigned, one can guess the type of orbital where the unpaired electron resides, the *singly occupied molecular orbital* (SOMO) ,. Hyperfine coupling also has an anisotropic contribution that stems from the interaction between the magnetic dipole associated with electron spin and the magnetic dipole associated with the nuclear spin (dipole-dipole interaction). If spin density on this nucleus exists solely in s orbitals, as is the case for hydrogen, alkali metal ions, and earth alkali metal ions, the dipole-dipole hyperfine interaction is a pure through-space interaction. If furthermore the electron spin is localized on a single atom, as is often the case for paramagnetic metal ions, or if its spin density distribution is known, the anisotropic part of the hyperfine interaction can be used to localize hydrogen atoms, which are not usually seen in crystal structures of macromolecules. Such localization is based on the known r^{-3} distance dependence and $3\cos^2\theta - 1$ orientation dependence of the dipole-dipole coupling, where r is

the spin-spin distance and θ the angle between the spin-spin vector and the external magnetic field. If electron spin can also reside in p, d, or f orbitals on the nucleus under consideration, interpretation of the anisotropic contribution to the hyperfine coupling is more complicated and usually requires a quantum-chemical calculation.

g factor and g tensor

Like NMR spectroscopy, EPR spectroscopy is performed in an external magnetic field $\vec{B}_0$. The interaction of $\vec{B}_0$ with electron spin is the electron Zeeman interaction and can be exactly predicted from first physical principles for a free electron, where it is isotropic and characterized by the free electron g factor $g_e = 2.0023193$. For a bound electron, electron Zeeman interaction is modified by spin-orbit coupling and generally becomes anisotropic and must then described by a g tensor. Often, biological systems are studied by EPR in glassy frozen solution and then only the three principal values g_x, g_y, and g_z of the g tensor can be determined. They are used as fingerprint information on the type of organic radical or transition metal ion coordination. The three principal values also provide information on the symmetry of the electron spin density distribution in the SOMO. If this distribution has a symmetry axis C_n with $n \geq 3$, then the g tensor must have axial symmetry, $g_x = g_y = g_\perp$, $g_z = g_{||}$, where g_z is usually taken along the symmetry axis. If the symmetry is tetrahedral or cubic, the the g tensor is isotropic with $g_x = g_y = g_z = g_{\text{iso}}$. The g tensor is also isotropic to a very good approximation if the d or f shell is half filled, a case that applies to Mn(II) and Gd(III) ions.

Coupling between electron spins

Unpaired electrons are rare and, in order to obtain utmost resolution, samples are usually diluted to concentrations where the interaction of electron spins of different molecules is negligible (see Chapter 3). Electron-electron interaction may still play a role if two unpaired electrons exist in the same molecule. First and foremost, this applies to high-spin transition metal and rare earth metal ions , with the most prominent cases being Fe(III) ($3d^5$ electron configuration, $S = 5/2$), Mn(II) ($3d^5$, $S = 5/2$), Co(II) ($3d^7$, $S = 3/2$), Cr(III) ($3d^3$, $S = 3/2$), and Gd(III) ($4f^7$, $S = 7/2$), where S is an electron group spin that arises from the very strong coupling between the unpaired electrons on the same atom. If the ligand field is sufficiently strong, some of these ions may exist in low-spin states with $S = 1/2$, notably Fe(III) and Co(II) . In these cases, the spin state, provides information on the strength of the ligand field. Metal ions with an odd number of unpaired electrons are called *Kramers ions* and are usually easy to measure by EPR.

If the number of unpaired electrons is even, group spin is an integer number. For such *non-Kramers* ions, the low-spin state has $S = 0$, is diamagnetic and not accessible to EPR spectroscopy. Furthermore, the high-spin states are often *"EPR silent"*, i.e., they do not feature transitions that can be detected at the frequencies and magnetic fields that are accessible with usual laboratory EPR spectrometers. Such systems can still be characterized at very high fields and frequencies, usually in dedicated high-field laboratories or with dedicated very high frequency home-built spectrometers. In many cases, however, sensitivity of these spectrometers is not sufficient for the amounts of proteins that can be easily obtained and for the concentrations at which these proteins are still stable. The most notable paramagnetic non-Kramers ions which are hard to characterize by EPR spectroscopy are Fe(II) ($3d^6$, $S = 2$), Co(III) ($3d^6$, $S = 2$), and Ni(II) ($3d^8$, $S = 1$). If point symmetry at a non-Kramers ion is nearly axial, it may be possible to detect EPR transitions at low fields and frequencies. Such spectra are usually difficult to interpret.

In organic molecules, two unpaired electrons coexist in excited triplet states, leading to electron group spin $S = 1$. Such triplet states may be encountered for cofactors of biological macromolecules that convert photons to chemical energy (photosynthetic reaction centers) or harvest photons (antenna systems for photosynthesis). Although triplet states are formally non-Kramers systems, they are usually observable at common EPR fields and frequencies.

This is because the electron spins are spatially more widely distributed, leading to on average smaller dipole-dipole coupling between them and because spin-orbit coupling is small for organic molecules. As a result, the *zero-field splitting* is of the order of only a few hundred Megahertz to a few Gigahertz and thus smaller than the electron Zeeman frequency at typical EPR fields.

In macromolecules, two unpaired electrons can also be localized at different sites. In this case, they can be pictured as not residing in the same molecular orbital, as is the case for the two unpaired electrons in a triplet state, but rather in two distinct SOMOs. The corresponding state of the system is a radical pair state with spins $S_1 = 1/2$ and $S_2 = 1/2$. Such radical pair states occur in biological systems that are devoted to electron transfer (redox chemistry), often related to bioenergetics. The two SOMOs may still have non-negligible overlap that leads to *exchange coupling* J between the electron spins. Generally, if the exchange coupling is much larger than the electron Zeeman interaction, the system with four energy levels is best described in terms of a singlet state ($S = 0$, one level) and a triplet state ($S = 1$, three sublevels). If the exchange coupling is much smaller than the electron Zeeman interaction, but larger than about 1 MHz, the system is best described as a radical pair. Even smaller exchange couplings are unresolved. For this case, a description in terms of isolated radicals usually works well.

Zero-field splitting

The sublevels of group spins $S > 1/2$ are split already at zero magnetic field. This zero-field splitting arises from the pairwise dipole-dipole coupling between the electron spins and from spin-orbit coupling. For organic triplets, the spin-orbit contribution is negligible and the dipole-dipole contribution can be computed by spatial integration over the orbital in which the two unpaired electrons reside [Rie07]. For high-spin metal ions, the spin-orbit contribution usually dominates, but the spin-spin contribution is not negligible.

Zero-field splitting is a purely anisotropic interaction, usually parametrized by the two quantities D and E with $|E| \leq |D|/3$. Axial symmetry implies $E = 0$ and is encountered if the system has a symmetry axis C_n with $n \geq 3$. If D is much larger than the electron Zeeman interaction, then the EPR spectrum does not longer depend on the size of D and the observed spectral features can be assigned effective g values. These effective g values still depend on the ratio $|E/D|$. If D is smaller than the electron Zeeman interaction, D and E can be obtained by lineshape analysis. The values may serve as fingerprint information on the electronic structure. Zero-field splitting can be predicted by quantum-chemical calculations, but is usually harder to get right than hyperfine coupling and g tensor.

For group-spin $S > 1$, the zero-field splitting has further contributions that are not parametrized by D or E. These contributions are resolved only at very high symmetry and are usually ignored in biological systems.

Exchange coupling

Exchange coupling can be ferromagnetic (triplet state lower in energy, anti-binding overlap of the two SOMOs) or anti-ferromagnetic (singlet state lower in energy, binding overlap of the two SOMOs). The latter case is more frequently encountered. Exchange couplings that can be measured by EPR spectroscopy are much smaller than the electron Zeeman interaction and cannot be predicted with useful accuracy by quantum-chemical calculations.

Dipole-dipole coupling

If several unpaired electrons are strongly coupled, their dipole-dipole couplings are subsumed in the zero-field splitting. For radical pairs with a distance r between the centers of spin density that is much larger than the spatial distribution of the individual spin densities, electron-electron dipole-dipole coupling can be analyzed in terms of the distance r. This is the basis for generating distance (distribution) restraints on spatial structure of macromolecules by EPR, usually in

conjunction with SDSL.

2.1.2 Hierarchy of interactions

Spectra are most easy to interpret in the *high-field approximation* where the electron Zeeman interaction is much larger than all other interactions. This regime is hard or impossible to reach for $S > 1/2$ systems with large zero-field splitting. Hyperfine couplings are usually much smaller than the electron Zeeman interaction, but often larger than or of the same order of magnitude as the nuclear Zeeman interaction. An overview of the approximate magnitude of the most important interactions is given in Fig. 2.1.

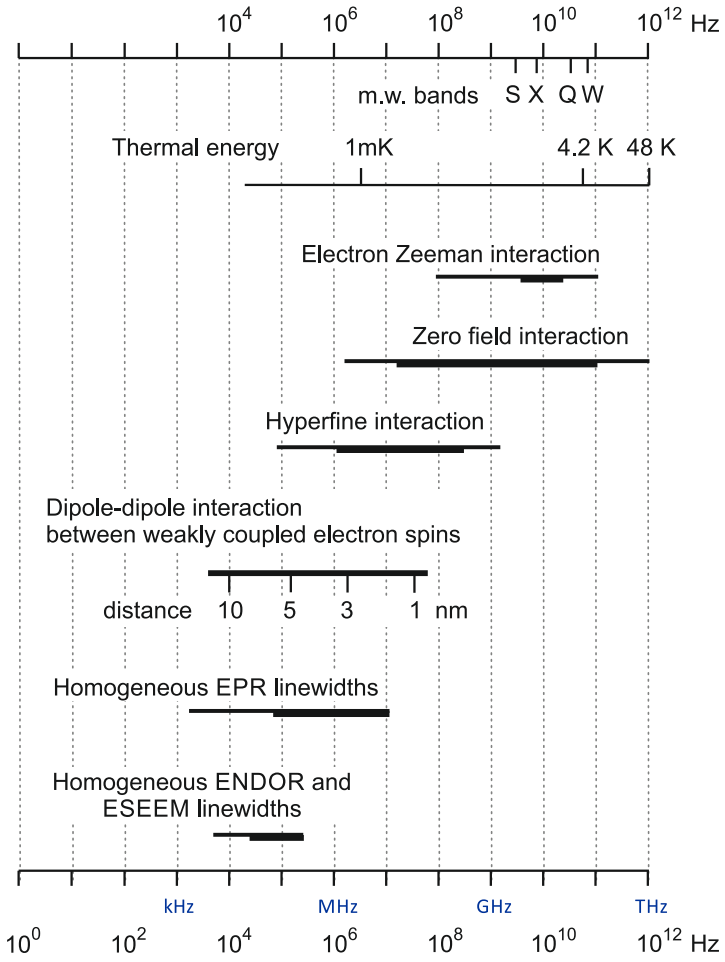


Figure 2.1: Magnitude of interactions of electron spins in frequency units and their relation to thermal energy and typical EPR frequency bands.

2.2 Experimental techniques

2.2.1 General considerations on the choice of experiment

The biological problem to be solved dictates what information should be obtained at what accuracy and with what uncertainty. In some cases, one may strive for utmost technically possible accuracy and minimal uncertainty, irrespective of the biological problem at hand. This is often the case in structure determination, where the underlying assumption is that the structure will

help to solve several biological problems, including ones that have not been posed at the moment when the structure is published.

The choice of experiments is restrained by locally available spectrometer hardware or measurement time that can be obtained at remote facilities or collaboration partners. It may be further restrained by the amount of sample and sample concentration that can be prepared, since different types of EPR experiments have different sensitivity. Especially in SDSL studies, the decision for another type of information may require not only additional measurement and data analysis effort, but also additional effort in sample preparation. In such cases it is advisable to base experiment planning on a modeling strategy. i.e. to first establish what the minimal set of restraints is that has to be collected in order to generate a model that can solve the problem at hand.

The following sections list experiments that can be used to characterize individual spin interactions, shortly discuss their advantages and disadvantages, and give an overview on the requirements on the spectrometer hardware and sample.

2.2.2 Spin state

For intrinsic paramagnetic centers, the spin state may not be known *a priori*. Usually, a solid-state continuous-wave (CW) EPR spectrum will reveal it. For instance, spectra of high-spin and low-spin Fe(III) or high-spin and low-spin Co(II) look very different. If the EPR spectrum cannot be assigned to a class of spectra with known spin state, it may be possible to determine S by nutation spectroscopy.

CW EPR spectroscopy

This experiment should be applied to any sample that is to be characterized by EPR spectroscopy, even if the sought information requires other experiments. Only that way one can ascertain that the expected paramagnetic centers exist at about the expected concentration and that no unexpected paramagnetic centers exist that could compromise the information obtained in other experiments.

In CW EPR, the sample is continuously irradiated with microwave at a fixed frequency (usually between 3 and 300 GHz), the magnetic field is varied (swept), and absorption of the microwave (mw) by the sample is recorded. For technical reasons, the signal is detected by a phase-sensitive detector that decodes a magnetic field modulation at 100 kHz, which is superimposed on the swept magnetic field. This mode of detection records the derivative of the absorption spectrum. Therefore, spin counting, i.e., the determination of the concentration of the paramagnetic species, requires double integration of the CW EPR spectrum. The modulation amplitude, measured in Gauss (1 G = 0.1 mT), should not exceed a third of the width of the narrowest line in the spectrum. Otherwise the lines are artificially broadened and resolution is lost. Typical values are $0.1 \ldots 1$ G for organic radicals in solution and $1 \ldots 10$ G for any type of system in the solid state. As long as the lines are not broadened, sensitivity is proportional to modulation amplitude.

Microwave power should be sufficiently low for electron spin relaxation being able to compete with absorption. Otherwise the electron spin transitions are saturated and the lines are also broadened, the signal may even be suppressed. Power is usually quoted as attenuation in decibel units (dB) with respect to a maximum power of 200 mW. A typical attenuation is 20 db (2 mW power). Organic radicals at cryogenic temperatures may require attenuation up to 40 dB, whereas metal ions may allow for attenuation as low as 10 dB or even application of the full power. As long as no saturation occurs, the signal doubles for every 6 dB decrease in attenuation.

Except for wrongly setting modulation amplitude or attenuation, the most frequent mistake in CW EPR spectroscopy is to measure in a too small field range. If the system is unknown, it makes much sense to start with a sweep from zero field to maximum possible magnetic field at X

band with 5 G modulation amplitude and 20 dB attenuation. This provides an overview spectrum. For the species seen in this overview spectrum one can then optimize field range, modulation amplitude and attenuation. An overview of important characteristics of the CW EPR experiment is given in Table 2.1.

Table 2.1: Characteristics of the CW EPR experiment. Sensitivity values depend on width of the spectrum and of individual features as well as on frequency.

Hardware requirements	CW EPR
Sensitivity for radicals	1 μM . . . 10 μM
Sensitivity for metal ions	10 μM . . . 100 μM
Aggregation state	liquid & solid

Nutation spectroscopy

During mw irradiation, magnetization of spin packets precesses in a rotating frame around an effective field made up of the irradiation field B_1 and the resonance offset field ΔB_0. This motion is called nutation. For on-resonance spin packets with $\Omega = 0$ corresponding to the $m_S \leftrightarrow m_S + 1$ of a spin S, the nutation frequency is given by

$$\omega_1 = \frac{g_{XY}\mu_\mathrm{B}B_1}{\hbar}\sqrt{S(S+1) - m_S(m_S+1)} \tag{2.1}$$

where g_{XY} is an effective g value in the plane perpendicular to the static magnetic field B_0, ans μ_B is the Bohr magneton. The equation applies if the individual $m_S \leftrightarrow m_S + 1$ transitions ($m_S = -S \ldots S-1$) are resolved and can thus be separately excited. If ω_1 for an $S = 1/2$ system with $g \approx g_e$ is known, a measurement of ω_1 can reveal the spin quantum number S and aid in the assignment of transitions [SJ01]. The known sample can either be measured in the same measurements session as the unknown with the same mw resonator or, if possible, an $S = 1/2$ reference can be added to the sample.

Table 2.2: Characteristics of the nutation experiment.

Hardware requirements	Pulse EPR
Sensitivity for radicals	10 μM . . . 100 μM
Sensitivity for metal ions	100 μM . . . 1 mM
Aggregation state	solid

Field-swept echo-detected EPR

In rare cases, CW EPR spectra may be missing information that can be obtained when measuring an absorption EPR spectrum. This is the case if spectral features in the solid state are much broader than the maximum attainable modulation amplitude. Such features are then not observable in the derivative spectrum obtained by CW EPR.

The absorption spectrum can be recorded by field-swept echo-detected EPR (FS ESE EPR). In general, this experiment is less sensitive than CW EPR and the lineshape may be distorted if the transverse relaxation time T_2 of the electron spins is short and orientation dependent or if there exist prominent nuclear echo modulations. If possible, both the FS ESE and CW EPR spectrum should be measured and compared. Measurement of an FS ESE spectrum is a precondition for applying any other pulsed EPR technique.

Table 2.3: Characteristics of the FS ESE EPR experiment.

Hardware requirements	Pulse EPR
Sensitivity for radicals	5 μM . . . 50 μM
Sensitivity for metal ions	50 μM . . . 0.5 mM
Aggregation state	solid

2.2.3 g tensor and zero-field splitting

The g tensor and zero-field splitting are best determined by fitting the spectrum obtained with the CW EPR experiment or, only if this fails, with the FS ESE EPR experiment (Section 2.2.2).

2.2.4 Hyperfine coupling

The large hyperfine couplings of paramagnetic transition metal ion (e.g. ^{63,65}Cu, ^{55}Mn(II), ^{59}Co(II)) are most easily obtained by fitting the CW EPR or FS ESE EPR spectrum, at least for mononuclear complexes. In multinuclear complexes, such as the oxygen-evolving complex of photosynthetic reaction centers, such fits may be ambiguous. In general, resolution can be improved and uncertainty of hyperfine couplings reduced by measuring nuclear-spin transition frequencies instead of electrons-pin transition frequencies. Since the former frequencies also depend on the nuclear Zeeman interaction, nuclear frequency spectra also aid assignment of hyperfine couplings. In particular for ^{14}N nuclei, the nuclear quadrupole coupling, which also contributes to nuclear spin transition frequencies, can further aid assignment, as it provides fingerprint information on the binding situation of the nitrogen atom. However, the nuclear quadrupole coupling also complicates spectrum analysis and may necessitate application of the two-dimensional HYSCORE experiment for assignment.

ELDOR-detected NMR

For large hyperfine couplings, the high-field approximation is violated for the nuclear spin. As a consequence, the normally forbidden transitions, where both the electron spin magnetic quantum number m_S and the nuclear spin magnetic quantum number m_I changes simultaneously, become weakly allowed in the solid-state. Selective mw irradiation at a power that leads to saturation or even spin inversion will then create a spectral center hole at the irradiation frequency and additional weaker side holes spaced from the center hole by the nuclear transition frequencies. Usually, the side holes are detected by a selective spin echo at a fixed frequency, while the hole burning pulse is applied at a different, swept frequency. This experiment is called ELDOR-detected NMR, as it measures the NMR frequencies of a paramagnetic system with an electron electron double resonance (ELDOR) technique.

Resolution of ELDOR-detected NMR is limited, since the nuclear frequencies are measured as differences between allowed and forbidden electron spin transitions that have larger linewidths than the nuclear transitions. Furthermore, the spectral hole burning approach at least doubles the line width compared to direct detection. Often, one accepts even stronger broadening by applying strong hole burning pulses much shorter than T_2, as this improves sensitivity. The experiment is thus restricted to large hyperfine couplings that are not resolved in or cannot be elucidated from CW EPR spectra, but still exceed the homogeneous EPR linewidth $1/T_2$ by a factor of at least four.

ENDOR

Nuclear spin transitions can be directly excited by applying radiofrequency (rf) irradiation. For sensitivity reasons, this is combined with a preparation that transfers the much larger polarization of electron spin transitions to the nuclear spin transitions and with indirect detection on the

Table 2.4: Characteristics of the ELDOR-detected NMR experiment. An AWG is an arbitrary waveform generator.

Hardware requirements	Pulse EPR, second mw frequency or AWG
Sensitivity for radicals	50 μM . . . 200 μM
Sensitivity for metal ions	200 μM . . . 1 mM
Aggregation state	solid

electron spins whose transitions have much larger energy and couple more strongly to the detector. Such schemes are called electron nuclear double resonance (ENDOR). ENDOR experiments can be performed with different experimental schemes. If certain conditions on relative spin relaxation rates are fulfilled, it is possible to simultaneously apply saturating mw and rf irradiation with fixed mw and swept rf frequency and observe desaturation of the electron spin transitions when the rf is resonant with a nuclear spin transition. This CW ENDOR experiment is technically challenging and may fail in the solid state if the conditions on relative relaxation rates cannot be met. Only a few groups worldwide still use it, although it is the only experiment that can measure nuclear frequencies of paramagnetic systems in the liquid state.[1].

Table 2.5: Characteristics of the CW ENDOR experiment.

Hardware requirements	CW EPR, high-power rf source, ENDOR resonator
Sensitivity for radicals	100 μM . . . 500 μM
Sensitivity for metal ions	500 μM . . . 2 mM
Aggregation state	liquid & solid

No particular balance between relaxation rates needs to be fulfilled for pulse ENDOR experiments [SJ01]. Depending on the magnitude of the hyperfine coupling, the Davies ENDOR experiment based on spectral hole burning or the Mims ENDOR experiment based on a stimulated echo sequence may be more sensitive. As a rule of thumb, Davies ENDOR works best for couplings larger than 5 MHz and Mims ENDOR for smaller ones. Sensitivity and resolution of pulsed ENDOR experiments increase drastically when going from X-band (9 GHz) to Q-band (34 GHz) mw frequencies and resolution further increases when going to even higher frequencies. The resolution increase is not only due to the broader dispersion of nuclear Zeeman frequencies at the higher magnetic fields associated with the higher mw frequencies. Spectral features also tend to sharpen when approaching the high-field regime for the nuclear spin, an effect that is especially notable in proton ENDOR. Often, the resolution of ENDOR experiments is limited by power broadening due to the application of relatively short rf pulses (typically 10 μs). For protons that have static line widths of about 100 kHz due to homonuclear dipole-dipole coupling, this is usually not limiting, but for rare nuclei with lower gyromagnetic ratio it is.

ESEEM & HYSCORE

The forbidden transitions that underlie the ELDOR-detected NMR experiment can also be utilized in Fourier-transform time-domain experiments. These experiments detect oscillations of electron spin echoes and are therefore denoted as electron spin echo envelope modulation (ESEEM) experiments. The direct time-domain equivalent of ELDOR-detected NMR is two-pulse ESEEM, which detects differences between frequencies of allowed and forbidden EPR transitions via

[1]Note, however, that paramagnetic systems may be accessible to conventional NMR in the liquid or solid state if the electron spin longitudinal relaxation rate is much larger than the nuclear Zeeman frequency. High-spin transition metal ions and rare earth ions often fall into this regime.

Table 2.6: Characteristics of the pulse ENDOR experiments.

Hardware requirements	Pulse EPR, high-power rf source, pulse ENDOR resonator
Sensitivity for radicals	50 μM . . . 200 μM
Sensitivity for metal ions	200 μM . . . 1 mM
Aggregation state	solid

oscillations of the Hahn echo decay. Since two-pulse ESEEM spectra are not better resolved than properly conducted ELDOR-detected NMR experiments, two-pulse ESEEM spectra contain more lines, and enhancement of forbidden transitions by high-turning angle pulses is more transparent and more efficient in ELDOR-detected NMR, the two-pulse ESEEM experiment is nowadays rarely applied.

