Supplementary Information

Benchmark test and guidelines for DEER/PELDOR experiments on nitroxide-labeled biomolecules

Olav Schiemann[a]*, Caspar A. Heubach[a], Dinar Abdullin[a], Katrin Ackermann[b], Mykhailo Azarkh[c], Elena G. Bagryanskaya[d], Malte Drescher[c], Burkhard Endeward[e], Jack H. Freed[f], Laura Galazzo[g], Daniella Goldfarb[h], Tobias Hett[a], Laura Esteban Hofer[i], Luis Fábregas Ibáñez[i], Eric J. Hustedt[j], Svetlana Kucher[g], Ilya Kuprov[k], Janet Eleanor Lovett[l], Andreas Meyer[m], Sharon Ruthstein[n], Sunil Saxena[q], Stefan Stoll[r], Christiane R. Timmel[s], Marilena Di Valentin[t], Hassane S. Mchaourab[j*], Thomas F. Prisner[e*], Bela Ernest Bode[b]*, Enrica Bordignon[g,#]*, Marina Bennati[m]*, Gunnar Jeschke[i]*

[a]*Institute of Physical and Theoretical Chemistry, University of Bonn, Wegelerstr. 12, 53115 Bonn, Germany,* [b]*EaStCHEM School of Chemistry, Biomedical Sciences Research Complex, and Centre of Magnetic Resonance, University of St Andrews North Haugh, St Andrews KY16 9ST, UK,* [c]*Department of Chemistry and Konstanz Research School Chemical Biology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany,* [d]*N.N. Vorozhtsov Novosibirsk Institute of Organic Chemistry, Lavrentieva aven 9, 630090 Novosibirsk, Russia,* [e]*Institute of Physical and Theoretical Chemistry and Center of Biomolecular Magnetic Resonance, Goethe University, Frankfurt am Main 60438, Germany,* [f]*Department of Chemistry and Chemical Biology, and ACERT, National Biomedical Center for Advanced Electron Spin Resonance Technology, Cornell University, Ithaca, New York 14853-1301, USA,* [g]*Faculty of Chemistry and Biochemistry, Ruhr University Bochum, 44801 Bochum, Germany,* [h]*Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot 76100, Israel,* [i]*Department of Chemistry and Applied Biosciences, ETH Hönggerberg, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland,* [j]*Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville TN 37232, USA,* [k]*School of Chemistry, University of Southampton, Highfield Campus, Southampton SO17 1BJ, UK,* [l]*SUPA School of Physics and Astronomy and BSRC, University of St Andrews, North Haugh, St Andrews KY16 9SS, UK,* [m]*Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany,* [n]*Department of Chemistry, Bar Ilan University, Ramat Gan 5290002, Israel,* [q]*Department of Chemistry, University of Pittsburgh, Pittsburgh, PA, 15260, USA,* [r]*Department of Chemistry, University of Washington, Seattle, WA 98195, USA,* [s]*Department of Chemistry, Centre for Advanced Electron Spin Resonance, University of Oxford, South Parks Road, Oxford OX1 3QR, UK,* [t]*Department of Chemical Sciences, University of Padova, via Marzolo 1, 35131 Padova, Italy,* [#]present address: *Department of Physical Chemistry, Sciences II, Quai Ernest Ansermet 30, 1205 Geneva, Switzerland.*

To whom correspondence should be addressed: schiemann@pc.uni-bonn.de, hassane.mchaourab@Vanderbilt.edu, prisner@chemie.uni-frankfurt.de, beb2@st-andrews.ac.uk, enrica.bordignon@unige.ch, marina.bennati@mpibpc.mpg.de, gjeschke@ethz.ch

S1. Methods and materials

S1.1. YopO expression, purification, labeling, and controls

Construct Design: Truncated $YopO_{89\text{-}729}$ C219A (YopO-WT) from *Yersinia enterocolitica* was cloned in frame into the pGex6p1 vector (GE Healthcare, Chicago, IL, USA), and transformed into *E. coli* DH5α cells for plasmid amplification. This construct served as the template for all further mutagenesis using the primers in Table S1. PCR products were purified, and subsequently transformed and amplified in *E. coli* DH5α cells. Successful mutation and construct identity were confirmed via Sanger sequencing (Figure S1).

Table S1. YopO primer pairs

Mutation	Sequence
S353C *fwd*	5'-CTGAGATTCATTACCTGTGAACCAGCGCACGTAATG-3'
S353C *rev*	5'-ACAGGTAATGAATCTCAGTCCTTGATTAGGCTTTATCTC-3'
S585C *fwd*	5'-CACAGCAAGGGCAGCCCGTGTCCTGTGAAACCT-3'
S585C *rev*	5'-GGAGAGTATTCAATTGCTGCGACAAGGTGCACTTAGC-3'
Y588C *fwd*	5'-CAAGGGCAGCCCGTGTCCTCTGAAACCTGTAGCTTCC-3'
Y588C *rev*	5'-GAATAGATAGTTGCGCTTTCGCACTCTCCTGCTGACACTGGAG-3'
V599C *fwd*	5'-GCTTCCTGAATCGATTAGCTGAGGCTAAGTGCACCTTG-3'
V599C *rev*	5'-GGAGAGTATTCAATTGCTGCGACAAGGTGCACTTAGCC-3'
Q603C *fwd*	5'-CGATTAGCTGAGGCTAAGGTCACCTTGTCGTGTCAATTG-3'
Q603C *rev*	5'-CTGCTGCTGCTGGAGAGTATTCAATTGACACGACAAGG-3'
N624C *fwd*	5'-GAGAGTGCGAAAGCGCAACTATCTATTCTGATTTGTCGTTCA-3'
N624C *rev*	5'-GAGCAACATCAGCCCAAGAACCTGAACGACAAATCAGAATA-3'
Q635C *fwd*	5'-CTGAGATTCATTACCTGTGAACCAGCGCACGTAATG-3'
Q635C *rev*	5'-TGCAGGGAACAACGAGCAACATCGGCCCAAGAACC-3'

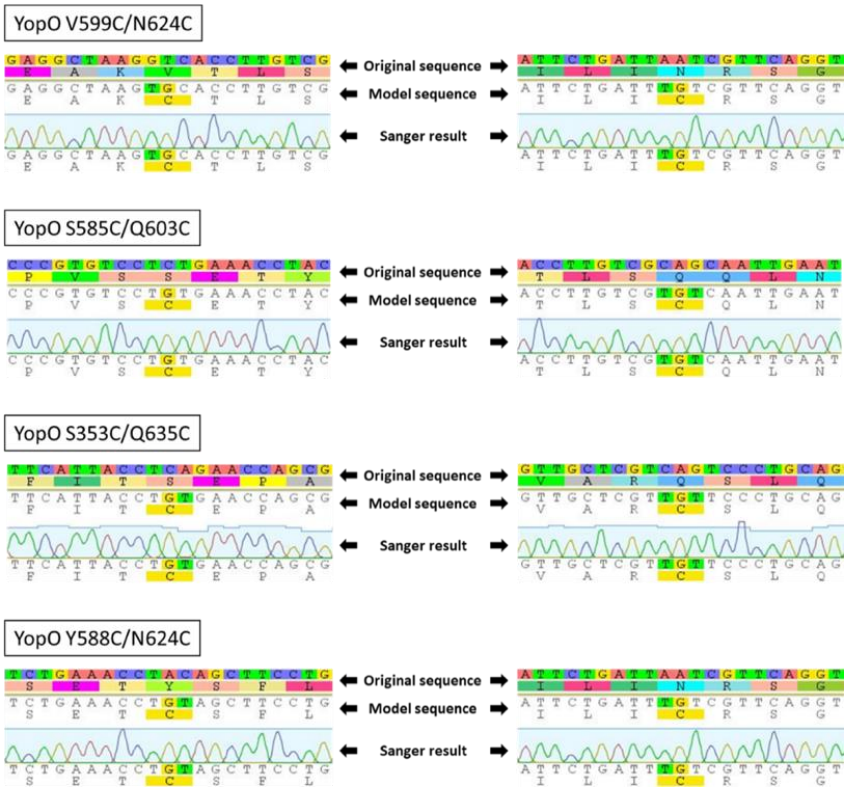

Figure S1. Sequencing results. The data shown are excerpts from Geneious (Biomatters Ltd). The YopO constructs are indicated next to the sequencing results.

YopO Expression: All YopO constructs were expressed in *E. coli* Rosetta DE3 cells. A single colony of the respective construct was picked from an agar plate (selection via ampicillin and chloramphenicol) and an overnight culture (20 mL lysogeny broth (LB) medium containing 0.3 mM ampicillin and 0.1 mM chloramphenicol) was set up (37 °C, 180 rpm). Main cultures were set up in 1 L LB medium (0.3 mM ampicillin and 0.1 mM chloramphenicol, 15 mL overnight culture) and incubated at 37 °C until an optical density at 600 nm (OD_{600}) of ~0.8 – 1.0 was reached. Protein expression was induced by addition of 0.1 mM isopropyl β-D-1-thiogalactopyranosid (IPTG). Then the cultures were cooled to 16 °C and grown for ~16 h before being harvested the next morning (4000 rcf, 20 min, 4 °C) and finally stored at -80 °C.

YopO Purification: The cell pellet was thawed and re-suspended in five-times v/w Lysis Buffer (50 mM Tris pH 8.0, 150 mM NaCl, 3 mM dithiothreitol (DTT)). The suspension was lysed twice at 32 kpsi in a cell disruptor (Constant Systems Limited, Northampton, UK) and the insoluble cell debris removed via centrifugation (48,500 rcf, 20 min, 4 °C). The supernatant was collected and incubated with equilibrated glutathione S-transferase (GST) sepharose beads for 1 h at room temperature under slight agitation. The GST-suspension was passed over a benchtop gravity column and the flow-through was run over the settled beads an additional time to maximize the protein yield. The GST-beads were washed with 50 mL Lysis Buffer, resuspended in 20 mL Cleavage Buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1 mM ethylenediaminetetraacetate (EDTA), 1 mM DTT, 100 U PreScission protease), and incubated overnight at 4 °C under slight agitation (~16 h). The next day, the beads were passed over the gravity column again and the eluate was collected, diluted with 130 mL No-Salt Buffer (50 mM Tris pH 8.0) and 190 µL DTT (2 M stock solution). Ion-exchange chromatography on an EnrichQ 10/100 column (Bio-Rad Laboratories GmbH, Feldkirchen, Germany) was performed against a linear gradient of High-Salt Buffer (50 mM Tris pH 8.0, 1 M NaCl). Protein-containing fractions were pooled and concentrated using a VivaSpin 20/10k molecular weight cutoff (MWCO, Sartorius, Göttingen, Germany) to below 3 mL. Subsequently, DTT was added to a final concentration of 3 mM and the protein was then loaded onto a HiLoad Superdex 200 16/600 gel filtration column (GE Healthcare, Chicago, IL, USA) equilibrated with gel filtration Buffer (50 mM Tris pH 8.0, 50 mM NaCl) (Figure S2). Protein purity was checked via SDS-PAGE (Figure S2) and the pure protein was pooled, concentrated to ~100 µM, flash frozen in 100 µL aliquots, and stored at -80 °C.

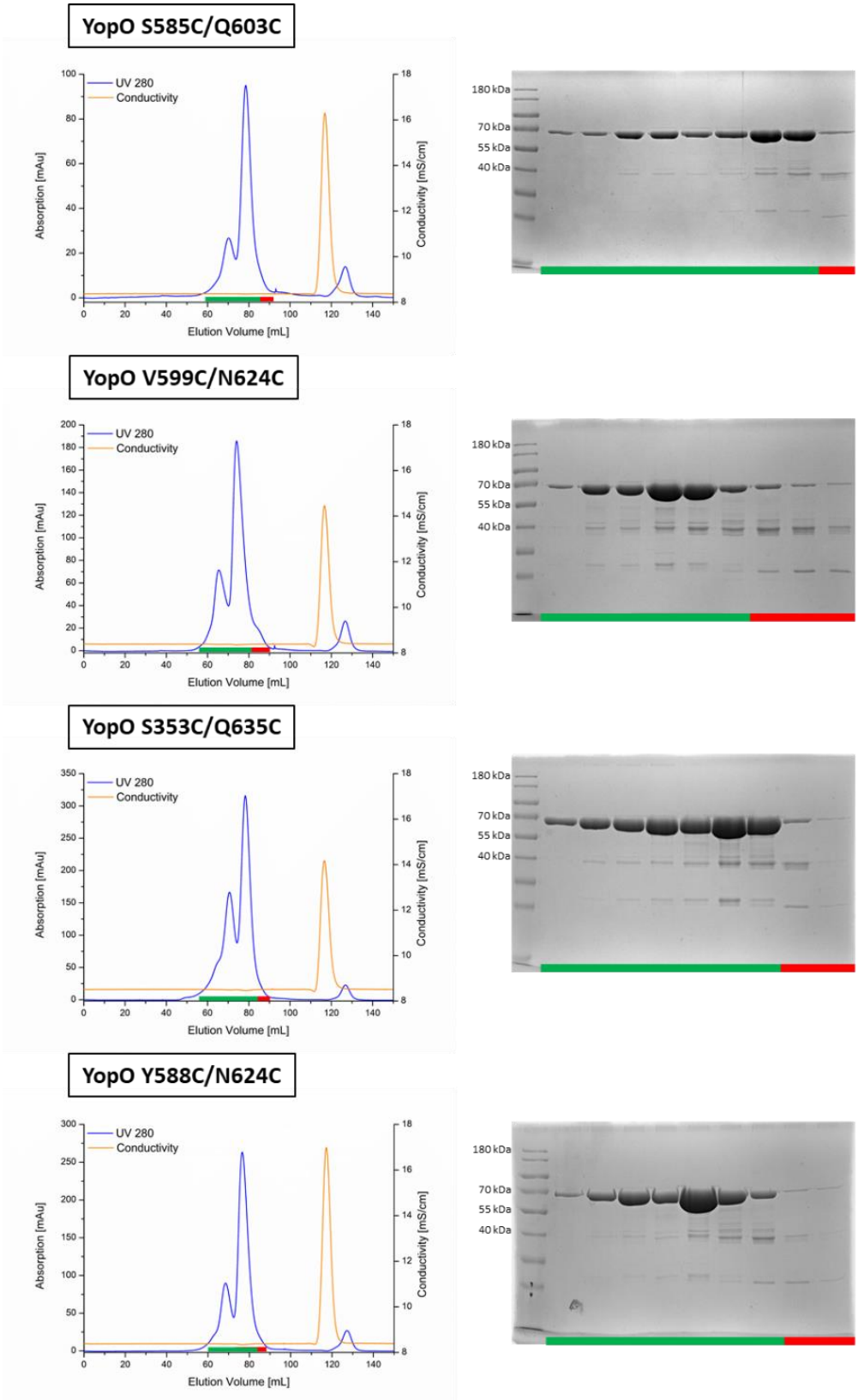

Figure S2. Purification of the YopO constructs. Gel filtration chromatograms of the indicated YopO construct (left) and the respective 10% SDS-PAGE gels (right). Gel samples are indicated with either green bars for further pooled and concentrated fractions or red bars for discarded fractions.

Labeling procedure: All labeling procedures were carried out in Labeling Buffer (50 mM Tris pH 8.0, 50 mM NaCl). Prior to each labeling experiment, 55 nmol of the respective YopO mutant were incubated with 3 mM DTT in a total volume of 2 mL for 1 h at room temperature. The reducing agent was subsequently removed using a PD-10 desalting column (GE Healthcare, Chicago, IL, USA). To the 3.5 mL PD-10 eluate containing the reduced protein, a 20-fold molar excess of (2,2,5,5-tetramethylpyrroline-1-oxyl-3-methyl)methanethiosulfonate (MTSL, Toronto Research Chemicals, North York, ON, Canada) per cysteine of a 100 mM stock solution in DMSO was added and the labeling reaction was allowed to incubate for 2 h at room temperature before being shifted to 4 °C overnight (16 h – 18 h). The next morning, a HiPrep 26/10 desalting column (GE Healthcare, Chicago, IL, USA) was run to remove any excess and unbound spin label and the first elution peak was pooled and buffer-exchanged thrice in deuterated buffer (100 mM TES pH 7.4, 100 mM NaCl) using a VivaSpin 6/10kDA MWCO (Sartorius, Göttingen, Germany). In a final step, the protein was spun for 1 min at 18k rcf in a microcentrifugal tube to remove any precipitates and the supernatant was carefully collected.
The protein amount was quantified using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) ($YopO_{89\text{-}729}$ ε_{280} = 49,390 L mol^{-1} cm^{-1}). The labeling efficiency was quantified to be about 90% for all constructs (Table S2), according to a combination of mass spectrometry, UV-vis, and quantitative EPR spin-counting experiments on an EMXnano spectrometer (Bruker BioSpin, Rheinstetten, Germany) (Figures S3 and S4, Table S3).

