Supporting information

Spin labelling for integrative structure modelling: A case study of the Polypyrimidine-Tract Binding Protein 1 domains in complexes with short RNAs

Christoph Gmeiner[a], Georg Dorn[b], Frédéric H. T. Allain[b], Gunnar Jeschke[a] and Maxim Yulikov[a,*]

[a] Laboratory of Physical Chemistry, ETH Zurich, Zurich, 8093, Switzerland
[b] Institute of Molecular Biology & Biophysics, ETH Zurich, Zurich, 8093, Switzerland
* E-mail: maxim.yulikov@phys.chem.ethz.ch

Contents

1 RBD structures and RNA sequences

RBD1 (aa 49-146):
GNDSKKFKGDSRSAGVPSRVIHIRKLPIDV**T**EGEVISLGLPFGKVTNLLMLKGKNQAFIEMNTEEAAN**T**MVNYYTSVTPVLRGQPIYIQFSNHKELKT
DSSPNQARAQAALQAVNSVQSGNLA

RBD2 (aa 172-298):
AGQSPVLRIIVENLFYPVTLDVLH**Q**IF**S**KFGTVLKIITFTKNNQFQALLQYADPVSAQHAKL**S**LDGQNIYNASSTLRIDFSKLTSLNVKYNNDKSRDYTR
PDLPSGDSQPSLDQTMAAAF GLSVPNVHGA LAPLAIPSA

RBD34 (aa 324-531):
GSHMGRIAIPGLAGAGNSVLLVSNLNPERVTPQSLFILFGVYGDVQRVKILFNKKENALVQMADGNQA**Q**LAMSHLNGHKLHGKPIRITLSKHQNV
QLPREGQEDQGLTKDYGNSPLHRFKKPGSKNFQNIFPPSATLHLSNIPPSVSEEDLKVLF**S**SNGGVVKGFKFFQKDRKMALIQMGSVEEAVQALIDL
HNHDLGENHHLRVSFSKSTI

Mutation positions in the RBDs are marked green.

RNA SL-E:
U9: 5'-UGCGUUUG**4-S-U**CUAUAUGU-3'
U11: 5'-UGCGUUUGUC**4-S-U**AUAUGU-3'
U13: 5'-UGCGUUUGUCUA**4-S-U**AUGU-3'

RNA SL-F:
U15: 5'-CAUAUUGCCGUCUU**4-S-U**UGGCAAUGUG-3'

Labelled positions in the RNA stem loops are marked magenta.

1.1 Comparison of simulated distance distribution using all existing RBD structures

Since different 3D structure of the individual RBDs of PTBP1 exist, we tested simulated distance distributions for all of them using the rotamer library approach for MTSSL, MAP and Gd(III)-DOTA in MMM. All distance distributions were analysed for mean distances <r>, peak maxima and widths σ(r) (Table S1, Fig. S1). In total, three solution-NMR structures (PDB code 2AD9, 1SJQ, 2N3O)[1, 2] exist for RBD1, two solution-NMR (PDB code 2ADB, 1SJR)[1, 2] and one X-ray crystallography structure[3] (PDB code 3ZZY) for RBD2 and two solution-NMR structures for RBD34 (PDB code 2ADC, 2EVZ)[1].
The simulated distance distributions of all RBD structures delivered similar results and comparable shapes of the distributions. Thus, we selected the three solution-NMR structures 2AD9, 2ADB and 2ADC[1] for our comparative study.

Table S1. Analysed simulated distance distributions of all existing RBD structures. Mean distances ⟨r⟩, peak maxima and standard deviations σ(r) were determined for the respective labelled pairs using the spin labels MTSSL, MAP and Gd(III)-DOTA.

Structure	Spin label	PDB code	⟨r⟩ / nm	r_{max} / nm	σ (r) / nm
RBD1	*MTSSL*	2AD9	3.24	3.45	0.35
T71C/T109C		1SJQ	2.99	3.25	0.39
		2N3O	3.27	3.22	0.42
	MAP	2AD9	3.05	3.05	0.51
		1SJQ	2.73	2.80	0.53
		2N3O	3.02	3.00	0.51
	Gd(III)-DOTA	2AD9	3.13	3.25	0.41
		1SJQ	2.70	2.90	0.52
		2N3O	3.31	3.35	0.47
RBD2	*MTSSL*	2ADB	2.19	2.40	0.34
S205C/S240C		1SJR	2.06	2.20	0.36
		3ZZY	2.01	2.10	0.35
	MAP	2ADB	2.03	2.20	0.39
		1SJR	1.82	1.95	0.36
		3ZZY	1.96	2.20	0.39
	Gd(III)-DOTA	2ADB	2.22	2.40	0.42
		1SJR	2.24	2.65	0.41
		3ZZY	2.17	2.40	0.41
RBD34	*MTSSL*	2ADC	4.16	4.10	0.30
Q388C/S475C		2EVZ	4.38	4.45	0.32
	MAP	2ADC	4.57	4.70	0.38
		2EVZ	4.42	4.50	0.48
	Gd(III)-DOTA	2ADC	4.51	4.60	0.40
		2EVZ	4.54	4.60	0.36

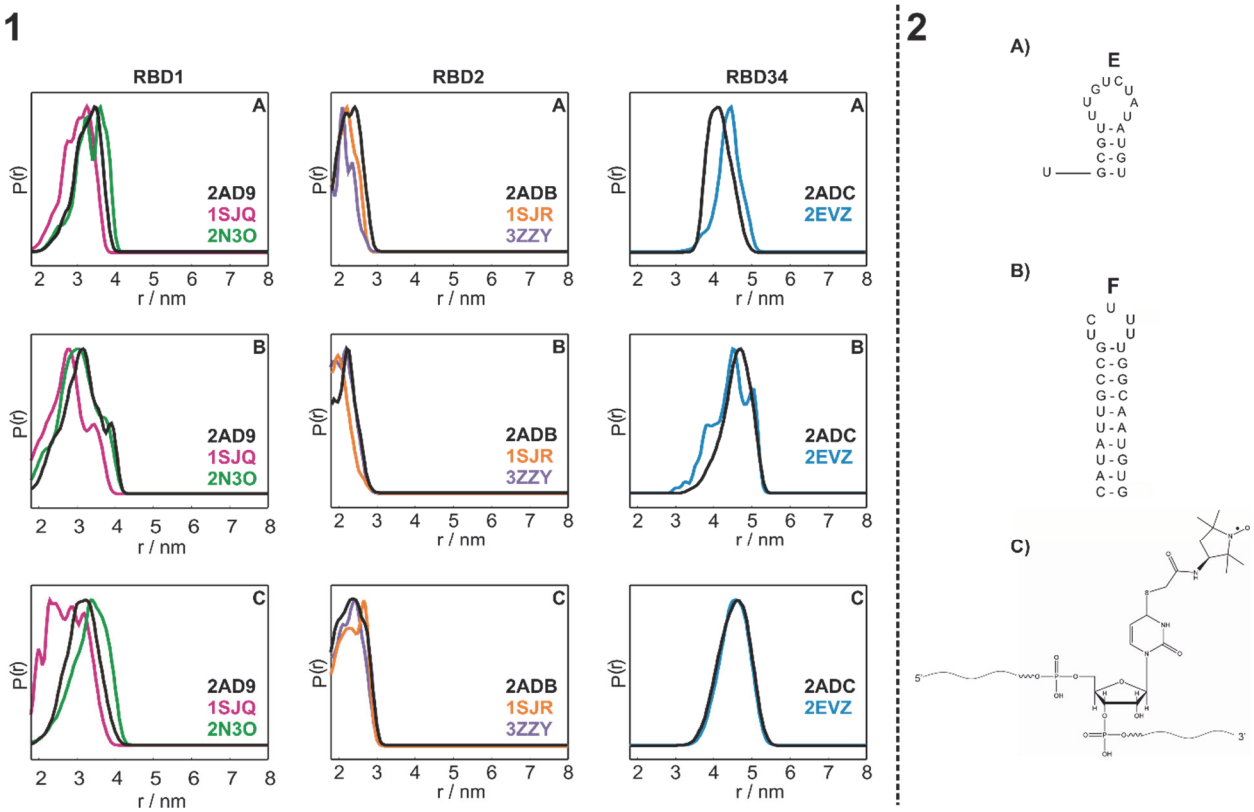

Figure S1. Simulated distance distributions of all existing RBD structures. 1) Superimposed simulated distance distributions for each individual RBD using MTSSL (A), MAP (B) and Gd(III)-DOTA (C). 2) Secondary structures of EMCV-IRES SL-E (A) and SL-F (B).[4, 5] C) SDSL of 4'-thiouridine with IAP.