Oscillations in the decay of a stimulated echo as a function of the delay between the second and third pulse result from the generation of nuclear coherence by the first two pulses. In this three-pulse ESEEM experiment, the utmost resolution is obtained as lines are broadened only by nuclear spin transverse relaxation and no power broadening due to rf irradiation is present. Yet, spectral crowding may be significant, as ESEEM requires violation of the high-field approximation, so that dispersion of nuclear Zeeman frequencies is narrow. Only in exceptional cases, such as strongly coupled nuclei with a low gyromagnetic ratio, can the experiment be applied at high fields. For protons, sensitivity drops already when going from X-band (9.6 GHz) to Q-band (34 GHz) frequencies.

The resolution problem due to spectral crowding is alleviated and assignment aided by the two-dimensional hyperfine sublevel correlation (HYSCORE) experiment that correlates the two transitions of the same nuclear spin with different electron spin magnetic quantum number m_S. Since the frequencies of these transitions differ by just the hyperfine coupling, interpretation of HYSCORE spectra for nuclei with spin $I = 1/2$, in particular protons, is straightforward. For powders or glassy frozen solutions, HYSCORE correlation patterns also separate the isotropic and anisotropic contributions to the hyperfine coupling. For nuclear spin $I > 1/2$ correlation patterns may become complicated if the nuclear Zeeman, nuclear quadrupole, and hyperfine interactions are all of the same order of magnitude.

Table 2.7: Characteristics of the ESEEM and HYSCORE experiments.

Hardware requirements	Pulse EPR
Sensitivity for radicals	50 μM . . . 200 μM
Sensitivity for metal ions	200 μM . . . 1 mM
Aggregation state	solid

2.2.5 Weak electron-electron coupling

Strong electron-electron coupling larger than about 20 MHz significantly influences the line shape of CW EPR spectra in many cases. Elucidation of the coupling tensor by line shape analysis may still be a formidable task, in particular in cases, where no reference spectra of the uncoupled subsystems are available. Pulse EPR techniques for separating interactions in this situation are still a matter of current research.

At a distance between two paramagnetic centers of 1.5 nm (15 Å), dipole-dipole coupling has dropped to 15.5 MHz and exchange coupling is yet by at least an order of magnitude lower, unless the two paramagnetic centers are linked by a chain of conjugated bonds or by spatially

overlapping π electron systems. The latter situation is sometimes encountered in electron-transfer proteins, where exchange coupling can be relevant up to distances of about 30 Å. Otherwise, electron-electron coupling can be interpreted as pure dipole-dipole coupling and the distance between the two paramagnetic centers can be inferred. Because of a lack of alternative distance measurement techniques on length scales of $15 \ldots 150$ Å that can achieve the same accuracy, EPR-based distance measurements are of great interest for characterizing structure and structural transitions of biological macromolecules. The distance range is limited by the width of EPR transitions, as the coupling is most easily measured as a difference between two EPR frequencies. This would appear to limit the distance range to about $40 \ldots 50$ Å for typical electron spin transverse relaxation times. By optimizing sample preparation and trading sensitivity against resolution with a constant-time experiment, the upper limit can be shifted to $80 \ldots 150$ Å in a number of cases.

DEER

In the DEER experiment (also known as PELDOR), all interactions including the dipole-dipole interaction are removed with a constant-time refocused echo sequence on the observer spin and the dipole-dipole coupling is selectively reintroduced by an inversion pulse at a second mw frequency that excites a fraction of the coupling partners of the observer spin. By variation of the timing of this inversion pulse, a dipolar oscillation of the observer echo is introduced, from which the distribution of distances can be elucidated.

Table 2.8: Characteristics of the DEER (PELDOR experiment.

Hardware requirements	Pulse EPR, second mw frequency or AWG
Sensitivity for radicals	10 μM . . . 100 μM
Sensitivity for metal ions	50 μM . . . 0.5 mM
Aggregation state	solid

Out-of-phase ESEEM

A special situation is encountered with radical pairs that have been generated by optical excitation with a laser. These radicals are spin-correlated, as they both stem from the same precursor molecule and angular momentum (spin) is conserved. Spin correlation decays in the weakly coupled radical pair due to angular momentum transfer to the environment (nuclear spins, the lattice), but immediately after generation of the pair it is virtually perfect. In biological systems, the pair is often generated by electron transfer from a donor to an acceptor after optical excitation of either the donor or acceptor and originates in a singlet state. The Hahn echo intensity for such a singlet radical pair exhibits a sine modulation with the electron-electron coupling. Since the signal arises 90° out of phase with a normal echo signal of a spin 1/2 species, this experiment is called out-of-phase ESEEM. Sensitivity depends on the quantum yield in generating the radical pair. Resolution may be limited if the radical pair lifetime is too short. Sometimes, the biological system can be modified by removing or reducing cofactors needed for further electron transfer in order to prolong the lifetime of the radical pair under study.

Table 2.9: Characteristics of the out-of-phase ESEEM.

Hardware requirements	Pulse EPR, laser for optical excitation
Aggregation state	solid

2.3 Choice of EPR frequency

2.3.1 Frequency bands

Microwave frequency bands are designated by a cryptic letter scheme that originated from military radar. Table 2.10 lists the band ranges and corresponding resonance fields for a species with $g = g_e$. To complete confusion, Ka band is usually called Q band in EPR spectroscopy (Q band EPR frequencies are at the lower edge of Q band, which extends from 33 to 50 GHz) and EPR probehead frequencies are not necessarily in the center of the band and may vary between probeheads for the same spectrometer. Commercial spectrometers are available for L band, S band, X band (most common), Q band, W band, and in mm band (263 GHz). Sample dimensions should not usually exceed a quarter of the wavelength when working with single-mode resonators.

Table 2.10: Microwave band designations. Resonance magnetic fields correspond to the center of the band an $g = g_e$.

Microwave band	L	S	X	Ku	K	Q (Ka)	V	W	mm
Frequency [GHz]	1-2	2-4	8-12	12-18	18-27	27-40	40-75	75-110	110-300
Magnetic field [T]	0.0535	0.107	0.357	0.535	0.803	1.195	2.05	3.3	7.3
Wavelength [mm]	200	100	30	20	13	9	5	3	1.5

2.3.2 Resolution

With increasing frequency, the resonance field increases and g value resolution thus also increases. For organic radicals, where g value dispersion is small, the best resolved EPR spectra are thus usually obtained at the highest accessible field/frequency combination. The same does not usually apply to paramagnetic metal ions. Especially in frozen glassy solution, but often even in microcrystalline powders, the principal values of the g tensor are also dispersed, an effect called *g strain*. When g strain dominates line width, hyperfine resolution decreases linearly with increasing field, since hyperfine couplings are field independent. For metal ions, hyperfine resolution in EPR spectra usually decreases already when going from X to Q band and sometimes already when going from S to X band.

The situation is different in ENDOR spectroscopy. Here, g strain hardly plays a role and the dispersion of nuclear Zeeman frequencies improves and spectral features typically sharpen for any type of sample.

2.3.3 Sensitivity

In NMR spectroscopy, sensitivity increases with increasing field and frequency. In EPR spectroscopy, we need to distinguish between absolute sensitivity, i.e. the minimum *number* of spins that can be detected, and concentration sensitivity, i.e., the minimum concentration at which a paramagnetic species can still be detected. With single-mode resonators, absolute sensitivity increases with increasing frequency up to the mm band. Concentration sensitivity increases only up to Q band for the reason that sample volume must be decreased when the sample dimensions become shorter than about a quarter of the wavelength. Furthermore, noise characteristics and other performance parameters of mw components typically worsen at higher frequencies. At Q-band frequencies the required sample volume ranges between 8 μL for standard tubes and 40 μL for slightly oversized samples that provide best concentration sensitivity. A sensitivity advantage at higher frequencies is expected only if available sample volume is smaller than that.

For techniques that depend on formally forbidden electron-nuclear transitions, such as ELDOR-detected NMR, ESEEM, and HYSCORE, sensitivity usually decreases with increasing

frequency once the nuclear Zeeman frequency is twice as large as the hyperfine coupling. Proton HYSCORE is often most sensitive in X band, whereas HYSCORE of heteronuclei with moderate hyperfine couplings and moderate gyromagnetic ratio is most sensitive in Q band and HYSCORE and ELDOR-detected NMR for strongly hyperfine coupled nuclei with low gyromagnetic ratio may be most sensitive in W band.

3 — Samples and measurement temperature

3.1 Concentration

Too large sample concentration may cause line broadening and, in extreme cases, may even reduce hyperfine splittings or cause collapse of hyperfine structure. Unless there are special reasons, sample concentration should not exceed 1 mM for organic radicals or 2 mM for transition metal complexes. With modern spectrometers, concentrations of 100 . . . 200 μM for radicals and 1 mM for metal ions are recommended for recording CW EPR or echo-detected field-swept EPR spectra and for hyperfine spectroscopy. The measurement of very long distances by DEER (>60 Å) may require even lower concentrations between 10 and 50 μM. Usually, detection of EPR spectra becomes difficult at concentrations below 1 μM.

For solvent-free samples, diamagnetic dilution may be necessary, i.e. the spin-labeled macromolecule may have to be diluted into the wildtype or into the same mutant labeled with a diamagnetic analog of the spin label. For detergent-solubilized membrane proteins, the same concentration requirements apply as for other samples. For membrane proteins reconstituted into liposomes, lipid-to-protein ratios of about 1000 are recommended.

3.2 Sample size

For EPR experiments on solid samples, typical sample tube outer diameters are 5 mm or lower in S-band, 4 mm or lower in X band, 1.6 mm in Q band, and 0.9 mm in W band, corresponding to sample volumes of about 250 μL (S), 150 μL (X), 10 μL (Q) and 1 μL (W). For pulse EPR experiments, larger sample volumes that lead to dielectric mw losses and thus reduce resonator quality factor are feasible if the sufficient mm power is available. With 150-200 W power at Q band and 1 kW power at W band, 3 mm sample tubes with a volume of about 40 μL can be used. For liquid samples in polar solvents, in particular water, tube diameter must be smaller. Typically, capillaries with 0.9 mm outer diameter at X band and 0.3 mm outer diameter at W band are used. If concentrations are very low, volume of aqueous samples can be increased with flat cells or bundles of capillaries, such as the AquaX system of Bruker.

3.3 Presence of oxygen

Oxygen has a triplet ground state and causes paramagnetic relaxation enhancement on other paramagnetic centers both in solution and in the solid state. The effect is usually mild in aqueous

solution, but may be more prominent in unpolar solvents or lipid membranes. Deoxygenation may thus be necessary if utmost resolution is required, but can often be avoided.

3.4 Maximizing transverse relaxation time

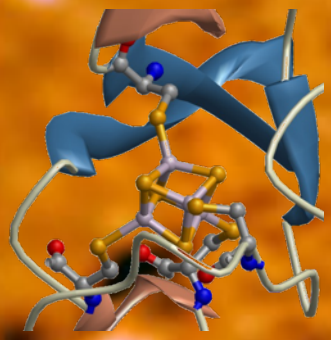
Pulsed EPR experiments depend critically on transverse relaxation time, since the typical duration of the pulse sequences is of the same order as the phase memory time T_m. This applies in particular to pulsed dipolar spectroscopy, such as DEER spectroscopy, where a constant-time observer pulse sequence is applied that is typically significantly longer than T_m. If the echo signal is observed at 10% of its maximum amplitude, a doubling of T_m will increase it to 31.6% of its maximum amplitude.

At the same time, longitudinal relaxation time T_1 should not be unnecessarily prolonged, as this reduces the permitted repetition rate of the experiment. Typical repetition times are $1 \ldots 4$ ms. In the solid state, electron spin T_1 usually increases monotonically with decreasing temperature. In contrast, T_m approaches a rigid-limit maximum, where transverse relaxation is dominated by fluctuations of the hyperfine field of matrix nuclei due to nuclear spin diffusion. Spin diffusion is driven by homonuclear couplings and is not temperature dependent in the rigid limit. The rigid limit of T_m is attained for nitroxide radicals at about 50 K and for metal centers usually between 15 and 6 K. For high-spin species the rigid limit may be attained only well below 4.2 K, which is hard to sustain for a long time with a helium flow cryostat. Note that T_m is very short at temperatures where the time scale of methyl group rotation is of the order of 1 ns to 10 μs. For nitroxides, this happens between about 90 and 160 K. For organic radicals with only a single methyl group attached to the carbon atom(s) in α-position to the paramagnetic center, this regime may be encountered at lower temperatures.

In the low temperature limit, the phase memory time depends on the concentration and gyromagnetic ratio of nuclear spins surrounding the electron spin. The gyromagnetic ratio can be reduced by substituting protons by deuterons. Often it is easily possible and sufficient to use deuterated water in the buffer and deuterated glycerol as a cryoprotectant. However, deuteration of the biological macromolecule(s) can further prolong T_m.

3.5 Cryoprotectant

Pulse EPR experiments and recording of a rigid-limit CW EPR spectrum need to be performed in glassy frozen solutions. If a buffer solution is frozen without a cryoprotectant, ice crystals will form that may distort structure, lead to precipitation and to breaking of the sample tube. This is avoided by adding a cryoprotectant, for soluble proteins usually between 25% and 50% of glycerol (v/v), and shock-freezing the samples by immersion of the tube into liquid nitrogen. With less than 25% of glycerol, good glasses cannot be obtained even when the sample tube is immersed into a very cold liquid (isopentane or ethanol cooled to about 120-130 K), which leads to better heat transfer and thus faster freezing. Breaking of the sample tube can usually be avoided down to about 10% glycerol by adding a small piece of thin teflon tubing (1 mm outer diameter) to the sample before freezing. For membrane proteins reconstituted into liposomes, 10% glycerol is amply sufficient and freezing without cryoprotectant may be feasible. Usually, addition of glycerol stabilizes the native state of proteins. However, it may influence formation of weak protein-protein complexes. For nucleic acids, dimethyl sulfoxide (DMSO, 50% v/v) is also a good cryoprotectant.



4 — Intrinsic paramagnetic centers

4.1 Radical enzymes

Some bioenergetics processes and biosynthesis of some metabolites involve one-electron transfer steps. One-electron transfer generates two neutral radicals, if it is coupled to proton transfer, or a cation radical from the electron donor and an anion radical from the electron acceptor, if it is not coupled to proton transfer. Often nature relies on moderately stabilized radicals, probably because this allows for better control of the reactions and less side products and protein damage. In most cases the radicals are stabilized by delocalization of the unpaired electron in a π system. This causes spin density on many nuclei and thus a rich hyperfine spectrum that can be used to recognize the type of radical. The radical can be localized on an amino acid or on a cofactor.

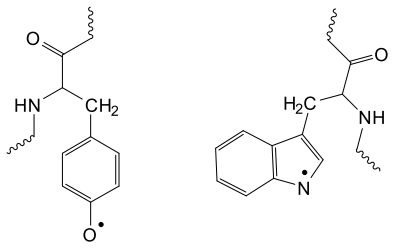


Figure 4.1: Structures of the tyrosyl and tryptophan radical.

4.1.1 Radicals derived from amino acids

The two amino acids that are best suited to stabilize radicals are the aromatic amino acids tyrosine (Y) and tryptophan (W) (see Fig. 4.1). Radicals centered on the Cα atom of glycine (G) have also been observed and thiyl radicals centered on the Sγ atom of cysteine (C) have been postulated and in some cases observed [Kol+02]. The latter radicals are often hard to detect, as in thiyl radicals frontier orbitals are degenerate or near-degenerate, which leads to large spin-orbit coupling and very broad EPR spectra that are furthermore highly sensitive to small changes in the radical

environment that lift orbital degeneracy. Radicals may also be derived from modified amino acids, for an overview see [Jes05].

The small g value differences of organic radicals are not well resolved at X-band frequencies. This makes it difficult to assign the radical type from X-band EPR spectra, as can be seen by comparing the simulated X-band CW EPR spectra of a tyrosyl (Fig. 4.2a) and tryptophan (4.2a) radical (parameters taken from [Ble+01]). In contrast, at the higher W-band frequency of 94 GHz g value differences are resolved and the spectra differ considerably.

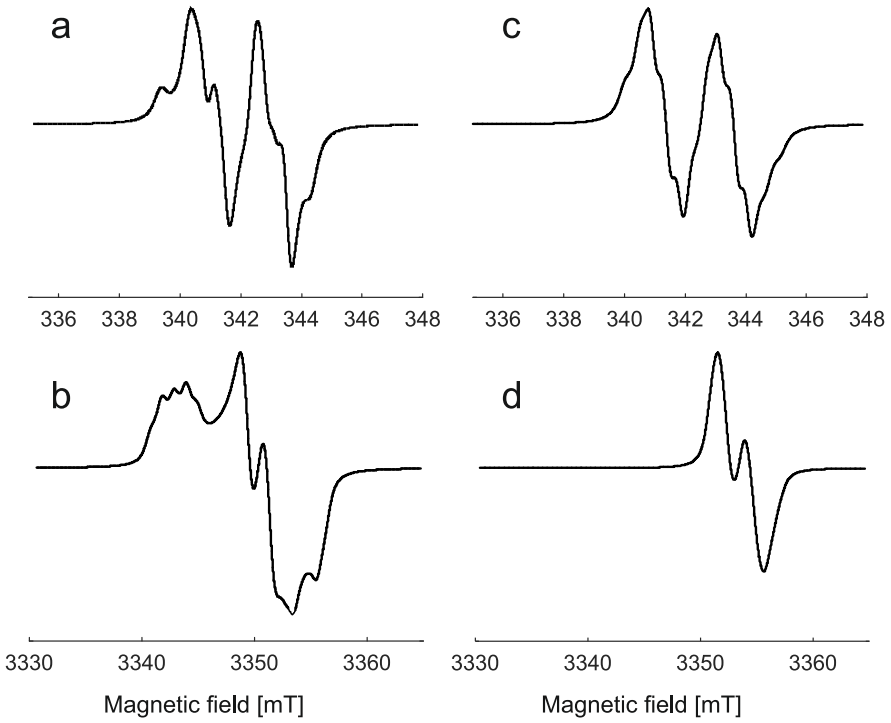


Figure 4.2: Simulated CW EPR spectra of a tyrosyl radical (a,b) with principal g values 2.0091, 2.0046, and 2.0021 and a tryptophan radical (c,d) with principal g values 2.0033, 2.0024, and 2.0021 at an X-band frequency of 9.5 GHz (a,c) and a W-band frequency of 94 GHz (b,d). The simulation scripts for EasySpin are available on the course homepage. Parameters were taken from [Ble+01], ignoring the rotation of the hyperfine tensor principal axes system (PAS) with respect to the g tensor PAS.

In many cases, different tryptophan or tyrosine residues are potential radical sites. Unless EPR measurements can be performed on single crystals, is not usually possible to make an assignment to a particular site on the basis of the EPR spectra. Such an assignment requires a mutation strategy, where potential radical sites are removed and EPR spectra are recorded to check for disappearance of the signals or for spectral changes. Often, spectral changes rather than disappearance of the EPR signal is observed, as the radical is stabilized at a different residue. This is because electron transfer usually occurs along a chain of residues. Mutation then blocks onward transfer and causes stabilization of the radical at an earlier site in the transfer chain.

4.1.2 Cofactor-derived radicals

Some cofactors in proteins are primed for one-electron transfer. In particular, quinones are ubiquitously utilized in redox reactions in living cells and the quinone/hydroquinone pair serves in redox pools in mitochondria and chloroplasts. The intermediate between the diamagnetic quinone and hydroquinone is the semiquinone radical (Fig. 4.3) that is fairly stable and can

exist in a anionic or neutral form. Hyperfine couplings in semiquinone radicals are sensitive to hydrogen bonding to the radical by amino acid residues in their environment. The maximum hyperfine coupling is smaller than in tyrosyl or tryptophan radicals, so that they can usually be distinguished from those radicals already at X-band frequencies (see Fig. 4.4(a)). Yet, their X-band EPR spectrum is not very characteristic, whereas the W-band EPR spectrum reveals the typical g tensor with nearly axial symmetry and $g_\perp > g_{||}$ (see Fig. 4.4(b)).

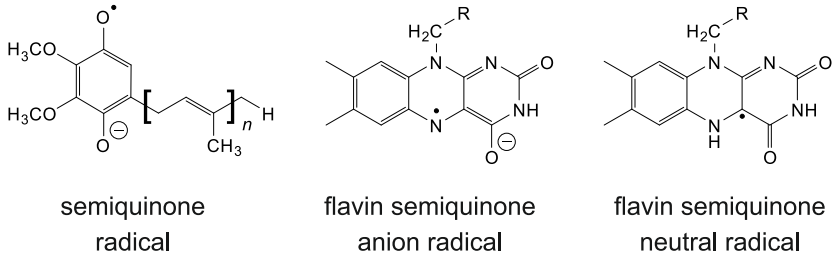


Figure 4.3: Structures of some frequently encountered cofactor radicals.

The flavin group (Fig. 4.3) is also used in many redox enzymes, often in the form of flavin adenine nucleotide (FAD) that forms a two-electron redox pair with $FADH_2$. The intermediate state of this pair is a flavine semiquinone that can exist in neutral form or as an anion radical. The g tensors of these two radicals are very similar and have nearly axial symmetry with $g_\perp > g_{||}$ similar to ubiquinone radicals, but with a slightly smaller $g_\perp$ value. The main difference between ubiquinone and flavin radicals lies in the hyperfine couplings. The additional hyperfine coupling to the proton on the N^5 atom in the neutral flavine radical as compared to the flavin anion radical is the main difference between the two. This difference is best recognized in ENDOR spectra, but can also be seen in the EPR spectrum taken at 94 GHz (see Fig. 4.4(d)).

There exists a variety of other cofactors that can form radicals. In photosynthetic reaction centers, radicals are primarily formed by one-electron transfer from a chlorophyll cofactor or a pair of chlorophyll cofactors to pheophytin cofactors. Subsequent electron transfers generate a tyrosyl radical and a quinone radical and finally lead to reduction of plastoquinone. In the water-splitting photosystem II, the transferred electrons are replaced by electrons generated in the oxygen-evolving complex. Intermediates of most electron transfer steps in this cycle have been detected by EPR spectroscopy.

4.2 Metalloproteins

4.2.1 Mononuclear metal centers

Since transition metals often can exist in several stable oxidation states and are easily coordinated by organic ligands, they are well suited to redox catalysis. In proteins, transition metal ions are often bound by cofactors that act as ligands, but coordination by only amino acid side chains is also possible. The variety of transition metal centers in proteins is large. Many of them go through a paramagnetic state or even several paramagnetic states during enzyme action. In the following, only a few examples can be discussed. A more comprehensive overview is given in [HKS96].