Table S2. Labeling efficiencies

Construct	Protein conc. [µM]	Spin conc. [µM]	Labeling Efficiency %
S585R1/Q603R1	100	175	87
V599R1/N624R1	107	187	87
Y588R1/N624R1	91	173	95
S353R1/Q635R1	105	190	90

Table S3. Calculated and experimental masses

Construct	Sum formula	Calc. mass [Da]	Exp. Mass [Da]
S585R1/Q603R1	$C_{3192}H_{5110}N_{898}O_{989}S_{18}$	72,468	72,465
V599R1/N624R1	$C_{3191}H_{5108}N_{898}O_{990}S_{18}$	72,470	72,466
S353R1/Q635R1	$C_{3192}H_{5110}N_{898}O_{989}S_{18}$	72,468	72,464
Y588R1/N624R1	$C_{3187}H_{5108}N_{898}O_{989}S_{18}$	72,406	72,405

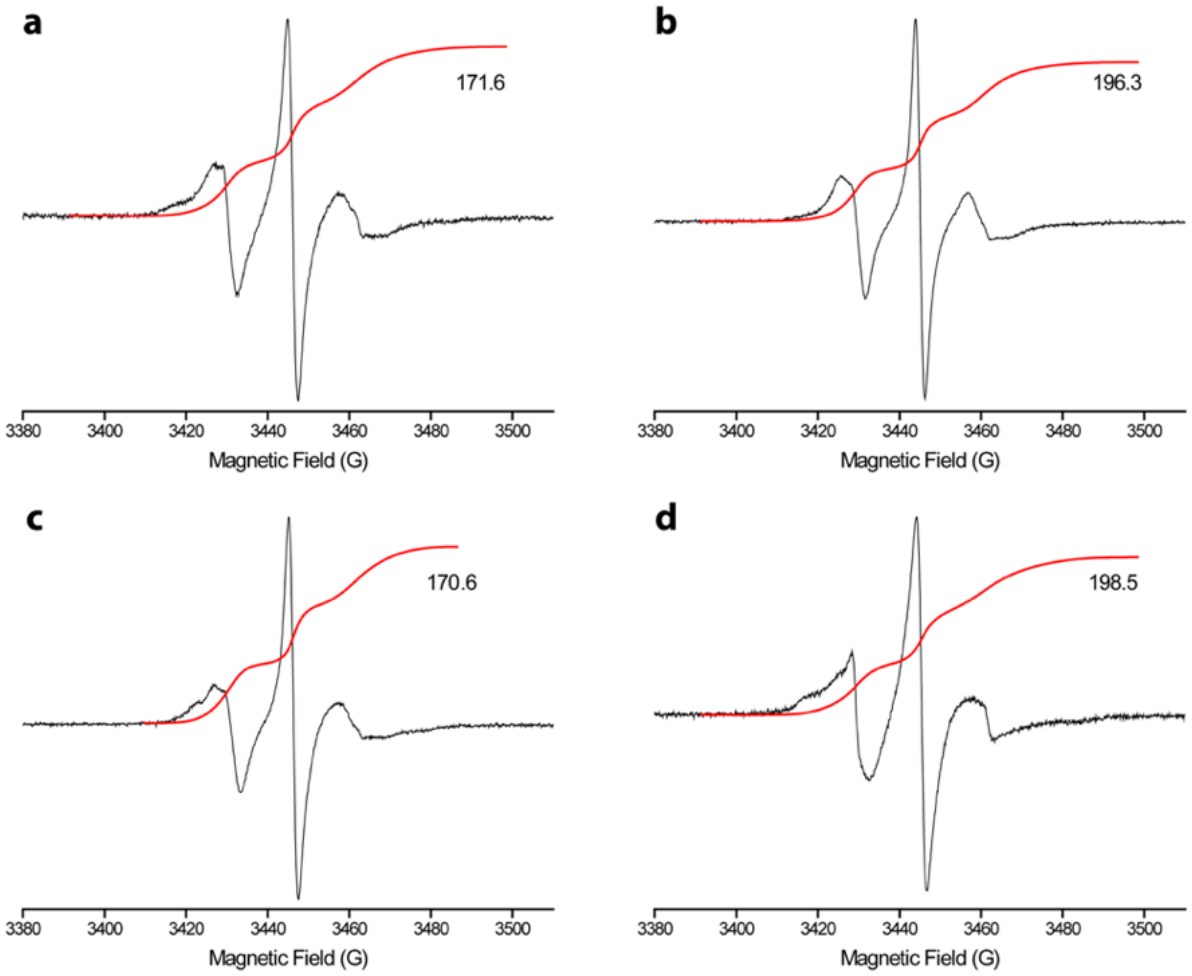

Figure S3. CW X-band EPR spectra of the YopO constructs. (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1. The red lines are the double integrals of the spectra with their values given next to them. The spectra were recorded at ~9 GHz on a Bruker EMXnano benchtop EPR spectrometer (Bruker BioSpin, Rheinstetten, Germany). The protein solution was filled into 10 µL capillaries (Disposable Capillaries, Hirschmann Laborgeräte, Eberstadt, Germany), sealed with superglue, and then transferred into 5 mm o.d. X-band EPR-tubes from Wilmad LabGlass (Vineland, NJ, USA). EPR spectrometer settings were: Microwave frequency 9.6 GHz, microwave power 10 mW (10 dB attenuation), conversion time 28.60 ms, receiver time constant 20.48 ms, center field 3445.0 G, sweep width 140 G, modulation amplitude 1.0 G, modulation frequency 100 kHz. 50 scans were averaged for each sample. The Q-value of the resonator was between 4500 and 5000 for all measurements. Spin quantitation was done using the on-board spin counting routine of the spectrometer.

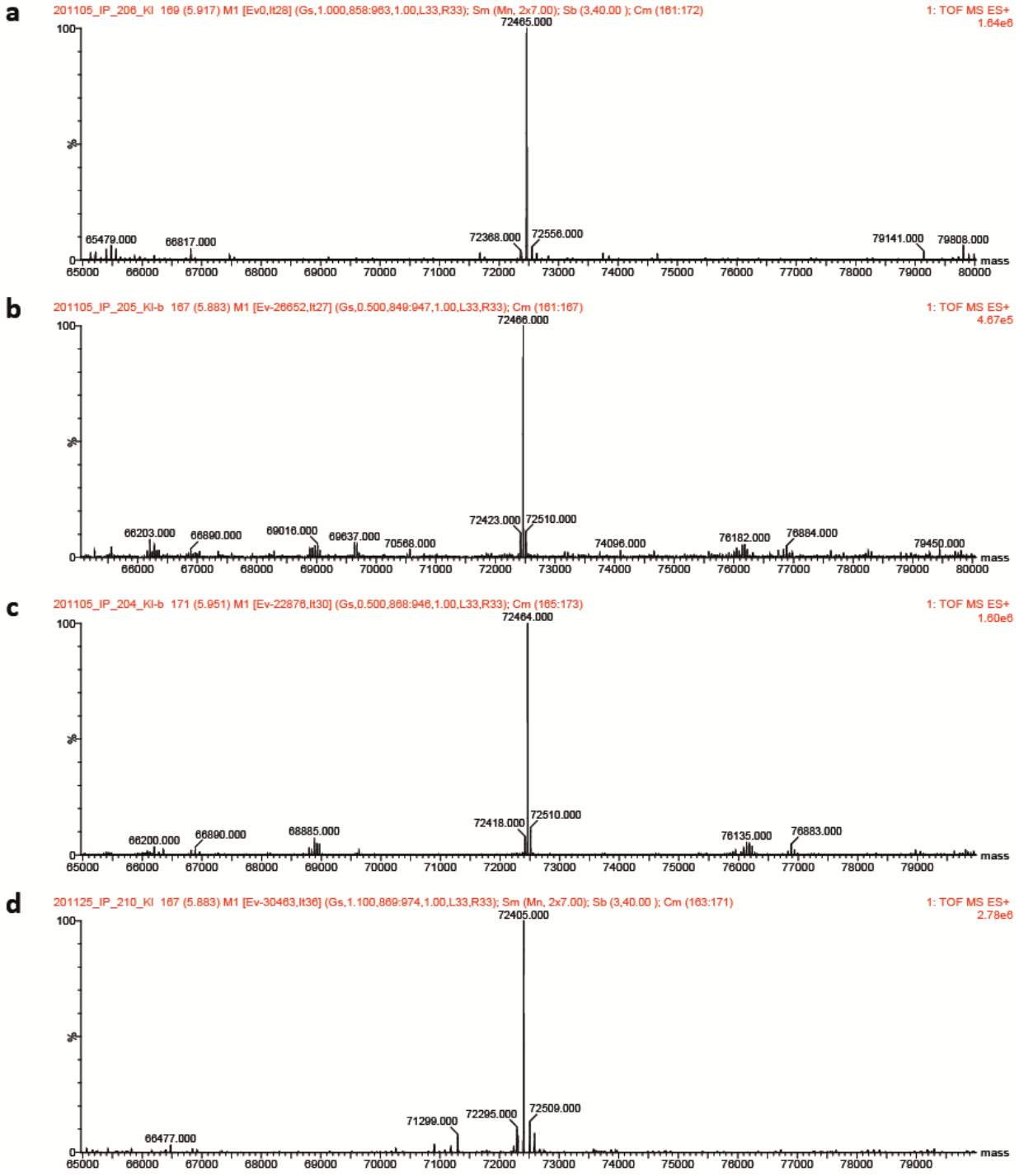

Figure S4. High-resolution ESI(+)-MS spectra. (a) YopO S585R1/Q603R1, (b) YopO V599R1/N624R1, (c) YopO S353R1/Q635R1, and (d) YopO Y588R1/N624R1. The mass spectra were recorded from aliquots of the protein solutions before the exchange with deuterated buffer.

Activity assay: In order to prove the structural integrity of all labeled YopO constructs, the autophosphorylation capability of MTSL-labeled YopO in the presence of actin was verified. 2 µM of labeled YopO were incubated in the presence of 6 µM G-actin (extracted from rabbit muscle acetone powder) in phosphorylation buffer (50 mM Tris pH 8.0, 10 mM $MgCl_2$, 1 mM ATP, 2 mM $MnCl_2$) for 1.5 h at 37 °C. For each construct, a sample without G-actin served as a negative control. The reaction mixture was quenched upon adding 8x SDS sample buffer and subsequent heating to 95 °C for 5 min. An SDS-PAGE was performed, and the acrylamide gel was fixed in 45% MeOH, 10% AcOH (2 times, 30 min, 100 mL each). Afterwards, the gel was washed with MilliQ water (3 times, 10 min, 100 mL each) and stained in the dark with 100 mL Pro-Q Phosphoprotein Diamond Stain (Thermo Fisher Scientific, Waltham, MA, USA) for 90 min. The excess staining solution was washed away with 20% acetonitrile, 50 mM NaOAc pH 4.0 (3 times, 30 min, 100 mL each), and MilliQ water (2 times, 5 min, 100 mL each). Phosphorylated protein was detected using a UV table equipped with a 590 nm long-pass emission filter. Subsequently, the gel was stained in Coomassie Brilliant Blue for the visualization of all protein (Figure S5).

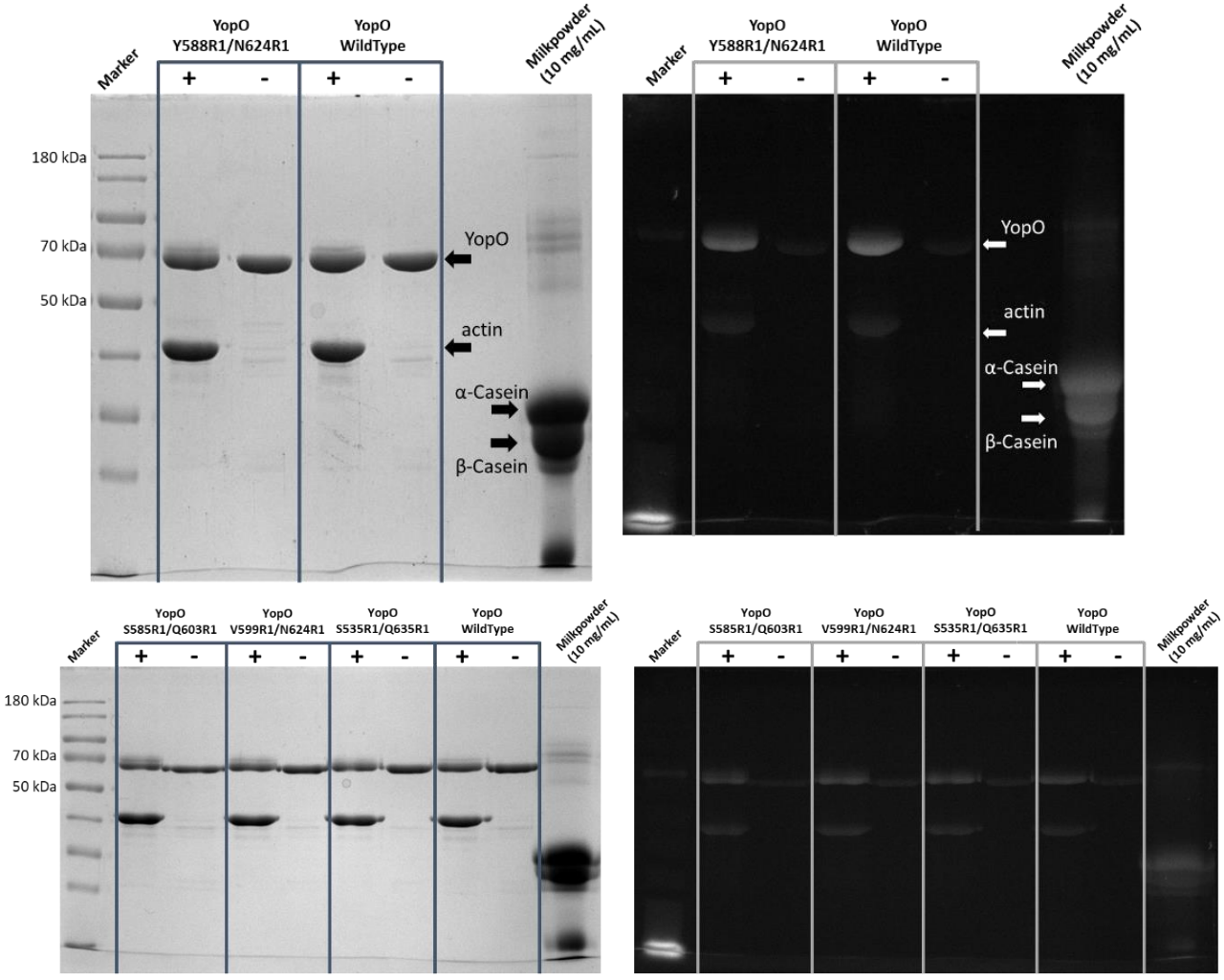

Figure S5. Activity assay. Coomassie-stained SDS gels (left) and the corresponding Pro-Q stained gels (right). Phosphorylated proteins e.g., YopO and α- + β-casein in the positive control appear significantly brighter in the Pro-Q stained gels.

S1.2 PELDOR/DEER measurements

The protein solution was diluted to a final spin concentration of 100 µM taking the labeling efficiency into account before adding a 1:1 volume of ethylene glycole-d_6. The final spin concentration was thus 50 µM for each sample. Final buffer conditions were 50 mM TES, pH 7.4, 50 mM NaCl in D_2O. The samples were aliquoted to 100 µL batches, flash-frozen in liquid nitrogen, and stored at -80 °C. The samples were shipped overnight to the respective laboratories participating in the ring test, each on dry ice within a styrofoam box. Without prior knowledge on the protein construct and the distances to expect, all laboratories performed high-power Q-band (150–170 W) PELDOR/DEER measurements with oversized samples (30 to 80 µL in 2.8 to 3 mm outer diameter tubes). Spectrometers, cryostats, setups, and data analyses differed slightly between the labs and are described below for each laboratory. Measurement parameters are given in Tables S4–S10 below. The original datasets are available with DOI: 10.5281/zenodo.5092869.

