Molecular and Structural Biology V:
Studying Macromolecules by NMR and EPR

Gunnar Jeschke, HCI F227, gjeschke@ethz.ch

Two Lectures
- general introduction to EPR techniques & intrinsic paramagnetic centers in biological systems (11.4.)
- spin labeling & structure modeling (18.4.)

One tutorial (2.5.)
- simulating EPR spectra with EasySpin (two short examples)
- analyzing DEER data in terms of a distance distribution (two examples)
- rotamer library simulation of spin labels and comparison to DEER data
- localization of a spin label site in a protein

Script ☞ in-depth discussion, reference for future research work (epr.ethz.ch/education.html)

Examination content will be specified at the end of semester

The focus is on information from EPR and its use in structural biology, not on inner working and theory of EPR

(see „The EPR part of the ETH Magnetic Resonance lecture script“ at epr.ethz.ch/education.html)

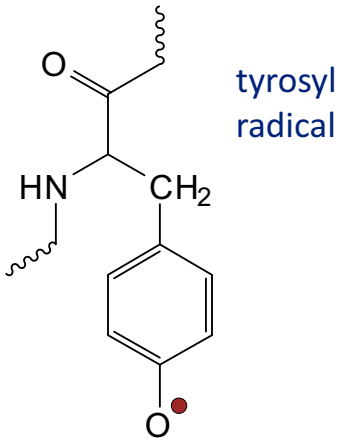
We need an unpaired electron

Three basic types of systems

Native

Radical enzymes

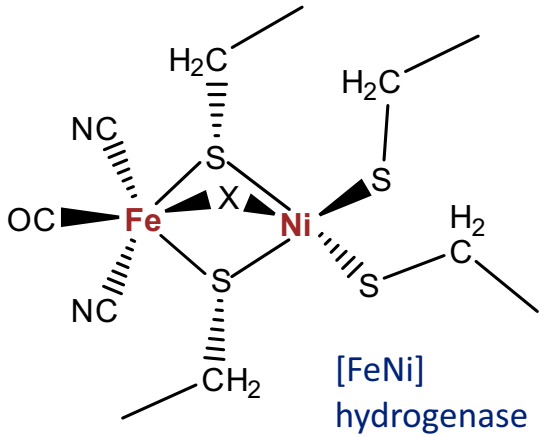
- electron transfer reactions in cell energetics and metabolism



tyrosyl radical

Metalloproteins

- electron transfer reactions and catalysis of reactions



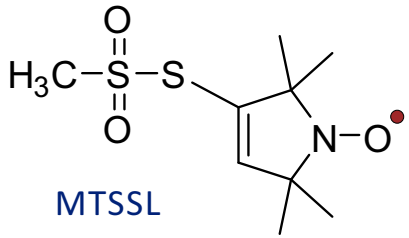
[FeNi] hydrogenase

Electronic structure, identity of nuclei, proton coordinates

Engineered

Spin labels

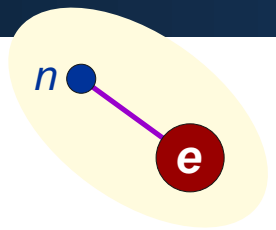
- studies of structure and dynamics on diamagnetic macromolecules



MTSSL

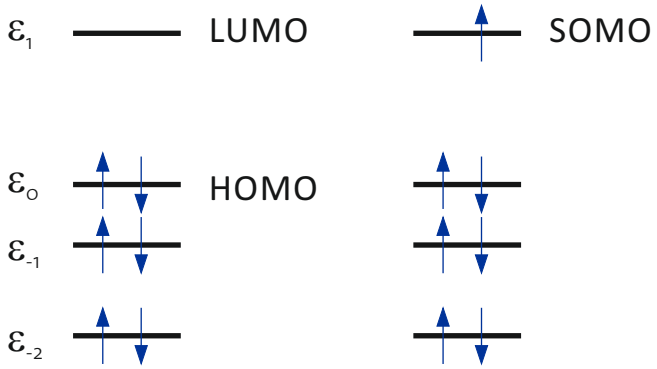
Probing the environment

Nanometer-range distance distributions

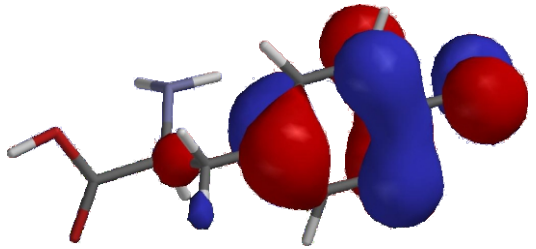


What is a SOMO?

Singly occupied molecular orbital



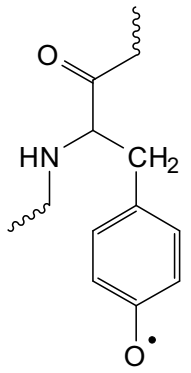
Visualization of the SOMO of a tyrosyl radical



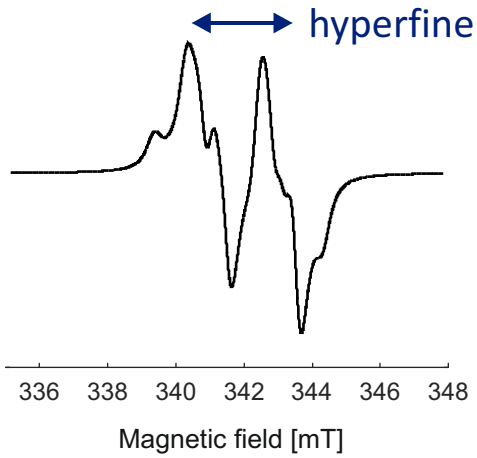
The SOMO can be probed by hyperfine couplings to nuclei

- ^{1}H hyperfine couplings related to spin density on the adjacent heavy atom
- *g* tensor related to global properties of the SOMO *via* spin-**orbit** coupling

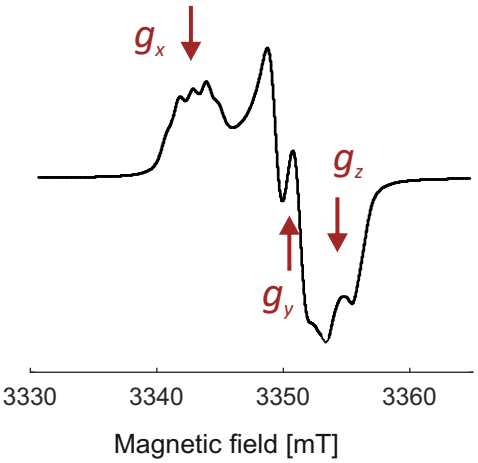
The SOMO and free radical EPR spectra

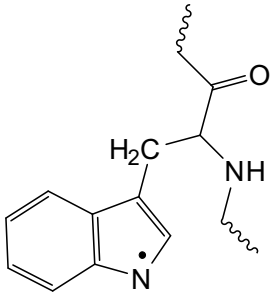


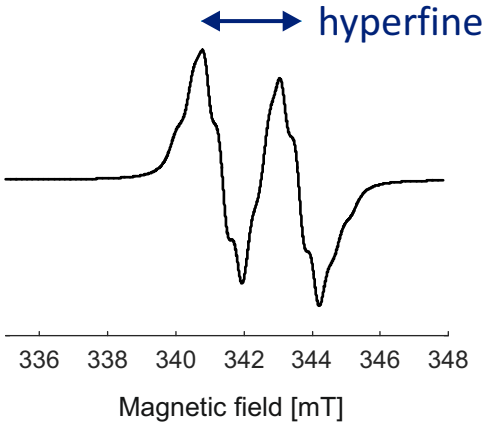
X band: 9.5 GHz

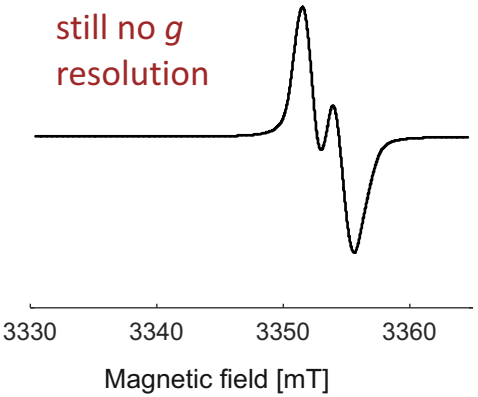


W band: 94 GHz





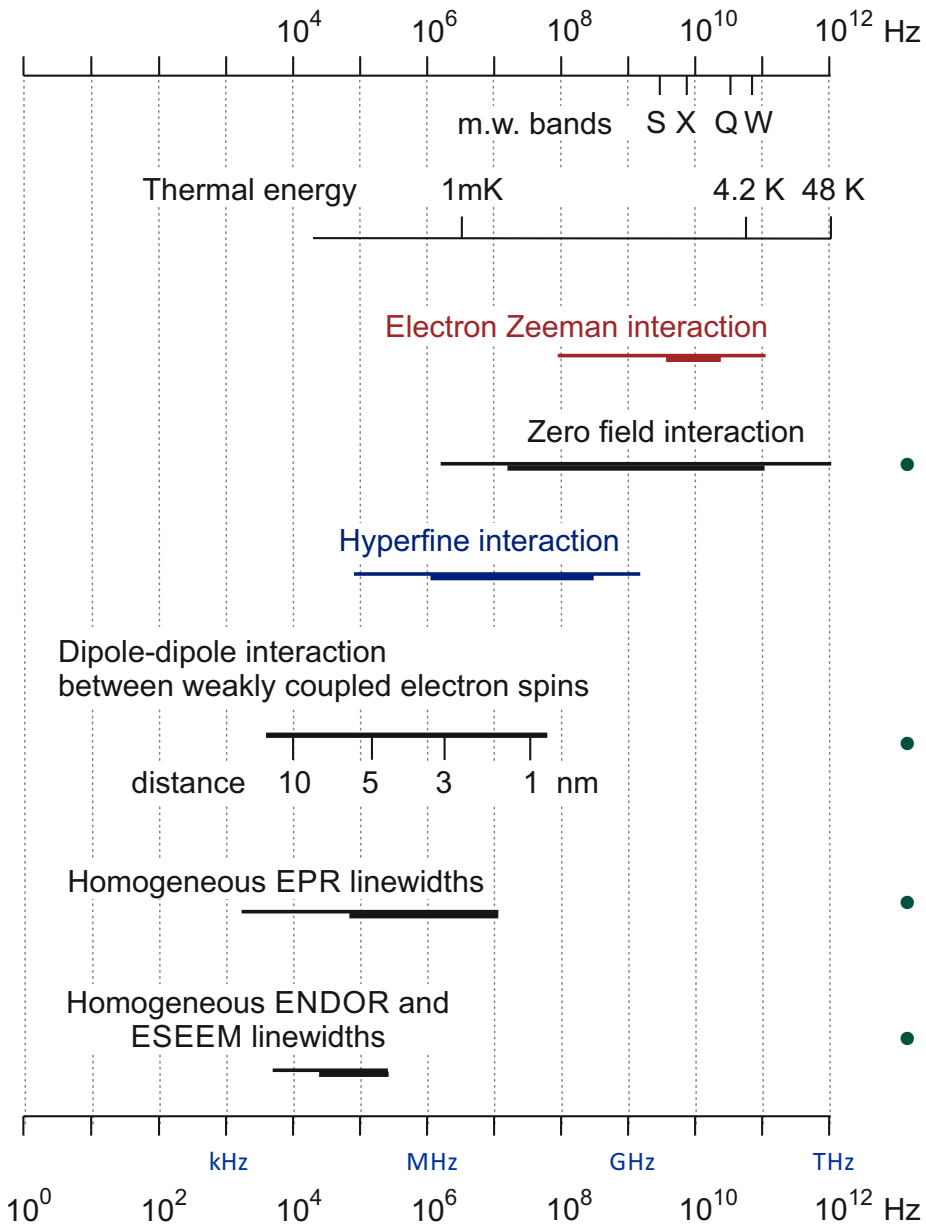




hyperfine-dominated

better *g* resolution, worse hyperfine resolution

An overview of microwave bands and interactions

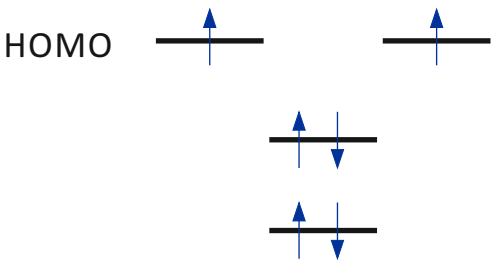


S band 2-4 GHz
X band 9-10 GHz
Q band 33-35 GHz
W band 94-95 GHz

- for electron group spin > 1/2 (more than one unpaired electron)
- most valuable source of EPR restraints on structure
- resolution limit depends on sample preparation
- hyperfine resolution better when measuring on the nuclear spin

More than one unpaired electron

Triplet state (*S* = 1)

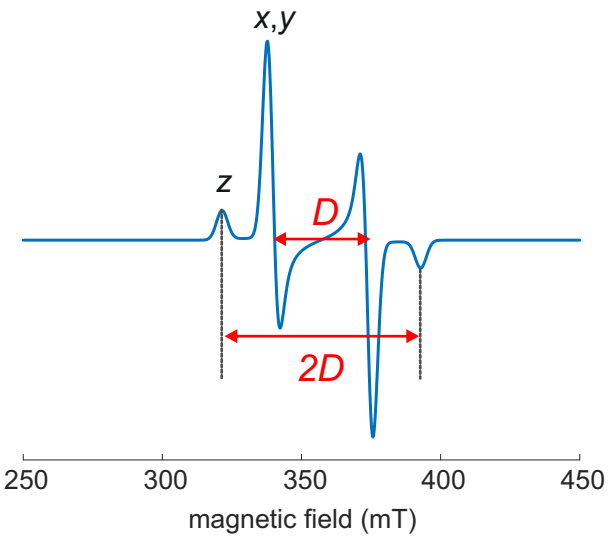


- the two magnetic moments couple through space (dipole-dipole coupling)
- they are both spatially distributed in their orbitals
- this causes **zero-field splitting** (typically 300 MHz… 2 GHz)
- for heavier atoms, especially transition metals, zero-field splitting has a spin-orbit contribution

chlorophyll, carotenoids

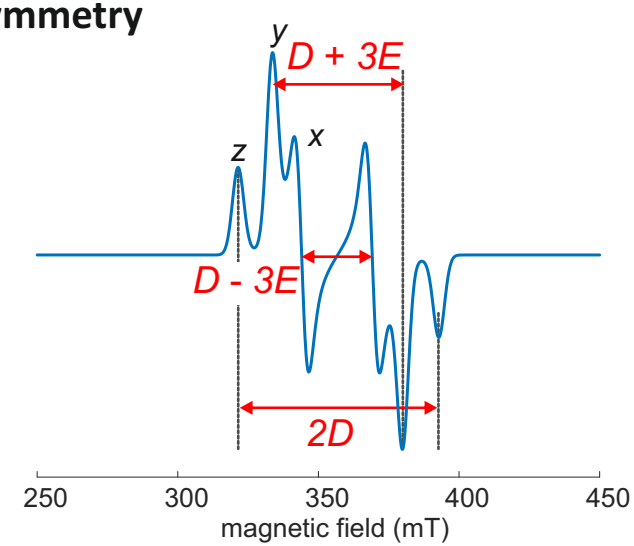
Axial symmetry

D = 1000 MHz
E = 0 MHz



Orthorhombic symmetry

D = 1000 MHz
E = 100 MHz

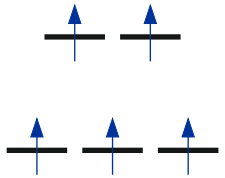


More than one electron in metal centers

Bare $3d^5$ metal ion (Fe^{III}, Mn^{II})

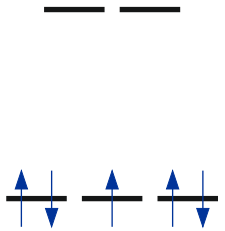
S = 5/2 (Hund's rule)

Weak ligand field



S = 5/2 (high spin)

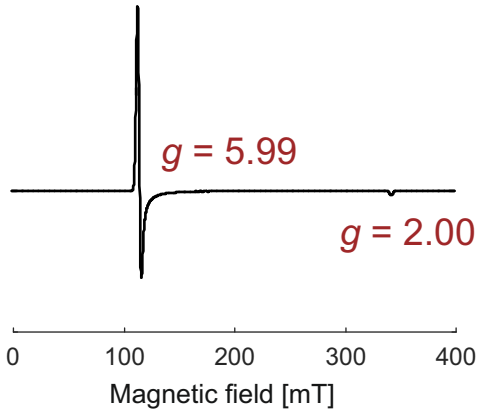
Strong ligand field



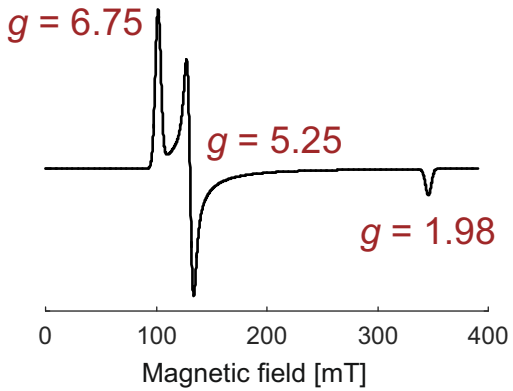
S = 1/2 (low spin)

High-spin Fe^{III}: large ZFS and effective g values

axial symmetry

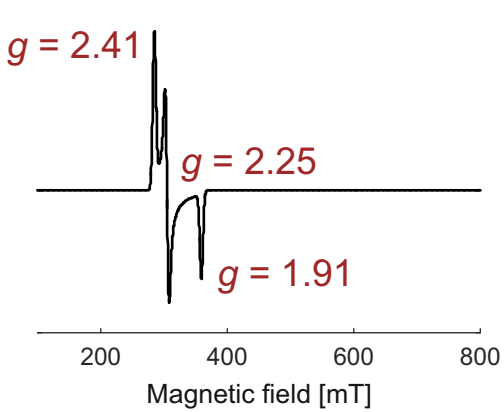


orthorhombic symmetry

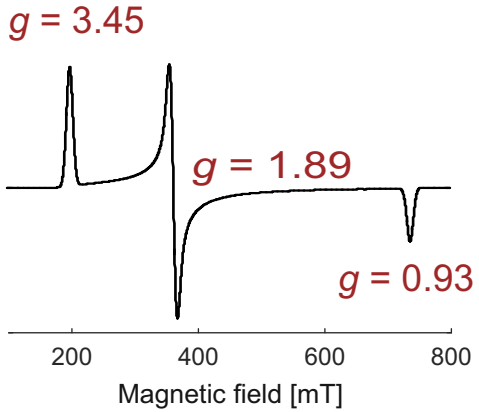


Low-spin Fe^{III}: smaller g dispersion

Type II ($P450_{cam}$)