Heme proteins

Heme proteins epitomize the versatility of mononuclear metal centers in biological systems. They are involved in ligand transport, most famously oxygen transport by hemoglobin, ligand storage, electron transfer, enzymatic catalysis, and regulation processes [SKT10]. The heme

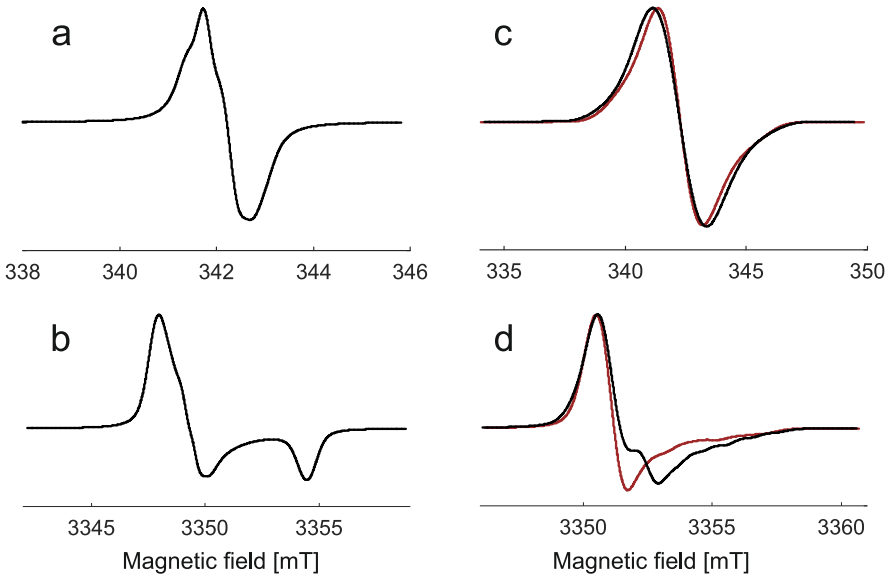


Figure 4.4: Simulated CW EPR spectra of an ubiquinone radical (a,b) with principal g values 2.00605, 2.00519, and 2.00213 [Gri+01] and a neutral flavin semiquinone radical (c,d) with principal g values 2.00425,2.00360, and 2.00227 [Bar+03] at an X-band frequency of 9.5 GHz (a,c) and a W-band frequency of 94 GHz (b,d). The red lines in (c,d) correspond to a flavine anion radical with principal g values 2.00436,2.00402, and 2.00228 [Bar+03]. The simulation scripts for EasySpin are available on the course homepage. Rotation of the hyperfine tensor principal axes system (PAS) with respect to the g tensor PAS was ignored for the ubiquinone radical, but considered for the flavin radical.

group consists of an iron ion bound to a porphyrin ligand. In the resting state, usually an Fe(II) ion is encountered (ferrous form), which can be in a low-spin state ($S = 0$, diamagnetic) or in a high-spin state ($S = 2$) that can usually be detected by EPR only at Terahertz frequencies. EPR-accessible Fe(III) (ferric form) is intermediately formed during function of heme proteins. The ferric form exists in a low-spin state ($S = 1/2$) when the iron ion is coordinated by two strong bases and otherwise in a high-spin state ($S = 5/2$). Occasionally, temperature-dependent spin crossover between the high-spin state and an intermediate spin state ($S = 3/2$) is observed. The spectra of high-spin and low-spin heme centers differ considerably. Other paramagnetic intermediates exist, most famously "compound 1", which is considered as an O=Fe(IV) (ferryl) center spin-coupled to a porphyrin cation radical. A more detailed discussion can be found in [Doo17].

Most low-spin Fe(III) complexes have a $(\mathrm{d}_{xy})^2(\mathrm{d}_{xz},\mathrm{d}_{yz})^3$ electron configuration [Wal99]. Among those, the ones with a near degeneracy of the d_{xz} and d_{yz} orbitals have a large g anisotropy (Fig. 4.5a) and their spectra can often be observed only at very low temperatures. Such low-spin Fe(III) centers are referred to as Type I. If the energies of the d_{xz} and d_{yz} orbitals are well separated, g anisotropy is much smaller (Fig. 4.5b) and the spectra can usually be observed already at a temperature of 78 K. Such low-spin Fe(III) centers are referred to as Type II. For Type I and Type II centers, $g_x^2 + g_y^2 + g_z^2 = 16$ can usually be assumed. Low-spin Fe(III) Type III centers have a $(\mathrm{d}_{xz},\mathrm{d}_{yz})^4(\mathrm{d}_{xy})^1$ ground state and $g_x \approx -2.2 \ldots -2.5$, $g_y \approx 2.2 \ldots 2.5$, and $g_z \approx -1.7 \ldots -1.9$.

High-spin Fe(III) complexes have zero-field splittings that are much larger than the electron Zeeman interaction at any accessible magnetic field. They can be described by an effective spin $S' = 1/2$ that has g values $g_\perp \approx 6$ and $g_{||} \approx 2$ if the zero-field splitting tensor has axial

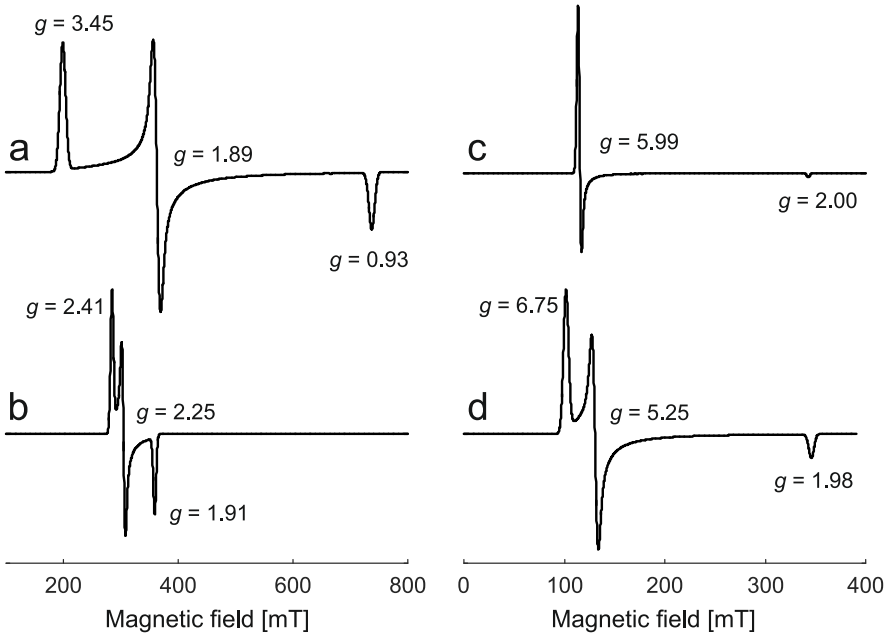


Figure 4.5: Simulated CW EPR spectra of heme proteins at 9.5 GHz in their ferric state (Fe^{III}). (a) Type-I low-spin heme (Myoglobin-CN) (b) Type-II low-spin heme ($P450_{cam}$) (c) High-spin heme with axial symmetry (human myoglobin). (d) High-spin heme with orthorhombic symmetry (H93Y mutant of human myoglobin, only one of two species simulated).

symmetry (Fig. 4.5c). In orthorhombic symmetry, the $g_\perp$ feature splits into a g_x feature with $g_x > 6$ and a g_y feature with $g_y < 6$ (Fig. 4.5d). The limiting value for g_y is 4.3 at $|E/D| = 1/3$. The large g_x and g_y values allow for an unequivocal assignment of the spin state.

Copper proteins

Metal centers in proteins are often characterized by unusual coordination geometries that facilitate electron transfer or catalysis. Cu(II) proteins are a case in point. Usually, Cu(II) prefers square planar coordination (equatorial plane), sometimes with one or two auxiliary axial ligands at somewhat longer coordination distances. This coordination geometry also exists in some copper enzymes and is then referred to as Type II (see Fig. 4.6a). The g tensor and the ^{63,65}Cu hyperfine tensors have axial symmetry. However, in multicopper oxidoreductases, Cu(II) experiences trigonal planar coordination by one cysteine ligand and two histidine ligands with a variable axial ligand. This coordination is achieved by the protein presenting the amino acid side chains in an appropriate geometry. Such Cu(II) sites are referred to as blue copper sites or Type I sites. They cycle between Cu(II) and Cu(I) states during electron transfer processes and can be further subdivided in structural classes. In the unperturbed form, the g and Cu hyperfine tensors are also axial, but the hyperfine coupling is much smaller than in square planar complexes (see Fig. 4.6b). In the perturbed form, with a shorter Cu-S bond and non-optimal angles of the coordination polyhedron, the tensors become rhombic and the Cu hyperfine coupling is resolved also in the x direction (see Fig. 4.6c).

4.2.2 Binuclear metal centers and metal clusters

Many transition-metal active centers in proteins are binuclear or multinuclear and some of them contain several strongly coupled paramagnetic metal centers. Hydrogenases are a case in point, with some of them containing [NiFe] centers (see Figure 4.7) and others [FeFe] centers. In the

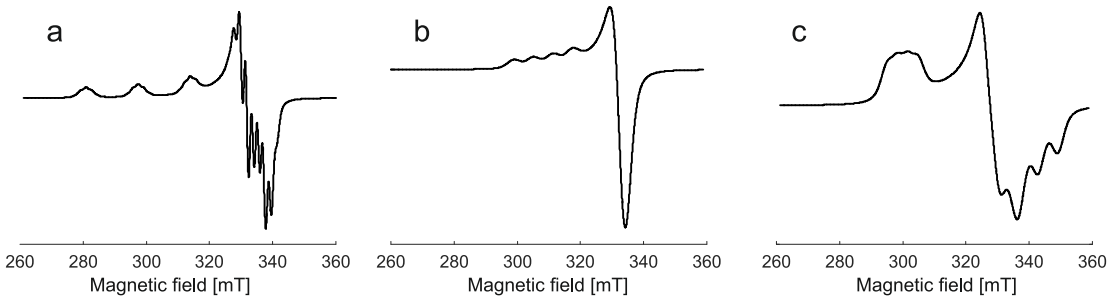


Figure 4.6: Simulated CW EPR spectra of copper proteins at 9.5 GHz. (a) Type-II (particulate methane monooxygenase from *Methylococcus capsulatus*) [Lem+00] The hyperfine structure between 330 and 340 mT is due to ^{15}N nuclei in a ^{63}Cu and ^{15}N isotope-labelled sample. (b) Undistorted Type-I center of poplar plastocyanin (parameters from [HKS96]) (c) Distorted Type-I center of stellacyanin (parameters from [HKS96]).

[NiFe] centers the Ni ion is coordinated by four cysteines, of which two form sulfur bridges to the Fe ion, and a fifth atom X, which from X-ray crystallography can be assigned only as either nitrogen or oxygen. The Fe ion is coordinated by the two bridging cysteines, two cyanide ions, and a carbon monoxide molecule. The [NiFe] center exhibits a variety of redox states of which many give EPR spectra [LRG07]. A highly active, reduced form is the Ni-C center with typical g values of 2.01, 2.15, and 2.21. Inhibition of the enzyme by flushing with carbon monoxide leads to a different EPR spectrum with g values of 2.02, 2.07, and 2.12 (Ni-CO state). Hyperfine spectroscopy of hydrogenases has provided rich information on electronic structure of the different states and on biosynthesis of the binuclear complex with its for a biological system unusual ligands.

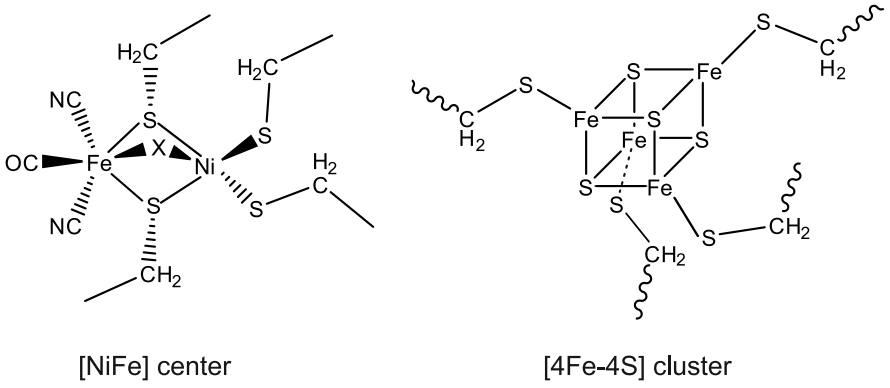
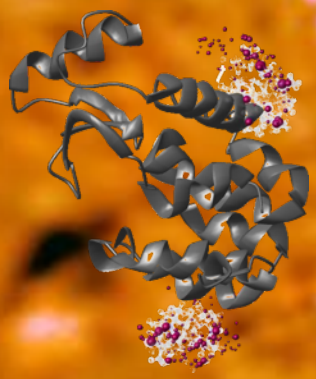


Figure 4.7: Schematic structures of the binuclear metal center in [NiFe] hydrogenase and of an [4Fe-4S] iron sulfur cluster.

The [NiFe] center in hydrogenases is in close enough proximity to a chain of *iron-sulfur clusters* to engage in electron transfer. Typically, the optimum distance is around 12 Å. At low temperatures (4.2 K), spin-spin coupling to the proximal [4Fe-4S] cluster is revealed by splittings in the EPR spectra of both the Ni-C and Ni-CO state [LRG07]. Iron-sulfur clusters form a class of ubiquitous relay stations in electron transfer chains, possibly, because life originated under anaerobic conditions on the surface of iron-sulfur minerals. Such clusters exist with different stoichiometries in different oxidation states, such as [2Fe-2S]$^{n+}$ with $n = 1, 2$, [3Fe-4S]$^{n+}$

with $n = 0, 1$, and [4Fe-4S]$^{n+}$ with $n = 1, 2, 3$. The spins of the individual iron centers are strongly coupled, so that the whole cluster may assume an $S = 1/2$ low-spin state even though the constituent Fe centers are high spin [95].



5 — Site-directed spin labeling

5.1 Choice of spin labels

The ideal label would be exactly the size and shape of the amino acid side group or nucleobase to be replaced and chemically stable under all conditions where the system is to be studied. Its EPR spectrum would be narrow, so that it could be fully excited by a microwave pulse, yet the spectral parameters would be sensitive to the environment (pH, polarity, oxygen concentration, dynamics). The label would have a very long transverse relaxation time T_2 in order to allow for high-resolution measurements and a very short longitudinal relaxation time T_1 in order to allow for their fast repetition. Last but not least it would be easy to synthesize with different reactive linker groups for attachment to proteins and RNA.

It turns out that not all these properties can be realized with the same type of label. All labels are at least slightly larger than amino acid side chains or nucleobases. Nitroxide labels come closest to this size. Any system with unpaired electrons will react under strongly reducing or strongly oxidizing conditions, with nitroxides being more susceptible to reduction than Gd(III) or trityl labels. Labels with a very narrow EPR spectrum are unlikely to exhibit high sensitivity to environmental parameters and have very low sensitivity to dynamics. For these reasons, the choice of label depends on the problem at hand.

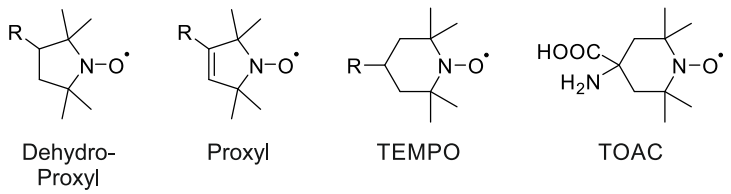


Figure 5.1: Structures of nitroxide radicals commonly used for spin labeling and spin probing. R stands for a functional group that either directs the probe to the site of interest or serves as the linker to a protein or nucleic acid site.

5.1.1 Nitroxide labels

The smallest and most versatile labels a nitroxides (Fig. 5.1). They are commercially available and for most of them synthesis is facile. Nitroxides are stable under ambient conditions and throughout the temperature range where proteins can be expected to be stable. The N-O group is however rather susceptible to reduction. Ascorbic acid at pH 7 (standard redox potential 0.06 V) reduces it to a hydroxylamine group (N-OH) that is EPR silent. In the absence of ascorbic acid and the presence of oxygen, the hydroxylamine is oxidized to the nitroxide, but this reaction may be slow. Importantly, nitroxides are reduced on a time scale of minutes to hours (depending on vicinal groups) in living cells, but not in the periplasm of gram-negative bacteria. In-cell EPR experiments with nitroxides that take several hours thus require that the cells are frozen. Nitroxide radicals are not stable at pH < 2 or pH > 12. In general, nitroxides based on five-membered rings (Proxyl- and Dehydroproxyl-Nitroxides) are more stable than those with six-membered rings (TEMPO derivatives). Chemical stability can also be improved by substituting the four methyl groups attach to the two α carbon atoms by ethyl or spirohexyl groups.

Nitroxide spin labels have relaxation times T_2 and T_1 in the low microsecond range at ambient temperature. In the rigid limit, T_2 can be as long as a few microseconds whereas T_1 is in the millisecond range at the optimum measurement temperature of about 50 K, where this limit is attained. For nitroxides with methyl groups attached to the α carbons, the phase memory time is very short between about 100 and 170 K, where methyl group rotation is neither fast enough for the anisotropic proton hyperfine coupling to average nor slow enough for it to be static on the time scale of spin echo experiments.

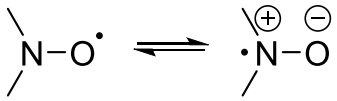


Figure 5.2: Dependence of the nitrogen hyperfine coupling of nitroxides on polarity of the environment. The mesomeric structure, where the unpaired electron is localized on N, is polar and thus stabilized by a polar environment.

Polarity dependence of the EPR spectrum

The EPR spectra of nitroxide radicals are highly sensitive to their environment. First, the spin density distribution between the nitrogen and oxygen atom of the N-O group is influenced by polarity of the environment (Fig. 5.2). Formally, the unpaired electron is localized on the oxygen atom for a neutral charge distribution. Since the by far most abundant oxygen isotope ^{16}O does not carry nuclear spin, this limiting structure does not contribute to hyperfine coupling. The limiting structure, where the unpaired electron is localized on the nitrogen atom, is charge-separated. Since the highly aboundant isotope of nitrogen is ^{14}N with nuclear spin 1, this limiting structure does contribute to hyperfine coupling. The mesomeric "equilibrium" between these two structures is shifted more towards the charge-neutral structure in unpolar environments and more towards the charge-separated structure in polar solvents. Hence, ^{14}N hyperfine coupling is stronger in polar than in unpolar environments.

For a given nitroxide label, the isotropic ^{14}N hyperfine coupling a_{iso} observable in liquid solution or the hyperfine anisotropy $2A_{zz}$ (Fig. 5.3) observable in frozen solution in the rigid limit are thus a measure for polarity of the nitroxide environment. In practice, a_{iso} is not easily observable for a nitroxide label bound to a macromolecule, which often tumbles too slowly for complete averaging of the hyperfine anisotropy (*vide infra*). Hence, the rigid-limit CW EPR spectrum that is best observed at 110 to 120 K is used for this purpose. Often, $2A_{zz}$ in the rigid

limit is interpreted as a measure for water accessibility of the spin-labeled residue. Note however that charged residues or cofactors in the vicinity may compromise such interpretation.

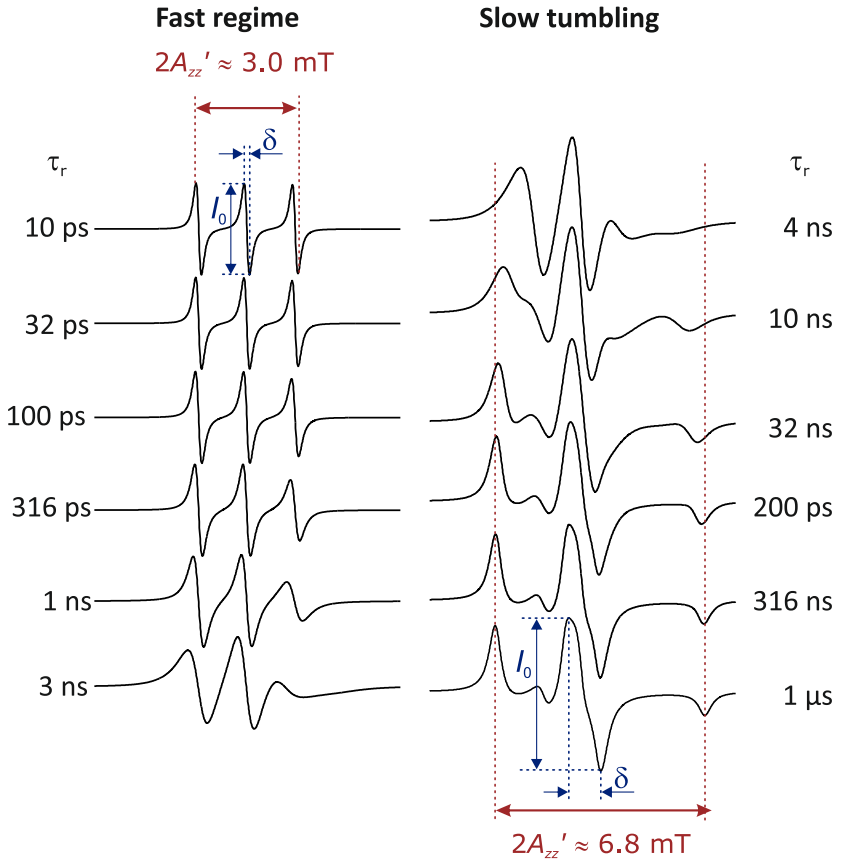


Figure 5.3: The shape of the nitroxide lineshape (here, simulations at an X-band frequency of 9.6 GHz are shown) depends on rotational correlation time τ_r and thus on restrictions of nitroxide mobility by the environment. Relative mobility for a given label type can be characterized by the parameters $2A'_{zz}$ (outer extrema splitting) or δ (width of the central line). The latter parameter is more robust for complicated types of motion or when subensembles with different mobility exist. The amplitude I_0 of the central line is analyzed in progressive power saturation experiments.

Mobility dependence of the EPR spectrum

Rotational motion of a nitroxide non time scales between about 100 ps and 100 ns strongly influences the spectra lineshape observed by CW EPR (Fig. 5.3). This dependence comes about by the combined anisotropy of the hyperfine and g tensor, which is of the order of 1/(3.5 ns). Since sidegroup motion in proteins falls into this time scale, nitroxide EPR spectra report on restraints on such motion by neighboring sidegroups or molecules.

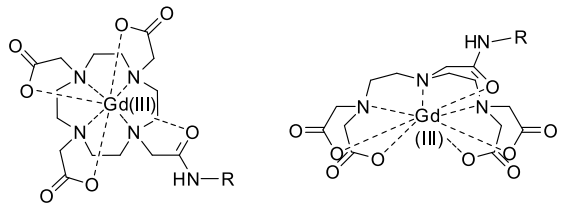
Much research has been invested onto understanding CW EPR lineshapes of nitroxide labels in proteins quantitatively. In favorable cases, the lineshape can be predicted by sophisticated molecular dynamics simulations if the structure is already known. In order to obtain restraints on structure, less sophisticated analysis is sufficient. In such analysis, a mobility parameter is derived from either the outer extrema splitting $2A'_{zz}$ or the width δ of the central line (Fig. 5.3). The latter parameter is preferable if subensembles with different mobility are suspected. Usually, one does not try to interpredt δ in terms of a time scale, but rather compares the δ parameters obtained for different labeling sites in the same protein or in different states of a protein or complex in order

to obtain information on relative mobility or, with a site scan, on secondary structure (*vide infra*).

Proximity of paramagnetic quenchers

The collision of other paramagnetic molecules or even their passing by at close proximity enhances electron spin relaxation. This effect can be detected by progressive power saturation experiments (*vide infra*). Since dioxygen has a triplet ground state, it acts as such a paramagnetic quencher and can be detected via changes in nitroxide relaxation. This technique is of interest for membrane proteins, since oxygen is much better soluble in lipids than in water. Labeling sites that are buried in the protein are not al all oxygen accessible and show no effect.

It is also possible to apply water-born paramagnetic quenchers. The most popular one is the charge-neutral complex NiEDDA of a high-spin Ni(II) ion. Nitroxides on the water accessible surface of a protein are strongly influenced by this quencher, while the influence on bilayer-immersed sites or buried sites depends on the immersion depth.