Lab A

Sample handling: One aliquot of the respective protein sample was thawed, gently homogenized using an Eppendorf pipette, and 35 µL of it were filled into a Q-band EPR tube (o.d. 3 mm, Aachener Quarzglas-Technologie Heinrich GmbH & Co. KG, Aachen, Germany). The samples were then flash-frozen in liquid nitrogen.

Pulsed EPR spectrometer: Pulsed EPR measurements were performed on a homebuilt Q-band (~34 GHz) spectrometer,[1] a Bruker ElexSys acquisition system (E580), a homebuilt Q-band resonator for 3 mm o.d. sample tubes,[1] and a 150 W TWT-amplifier (model 187 Ka, Applied Systems Engineering, Fort Worth, TX, USA). The temperature was adjusted to 50 K using an ER 4118CF helium gas-flow cryostat (Oxford Instruments, Abingdon, UK) and an ITC 503 temperature controller (Oxford Instruments). All data was acquired using quadrature detection.

Pulse settings for PELDOR/DEER: PELDOR/DEER experiments were performed using the standard four-pulse DEER sequence $\pi/2(\nu_A)–\tau_1–\pi(\nu_A)–t_1–\pi(\nu_B)–(\tau_1+\tau_2-t_1)–\pi(\nu_A)–\tau_2–$echo. The magnetic field was set to the maximum of the nitroxide spectrum and the pump pulse was applied at ν_B (approximately 34.4 GHz). The detection frequency ν_A was set 100 MHz lower than ν_B. The detection and pump frequency were set symmetric around the center of the cavity. Detection pulse lengths and the microwave power were adjusted to obtain $\pi/2$ and π-pulses ($\pi/2$ = 12 ns, π = 12 ns or $\pi/2$ = 12 ns, π = 24 ns). The length of the pump pulse was adjusted by a transient nutation experiment in such a way that a maximal inversion of the Hahn echo was obtained (12 ns). The initial time value of the interpulse delay τ_1 was set to 400 ns and the dead-time delay t_1 to 280 ns. In order to suppress deuterium ESEEM in the PELDOR/DEER time trace, a modulation averaging procedure was applied, incrementing τ_1 eight times by 16 ns and summing the individual traces (nuclear modulation averaging). The dipolar evolution time τ_2 was set with regard to the distance-dependent modulation frequency and the need to resolve at least 1.5 oscillations in the time trace. The time step Δt was chosen depending on the oscillation frequency so that a sufficient resolution of the dipolar oscillations was obtained. The number of scans was set as to obtain a decent signal-to-noise ratio of the time trace, permitting reliable data analysis. The integration gate was positioned symmetrically around the refocused Hahn echo. The gate width was adjusted to 24 ns. For every point on the time trace, 5 shots were averaged and the shot repetition time (SRT) was adjusted in such a way that ~90% of the full echo intensity could recover before starting the next cycle of the pulse sequence. Two-step phase cycling of the $\pi/2$-pulse was used to eliminate

undesired echoes and receiver baseline offsets. Table S4 summarizes the parameters used for the individual samples.

Table S4. PELDOR/DEER acquisition parameters from lab A

	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
$(\pi/2)_A$ (ns)	12	12	12	12
π_A (ns)	24	24	12	24
π_B (ns)	12	12	12	12
τ_1 (ns)	400	400	400	400
τ_2 (ns)	3000	5000	8000	8000
Δt (ns)	8	12	20	20
SRT (ms)	5	5	6	5
Shots per Point	5	5	5	5
No. of averages	20	20	78	200

Lab B

Sample handling: One aliquot of the respective protein sample was thawed, gently homogenized using an Eppendorf pipette, and 40 µL of it were filled into a Q-band EPR tube (Aachener Quarzglas-Technologie Heinrich GmbH & Co, Aachen, Germany, o.d. 3 mm). The samples were then flash-frozen in liquid nitrogen.

Pulsed EPR spectrometer: Pulsed EPR measurements were performed at Q-band (~34 GHz) on a Bruker ELEXSYS E580Q-AWG (arbitrary waveform generator) dedicated pulse Q-band spectrometer equipped with a home-made Q-band resonator (provided by G. Jeschke, ETH Zurich) and a 150 W TWT amplifier. The temperature was adjusted to 50 K using a CF935 helium gas-flow cryostat (Oxford Instruments, Abingdon, UK) and a MercuryITC temperature controller (Oxford Instruments). All data was acquired using quadrature detection.

Pulse settings for PELDOR/DEER: PELDOR/DEER experiments were performed using the standard four-pulse sequence $\pi/2(\nu_A)$–τ_1–$\pi(\nu_A)$–(τ_1+t)–$\pi(\nu_B)$–(τ_2-t)–$\pi(\nu_A)$–τ_2–echo. The magnetic field was set to the maximum of the nitroxide spectrum and pump and observer positions were chosen symmetrically with respect to the center of the cavity. The detection frequency was set 100 MHz lower than the pump frequency. Detection pulse lengths and the microwave power were adjusted to obtain $\pi/2$ and π-pulses ($\pi/2$ = 16 ns, π = 16 ns). The optimal power of the pulses was adjusted by a transient nutation experiment in such a way that a maximal inversion of the Hahn echo was obtained. The interpulse delay τ_1 was set to 400 ns. In order to suppress deuterium ESEEM, a modulation averaging procedure was applied, incrementing τ_1 eight times by 16 ns and summing the individual traces (nuclear modulation averaging). The dipolar evolution time τ_2 was set to 6000 or 10000 ns. The time step Δt was 8 ns (16 ns for Y588R1/N624R1). The number of scans was set as to obtain a decent signal-to-noise ratio of the time trace, permitting reliable data analysis. The integration gate was positioned symmetrically around the refocused Hahn echo. The gate width was adjusted to the length of the longest pulse in the sequence, herein the π-pulse at the detection frequency ν_A. For every point on the time trace, 1 or 10 shots were averaged and the shot repetition time (SRT) was fixed to 5.1 ms for all samples (1.02 ms for Y588R1/N624R1). Due to the use of the AWG, 16-step phase cycling[2] was applied to the observer and pump pulses. Table S5 summarizes the parameters used for the individual samples.

Table S5. PELDOR/DEER acquisition parameters from lab B

	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
$(\pi/2)_A$ (ns)	16	16	16	16
π_A (ns)	16	16	16	16
π_B (ns)	16	16	16	16
τ_1 (ns)	400	400	400	400
τ_2 (ns)	6000	6000	10000	8000
Δt (ns)	8	8	16	8
SRT (ms)	5	5	1	1
Shots per Point	1	1	10	1
No. of averages	20	14	55	65

Lab C

Sample handling: Aliquots of the respective protein samples were thawed, gently mixed by pipetting, and filled into a quartz EPR tube (o.d. 3 mm, H Baumbach & Co Ltd, Ipswich, UK). The vial containing mutant S353R1/Q635R1 arrived broken; ~40 µL of this sample could be rescued and transferred to an EPR tube. All samples were flash-frozen in liquid nitrogen.

Pulsed EPR spectrometer: Pulsed EPR measurements were performed at Q-band (~34 GHz) on a Bruker ELEXSYS E580 EPR spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with a Flexline probehead, a 3 mm cylindrical resonator (ER5106QT-2w in TE012 mode, Bruker), and a pulse travelling wave tube (TWT) amplifier (Applied Systems Engineering, Fort Worth, TX, USA) with nominal output of 150 W. The temperature was adjusted to 50 K using a cryogen-free variable temperature cryostat (Cryogenic Ltd, London, UK) operating in the 3.5 to 300 K temperature range.

Pulse settings for PELDOR/DEER: PELDOR/DEER experiments were performed using the standard four-pulse sequence $\pi/2(\nu_A)$–τ_1–$\pi(\nu_A)$–(τ_1+t)–$\pi(\nu_B)$–(τ_2-t)–$\pi(\nu_A)$–τ_2–echo. The magnetic field was set to the maximum of the nitroxide spectrum and the pump pulse was applied at ν_B (~34 GHz) in the center of the cavity. The detection frequency ν_A was set 80 MHz lower than ν_B. Pulse lengths were 16 and 32 ns for $\pi/2$ and π detection, and 12 ns for the ELDOR π pump pulse. Shot repetition times (SRT) were set to 4 or 5 ms; the interpulse delay τ_1 was set to 380 ns, and a nuclear modulation averaging procedure was applied (except for mutant Y588R1/N624R1), incrementing τ_1 eight times by 16 ns and summing the individual traces. The dipolar evolution time τ_2 was set depending on the distance-dependent modulation frequency and varied between 4 and 9.5 µs. The time step Δt was chosen depending on the oscillation frequency so that a sufficient resolution of the dipolar oscillations was obtained and varied between 12 and 20 ns. The number of scans was set as to obtain a decent signal-to-noise ratio of the time trace, permitting reliable data analysis. The integration gate was positioned symmetrically around the refocused Hahn echo. The gate width was 40 ns for mutant Y588R1/N624R1 and 32 ns, corresponding to the detection π-pulse length, for the other samples. For every point on the time trace, 50 shots were averaged. Two-step phase cycling of the $\pi/2$-pulse was used to eliminate undesired echoes and receiver baseline offsets. Table S6 summarizes the parameters used for the individual samples.

Table S6. PELDOR/DEER acquisition parameters from lab C

	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
$(\pi/2)_A$ (ns)	16	16	16	16
π_A (ns)	32	32	32	32
π_B (ns)	12	12	12	12
τ_1 (ns)	380	380	380	380
τ_2 (ns)	4000	6000	9500	8000
Δt (ns)	12	20	20	12
SRT (ms)	5	5	4	5
Shots per Point	50	50	50	50
No. of averages	1	15	286	29

Lab D

Sample handling: One aliquot of the respective protein sample was thawed, gently homogenized using an Eppendorf pipette, and 80 µL of it were filled into a Q-band EPR tube (o.d. 3 mm, Wilmad LabGlass, Vineland, NJ, USA). The samples were then flash-frozen in liquid nitrogen.

Pulsed EPR spectrometer: Pulsed EPR measurements were performed at Q-band (~34 GHz) on a Bruker ELEXSYS E580 EPR spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with a Flexline probehead, an ER5106QT-II resonator (Bruker), and a 150 W TWT-amplifier (model 187 Ka, Applied Systems Engineering, Fort Worth, TX, USA). The temperature was adjusted to 50 K using a CF935 helium gas-flow cryostat (Oxford Instruments, Abingdon, UK) and an iTC503S temperature controller (Oxford Instruments). All data was acquired using quadrature detection.

Pulse settings for PELDOR/DEER: PELDOR/DEER experiments were performed using the standard four-pulse sequence $\pi/2(\nu_A)$–τ_1–$\pi(\nu_A)$–(τ_1+t)–$\pi(\nu_B)$–(τ_2-t)–$\pi(\nu_A)$–τ_2–echo. The magnetic field was set to the maximum of the nitroxide spectrum and the pump pulse was applied at ν_B (33.7 GHz) in the center of the cavity. The detection frequency ν_A was set 80 MHz lower than ν_B. Detection pulse lengths and the microwave power were adjusted to obtain $\pi/2$ and π-pulses ($\pi/2$ = 16 ns, π = 32 ns). The length of the pump pulse was adjusted by a transient nutation experiment in such a way that a maximal inversion of the Hahn echo was obtained (12 ns–16 ns). For the interpulse delay τ_1, the initial time value was set to the first maximum in the two-pulse ESEEM trace (232 ns). In order to suppress deuterium ESEEM in the PELDOR/DEER time trace, a modulation averaging procedure was applied, incrementing τ_1 eight times by 16 ns and summing the individual traces (nuclear modulation averaging). The dipolar evolution time τ_2 was set with regard to the distance-dependent modulation frequency and the need to resolve at least 1.5 oscillations in the time trace. The time step Δt was chosen depending on the oscillation frequency so that a sufficient resolution of the dipolar oscillations was obtained. The number of scans was set as to obtain a decent signal-to-noise ratio of the time trace, permitting reliable data analysis. The integration gate was positioned symmetrically around the refocused Hahn echo. The gate width was adjusted to the length of the longest pulse in the sequence, herein the π-pulse at the detection frequency ν_A. For every point on the time trace, 10 shots were averaged and the shot repetition time (SRT) was adjusted in such a way that ~80% of the full echo intensity could recover before starting the next cycle of the pulse sequence. Two-step phase cycling of the $\pi/2$-pulse was used to eliminate undesired echoes and receiver baseline offsets. Table S7 summarizes the parameters used for the individual samples.

Table S7. PELDOR/DEER acquisition parameters from lab D

	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
$(\pi/2)_A$ (ns)	16	16	12	16
π_A (ns)	32	32	24	32
π_B (ns)	14	14	16	16
τ_1 (ns)	232	232	260	232
τ_2 (ns)	4000	5500	12000	9000
Δt (ns)	4	8	16	16
SRT (ms)	4	4	4	4
Shots per Point	10	10	10	10
No. of averages	33	124	517	59

Lab E

Sample handling: One aliquot of the respective protein sample was thawed, gently homogenized using an Eppendorf pipette, and 80 µL of it were filled into a Q-band EPR tube (o.d. 3 mm, Wilmad LabGlass, Vineland, NJ, USA). The samples were then flash-frozen in liquid nitrogen.

Pulsed EPR Spectrometer: Pulsed EPR and PELDOR/DEER experiments at 34 GHz (Q-band) microwave (MW) frequency were performed on a Bruker ELEXSYS E580 EPR spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with a Bruker ER5106QT-2 resonator and a CF935 helium gas flow cryostat (Oxford Instruments, Abingdon, UK) cooled down to 50 K. The temperature was controlled using an iTC503S temperature controller (Oxford Instruments). MW pulses were amplified by a pulsed 150 W TWT amplifier (Model 187Ka, Applied Systems Engineering, USA).

Pulse settings for PELDOR/DEER: PELDOR/DEER experiments were performed using the standard four-pulse sequence $\pi/2(\nu_A)$–τ_1–$\pi(\nu_A)$–(τ_1+t)–$\pi(\nu_B)$–(τ_2-t)–$\pi(\nu_A)$–τ_2–echo. The magnetic field was set to the maximum of the nitroxide spectrum and the pump pulse was applied at ν_B (33.4 – 33.5 GHz for the different samples) in the center of the cavity. The detection frequency ν_A was set 100 MHz lower than ν_B. Detection pulse lengths and the microwave power were adjusted to obtain $\pi/2$ and π-pulses ($\pi/2$ = 12 – 16 ns, π = 24 – 32 ns). The length of the pump pulse was adjusted by a transient nutation experiment in such a way that a maximal inversion of the Hahn echo was obtained (14 – 16 ns). The initial interpulse delay τ_1 was set to 400 or 450 ns. In order to suppress deuterium ESEEM in the PELDOR/DEER time trace, a modulation averaging procedure was applied, incrementing τ_1 eight times by 16 ns and summing the individual traces (nuclear modulation averaging). The dipolar evolution time τ_2 was set with regard to the distance-dependent modulation frequency and the need to resolve at least 1.5 oscillations in the time trace. The time step Δt was chosen depending on the oscillation frequency (typically $\geq$ 50 time points per oscillation). The number of scans was set as to obtain a decent signal-to-noise ratio of the time trace, permitting reliable data analysis. The integration gate was positioned symmetrically around the refocused Hahn echo. The gate width was adjusted to the length of the longest pulse in the sequence, herein the π-pulse at the detection frequency ν_A. For every point on the time trace, 10 shots were averaged and the shot repetition time (SRT) was adjusted in such a way that ~80% of the full echo intensity could recover before starting the next cycle of the pulse sequence. Two-step phase cycling of the $\pi/2$-pulse (+x, -x) was used. Table S8 summarizes the parameters used for the individual samples.

Table S8. PELDOR/DEER acquisition parameters from lab E

	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
$(\pi/2)_A$ (ns)	14	14	14	16
π_A (ns)	28	28	24	32
π_B (ns)	14	16	14	14
τ_1 (ns)	450	450	400	450
τ_2 (ns)	4200	6200	9000	8000
Δt (ns)	8	16	30	20
SRT (ms)	5	5	4	5
Shots per Point	10	10	10	10
No. of averages	177	243	218	286

Lab F

Sample handling: One aliquot of the respective protein sample was thawed, gently homogenized using an Eppendorf pipette, and 60 µL of it were filled into a Q-band EPR tube (o.d. 3 mm, Technical Glass Products, USA). The samples were then flash-frozen in liquid nitrogen.