2 Mutagenesis of RBDs, protein expression and purification

Site-directed mutagenesis of the cysteine-less RBDs was performed using a conventional PCR program (Table S2), with 24 cycles of melting, annealing, and amplification of the respective plasmids, pTYB11 (New England Biolabs) for RBD1 and RBD2 and pET28a (Novagen) for RBD34, respectively. Plasmids were then incubated with 0.5 μL *DpnI* for 4 h at 37°C and were transformed into chemically competent *Escherichia coli* TOP10 cells (Thermo Fisher Scientific) that are not resistant to any antibiotics. Cells which took up a plasmid that encodes either a carbenicillin (RBD1 and RBD2) or kanamycin (RBD34) resistance, were selected on agar plates containing the respective antibiotic. After plasmid purification, plasmids were sent for sequencing (customer service, Microsynth) and the correct plasmids were transformed into codon optimized *E. coli* cells (BL21 Codon Plus (DE3) RIL, Agilent) containing a chloramphenicol resistance. Transformation was done on ice by adding 1-1.5 μL of the pure plasmid to 100 μL cells. After 10 minutes incubation at 4°C, the plasmids were incorporated by weakening the cell wall through heat shock at 42°C for 1 minute. Recovery step was performed for 2 minutes on ice and afterwards cells were incubated with 300 μL LB-medium for 30 minutes at 37°C. Transformed cells were plated on agar containing the respective antibiotic for the plasmid as well as chloramphenicol for the BL21 Codon Plus (DE3) RIL cells.

Table S2. PCR program for site-directed mutagenesis with repeating the steps 2 to 4 for 24 cycles.

Program	Temp. / °C	t / min
1	95	5
2	95	1
3	60	1
4	68	17
5	68	10
6	4	-

Protein expression was performed in 2 L LB Broth medium (25 $g \cdot L^{-1}$; Chemie Brunschwig AG), containing carbenicillin and chloramphenicol (RBD1 and RBD2) or kanamycin and chloramphenicol (RBD34). After reaching an OD600 value of 0.6 to 0.8 of the cell suspension, the induction phase was started with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 20°C. For harvest, cells were centrifuged (4°C, 15', 7.000 g) and cell pellets were resuspended in a lysis buffer A (100 mM sodium phosphate, 1 M NaCl, 1 mM EDTA, pH: 8.0) for RBD1 and RBD2 encoding an N-terminal Chitin binding domain as affinity tag (target sequence of RBD1 is separated by an Intein-sequence) , or in lysis buffer B (50 mM sodium phosphate, 1 M NaCl, 7 mM imidazole, pH: 8.0) for RBD34 which encodes for a N-terminal hexa-His affinity-tag. Cell lysis was ensured by cell disruption using a Microfluidizer LM-10 (Instrumat AG) and cell debris was separated from supernatant by centrifugation (30.000 g, 4°C and 45 min).
Supernatants were loaded on Chitin binding beads (New England, BioLabs) in case of RBD1 and RBD2 and on Ni-NTA agarose beads (Qiagen AG) for RBD34, respectively. Bound double mutated RBD1 and RBD2 samples were washed with 300 mL of a high salt buffer (100 mM sodium phosphate, 1.5 M NaCl, pH: 8.0). Afterwards, intein-cleavage was performed by incubation with a cleavage buffer (100 mM sodium phosphate, 200 mM NaCl, 50 mM DTT, pH: 8.0) for 48 hours at 4°C. Elution was performed with an elution buffer (100 mM sodium phosphate, 200 mM NaCl, pH: 8.0) and concentrated proteins were purified by size exclusion chromatography (Superdex 75 column, GE Healthcare Life Science) using the spin labelling buffer A (50 mM MOPS, 25 mM NaCl, pH: 6.5), or B (50 mM MOPS, 500 mM NaCl, pH: 6.5), respectively. In case of RBD34, the bound double mutant was washed with 5 bed volumes of lysis buffer and elution was performed stepwise with an increasing imidazole concentration from 20 to 200 mM in steps of 20-40-60-80-100 and 200 mM imidazole in the buffer. All fractions were analysed by SDS PAGE and pure ones, containing the double mutated RBD34, were pooled and dialysed overnight at 4°C into a low salt buffer (10 mM sodium phosphate, 20 mM NaCl, pH: 6.5. The hexa-His affinity-tag was removed by Thrombin cleavage (Sigma Aldrich) in an overnight reaction at 4°C . Since the pI of RBD34 is around 9.4, cation-exchange chromatography (CEX) using a HiTrap™ SF HP (GE Healthcare Life Science) was performed to separate cleaved protein from Thrombin. Here we used a loading buffer without salt (20 mM sodium phosphate, pH: 6.5) and for elution a high-salt buffer (20 mM sodium phosphate, 1 M NaCl, pH: 6.5). The final purification step contained a size exclusion chromatography (SDSL labelling buffer A or B) using a Superdex 75 column (GE Healthcare Life Science). Protein concentrations were determined by measuring UV absorption with a NanoDrop Spectrophotometer ND-1000 (Witec AG) and using the respective theoretical extinction coefficients of $\varepsilon = 4.47\ L \cdot mmol^{-1} \cdot cm^{-1}$ (RBD1), $7.45\ L \cdot mmol^{-1} \cdot cm^{-1}$ (RBD2) and $2.93\ L \cdot mmol^{-1} \cdot cm^{-1}$ (RBD34).[6]

3 Sample analysis

3.1 Electrophoresis

Protein purity was analysed using a Mini-PROTEAN Tetra Vertical Electrophoresis System (BioRad Laboratories). A prestained protein ladder Dual Xtra with a size range of 2-250 kDa was used as reference (BioRad Laboratories). We mixed 20 μL of sample with 5 μL reducing (including β-mercaptoethanol) or non-reducing (without β-mercaptoethanol) loading dye and incubated all samples at 98°C for 5 minutes. RBDs were analysed on a 15% SDS gel. Comparing reducing and non-reducing loading dye treated samples, we could directly analyse all RBDs for cysteine-mediated oxidative dimerisation since formed disulphide bridges would be broken under reducing conditions. Parameters for the electrophoresis were set to 150 V for 60 minutes.
Complex formations of RBDs and the RNA SLs were analysed by native PAGE. We compared protein/RNA complexes with free, unbound RNA SLs. Since separation in native PAGE is based on charge, a clear shift of the bands indicates a successful complex formation because free RNA is stronger negatively charged than RNA in complex with RBD. Electrophoresis was performed at 80 V for 75 minutes at room temperature. We could detect incomplete complex formation when mixing RNA and RBD in a ratio 1:1 because a second band, which was comparable to free unbound RNA, appeared. The incomplete complex formation influenced also orthogonal DEER measurements resulting in a lower amount of coupled electron spins in the protein/RNA complexes (Fig. 8-10).

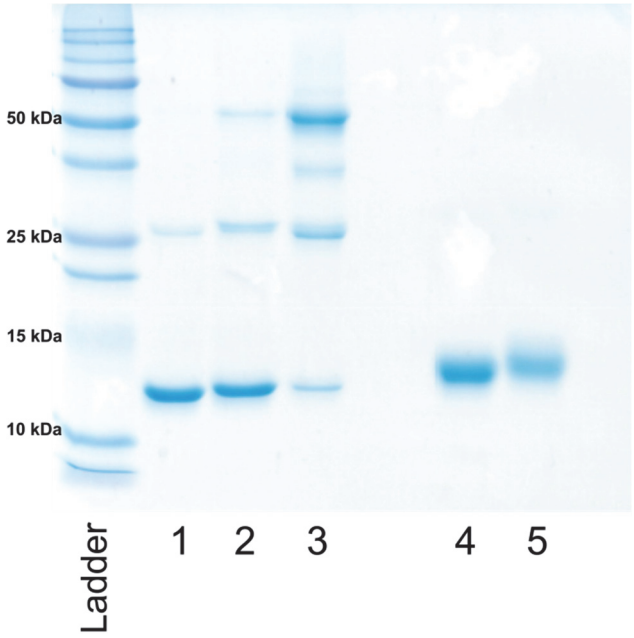

Figure S2 15% SDS PAGE showing all labelled RBD1 samples compared to a precision plus protein dual xtra standard in a range of 2 – 250 kDa (BioRad). Lane 1: T71C/T109C R1. Lane 2: T71C/T109C IAP. Lane 3: T71C/T109C MAP. Lane 4: T71C/T109C Gd(III)-DOTA. Lane 5: T71C/T109C Gd(III)-DTPA. Gd(III)-labelled RBD1 samples show a high purity without any dimerisation bands in the higher kDa region (> 50 kDa). In case of T71C/T109C R1 and IAP, a small fraction of dimerised protein could be detected, which did not significantly contribute to DEER measurements. T71C/T109C MAP shows dimerisation and even multimerisation bands in higher regions such as 25 and 50 kDa. This could be also seen in nitroxide-nitroxide DEER because a second peak at longer distances was detected.

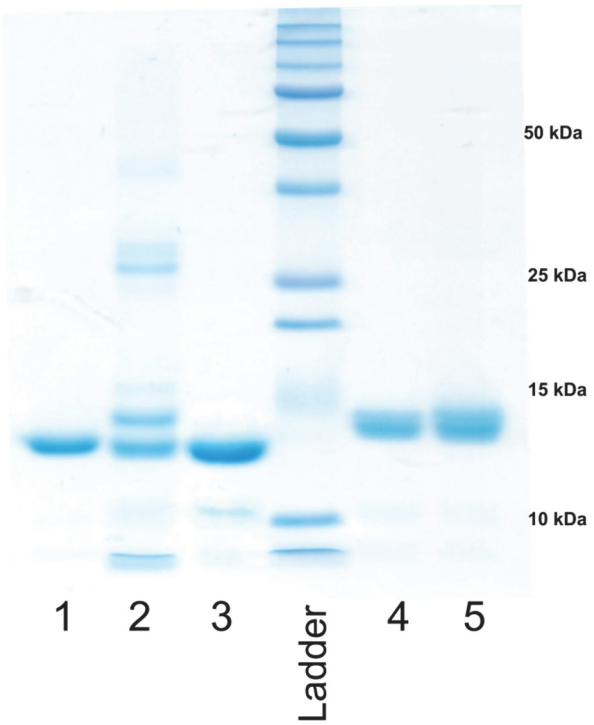

Figure S3 15% SDS PAGE showing all labelled RBD2 samples compared to a precision plus protein dual xtra standard in a range of 2 – 250 kDa (BioRad). Lane 1: S205C/S240C R1. Lane 2: S205C/S240C IAP. Lane 3: S205C/S240C MAP. Lane 4: S205C/S240C Gd(III)-DOTA. Lane 5: S205C/S240C Gd(III)-DTPA. All samples except of S205C/S240C IAP reveal a high purity. The latter one shows dimerised and degraded fractions leading to a incorrect distance distribution measured by nitroxide-nitroxide DEER (Fig. S.17).