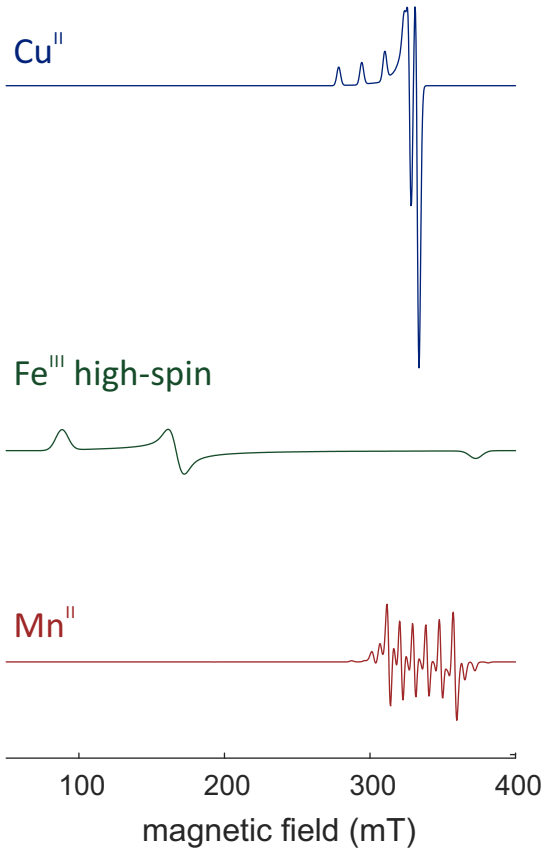


Type I (Myoglobin-CN)



Fingerprinting metal ions

X-band EPR of common metal ions



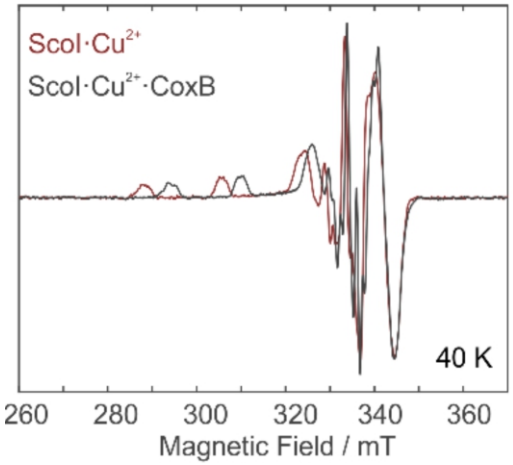
$S = 1/2$, $I = 3/2$

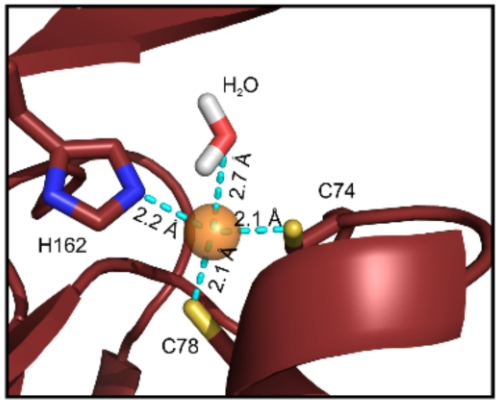
$S = 5/2$
ZFS >> Electron Zeeman

$S = 5/2$, $I = 5/2$
ZFS << Electron Zeeman

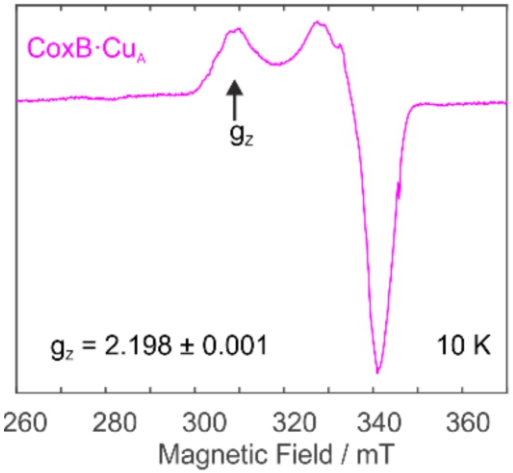
- the half-filled shell (3d5) of Mn^{II} makes the g tensor and hyperfine coupling almost isotropic

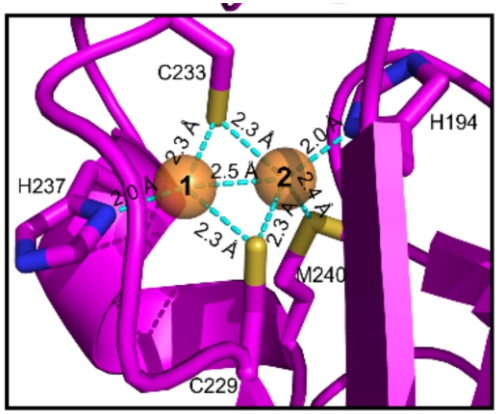
Different types of Cu^{II} centers





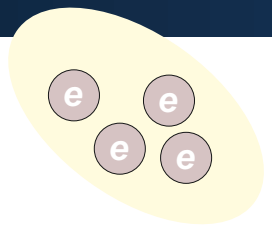
CXXXC...H motif





H...CXXXCXXXHXXM motif

F. CANONICA *et al.* Science Advances



Some are invisible

Kramers and non-Kramers ions

Denn die einen sind im Dunkeln
Und die andern sind im Licht.
Und man siehet die im Lichte
Die im Dunkeln sieht man nicht.

Bertolt Brecht

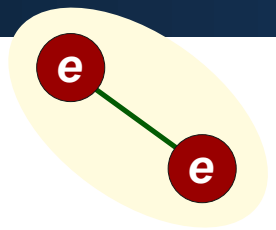
Large ZFS may split all levels by more than the microwave frequency

- to first order ZFS contribution is proportional to m_S^2
- for half-integer spin (S = 1/2, 3/2, 5/2, 7/2), there are $\pm m_s$ pairs of levels that are degenerate in zero field: **Kramers ions**
- whatever spectrometer you have, there is a field/frequency combination where you can excite transitions of Kramers ions
- for integer spin (S = 1, 3, 5), all levels are split to first order by ZFS at zero field (unless symmetry is axial): **non-Kramers ions**
- if ZFS is larger than microwave frequency plus electron Zeeman interaction at maximum field, no transition can be excited for non-Kramers ions

Non-Kramers ions may be "EPR silent"

- typical cases: **Fe**^{II} ($3d^6$, S = 2), **Co**^{III} ($3d^6$, S = 2), **Ni**^{II} ($3d^8$, S = 1) in their high-spin states
- low-spin states of ions with an even number of unpaired electrons are diamagnetic (S = 0)

$\Rightarrow$ **usually, metal ions are only seen when they have an odd number of unpaired electrons**



Weakly coupled electron spins

Exchange coupling

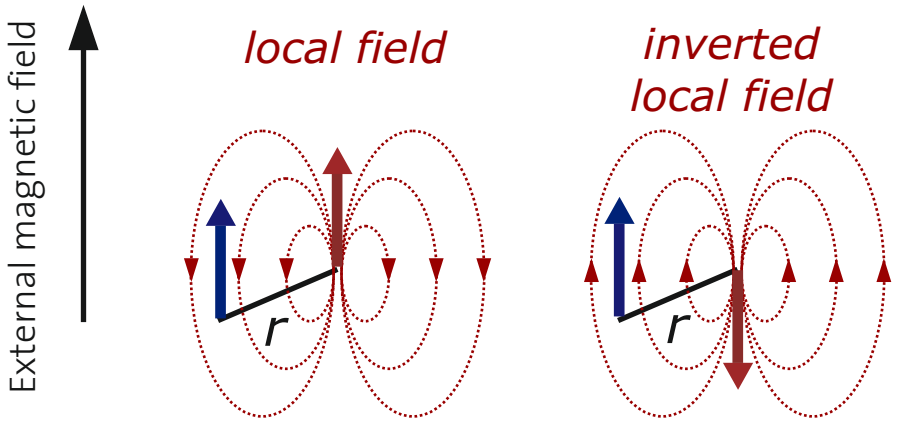
- arises from overlap of the SOMO's of two electrons
  - binding overlap ↔ antiferromagnetic coupling ↔ $\Delta E_{\alpha\beta}=\Delta E_{\beta\alpha} < \Delta E_{\alpha\alpha}=\Delta E_{bb}$
  - anti-binding overlap ↔ ferromagnetic coupling ↔ $\Delta E_{\alpha\beta}=\Delta E_{\beta\alpha} > \Delta E_{\alpha\alpha}=\Delta E_{bb}$

- strong exchange coupling ($J > g\mu_B B_o/\hbar$)
  - antiferromagnetic: diamagnetic singlet ground state
  - ferromagnetic: paramagnetic triplet ground state

Exchange coupling decreases exponentially with distance *r*

Unless orbitals strongly overlap, exchange coupling is negligible at *r* > 15 Å

Dipole-dipole coupling

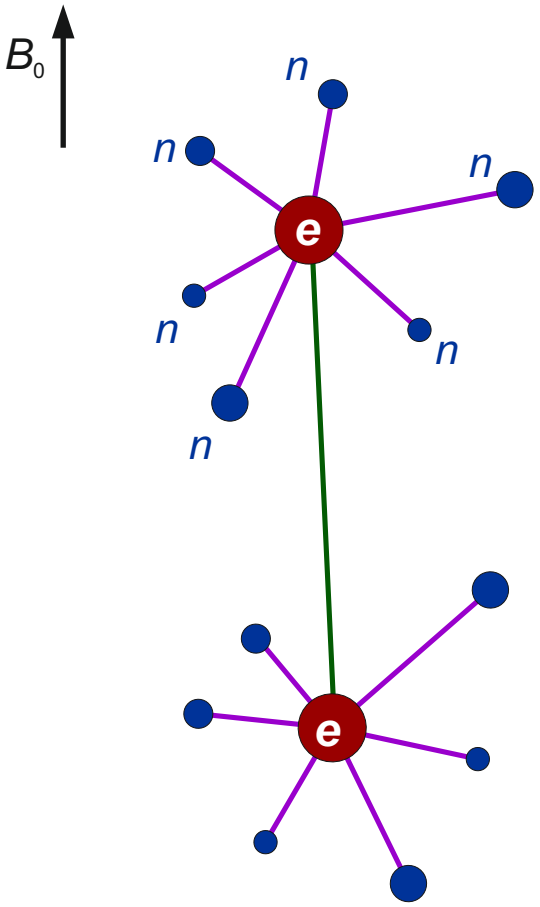


for weak *g* anisotropy

$$\omega_{dd} = \frac{1}{r_{12}^3}\frac{\mu_0}{4\pi\hbar} g_1^2 g_2^2 \mu_B^2$$

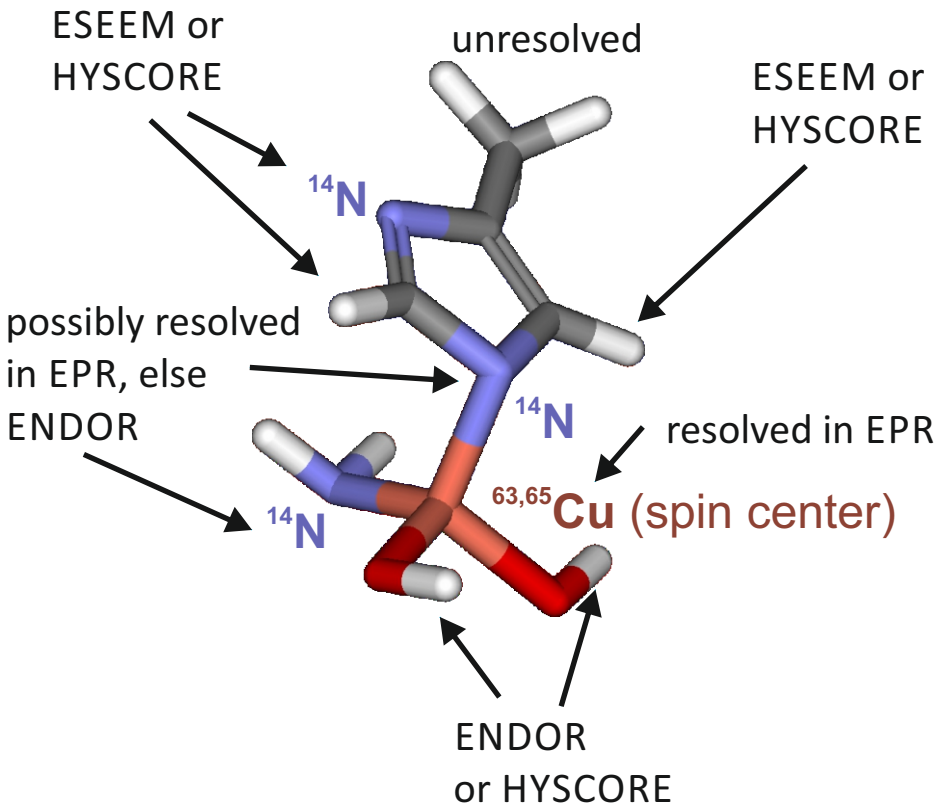
$\omega_{dd}/2\pi \approx 52.04$ MHz at r_{12} = 1 nm

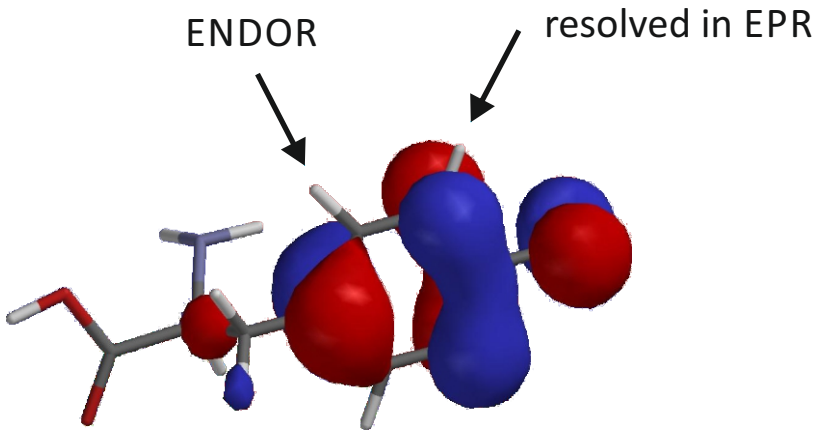
Interactions and the information that they provide



Name	Information
electron Zeeman	**fingerprinting** of radical type or metal coordination
hyperfine	distribution of the SOMO (**reactivity**) **distance of protons** from the center of spin density
nuclear Zeeman	i**dentification of nuclei** that give rise to hfi
nuclear quadrupole	**binding situation** of the nucleus for $I > 1/2$ (chemical shift is not available)
zero-field	**fingerprinting** of triplet type or metal coordination spin state for metal ions (**low or high spin**)
exchange	**orbital overlap** (important for **electron transfer**)
dipole-dipole	distances in the nanometer range (15 - 100 Å) $\Rightarrow$ **structure**

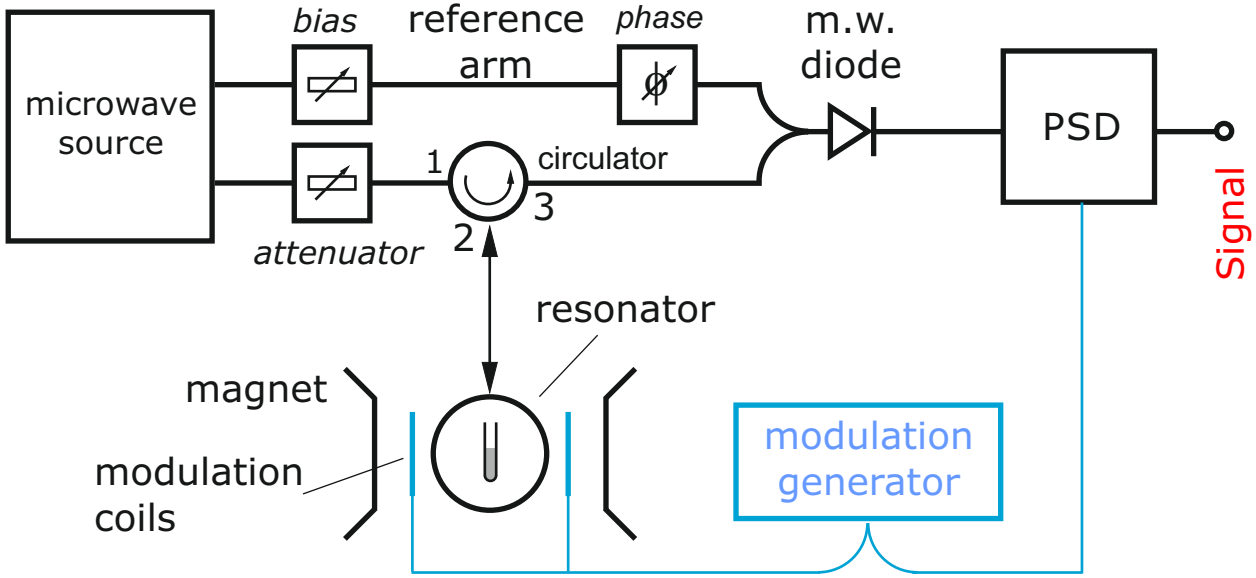
Measuring hyperfine couplings



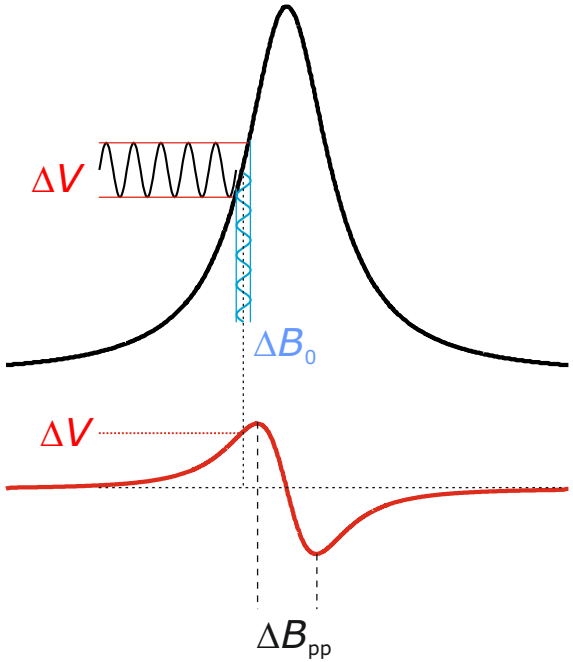


oxygen is normally invisible, but can be made visible with ^{17}O if the problem justifies the expense

What is CW EPR?



Field modulation



Points to remember

- signal increases linearly with modulation amplitude, until it starts to broaden (use 2 G amplitude at the beginning)
- signal increases proportionally to the square root of microwave power (factor 2 per 6 dB less attenuation) until it starts to broaden, level off, and eventually to *decrease* again (use 20 dB attenuation at the beginning)

When can and should CW EPR be applied?