Pulsed EPR spectrometer: Pulsed EPR measurements were performed at Q-band (~34 GHz) on a Bruker ELEXSYS E580 EPR spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with an arbitrary waveform generator (AWG), a Flexline probehead, an ER5106QT-II resonator (Bruker), and a 150 W TWT-amplifier (model 187 Ka, Applied Systems Engineering, Fort Worth, TX, USA). The temperature was adjusted to 50 K using a CF935O helium gas-flow cryostat (Oxford Instruments, Abingdon, UK) and a MercuryITC temperature controller (Oxford Instruments).

Pulse settings for PELDOR/DEER: PELDOR/DEER experiments were performed using the standard four-pulse sequence $\pi/2(\nu_A)$–τ_1–$\pi(\nu_A)$–(τ_1+t)–$\pi(\nu_B)$–(τ_2-t)–$\pi(\nu_A)$–τ_2–echo. The magnetic field was set to the maximum of the nitroxide spectrum and the pump pulse was applied at ν_B (34 GHz) in the center of the cavity. The detection frequency ν_A was set 70 MHz lower than ν_B for Y588R1/N624R1 and 80 MHz for all other samples. Pulse lengths were 12 and 24 ns for $\pi/2$ and π detection, and 24 – 28 ns for the ELDOR π pump pulse. Shot repetition times (SRT) were set to 4 or 5 ms; the interpulse delay τ_1 was set to 400 ns, and a nuclear modulation averaging procedure was applied incrementing τ_1 eight times by 16 ns and summing the individual traces. The dipolar evolution time τ_2 was set to 10 µs for Y588R1/N624R1 and 5 µs for all other samples. The time step Δt was set to 8 ns. The number of scans was set as to obtain a decent signal-to-noise ratio of the time trace, permitting reliable data analysis. The integration gate was positioned symmetrically around the refocused Hahn echo. The gate width was 24 ns for Y588R1/N624R1 and 28 ns, corresponding to the detection π-pulse length, for all other samples. For every point on the time trace, 5 to 25 shots were averaged. Eight-step phase cycling was used to eliminate undesired echoes from coherent microwave source and receiver baseline offsets.[3] Table S9 summarizes the parameters used for the individual samples.

Table S9. PELDOR/DEER acquisition parameters from lab F

	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
$(\pi/2)_A$ (ns)	12	12	12	12
π_A (ns)	24	24	24	24
π_B (ns)	28	28	24	28
τ_1 (ns)	400	400	600	400
τ_2 (ns)	5000	5000	10000	5000
Δt (ns)	8	8	8	8
SRT (ms)	5	5	4	5
Shots per Point	25	5	25	5
No. of averages	8	38	11	40

Lab G

Sample handling: One aliquot of the respective protein sample was thawed, gently homogenized using an Eppendorf pipette, and 30 µL of it were filled into a Q-band EPR tube (o.d. 2.8 mm, Wilmad LabGlass, Vineland, NJ, USA). The samples were then flash-frozen in liquid nitrogen. Sample Y588R1/N624R1 was prepared the same way, however only 10 µL were filled into the EPR tube (o.d. 1.6 mm, Wilmad LabGlass).

Pulsed EPR spectrometer: Pulsed EPR measurements were performed at Q-band (~34 GHz) on a Bruker ELEXSYS E580 EPR spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with a Flexline probehead, an ER5106QT-2 resonator (Bruker), and a 150 W TWT-amplifier (model 187 Ka, Applied Systems Engineering, Fort Worth, TX, USA). The temperature was adjusted to 50 K using a CF935 helium gas-flow cryostat (Oxford Instruments, Abingdon, UK) and an iTC502 temperature controller (Oxford Instruments). All data was acquired using quadrature detection. Sample Y588R1/N624R1 was measured at ~33.7 GHz using an EN5170D2 resonator (Bruker).

Pulse settings for PELDOR/DEER: PELDOR/DEER experiments were performed using the standard four-pulse sequence $\pi/2(\nu_A)$–τ_1–$\pi(\nu_A)$–(τ_1+t)–$\pi(\nu_B)$–(τ_2-t)–$\pi(\nu_A)$–τ_2–echo. The magnetic field was set to the maximum of the nitroxide spectrum and the pump pulse was applied at ν_B (34 GHz) in the center of the cavity. The detection frequency ν_A was set lower than ν_B (see offsets in Table S10). Detection pulse lengths and the microwave power were adjusted to obtain $\pi/2$ and π-pulses ($\pi/2$ = 16 ns, π = 32 ns). The length of the pump pulse was adjusted by a transient nutation experiment in such a way that a maximal inversion of the Hahn echo was obtained (12 ns – 16 ns). For the inter-pulse delay τ_1, the initial time value was set to the first maximum in the two-pulse ESEEM trace (232 ns). In order to suppress deuterium ESEEM in the PELDOR/DEER time trace, a modulation averaging procedure was applied, incrementing τ_1 eight times by 16 ns and summing the individual traces. The dipolar evolution time τ_2 was set with regard to the distance-dependent modulation frequency and the need to resolve at least 1.5 oscillations in the time trace. The time step Δt in the time traces was chosen depending on the oscillation frequency so that a sufficient resolution of the dipolar oscillations was obtained. The number of scans was set as to obtain a decent signal-to-noise ratio of the time trace, permitting reliable data analysis. The integration gate was positioned symmetrically around the refocused Hahn echo. The gate width was adjusted to the length of the longest pulse in the sequence, herein the π-pulse at the detection frequency ν_A. For every point on the time trace, 20 shots were averaged and the shot repetition time (SRT) was adjusted in such a way that ~80% of the full echo intensity could recover before starting the next cycle of the pulse sequence. Two-step phase cycling of the $\pi/2$-pulse was used to eliminate

undesired echoes and receiver baseline offsets. Table S10 summarizes the parameters used for the individual samples.

Table S10. PELDOR/DEER acquisition parameters from lab G

	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
$(\pi/2)_A$ (ns)	16	16	32	16
π_A (ns)	32	32	32	32
Offset (MHz)	80	80	70	80
π_B (ns)	16	16	12	16
τ_1 (ns)	232	232	220	232
τ_2 (ns)	1400	2800	7000	8000
Δt (ns)	8	12	32	24
SRT (ms)	5	5	4	5
Shots per Point	20	20	20	20
No. of averages	8	16	64	16

S1.2.1. Quantification of the signal-to-noise ratio

The signal-to-noise ratios (Table S11) were determined with respect to modulation depth Δ (Table S12) by considering the standard deviation of the imaginary part after phase correction as the noise level.[4] For the purely real data from laboratory E, the standard deviation of the Tikhonov fit residual from ComparativeDeerAnalyzer was used instead. The latter procedure was also used for the data of S353R1/Q635R1 from laboratory D, where noise level differed strongly between real and imaginary part. In the other cases, the two estimates were in very good to reasonable agreement, with the fit residual SNR being somewhat lower due to an imperfect fit at the end of some data traces.

Table S11. Summary of the signal-to-noise ratios (SNR)[a]

Lab	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
		SNR		
A	187	322	144	233
B	210	157	64	135
C	166	400	201	195
D	474	721	355	144
E	287	232	87	139
F	148	95	35	158
G	93	93	53	56

[a] The SNR was calculated as $\frac{\Delta}{\sigma_N}$. Δ is the modulation depth and σ_N is the standard deviation of the imaginary part from zero, except for lab E and S353R1/Q635R1 from lab D, where it is the standard deviation of the Tikhonov fit residual in ComparativeDeerAnalyzer.

Table S12. Summary of the modulation depths Δ

Lab	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
		Δ (%)		
A	41	47	48	45
B	36	43	41	40
C	35	38	36	36
D	26	36	31	35
E	30	36	30	27
F	26	25	29	19
G	22	27	32	32
average:[a]	31	36	35	33
error:[b]	7	8	7	9

[a] Average value of the parameter in the respective column. [b] Standard deviation of the average value in this column.

S1.2.2 Assessment of data analysis quality with different approaches as a function of signal-to-noise ratio

For comparison of data analysis approaches and for an assessment of the influence of SNR on fidelity of the distance distributions, we designed a set of 75 test cases, independent on the presented experimental ring test. The moderate size allows for inspecting each individual output by eye and at the same time is sufficient for obtaining statistically valid results. The first 25 distance distributions were designed to present challenging tests with respect to resolution of several Gaussian peaks, superposition of narrow and broad components, and shape recognition of broad asymmetric distributions. The remaining 50 distributions were derived from an ensemble model of the RNA-binding protein PTBP1 in complex with an internal ribosome entry site of encephalomyocarditis virus by selecting spin label site pairs from a set of labeling sites used in an ongoing experimental study. Length of the time traces was adjusted to the maximum distance by the equation $t_{max} = 2(r_{max}/40)^3$ μs/Å that corresponds to slightly more than two full oscillations of the dipolar frequency ($\theta = 90°$ in Eq. (1) in the main text), but was limited by an upper bound of 12 μs and a lower bound of 2 μs. We allowed 5% of the total distance distribution to exceed r_{max}. Origin of the time axis was at -100 to -144 ns. The modulation depth ranged between 0.2 and 0.45. Monoexponential background decay with time constants between 20 and 45 μs was assumed. Within these ranges, modulation depth and background decay time constant were selected randomly. Normally distributed random numbers were added to achieve SNRs of 4, 8, 10, 16, 32, 50, and 100 with respect to modulation. Two other sets of test signals were simulated from the same noiseless time-domain traces with random SNR between 5 and 15 (high-noise set) and between 100 and 300 (low-noise set). The complete set of test data, including ground-truth distance distributions and simulated time-domain traces at the different SNRs is available in Spinach 2.6.5608 and later in folder \examples\deernet\noise_test. The whole package can be downloaded at spindynamics.org.

The low-noise and high-noise sets with random SNR were analyzed by DEERNet 2.0, by Tikhonov regularization in DeerLab using the AIC for determining the regularization parameter, and by multi-Gauss fitting in DD using the AIC for determining the number of Gaussian fit components (Table S13).

Table S13. Comparison of analysis approaches on 75 test data sets with known ground truth

	Neural network	Tikhonov regularization	Multi-Gauss Fit
Low-noise $\bar{\Delta}_2$ [a]	0.011 (0.004)[b]	0.020 (0.014)	0.010 (0.005)
Low-noise $\bar{\Delta}_o$ [c]	0.051[c] (0.028)	0.065 (0.075)	0.037 (0.027)
High-noise $\bar{\Delta}_2$ [a]	0.031 (0.016)	0.023 (0.009)	0.027 (0.015)
High-noise $\bar{\Delta}_o$ [c]	0.144 (0.069)	0.124 (0.056)	0.117 (0.063)

[a] Root mean square average [b] Values in parentheses are standard deviations [c]By geometric mean of overlap indices.

Fidelity of the output distance distributions $P_{\mathrm{out}}(r)$ was assessed by interpolating these distributions to the same distance axis as the ground-truth distribution $P_{\mathrm{gt}}(r)$, normalizing both $P_{\mathrm{out}}(r)$ and $P_{\mathrm{gt}}(r)$ to unity sum of elements, and computing the two-norm

$$\Delta_2 = \left\| P_{\mathrm{out}}(r) - P_{\mathrm{gt}}(r) \right\|_2 \qquad \text{(S1)}$$

as well as the overlap

$$o = \sum \min\left\{ P_{\mathrm{out}}(r), P_{\mathrm{gt}}(r) \right\} \qquad \text{(S2).}$$

Overlap o ranges between 0 (disjoint distributions) and 1 (identical distributions). In ensemble structure modeling, the overlap deficiency $\Delta_o = 1 - o$ is minimized.[8] For assessing the complete test set, we used the root mean square of Δ_2 and the geometric mean of o, defining the mean overlap deficiency as $\bar{\Delta}_o = 1 - \left(\prod_{k-1}^{n} o_k\right)^{1/n}$, where $n = 75$ is the number of individual data sets.

In general, all three automated approaches perform on a similar level. Inspection of individual cases shows that limitations in DEERNet 2.0 performance arise from problems with background separation at low SNR. For Tikhonov regularization, performance is limited by underestimates of the regularization parameter by the AIC at high SNR, which causes a breakup of the distribution into many narrow peaks and leads to unrealistically large uncertainty estimates. For multi-Gauss fitting, performance is limited by occasionally appearing very narrow Gaussian components, which are more detrimental with the two-norm metric than with the overlap metric. These narrow components can also lead to unrealistically large uncertainty estimates. The comparative analysis described in Section S2 combines neural network and regularization analysis with the aim to reveal problematic cases.

We studied the dependence of distance distribution accuracy on SNR with DEERNet Rev 5608, DeerLab Tikhonov regularization with the regularization parameter fitted to DEERNet resolution (ComparativeDeerAnalyzer), and MultiGauss fitting by DD using two metrics. Down to an SNR of 16, the increase in the 2-norm and decrease in overlap are small compared to the variation of the two metrics among the 75 test data sets (Figure S6). We still consider performance as acceptable at an SNR of 10, where overlap deficiency and 2-norm increase by a factor of two with respect to the low-noise limit (Figure S6b). In structure or ensemble modeling, overlap between the distribution simulated from the structure and the experimental distribution is limited by additional uncertainties, such as excitation band overlaps, phase drifts, and errors in spin label modeling. Therefore, geometric mean overlap deficiencies lower than about 0.2 may indicate

overfitting. Note that at an SNR of 4, all approaches perform much worse than that. Whereas measuring to an SNR much larger than 20 appears to be unnecessary, data with an SNR substantially below 10 should not be analyzed in terms of distance distributions.

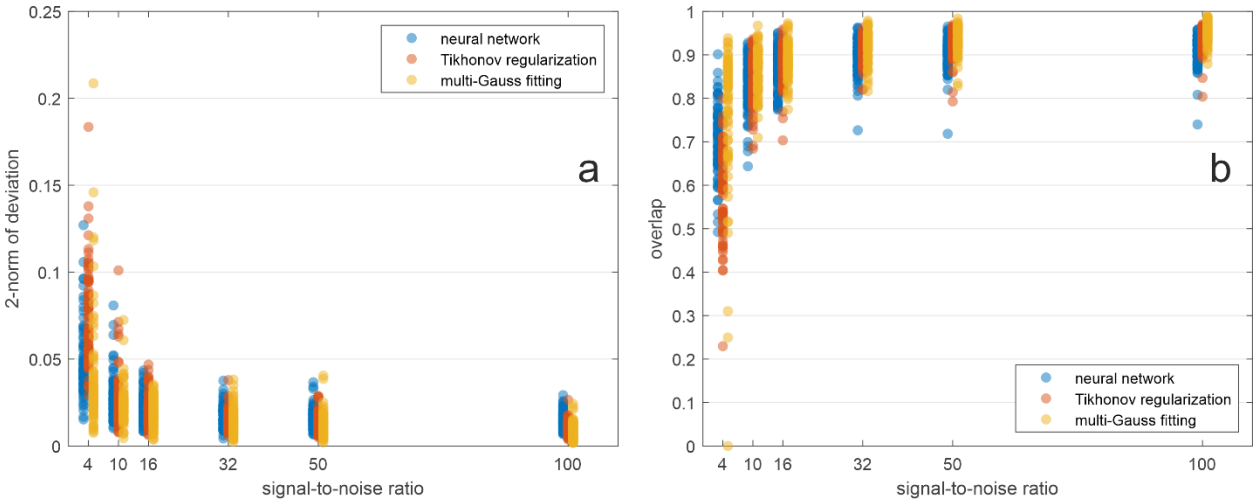

Figure S6. Dependence of distance distribution accuracy on signal-to-noise ratio (simulations). The set of 75 time-domain traces was simulated from 25 parametric distance distributions and 50 distance distributions generated from a protein ensemble model with modulation depths between 0.2 and 0.45 and exponentially decaying background with time constants between 25 and 50 µs. Data were analyzed with DEERNet rev 5608 neural network (blue), DeerLab Tikhonov regularization by ComparativeDeerAnalyzer (red), and DD multi-Gauss fitting (ochre). (a) Scatter plot of the Euclidian norm (2-norm) of the deviation between ground truth and recovered distance distributions. (b) Overlap between ground truth and recovered distance distribution.