[Gd(DOTA)] [Gd(DTPA)]$^-$

Figure 5.4: Structures of Gd complexes with DOTA or DTPA ligands that can be used as spin labels. R stands for the functional group that serves as the linker to a protein or nucleic acid site.

5.1.2 Gd(III) and Cu(II) labels

Gd(III) is a high-spin Kramers ion with electron group spin $S = 7/2$ and a half-filled f shell that leads to a nearly isotropic g value. The metal ion can accommodate up to nine ligands, which allows for designing chelate complexes with extremely high thermodynamic stability that are also kinetically stable. The EPR spectra are several Gigahertz wide, but the central $m_S = -1/2 \leftrightarrow +1/2$ transition is not affected to first order by the zero-field splitting and is thus much narrower. It is mainly broadened by second-order effects of the zero-field splitting and thus becomes narrower at higher magnetic fields where the electron Zeeman interaction dominates more strongly. In practice, Gd(III) can be used as a spin label at Q band frequencies and higher. Typical chelate ligands, such as DOTA or DTPA, coordinate eight of the nine ligation sites (Fig. 5.4).

Gd(III) labels are significantly bulkier than nitroxide labels, which limits accuracy of relating Gd(III)-Gd(III) distances to backbone-backbone distances in the macromolecule. Furthermore, polarity and mobility information is not easily obtained from Gd(III) CW EPR spectra, which may even be difficult to detect under ambient conditions. Since Gd(III) EPR transitions are harder to saturate than those of nitroxides, Gd(III) is also not well suited for detecting proximity of paramagnetic quenchers.

In order to attain the rigid-limit of transverse relaxation of Gd(III) labels, the sample needs to be cooled to between 6 and 10 K. Even at these temperatures, the longitudinal relaxation time of a few hundred microseconds is still shorter than the one for nitroxide labels at 50 K. This is actually an advantage, since faster repetition of the experiments at temperatures where the Boltzmann polarization of the EPR transitions is stronger, enhances sensitivity.

Although Gd(III) labels are not as versatile as nitroxide labels, they are the labels of choice in certain situations. First, Gd(III) labels are chemically even more stable than nitroxide labels, in particular under mildly reducing conditions in living cells. Second, in macromolecular complexes it may be necessary to introduce different types of labels in order to be able to unequivocally assign measured label pair distances. This strategy is called orthogonal spin labeling and it profits from the strong spectroscopic differences between nitroxide an Gd(III) labels (low spin versus high spin, different g values and spectra widths, different relaxation behavior).

The same ligands that bind Gd(III) also bind other lanthanide ions that induce pseudo-contact shifts (PCS) and paramagnetic relaxation enhancement (PRE)*paramagnetic relaxation enhancement* in NMR spectroscopy. The same type of labeling can thus provide a large number of NMR restraints on top of the long-range EPR distance distribution restraints. Furthermore, PCS data for nuclei in the vicinity of the labeling site can be used to derive the spatial distribution of the label with respect to the protein backbone [Sut+17].

Cu(II) ions have spin $S = 1/2$ and can be bound by the same DOTA ligand as Gd(III) ions. Alternatively, an EDTA ligand with a thiol-reactive maleiimido group is also commercially available. Due to g and hyperfine anisotropy, Cu(II) EPR spectra are usually wider than 1 GHz already at X band. This limits sensitivity on pulsed EPR experiments with Cu(II) labels and it shifts the time scale for assessing mobility by about an order of magnitude towards shorter times. Cu(II) complexes are not sufficiently sensitive to polarity of the environment or to experiments with paramagnetic quenchers at ambient temperature in order to derive information from them. Nevertheless, Cu(II) can be valuable as a label that is spectroscopically orthogonal to nitroxides.

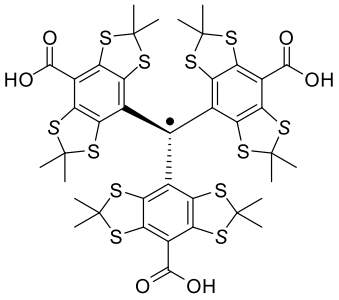


Figure 5.5: Structure of "Finland" trityl that is designed to minimize hyperfine broadening of the EPR spectrum. A reactive linker can be attached via one of the carboxylate groups by an ester or amide link.

5.1.3 Trityl labels

The anisotropic ^{14}N hyperfine coupling of nitroxide labels is a source of information, but also a source of electron spin relaxation and spectral broadening. Trityl spin labels were designed in order to minimize hyperfine broadening, which led to very narrow EPR lines at low and moderate magnetic fields and to slower electron spin transverse relaxation at ambient temperature. Trityl radicals are also less sensitive to mildly reducing conditions than nitroxide radicals.

These properties make trityl radicals very attractive for distance distribution measurements. Yet, they are much bulkier than nitroxide radicals and thus more strongly perturbing. Currently, wider application of trityl radicals is held back by their rather tedious synthesis, in particular, regarding modification with a reactive linker. Unless nitroxide and Gd(III) labels, trityl labels with a thiol-reactive linker are not yet commercially available.

5.1.4 Paramagnetic metal ion substitution

A special type of spin labeling is possible for systems that bind diamagnetic metal ions. For instance, Mg(II) is often involved in nucleotide binding by proteins and also plays a role in some enzymes. Due to having the same charge and a very similar ion radius, Mn(II) with spin $S = 1/2$ can usually substitute for Mg(II) without loss of function. Some Mg(II)-dependent enzymes are indeed promiscuous and use either Mg(II) or Mn(II) depending on their environment. Like Gd(III), Mn(II) is a high-spin ion with comparatively narrow $m_S = -1/2 \leftrightarrow +1/2$ transitions, but it features the additional complication of ^{55}Mn hyperfine coupling ($I = 5/2$ with a coupling that is nearly isotropic with usual values of $a_{\mathrm{iso}} \approx 8$ mT. Substitution of Mn(II) for Mg(II) in nucleotide-binding domains of a helicase and of an ABC transporter and distance measurements between Mn(II) ions in these systems have been demonstrated [Wie+17a].

Lanthanide ions have coordination properties similar to Ca(II) and can sometimes substitute for this ion. Cobalt ions can substitute for iron ions in heme proteins without loss of the capacity for oxygen binding[Bow+97].

5.2 Choice of linker chemistry

Site-directed spin labeling (SDSL) requires that the label can be attached to the macromolecule very specifically, with a high yield, and under mild conditions where the protein is stable. Among all amino acids, cysteines are best suited as attachment sites, as they offer the unique chemistry of a thiol group and are rather rare, so that it is often possible to generate a functional cysless mutant as a staring point of an SDSL study. If a cysless mutant does not express or is not functional, it may still be possible to use a cysteine-based SDSL strategy. This requires that those native cysteines that cannot be removed are buried and do not label. In other cases, spin labeling may require the introduction of unnatural amino acids, a strategy that may work more or less well depending on the type of protein. In some cases, it may also be possible to label a cofactor, substrate or inhibitor instead of the macromolecule itself.

For RNA, thiol-reactive linkers can be applied to labeling thiouracil or to 5'-terminal thiophosphate groups. In this case, the thiol groups cannot be introduced biochemically into recombinant RNA. It is, however, possible to synthesize small pieces of thiouracil-modified RNA, spin label them, and use splinted T4 DNA ligation in order to build the required RNA construct [Dus+14].

Attachment of a label to a sidegroup or to a thiophosphate end group introduces a flexible linker between the location of the unpaired electron, where restraint information is obtained, and the backbone of the macromolecule for which restraint information is usually sought. This limits accuracy and precision of models generated from EPR-based restraints, but may also be necessary to avoid perturbation by a rigid label that does not fit into the structural context of the labeling site. In favorable cases, the problem can be avoided by the TOAC label (Fig. 5.1), which is an α amino acid with a spiro-linked six-membered ring nitroxide that is commercially available with protection groups suitable for solid-state peptide synthesis. The spiro link leads to a rather rigid attachment of the nitroxide to the backbone and the sidegroup is not larger than the one of aromatic amino acids. However, the spiro link also means that TOAC is not a usual amino acid with a single substituent at the Cα atom. It has a Ramachandran plot that strongly differs from the one of other amino acids and may thus perturb backbone conformation. Furthermore, TOAC is not easily introduced into large proteins. This label has therefore mostly been used for studies with peptides, to a large advantage with antibiotic peptides that contain the rare amino acid 2-Aminoisobutyric acid (Aib), which has two methyl substituents at the Cα atom. Like Aib, TOAC is strongly helix inducing.

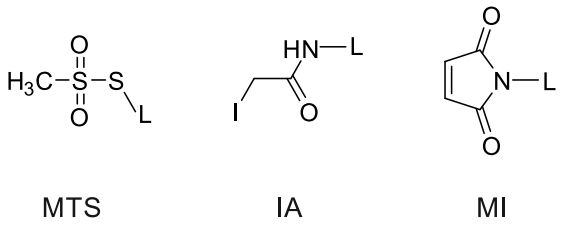


Figure 5.6: Structures of the label linkers methanethiosulfonate (MTS), iodoacetamide (IA), and maleimide (MI). L stands for the label.

5.2.1 Methanethiosulfonate linkers

The methanethiosulfonate (MTS) linker (Fig. 5.6) has a bulky methylsulfonate leaving group and reacts very specifically at ambient temperature or even at 4°C with thiol groups. Depending on steric restraints at the labeling site, a stochiometric amount or up to tenfold excess of the label is usually applied. The label is then attached via a disulfide bridge, which has the advantage that it can be easily removed under reducing conditions. However, this advantage is rarely used and the weak and reduction-sensitive disulfide bridge is usually a (mild) disadvantage. The combination of MTS with the Proxyl nitroxide (Fig. 5.1) is known under the abbreviation MTSSL (MTS spin label) or MTSL and is by far the most widely used spin label in protein studies.

5.2.2 Maleimide and Iodoacetamide linkers

The double bond of the Maleiimide group adds a thiol group and thus forms a stable thioether link. The reaction is highly specific to thiols in the pH range 6.5 to 7.5 and proceeds under mild conditions. At pH > 8.0 there is a substantial risk of labeling primary amine groups in lysine and arginine. MI labels are a bit more bulky and a bit more flexible than MTS labels.

The Iodoacetamide (IA) linker has a iodine leaving group and attaches rather selectively, but not as selectively as MTS, to thiol groups. It forms a stable thioether bond. The linker is slightly more flexible than the one of MTS. With the IA linker only a slight excess of reagent should be used at pH 8.3. Otherwise or in the absence of free thiol groups or if the thiol groups are not accessible, it is possible that other amino acids are labeled.

5.2.3 Linkers for unnatural amino acids

In general, labeling of unnatural amino acids leads to labeled side chains with a larger size than cysteine-based labeling. A moderately sized and moderately flexible linker can be introduced via *p*-acetyl-phenylalanine by ketoxime formation (Fig. 5.7a) [Fle+09]. The reaction requires tenfold excess of the label and proceeds at 37°C for 12-48 h at pH 4 in phosphate buffer. For a labeling reaction, these are relatively harsh conditions that exclude application to proteins that are only marginally stable *in vitro*.

Click reactions between an azide group and an alkyne offer a milder alternative for an attachment reaction, which is completely bioorthogonal and thus very selective. The azide group can be included in the spin label (Fig. 5.7b) or in the unnatural amino acid (Fig. 5.7c). Such click reactions are Cu(I)-catalyzed, i.e., they proceed under conditions where the nitroxide spin label is eventually recuced. This type of labeling thus requires careful optimization of the reaction time.

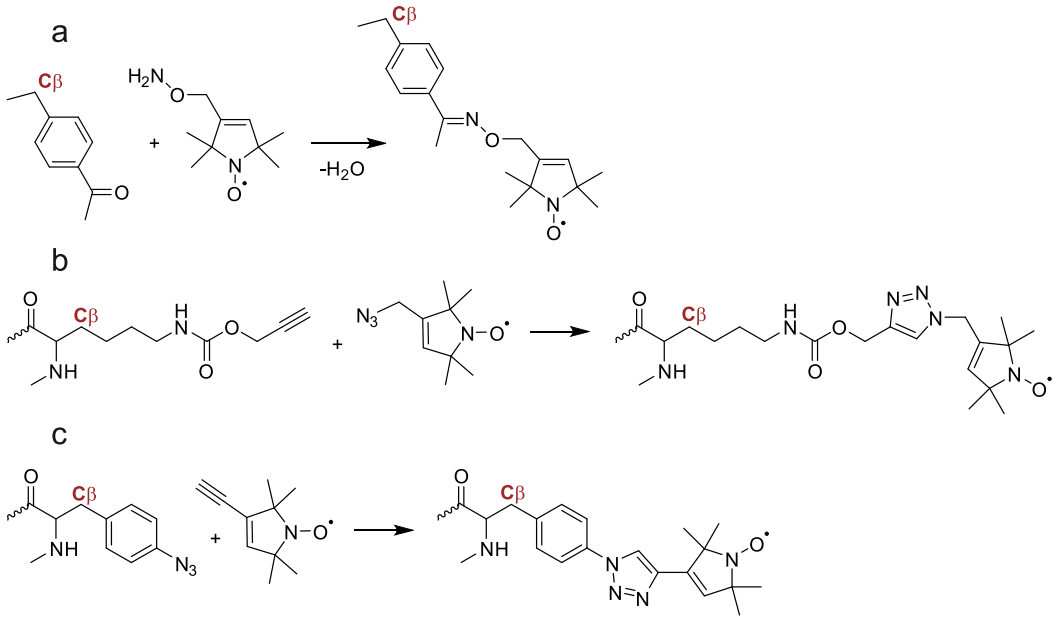


Figure 5.7: Site-directed spin labeling with unnatural amino acids. a) Labeling of *p*-acetyl-phenylalanine by ketoxime formation. [Fle+09]. b) Labeling of an alkyne-functionalized arginine analog by a Cu(I)-catalyzed click reaction with an azide-functionalized spin label [Kuc+17]. c) Labeling of an azide-functionalized phenylalanine analog by a Cu(I)-catalyzed click reaction with an alkyne-functionalized spin label [Kuc+17].

5.3 Choice of labeling sites

The main requirement on a labeling site is that introduction of the label is unlikely to perturb structure and function of the system. In general, some structure analysis technique or functional assay should be available for testing that such perturbations are indeed avoided or only minor. It is thus advisable to avoid labeling sites near potential interaction sites or on interaction surfaces as well as amino acids that are likely to be involved in salt bridges (Arg, Lys, Glu, Asp) or in metal binding (His). Amino acids with aromatic side chains (Phe, Tyr, Trp) or helix-breaking amino acids (Gly, Pro) are also often structure-determining and thus poor labeling targets. The most commonly used spin label, MTSSL, has a sidechain polarity similar to the one of Ile or Leu. Substitution of these amino acids, of Ala, Val, Met, Thr, Cys, and Ser is usually benign.

Further considerations depend on availability of secondary or tertiary structure information. Surface-exposed helical sites are a popular choice for distance measurements, as they combine a minor perturbation risk with an attachment in a rigidified section of the peptide chain. In studies where an atomistic structure of the system or parts of the system is available from previous x-ray diffraction, NMR, or cryo-EM work, the search for suitable labeling sites can be perform with an *in silico* spin labeling site scan.

5.3.1 *In silico* spin labeling site scans

Natural amino acid side chains and spin label side chains can adapt to steric requirements by selection of rotameric states. Each of the side chain torsion angles χ_i can usually take one of a few discrete values that correspond to energy minima (Fig. 5.8). By combining favorable torsion angles at each rotatable bond, the sidechain can often fit in restricted space. On the other hand, if much space is available, as is often the case for surface-exposed helical sites, several rotameric

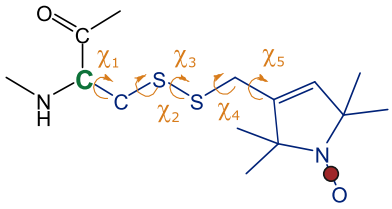


Figure 5.8: Flexibility of the MTSSL side chain by rotamer selection. Each of the torsion angles $\chi_1 \ldots \chi_5$ can adopt one of two or three distinct states. The backbone anchor point is the Cα atom (green). The location of the unpaired electron (red) is spatially distributed to to the existence of different rotameric states.

states can be simultaneously populated in thermal equilibrium.

A library of all expected rotameric states can be used to test whether an intended labeling site is well or not so well accessible to a spin label [PBJ11]. For this, the estimated free energies of the rotameric states of an isolated spin label are combined with the energies resulting from pairwise non-bonding interaction between label atoms and atoms of the labeled macromolecule. From the total free energy estimate, a partition function estimate is computed which is a measure for the number of rotameric states that are thermally populated. Partition function estimates well below unity indicate that the site is hard to label. Furthermore, the rotamer library approach provides an estimate of the conformation distribution of the spin label side chain and thus an estimate of the spatial distribution of the unpaired electron site with respect to the backbone of the macromolecule. Accuracy of the prediction is limited by the neglect of electrostatic interactions, of flexibility of the macromolecule backbone, and of slight changes in spin label bond angles upon attachment.

The rotamer library approach is sufficiently fast to test all potential labeling sites even in large macromolecular complexes by an *in silico* spin labeling site scan [PJ10]. Such site scans may turn out false negatives, i.e., labeling sites that are predicted to be too crowded may sometimes be labeled even without distortion. It is less likely that the site scans turn out false positives.

5.3.2 Choice of site pairs

For distance distribution measurements, site pairs need to be selected. Usually one aims at distances in the range between 20 and 50 Å or, if this is not feasible, between 15 and 80 Å. The measurement of even longer distances may require full deuteration of the system and may fail even then, in particular, for membrane proteins, where local concentrations may be too high for measuring very long distances even if bulk concentrations are close to the sensitivity limit of EPR distance measurements.

In cases where a structure in one state of the system is already known and the measurements are intended to reveal conformational changes in other states, the distance distributions in the known state can usually predicted fairly well, with about 3 to 5 Å uncertainty of the mean distance, by the rotamer library approach. If no structural information is known, the expected distance range can be roughly estimated from the conformation statistics of an unstructured peptide coil. Two sites in the same peptide chain that are separated by between 15 and 60 residues are likely to have a mean distance that falls into the range accessible by DEER measurements [Jes16].

Further considerations on site pair choice may arise from the modeling approach that is used in order to process the restraints. This is discussed in Chapter 6.

5.4 Information from EPR and DNP

5.4.1 Experimental techniques

Information on polarity of the environment and on sidegroup mobility as a measure of crowdedness of the environment can be obtained from analyzing CW EPR spectra. Obtaining further information requires specialized experiments.

Progressive power saturation

The influence of the environment on relaxation of the electron spin of a label can be assessed at ambient temperature by the progressive power saturation technique. At low microwave power, the CW EPR signal amplitude is proportional to the square root of power P_{mw} and thus to the amplitude B_1 of the microwave field. When power is increased, this behavior is observed as long as relaxation of the electron spins back to the spin ground state is much faster than excitation. Once excitation can efficiently compete with relaxation, populations of the ground and excited spin state change significantly and polarization of the transition decreases. In this situation, signal amplitude decreases less than linearly with the amplitude of the microwave field. Eventually, excitation outcompetes relaxation and the signal amplitude begins to decrease. If excitation dominates by far, the transition is saturated, i.e., the population of the spin ground and excited states is equalized. In this limit, the signal vanishes. The observation of the signal amplitude as a function of microwave power thus provides quantitative information on the product of the relaxation times T_1 and T_2.

In practice, one measures CW EPR spectra at decreasing microwave attenuation and thus increasing power, where a 3 dB decrease of attenuation corresponds to a doubling signal amplitude I_0 (see Figure 5.3) in the *linear regime* where $I_0 \propto B_1$ and $I_0 \propto P_{\mathrm{mw}}^{1/2}$. I_0 is plotted against $P_{\mathrm{mw}}^{1/2}$ and fitted by the equation

$$I_0\left(P_{\mathrm{mw}}\right) = \frac{A\sqrt{P_{\mathrm{mw}}}}{\left[1 + \left(2^{1/\epsilon} - 1\right) P_{\mathrm{mw}}/P_{1/2}\right]^{\epsilon}} , \tag{5.1}$$

where the power at half saturation $P_{1/2}$ corresponds to the situation where the signal is half as large as it would have been at the same power for infinitely fast relaxation. The inhomogeneity parameter ϵ takes the value 1.5 in the homogeneous limit and 0.5 in the inhomogeneous limit and is treated as a fit parameter.

The experiment is done in the absence and presence of a paramagnetic quencher, and for consistency of results with different microwave resonators, a reference measurement is performed on the reference sample DPPH (dilute diphenylpicrylhydrazyl powder in KCl). An accessibility parameter can then be defined as

$$\Pi = \frac{\Delta P_{1/2}/\delta_0}{\left[\Delta P_{1/2}/\delta_0\right]_{\mathrm{ref}}} . \tag{5.2}$$

where the denominator is obtained from the measurements on the reference substance and δ_0 is the linewidth in the linear regime, i.e. at low microwave power.

For metal complex quenchers, such as NiEDDA, the experiment is performed with the spin-labeled system in pure buffer and in buffer containing the quencher. For oxygen accessibility experiments, it is most convenient to use a capillary of the gas-permeable plastic TPX and perform the experiment once in a stream of nitrogen and once in an air stream or stream of pure oxygen. Example data of such a measurement for a membrane protein residue in the interior of a detergent micelle are shown in Fig. 5.9.

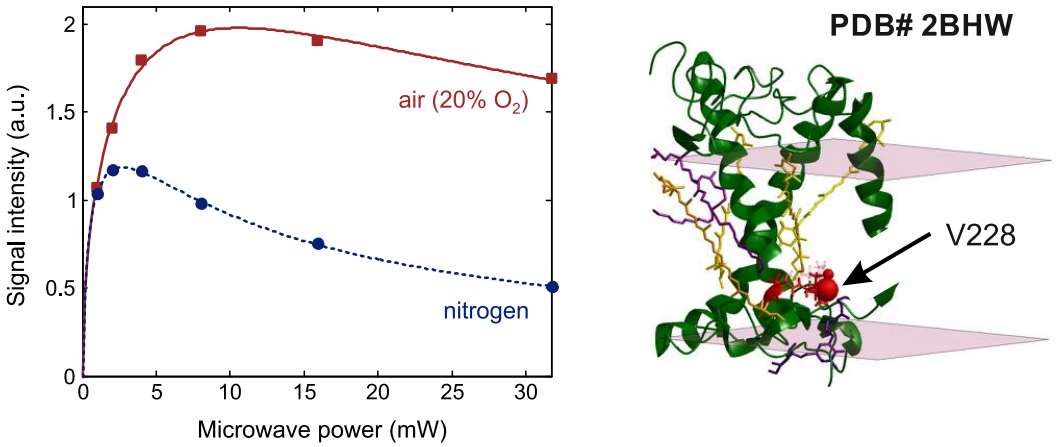


Figure 5.9: Progressive power saturation measurement on residue V228 in major plant light-harvesting complex LHCII mutated to cysteine and spin-labeled with IA-Proxyl in the absence and presence of oxygen. Due to the localization of this residue on the protein surface and exposure to the alkyl chain region of detergent micelle, a strong desaturation is observed in the presence of oxygen. The ribbon cartoon shows the putative localization of the residue in a lipid bilayer.

Overhauser DNP

Electron spin transitions have a factor of about 660 higher energy than proton spin transitions. At temperatures and magnetic fields, where both energies are much smaller than the thermal energy $k_{\mathrm{B}}T$, electron spin transitions are by the same factor stronger polarized than proton transitions. At ambient temperature, this applies throughout the magnetic field range achievable with magnetic resonance spectrometers. Therefore, proton spin transitions can be enhanced by polarisation transfer from electron spin transition. This approach is called *dynamic nuclear polarization*.