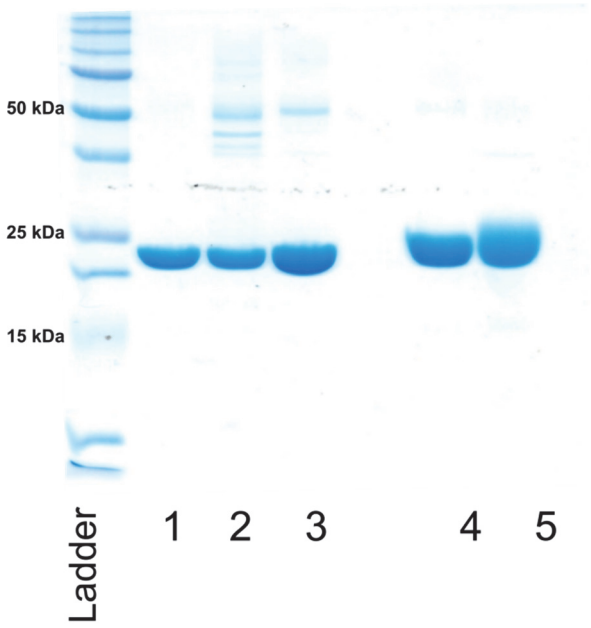

Figure S4 15% SDS PAGE showing all labelled RBD34 samples compared to a precision plus protein dual xtra standard in a range of 2 – 250 kDa (BioRad). Lane 1: Q388C/S475C R1. Lane 2: Q388C/S475C IAP. Lane 3: Q388C/S475C MAP. Lane 4: Q388C/S475C Gd(III)-DOTA. Lane 5: Q388C/S475C Gd(III)-DTPA. Dimerisation bands could be detected only for Q388C/S475C IAP and MAP which did not significantly contribute to the distance distribution, since all measured samples agreed very well with the predicted distances.

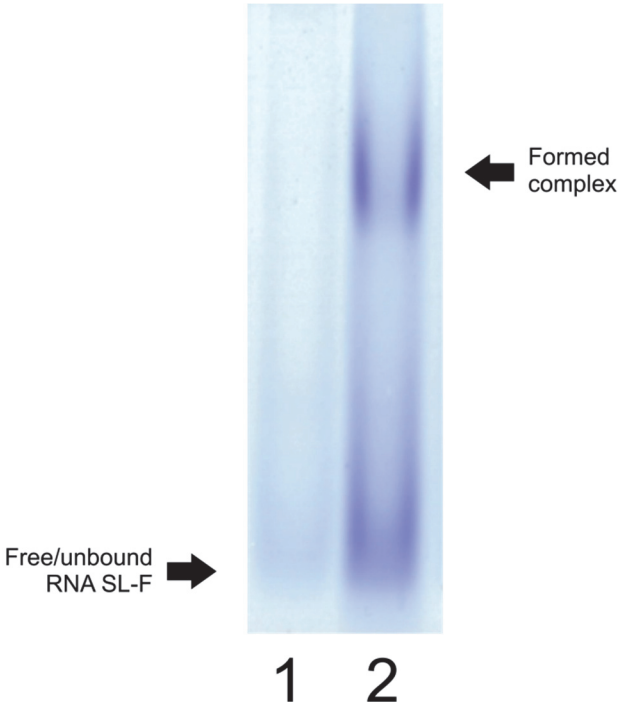

Figure S5 Native PAGE of single RNA SL-F U15 labelled with IAP (1) and in complex with RBD2 S205C/S240C labelled with Gd(III)-DOTA (2). A second peak at the height of the free/unbound SL-F in lane 2 indicates incomplete 1:1 complex formation between protein and RNA. Successful complex formation can be seen by a shift of the RNA towards the upper range of the gel because of the reduced negative charge of the RNA (2).

3.2 CD spectroscopy – all spectra

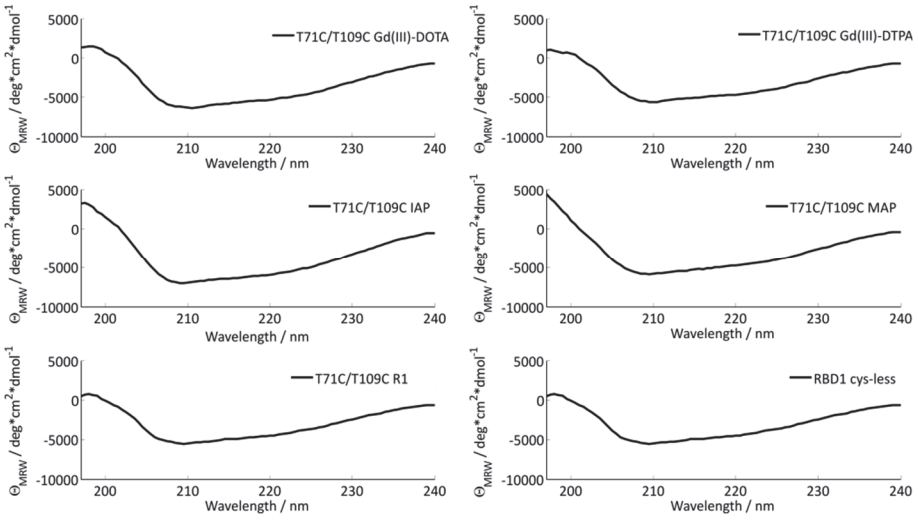

Figure S6 CD spectra of all spin-labelled RBD1 samples. The overall shapes of all spectra look similar, indicating that SDSL to the α-helices did not influence the secondary structure. Differences in signal arose from uncertainties in protein concentration measurements. Peak minima at 208 and 222 nm and a maximum at 195 nm describe main characteristics for helical proteins.

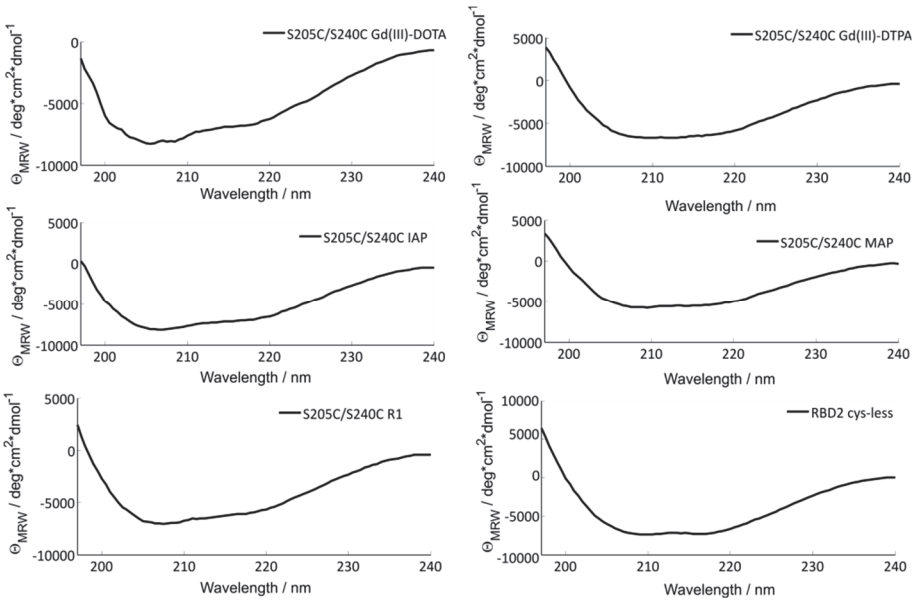

Figure S7 CD spectra of all spin-labelled RBD2 samples. The overall shapes of all spectra look similar, indicating that SDSL to the α-helices did not influence the secondary structure. Differences in signal arose from uncertainties in protein concentration measurements. Peak minima at 208 and 222 nm and a maximum at 195 nm describe main characteristics for helical proteins.