S1.2.3. Assessment of the background decay

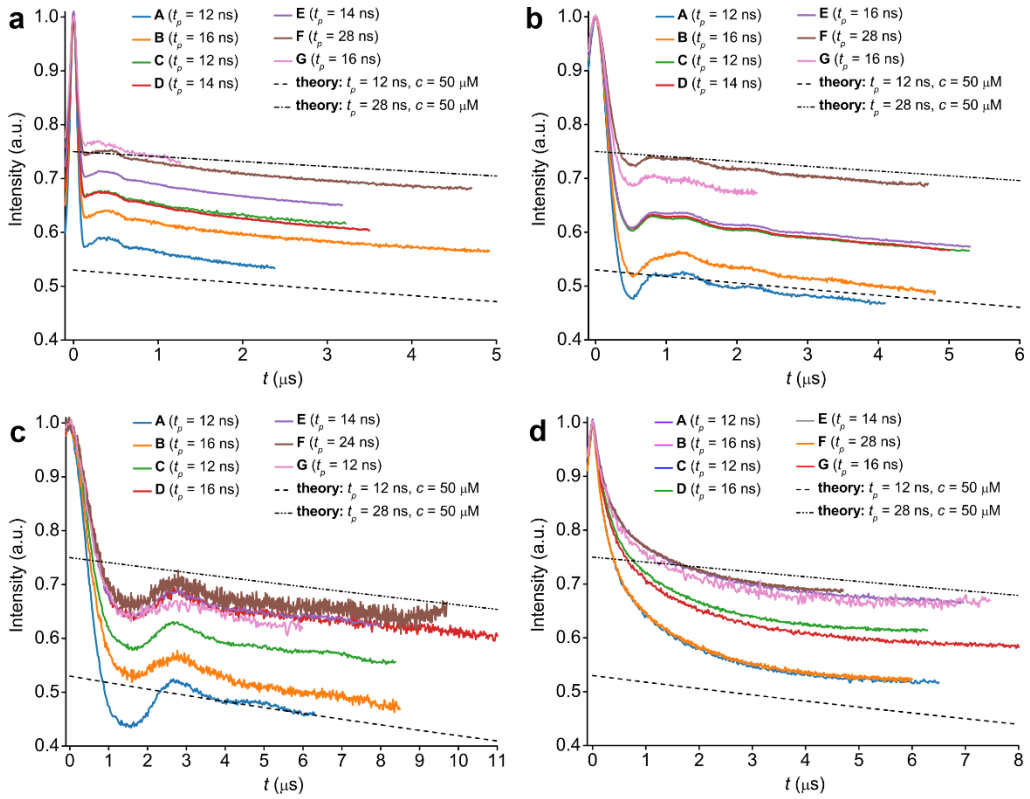

Figure S7. Evaluation of the background decays. (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1. The time traces are color-coded according to the labs A-G that measured them. Taking the sample's bulk spin concentration of 50 μM and the pump pulse lengths of 12 up to 28 ns into account, the dashed and dashed-dotted decay curves for the homogenous 3D background were calculated, using equation (3) in the main text.

S1.3. Data analysis with Tikhonov regularization as implemented in DeerAnalysis

Lab A: Distance distribution analysis of lab A was performed in DeerAnalysis 2019. The primary DEER data were zero-time corrected and cropped to remove the "2+1"-artefact at the end of the trace (S585R1/Q603R1 from 2696 to 2380 ns, V599R1/N624R1 from 4692 to 4100 ns, Y588R1/N624R1 from 7820 to 6500 ns, S353R1/Q635R1 from 7680 to 6500 ns). The data were then background-corrected by a monoexponential background function. Distance distributions were obtained by Tikhonov regularization and selecting the regularization parameter by the Generalized Cross Validation criterion, or manually by selecting the corner of the L-curve when automatic selection was leading to obvious over- or undersmoothing of the distance distribution. The results were validated with respect to the start of the background fit and to white noise using default settings of DeerAnalysis.

Lab B: DEER data were analyzed in DeerAnalysis2019. The zero time was set to 120 ns. The "2+1" artifact was removed by cutting 800 ns from sample A, 1200 ns from the traces of sample B, C, and D. 1/3 – 2/3 of the trace length was used for the background fitting with homogeneous background function (D=3). For data validation, the background length was varied within 1/3 – 2/3 of the trace length and the background dimension within D = 2 to D = 3, in total 30 trials.

Lab C: PDS experiments were analyzed using DeerAnalysis2015. PELDOR data were first cut to remove an end-of-trace artefact (sample A at 3200 ns, sample B at 5300 ns, and sample C at 6300 ns) before background-correction using a 3-dimensional homogeneous background function. Zero time and background start time for the first Tikhonov regularization were chosen using the default ("!") function in DeerAnalysis except for sample C, where a background start time of 3600 ns was chosen by eye to avoid a very steep background suggested by the default option. Tikhonov regularization was followed by statistical analysis using the validation tool in DeerAnalysis2015, varying background start from 5 to 80% of the trace length in 16 trials and including the addition of 50% random noise in 50 trials, resulting in a total of 800 trials. Validation trials were pruned with a prune level of 1.15, where trials exceeding the root mean square deviation of the best fit by at least 15% are discarded. Resulting background start time for the best fit was then used as starting point for a second round of Tikhonov regularization. The distance distribution including 95% confidence estimate obtained from the data validation and the best-fit background-corrected trace obtained from the second Tikhonov regularization were used to generate the plots.

Lab D: PELDOR/DEER time traces were transformed into distance distributions by the DeerAnalysis 2019 toolbox. Phase correction was done by the program automatically upon loading the data. The zero-time was set to coincide with the maximum of the time trace and the "2+1"-artefact at the end of the trace was cut off. Background removal was performed assuming a 3D-homogeneous distribution of nano-objects. The starting time of the background fit suggested by DeerAnalysis was used as initial guess and was altered as to obtain an artefact-free Pake pattern. Distance distributions were computed by Tikhonov regularization with a regularization parameter α located within or close to the corner of the L-curve. A statistical uncertainty analysis was done by the validation routine of DeerAnalysis, varying exclusively the starting point of the background fit, and keeping the remaining parameters (background dimensionality, background density, and modulation depth) at the respective default value.

Lab E: The PELDOR time traces were analyzed using Tikhonov regularization in combination with homogeneous three-dimensional background (i.e., exponential decay) fitting as implemented in DeerAnalysis 2019. The regularization parameter was chosen using the L-curve criterion as implemented in the software. The zero-point of the time trace and the starting-point of background fitting were selected using the automated software feature. The final 700 – 900 ns of the time traces were truncated after observing artifacts due to the occurrence of 2+1 effect. The validation tool was used to estimate the uncertainty in the distance distributions. For this, the white noise was varied in five steps at a level setting of 1.25, the background starting point was varied in 11 steps in a region spanning 0.2 – 0.6 of the time trace's length, and the background dimensionality was varied in 11 steps between values of 1.00 and 3.00. This set-up resulted in a total of 605 simulated time traces for validation. Subsequently, simulations which did not meet the prune-level criterion (i.e. rmsd values $\leq$ 1.15-times the rmsd value of the best fit) were not taken into account for the uncertainty estimates of the distance distributions.

Lab F: All DEER time traces were analyzed with the DeerAnalysis2018 software package for Matlab. The built-in functions of DeerAnalysis were used to perform phase correction of the raw data, background correction, and setting of the zero-time. For the background correction, a function that corresponds to a three-dimensional homogeneous distribution of spin-labeled objects was taken and fitted to the last 2/3 of the whole DEER time trace. Dividing the primary DEER

time trace through the background function yielded a form factor, which was further analyzed with Tikhonov regularization. The distance distribution at the optimum alpha-parameter (corner criterion) was subjected to validation. The validated distance distribution was a result of 81 trials where the background start and noise level were varied.

Lab G: The PELDOR time traces have been analyzed using DeerAnalysis Version 2019. The loaded time traces are phase corrected (build in) and the 'time shift' was set to 152ns (predetermined setting for the used pulses/delays). In some cases, when the time traces rises at the end of the time trace, the time trace was shortened to remove/reduce the uprising end. The Tikhonov regularization L-curve was calculated and the optimum to the Lc criteria selected. The distance distribution was verified by the standard setting (white noise and start of background variation with default settings) and as an error estimation used. The background dimension was always set to 3 dimensional. All other settings in DeerAnalysis have been used by its standard values.

The individual steps of the data analysis (background correction, L-curve and alpha-value, Fourier transforms) of each individual laboratory and an evaluation of the experimental background are collectively shown in Table S14 and Figures S8 – S10. The large spread in regularization parameters seen in Table S14 results partially from different SNR and trace length and partially from the use of different selection criteria. The different positioning of the L curve of laboratory C in Figure S10 results from a different scaling of misfit and roughness in DeerAnalysis2015 compared to later versions.

Table S14. Summary of α-values

Lab	S585R1/Q603R1	V599R1/N624R1	Y588R1/N624R1	S353R1/Q635R1
		α-value		
A (DeerAnalysis2019)	10	16	16	126
B (DeerAnalysis2019)	158	100	1000	3000
C (DeerAnalysis2015)	1	1	10	1000
D (DeerAnalysis2019)	794	158	794	1259
E (DeerAnalysis2019)	158	50	126	398
F (DeerAnalysis2018)	630	1585	31623	6310
G (DeerAnalysis2019)	63	126	200	316

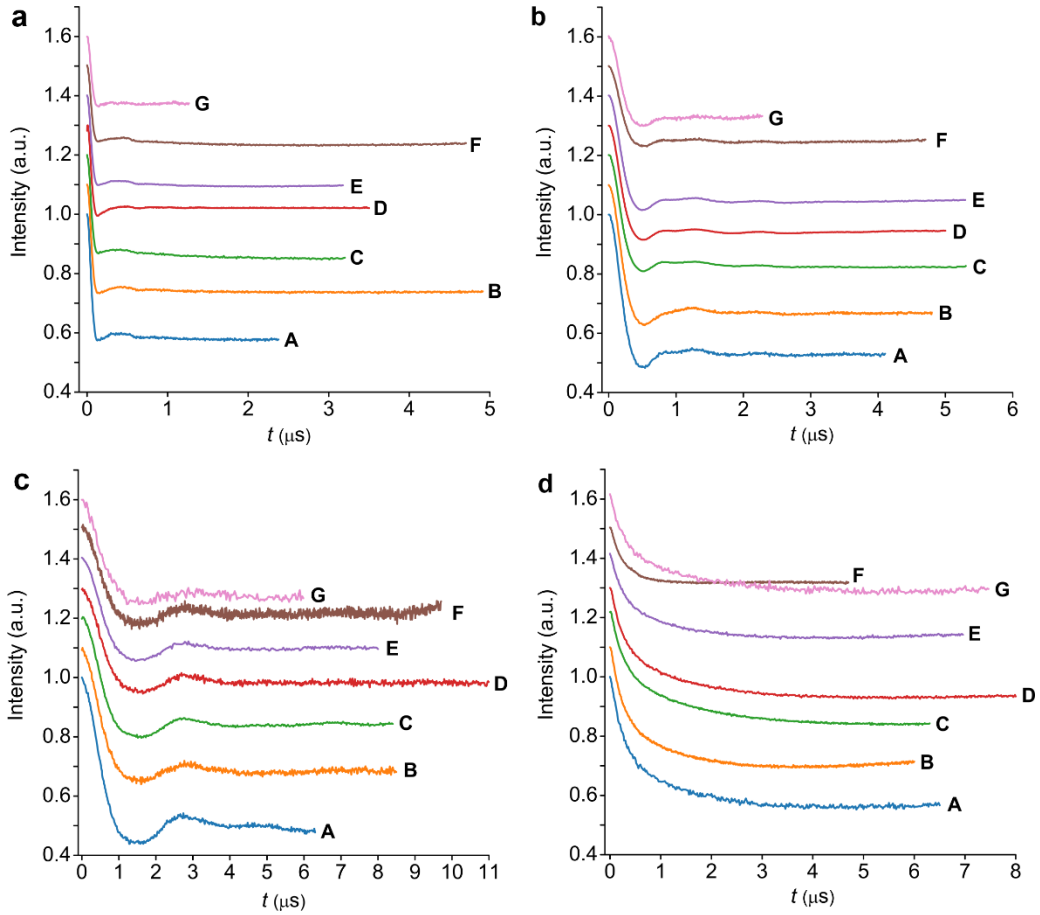

Figure S8. Background corrected PELDOR/DEER time traces. Background corrected time traces for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1. The time traces are color-coded according to the labs A-G that measured them. The time traces are vertically offset for better visibility. The background corrections were done with DeerAnalysis (versions 2015 - 2019).[5]

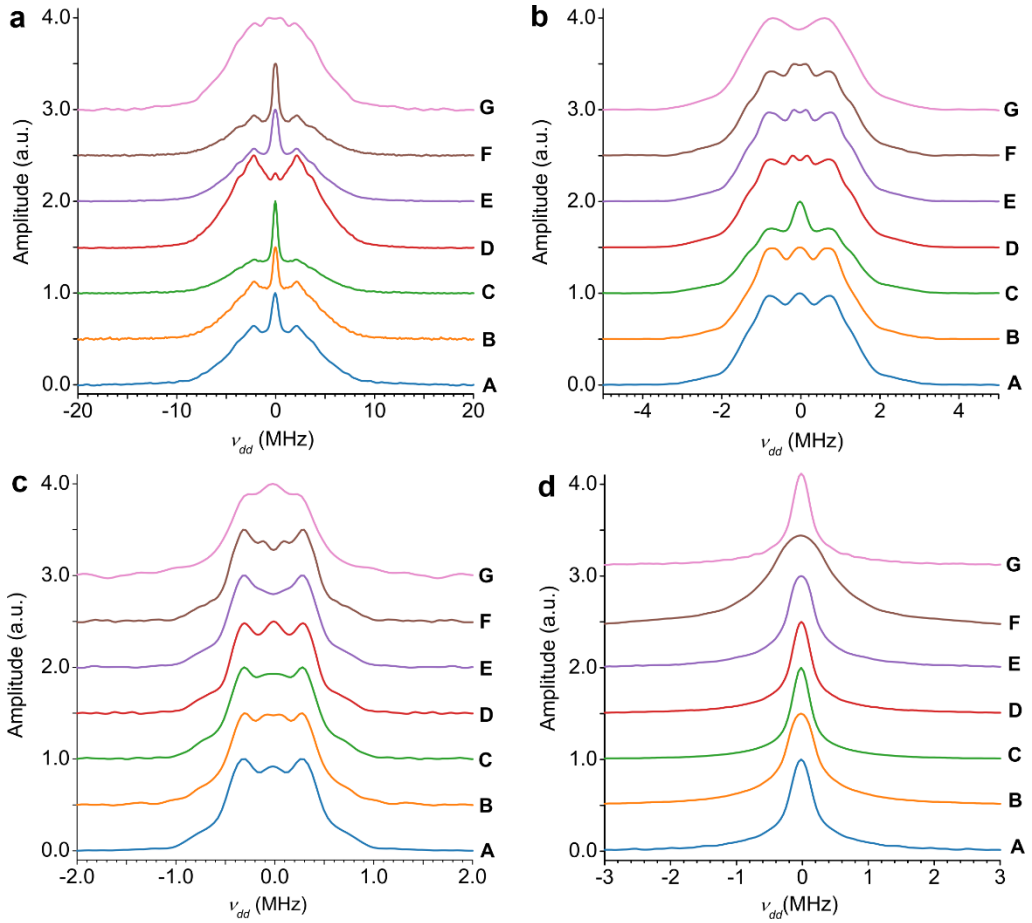

Figure S9. Fourier transformed data. Fourier-transformed spectra for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1. The spectra are color-coded according to the labs A-G. The spectra are vertically offset for better visibility. The Fourier-transformed were generated in DeerAnalysis (versions 2015 – 2019).[5]