CW EPR is the first experiment to be applied to any unknown sample

Hardware requirements: basic CW EPR spectrometer (widely available, cheap)

Sensitivity : radicals >1 μM to 10 μM
metal ions >10 μM to 100 μM

Aggregation state : liquid & solid

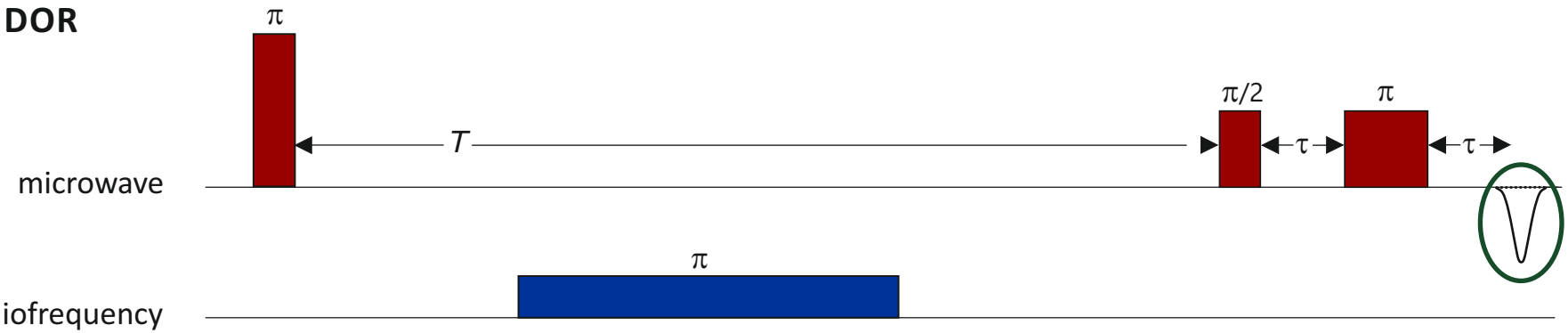
Special requirements : liquid polar solvents (aqueous buffer) require special sample geometry for best sensitivity
(flat cells or bundles of capillaries)
if utmost sensitivity is not an issue, a capillary will do nicely

Information : type of paramagnetic center (may require high field)
large hyperfine couplings
rough idea on relaxation by playing with microwave attenuation
spin quantification (comparison of double integral with the one of a reference sample)

What is ENDOR?

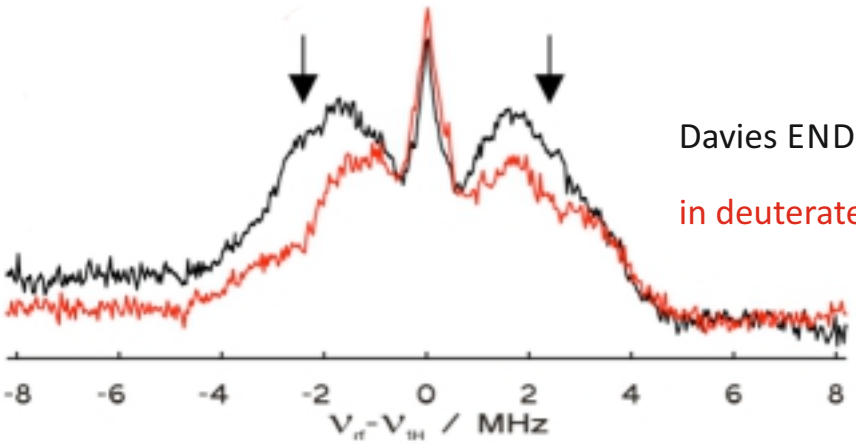
Electron nuclear double resonance

Davies ENDOR



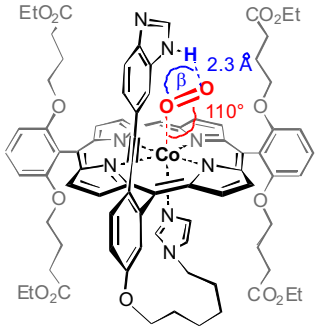
"frequency-swept NMR"

The **echo amplitude** is observed as a function of **radio frequency**

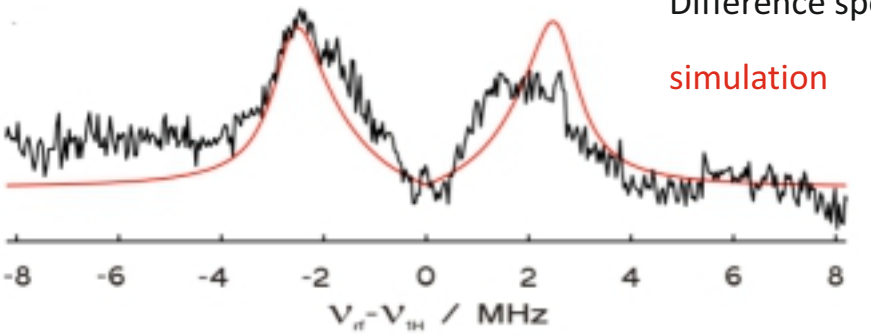


Davies ENDOR in water

in deuterated water



hemoglobin
model compound



Difference spectrum

simulation

When can and should ENDOR be applied?

ENDOR is applied if hyperfine couplings are unresolved in CW EPR and too large for ESEEM/HYSCORE

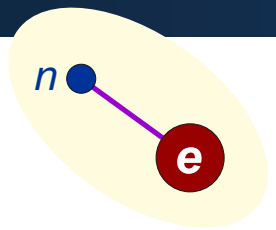
Hardware requirements: pulse EPR, radiofrequency channel

Sensitivity : radicals >50 μM to 200 μM
metal ions >200 μM to 1 mM

Aggregation state : solid (liquid state requires rarely available CW ENDOR)

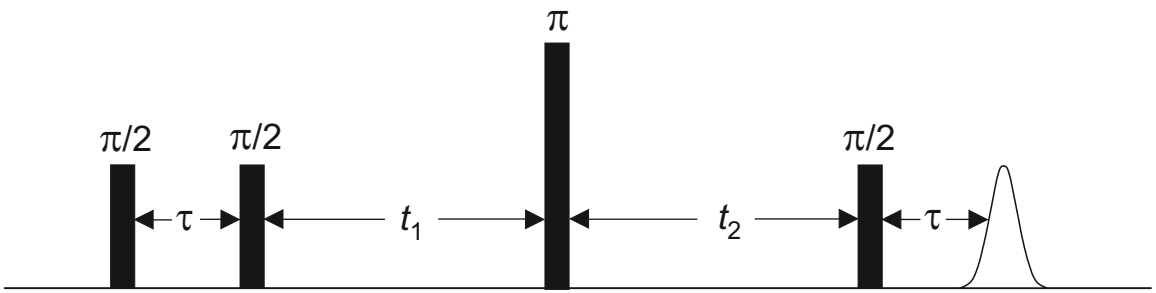
Special requirements : longitudinal relaxation time of at least 10 μs
signals of different isotopes overlap at X band, high field may be required in some cases

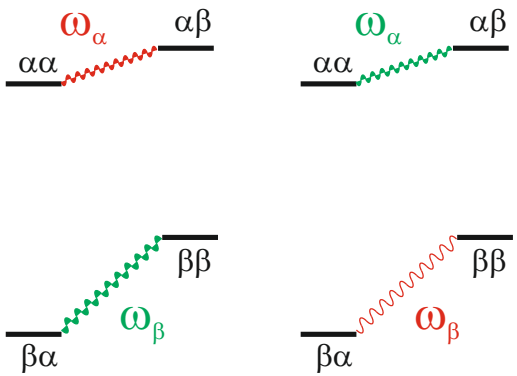
Information : large and moderately sized hyperfine couplings
nuclear Zeeman frequency
nuclear quadrupole coupling



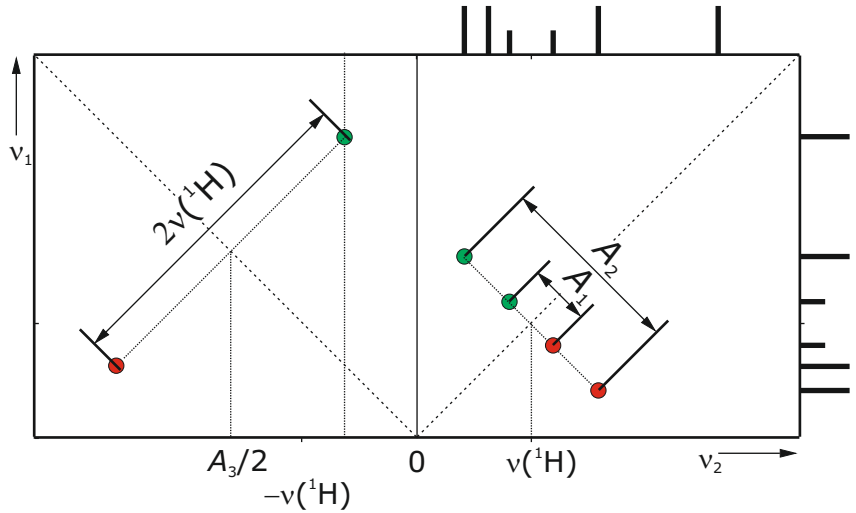
What is HYSCORE?

Hyperfine sublevel correlation

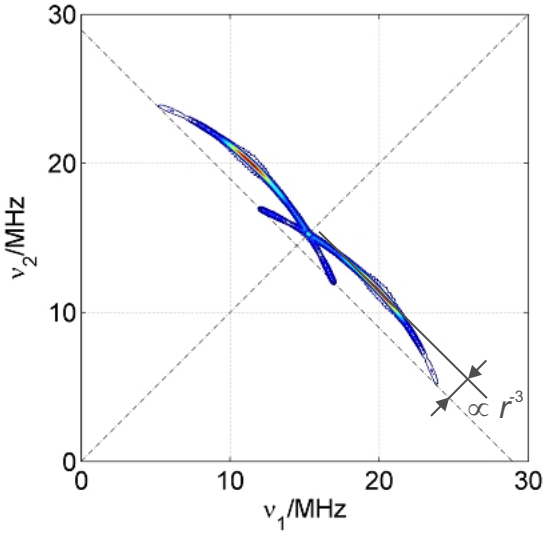




large couplings small couplings



curved correlation ridges
contain information
on hyperfine anisotropy
⇒ **^{1}H-electron spin distance**



- correlates frequencies of the same nucleus for the α and β state of the electron spin
- the 1D version without the π pulse is called 3-pulse ESEEM

When can and should HYSCORE/ESEEM be applied?

HYSCORE is applied if hyperfine couplings are unresolved in CW EPR

Hardware requirements: pulse EPR

Sensitivity : radicals >50 μM to 200 μM
metal ions >200 μM to 1 mM

Aggregation state : solid

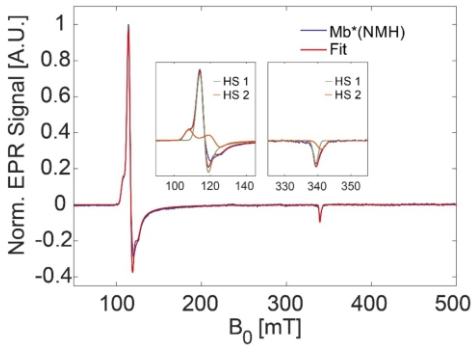
Special requirements : transverse relaxation time of at least 100 ns
anisotropic hyperfine couplings
hyperfine coupling of the same order of magnitude as twice the nuclear Zeeman frequency

Information : small and moderately sized hyperfine couplings
nuclear Zeeman frequency
nuclear quadrupole coupling
separation of isotropic and anisotropic hyperfine contributions (^{1}H distances)

Haem-carbene bond in an artificial enzyme

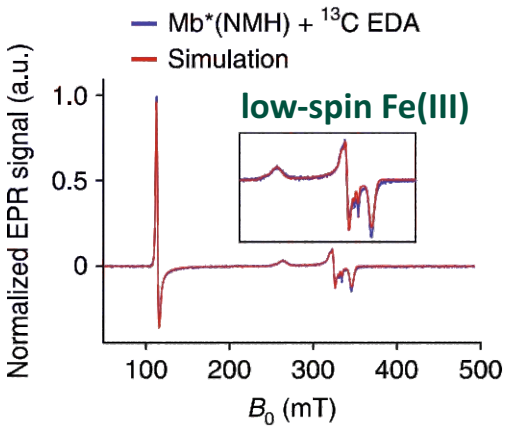
Hyperfine coupling detects Fe-C bond in solution

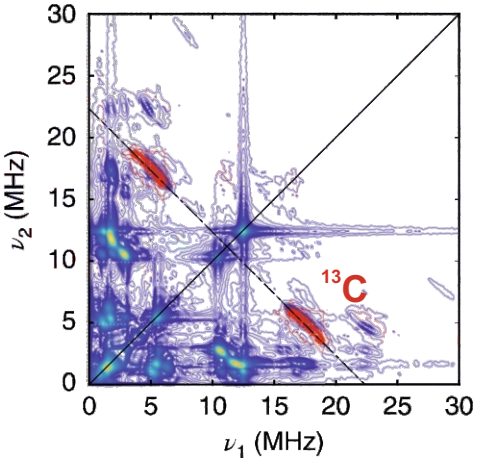
CW EPR before adding the substrate EDA



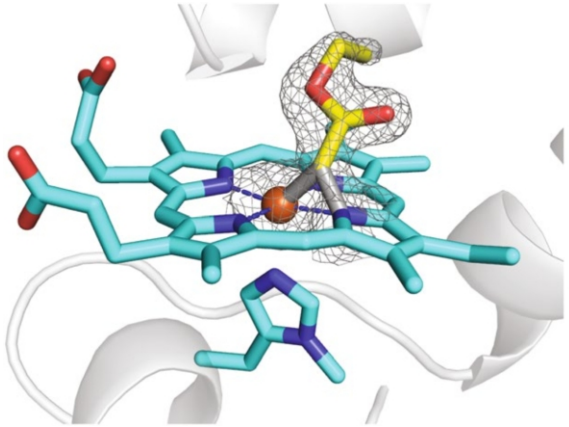
high-spin Fe(III)

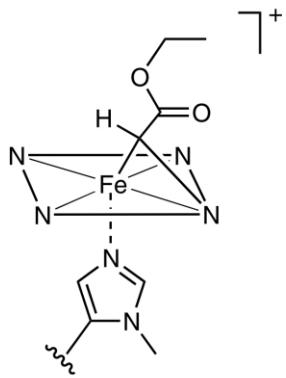
After adding the substrate EDA





Intermediate captured in a crystal



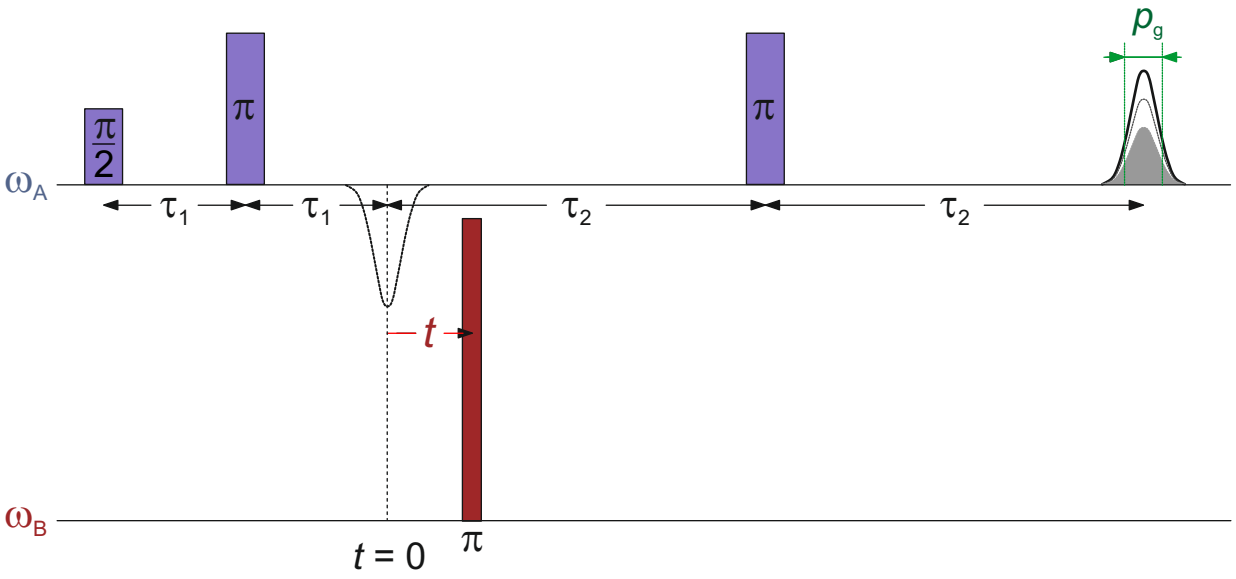


the same spectrum contains further information on electronic and spatial structure:

- ^{14}N hyperfine and nuclear quadrupole couplings (lower frequencies)
- ^{1}H hyperfine couplings (signals not shown)

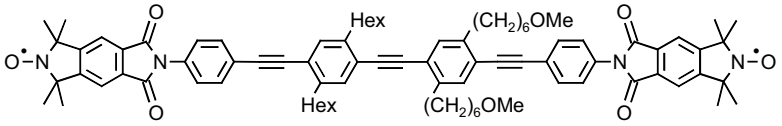
T. HAYASHI *et al.* Nat. Catalysis **2018**, *1*, 578-584

What is DEER?

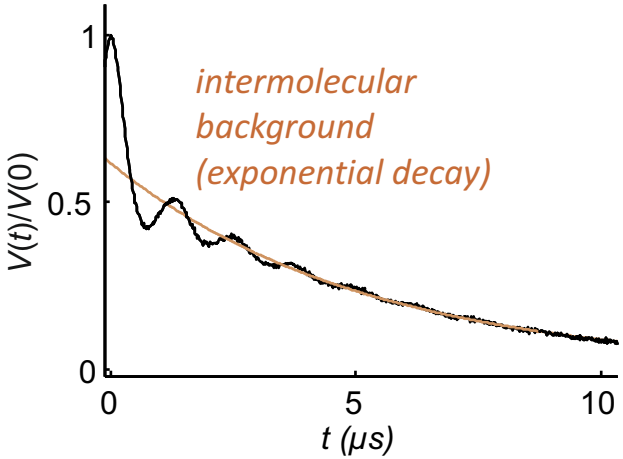


The **echo amplitude** is observed as a function of **time *t***

Model compound

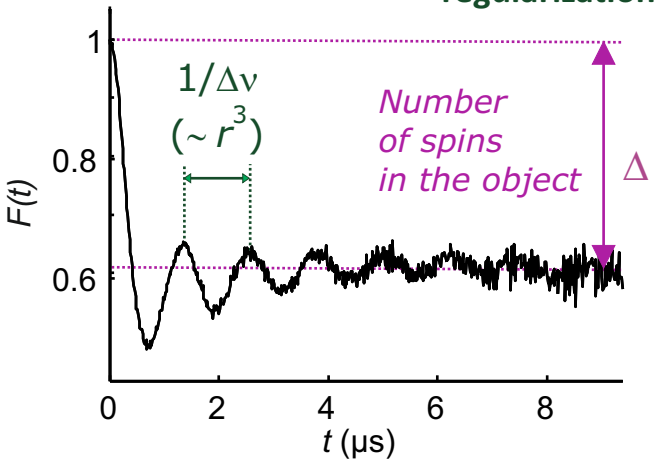


Primary data

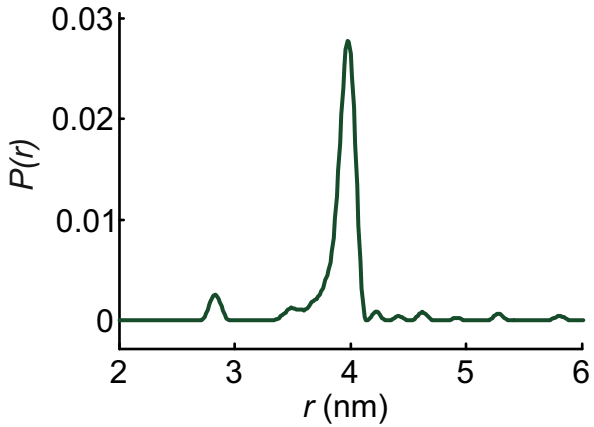


Form factor

Tikhonov regularization



Distance distribution



MARTIN RE et al., *Angew. Chem. Int. Ed.* **1998**, *37*, 2834

PANNIER M, VEIT S, GODT A, JESCHKE G, SPIESS HW, *J. Magn. Reson.* **2000**, *142*, 331

JESCHKE G et al. *J. Magn. Reson.* **2002**, *155*, 72

JESCHKE G et al. *Appl. Magn. Reson.* **2006**, *30*, 473

When can and should DEER be applied?