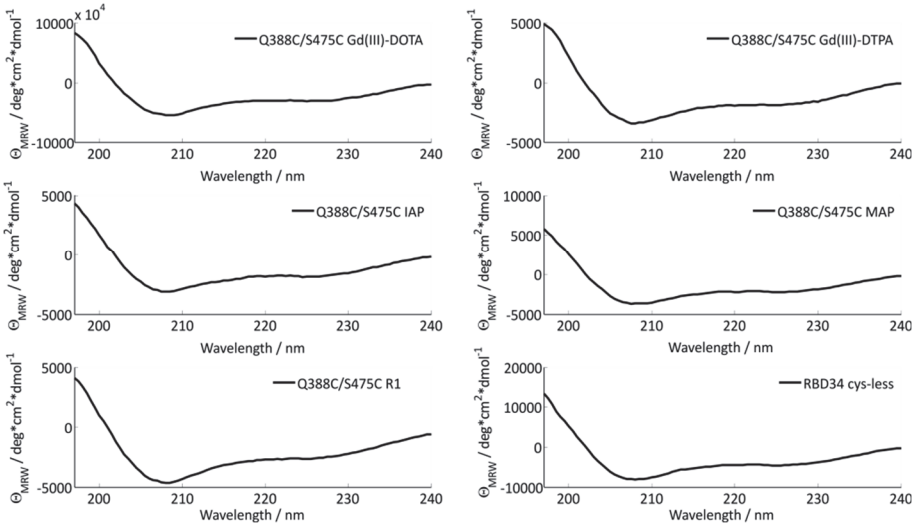

Figure S8 CD spectra of all spin-labelled RBD34 samples. The overall shapes of all spectra look similar, indicating that SDSL to the α-helices did not influence the secondary structure. Differences in signal arose from uncertainties in protein concentration measurements. Peak minima at 208 and 222 nm and a maximum at 195 nm describe main characteristics for helical proteins.

3.3 Mass spectrometry

All mutated and labelled RBD samples were analysed by electrospray ionization – mass spectrometry (ESI-MS) in a range of 10 to 20 kDa for RBD1 and RBD2, and 20 to 30 kDa for RBD3/4, respectively. Sampling cone energy was set at 40V and m/z data were deconvoluted into MS data using the MaxEnt1 Software with a resolution of the output mass of 0.5 Da/channel and Uniform Gaussian Damage Model at the half height of 0.7 Da. MS service was performed at the Functional Genomics Centre Zurich (FGCZ) by Dr. Serge Chesnov.

Table S3. Detected and calculated mass of individual spin labelled RBDs.

Labelled domains	Detected mass / kDa	Theoretical mass / kDa	Δ / Da
RBD1			
T71C/T109C			
R1	13.738	13.732	6
IAP	13.765	13.758	7
MAP	13.838	13.838	0
Gd(III)-DOTA	14.725	14.730	5
Gd(III)-DTPA	14.703	14.707	4
RBD2			
S205C/S240C			
R1	15.564	15.563	1
IAP	15.590	15.590	0
MAP	15.670	15.670	0
Gd(III)-DOTA	16.557	16.562	5
Gd(III)-DTPA	16.534	16.540	6
RBD34			
Q388C/S475C			
R1	23.705	23.702	3
IAP	23.731	23.728	3
MAP	23.808	23.808	0
Gd(III)-DOTA	24.697	24.700	3
Gd(III)-DTPA	24.672	24.678	6

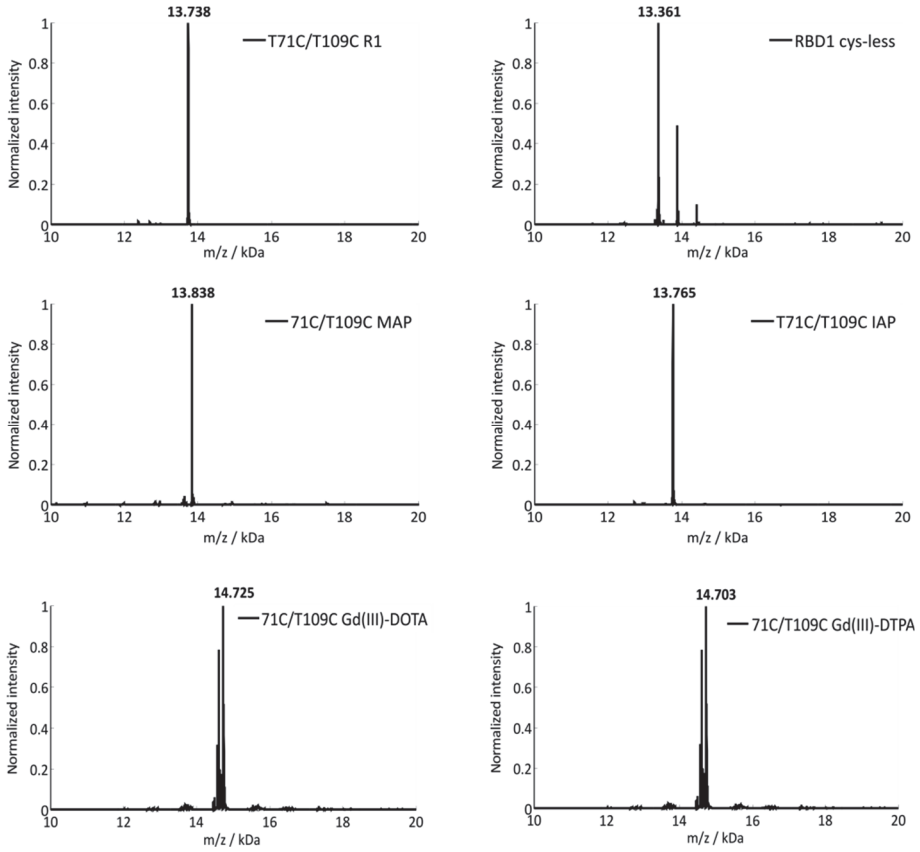

Figure S9 Normalised mass spectra of cysteine-less RBD1 and all spin-labelled T71C/T109C samples in the range of 10 to 20 kDa. Comparing with Table S3 one can see that all expected mass could be detected. All main peaks represent doubly labelled RBD1 which is consistent with DEER measurements, where the expected modulation depth was observed.

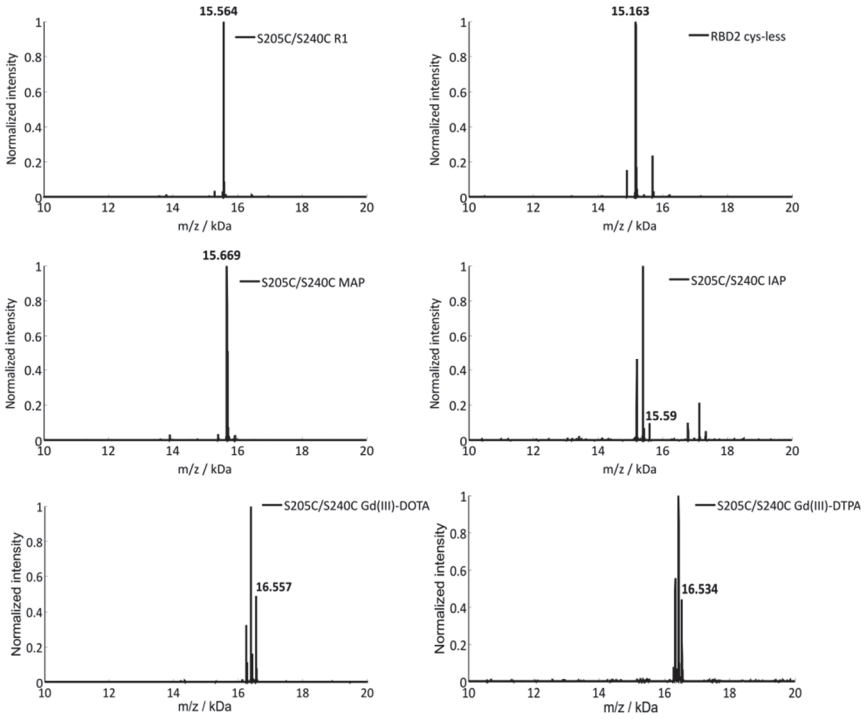

Figure S10 Normalised mass spectra of cysteine-less RBD2 and all spin-labelled S205C/S240C samples in the range of 10 to 20 kDa. Comparing with Table S3 one can see that all expected masses could be detected. S205C/S240C IAP represents only a small fraction of doubly-labelled RBD2, whereas the main peak represents a single labelled version of this double mutant. This result is consistent with the nitroxide-nitroxide DEER measurements, where a significantly lower modulation depth was found. Also Gd(III)-labelled RBD2 samples show a higher amount of single-labelled than doubly-labelled RBD2.

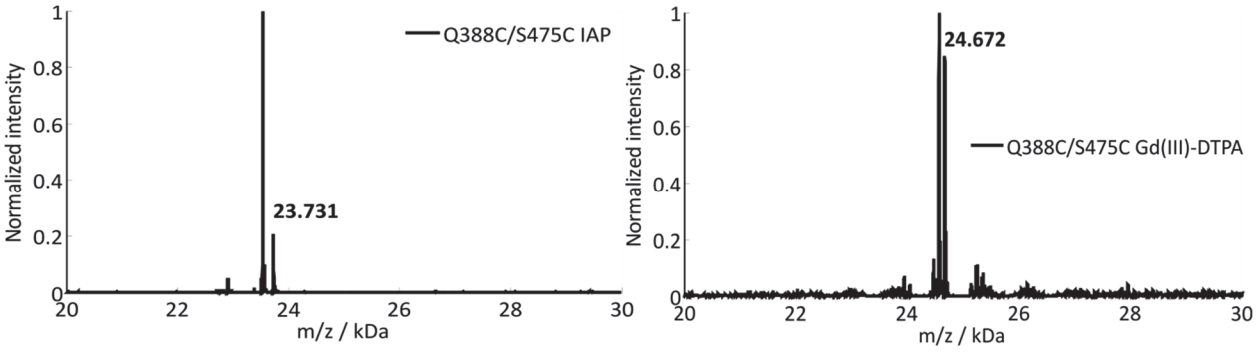

Figure S11 Normalised mass spectra of the double mutant RBD34 Q388C/S475C labelled with IAP and Gd(III)-DTPA in the range of 20 to 30 kDa. Comparing with Table S3 one can see that the expected masses could be detected. However, only a small fraction of doubly labelled Q388C/S475C IAP could be analysed whereas the main peak represents a singly-labelled version of this double mutant. This result is consistent with the nitroxide-nitroxide DEER measurements, where a significantly lower modulation depth was found.