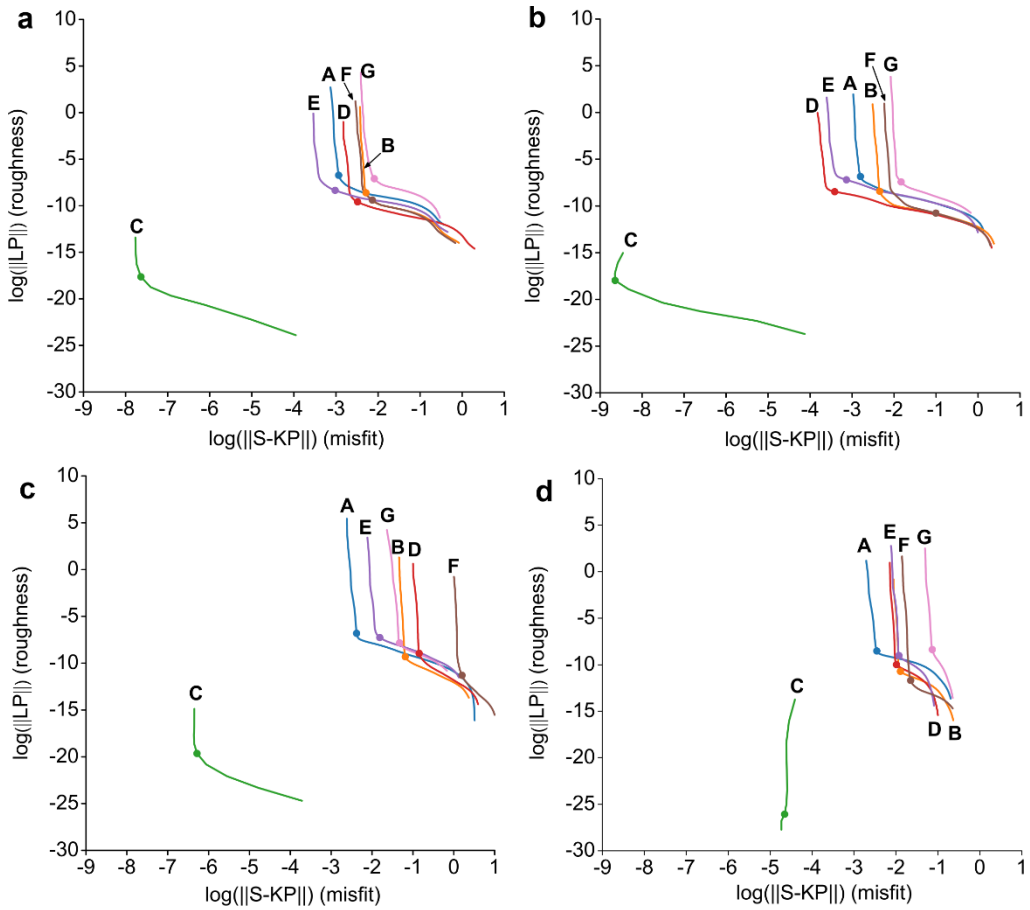

Figure S10. L-curves and α-values. L-curves for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1. The L-curves are color-coded according to labs A-G. The used α-values are depicted as full circles and are collected in Table S13. The L-curves were generated in DeerAnalysis (versions 2015 – 2019).[5]

S1.4. Determination of mean distance and standard deviation

Fitting of a parametrized distance distribution consisting of a single Gaussian peak and homogeneous distribution of remote spins in three dimensions proved to be the most robust and reliable way of determining mean distance and standard deviation of the distribution. These fits were performed in DeerAnalysis2021, after default phase and zero-time correction as they are automatically performed upon loading of data. For samples V599R1/N624R1, Y588R1/N624R1, and S353R1/Q635R1 from lab B, the first data point had too large amplitude and was removed before processing, because it upset automatic zero-time determination.

S2. Comparison between data analyses

S2.1. DEERNet 2.0 data analysis

Primary experimental data from all labs were processed by DEERNet 2.0[6] as provided. This analysis does not involve selection of any adjustable processing parameters. Uncertainty is computed as 95% confidence interval from variation between 24 separately trained neural networks. The results are shown in Figure S11. The software is part of Spinach V2.6 and is available as a Matlab toolbox at spindynamics.org.

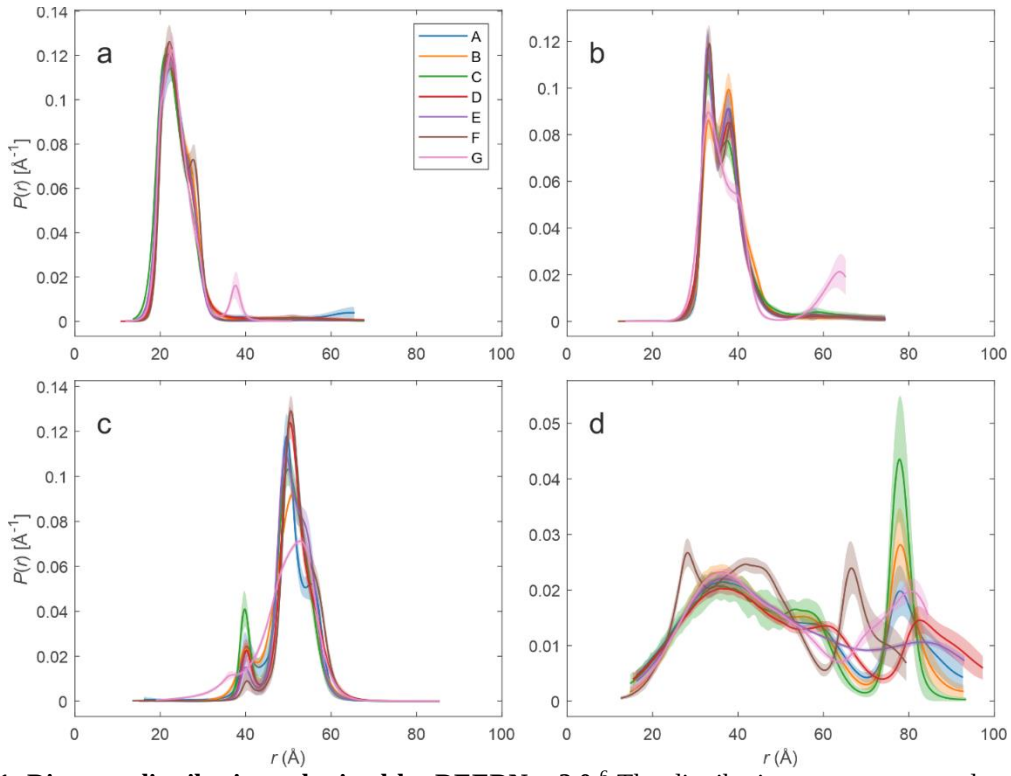

Figure S11. Distance distributions obtained by DEERNet 2.0.[6] The distributions were computed automatically from the full data primary experimental data sets provided in the ring test by labs A-G (color-coded). Distributions for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1.

S2.2. DD data analysis

Primary experimental data from all labs were processed by Multi-Gauss fitting including a homogeneous background with DD[4] as provided. The Bayesian information criterion was selected for determining the optimal number auf Gaussians. For the data of sample 1 from labs A and D, the number auf Gaussians was reduced from 5, as suggested by the criterion, to 4 in order to avoid an unrealistically large uncertainty. Uncertainties were computed as 95% confidence intervals. The results are shown in Figure S12. The software is available as a Matlab toolbox at https://lab.vanderbilt.edu/hustedt-lab/software/dd/.

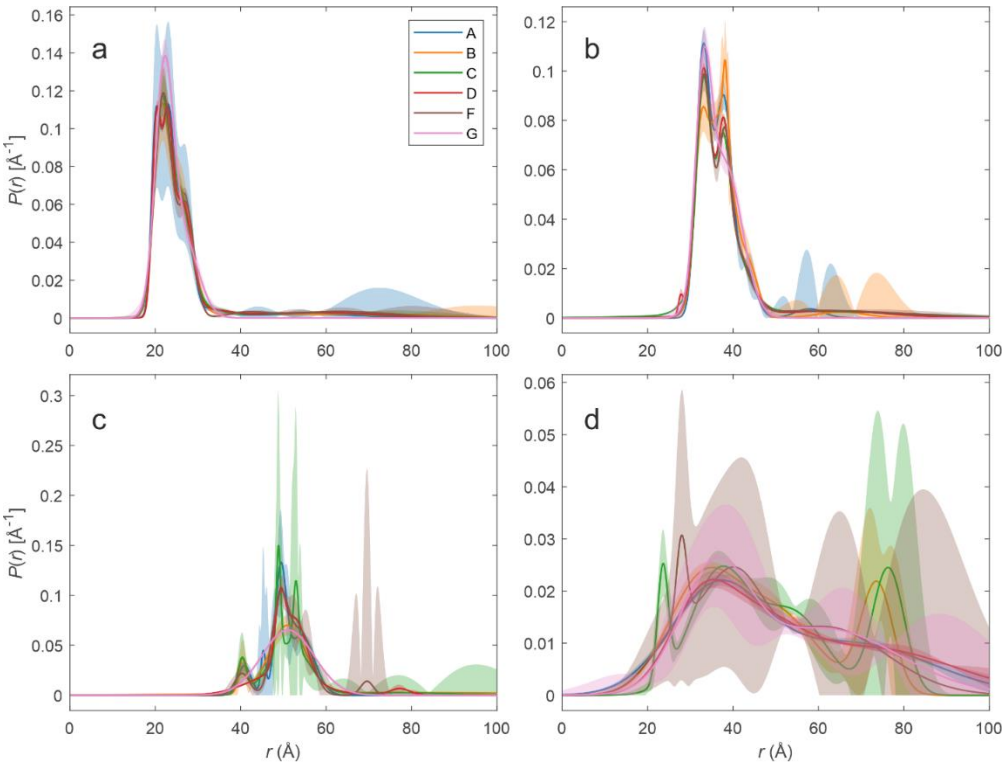

Figure S12. Distance distributions obtained by DD Multi-Gauss Fitting.[4] The distributions were computed automatically from the full primary experimental data sets provided in the ring test by labs A-D and F-G (color-coded). The number of Gaussians was determined by the Bayesian information criterion, except for data from labs A and D for sample 1, where the number of Gaussians was limited to 4 in order to avoid unrealistically large uncertainties. Data from lab E is missing since it contained only the real part of the signals, whereas DD depends on the imaginary part for the noise estimate required for determining the optimal number of Gaussians. Distributions for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1.

S2.3. DeerLab data analysis

Primary experimental data from all labs were processed by Tikhonov regularization with DeerLab v0.13.1.[7] The last 1000 ns were cut off to avoid unbalancing determination of the regularization parameter by end artifacts. For the shorter data from laboratory G, only the last 300 ns were cut off. The background and distance distribution were simultaneously fitted using separable non-linear least-squares. The Akaike information criterion (AIC) was used for determining the optimal regularization parameter, as is default in DeerLab. Uncertainties were computed by 1000 bootstrapping trials. The results are shown in Figure S13. The software is available as a Python package (see https://jeschkelab.github.io/DeerLab/beginners_guide.html).

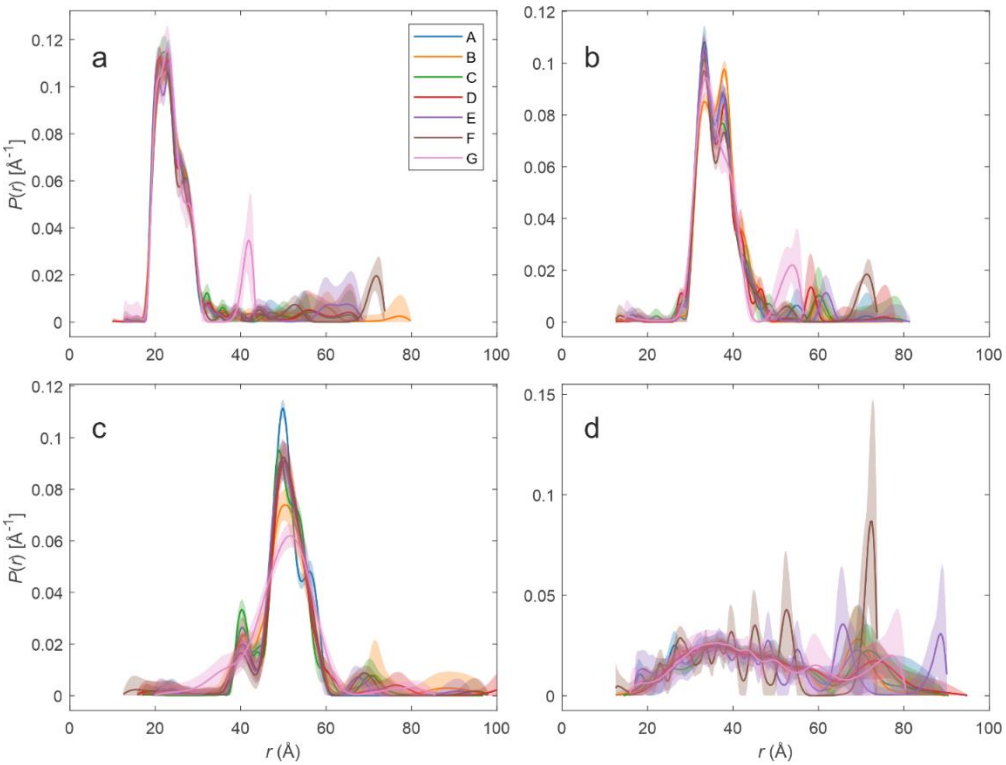

Figure S13. Distance distributions obtained by DeerLab[7] by Tikhonov regularization. The distributions were computed automatically from the primary experimental data sets provided in the ring test by labs A -G (color-coded). The regularization parameter was selected by the Akaike information criterion. Uncertainty estimates are 95% confidence intervals from 1000 bootstrapping samples. Distributions for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1.

S2.4. Comparative DEER data analysis

Primary experimental data from all labs were processed by ComparativeDeerAnalyzer as provided, except for samples 2-4 from lab B, where a spurious first data point was removed before processing, because it upset automatic zero-time determination. The processing involves automated neural network analysis with DEERNet[6] from Spinach revision 5501, automated determination of the regularization parameter that maximizes overlap between the distance distributions obtained by DEERNet, Tikhonov regularization with DeerLab routines,[7] and assessment of regularization uncertainty by combining DEERNet background uncertainty with 5 noise samples for each of 11 background trials. If DEERNet background is not minimal at zero time, mean background for regularization is determined by fitting an exponential decay to the last three quarters of the data trace. A mean distance distribution is computed together with uncertainty bands that include the uncertainties from both approaches. The results are shown in Figure S14. The software is available as Windows executable at https://epr.ethz.ch/software.html and is included as default processing mode in the Matlab program DeerAnalysis2021, which is available at the same homepage.

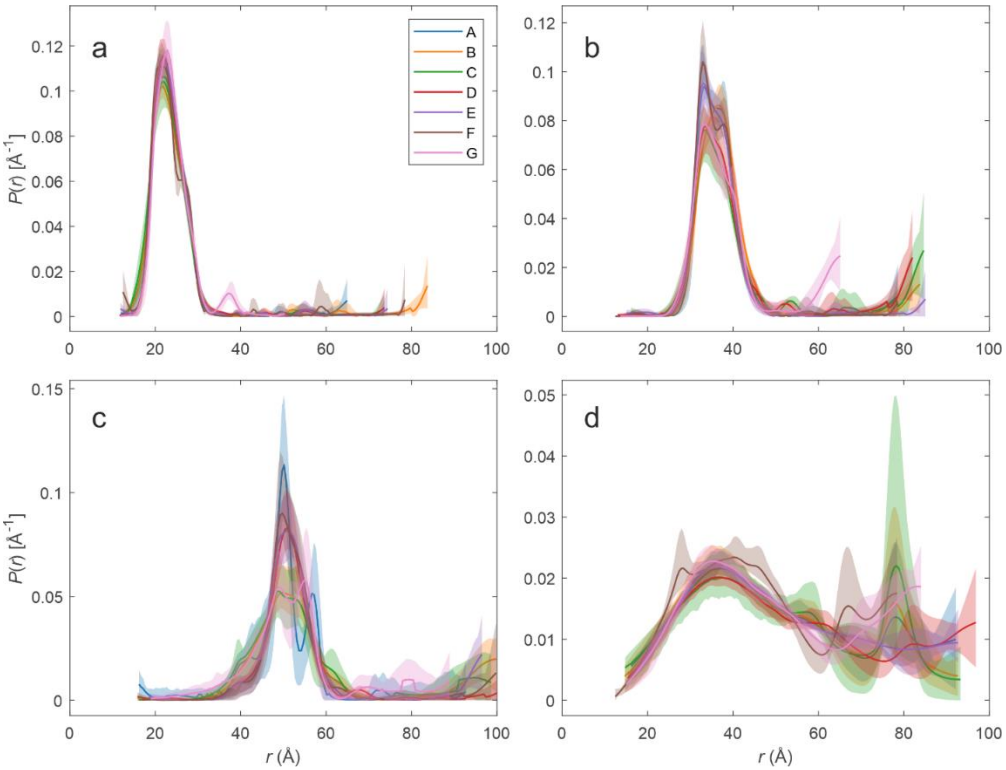

Figure S14. Distance distributions obtained by comparative DEER analysis between DEERNet and Tikhonov regularization. The distributions were computed automatically from the full primary experimental data sets provided in the ring test by labs A-G (color-coded). The regularization parameter for Tikhonov regularization was selected by maximizing overlap between the normalized probability density distributions from neural network and regularization analysis. Uncertainty estimates refer to the maximum of the upper-bound estimates and minimum of the lower-bound estimates from both separate analyses. Distributions for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1.