DEER is applied to measure distances in the range from 15 Å up to 60 (membrane proteins) or even 100 Å
(world record at 160 Å in fully deuterated GroEL)

Hardware requirements: pulse EPR, second microwave frequency (ELDOR) or arbitrary waveform generator (AWG)

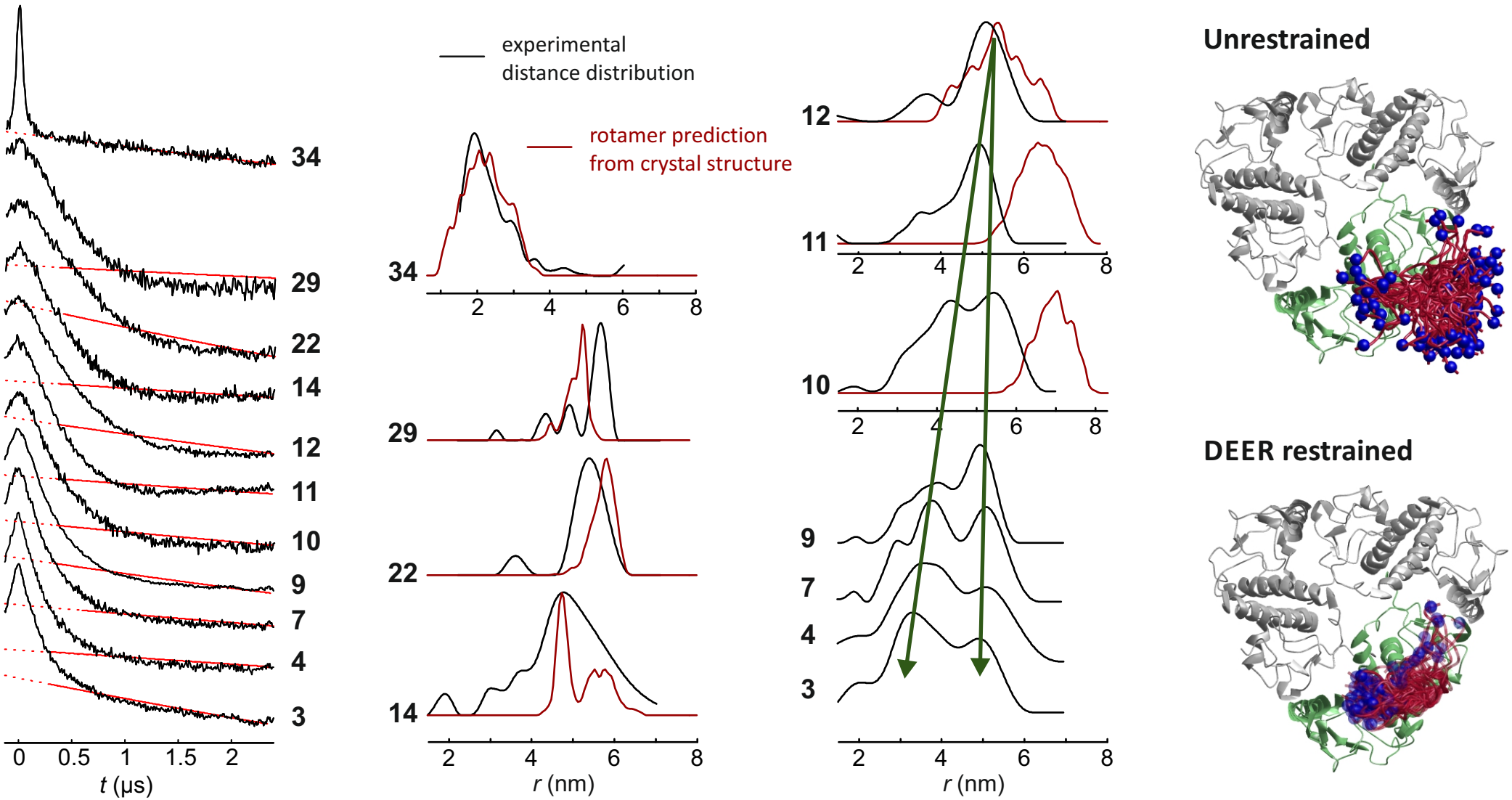
Sensitivity : radicals >10 μM to 100 μM
metal ions >50 μM to 0.5 mM

Aggregation state : solid

Special requirements : transverse relaxation time of at least 500 ns (unless distance is very short)
absence of exchange coupling for straight distance determination
orientation of spin-spin vector to magnetic field uncorrelated to spectral selection (for straight distance analysis)

Information : distance distributions or, at the long limit, mean distances between electron spins
number of spins in the same macromolecule or complex
orientation of the spin-spin vector in the molecular frame (high field, larger effort)

DEER example: Localization of the N-terminal domain in LHCII



N. Fehr *et al.* J. Biol. Chem. **2015**, *290*, 26007-20

The art of sample preparation

Concentration

- too high concentration in liquid state : exchange broadening (stay < 1 mM… 200 μM for radicals, < 2 mM for metal ions)
- too high concentration in solid state : dipolar broadening, shorter phase memory time (stay below 200 μM/1 mM)
- at very high (local) concentration, hyperfine structure may collapse (exchange narrowing)

Oxygen

- $^3O_2$ is a paramagnetic line-broadening agent, especially in unpolar solvents, detergent micelles, and lipid bilayers
- weaker effects in the solid state, but relaxation times may shorten

Cryoprotectant

- biomacromolecules don't like ice crystals, structure distortion and precipitation may occur
- 10% glycerol may suffice for liposomes, 25% for soluble proteins, 50% makes freezing simple
- DMSO can be used for DNA/RNA

Sample freezing

- immersion of the tube in liquid nitrogen: freeze-quench to 80 K in a few seconds, limited by gas bubbles (poor heat conduction)
- immersion of the tube into iso-pentane or ethanol cooled to 120 K: freeze quench to below glass transition in shorter time
- spraying of the sample onto a silver wheel that rotates in liquid nitrogen, collection of the "snow": about 40 ms freeze time

Optimizing relaxation time for pulsed EPR

Long T_2 (T_m), but short T_1

- transverse relaxation limits resolution and pulsed EPR sensitivity
- too long T_1 requires long waiting times between experiments, optimum 100 µs to 1 ms
- T_2 attains a low-temperature limit (~50 K for radicals, ~10 K for S = 1/2 metal ions)

Prolonging the low-temperature limit of the phase memory time T_m

- nuclear spin diffusion generates fluctuating hyperfine fields ⇒ dominating phase memory loss mechanism for electron spin in the low-temperature limit
- concentration of nuclei with high gyromagnetic ratio must be reduced: deuteration helps
  use D_2O in the buffer
  use d_8-glycerol as cryoprotectant
  deuterate recombinant protein by using D_2O in the growth medium
  deuterate recombinant protein even better by feeding deuterated glucose in minimal medium
  reconstitute membrane protein into deuterated lipids (or solubilize in deutrated detergent)
- check, whether concentration limits T_m by instantaneous diffusion
  (for DEER to measure very long distances, 100 µM may be too much)
- if all is done and it still does not suffice, work in the absence of oxygen (if you can)

increasing expense and effort

Spin labeling

Labeling and sample preparation

- where to label?
- which label?
- what matrix (solvent, detergent, liposome, nanodisc, deuteration)?
- what concentration?

Measurement

- what frequency?
- how much sample?
- which experiment and what parameters?
- optimization of spectrometer and probehead

Data analysis

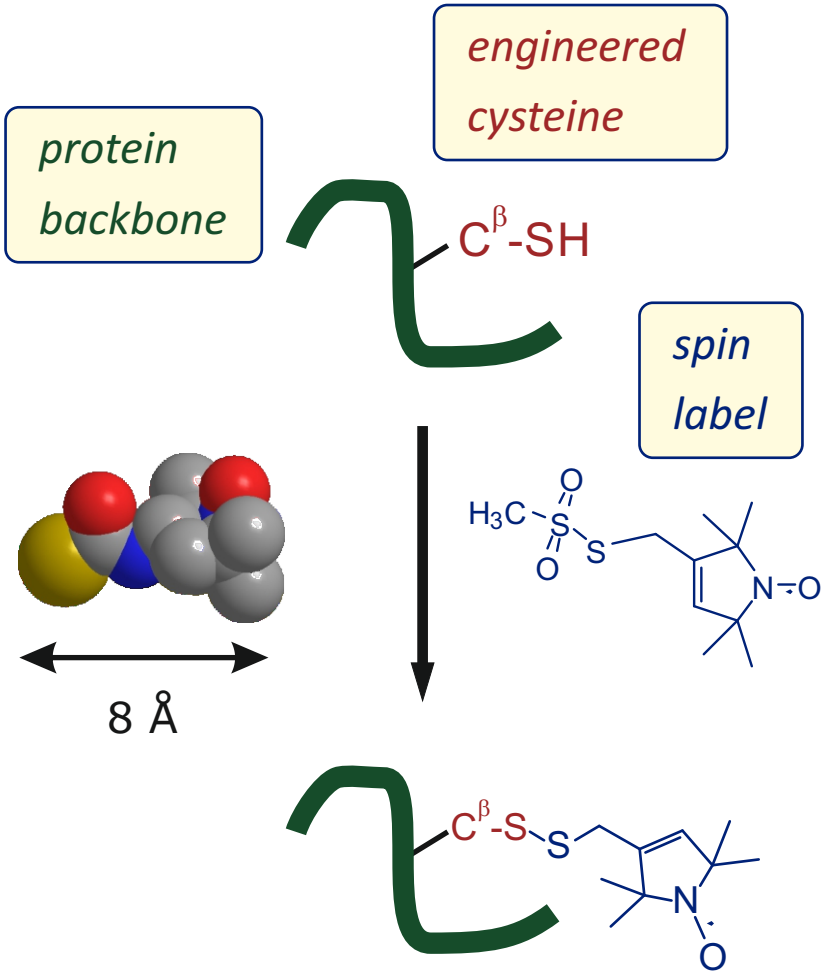
- what can be trusted?
- model-free or model-based?
- restraints and their error bars
- information beyond distances

Structure modeling

- label conformation problem
- sparse constraint problem (which approach?)
- extent of coarse graining
- uncertainty of models

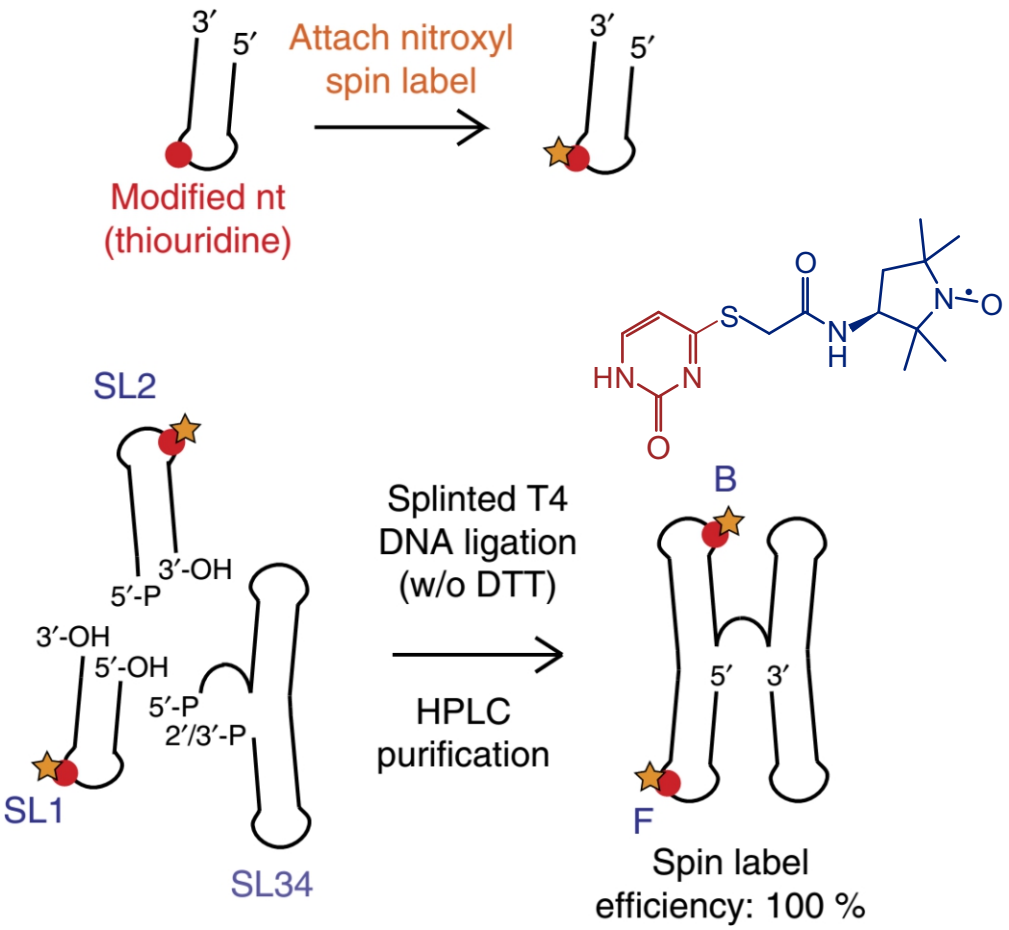
Site-directed spin labeling of proteins and RNA

Proteins



W.L. Hubbell, C. Altenbach, et al.

RNA

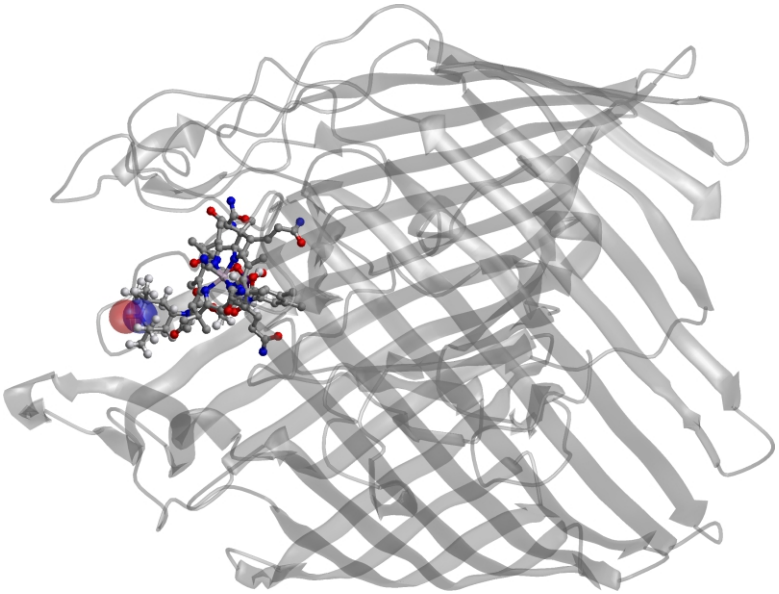


O. Duss, M. Yulikov, G. Jeschke, F. H.-T. Allain *Nature Comm.* **5**, 3669 (2014)

Alternative types of labeling

Cofactor labeling

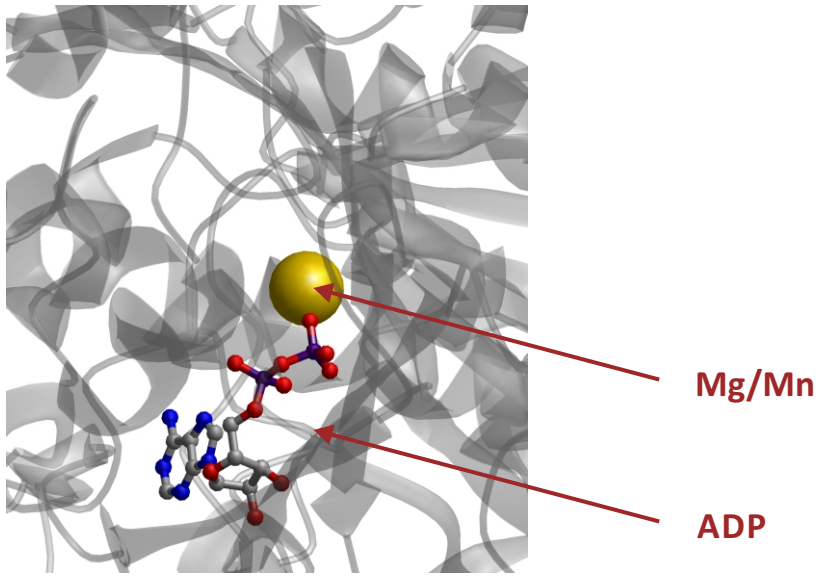
TEMPO-labeled cobalamin (vitamin B12) bound to BtuB



B. JOSEPH ET AL. *Angew. Chem. Int. Ed.* **2015**, 54, 6196 –6199

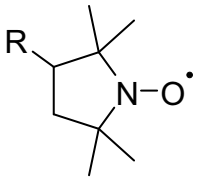
Metal ion substitution

Mn^{II} substitution for Mg^{II} in hnDnaB helicase

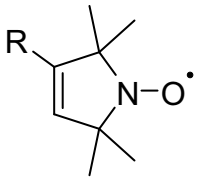


T. WIEGAND ET AL. *Angew. Chem. Int. Ed.* **2017**, 56, 3369 –3373

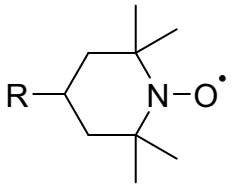
Nitroxide labels



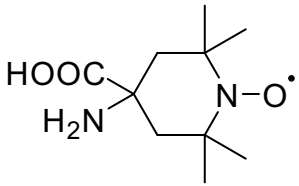
Dehydro-Proxyl



Proxyl



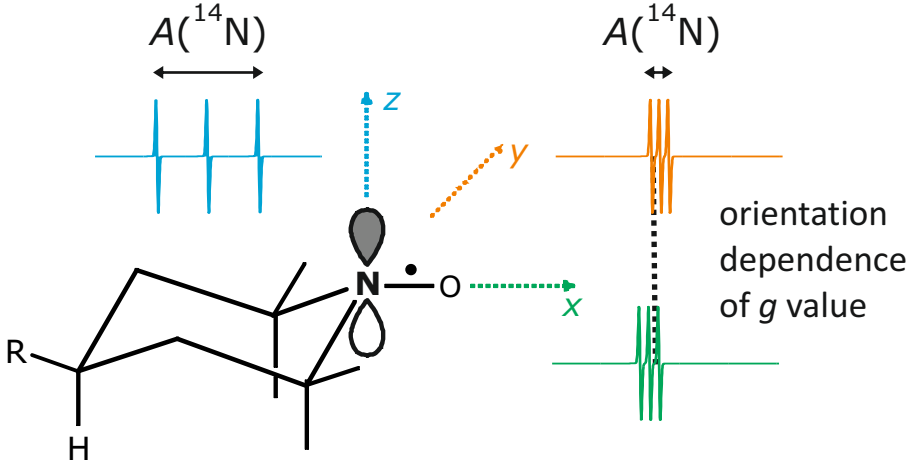
TEMPO



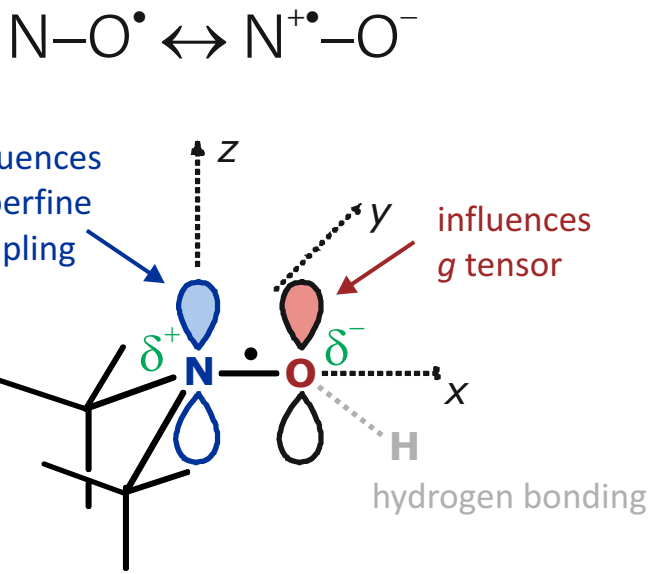
TOAC

- Proxyl preferred because of stability and relative rigidity
- methyl group replacement by ethyl or spirohexyl groups is advantageous for relaxation and stability - but tedious

The nitroxide spectrum depends on orientation...