4 EPR spectroscopy

4.1 Continuous-Wave EPR spectroscopy

Nitroxide-labelled RBDs

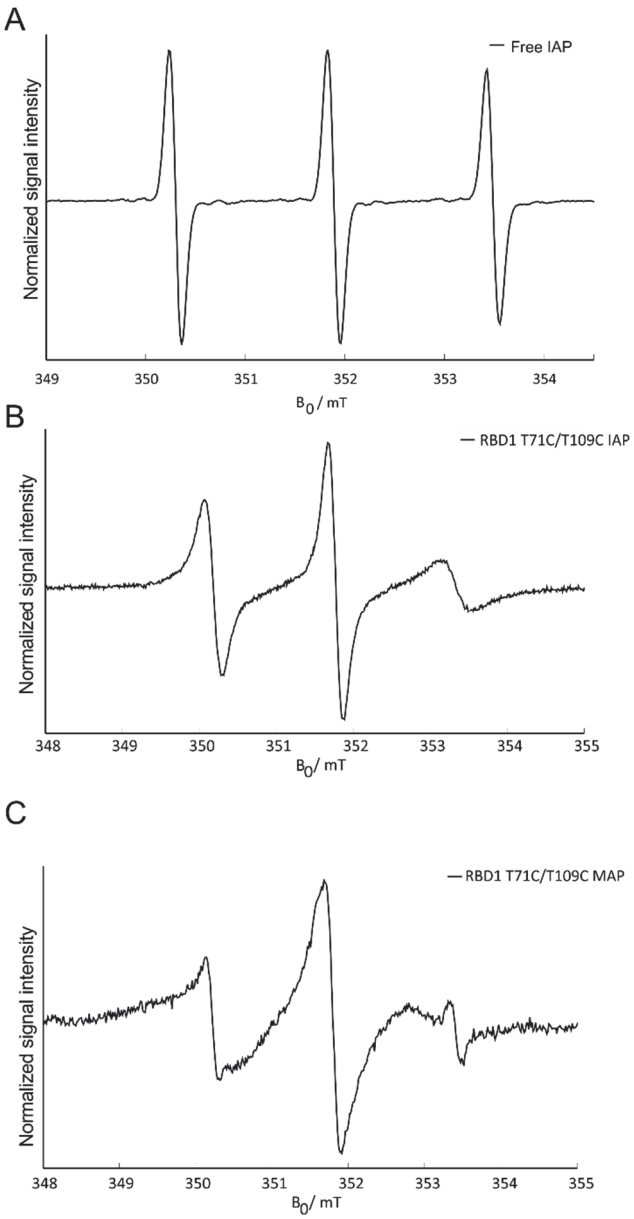

Figure S12 CW EPR spectra of free IAP (A) and RBD1 T71C/T109C labelled with IAP (B) and MAP (C). IAP and MAP spectra show spin label mobility which is consistent with a nitroxide side group rather than with free spin label. However, in the MAP spectra also a sharp peak is clearly visible in the high- and low-field component which probably stems from a small fraction of free label.

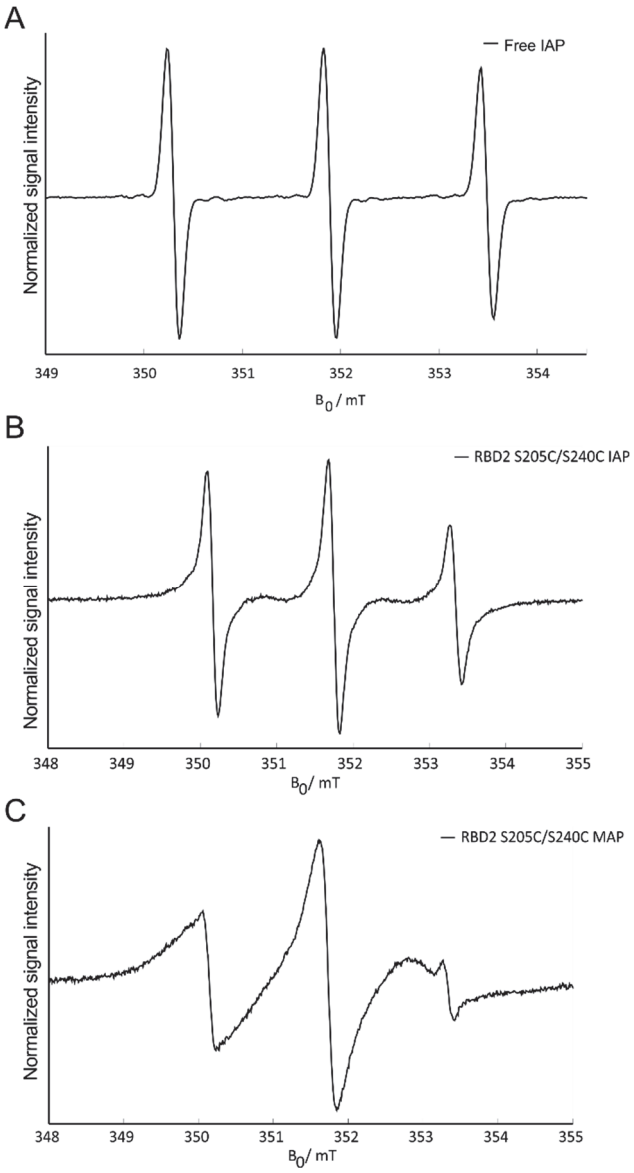

Figure S13 CW EPR spectra of free IAP (A) and RBD1 S205C/S240C labelled with IAP (B) and MAP (C). IAP spectra shows a sharp peak is clearly visible in the high- and low-field component which probably stems from a small fraction of free label.

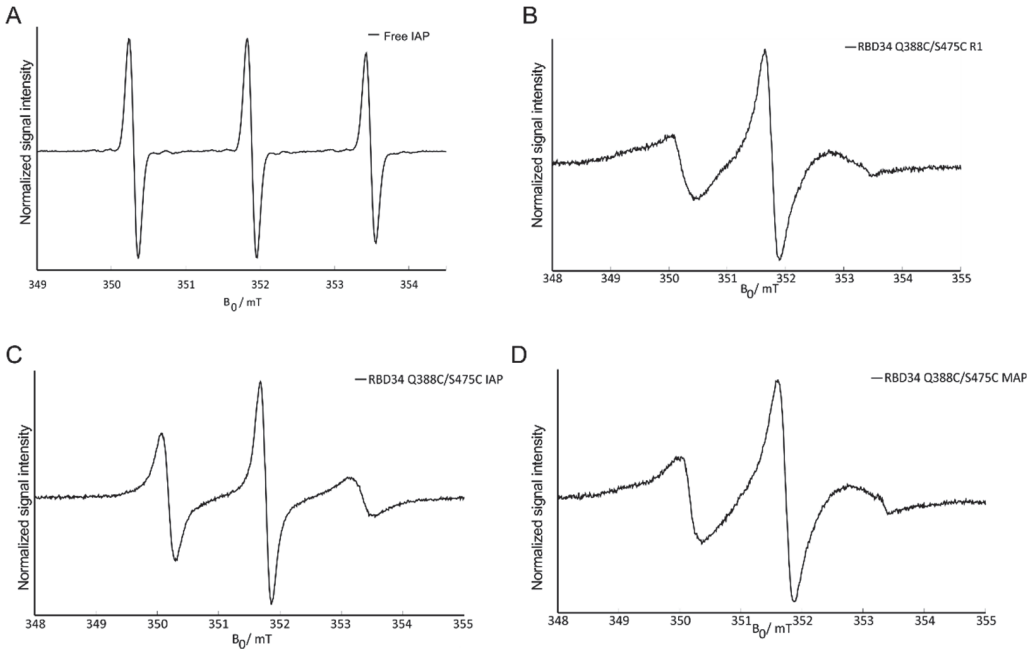

Figure S14 CW EPR spectra of free IAP (A) and RBD34 Q388/S475C nitroxide labelled samples; R1 (B), IAP (C) and MAP (D). Spectra in B-D show spin label mobility which is consistent with a nitroxide side group rather than with free spin label. While R1 and MAP spectra exhibit quite a similar broadening of the line shape and thus similar reduced spin flexibility, the IAP spectra exhibits slightly sharper peaks indicating that the spin label has higher mobility in this sample.

Nitroxide-labelled RNAs

Attachment of nitroxide spin labels to the cysteine residue in the RBD samples leads to a clear decrease of their flexibility, as detected by significant broadening of the line shapes and consequently lower amplitudes of the low- and high-field components (Fig. S12-14). In case of RNA molecules, which are here significantly smaller than the RBDs, the attached IAP shows high spin mobility. Thus, we performed CW EPR spectroscopy also in buffer containing 30% sucrose to reduce the contribution of small molecule rotation, as it was shown earlier.[7] Since no significant change of the free IAP spectra is detectable between pure buffer and buffer containing 30% sucrose (Fig. S16), such comparison can be used to distinguish between small labelled molecules showing high spin flexibility and free spin label. Here we found that all label is attached to the RNA.