In liquid solution, DNP can arise on saturation of electron spin transitions due to a difference in relaxation rate on the two forbidden electron-nuclear transitions where the electron and nuclear spin flip in parallel or antiparallel. This DNP mechanism is called Overhauser DNP. Its efficiency depends on the dominant coupling mechanism between the electron and nuclear spin and on their relative spatial dynamics, which must be on an appropriate time scale in order to create the difference between relaxation rates. It turns out that at ambient temperature X-band frequencies and lower, polarization transfer from spin labels to protons in water molecules is highly efficient. After transfer of the polarization from the spin label to protons of nearby water molecules, the polarization needs to spread throughout the sample to the bulk water. Simultaneously, nuclear spin polarization is lost by longitudinal relaxation of the water protons. This competition between spread and loss of polarization makes Overhauser DNP sensitive to the water accessibility of a spin label.

The experiment requires an EPR resonator fitted with a radiofrequency coil for detection of the ^{1}H NMR signal of water and a low-frequency NMR console suitable for an FT NMR experiment on the water protons at a frequency of about 14 MHz. This experiment is then performed in the absence of microwave irradiation and in the presence of such irradiation with increasing power. The change of the water proton NMR signal as a function of microwave power carries information on the water accessibility of the label.

An example is shown for residues BtuA 141 and BtuB 141 near the cytoplasmic gate of the vitamin B12 transporter of *E. coli* (Fig. 5.10). According to crystal structures, one expects an increase of water accessibility near the cytoplasmic gate upon binding of the substrate-binding protein BtuF to the transporter. The experiment addressed the question whether such a change

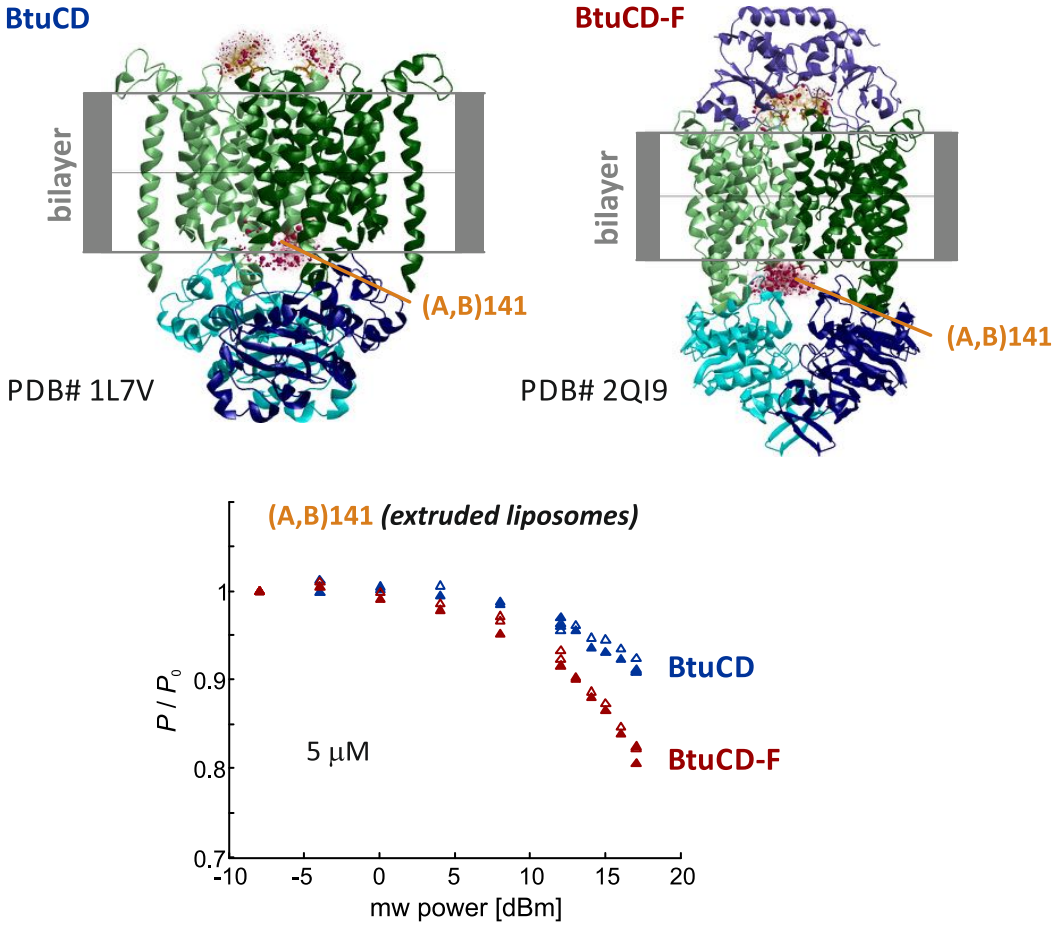


Figure 5.10: Overhauser DNP measurement of a change in water accessibility upon binding of the substrate-binding protein BtuF to the vitamin B12 transporter transmembrane section BtuCD. The change is probed near the cytoplasmic gate of the transporter by the spin-labeled residue 141 in the nucleotide-binding domains BtuAB. In the absence of BtuF, the ODNP effect is smaller than in its presence.

is indeed observable for the transporter reconstituted into liposomes. Because of the low concentration of spin-labeled protein in the sample, the ODNP effect is small, and because the enhancement is negative, a loss of ^{1}H NMR signal is observed. This loss is about 8% in the absence and about 20% in the presence of BtuF, consistent with the expected increase in water accessibility of these residues.

Pulsed dipolar spectroscopy

Distance information from spin labels is best obtained in a range where the label-to-label distance is much larger than the spatial distribution of the label position and where the exchange interaction between labels can be safely neglected. This applies roughly at mean distances larger than 15 Å and one will often aim at mean distances larger than 20 Å. In this range, pulsed dipolar spectroscopy techniques, such as DEER, double-quantum EPR (DQ-EPR), and the single frequency technique for refocusing (SIFTER) dipolar couplings, provide more accurate and better resolved data than analysis of the dipolar broadening in CW EPR spectra.

Among these techniques, four-pulse DEER [Jes12b] provides data for which both the intramolecular and intermolecular contributions can be understood from first principles. The

signal takes the form

$$V(t)/V(0) = F(t)B(t)\,, \tag{5.3}$$

where t is the dipolar evolution time, the form factor F due to an intramolecular contribution with fixed distance r is given by

$$F\left(t,r\right) = F\left(0\right)\int_0^1 1-\lambda\left(z\right)\left\{1-\cos\left[\omega_\perp(r)\left(3z^2-1\right)t\right]\right\}\mathrm{d}z\,, \tag{5.4}$$

with the inversion efficiency $\lambda(z)$ and the dipole-dipole coupling

$$\omega_\perp = \frac{1}{r^3}\frac{g_1 g_2 \mu_0 \mu_\mathrm{B}^2}{4\pi\hbar}\,. \tag{5.5}$$

The parameter $z = \cos(\theta)$ accounts for the dependence of the dipole-dipole coupling on the angle θ between the spin-spin vector and the magnetic field. For flexible spin labels, it is usually a good approximation to assume that orientation selection is absent, i.e. all orientations z contribute with the same inversion efficiency λ to the signal.

In the absence of excluded volume effects, the background factor $B(t)$ due to intermolecular contributions takes the form

$$B(t) = \exp\left(-(kt)^{D/3}\right)\,, \tag{5.6}$$

where the dimensionality D of the homogeneous distribution of molecules is $D = 3$ for soluble proteins and detergent-solubilized membrane proteins and lies between 2 and 3 for membrane proteins reconstituted into unilamellar liposomes. Excluded volume effects occur because the protein molecules are not homogeneously distributed point particles. The background function given by Eq. (5.6) is still a good approximation when considering such effects, but D may be somewhat lower than expected.

For single-frequency techniques, such as DQ-EPR and SIFTER, the factorization into a form factor and a background contribution implied by Eq. (eq:factorization) cannot be taken for granted, and if it is taken as an approximation, $B(t)$ cannot be derived from first principles even in the absence of excluded volume effects. Such techniques can still provide high-quality data if the concentration is sufficiently low in order to allow for empirical background correction. Multi-pulse DEER experiments, such as five-pulse DEER [BGF13], may provide data that have additional contributions from other coherence transfer pathways. In such a case, the experimental scheme needs to be adapted and a signal pre-processing needs to be introduced in order to recover form factor data that has the form of Eq. (5.4). For four-pulse DEER, $F(t)$ is obtained by fitting $B(t)$ in a time range after at least the first half wave of the dipolar oscillation with frequency $\omega_\perp$, dividing the normalized data $V(t)/V(0)$ by the estimate of $B(t)$, and renormalization. Subsequently, the constant contribution $1-\lambda$ is subtracted and the data is again normalized to give the dipolar evolution function

$$D(t) = \int_{r_\mathrm{min}}^{r_\mathrm{max}} P(r)\int_0^1 \cos\left[\omega_\perp(r)\left(3z^2-1\right)t\right]\mathrm{d}z\mathrm{d}r\,, \tag{5.7}$$

where we have introduced a normalized distance distribution $P(r)$ defined in an accessible distance range from r_min to r_max with $\int_{r_\mathrm{min}}^{r_\mathrm{max}} P(r)\mathrm{d}r = 1$.

An experiment for distance measurements is the better the closer its output data approximate $D(t)$. Among two experiments that have a similar data fidelity in that sense, the better one provides data for a longer time range $t = 0\ldots t_\mathrm{max}$ at given signal-to-noise ratio. In many cases, four-pulse DEER performed in the low-temperature limit of T_2 on a high-power Q-band EPR spectrometer and on samples that are deuterated as far as possible is the best distance measurement experiment in that sense.

5.4.2 Secondary structure

In a spin labeling site scan, the polarity parameter A_{zz}, the mobility parameter δ, and, for membrane proteins, the accessibility parameters Π_{O_2} and Π_{EDDA} are measured for a stretch of consecutive amino acid residues in a peptide chain. If this chain exhibits secondary structure, some or all of these parameters exhibit a periodic dependence on residue number. For instance, for an α helix one expects a perodicity of 3.6 residues, which will be observed for an amphiphatic helix with one water-exposed side in A_{zz} and Π_{EDDA}. For a helix with one surface-exposed and one buried side one expects periodicity in δ, A_{zz} and Π_{EDDA} if the surface is water exposed and in δ and Π_{O_2} if it is lipid exposed.

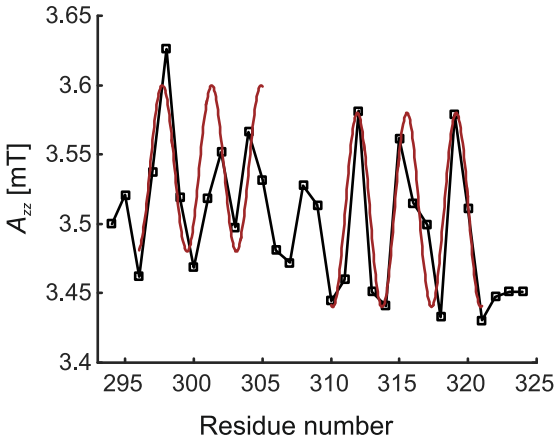


Figure 5.11: Dependence of the polarity parameter A_{zz} on residue number in a spin labeling site scan on the external loop eL4 of the Na^+/proline symporter PutP of *Escherichia coli*. The red oscillatory patterns are guides to the eyes assuming a periodicity of the polarity of one turn per 3.6 residues for residues 296-305 and 310-321 as expected for an α helix. Data taken from [Rab+14]

As an example, polarity data from a site scan on the periplasmic loop eL4 of the sodium/proline symporter PutP of *Escherichia coli* is shown in Fig. 5.11 [Rab+14]. Two α-helical segments eL4a (residues 296-306) and eL4b (312-321) could be discerned by conisdering not only the polarity information, but also variation of the mobility parameter δ and of Π_{O_2} and Π_{EDDA} within the loop. Use of such secondary structure information for modeling is discussed in Chapter 6.

5.4.3 Lipid bilayer immersion

Lipid bilayers feature concentration profiles for water and oxygen along the bilayer normal. Whereas the water concentration is high outside the membrane, moderate in the headgroup region, and low near the bilayer center, the oxygen concentration is high in the center, moderate in the headgroup region, and comparatively low in water outside the bilayer. Measurements of water accessibility and oxygen accessibility can thus provide semi-quantitative estimates of bilayer immersion depth of spin-labeled sites in membrane protein. The data has to be interpreted with caution, since residues can also be buried or can face a channel or open gate of a transporter. Buried residues can be recognized, as both their water accessibility and their oxygen accessibility is low.

Fig. 5.12 illustrates the concept with ODNP data on water accessibility of residues in the bilayer-inserting model peptide WALP23. Stronger (negative) DNP enhancements are observed in the headgroup region (residues 2-5 and 22) than near the center of the bilayer (residue 11). The data is in surprisingly good agreement with accessibilities computed with a rotamer library

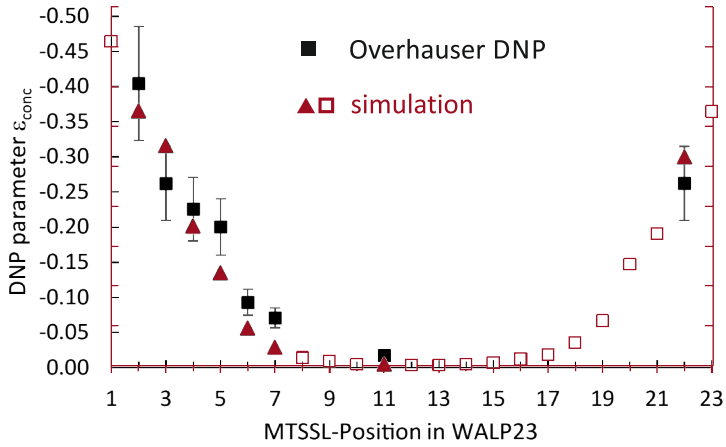


Figure 5.12: Dependence of Overhauser DNP enhancement on labeling site in the membrane-inserting peptide WALP23 inserted into the lipid bilayer of unilamellar vesicles. Data taken from [Seg+16].

model and the water distribution model [Kuc+08] for a DOPC bilayer.

5.4.4 Distance distributions and spin counting

Distance distributions

The dipolar evolution function $D(t)$ given in Eq. (5.7) can be inverted to the distance distribution $P(r)$. From a mathematical point of view, this inversion is an ill-posed problem where small deviations of the input data $D(t)$ from the ideal data can cause very large deviations of the output data $P(r)$. This ill-posed problem is stabilized by Tikhonov regularization. Instead of just minimizing the mean-square deviation $\rho = ||D(t) - S(t)||^2$ of the simulated data $S(t)$ from the experimental data $D(t)$, Tikhonov regularization minimizes

$$G_\alpha = \rho + \alpha\eta \,, \tag{5.8}$$

where $\eta = \left|\left|\mathrm{d}^2P(r)/\mathrm{r}^2\right|\right|$ is the square norm of the second derivative of the distance distribution, which is a measure of the roughness of this distribution.

The *regularization parameter* α determines the weight that is given to the roughness criterion. Since the result depends on this parameter, it is important to set it properly. In general, α needs to be large for a fast decaying $D(t)$ that corresponds to a broad distance distribution and for poor signal-to-noise ratio. A good choice for α can be inferred from a parametric plot of $\log\eta$ versus $\log\rho$ as a function of α. For data of good or moderate quality, this plot is L-shaped and it is therefore called an L-curve. The best compromise between smoothing of $P(r)$ to avoid artificial peak splitting on the one hand and excessive artificial broadening of $P(r)$ on the other hand corresponds to the corner of the L-curve. Furthermore, $P(r)$ can be stabilized by the non-negativity criterion $P(r) \geq 0$ that strictly applies since $P(r)$ is a probability density distribution.

Tikhonov regularization is a relatively robust way of converting $D(t)$ to a distance distribution $P(r)$ without assuming a specific model. However, regularization can deal only poorly with distributions that contain both very narrow and very broad peaks. The compromise regularization parameter α found at the corner of the L-curve for such cases causes significant broadening of the narrow peaks while the broad peaks are still artificially split. This behavior needs to be considered in data interpretation. If a good model for the distance distribution with a small number of parameters can be formulated, such a model may better stabilize the problem. Models

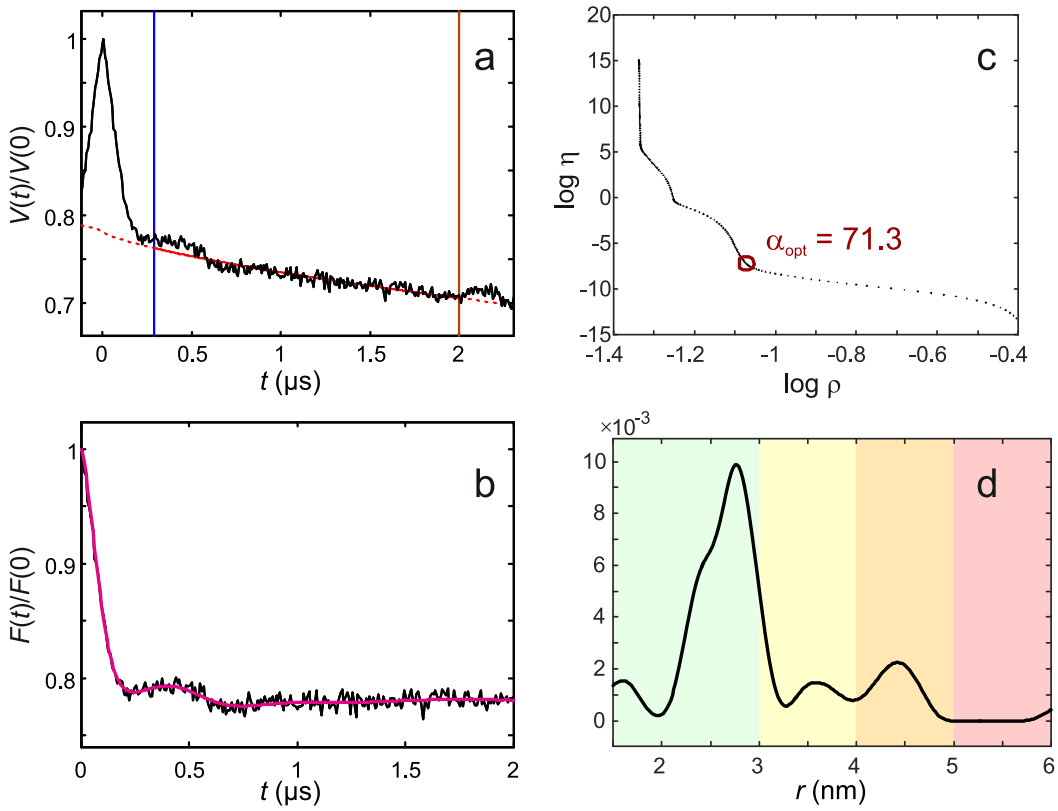


Figure 5.13: Conversion of DEER data to a distance distribution by Tikhonov regularization with the non-negativity restraint $P(r) \geq 0$. The data was obtained from 20 μL of a preparation of about 30 μM of the MTSSL labeled double mutant 87R1/126R1 of the pro-apoptotic protein Bax in mitochondria-like vesicles. It was measured at 50 K with a high-power Q-band pulse EPR spectrometer and a probehead for oversized samples. a) Original DEER data. The background function $B(t)$ (red) is fitted in a range between the first half wave of the dipolar decay (blue vertical line) and the appearance of and end artefact in the data (brown vertical line). The end artefact is cut off. b) Form factor $F(t)$ obtained after background correction (black) and its fit (magenta) by the result of Tikhonov regularization. (c) L-curve for the form factor shown in (b) obtained at high resolution in α but without the non-negativity constraint. Such L-curves may have several corners. Usually, but not always the last corner is the "right one". (d) Distance distribution with color coding for reliability ranges. In the green range, the shape of $P(r)$ can be trusted, in the yellow range its width but not it shape, in the orange range the mean distance, but neither its shape nor its width, and the red range only contains information on the existence of contributions at long distances that cannot be further quantified. Data taken from [Ble+14a].

with a small number of Gaussian peaks are a popular choice, but should only be applied with rigorous statistical testing or in conjunction with Tikhonov analysis, as they may suppress real features of $P(r)$ such as peak asymmetry.

Data processing for a benign case is illustrated in Fig. 5.13. As the distances are relatively short, data could be obtained at low concentration with a sufficiently long maximum dipolar evolution time t to discern the background. Although the L-curve exhibits several corners, inspection of the two corners at low α (left side of the curve) showed that these regularization parameters led to unrealistic splittings into very narrow peaks. The distance distribution mainly falls into a range where its shape, width, and mean distance can be considered as reliable. The data was processed with DeerAnalysis 2016 `http://www.epr.ethz.ch/software.html`.

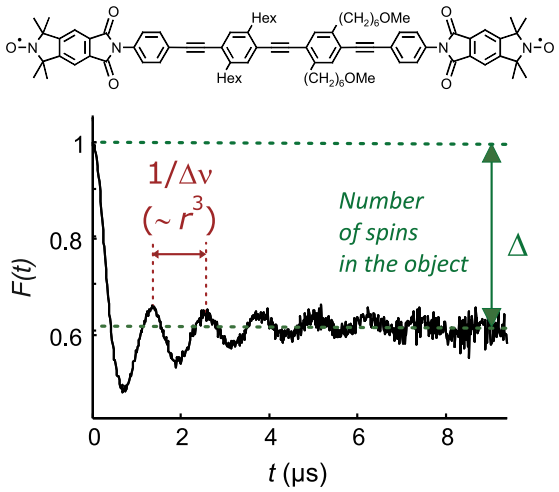


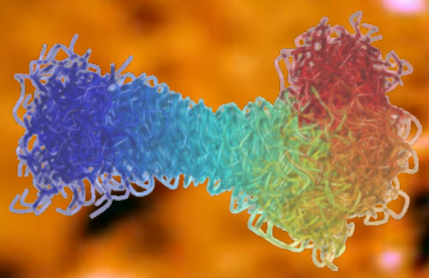
Figure 5.14: DEER data for a biradical with 100% labeling efficiency ($f = 1$). The modulation depth Δ coincides with the inversion efficiency λ. For this case with a very narrow distance distribution, the mean distance can be directly computed from the period of the dipolar oscillation.

Spin counting

In the case where exactly two spins reside in the same molecule, i.e., when the form factor after background correction corresponds to a spin pair, the modulation depth Δ of the form factor (Fig. 5.14) coincides with the inversion efficiency λ in Eq. (5.4). For systems with any number $n > 0$ of spin labels in the molecule, the unmodulated fraction $1 - \Delta$ scales with $(1 - \lambda)^{n-1}$, as long as the labeling efficiency f at each site is unity [Bod+07]. In a homooligomer, where all spin labels are attached to equivalent sites in the individual protomers, one can assume that the labeling efficiency f is the same at all sites. In this case, $1 - \Delta$ scales with $(1 - f\lambda)^{n-1}$. If the background correction is reliable, so that a good estimate of Δ can be obtained and if a good estimate of f is known, the number of protomers in the oligomer can be obtained as

$$n = 1 + \frac{\ln 1 - \Delta}{\ln (1 - f\lambda)} \,. \tag{5.9}$$

Such spin counting can also be applied to assess binding equilibria of weak protein complexes [Hil+05]. Because of the logarithms on the right-hand side of Eq. (5.9), it becomes more and more difficult with increasing n to determine n with integer resolution. In practice, it is easy to distinguish two-spin from three-spin systems and possible to distinguish three-spin from four-spin systems, whereas determination of $n > 4$ would require data of very high quality, utmost care in background correction, and high accuracy in determining labeling efficiency f.