S3. In silico labeling

The distance distributions in Figure S15 were computed based on the two crystal structures of YopO (PDB-ID 2h7o; PDB-ID 4ci6) (Figure 1 and Figure S16).

mtsslWizard: With mtsslWizard (http://www.mtsslsuite.isb.ukbonn.de/),[9] the maximum number of conformers was set to 200 per labeling site, using a van-der-Waals cutoff of 2.5 Å and 5 allowed clashes ("loose" mode).

MMM: For computing distance distributions with MMM 2021.1 (https://epr.ethz.ch/software.html),[10] the default settings for labeling with MTSL were used.

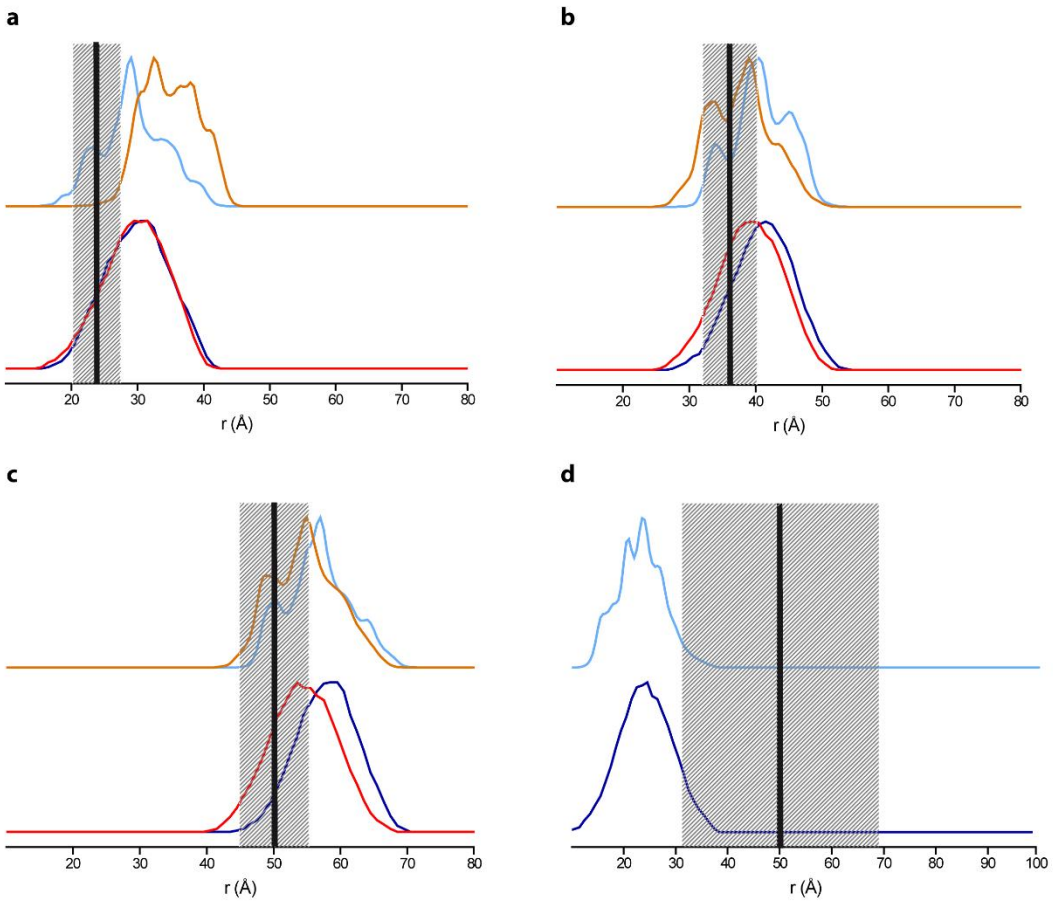

Figure S15. *In silico* derived distance distributions. Shown are the *in silico*-derived distance distributions for (a) S585R1/Q603R1, (b) V599R1/N624R1, (c) Y588R1/N624R1, and (d) S353R1/Q635R1. Top, the distribution obtained with MMM[10]; bottom, the distribution obtained with mtsslWizard[9]. The distributions in orange/red are based on the crystal structure with PDB-ID 2h7o. The distributions in light blue/dark blue are based on the crystal structure with PDB-ID 4ci6. The experimental mean distances are indicated by vertical black lines and the average widths by grey shaded areas.

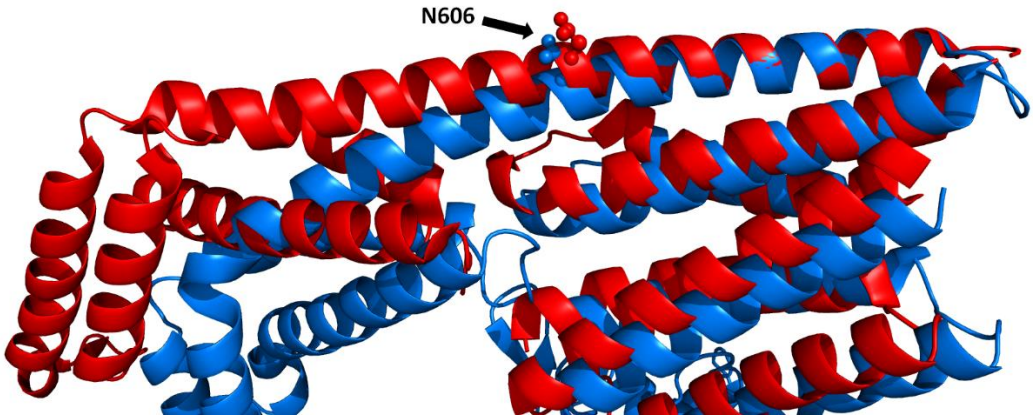

Figure S16. Comparison of the structures of α-helix 14. Red, structure as found in PDB-ID 2h7o; blue, structure as found in PDB-ID 4ci6.

S4. Detailed guidelines for good practice

S4.1. Parameters in the 4-pulse DEER sequence for optimal signal-to-noise ratio

Inter-pulse delay τ_1. This delay should provide a compromise between the echo decay due to the phase memory time T_m, and the negative time of t that allows for recording the DEER traces prior to its maximum, which gives a good measure for the t_0 point. This timing also depends on the spectrometer dead time; for example, for a commercial Bruker QT2 resonator at 34 GHz, the dead time is about 150 – 200 ns and a typical value of τ_1 is around 260 – 400 ns. T_m can strongly vary between biological samples and it is recommended to record a two-pulse echo decay ($\pi/2$-τ-π-τ-echo) or better, a refocused echo decay ($\pi/2$-τ_1-π-τ_1+τ_2-π-τ_2-echo) as a function of τ_2. Note that the 2-pulse echo decay trace should be analyzed as function of $2 \cdot \tau$ according to the expression $exp - (\frac{2\tau}{T_m})^{\xi}$ with ξ being an exponent between 0.5 and 3 that often ranges only between 1 and 2.[11]

The delay τ_1 might generate nuclear electron-spin echo modulation (ESEEM) effects that can be mistaken as dipolar modulations. At X-band, ^{1}H-modulation produces an artificial distance of about 15 Å, ^{2}H-ESEEM a distance of about 28 Å. At Q-band, the latter shifts to about 18 Å and is weaker. These effects can be severe if the two microwave sources are coherent and the excitation bandwidths of the pump and observer pulses overlap. Typically, an 8 to 10-step 'nuclear modulation averaging' should be applied, incrementing τ_1 by $\Delta\tau$ after each shot and adding up the traces. $\Delta\tau$ is calculated from the inverse of the observed ESEEM frequency divided by the number of steps.[12] Note that modulation averaging might not completely eliminate ESEEM effects.

Delay τ_2 *and* t_{max}. The delay τ_2 determines the maximum distance that can be detected as well as the SNR. Its length is limited by the sample's T_m. It determines the maximal length of the DEER trace (t_{max}) and should allow for the observation of as many modulation periods as possible. Ideally, t_{max} should be extended to record the signal after all modulation is damped, up to a length of at least a third of the trace.

SNR. The DEER SNR is defined as the ratio between the modulation depth Δ and the noise level, and is calculated by dividing Δ by the standard deviation of the noise of the trace. The modulation depth Δ can be obtained by extrapolating the background-fitting function to zero time, and the noise level can be estimated for quadrature-detected data from the imaginary part[4] as described above (recommended) or from the fit residual for purely real data (discouraged). Please note that some references define the SNR purely as the noise level and factor the modulation depth into the DEER sensitivity.

When considering SNR, longer t_{max} will result in a lower SNR, due to relaxation loss of the observer echo amplitude. The setting should be decided on a case-by-case basis with all relevant parameters considered and pre-knowledge on the system (concentration, labeling efficiency) along with an estimate of the longest expected intramolecular distance of the molecule. In some cases, an idea on the expected mean distance can be derived from a model of the structure under investigation, but be aware that the solution structure may differ considerably from crystal structures.

If relaxation time permits, the evolution time t_{max} should be set long enough for the dipolar modulation to decay to the noise level within $2t_{max}/3$. In order to determine at least a mean distance, one needs to ensure

$$t_{\max} \geq 1.9 \times 10^{-5} \cdot r_{max}^3 \frac{\mu s}{\text{Å}^3}, \qquad \text{(S1)}$$

corresponding to observation of at least one full dipolar oscillation. In order to estimate at least the width of the distribution, one needs to ensure

$$t_{\max} \geq 3.8 \times 10^{-5} r_{max}^3 \frac{\mu s}{\text{Å}^3}, \quad \text{(S2)}$$

corresponding to observation of at least two full dipolar oscillations with angular frequency ω_{AB} as given by Eq. (1) in the main text with θ = 90° corresponding to the singularities of the Pake pattern, where r_{max} is the maximum distance where the analysis can be considered as reliable. These are rules of thumb. Particulars depend on the distance distribution itself, on signal-to-noise ratio, and on the slope of the intermolecular background. In general, reliability of a distance distribution can only be assessed after processing of the data and both the uncertainty band and the limits given by equations (S1) and (S2) should be taken into account.
Analysis of test data sets with known ground truth has revealed that below an SNR of 10 with respect to modulation depth, data analysis becomes unreliable (Section S1.2.2, Figure S6). Together with long-term stability of a spectrometer, which limits measurement time to about 48-72 h, this sets an upper limit to $t_{\max}$ and, therefore, to $r_{\max}$. This upper limit depends on phase memory time T_{m}. Note that in case of a broad distribution of conformers, different subensembles might experience different T_{m}. A long τ_2 may bias their weighting. If this is suspected, the dependence of the form factor on τ_2 should be checked.[13]
If there is no *a priori* estimate of the distance, $r_{\max}$ can be estimated from a preliminary measurement with the longest τ_2 that is attainable at reasonable SNR. In a second measurement, τ_2 can then be adjusted appropriately to achieve a good SNR. In the rare cases where a very good SNR can be achieved in 1-2 hours, it is advisable to increase τ_2, as this leads to more reliable background separation and better resolution of the shape of the distance distribution.

Time increment Δt. This increment determines the highest modulation frequency that can be observed. In the limit of short distances, for instance for 15 Å, $\nu_{\perp} \backsimeq 18$ MHz, the Nyquist criterion for the shoulders of the Pake pattern at $\nu_{||} = 2\nu_{\perp}$is satisfied by Δt < 13.9 ns. Oversampling i.e., using a shorter Δt, is generally advantageous in order to avoid aliasing of noise. Although longer Δt reduce processing time during data analysis, we discourage such optimization of analysis time at the cost of a reduction of the SNR. If necessary, processing time can be reduced by prolonging Δt not during the measurement, but in a pre-processing step that involves digital filtering, which avoids noise aliasing. A too long Δt can distort the maximum of the trace and make determination of t_0 less reliable. Therefore, Δt should not exceed 32 ns.

Integration time t_{gate}. This window should be as long as the length of the observer π-pulse with the gate centered at the echo maximum. Note that sensitivity is reduced for much longer or much shorter integration gates.

Sequence repetition time (SRT). This time determines the total signal averaging time. It depends on the spin lattice relaxation time T_1, which in turns depends on the sample temperature. For maximum sensitivity, $1.3 \cdot T_1$ gives the best compromise between fast averaging and signal recovery. However, in order to make contributions to the echo quantitative, the SRT should be a

factor 3 up to $5 \cdot T_1$. While T_1 continuously increases as the temperature decreases, T_m remains approximately constant below 50 K and shortens dramatically above 70 – 80 K.[1] For nitroxides in water/glycerol matrices, a temperature around 50 K is therefore optimal. The typical repetition time at 50 K is 4-5 ms. For different solvents or matrices, it may be useful to perform an inversion recovery experiment (π-T-π/2-τ-π-τ-echo). Alternatively and faster, the observer echo can be observed in the digitizer window while SRT is varied. It is acceptable to reduce SRT until one loses a few percent, but certainly not more than 20% of the maximum echo intensity.

Phase cycle for removal of receiver offset. An unwanted offset in the receiver would distort the intensity of the detected echo signal and would not report the correct modulation depth. This can be eliminated by a phase cycle [+(+x)-(-x)] of the first observer pulse and is always recommended.

Total acquisition (or accumulation) time. The total acquisition time of the DEER experiment is approximately given by the product of the SRT, the number of shots per point (SPP), the number of points in the DEER trace (*N*) and the number of averaged scans (counting also phase cycling and 'nuclear modulation averaging'). Standard measurement times for a DEER experiments are often between 12 and 24 hours. Depending on spectrometer stability, longer accumulation times might not be beneficial and might introduce signal distortions and sensitivity loss due to phase or temperature drift. Longer times might be accompanied by a pseudo-2D data acquisition, for which individual scans or scan-batches are saved individually. Such a procedure also allows for statistical error analysis.

Pulse lengths and difference between observer and pump frequency. The shortest accessible distance is determined by the length of both, the observer and pump pulses, with 32 ns observer pulses setting a limit around 18-20 Å[14] and all 12 ns pulses[1] a limit around 15 Å. Ideally, there should be no spectral overlap between the excitation profiles of the pump and probe pulses. However, as can be seen in Figure 1B, some interference between the pump and probe pulses might occur, depending on the chosen pulse lengths and frequency offsets. This leads to a reduction in the remote echo intensity as soon as the pump pulse is turned on and, more important, to an additional dipolar signal contribution at the end of the trace that complicates data analysis. The pump pulse excitation efficiency determines the modulation depth of the observed dipolar time trace for spin pairs and the slope of the intermolecular decay function. Because the excitation bandwidth $\Delta\nu_{ex}$ has to be larger than the strength of the dipolar coupling between the two spins, the pulse length t_p sets a lower limit to the accessible distance range. While a short pump pulse provides larger modulation depth and excites spin pairs at shorter distances, a too short pulse will have an excitation bandwidth overlapping with that of the probe pulses. A minimal pulse length of 12 ns for all pulses combined with 80-100 MHz frequency difference presents an acceptable compromise.[1]

S4.2. Resonator and excitation profiles

For optimal performance of the experiment, proper adjustment of the turning angles of all pulses along with their excitation profiles is essential. Both the turning angles and excitation profiles depend on the microwave excitation power and on the resonator profile. The latter is determined by the quality factor *Q*, described by the resonator resonance frequency ν_r and its bandwidth $\Delta\nu_r$ according to: $= {}^{\nu_r}/_{\Delta\nu_r}$.

The microwave resonator bandwidth has to be broad enough to accommodate both the pump and probe frequencies (ν_A and ν_B) without disturbing the shape of the rectangular pulses too strongly. This is fulfilled if the rise-time of the resonator $t_r = Q/(2\pi \cdot \nu_r)$ is shorter than the pulse length t_p.
The pulse turning angle can best be optimized by performing Rabi nutation experiments (β-T-π/2-τ-π-τ-echo, with the first pulse being varied in flip angle β by incrementing its length) at the specific pump and observe frequencies (ν_B and ν_A, respectively). For a rectangular pump pulse (π-pulse) of length t_p, the excitation bandwidth $\Delta\omega_{ex}$ is roughly given by $1/t_p$. The excitation profile for the remote echo detection sequence is somewhat more complicated,[14] but the profile of the observer π-pulse can be taken as an estimate for the detected spin packets. An example with pump and observer pulse lengths of 16 ns and a frequency offset of 100 MHz at Q-band frequencies is shown in Figure 1B.