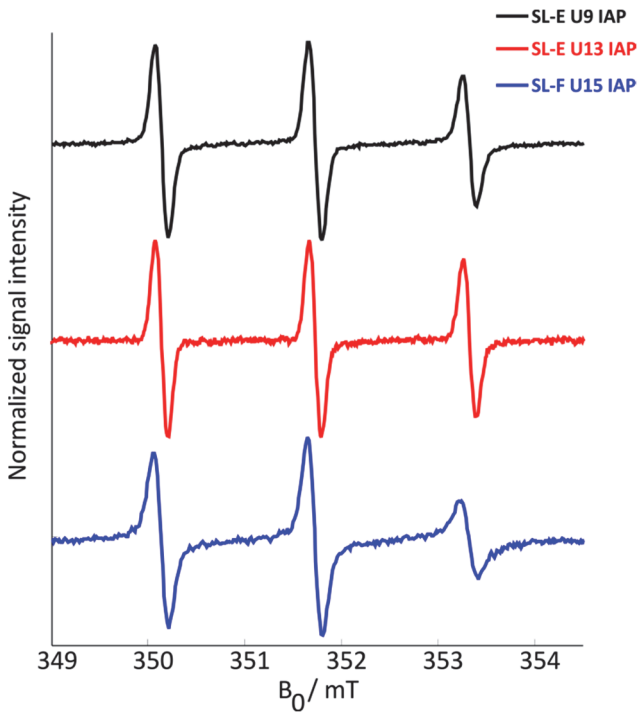

Figure S15 Normalised CW-EPR spectra of IAP labelled RNA SL-E U9 (black) and U13 (red) and RNA SL-F U15 (blue). Since SL-F is larger as SL-E, it shows a stronger reduction of the spin label mobility because the contribution of small molecule rotation is weaker. In case of RNA SL-E, the spin label has a high flexibility which can be recognized by sharp peaks of the high-field component.

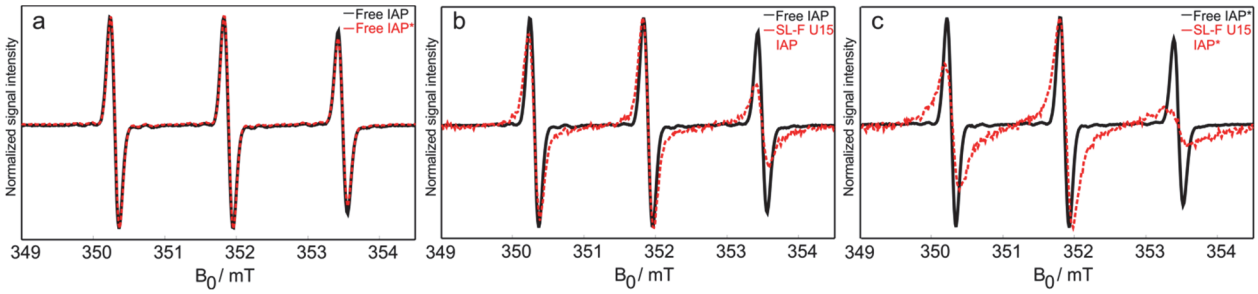

Figure S16 a) CW EPR spectra of free IAP (100 μM) in low salt buffer (20 mM sodium phosphate, 10 mM NaCl, pH: 6.5, in H2O) and in buffer containing 30% sucrose (*). b) IAP labelled RNA (SL-F U15) in comparison with free IAP in low salt buffer (20 mM sodium phosphate, 10 mM NaCl, pH: 6.5, in H2O). c) IAP labelled RNA (SL-F U15) in comparison with free IAP in 30% sucrose (*).

4.2 Protein-protein distance distributions obtained by 4-pulse DEER

Nitroxide-nitroxide DEER of RBD2 S205C/S240C

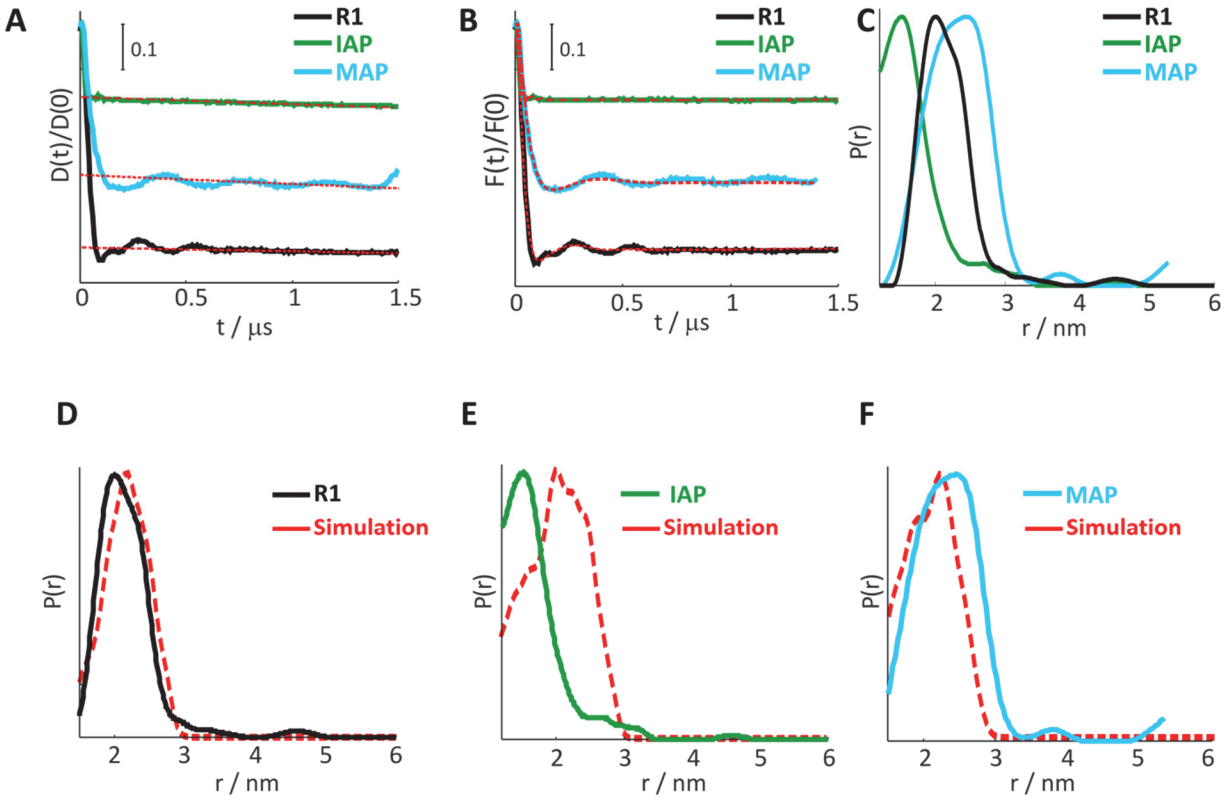

Figure S17 Nitroxide-nitroxide DEER results of RBD2 S205C/S240C. Experimental data including background correction (red dashed, A), background-corrected data including form factor fit (red dashed, B) and overlaid distance distributions (C) are shown for R1, IAP and MAP side chains. Distance distributions were compared to the solution-NMR structure (PDB 2ADB) and are in good agreement with the simulation (D-F, red dashed). Only for S205C/S240C IAP we find a deviation of approximately 0.5 nm, indicating that the sample quality was insufficient. The modulation depth was almost 0.5 for S205C/S240C R1 whereas IAP and MAP showed a reduced depth. The distance distribution obtained by MAP was slightly broader than in the case of S205C/S240C R1 or IAP because of the larger maleimido-tag.