6 — Structure modeling with EPR restraints

Most structure modeling programs do not provide specific functionality for EPR-based restraints. Although it is possible to incorporate such restraints by workarounds [Bib+17; Dus+14], this is not necessarily optimal, since distance distribution restraints provide information on the width of a conformation ensemble that is foreign to the architecture of modeling programs that are geared to solving a single atomistic structure. This Chapter explains functionality of the modeling program MMM (Multiscale Modeling of Macromolecules) that takes advantage of the information content of EPR restraints on the one hand and also considers their uncertainty on the other hand.

6.1 Hybrid structure determination

Traditional modeling software aims to compute a structure where each atom position is exactly defined based on restraints from a single technique. This strategy is based on the Anfinsen dogma, which states that the native structure of a protein is fully determined by the amino acid sequence, as this sequence encodes a single conformation with minimum free energy. Deviations from Anfinsen's dogma, which are often related to missing restraints, are treated in different ways. For x-ray diffraction, electron density of disordered domains may be missing completely, in which case this part of the structure is not specified. Weak disorder can be accounted for by specifying a small number of alternate locations for an atom. The PDB format for macromolecular structure was originally designed with this concept in mind.

In solution NMR structure determination, lack of restraints for loop regions or termini is more often observed than in structure determination by x-ray diffraction. This is not usually due to an intrinsic weakness of the NMR approach, but rather due to intrinsic flexibility of these domains in solution that is artificially reduced during crystallization by formation of crystal contacts. NMR structure modeling programs therefore opt for computing a model ensemble. The PDB format was extended to represent such information by a set models that are each specified as a fully determined atomistic structure.

Since the number and precision of the restraints is usually insufficient to fully specify an atomistic structure fro x-ray diffraction or NMR data, the programs augment experimental information with general information on protein or nucleic acid structure that can be safely assumed. Such information are bond lengths, bond angles, Ramachandran-forbidden torsion angles in the backbone, and the absence of clashes.

Hybrid structure determination is applied if restraints from a single experimental technique

are insufficient for specifying a model even on the level of a small ensemble of structures with atomistic resolution. This lack of restraints may have two principal reasons. First, it may be genuine in the sense that the missing restraints would be, in principle, accessible but could not be determined because of experimental problems (often resolution or assignment problems) or because the required effort would have been too large. Second, the structure or parts of it may be intrinsically disordered to an extent that invalidates the idea of restraints at atomistic resolution. In the second case, the information is in the width of the ensemble of conformations, in inaccessible regions of the total space of conformations, and in structural correlations that are persistent throughout the ensemble or throughout subensembles. As of now, this problem is poorly understood even on the level of representing such information once it has been obtained. As a result, it is not yet clear what the best ways are for obtaining the required information and for translating it into models. Such systems are therefore still modeled by large ensembles of conformations with each of them being specified at atomistic resolution.

Whatever the reason for the lack of restraints is, hybrid structure determination tries to solve it by combining information from different experimental techniques. If the experimental restraints still do not suffice, they may be augmented by homology information or molecular force field energy terms. Such an approach blurs the border between experimentally specified and theoretically predicted structure. It is useful, but the models should be interpreted with more caution than models that are sufficiently restrained by experimental data.

Some of the approaches discussed in this Chapter and implemented in MMM include potentially unsafe information that does not stem from experiments. This applies to restraint-augmented homology modeling (Section 6.2) and to modeling of large-scale conformational change on the basis of normal-mode analysis of elastic network models (Section 6.3). The other approaches are puristic in the sense that they assume only safe information, i.e. canonical bond lengths, bond angles, Ramachandran-forbidden torsion angles, and the absence of clashes.

6.1.1 Specifics of EPR restraints

Restraints obtained by EPR techniques may relate to intrinsic cofactors or to artificially introduced spin labels. Cofactors are usually located in domains whose structure is indeed defined with atomistic or near-atomistic resolution. EPR restraints between two cofactors or hyperfine restraints between a cofactor and the nuclear spin of an atom in the biological macromolecule are easily included into the restraint list for traditional structure determination programs. This also applies to the substitution of diamagnetic metal ions by paramagnetic ones.

Uncertainty and inaccuracy of spin-label based restraints

For spin labels in the usual sense, the label itself must be included into the modeling. This is complicated by the fact that linkers are usually flexible, so that the position of the electron spin is blurred by conformation disorder of the spin label side chain. In early stages of modeling it is possible to assume an approximate mean position of the electron spin with respect to the macromolecular backbone. During refinement or in a quality check of the model, conformation distribution of the label needs to be accounted for.

Prediction of the distribution of conformations of the label itself introduces inaccuracy. Mean distances between two MTS spin labels can be predicted with a standard deviation of about 2 Å if the backbone structure and other sidechain conformations are known at high resolution and of about 3 Å if they are known with moderate resolution [Jes13]. Fig. 6.1 illustrates this uncertainty. The position of the electron spin of MTSSL attached to residue 131 of T4 Lysozyme was determined by measuring five distances to MTSSL attached to residues 68, 72, 86, 89, and 109 using the GPS-like approach explained in Section 6.4.1. The uncertainty of the position that results from experimental uncertainty of the distances is visualized by the red surface which includes the space where the electron spin is located with 50% probability if the positions of

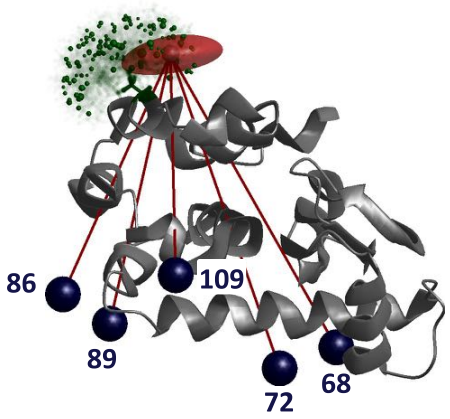


Figure 6.1: Uncertainty and inaccuracy of determining a spin label position with respect to a macromolecular backbone. The location of MTSSL attached to residue 131 of T4 Lysozyme (131R1) was determined by five DEER measurements of distance distributions between this sites and MTSSL attached to one of the residues 68, 72, 86, 89, and 109. The red surface includes locations where the electron spin of 131R1 is located with 50% probability. Its extension denotes experimental uncertainty. The green spheres denote the electron spin location of predicted rotamers of 131R1 with their volume being proportional to rotamer population. Displacement of the red surface from the green spheres visualizes inaccuracy of the rotamer prediction.

the electron spins at sites 68, 72, 86, 89, and 109 are assumed to be known exactly. The small green spheres visualize locations of the electron spin at site 131 predicted by the rotamer library approach with sphere volume being proportional to the predicted population of the rotamer. The displacement of the red surface from the swarm of green spheres corresponds to the inaccuracy of rotamer library modeling for this particular case. Such inaccuracy adds to the uncertainty of models derived from EPR restraints and must be considered in their interpretation.

Sparsity of spin-label based restraints

Another limitation of EPR restraints is that they are usually sparse. Sparsity means that the number of restraints is much too small to fully specify the model and is thus related to an uncertainty that does not stem from intrinsic flexibility of the macromolecule. EPR restraints are sparse due to the large effort in generating each of them, which involves mutation, protein expression, spin labeling and measurement. Sparsity combined with inaccuracy on a 2-3 Å scale precludes full specification of atomistic structures from only EPR restraints and thus requires hybrid modeling approaches based on auxiliary restraints.

6.1.2 Types of auxiliary information

Atomistically resolved domain structures

Many biomolecular systems feature rigid domains that conform to the Anfinsen dogma. Their structure can be specified at atomistic resolution and can often be determined by x-ray diffraction, high-resolution or solid-state NMR, or cryo-electron microscopy. Often, these rigid domains constitute the major part of the system. Even if a domain is not strictly rigid, an atomistically resolved structure of it still provides information on one of its conformations that is most stable under certain conditions or after certain modifications. The atom coordinates of the structures of such domains can be used as auxiliary restraints for model building with EPR restraints. Often, the domains are treated as strictly rigid, but it is also possible to allow for well-defined conformational change (Section 6.3).

Secondary structure propensities

In some cases, in particular for intrinsically disordered domains, NMR spectroscopy can provide secondary structure propensities of some residues. A secondary structure propensity specifies the probability that this residue is found in an α-helical or β-strand context in an ensemble of conformations. Secondary structure propensities can be used to restrain flexible domains (Section 6.6).

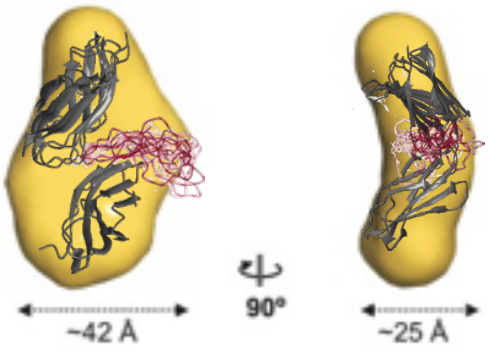


Figure 6.2: Molecular shape of a pair of FnIII domains of integrin $\alpha 6\beta 4$ computed from a SAXS curve (yellow) and model based on x-ray structures of the individual domains, 15 DEER distance restraints, and the SAXS curve (ribbon model, individual domains in grey and linker in crimson). Based on data from [Alo+15]

Small-angle scattering curves

The global shape of a macromolecule or macromolecular complex determines its scattering curve in small-angle x-ray scattering (SAXS) or small-angle neutron scattering (SANS). In the latter case, selective deuteration or contrast matching can be used to also determine the shape of parts of a complex. Macromolecular shapes obtained by SAXS and SANS are of rather low resolution and their computation from the scattering curves is an inverse problem that is ill-posed involves some uncertainty. The direct problem of computing the scattering curves from a given model is not ill-posed. Therefore, the fit quality of scattering curves can be used to distinguish between poor and good models. Scattering curve processing and fitting is best performed with the ATSAS package [Pet+12]. Fig. 6.2 shows the superposition of a model derived for a pair of FnIII domains of integrin $\alpha 6\beta 4$ from x-ray structures of the two individual domains, 15 DEER distance restraints, and the SAXS curve [Alo+15] with the molecular shape computed from the SAXS curve only. The distribution of the linker conformation (crimson) is due to lack of restraints rather than due to intrinsic flexibility.

Cross-linking restraints

Residues that are in close spatial proximity can be crosslinked, the macromolecule can then partially be digested using trypsin, and the mixture can be analyzed by mass spectrometry (MS). This technique can reveal which residues are crosslinked [Lei+10; Rap11]. Different versions exist that can work either with wild-type protein or with cysteine double mutants. The latter technique provides fewer but more precise restraints, whereas the former one provides a larger number of rough restraints (proximities below 20-30 Å) sometimes without the effort of producing recombinant protein or at least without the effort of generating mutants.

Cross-linking may not provide all expected proximities (false negatives) [Lei+10] and may also produce false positives if a protein or protein complex is not sufficiently stable. For systems

with intrinsic disorder, the total set of cross-linking restraints reveals proximities that exist in any of the conformations, but not necessarily in all of them. Therefore, the total set of cross-linking restraints may be inconsistent when interpreted in terms of a single structural model rather than an ensemble. At least in the early stages of modeling, cross-linking restraints are therefore best treated as uncertain by requiring that a certain fraction of them is fulfilled in each single conformation while the others may be violated.

6.2 Restraint-augmented homology modeling

Proteins with more than about 20% sequence homology often have similar tertiary structures. Whether or not this is the case can be tested with a relatively small number of experimental restraints, in particular, with long-range distance distribution restraints. Such restraints can also be incorporated during homology (or compartive) modeling with the MODELLER software that encodes the homology information also in terms of spatial restraints [FS03]. MODELLER also allows for specifying additional secondary structure restraints, as they can be obtained from spin-labeling site scans (see Section 5.4.2). MMM runs MODELLER via Python scripts that are prepared in MMM from sequence alignment information and experimental restraints.

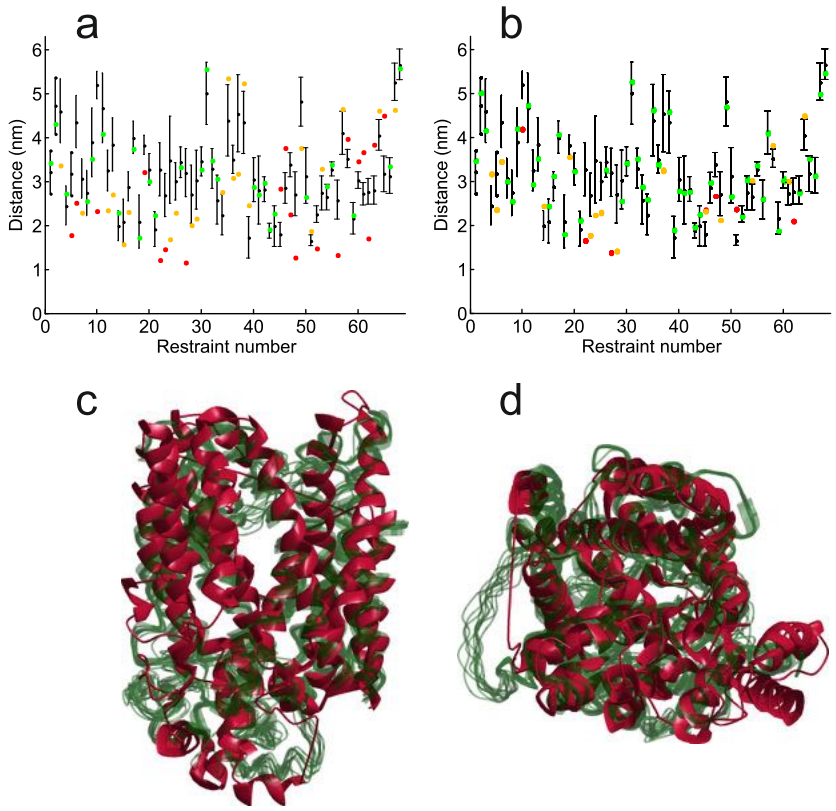


Figure 6.3: Homology modeling of the Na^+/proline symporter PutP based on the x-ray structure PDB# 2XQ2 of the Na^+/glucose symporter vSGLT and 68 DEER distance restraints. (a) Restraint matching by aligned residues in vSGLT. The experimental mean distances are visualized by black circles and the width of their distributions by black vertical lines. Green circles indicate matched restraints, orange circles slightly mismatched restraints, and red circles strongly mismatched restraints. (b) Restraint matching by the homology model (10 best scoring models). (c) View on the vSGLT structure (crimson) and the homology model (transparent green) parallel to the assumed lipid bilayer orientation. (d) View perpendicular to the assumed bilayer orientation. DEER data from a collaboration with Heinrich Jung and Daniel Hilger.

Fig. 6.3 shows homology modeling of the Na^+/proline symporter PutP of *Escherichia coli* based on the crystal structure of the Na^+/glucose symporter vSGLT of *Vibrio parahaemolyticus* from the same family of sodium/solute symporters (PDB# 2XQ2). Distances for simulated spin-label pairs of aligned residues in the vSGLT structure are in reasonable, but not good agreement with distances measured for PutP (panel a). Agreement improves significantly in a homology model that takes into account these restraints, although some of the restraints are still not fulfilled (red points) or marginally not fulfilled (orange points). The same approach with a smaller number of DEER restraints and additional restraints on secondary structure from a CW EPR spin-labeling site scan was applied to modeling of the external eL4 loop of PutP [Rab+14].

6.3 Large-scale conformational change

Many proteins, in particular some enzymes and transporters, undergo large-scale conformational change between different states. Quite often, structures at atomistic resolution can be obtained for one or a few of these states, but not for all of them. Since EPR spectroscopy does not require crystallization, distance distribution restraints can usually be obtained for all states. For membrane proteins it is a further advantage that EPR spectroscopy is compatible with reconstitution into liposomes, so that the restraints can be gathered in an environment that is closer to a cell membrane that a protein crystal or detergent-solubilized protein.

Modeling of the states for which no crystal structure exists can then be based on driving the conformational change *in silico* with forces that result from the distance restraints. As the set of distance restraints is sparse, the trajectory of conformational change must be further restrained by other information. Such information can be derived from the known structure of the initial state by an elastic network model.

6.3.1 Elastic network models

Elastic network models (ENM) are simple bead-and-spring models of the protein backbone. On residue level, the beads are located at the Cα atom positions and the force constant of the Hookean springs depends on distance. It has been demonstrated that such models provide a good approximation to the normal modes of backbone deformation [Bah+10]. Usually, dependence of the force constants γ_{ij} on Cα-Cα distance r_{ij} is defined by a power scaling r_{ij}^{-n}. It is advantageous to assign a different force constant to direct neighbors ($i, i+1$ pairs, where i is the residue index), and next neighbors ($i, i+2$ pairs) in the peptide chain and a cutoff distance above which force constants are set to zero.

6.3.2 Zheng-Brooks approach

Zheng and Brooks developed an approach for modeling conformational transitions by an ENM that implements Cα-Cα distance restraints as forces that deform the network [06]. This approach was later adapted for restraints between spin labels [Jes12a] and optimized by introducing thermal steps along a combination of normal modes and by parametrization of the ENM for this application [Jes12c]. The approach works well for simple hinge motion, but is somewhat deficient for more complex motion. The approach has been applied, among else, to the conformational change observed in a cyclic nucleotide-gated (HCN) ion channel upon binding of 3',5'-cyclic adenosine monophosphate [Pul+14].

6.4 Site or cofactor localization

A point in space can be determined by measuring distances to at least four known reference points that do not all lie in the same plane. Such *multilateration* is the basis of the global positioning

system (GPS), wherein the reference points are geostationary satellites. In protein structure determination, the reference points or *beacons* are spin-labeled sites in domains whose structure is known with atomistic resolution whereas the unknown point may a residue or nucleotide in a structurally unresolved domain or a cofactor that is not seen in the structure of the resolved domain. The beacon positions can be computed from the known structure by a rotamer library approach.

6.4.1 Multilateration with fixed beacons

In the simplest version the beacon positions are assumed to be known exactly and only the distances between the unknown site and each beacon are required [Hag+13]. As the the beacon positions are in fact somewhat uncertain due to the inaccuracy of rotamer library prediction and the distances have some experimental error, localization needs to be performed by a least-squares approach. The uncertainty of the site position can be estimated by taking advantage of the full distance distribution information or by parametrization of the distribution in terms of Gaussian restraints specified by mean distances $\langle r_i \rangle$ and standard deviations σ_i. For this estimate, probability densities for finding the site are computed on a cubic grid around the most probable position by multiplying the probability densities $\rho_i = \exp[-(\langle r_i \rangle - r_{\mathrm{xyz},i})^2/(2\sigma_i^2)]$, where $r_{\mathrm{xyz},i}$ is the distance of the grid point from the i^{th} beacon. An isodensity surface is then computed that includes a certain fraction of the total probability. Usually, one chooses the surface with a 50% probability of finding the unknown site inside. This is the same choice as used for thermal ellipsoids in the representation of crystal structures.

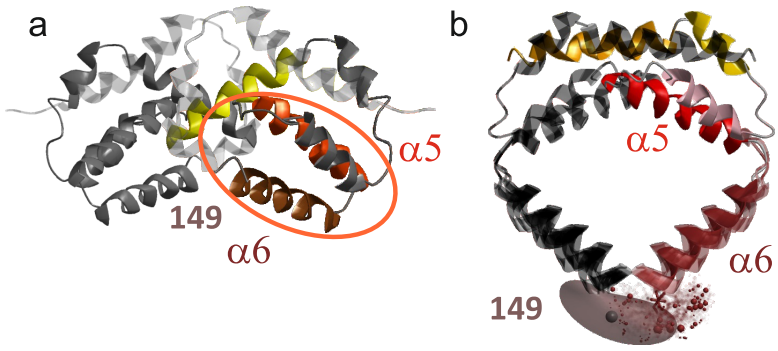


Figure 6.4: Localization of residue 149 of the pro-apoptotic protein Bax in the active, membrane bound form. (a) In the inactive form in solution, helices α5 and α6 form a hairpin. Residue 149 is located 3 residues downstream from the end of helix α6. Based on the NMR structure PDB# 1F16 of Bax. (b) In the active form, residue 149 was localized by distance measurements to six reference points in the dimerization domain resolved in x-ray crystal structure PDB# 4BDU. The brown surface visualizes uncertainty of the localization. Helix α6, which is not contained in the crystallized construct, could then be modeled by using this localization and secondary structure information. Data taken from [Ble+14a].

As an example consider localization of residue 149 in the pro-apoptotic protein Bax in the active, membrane bound form (Fig. 6.4). A structure of full-length Bax was available only for the inactive form observed in solution in the absence of the activator Bid. In the presence of Bid and of mitochondria-like liposomes, Bax dimerizes, inserts into the liposome bilayer, further oligomerizes and form a pore in the liposome. A crystal structure of a construct of Bax truncated after residue 122 and joined to green fluorescent protein (GFP) at the N terminus was available. First, several distances within the dimerization domain (residues 1-122) with supposedly known structure were measure and compared to distances predicted by the rotamer library approach [Ble+14a]. This confirmed that activated Bax inserted into mitochondria-like liposomes has the

same dimerization-domain structure as observed in crystals of the Bax-GFP construct within the precision of the EPR approach. Then, six distances between beacons in the dimerization and MTSSL attached to residue 149 near the end of helix $\alpha 6$ were measured and the spin label at residue 149 was localized. Assuming that secondary structure is maintained during activation, as indicated by circular dichroism spectra, helix $\alpha 6$ could then be modeled. The same strategy was applied to residues 168 and 193. In these cases uncertainty was large, but it could still be inferred that the dimerization domain lines the pore in the liposome.

6.4.2 Localization by distance matrix geometry

It cannot always be assumed that the known structure of a "rigid" domain is completely invariant during ligand binding or other state changes. If distances between the beacons inside this domain somewhat deviate from the distances predicted from the known structure, it may be prudent to relax the beacon configuration during localization of the unknown site. This is possible with distance matrix geometry [12]. Distance matrix geometry is based on the finding that a sufficient number of pair distances between N points fully specifies the geometry of the point arrangement, except for translation (3 degrees of freedom) and rotation (another 3 degrees of freedom) of the object as a whole and inversion symmetry. The problem of inversion symmetry vanishes for chiral objects with known chirality and $N \geq 3$. The number of free parameters is $3N - 6$ and the number of available distances is $N(N-1)/2$, i.e., the problem can be fully determined for $N = 2$ and $N = 3$ and can always be overdetermined for $N > 3$.

In distance matrix geometry [CH88], matrices of lower and upper bounds for all distances r_{ij} with $i, j = 1 \ldots N$ and $i \neq j$ are generated. The bounds can be tightened by making sure that all triples of points fulfill triangle inequalities. By a Monte Carlo approach, one then generates a set of distance matrices that fulfill all lower and upper bounds. If the distances are supposed to be exact and are all known, the distance matrix can be directly set and the approach becomes deterministic. From the distance matrix, a metric matrix can be computed and its eigenproblem be solved. If the distance matrix corresponds to an object in 3D space, the metric matrix has at most three positive eigenvalues. The eigenvectors corresponding to these eigenvavlues scaled by the square root of the corresponding eigenvalue determine the x, y, and z coordinates of the N points. Errors in the distance matrix lead to a solution that only approximates the known distances. Hence, the initial solution is used as a starting point for a least-squares refinement that can include other restraints. For small numbers $N < 100$ computation of a coordinate set is very fast.