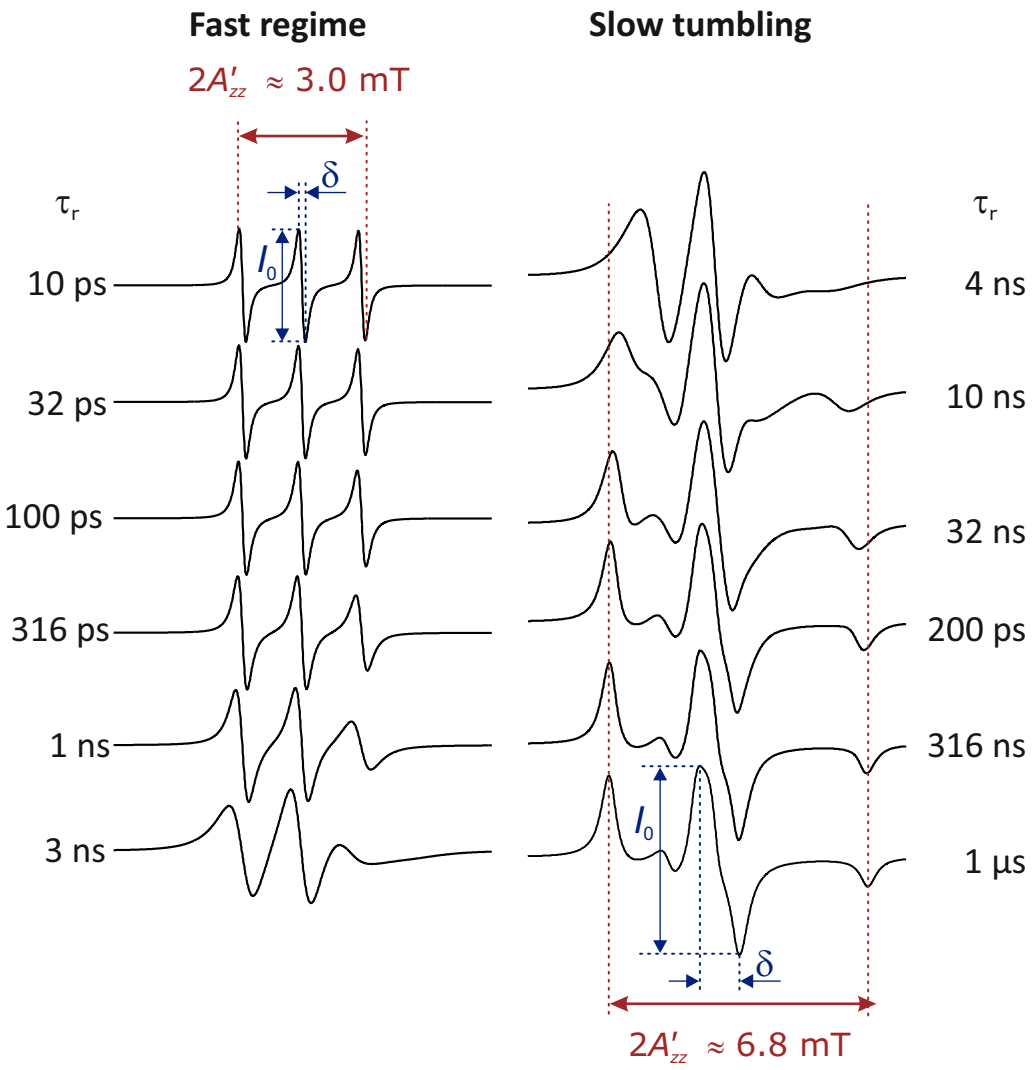


...and on polarity of the environment



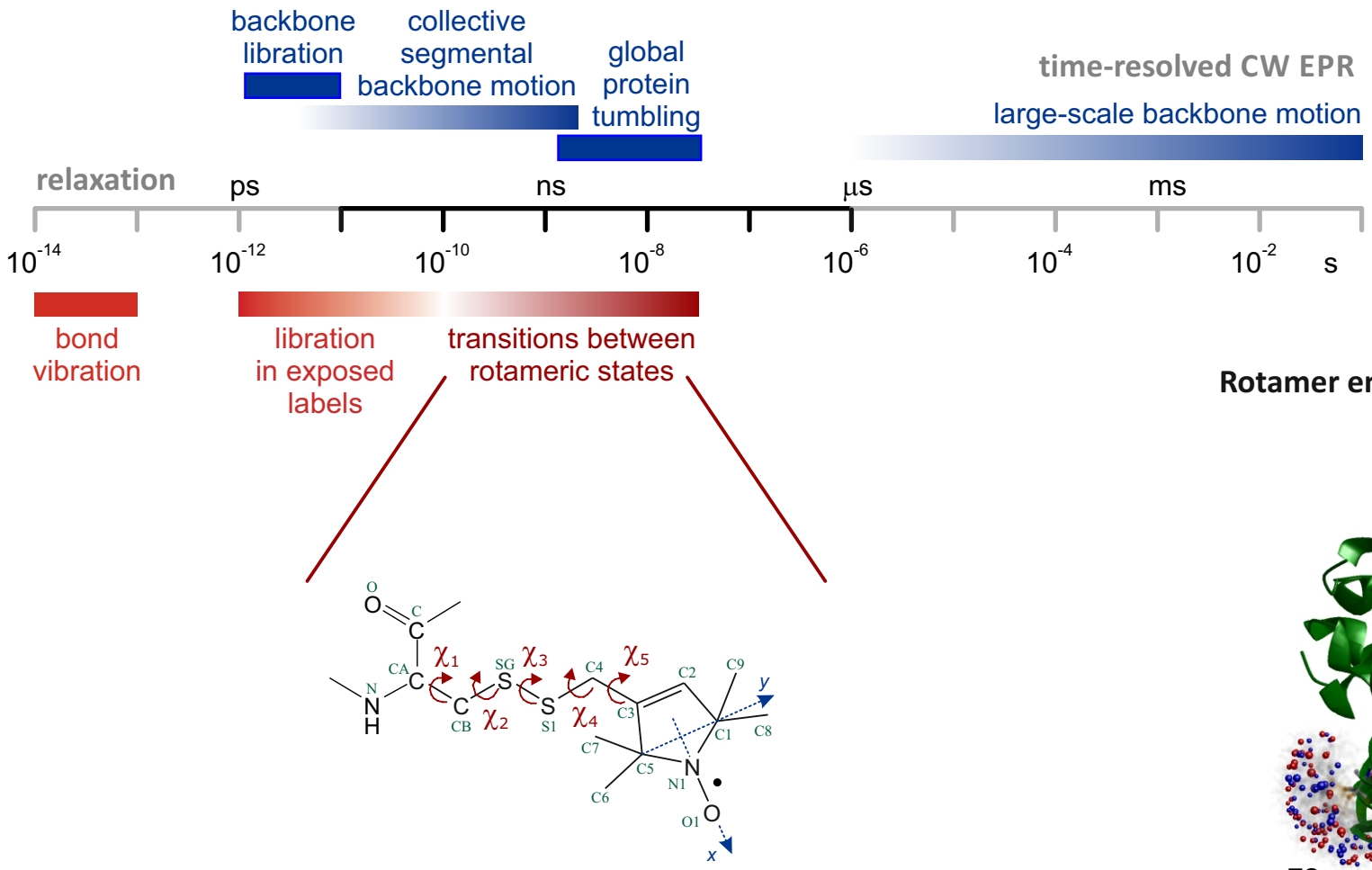
Nitroxide spectra and dynamics

X-band CW EPR spectra for isotropic Brownian rotational diffusion

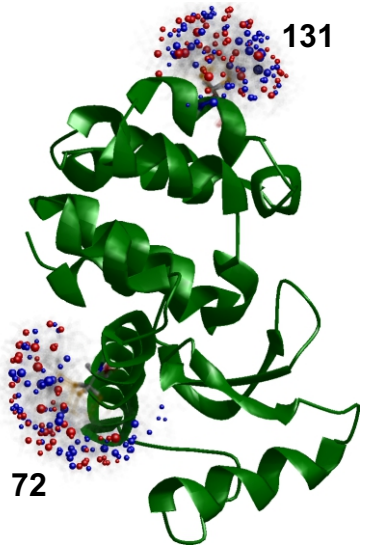


- nitroxide spectra are sensitive on the time scale of sidegroup dynamics
- the actual dynamics is more complex than isotropic rotational diffusion
- in many cases, semi-quantitative analysis in terms of spectral parameters A'_{zz} or δ suffices

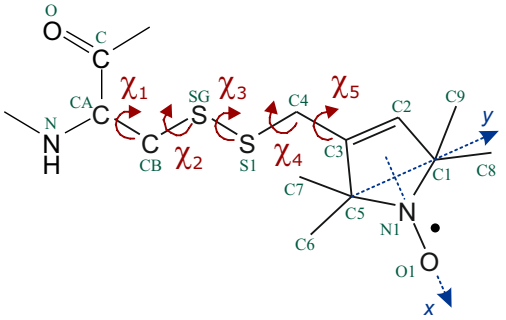
Nitroxide motion - What really happens



Rotamer ensembles in T4 Lysozyme

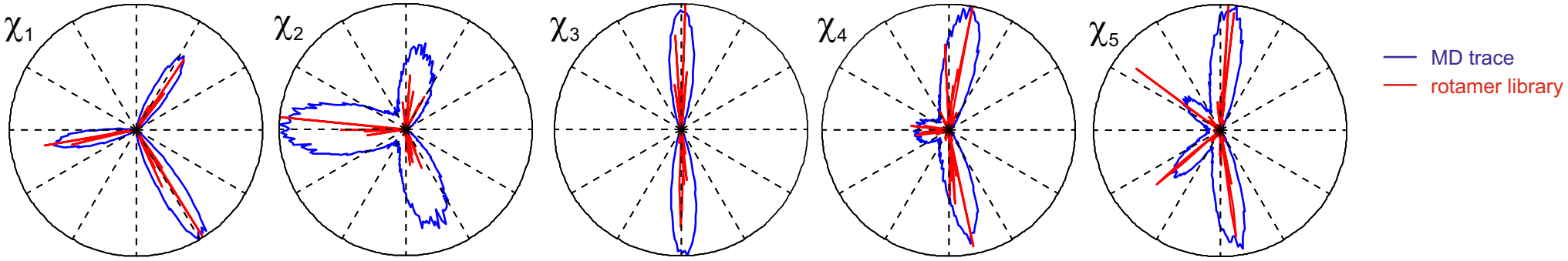


Nitroxide rotamer libraries



Spin label conformations are (semi-)discrete

MD simulation of unrestricted MTSSL spin label side chain



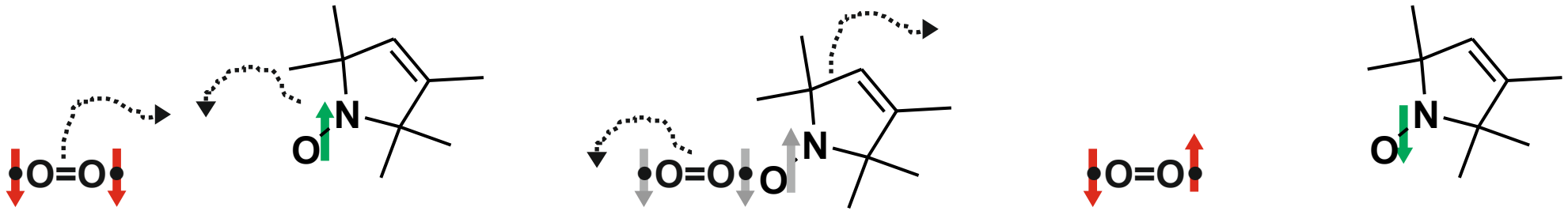
Principle of rotamer library prediction of spin label conformations

- rotamer populations for the unrestricted label $\xrightarrow{\text{Boltzmann inversion}}$ relative free energies of unbound rotamers
- + non-bonded label-macromolecule interaction from Lennard-Jones potential $\Rightarrow$ relative free energies of bound rotamers
- via Boltzmann distribution: ensemble of rotamers with populations and partition function

Paramagnetic quenchers relax nitroxides

Diffusing paramagnetic species

before **during** **and after collision**

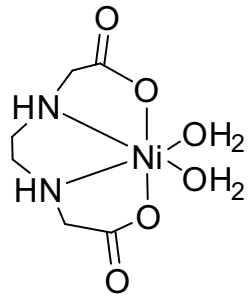


observer spin up ↑

electrons in overlapping orbitals are indistinguishable

observer spin down ↓

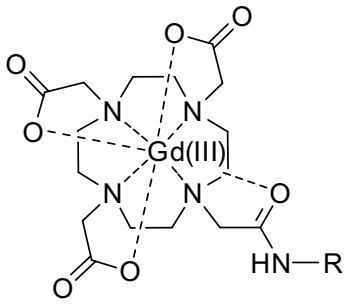
- most easily detected via change in T_1 by progressive power saturation
- at high concentration, shortening of T_2 leads to line broadening ($T_2 \leq T_1$)
- the environment (macromolecule, lipids) may shield the nitroxide from such collisions ⇒ accessibility measurements



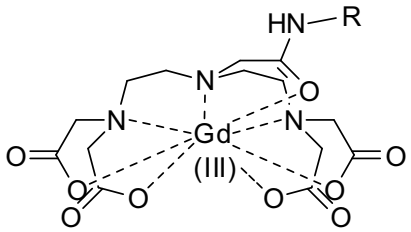
NiEDDA

ALTENBACH et al., *Proc. Natl. Acad. Sci. USA* **91**, 1667-1671 (1994)

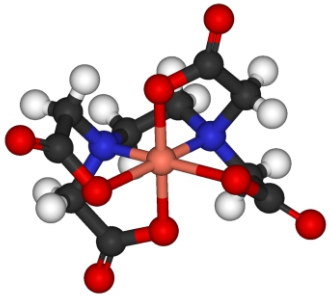
Gd^{III} and Cu^{II} labels



[Gd(DOTA)]



[Gd(DTPA)]$^{-}$

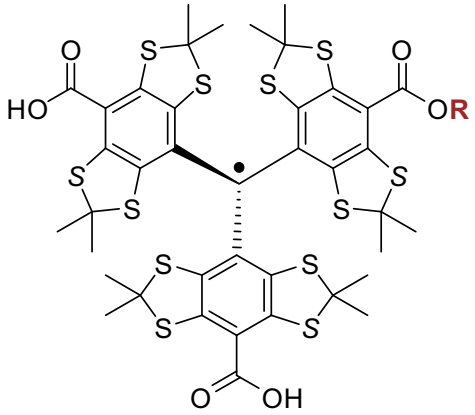


[Cu(EDTA)]$^{2-}$

- ✗ broader EPR spectra, faster relaxation
- ✗ larger size
- ✗ not suitable for assessing dynamics, polarity, and accessibility

- ✓ chemically more stable (especially Gd^{III})
- ✓ spectroscopically orthogonal

Trityl labels



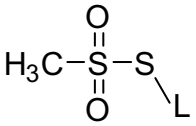
- ✗ hard to synthesize, not commercially available
- ✗ larger size than nitroxides
- ✗ not suitable for assessing dynamics, polarity, and accessibility

- ✓ chemically more stable than nitroxides
- ✓ spectroscopically orthogonal
- ✓ very narrow spectral line up to high fields

Linker chemistry for spin labels

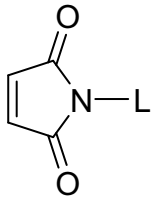
Thiol-specific linkers

MTS



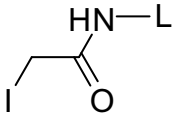
most selective, short,
but labile attachment