Nitroxide-nitroxide DEER of RBD34 Q388C/S475C

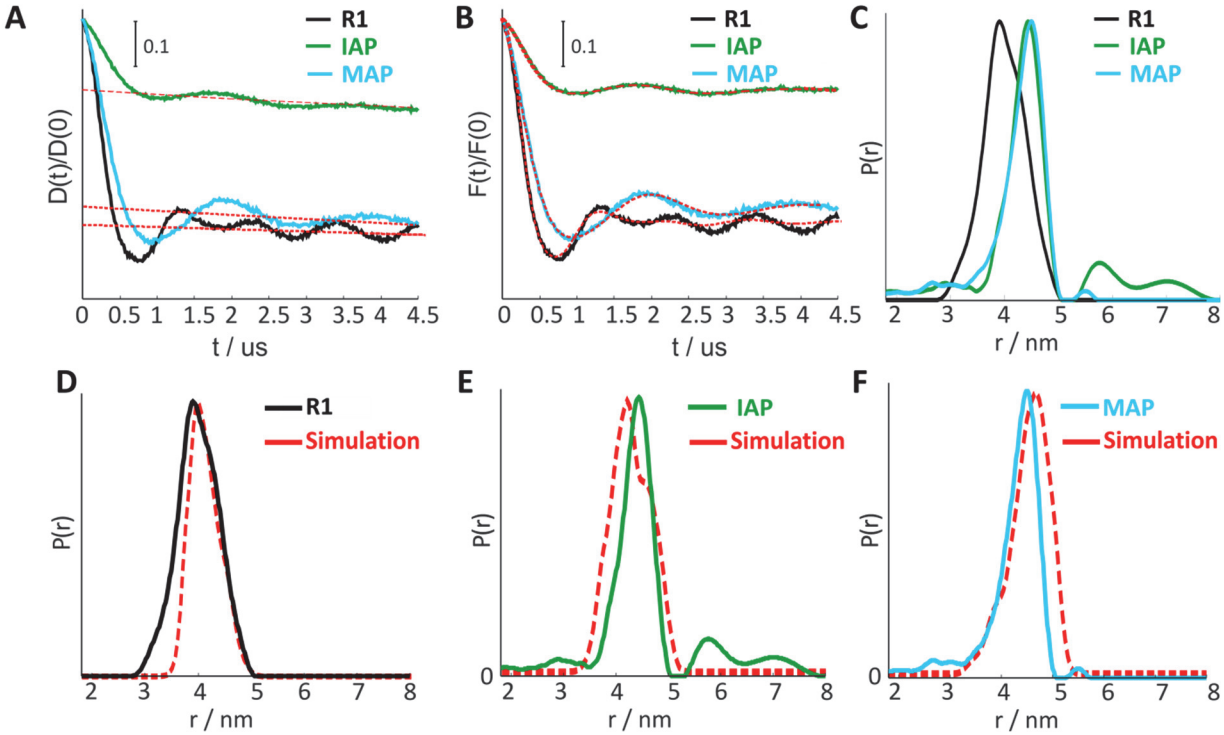

Figure S18 Nitroxide-nitroxide DEER results of RBD34 Q388C/S475C. Experimental data including background correction (red dashed, A), background-corrected data including form factor fit (red dashed, B) and overlaid distance distributions (C) are shown for R1, IAP and MAP. Distance distributions were compared to the solution-NMR structure (PDB 2ADC) and are in good agreement with the simulation (D-F, red dashed). Q388C/S475C R1 and MAP results revealed the highest modulation depth whereas IAP showed a significant reduced depth.

Gd(III)-Gd(III) DEER of RBD1 T71C/T109C

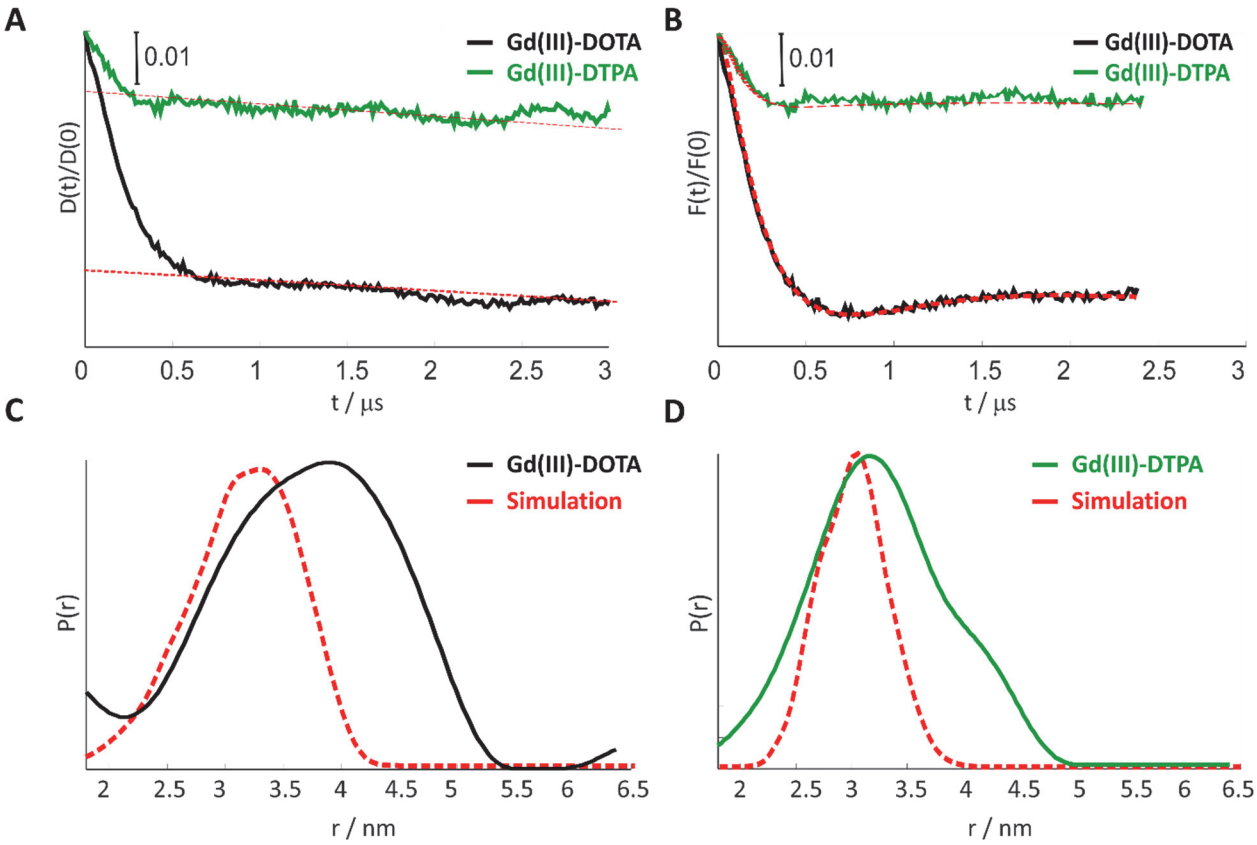

Figure S19 Gd(III)-Gd(III) DEER results of RBD1 T71C/T109C. A) Experimental data with background fitting (red, dashed), B) background-corrected data including form factor fitting (red, dashed) and distance distributions of Gd(III)-DOTA (C) and Gd(III)-DTPA (D). Distance distributions were compared to the solution-NMR structure (PDB 2AD9) and are in agreement with the simulations (red, dashed). Gd(III)-DOTA had a significant higher modulation depth than Gd(III)-DTPA but revealed also a slightly broader distance distribution with a peak maximum shifted to longer distances by 0.7 nm. Both measurements showed a broader width as the predictions.

Gd(III)-Gd(III) DEER of RBD2 S205C/S240C

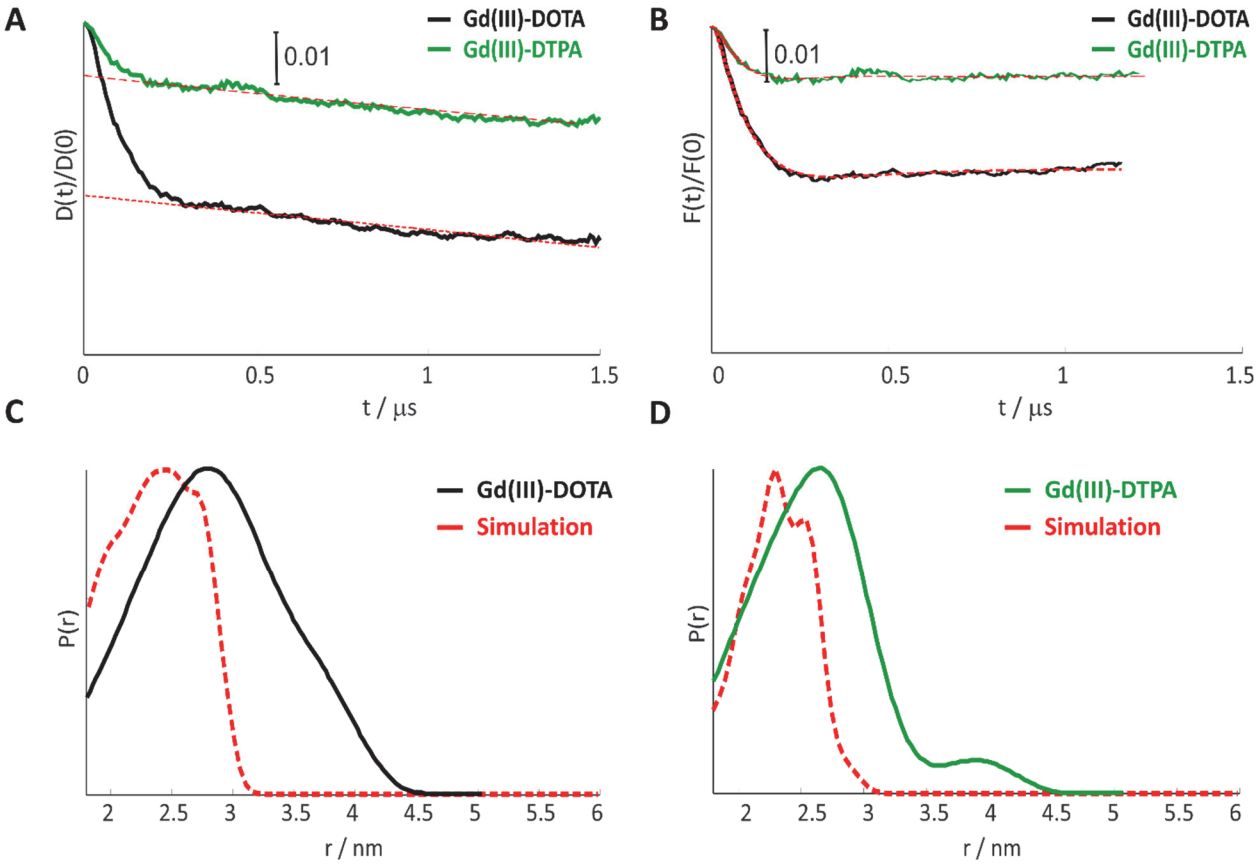

Figure S20 Gd(III)-Gd(III) DEER results of RBD2 S205C/S240C. A) Experimental data with background fitting (red, dashed), B) background-corrected data including form factor fitting (red, dashed) and distance distributions of Gd(III)-DOTA (C) and Gd(III)-DTPA (D). Distance distributions were compared to the solution-NMR structure (PDB 2ADB) and showed an agreement with the simulation (red, dashed). Gd(III)-DOTA had a significant higher modulation depth than Gd(III)-DTPA. Both revealed a slightly broader distance distribution than the simulations with a peak maximum shifted to longer distances.