As a localization approach, distance matrix geometry does not only provide the coordinates of the unknown site, but also corrected relative coordinates of the beacons. The corrected beacon coordinates must then be least-square fitted to the original beacon coordinates in order to localize the unknown site with respect to the known domain structure. Alternatively, the known domain structure could be first relaxed by the Zheng-Brooks approach (Section 6.3.2) or by restraint-augmented homology modeling (Section 6.2) using all inter-beacon distances as restraints.

6.5 Rigid-body docking

Relative positioning of two rigid bodies has only six degrees of freedom, three translational and three rotational ones. If the bodies are equivalent, as in a protein homodimer, and C_2 symmetry can be assumed, the number of degrees of freedom reduces to four (Fig. 6.5). Such problems can be overdetermined by EPR distance distribution restraints, as it is usually possible to measure 10-20 restraints with reasonable effort. This is helped by the fact that only single mutants of the individual domains need to be spin labeled and in the heterodimer, a set of n singly-labeled

constructs of the first body and m singly labeled constructs of the second body can provide up to nm distances if all of them fall into the DEER-accessible range.

Since the number of degrees of freedom is so small, the problem can be solved by a full grid search. The interval of the Euler angles (rotational degrees of freedom) is finite in any case and intervals of the translational degrees of freedom can be restricted to the range where the two bodies can still be in contact, as such contact is expected for a macromolecular complex. The grid search can be performed against distance restraints, but it is also possible to directly fit primary DEER data [07]. The latter approach avoids solving the ill-posed problem of converting the form factor to the distance distribution and defers background correction of the DEER data to a point where it is stabilized by a model.

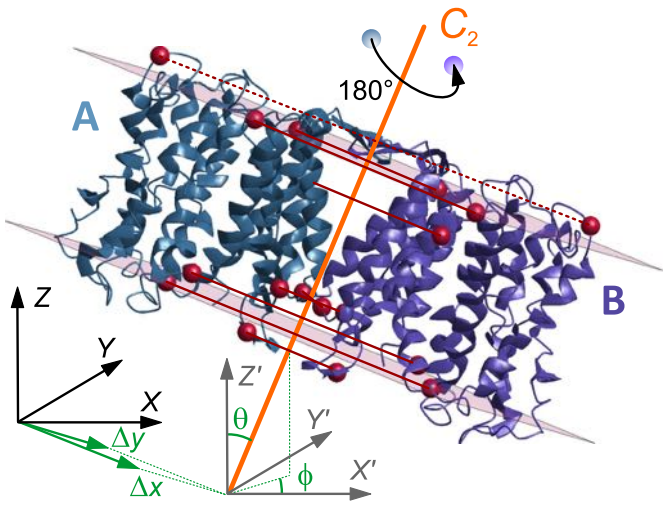


Figure 6.5: Rigid-body docking for determination of the homodimer arrangement of the Na^+/H^+ antiporter NhaA of *Escherichia coli*. The structure of the protomer was known by x-ray diffraction (PDB# 1ZCD) and was assumed to be rigid. Nine distance restraints were used to determine four free parameters, the two polar angles θ and ϕ that specify the orientation of the C_2 symmetry axis in the coordinate frame of the crystal structure and the two parameters Δx and Δy that specify the translation of the center of gravity of the protomer from the C_2 axis within the plane perpendicular to this axis. The second protomer is then generated by a 180° rotation about the C_2 axis. Crimson spheres denote labeling sites, pink planes the headgroup planes of the lipid bilayer, and dotted red lines the distance restraints. Data taken from [07].

The approach is visualized in Fig. 6.5 for the homodimer case of the Na^+/H^+ antiporter NhaA of *Escherichia coli*. As the four free parameters, the polar angles θ and ϕ of the C_2 symmetry axis with respect to the crystal structure coordinates of the protomer and the two distances Δx and Δy of the center of gravity of the protomer from the C_2 axis in a plane perpendicular to this axis were chosen. The original grid search was performed for θ in the interval $0 \ldots 90°$, for ϕ in the interval $0 \ldots 360°$ both in steps of 5°, and for Δx and Δy in the interval $0 \ldots 100$ Å in steps of 5 Å and the mean-square deviation between model distances and experimental mean distances was minimized over the grid. A clear global minimum was found. The rigid-body arrangement was then refined by minimization of the mean-square deviation of simulated from experimental primary DEER data. In the solution, the two protomers were in contact but did not clash. Side chain orientations in the interface were optimized by Gromos. The final structure was in agreement with low-resolution cryo-EM data for the dimer. A later higher-resolution cryo-EM structure of the dimer was in good agreement with the EPR-based model, except for a periplasmic β sheet that was displaced in the structure of the protomer by crystal contacts [09].

This displacement illustrates the danger of transferring crystal artefacts into the final model when working with a rigid-body approach.

6.6 Intrinsically disordered domains

Intrinsically disordered domains (IDDs) and intrinsically disordered proteins (IDPs) must be represented by ensembles of conformations. If all restraints are only mean values over the ensemble, one cannot have certainty on conformational variability of the ensemble. There exist two basic approaches for coping with this situation. The minimal ensemble approach takes the Anfinsen dogma as its point of departure and aims to find the minimum number of distinct conformations that can fulfill all experimental restraints. This approach may be appropriate for narrow ensembles, but it has the disadvantages that it corresponds to a minimum entropy rather than maximum entropy solution of the problem and that it has the counter-intuitive consequence that additional restraints will tend to broaden rather than narrow the ensemble. The maximal ensemble approach aims to find a set of conformations that are a good representation of the total conformation space that is accessible given the experimental restraints. While this provides a maximum entropy solution, it requires efficient sampling of a vast space of conformations and a decision on the number of ensemble members that suffices to represent the accessible subspace. With restraints that are mean values over the ensemble, the maximal ensemble approach requires that first a large set of conformations is computed that is representative of the whole conformation space. The ensemble can then be culled based on the restraints.

The distance *distribution* restraints available by pulse dipolar spectroscopy allow for a different implementation of the maximal ensemble approach. Now each individual conformation can be assigned a probability based on how the model distances are situated with respect to the experimental distance distributions [Jes16]. In fact, distance distribution restraints can be tested as soon as information on the required site coordinates is available and further computation of an individual conformation can be aborted as soon as it becomes clear that it violates the restraints than can already be evaluated. This allows for more efficient sampling of the large space of conformations. Furthermore, the final ensemble can be tested against the experimental distance distributions regarding both their mean value and width. This makes it easier to avoid both an underestimate and an overestimate of the width of the ensemble of conformations.

For IDPs, except for experimental restraints only the bond length and angles and the Ramachandran angle statistics are known. Furthermore, conformations can be rejected if the peptide chain clashes with itself. IDDs are further restrained by their attachment to an ordered domain whose structure may be known. For terminal domains, the attachment restraint is weak as it only causes a superposition of one terminal residue of the IDD. Part of the conformation space is excluded by avoidance of clashes with the ordered domain(s). Internal IDDs are restrained by two attachment points (anchor residues). This excludes large parts of conformation space, in particular for short IDDs.

6.6.1 Restraint types

Internal distance distribution restraints

For an IDP, distance distribution restraints (DDRs) between two sites in the IDP are the most informative restraints. They can be evaluated as soon as coordinates for the second site become available.

Beacon restraints

For an IDD, DDRs between one beacon site in an ordered domain of known structure and one site in the IDD are the most informative restraints. This is because they will tend to be more

narrowly distributed than domain-internal DDRs and will thus reject more conformations. Such restraints can be evaluated as soon as coordinates for the IDD-internal site become available.

Oligomer restraints

If the IDD is attached to a homooligomer with C_n symmetry, DDRs can be measured by spin-labeling a single site per protomer within the IDD. Such restraints can be evaluated as soon as coordinates for the IDD-internal site become available and they can be obtained with less effort than internal or beacon restraints. However, such restraints encode only the distance of the labeled site from the C_n axis without providing information on positioning along this axis. Therefore, oligomer restraints always need to be augmented by other restraints [Feh+15].

Bilayer immersion depth restraints

Restraints on bilayer immersion depth from water accessibility or oxygen accessibility measurements provide only a mean over the ensemble, are less precise than DEER distance restraints, and can be obtained only for membrane proteins. For homooligomeric membrane proteins, they can be a valuable complement to to oligomer restraints, as they provide information on site position along the C_n symmetry axis [Feh+15].

Secondary structure propensities

Secondary structure restraints or propensities can be used already during Monte Carlo conformer generation as they influence the Ramachandran angle statistics underlying backbone generation.

6.6.2 Flexibility and uncertainty

The width of a maximal ensemble for an IDD or IDP has two contributions. One of them stems from intrinsic disorder and characterizes flexibility of the chain. The other one stems from a lack of restraints and reduces when adding more restraints. With DDRs from spin labels, uncertainty also has a contribution that stems from conformation distribution of the label itself. In other words, even if the width of the ensemble converges by adding more restraints, the converged width may still be an overestimate of the true width.

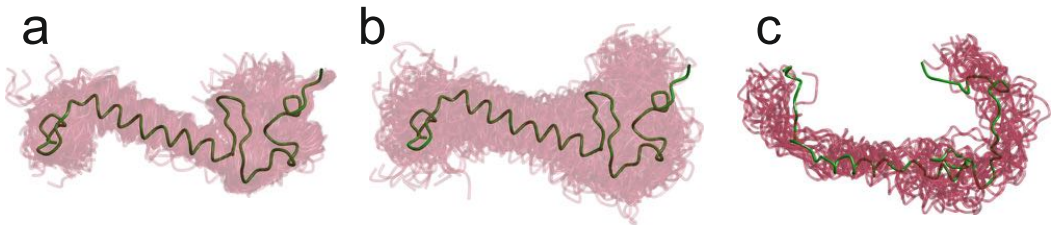
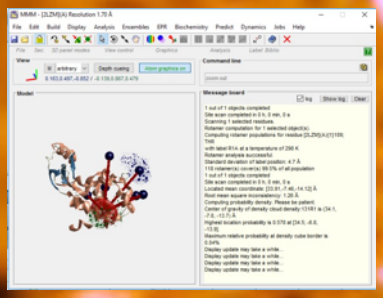


Figure 6.6: Evaluation of modeling on an IDP and of a well-defined structure from simulated EPR-based distance distribution restraints (DDRs) for the kinase inhibitory domain of p27. (a) Ensemble model for p27KID deposited in the Protein Ensemble Database (identifier PED2AAA). The whole ensemble is visualized by transparent crimson coil models and its central conformation as a guide to the eyes by a green coil model. 56 DDRs were computed from this ensemble by the rotamer library approach. (b) Ensemble recovered from the 56 DDRs and from secondary structure information on the helical section (crimson) seen in the central conformation of ensemble PED2AA (green). (c) Well-defined structure of p27Kip1 bound to Cdk2 complex (PDB# 1JSU, green) and ensemble recovered from 56 DDRs that were simulated from the well-defined structure and secondary structure restraints on the helical section.

This is illustrated in Fig. 6.6 for two examples. 56 DDRs were simulated from an ensemble structure of the kinase inhibitory domain of p27 (residues 25-93) and used in ensemble modeling together with secondary structure restraints on the helical section (residues 38–58). The ensemble

recovered from these restraints (panel (b)) is clearly broader than the one used for simulating the restraints. This is only partially due to a lack of restraints. Panel (c) shows the same exercise for the well-defined structure of p27Kip1 bound to Cdk2 complex where the restraints were generated from a single conformation. The EPR-derived ensemble is clearly narrower than the one derived from restraints for the ensemble model of p27KID (panel (b)), yet it is also significantly less defined than the input structure. This demonstrates that even at a level of 56 DDRs per 69 residues, atomistic resolution cannot be expected from an EPR-derived model.



7 — Tutorial

7.1 EPR spectrum assignment using EasySpin

7.1.1 Task

This task is related to an industrial problem, but the same radical species is also observed in proteins with flavin cofactors. Section 4.1.2 and Fig. 4.4 provide additional information. The experimental CW EPR spectrum was measured in X band on an industrial riboflavin (vitamin B_2) sample taken from the production process before final purification. Comparison with literature spectra indicated that it may be assigned to a flavin semiquinone radical. The question arose whether the radical is primarily in its charge-neutral or in its anionic state. It is solved here by comparison with EasySpin simulations, based on hyperfine parameters and g tensors of the two radicals published in [Bar+03].

7.1.2 EasySpin

The spectrum simulations are based on the EasySpin package downloadable at `http://www.easyspin.org/`. This package provides, among else, functions for loading experimental spectra stored by commonly used spectrometers (`eprload`) and for simulating CW EPR spectra of solid samples (`pepper`). In an EasySpin simulation, parameters of the spin system have to be provided in a Matlab structure `Sys`. This is illustrated here for the neutral flavin semiquinone radical:

```
% flavine neutral radical parameters
Sys_n.g = [2.00425,2.00360,2.00227];
Sys_n.Nucs = '14N,14N,1H,1H,1H,1H';
Sys_n.A = [0.2,0.2,52.5;...
    2.0,2.0,28.9;...
    0.2,38.6,25.8;...
    8.1,9.3,8.4;...
    10.0,8.5,8.5;...
    1.4,5.4,6.0];
Sys_n.AFrame = [45 7 0;...
    12 13 0;...
    8 5 0;...
    8 5 0;...
    8, 5 -60;...
    8, 5, 30]*pi/180;
```

```
Sys_n.lwpp = lwpp;
```

The system variable is called `Sys_n` here to indicate the neutral radical. The g tensor principal values (`Sys_n.g`) were taken from line G in Table 2 of Ref. [Bar+03]. `Sys_n.Nucs` defines the isotopes of the nuclei for which hyperfine tensors are provided and `Sys_n.A` defines the principal values of the hyperfine tensors, which are also taken for Table 2 of Ref. [Bar+03]. The must be provided in units of MHz. Note that in the literature, sometimes other units are used. In Matlab, `Sys_n.A` is a matrix. Within each line, values are separated by commas or space, lines are separated by semicolons. Each line lists the three principal values for a single nucleus, in the same order as the isotopes are given in `Sys_n.Nucs`. In order to make the code more readable, the lines are separated by linebreaks, which in Matlab requires three dots (`...`) to indicate to the interpreter that the code line continues.

The hyperfine tensors have principal axes systems (PAS) that differ from the one of the g tensor. They can be related to the g tensor PAS by Euler rotations. The angles for the Euler rotations are also given in Table 2 of Ref. [Bar+03]. `Sys_n.AFrame` defines these Euler angles in the same way as `Sys_n.A` defines the principal values. Usually, Euler angles are given in degrees in the literature, while EasySpin expects them in radians. This is taken care of by multiplying the whole matrix by π (Matlab constant `pi`) and dividing it by 180. Euler angles that are not reported were set to zero.

`Sys_n.lwpp` defines the peak-to-peak linewidth. Table 2 of Ref. [Bar+03] reports a linewidth tensor, but here we are content with an isotropic linewidth. This is an adjustable parameter that we use to fit the simulated to the experimental spectrum. As we want to use the same linewidth for both the neutral and anion radical and are too lazy to change it in two lines, we define a variable `lwpp` that is set at the beginning of the Matlab script.

Next, we load and plot the spectrum.

```
% load the spectrum
[B0_exp,spc_exp,Par_exp] = eprload('riboflavin_EPR');

% plot the experimental spectrum
figure(1); clf;
plot(B0_exp/10,spc_exp,'k'); % field axis in mT instead of G
hold on; % avoid that the plot is cleared when adding simulated spectra
```

For a spectrum in Bruker's Xepr format (.DTA for the data, .DSC for the description), EasySpin can import not only the magnetic field axis `B0_exp` and spectral intensity `spc_exp`, but also the measurement parameters `Par_exp`, which are returned as a Matlab structure. We shall use them below. The following lines open a figure, plot it, and tell Matlab that further plots are to be added to this figure instead of replacing the existing plot.

Now, we define experimental parameters so that the experimental and simulated spectra match:

```
% set simulation parameters to spectrum parameters
Exp.Range = [Par_exp.XMIN/10,(Par_exp.XMIN+Par_exp.XWID)/10]; % ...
    conversion from G to mT
Exp.mwFreq = Par_exp.MWFQ/1e9; % conversion from Hz to GHz
Exp.nPoints = Par_exp.XPTS;
Exp.mwFreq = Exp.mwFreq + frq_shift; % compensate frequency for ...
    imprecise field measurement
```

Bruker provides in the .DSC file the start of the field axis `XMIN` and its width `XWID`, both in units of Gauss. We convert this to a field range in units of milliTesla, as required by EasySpin. Bruker provides the microwave frequency `MWFQ` in Hz, we convert it to GHz for EasySpin. We

use the same number of data points `XPTS` for the simulated as for the experimental spectrum. That way we could easily compute an RMSD if we wish. The last line of this script is related to the absence of an internal reference for the g value. The magnetic field axis of X- and Q-band EPR spectra stems from a Hall probe measurement that can be off for a few Gauss even if the Hall probe was carefully calibrated on spectrometer installation. If precise g values are required, an internal reference or field calibration with an external reference are required. Here we are mostly interested in the line shape. We compensate for the field shift by an equivalent frequency shift, where `frq_shift` is an adjustable parameter. This compensation of field shifts by frequency shifts is possible if the g anisotropy is small, as is the case for organic radicals.

```
% set options for fast simulation at reasonable quality
Opt.Method = 'perturb';
```

Finally, we tell EasySpin that it should use a perturbation approach rather than exact diagonalization of the Hamiltonian. It is possible to call the `pepper` function without options and rely on the defaults, but in this particular case with that many nuclei computation time would be very long.

7.1.3 Problem

Vary parameters `lwpp` and `frq_shift` until you get a reasonable fit with one of the two parameter sets. Are neutral radicals or anion radicals predominating in this sample? Do you have an idea how the fit could be further improved? Fig. 7.1 illustrates the fit quality that can be achieved without using additional features of EasySpin.

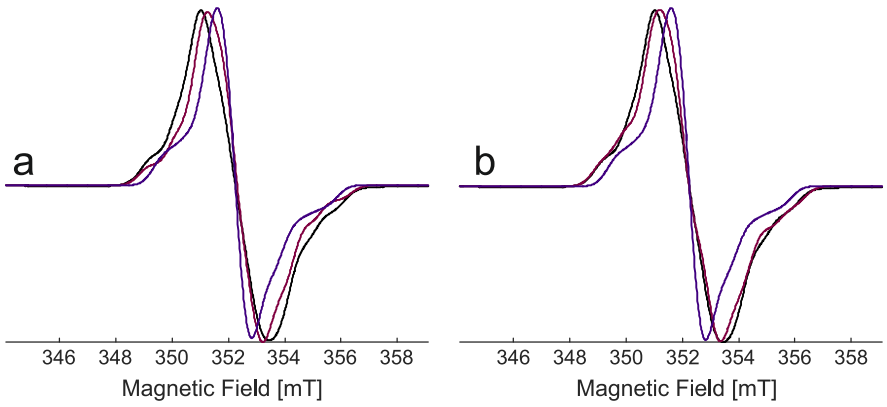


Figure 7.1: Experimental X-band CW EPR spectrum of a riboflavin sample (black) and simulated spectra of flavin semiquinone neutral and anion radicals. (a) Spectra simulated with parameters taken from Ref. [Bar+03]. (b) For one spectrum, some parameters were adapted.

7.2 Analysis of DEER data in terms of a distance distribution

7.2.1 Task

An experimental data set of T4 Lysozyme spin-labeled by MTSL at residues 82 and 131 is given (data courtesy Enrica Bordignon and Irina Ritsch). The data should be analyzed in DeerAnalysis in terms of a distance distribution and its uncertainty.

7.2.2 DeerAnalysis

DeerAnalysis is an open-source Matlab-based program for extracting distance distributions from DEER data and for validating them. Using the `Load` click button in the `Data sets` panel on the top right in the graphical user interface (GUI), the data set `DEER_82R1_V131R1...` is loaded. In the `Original data` panel on the top left, click the blue exclamation mark (`!`) button of the `Backgr. (ns)` fit row. To better see the background fit, uncheck the `imaginary` checkbox in the same panel.

The oscillations decay within 2 μs, yet the form factor is not flat at the end. This is an artefact due to excitation band overlap. In the orange `Cutoff` edit field in the `Original data` panel type `4000`. This cuts the data at 4 μs, which is fully sufficient and eliminates the artefact. Repeat the background fit.

An estimate of the distance distribution is seen in the corresponding panel on the top right. This is computed by fast approximate Pake transformation. For the sample at hand, this works well, but Tikhonov regularization with non-negativity constraint is preferable. To perform Tikhonov regularization, selecte the corresponding radiobutton in the `Distance analysis` panel on the bottom right and check the `L curve` checkbox. Click the `Compute` button to the right of this checkbox. After a while, the L curve is displayed in the plot below the `Distance ... distribution` panel (Fig. 7.2(a)). Initial corner location (red circle) is by minmum distance to the lower left corner of the L curve range. In the case at hand, you find a clear corner and can select this one by clicking on it. The marker circle moves to the point that you clicked and turns blue. The `Reg. par.` edit field in the `Distance distribution` panel should now show a value between 200 and 350 (depending on where exactly you clicked). Now uncheck the `L curve` checkbox in the `Distance distribution` panel. The distance distribution corresponding to the selected regularization parameter is displayed.

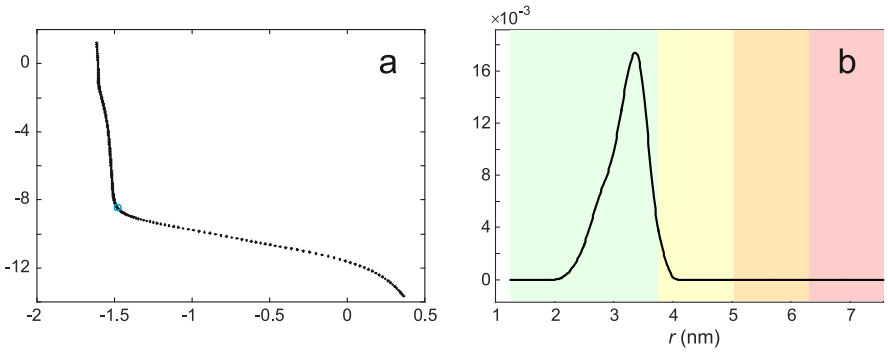


Figure 7.2: Distance distribution analysis for T4 Lysozyme double mutant 82R1/131R1. (a) Tikhonov regularization L curve (black) with manual corner selection (blue). (b) Distance distribution corresponding to the corner of the L curve. Virtually all probability density is in the green range, where mean distance, width, and shape are reliable. In the yellow range, only mean distance and width are reliable, in the orange range only the mean distance is. Contributions in the red range would merely indicate that there are longer distances than can be quantified with the given form factor length.

The distance distribution is almost entirely in the green range and hence the mean distance ($\langle r \rangle$3.20 nm), width ($\sigma_r = 0.34$ nm), see `Distance analysis` panel) and shape are expected to be reliable (Fig. 7.2(b)). However, this color code is based only on form factor length and does not take into account signal-to-noise ratio or other uncertainties. To get a better idea about uncertainty, click on the `Validation` button in the `Distance analysis` panel. In the validation window that opens click on the `Compute` button. The program will perform Tikhonov

regularization for a set of different background fit ranges. After a while, the uncertainty of the distance distribution due to noise is visualized in the `Distance distribution` plot. The red line corresponds to the mean distribution plus two standard deviations, the blue curve to the mean distribution minus two standard deviations. In the case at hand, the uncertainty is very low and it would remain low if you would also test for the noise influence by checking the `White noise` checkbox and selecting 11 trials in the edit field to its right. When you close the validation window you can also visualize this error estimate by checking the corresponding checkbox in the `Distance distribution` panel of the main window. Please save the distance distribution by clicking on the `Save` button. We need it for the next task.