Maleiimide



selective at pH 6.5… 7.5
somewhat bulky

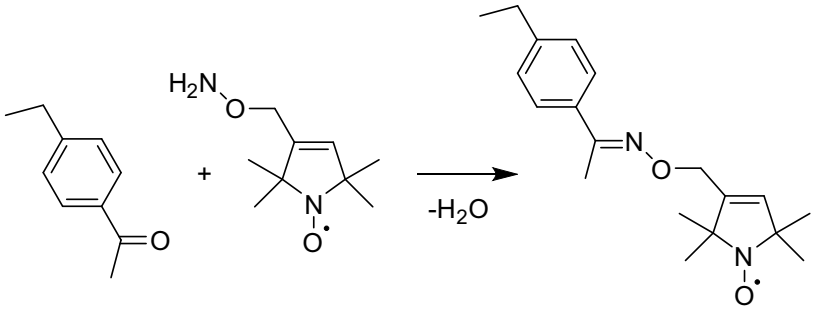
Iodoacetamide



may label primary amines
if thiol groups are inaccessible or missing

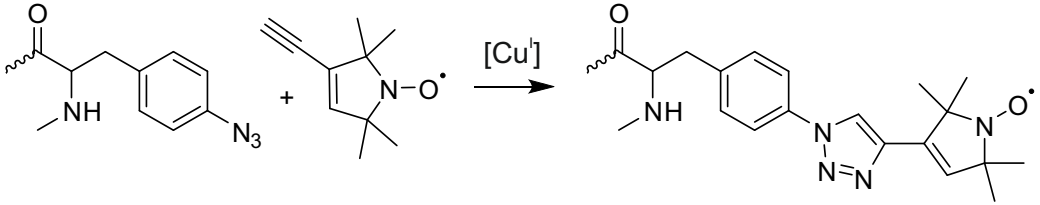
Linkers to unnatural amino acids

Ketoxime chemistry



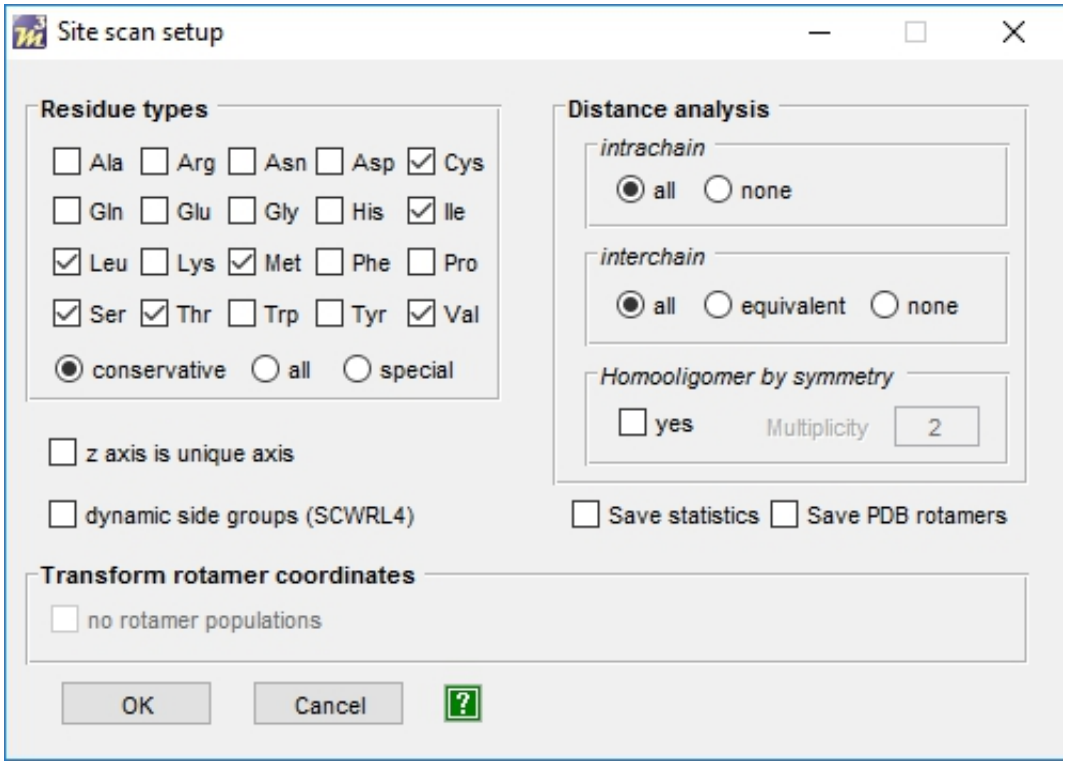
12-48 h at pH 4, not all proteins like that

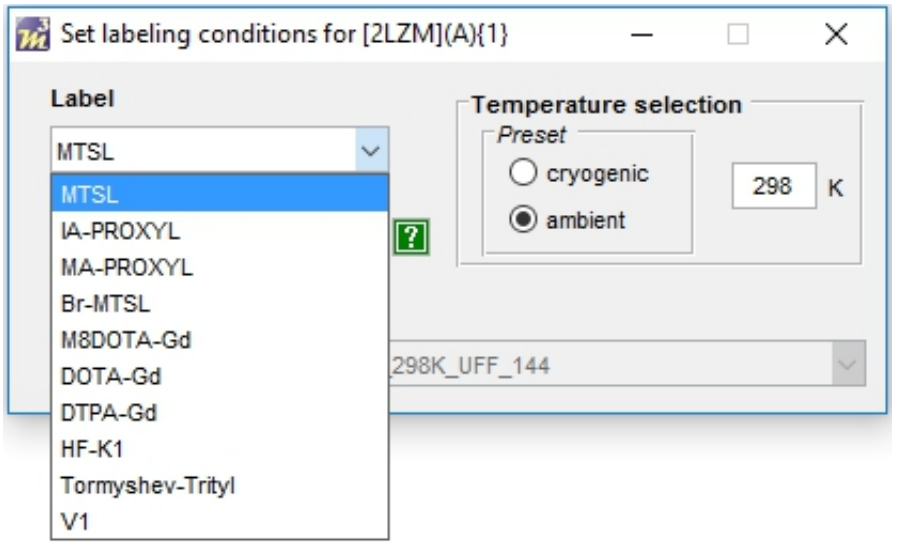
Click chemistry



catalyst may reduce nitroxide label

Choice of labeling sites and site scan





- well accessible sites with many rotamers and large partition function are preferable
- helix surface sites are often suitable

Site analysis 2LZM/A1

15 loop sites, rmsd min/mean/max 0.01/0.40/0.57 nm
37 helix sites, rmsd min/mean/max 0.01/0.32/0.60 nm
7 strand sites, rmsd min/mean/max 0.01/0.33/0.51 nm

Residue	label	location	NO rmsd	rotamers	partition function
50 Ile	R1A	helix	0.03 nm	1	0.09705
66 Leu	R1A	helix	0.13 nm	9	0.37824
118 Leu	R1A	helix	0.18 nm	13	0.09008
13 Leu	R1A	loop	0.18 nm	17	0.05082

Progressive power saturation

Microwave power P_{mw} is increased and the amplitude I_0 of the central line measured

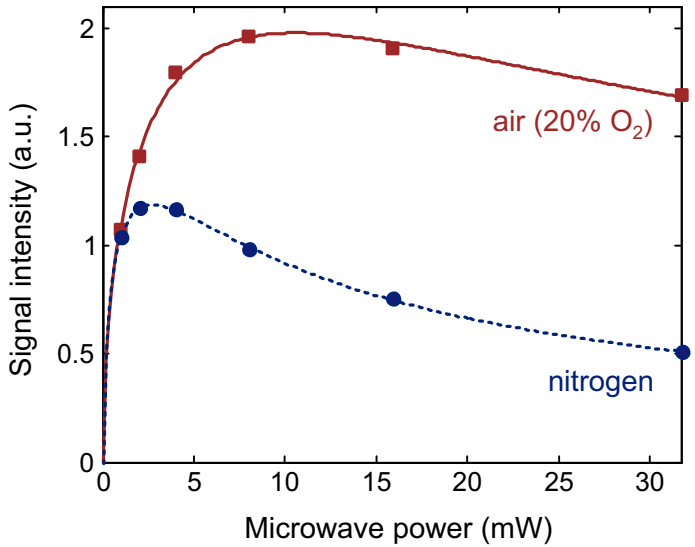
$$I_0\left(P_{\mathrm{mw}}\right) = \frac{A\sqrt{P_{\mathrm{mw}}}}{\left[1 + \left(2^{1/\epsilon} - 1\right) P_{\mathrm{mw}}/P_{1/2}\right]^{\epsilon}}$$

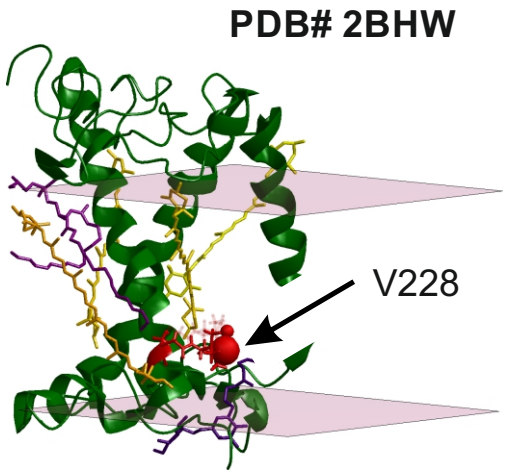
- the half-saturation power $P_{1/2}$ quantifies the relaxation enhancement
- amplitude *A* and homogeneity parameter ε are of no concern

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

- the accessibility parameter Π removes line broadening effects (δ_0) and normalizes to power conversion of the given spectrometer/probehead (reference measurement)

Example: High oxygen accessibility of a lipid-exposed residue in plant light-harvesting complex LHCII



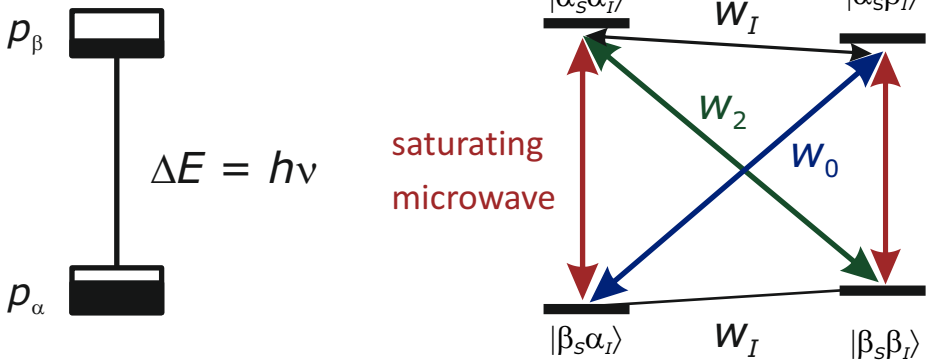


- protein complex is detergent-solubilized
- sample is contained in a gas-permeable plastic (TPX) capillary
- sample equilibrates with the composition of an external gas stream in less than a minute

Overhauser Dynamic Nuclear Polarization (DNP)

Transferring electron polarization to nuclear transitions

Boltzmann distribution



$$\sigma = w_2 - w_0$$

$$\rho = w_2 + 2w_I + w_0$$

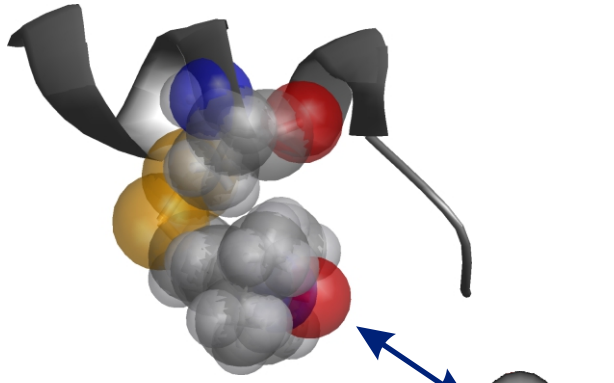
$$w_t = 1/T_{1n}^{(0)} + \rho$$

- σ is maximum if relative diffusion rate matches nuclear resonance frequency

$$\langle I_z \rangle = \langle I_z^{(0)} \rangle \left(1 + \underbrace{\frac{\sigma}{\rho}}_{\xi} \underbrace{\frac{\rho}{w_t}}_{f} \underbrace{\frac{\langle S_z^{(0)} \rangle - \langle S_z \rangle}{\langle S_z^{(0)} \rangle}}_{s} \underbrace{\frac{\langle S_z^{(0)} \rangle}{\langle I_z^{(0)} \rangle}}_{\gamma_S/\gamma_I} \right)$$

Polarization enhancement

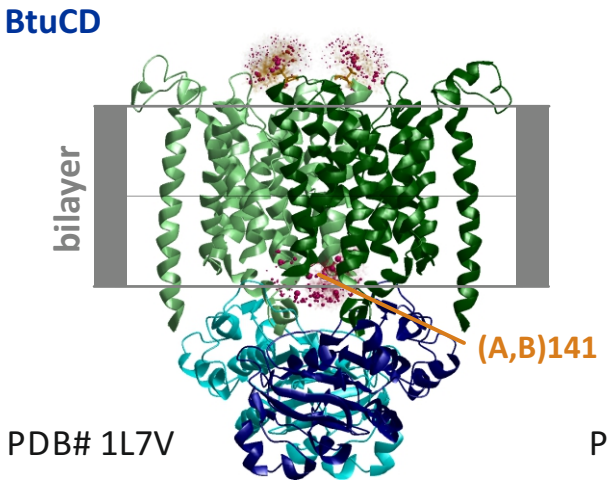
$$\varepsilon = \frac{\langle I_z \rangle}{\langle I_z^{(0)} \rangle} = 1 - \xi \cdot f \cdot s \cdot \frac{\gamma_S}{\gamma_I}$$

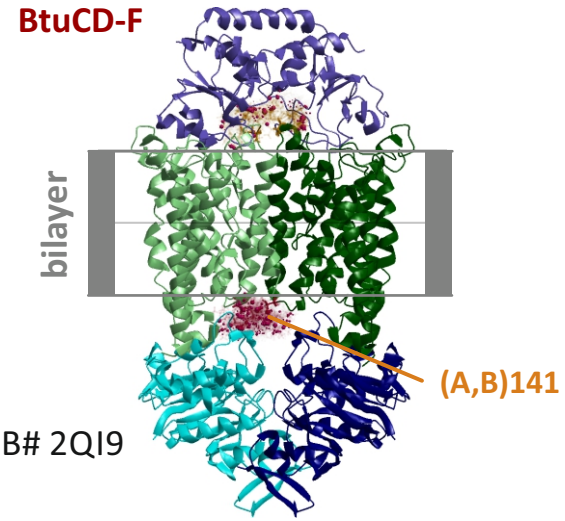


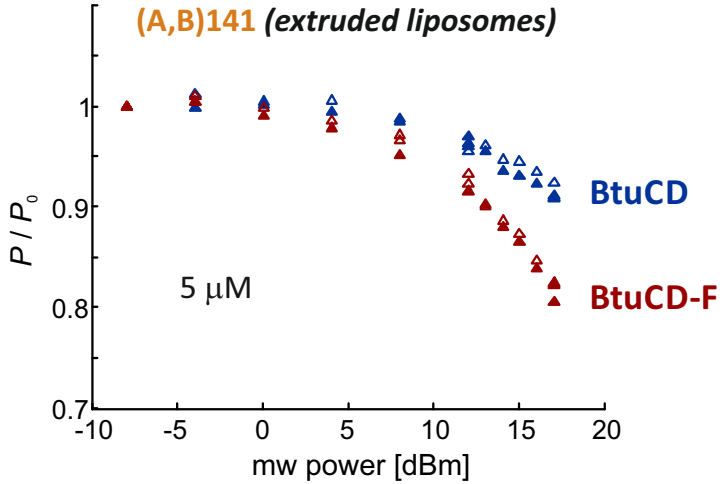
- works at physiological temperature
- no deuteration required

Overhauser DNP water accessibility measurements

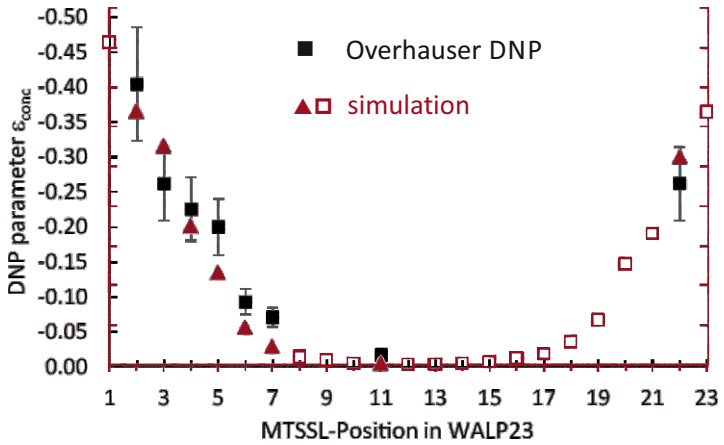
Opening of an inner gate of an ABC transporter on binding of the substrate-binding protein

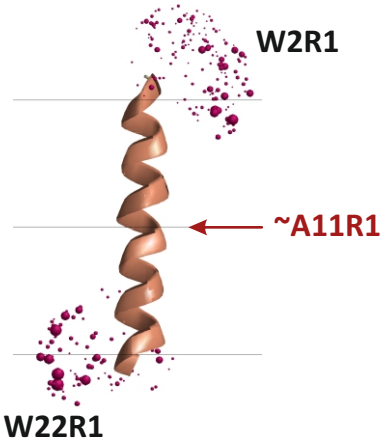






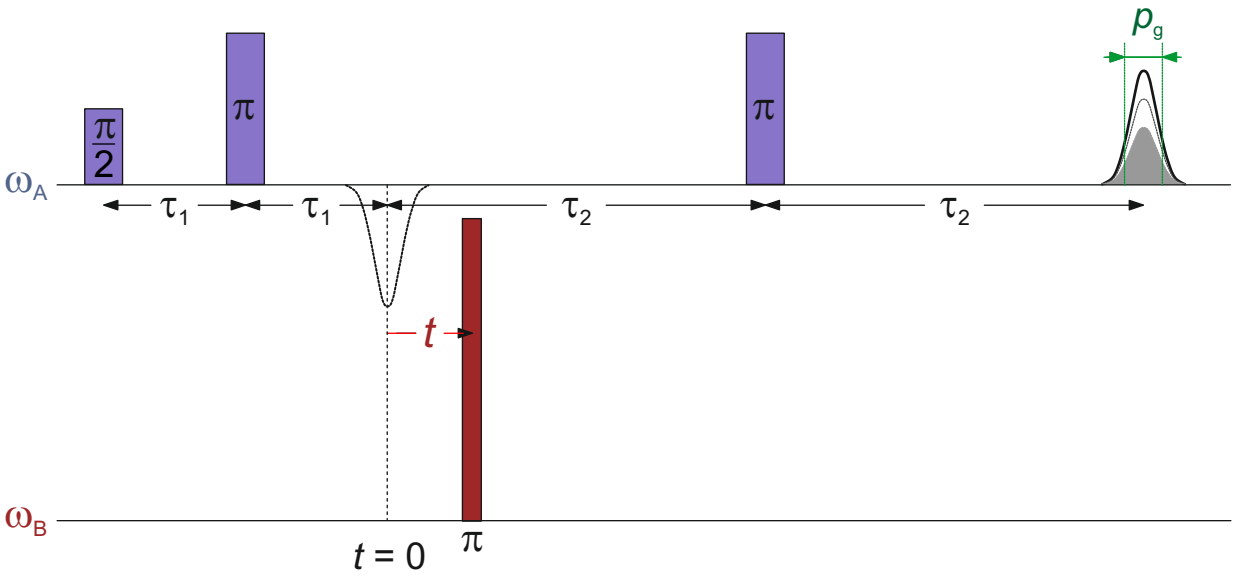
Dependence of water accessibility on immersion depth in a lipid bilayer





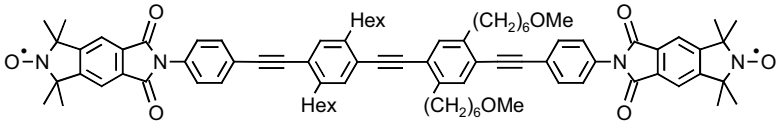
- a few µL of sample at a concentration of 10-100 µM suffice