S4.3. Spin labeling

Labeling efficiency may strongly depend on the choice of labeling sites. Labeling of buried residues should be avoided, since local structure can be distorted even if sufficient labeling efficiency is achieved. Spin labeling site scans with MMM[10] or equivalent software can reveal accessible sites. Functional or structurally important residues, such as those involved in metal binding or salt bridges, should also be avoided as labeling sites. The reducing agent e.g., dithiothreitol (DTT), has to be removed before the labeling step to avoid reducing the label and cleaving the formed disulfide bridge between label and protein. MTSL dissolved in DMSO is added in excess but should not exceed 10% of the total volume of the protein solution, because DMSO may affect the stability of proteins. After incubation (at 4 °C or ambient temperature and from a few hours up to overnight, depending on the protein), free label should be removed either via dialysis or chromatography. The protein concentration should then be determined by UV-vis spectroscopy, and the purity of the protein by mass spectrometry, HPLC, and/or SDS-PAGE. The labeling efficiency can be obtained by mass spectrometry, but requires that labeled and unlabeled proteins ionize with equal efficiency. Mass spectrometry is cumbersome for large proteins and can become difficult for membrane proteins. Instead, cw EPR experiments in combination with a reference sample of known spin concentration can be used to evaluate the amount of spin label in the sample. Alternatively, spin-counting experiments using the on-board tools provided by many spectrometers can be performed, which do not require a reference sample. Relating the spin label concentration to the protein concentration yields the labeling efficiency. In addition, the width and shape of the X-band cw EPR spectra measured in the liquid state will indicate whether all spin label is bound. It is good practice to test for unspecific labeling of the wt. If necessary, wt cysteine residues must be exchanged. Commonly, alanine or serine mutations are chosen, but a more intricate design of constructs can be advantageous (www.dezyme.com).

S4.4. Check for integrity of protein structure and function

The modification of the protein by mutagenesis and spin-labeling can affect its structure or function. Therefore, several control experiments are recommended to ensure structural or functional integrity of the spin-labeled mutants. Circular dichroism spectroscopy and UV-Vis based melting studies can confirm that the mutations and spin-labeling did not alter the secondary and tertiary structure of the protein. Small-angle X-ray Scattering (SAXS) can reveal changes in the radius of gyration and thus also changes in disordered domains. In some cases, functional assays are preferable. In case that the spin-labeled protein interacts with a partner protein, other biomolecules, or small ligands, cross-linking experiments, electrophoresis mobility shift assays,

pull-down experiments, binding kinetics, and Isothermal Titration Calorimetry (ITC) can be performed. In case that the integrity of higher-order assemblies or the morphology of the studied system are to be assessed, microscopy tools such as Transmission Electron Microscopy (TEM) or Atomic Force Microscopy (AFM) can be helpful. The latter can also be used to verify reconstitution of the spin-labeled protein into membrane systems.

S4.5. Cryoprotection

In order to obtain homogeneously distributed biomolecules in the frozen sample,[15] a cryo-protectant is added to the aqueous solution, usually glycerol but sometimes also ethylene glycol or, for nucleic acids, DMSO. The cryo-protectant prevents the formation of ice crystals, which otherwise severely reduce T_m and thereby compromise the resolution and accessible range of distances. The amount of cryoprotectant added varies and depends on the protein and its environment, e.g., detergent, nanodiscs, or liposomes for membrane proteins. Measurements of the background decay showed that at 50% v/v a perfect glass is formed.[16] A lower amount increased background decay rates and decreased the phase memory time. The optimal amount is a trade-off between high resolution and maximum accessible distance *versus* viability of the system in the presence of high amounts of cryoprotectant. In many cases, 20% or even 10% can be sufficient. Agarose gel can also provide cryoprotection without compromising protein solvation.[17] In cases where the addition of cryo-protectant should be avoided altogether, rapid freezing using a rapid freeze quench (RFQ) set-up is recommended.[18,19] Extended phase memory times and a reduced background decay were obtained with RFQ, as compared to plunging the EPR tube into liquid nitrogen or, better, into isopentane cooled down by liquid nitrogen. Immersion of the tube into nitrogen-cold liquids does not lead to fast freezing on molecular timescales and, depending on the EPR tube, can take seconds, in contrast to RFQ where the freezing proceeds on the order of milliseconds or less. It has been reported that the two different freezing methods (rates) can yield altered distance distribution shapes, mainly manifested in the width as the samples represent different conformational ensembles.[18,19]

S4.6. Prolonging phase relaxation and diamagnetic dilution

Given good cryo-protection or very rapid freezing, the phase memory time T_m is affected most by the spin concentration and the degree of deuteration of both the solvent (D_2O/glycerol-d_8/ethylene glycol-d_6) and the biomolecule. The concentration influences T_m via electron–electron dipolar interactions contributing to spectral diffusion and instantaneous diffusion mechanisms.[20-22] Therefore, low concentrations are beneficial for a long T_m. As long as instantaneous diffusion is the dominant decoherence mechanism, reducing the concentration may even increase the SNR, since the loss in the number of spins is linear, but the gain in signal due to reduced decoherence is exponential in concentration. Further, a low concentration reduces the background decay contribution and thus the damping of the dipolar modulations. Thereby, it minimizes uncertainties in the identification of the background decay and thus uncertainties in the ensuing distance distribution. However, once decoherence is dominated by nuclear spin diffusion, further decreasing the concentration leads to a linear decrease in SNR. Typical spin concentrations that yield a good PELDOR/DEER SNR (>50) with a high power (150 W) Q-band setup within a reasonable time (<12 h) are 20 – 50 μM for 8 – 10 μs long evolution time. Since T_m depends on the local concentration of spin-labeled molecules, there are instances where diamagnetic dilution can be helpful, meaning the spin-labeled molecules are mixed with non-labeled molecules. This strategy also applies to aggregates and fibrils[23] and to proteins in membranes or membrane-

mimetic environments (except nanodiscs) where the local concentration of paramagnetic species should be reduced.[24] Diamagnetic dilution can also counter adverse effects of aggregation,[25] for oligomeric proteins it can reduce multi-spin effects[21] or help disentangle intra- and inter-molecular distances.[26] In the concentration range where nuclear spin diffusion dominates decoherence, deuterated solvents are advantageous as they lead to slower nuclear spin diffusion.[27-31] This increases the SNR or the accessible distance range and can be further extended by deuteration of the biomolecule.[32,33]

S4.7. Intermolecular background

If the biomolecules are not homogeneously distributed in three dimensions, the intermolecular background decay is not mono-exponential. This increases the uncertainty when separating $\boldsymbol{V}_{\mathrm{intra}}$ from $\boldsymbol{V}_{\mathrm{inter}}$ tremendously. In this situation, it is best practice to perform a PELDOR/DEER measurement under the same conditions using singly labeled samples (ideally the mixture of the corresponding single mutants) with twice the protein concentration of the doubly labeled samples (i.e., the same spin concentration). This may also be required for membrane proteins and fibrils,[20] where the distribution is not homogeneous in three dimensions and for in-cell samples, where the distribution may not be homogeneous in general. The same approach can address excluded volume effects in soluble biomolecules under standard buffer conditions.[34] Finally, this approach can allow identifying protein aggregation in the sample, manifesting itself in the appearance of a long and broadly distributed distance.[35]

S4.8. Pre-processing PELDOR/DEER data

PELDOR/DEER data should be detected in quadrature since the accuracy of phasing during the experiment setup is poor for weak signals. Since the standard kernel expects a real signal, phase correction is required. The ratio between the root-mean-square imaginary part (after phase correction) and root-mean-square noise determined by other means is a quality measure that should be reported if available. Ratios well above unity indicate a phase drift or detector saturation. In such situations, the real part of the dataset may deviate from kernel assumptions to an unknown extent. Such data should not be used.

The time axis in experimental DEER data is shifted relative to the theoretical dipolar evolution time. The offset is determined during data pre-processing or fitting, and must be reported. This may be difficult for noisy data – a good practice is to determine the offset by first running a high-quality sample with a well-defined distance of 20-30 Å on the same hardware with the same pulse program. At least 100 ns of negative dipolar evolution time is recommended, more if distances above 50 Å are present.

S5. References

(1) Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G. High sensitivity and versatility of the DEER experiment on nitroxide radical pairs at Q-band frequencies. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10762–10773.

(2) Tait, C. E. and S. Stoll. Coherent pump pulses in Double Electron Electron Resonance spectroscopy. *Phys. Chem. Chem. Phys.* **2016**, *18*, 18470–18485.

(3) Edwards, T. H.; Stoll, S. A Bayesian approach to quantifying uncertainty from experimental noise in DEER spectroscopy. *J. Magn. Reson.* **2016**, *270*, 87–97.

(4) Brandon, S.; Beth, A. H.; Hustedt, E. J. The global analysis of DEER data. *J. Magn. Reson.* **2012**, *218*, 93–104.

(5) Jeschke, G., Chechik, V.; Ionita, P.; Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., Jung, H. DeerAnalysis 2006 – A Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006**, *30*, 473–498.

(6) Keeley, J.; Choudhury, T.; Galazzo, L.; Bordignon, E.; Feintuch, A.; Goldfarb, D.; Eggeling, A.; Fábregas Ibáñez, L.; Jeschke, G.; Kuprov, I. Neural networks in double electron-electron resonance: a practical guide. *arXiv* **2021**, arXiv:2106.07465.

(7) Fábregas Ibáñez, L.; Jeschke, G.; Stoll, S. DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data. *Magn. Reson.* **2020**, *1*, 209–224.

(8) Jeschke, G. MMM: Integrative ensemble modeling and ensemble analysis, *Protein. Sci.* **2021**, *30*, 125–135.

(9) Hagelueken, G.; Abdullin, D.; Schiemann, O. mtsslSuite: Probing biomolecular conformation by spin labeling studies. Meth. Enzym. **2015**, *563*, 595–621.

(10) Jeschke, G. MMM: A toolbox for integrative structural modelling. *Prot. Sci.* **2018**, *27*, 76–85.

(11) Eaton, S. S.; Eaton, G. R. Relaxation Times of Organic Radicals and Transition Metal Ions. *Biol. Magn. Reson.* **2000**, *19*, 29–154.

(12) Vandermeulen, G. W. M.; Hinderberger, D.; Xu, H.; Sheiko, S. S.; Jeschke, G.; Klok, H.-A. Structure and Dynamics of Self-Assembled Poly(ethylene glycol) Based Coiled-Coil Nano-Objects. *ChemPhysChem* **2004**, *5*, 488–494.

(13) Baber, J. L.; Louis, J. M.; Clore, G. M. Dependence of Distance Distributions Derived from Double Electron–Electron Resonance Pulsed EPR Spectroscopy on Pulse-Sequence Time. *Angew. Chem. Int. Ed.* **2015**, *54*, 5336–5339.

(14) Banham, J. E.; Baker, C. M.; Ceola, S.; Day, I. J.; Grant, G. H.; Groenen, E. J. J.; Rodgers, C. T.; Jeschke, G.; Timmel, C. R. Distance measurements in the borderline region of applicability of CW EPR and DEER: A model study on a homologous series of spin-labelled peptides. *J. Magn. Reson.* **2008**, *191*, 202–218.

(15) Jeschke, G.; Polyhach, Y. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1895–1910.

(16) Hubbell, W. L.; López, C. J.; Altenbach, C.; Yang, Z. Technological advances in site-directed spin labeling of proteins. *Curr. Opin. Struct. Biol.* **2013**, *23*, 725–733.

(17) Emmanouilidis, L.; Esteban-Hofer, L.; Damberger, F. F.,; de Vries, T.; Nguyen, C. K. X.; Fábregas Ibáñez, L.; Mergenthal, S.; Klotzsch, E.; Yulikov, M.; Jeschke, G.; Allain, F. H.-T. NMR and EPR reveal a compaction of the RNA-binding protein FUS upon droplet formation. *Nature Chem. Biol.* **2021**, *17*, 608–614.

(18) Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H. Effect of freezing conditions on distances and their distributions derived from Double Electron Electron Resonance (DEER): A study of doubly-spin-labeled T4 lysozyme. *J. Magn. Reson.* **2012**, *216*, 69–77.

(19) Schmidt, T.; Jeon, J.; Okuno, Y.; Chiliveri, S. C.; Clore, G. M. Submillisecond Freezing Permits Cryoprotectant-Free EPR Double Electron−Electron Resonance Spectroscopy. *ChemPhysChem* **2020**, *21*, 1224–1229.

(20) Eaton, S. S.; Eaton, G. R. Relaxation Mechanisms. *eMagRes* **2016**, *5*, 1543–1556.

(21) Edwards, D. T.; Huber, T.; Hussain, S.; Stone, K. M.; Kinnebrew, M.; Kaminker, I.; Matalon, E.; Sherwin, M. S.; Goldfarb, D.; Han, S. Determining the Oligomeric Structure of

Proteorhodopsin by Gd^{3+}-Based Pulsed Dipolar Spectroscopy of Multiple Distances. *Structure* **2014**, *22*, 1677–1686.

(22) Raitsimring, A. M.; Salikhov, K. M.; Umanski, U.; Tsvetkov, Y. D. The instantaneous diffusion of the ESE paramagnetic centers in solids. *Fiz. Tverd. Tela* **1974**, *16*, 756–763.

(23) Pornsuwan, S.; Giller, K.; Riedel, D.; Becker, S.; Griesinger, C.; Bennati, M. Long-Range Distances in Amyloid Fibrils of α-Synuclein from PELDOR Spectroscopy. *Angew. Chem. Int. Ed.* **2013**, *52*, 10290–10294.

(24) Dastvan, R.; Bode, B. E.; Karuppiah, M. P. R.; Marko, A.; Lyubenova, S.; Schwalbe, H.; Prisner, T. F. Optimization of Transversal Relaxation of Nitroxides for Pulsed Electron−Electron Double Resonance Spectroscopy in Phospholipid Membranes. *J. Phys. Chem. B* **2010**, *114*, 13507–13516.

(25) Endeward, B.; Butterwick, J. A.; MacKinnon, R.; Prisner, T. F. Pulsed Electron−Electron Double-Resonance Determination of Spin-Label Distances and Orientations on the Tetrameric Potassium Ion Channel KcsA. *J. Am. Chem. Soc.* **2009**, *131*, 15246–15250.

(26) Bleicken, S.; et al. Structural Model of Active Bax at the Membrane. *Mol. Cell* **2014**, *56*, 496–505.

(27) Brown, I. in Time Domain Electron Spin Resonance (eds L. Kevan & R.N. Schwartz) Ch. 6, (Wiley, 1979).

(28) Canarie, E. R.; Jahn, S. M; Stoll, S. Quantitative Structure-Based Prediction of Electron Spin Decoherence in Organic Radicals. *J. Phys. Chem. Lett.* **2020**, *11*, 3396–3400.

(29) Huber, M.; Lindgren, M.; Hammerstroem, P.; Martensson, L. G.; Carlsson, U.; Eaton, G. R.; Eaton, S. S. Phase memory relaxation times of spin labels in human carbonic anhydrase II: pulsed EPR to determine spin label location. *Biophys. Chem.* **2001**, *94*, 245–256.

(30) Milov, A. D.; Tsvetkov, Y. D.; Salikhov, K. M. Phase relaxation of H-atoms stabilized in glassy matrices. *Fiz. Tverd. Tela* **1973**, *15*, 1187–1195.

(31) Zecevic, A.; Eaton, G. R.; Eaton, S. S.; Lindgren, M. Dephasing of electron spin echoes for nitroxyl radicals in glassy solvents by non-methyl and methyl protons. *Mol. Phys.* **1998**, *95*, 1255–1263.

(32) El Mkami, H.; Ward, R.; Bowman, A.; Owen-Hughes, T.; Norman, D. G. The spatial effect of protein deuteration on nitroxide spin-label relaxation: Implications for EPR distance measurement. *J. Magn. Reson.* **2014**, *248*, 36–41.

(33) Schmidt, T.; Wälti, M. A.; Baber, J. L.; Hustedt, E. J.; Clore, G. M. Long Distance Measurements up to 160 Å in the GroEL Tetradecamer Using Q-Band DEER EPR Spectroscopy. *Angew. Chem. Int. Ed.* **2016**, *55*, 15905–15909.

(34) Kattnig, D. R.; Reichenwallner, J.; Hinderberger, D. Modeling Excluded Volume Effects for the Faithful Description of the Background Signal in Double Electron–Electron Resonance. *J. Phys. Chem. B* **2013**, *117*, 16542–16557.

(35) Yardeni, E. H.; Mishra, S.; Stein, R. A.; Bibi, E.; Mchaourab, H. S. The Multidrug Transporter MdfA Deviates from the Canonical Model of Alternating Access of MFS Transporters. *J. Mol. Biol.* **2020**, *432*, 5665–5680.