7.3 *In silico* spin labeling with a rotamer library

7.3.1 Task

Here we predict the distance distribution between MTSL attached to residues 82 and 131 in T4 Lysozyme and compare the prediction with experiment. We use a rotamer library for MTSSL that is provided with the open-source Matlab-based program MMM.

7.3.2 MMM

After starting MMM, load the T4 lysozyme crystal structure with PDB identifier 2LZM via the `New from PDB web...` item of the `File` menu. You could select residues 82 and 131 via the `Hierarchy` window accessible in the `Display` menu, but it is faster and easier to do so via the `Command line` of the main MMM window by typing `select 82,131`. Now use `Site ... scan/selected residues` from the `EPR` menu. In the `Site scan setup` window, you can just click `OK`. In the `Set labeling conditions for selected residues` window you could select a different spin label, but since MTSL is the default choice, you can also just click `OK`. The program asks you for a file name for storing site scan results in an HTML file. Change the name or confirm the suggestion by just clicking `OK`. After the computation is finished, the results can be view in a browser. You will see that both sites are well accessible (more than 100 rotamers, partition function larger than 1).

At that point, the labels are computed, but not yet attached. You can attach them via the `Attach precomputed rotamers` item in the `EPR` menu. The `Set labeling ... conditions for selected residues` window appears again, you can just click `OK`. Now click on the red cross in the tool button line to unselect everything. You can visualize the labels as follows. Type `show 82,131 label` in the `Command line`. This will show only the N-O bond midpoint positions of all rotamers and visualize their populations. The spin label sidechains can be visualized by the command `show 82,131 ball&stick`. In this case, population is encoded by transparency.

We now want to compare the predicted distance distribution and the experimental distance distribution. For that, open the `DEER` window of MMM via the `DEER` item of the `EPR` menu. In the `Labels & selected atoms` panel, you will see a list of spin labeled sites, in this case only `[2LZM](A){1}82` and `[2LZM](A){1}131`, where `(A)` denotes the chain and `{1}` the model. Select `[2LZM](A){1}82` and click the exclamation mark `!` button above. Now select `[2LZM](A){1}131` and click the plus `+` button. The predicted distance distribution and form factor are displayed in the corresponding plots to the right. Load the experimental data and distance distribution vial the `Load DeerAnalysis` button. You will see that the shape and width of the distributions fit quite well, but the predicted distances are longer than the experimental result.

You can check just how well the shape and width fit and how large the distance shift is by checking the `Flexible` checkbox in the `Fit options` panel. In this case, a 0.26 nm (2.6 Å)

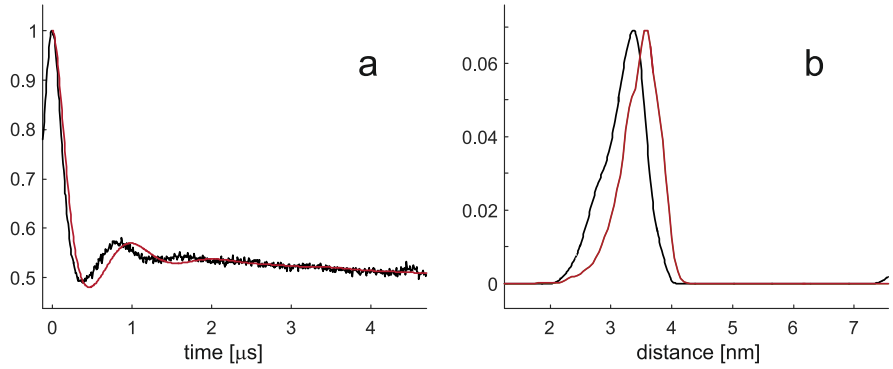


Figure 7.3: Distance distribution prediction for T4 Lysozyme double mutant 82R1/131R1. (a) EXperimental (black) and predicted (red) primary DEER data. (b) Experimental (black) and predicted (red) distance distribution.

shorter backbone-backbone distance in solution compared to the crystal structure would explain the discrepancy. Note however that the prediction is not accurate enough to conclude that the backbone conformation indeed differs.

Now close the DEER window of MMM. This is a modal window, i.e., you cannot access any other MMM or Matlab window while it is open. We shall use the existing label visualization for the next task. To see what is what we color the labels by executing the commands `color ... 82 darkblue` and `color 131 darkgreen`.

7.4 Site localization

7.4.1 Task

Here we localize the MTS label at residue 131 by relying on 4 distances to other residues extracted from data sets provided by Christian Altenbach (UCLA) and 2 data sets provided by Enrica Bordignon and Irina Ritsch and on a GPS-like approach. We compare this localization to the rotamer library prediction.

7.4.2 The multilateration module of MMM

The data sets were DEER distance measurements from MTS-labeled residue 131 to residues 68, 72, 82, 86, 88, and 109. All of them were processed as described in Section 7.2.2. With these data, a restraint file was constructed:

```
% Localization test restraints for MTSL at residue 131
% with respect to T4 Lysozyme structure 2LZM
% Data sets for DEER distance constraints courtesy Christian Altenbach
# PDB 2LZM
# LOCATE 131R1 0.5 MTSL 298 full
68  3.78    0.22
72  3.57    0.28 % this data set by Irina Ritsch, Enrica Bordignon
82  3.20    0.34 % this data set by Irina Ritsch, Enrica Bordignon
86  2.73    0.36
89  3.26  0.27
109 3.02    0.26
# END
```

In MMM restraint files, the hash character precedes keywords and there is space between the `#` and the keyword. The `# PDB` keyword defines the structure where the beacon residues are located. In our case it is T4 Lysozyme structure 2LZM, which is already loaded. The `# LOCATE` keyword defines what should be localized with what probability and with what label. The localized site gets a name (here `131R1`). THe isosurface of localization will include 50% of the probability for finding the label (this is the next argument `0.5`. The rotamer library for `MTSL` is used to specify beacon positions and rotamers are computed at `298` K. The temperature is supplied for historical reasons, newer MMM modules don't require it. The argument `full` specifies the display mode. All beacon positions are displayed by blue spheres and the distance from the beacon to the localized site by crimson lines.

The locate key is followed by lines that specify the restraints. Each line has three arguments, the residue address (here residue number) of the beacon, the mean distance (in nm) from the beacon to the site to be localized, and the standard deviation (in nm) of this distance. The `# END` key stops reading of the restraint file.

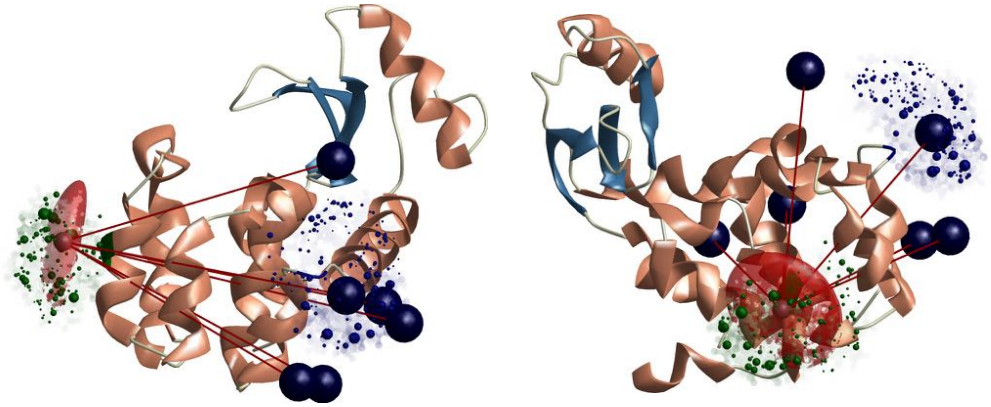
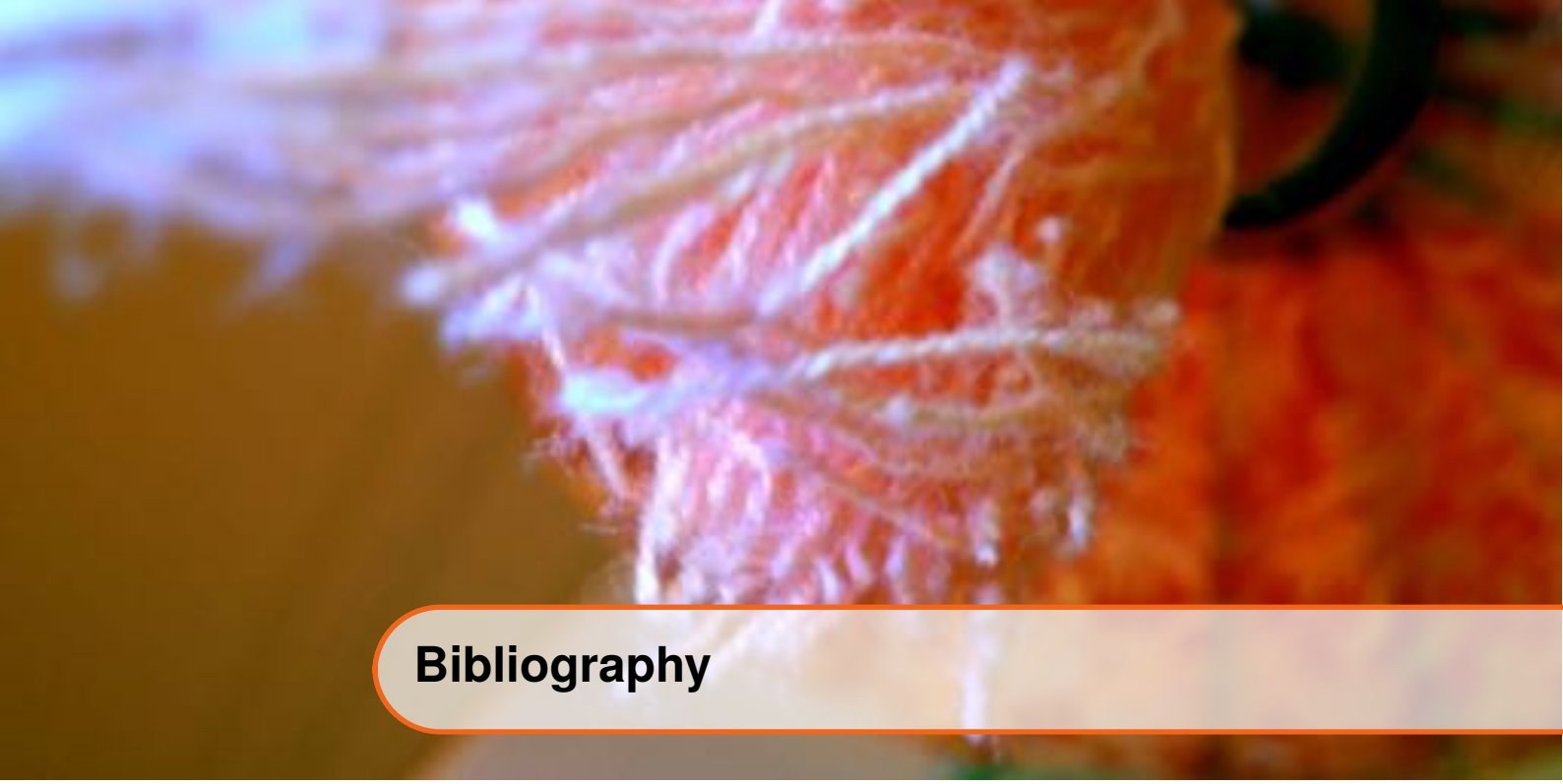


Figure 7.4: Localization of T4L residue 131 with 6 distance restraints. Two views are shown.

The result is visualized in Fig. 7.4. Compared to the localization of the same residue shown in Fig. 6.1 it looks slightly better, because we added one more restraint. The two views show that the position of the beacons matters. In the direction that is along the mean site-site vector the label position is much better defined than perpendicular to it. This is because the beacon sites are not broadly spread in the perpendicular direction.



Bibliography

Books

[CCM16] V. Chechik, E. Carter, and D. M. Murphy. *Electron Paramagnetic Resonance*. 1st Ed. Oxford: Oxford University Press, 2016 (cited on page 8).

[CH88] G. M. Crippen and T. F. Havel. *Distance Geometry and Molecular Conformation*. Taunton: John Wiley & Sons, Inc., 1988 (cited on page 56).

[HE07] M. A. Hemminga and Lawrence J. Berliner (Eds.) *ESR Spectroscopy in Membrane Biophysics*. Vol. 27 of Biol. Magn. Reson. New York: Springer, 2007 (cited on page 8).

[Rie07] Philip Rieger. *Electron Spin Resonance. Analysis and Interpretation*. The Royal Society of Chemistry, 2007, P001–173. ISBN: 978-0-85404-355-2. DOI: 10.1039/9781847557872. URL: http://dx.doi.org/10.1039/9781847557872 (cited on page 11).

[SJ01] A. Schweiger and G. Jeschke. *Principles of Pulse Electron Paramagnetic Resonance Spectroscopy*. Oxford: Oxford University Press, 2001 (cited on pages 14, 16).

Articles

[Alo+15] Noelia Alonso-Garcia et al. "Combination of X-ray crystallography, SAXS and DEER to obtain the structure of the FnIII-3,4 domains of integrin $\alpha 6\beta 4$". In: *Acta Cryst. D* 71.4 (2015), pages 969–985 (cited on page 52).

[Bah+10] Ivet Bahar et al. "Normal Mode Analysis of Biomolecular Structures: Functional Mechanisms of Membrane Proteins". In: *Chem. Rev.* 110.3 (2010), 1463–1497 (cited on page 54).

[Bar+03] B. Barquera et al. "X- and W-band EPR and Q-band ENDOR studies of the flavin radical in the Na+ -translocating NADH:quinone oxidoreductase from Vibrio cholerae." In: *J Am Chem Soc* 125.1 (2003), pages 265–75 (cited on pages 26, 61–63).

[Bib+17] Stefan Bibow et al. “Solution structure of discoidal high-density lipoprotein particles with a shortened apolipoprotein A-I”. In: *Nat. Struct. Mol. Biol.* 24.2 (2017), 187–193 (cited on page 49).

[Ble+14a] Stephanie Bleicken et al. “Structural Model of Active Bax at the Membrane”. In: *Mol. Cell* 56.4 (2014), 496–505 (cited on pages 46, 55).

[Ble+14b] S. Bleicken et al. “Structural model of active Bax at the membrane.” In: *Mol Cell* 56.4 (2014), pages 496–505 (cited on page 2).

[Ble+01] G. Bleifuss et al. “Tryptophan and tyrosine radicals in ribonucleotide reductase: a comparative high-field EPR study at 94 GHz.” In: *Biochemistry* 40.50 (2001), pages 15362–8 (cited on page 24).

[Bod+07] Bela E. Bode et al. “Counting the monomers in nanometer-sized oligomers by pulsed electron - Electron double resonance”. In: *J. Am. Chem. Soc.* 129.21 (2007), 6736–6745 (cited on page 47).

[BGF13] Peter P. Borbat, Elka R. Georgieva, and Jack H. Freed. “Improved Sensitivity for Long-Distance Measurements in Biomolecules: Five-Pulse Double Electron-Electron Resonance”. In: *J. Phys. Chem. Lett.* 4.1 (2013), 170–175 (cited on page 43).

[Bow+97] JH Bowen et al. “EPR studies of the dynamics of rotation of dioxygen in model cobalt(II) hemes and cobalt-containing hybrid hemoglobins”. In: *J. Phys. Chem. B* 101.43 (1997), 8683–8691 (cited on page 36).

[09] “Conformations of NhaA, the Na/H Exchanger from Escherichia coli, in the pH-Activated and Ion-Translocating States”. In: *J. Mol. Biol.* 386.2 (2009), pages 351–365 (cited on page 57).

[Doo17] S. Van Doorslaer. “Understanding heme proteins with hyperfine spectroscopy.” In: *J. Magn. Reson.* xx.yy (2017), zzz–zzw (cited on page 26).

[Dus+14] Olivier Duss et al. “EPR-aided approach for solution structure determination of large RNAs or protein-RNA complexes”. In: *Nat. Comm.* 5 (2014) (cited on pages 36, 49).

[Feh+15] Niklas Fehr et al. “Modeling of the N-terminal Section and the Lumenal Loop of Trimeric Light Harvesting Complex II (LHCII) by Using EPR”. In: *J. Biol. Chem.* 290.43 (2015), 26007–26020 (cited on page 59).

[Fle+09] Mark R. Fleissner et al. “Site-directed spin labeling of a genetically encoded unnatural amino acid”. In: *Proceedings of the National Academy of Sciences* 106.51 (2009), pages 21637–21642 (cited on pages 37, 38).

[Gri+01] S. Grimaldi et al. “QH*- ubisemiquinone radical in the bo3-type ubiquinol oxidase studied by pulsed electron paramagnetic resonance and hyperfine sublevel correlation spectroscopy.” In: *Biochemistry* 40.4 (2001), pages 1037–43 (cited on page 26).

[Hag+13] Gregor Hagelueken et al. “mtsslSuite: In silico spin labelling, trilateration and distance-constrained rigid body docking in PyMOL”. In: *Mol. Phys.* 111.18-19 (2013), 2757–2766 (cited on page 55).

[07] “High-Resolution Structure of a Na+/H+ Antiporter Dimer Obtained by Pulsed Electron Paramagnetic Resonance Distance Measurements”. In: *Biophys. Journal* 93.10 (2007), pages 3675–3683 (cited on page 57).

[Hil+05] D Hilger et al. “Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na+/H+ antiporter of E-coli”. In: *Biophys. J.* 89.2 (2005), 1328–1338 (cited on page 47).

[HKS96] Richard H. Holm, Pierre Kennepohl, and Edward I. Solomon. “Structural and Functional Aspects of Metal Sites in Biology”. In: *Chemical Reviews* 96.7 (1996), pages 2239–2314 (cited on pages 25, 28).

[Jes05] G. Jeschke. “EPR techniques for studying radical enzymes.” In: *Biochim Biophys Acta* 1707.1 (2005), pages 91–102 (cited on page 24).

[Jes12a] G. Jeschke. “Characterization of Protein Conformational Changes with Sparse Spin-Label Distance Constraints”. In: *Journal of Chemical Theory and Computation* 8.10 (2012), pages 3854–3863 (cited on page 54).

[Jes12b] Gunnar Jeschke. “DEER Distance Measurements on Proteins”. In: *Annu. Rev. Phys. Chem.* 63 (2012), 419–446 (cited on page 42).

[Jes12c] Gunnar Jeschke. “Optimization of Algorithms for Modeling Protein Structural Transitions from Sparse Long-Range Spin-Label Distance Constraints”. In: *Zeitschr. Phys. Chem.* 226.11-12 (2012), 1395–1414 (cited on page 54).

[Jes13] Gunnar Jeschke. “Conformational dynamics and distribution of nitroxide spin labels”. In: *Progr. Nucl. Magn. Reson. Spectrosc.* 72 (2013), 42–60 (cited on page 50).

[Jes16] Gunnar Jeschke. “Ensemble models of proteins and protein domains based on distance distribution restraints”. In: *Proteins* 84.4 (2016), 544–560 (cited on pages 39, 58).

[Kol+02] M. Kolberg et al. “Protein thiyl radicals directly observed by EPR spectroscopy.” In: *Arch Biochem Biophys* 403.1 (2002), pages 141–4 (cited on page 23).

[Kuc+08] Norbert Kucerka et al. “Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data”. In: *Biophys. J.* 95.5 (2008), 2356–2367 (cited on page 45).

[Kuc+17] Svetlana Kucher et al. “Orthogonal spin labeling using click chemistry for in vitro and in vivo applications”. In: *Journal of Magnetic Resonance* 275 (2017), pages 38–45 (cited on page 38).

[Lei+10] Alexander Leitner et al. “Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics”. In: *Mol. Cell. Proteomics* 9.8 (2010), 1634–1649 (cited on page 52).

[Lem+00] Sebastião S. Lemos et al. “Comparison of EPR-Visible Cu2+ Sites in pMMO from Methylococcus capsulatus (Bath) and Methylomicrobium album {BG8}”. In: *Biophys. J.* 79.2 (2000), pages 1085–1094 (cited on page 28).

[12] “Locating a Lipid at the Portal to the Lipoxygenase Active Site”. In: *Biophys. J.* 103.10 (2012), pages 2134–2144 (cited on page 56).

[LRG07] Wolfgang Lubitz, Eduard Reijerse, and Maurice van Gastel. “[NiFe] and [FeFe] Hydrogenases Studied by Advanced Magnetic Resonance Techniques”. In: *Chemical Reviews* 107.10 (2007), pages 4331–4365 (cited on page 28).

[06] “Modeling Protein Conformational Changes by Iterative Fitting of Distance Constraints Using Reoriented Normal Modes”. In: *Biophys. J.* 90.12 (2006), pages 4327–4336 (cited on page 54).

[95] "Orbital interactions, electron delocalization and spin coupling in iron-sulfur clusters". In: *Coord. Chem. Rev.* 144 (1995), pages 199–244 (cited on page 29).

[Pet+12] Maxim V. Petoukhov et al. "New developments in the ATSAS program package for small-angle scattering data analysis". In: *J. Appl. Cryst.* 45.2 (2012), 342–350 (cited on page 52).

[PBJ11] Yevhen Polyhach, Enrica Bordignon, and Gunnar Jeschke. "Rotamer libraries of spin labelled cysteines for protein studies". In: *Phys. Chem. Chem. Phys.* 13.6 (2011), 2356–2366 (cited on page 39).

[PJ10] Yevhen Polyhach and Gunnar Jeschke. "Prediction of favourable sites for spin labelling of proteins". In: *Spectroscopy* 24.6 (2010), 651–659 (cited on page 39).

[Pul+14] Michael C. Puljung et al. "Double electron–electron resonance reveals cAMP-induced conformational change in HCN channels". In: *Proc. Nat. Acad. Sci. USA* 111.27 (2014), pages 9816–9821 (cited on page 54).

[Rab+14] Michael Raba et al. "Extracellular Loop 4 of the Proline Transporter PutP Controls the Periplasmic Entrance to Ligand Binding Sites". In: *Structure* 22.5 (2014), 769–780 (cited on pages 44, 54).

[Rap11] J Rappsilber. "The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes". In: *J. Struct. Biology* 173.3 (2011), pages 530–540 (cited on page 52).

[Seg+16] Takuya F. Segawa et al. "Water accessibility in a membrane-inserting peptide comparing Overhauser DNP and pulse EPR methods". In: *J. Chem. Phys.* 144.19 (2016) (cited on page 45).

[SKT10] L. J. Smith, A. Kahraman, and J. M. Thornton. "Heme proteins–diversity in structural characteristics, function, and folding." In: *Proteins* 78.10 (2010), pages 2349–68 (cited on page 25).

[Sut+17] Elizaveta A. Suturina et al. "Model-free extraction of spin label position distributions from pseudocontact shift data". In: *Chem. Sci.* 8.4 (2017), 2751–2757 (cited on page 35).

[Wal99] F. A. Walker. "Magnetic spectroscopic (EPR, ESEEM, Mössbauer, MCD and NMR) studies of low-spin ferriheme centers and their corresponding heme proteins." In: *Coord. Chem. Rev.* 185-186 (1999), pages 471–534 (cited on page 26).

[Wie+17a] Thomas Wiegand et al. "Solid-state NMR and EPR Spectroscopy of Mn2+-Substituted ATP-Fueled Protein Engines". In: *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION* 56.12 (2017), 3369–3373 (cited on page 36).

[Wie+17b] T. Wiegand et al. "Solid-state NMR and EPR Spectroscopy of Mn2+ -Substituted ATP-Fueled Protein Engines." In: *Angew Chem Int Ed Engl* 56.12 (2017), pages 3369–3373 (cited on page 2).