What is DEER? A reminder

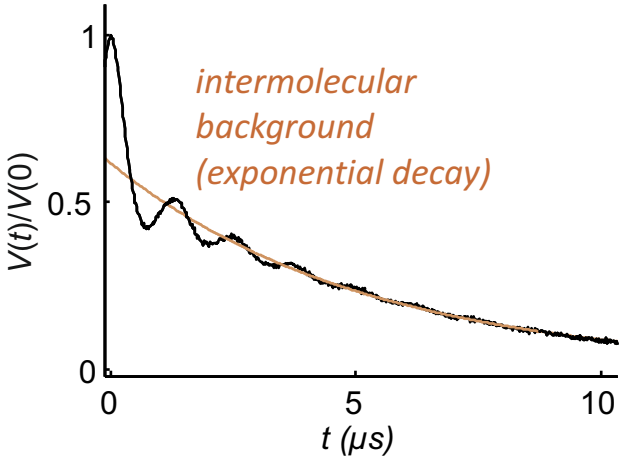


The **echo amplitude** is observed as a function of **time *t***

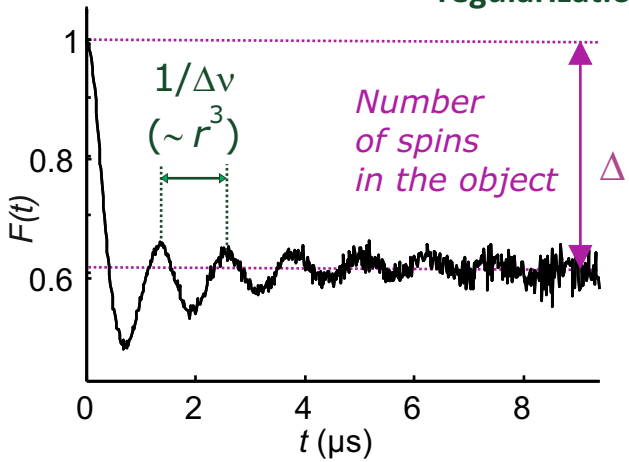
Model compound



Primary data



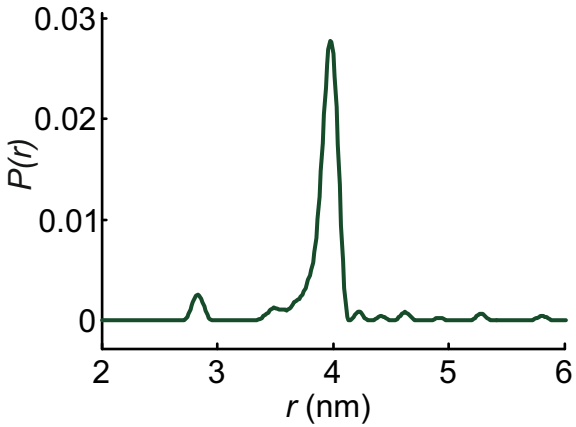
Form factor



Tikhonov regularization

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

Distance distribution



MARTIN RE et al., *Angew. Chem. Int. Ed.* **1998**, *37*, 2834

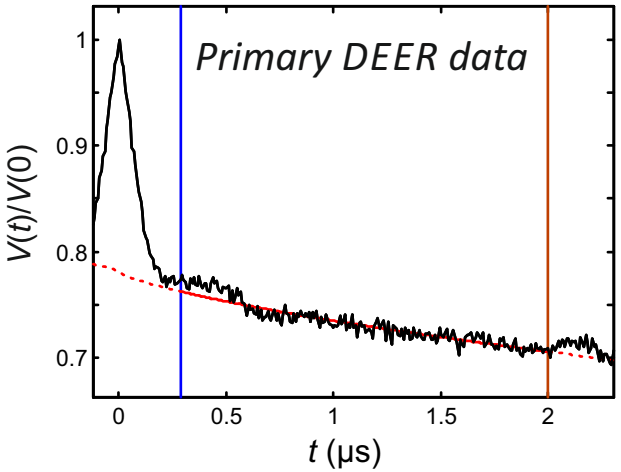
PANNIER M, VEIT S, GODT A, JESCHKE G, SPIESS HW, *J. Magn. Reson.* **2000**, *142*, 331

JESCHKE G et al. *J. Magn. Reson.* **2002**, *155*, 72

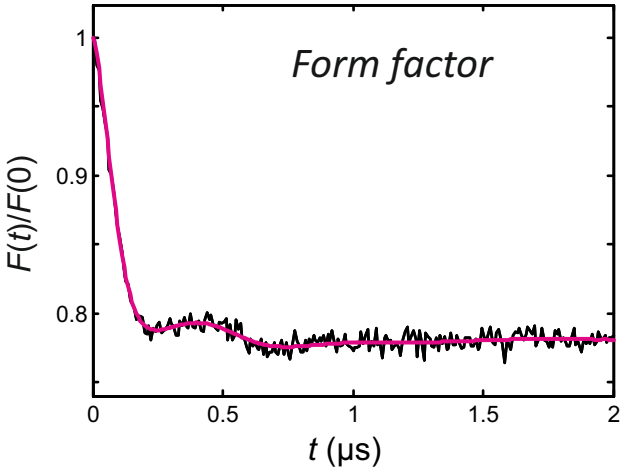
JESCHKE G et al. *Appl. Magn. Reson.* **2006**, *30*, 473

Long-range distance distribution restraints by DEER

~20-40 μM Bax 87R1/126R1 in mitochondria-like lipid vesicles (34 GHz, 150 W, 20 μL oversized sample)



background correction



- enforce $P(r) \geq 0$

Mean square deviation

$$\rho = \left\| KP(r) - D(t) \right\|^2$$

Roughness

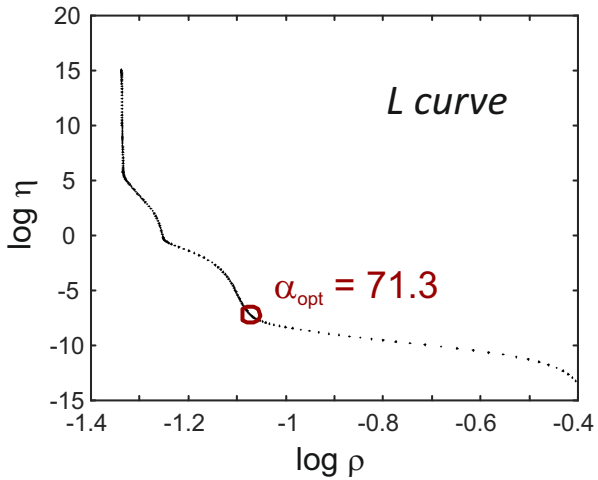
$$\eta = \left\| \frac{d^2}{dr^2} P(r) \right\|^2$$

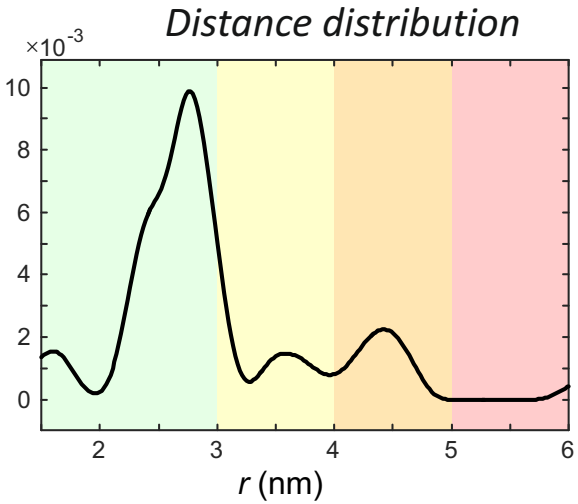
Minimize

$$G_\alpha(P) = \rho + \alpha\,\eta$$

(Tikhonov regularization)

DeerAnalysis 2019
www.epr.ethz.ch/software/index

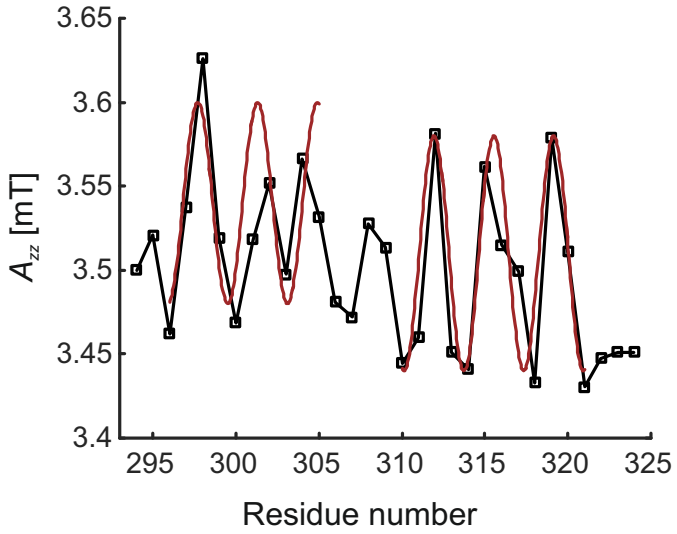


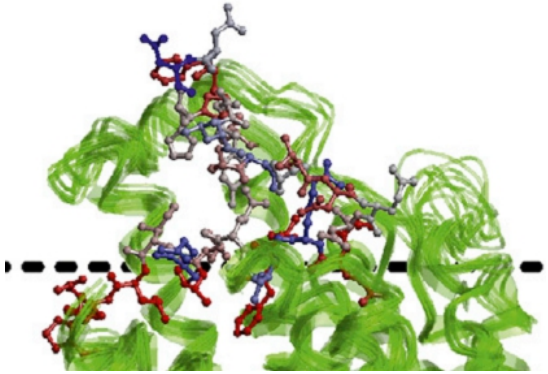


Secondary structure information from spin-labeling site scans

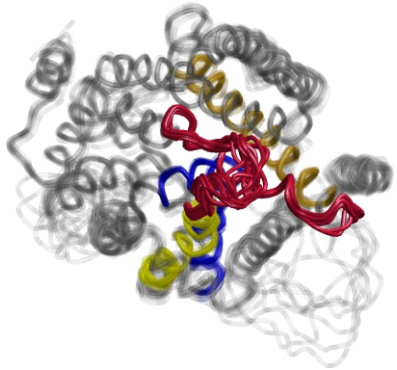
A_{zz}, δ_0, Π_{O_2} and Π_{NiEDDA} vary periodically in a site scan through an α-helix or β-sheet

External loop eL4 of Na^+/proline symporter PutP of E. coli

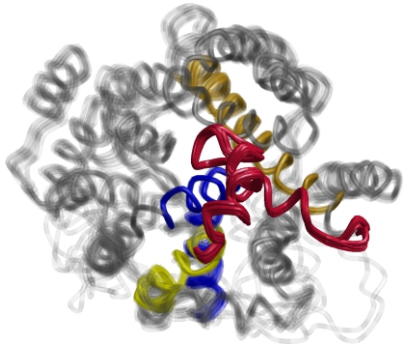




without



with



secondary structure restraints

- loop model based on homology (Na^+/glucose symporter vSGLT), secondary structure information, and a few DEER distance restraints

NATHANIEL VARGAS, Glendale

Hybrid structure determination

If a single method does not provide sufficient restraints

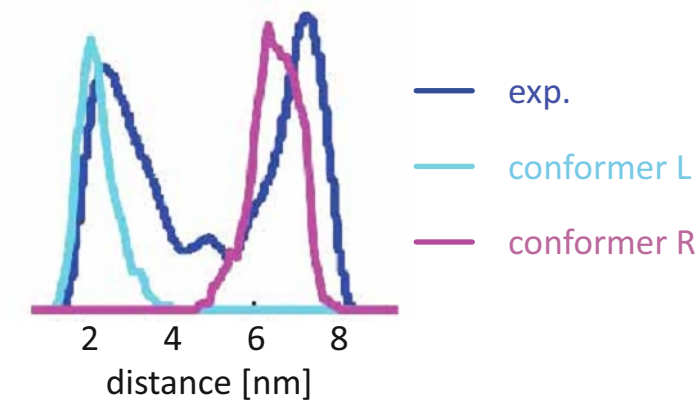
- combine experimental data from different techniques
- stabilize the solution by computational chemistry information

This may blur the boundary between experimental structure and *ab initio* model

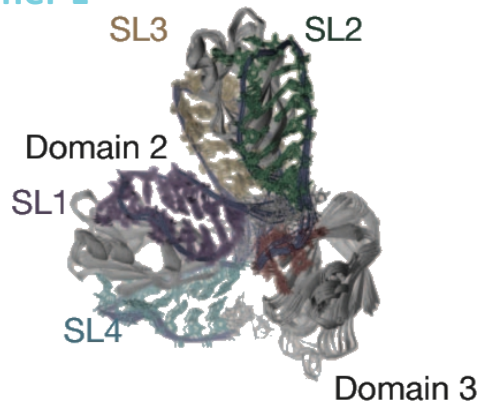
! for each restraint subset, be aware of its uncertainties

! be careful about your assumptions: the solution may not be a single conformation

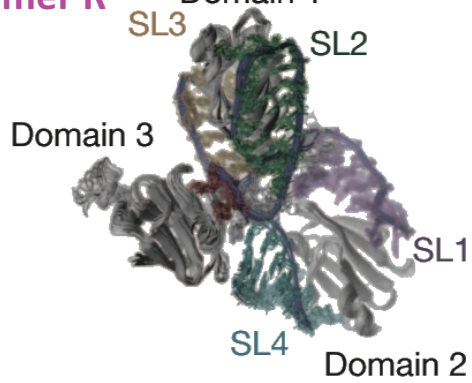
Stem loops **SL1**/**SL3**



Conformer L



Conformer R



O. DUSS, E. MICHEL, M. YULIKOV, M. SCHUBERT, G. JESCHKE, F. H.-T. ALLAIN *Nature* **509**, 588-592 (2014)

Hybrid structure modeling - Types of experimental information

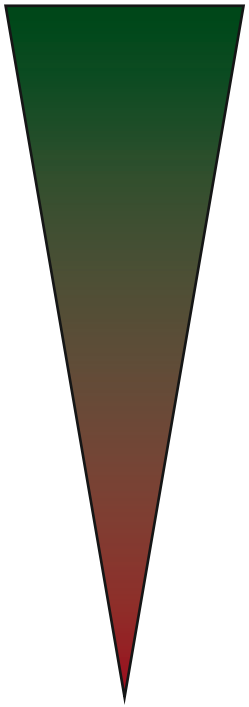
- atomic resolution structures of domains or complex components (x-ray, NMR, cryoEM)
  in the same state or in a different state
- EPR distance distribution restraints
  information on width of an ensemble of conformations or the presence of distinct conformations
- small-angle scattering curves (SAXS, SANS)
  low resolution, restrain global shape
- other EPR restraints
  (secondary structure, accessibility/bilayer immersion depth)
- other NMR restraints
  (secondary structure propensities, pseudo-contact shift information on label distribution)
- cross-linking restraints
  only subsets may apply if distinct conformations exist

Hybrid structure modeling - Types of *ab initio* information

Assumptions that one can make...

- bond length, bond angles
- clash avoidance (repulsive part of non-bonded interaction in a molecular force field)
- Ramachandran-allowed backbone torsion angles
- fragment-library information (Rosetta)
- homology information
- secondary structure prediction
- molecular force fields beyond repulsive part of non-bonded interactions
- *ab initio* folded structure

... and their reliability

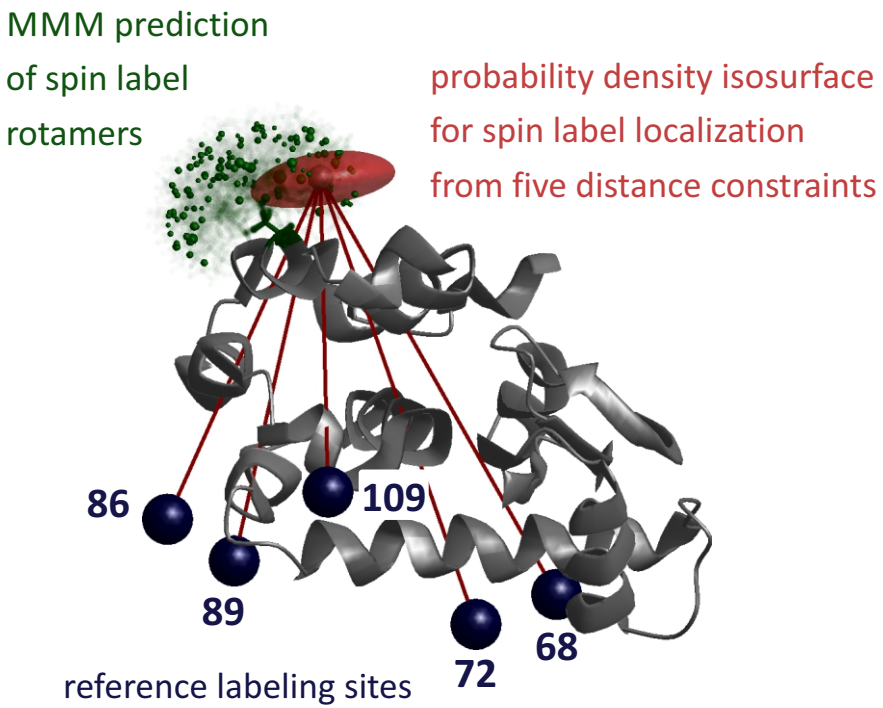


The larger the system and the more distributed its conformation, the more critical are assumptions with low reliability

The larger the system and the more distributed its conformation, the more assumptions must be made

Uncertainty and inaccuracy of spin-label based restraints

Exercise: GPS-like localization of 131R1 in T4 lysozyme



Accuracy test of label-to-label distance predictions (Å)

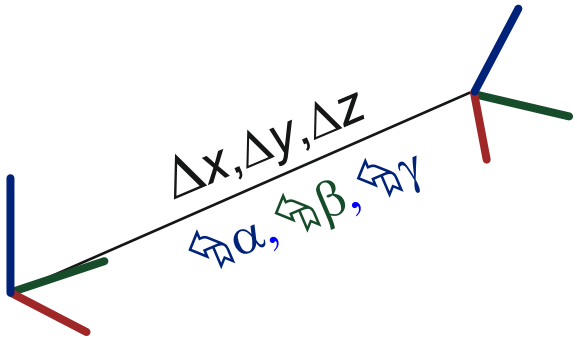
Rotamer library	*30 pairs T4L*	*62 pairs mixed*
MD/Charmm	2.3	3.0
MC/UFF	2.4	3.0
MC/UFF CαSδ	**1.7**	**2.6**

- approaches by others (mtsslWizard, PRONOX) perform on a similar level
  JESCHKE G *Progr. Nucl. Magn. Reson. Spectr.* **2013**, *72*, 42-60.
- MD simulation usually performs slightly worse, after special parametrization slightly better
  ISLAM SM, ROUX B *J. Phys. Chem. C* **2015**, *119*, 3901-3911.

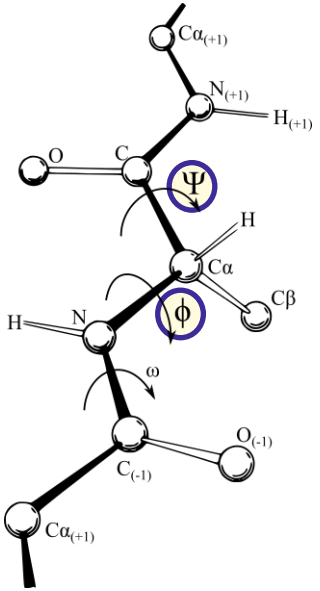
The error can be reduced by overdetermination, but EPR distance restraints are usually sparse

Sparse distance restraints & structure modeling

Concept: (Semi)rigid bodies joined by flexible linkers



- 6 degrees of freedom (3 translation/3 rotation) for each rigid body beyond the first one



- 2(*N*-1) free torsion angles (ϕ, ψ) for an *N*-residue peptide
  Beware of *cis* peptides!

- side groups are predicted by SCWRL4

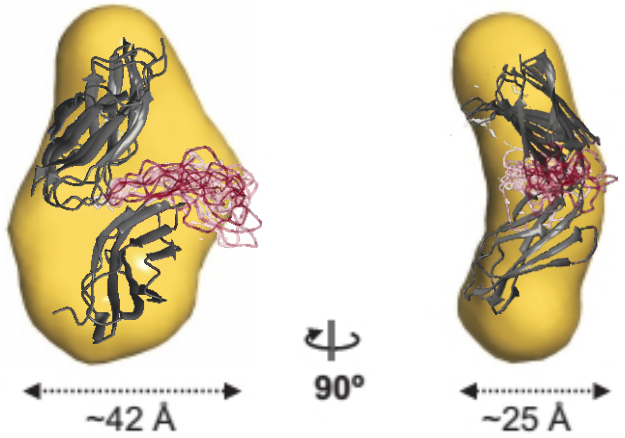
KRIVOV GG, SHAPOVALOV MV, DUNBRACK RL *Proteins* **2009**, *77*, 778-795.

Example: Combining x-ray crystallography, SAXS, and DEER

Second pair of FnIII domains of integrin α6β4

The problem

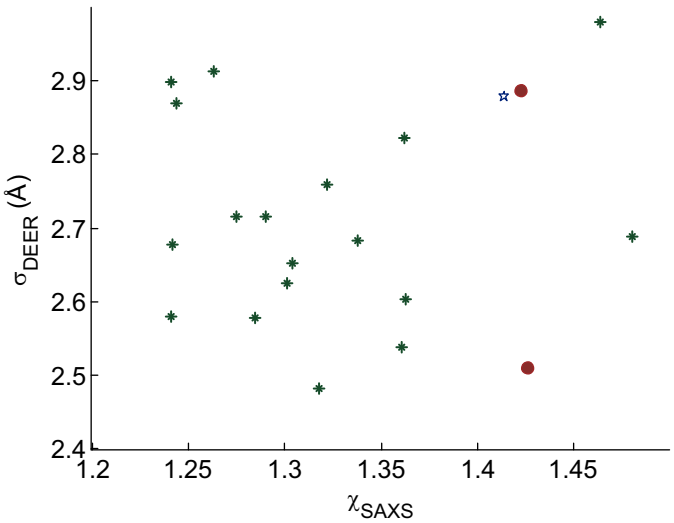
- the two individual domains crystallize, but the domain-linker-domain construct does not
- the SAXS shape does nor reveal orientation of the globular domains



linker ensemble width dominated by lack of restraints

The approach

- six rigid-body parameters from 13 DEER restraints
- two restraints to center of 21-residue linker
- Monte-Carlo linker modeling based on residue-specific Ramachandran plots
- CRYSOL for testing models against SAXS curves
- SAXS curves used for detecting structural changes by spin labeling



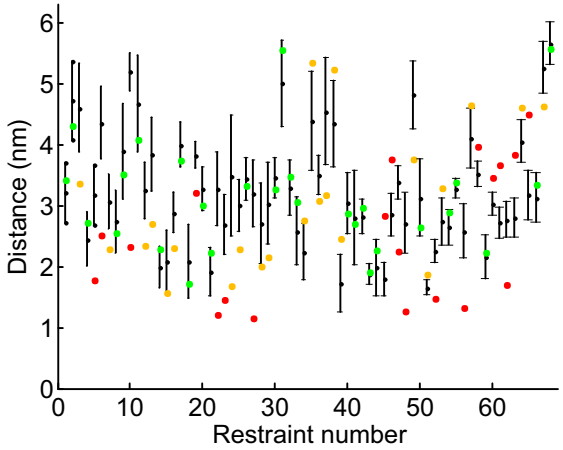
N. ALONSO-GARCÍA *et al. Acta Cryst. D* **2015**, *71*, 969-985

Restraint-augmented homology modeling

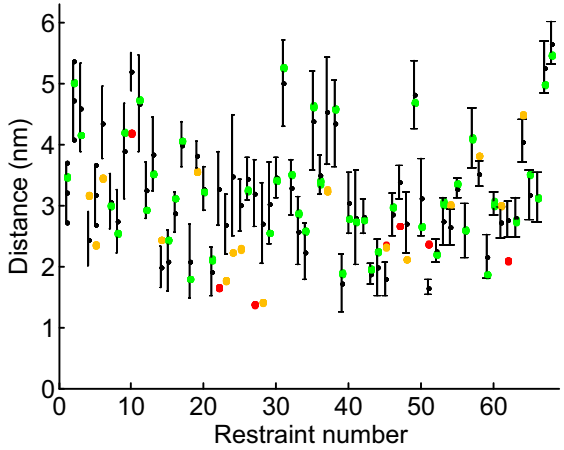
Modeling of Na^+/proline symporter PutP based on homology and DEER restraints

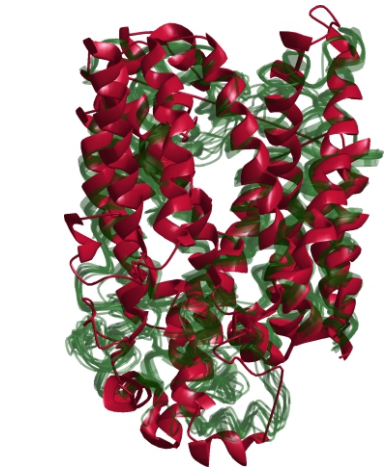
- crystal structure of the Na^+/glucose symporter vSGLT was known
- only about 20% sequence homology, different number of transmembrane helices
- 68 DEER distance restraints for “helix end” pairs

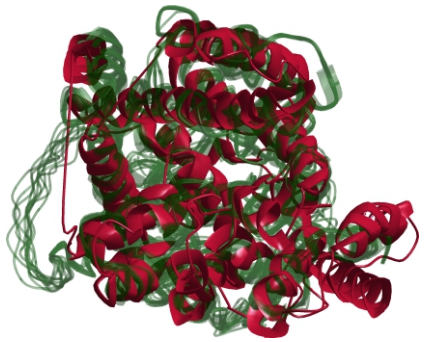
*Restraint matching for aligned residue pairs in **vSGLT***



*Restraint matching of the **DEER-augmented homology model***

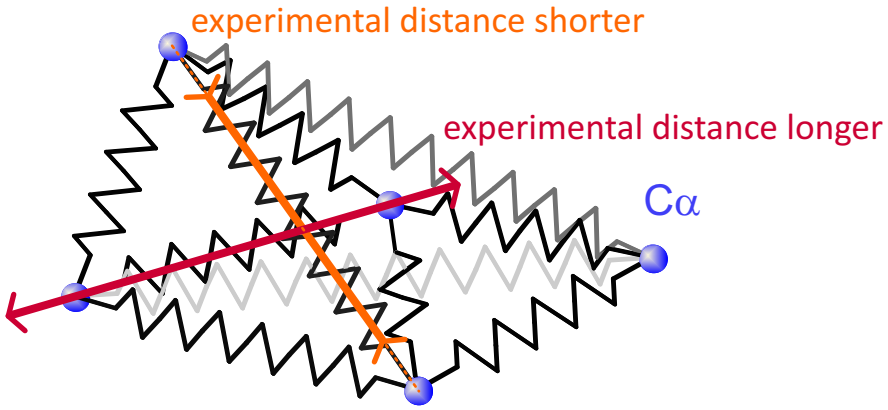






Large-scale conformational change by elastic network models

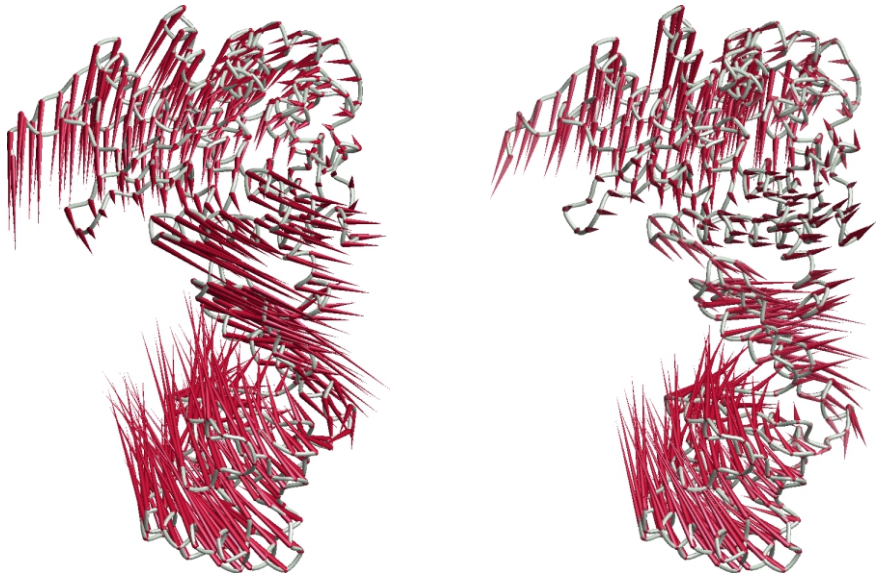
Residue-level elastic network model (ENM)



- force constants of the springs depend on Cα-Cα distance
- network is deformed along its normal modes by forces that are proportional to the mismatch of distance restraints
  ZHENG W, BROOKS BR *Biophys. J.* **2006**, *90*, 4327-36.
- label-label distances can be used as well
  JESCHKE G *J. Chem. Theor. Comp.* **2012**, *8*, 3854-63.

Hinge motion of chaperonin GroEL
with simulated DEER data

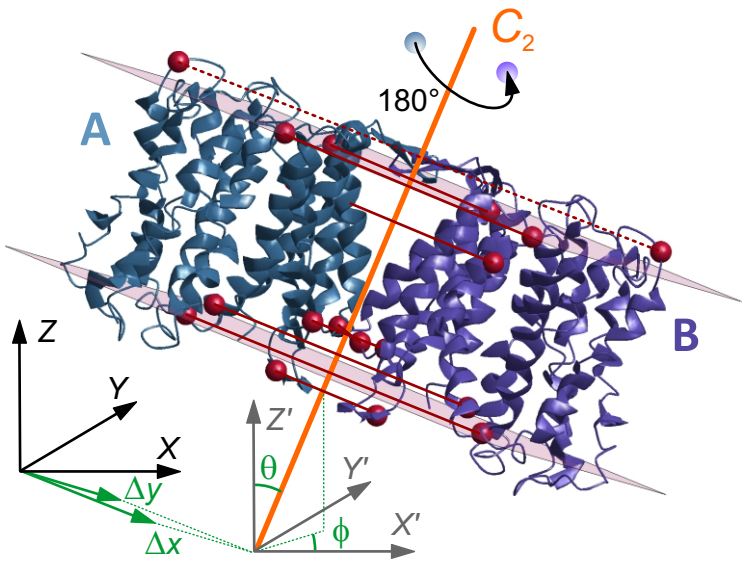
x-ray (1AON/1OEL) *ENM with 20 restraints*



- type of motion recognized, but model does not have atomic resolution
- may not work as well as for other types of motion

Rigid-body docking

Dimer structure of Na^+/H^+ antiporter NhaA

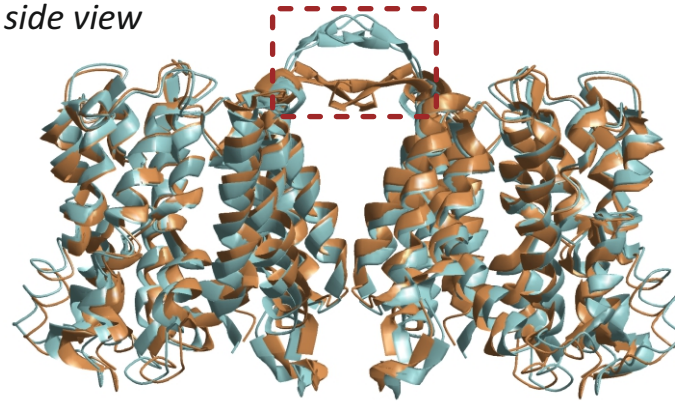


- 9 distance restraints determine 4 free parameters
- full grid search in parameter space
- protomer structure assumed to be rigid (PDB# 1ZCD)

D. Hilger, Ye. Polyhach, E. Padan, H. Jung, G. Jeschke, *Biophys. J.* **2007**, *93*, 3675-3683.

What you assume to be rigid, may move

side view

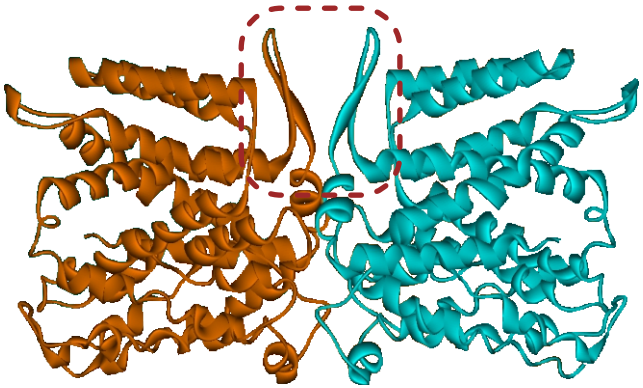


our model

new cryo-EM structure

M. Appel, D. Hizlan, K. R. Vinothkumar, C. Ziegler, W. Kühlbrandt, *J. Mol. Biol.* **2009**, *386*, 351-365.

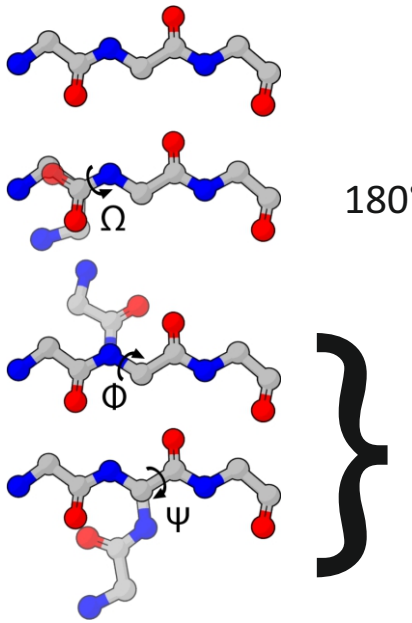
crystal packing effect in structure 1ZCD



Modeling of intrinsically disordered domains

Reliable information

- bond lengths and bond angles
- preferences for backbone dihedral angles
- side chain rotamer preferences as encoded in SCWRL4
- distance distribution restraints
- secondary structure propensities (NMR chemical shifts or periodicity of EPR parameters or)
- lipid bilayer immersion depths (membrane proteins)

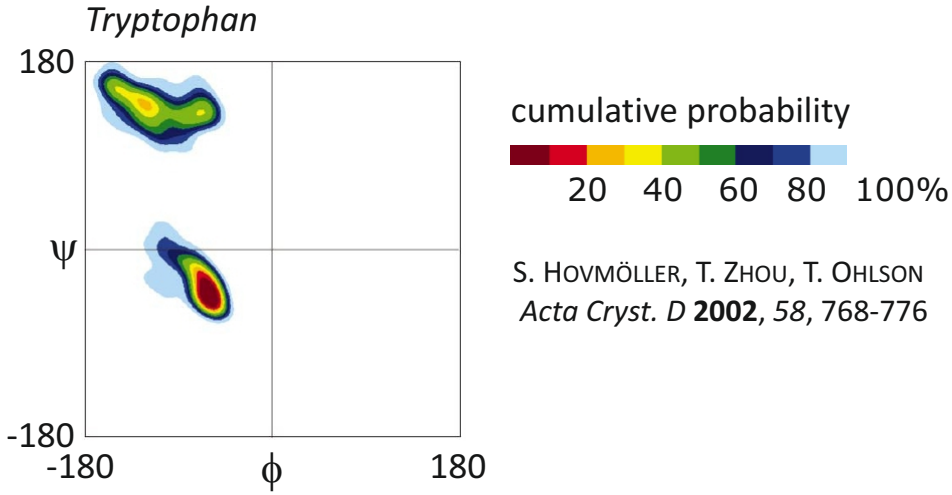


180° unless cis peptide

probability distribution known from residue-specific Ramachandran plots

Free parameters

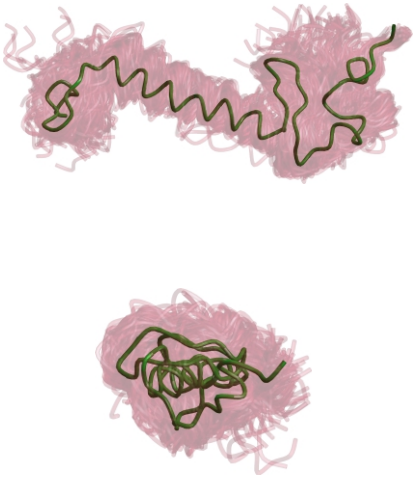
- 2(*N*-1) torsion angles fi,yi for a peptide with *N* residues
- without constraints sampling of solution space is unfeasible for *N* > 15… 20
- even with constraints loop closure between two anchor residues requires steering the loop to the second anchor



S. HOVMÖLLER, T. ZHOU, T. OHLSON
Acta Cryst. D **2002**, *58*, 768-776

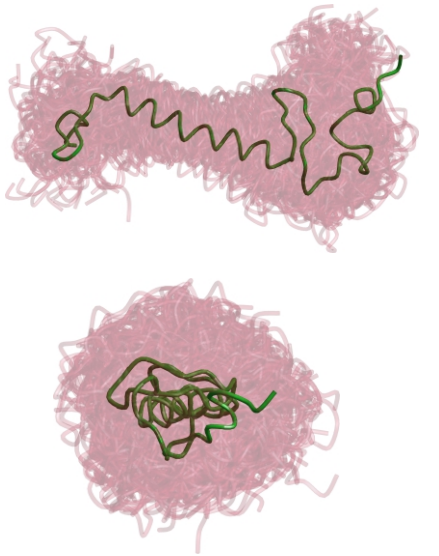
Intrinsically disordered domains: Uncertainty *versus* flexibility

p27KID: Ensemble from NMR-restrained MD and its central structure

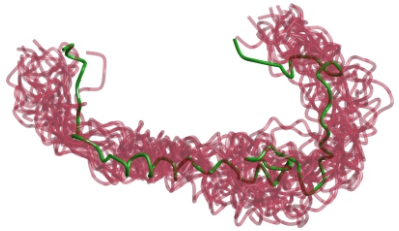


- uncertainty about spin label conformation and lack of restraints translate into larger ensemble width

p27KID: Ensemble recovered from 56 simulated DEER restraints and 21 secondary structure restraints



p27Kip1: Crystal structure in complex with Cdk2 and ensemble obtained from 56 DEER and 21 secondary structure restraints



- global shape is reproduced, but at low resolution

$\Rightarrow$ the width of EPR-derived ensembles is an upper bound on the conformation space that is actually sampled

G. JESCHKE, *Proteins* **2016**, *84*, 544-560