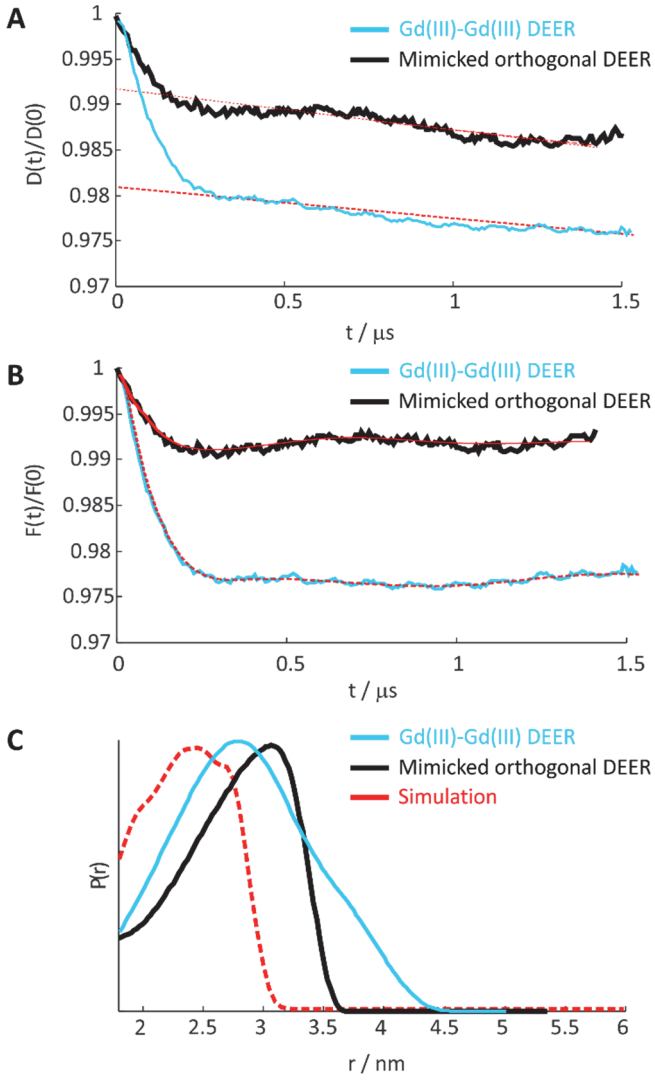

Figure S21 Mimicked orthogonal DEER experiment with Gd(III)-DOTA doubly labelled RBD2 S205C/S240C. A) Experimental data with background correction (red, dashed), B) background-corrected data including form factor fit (red, dashed) and C) distance distributions compared to the simulation (red, dashed) based on the solution-NMR structure of RBD2 (PDB 2ADB). The pump pulse was set 310 MHz away from the Gd(III) centre, which corresponds to the maximum of the nitroxide spectrum in an orthogonally labelled system. Therefore, a much smaller modulation depth was obtained compared to the conventional Gd(IIII)-Gd(III) 4-pulse DEER of the same sample (blue). Both experimentally obtained distance distributions (cyan, black) showed a slight shift to longer distances with respect to the simulation (red, dashed).

4.3 Orthogonal DEER experiments on RBD/RNA complexes

RBD1/SL-E complex (T109C Gd(III)-DOTA) + SL-E (U13 IAP)

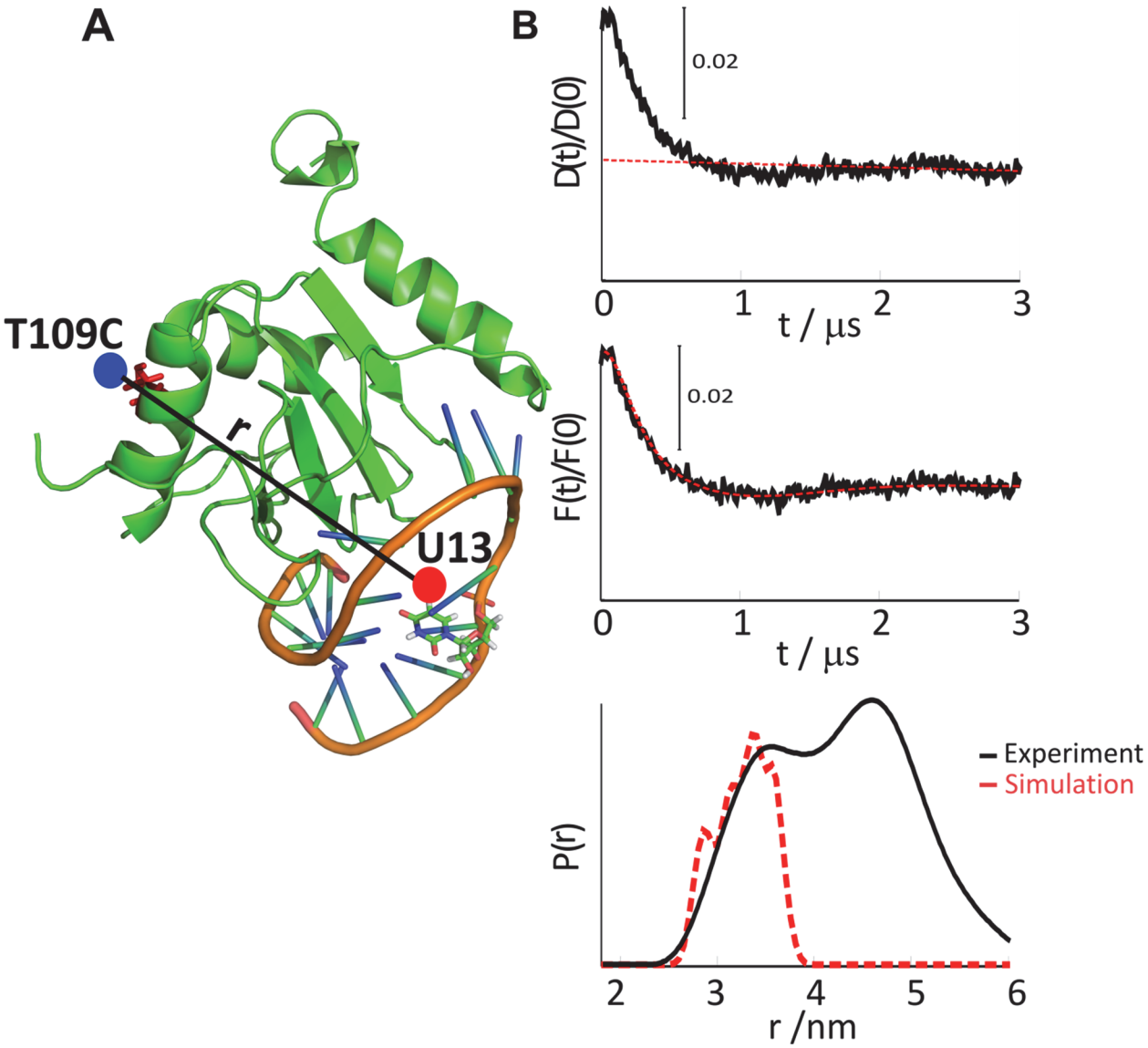

Figure S22 Orthogonal DEER measurement of RBD1 T109C Gd(III)-DOTA and RNA SL-E U13 IAP. The nitroxide spin label on the RNA was placed in the stem (red, A) based on the CLIR-MS/MS model. Gd(III)-DOTA was positioned on the α-helix (blue, A) of RBD1 pointing away from the interaction site. Experimental data including background fitting, background-corrected data with form factor fitting and the distance distribution are shown in B. The experimental distance distribution is bimodal, with the shorter distance agreeing very well with the model (B, red dashed). However, a second distance peak appears in a range of 4 to 5 nm. This bimodality could indicate an influence of the spin label on the conformation of the stem. The nucleotide U13 was assumed to participate in a non-canonical U-U base pair based on secondary structure predictions using MC-fold.[8] However, secondary structure predictions using the ViennaRNA package[9] predicts a more open loop in which U13 is not base paired. Thus, this could be an indication for an additional potential conformation of RBD1 bound to SL-E.

RBD2/SL-F complex (S205C/S240C Gd(III)-DOTA) + SL-F (U15 IAP)

A) Orthogonal DEER of RBD2 (S205C/S240C Gd(III)-DOTA) and SL-F (U15 IAP) with a complex ratio of 1:3

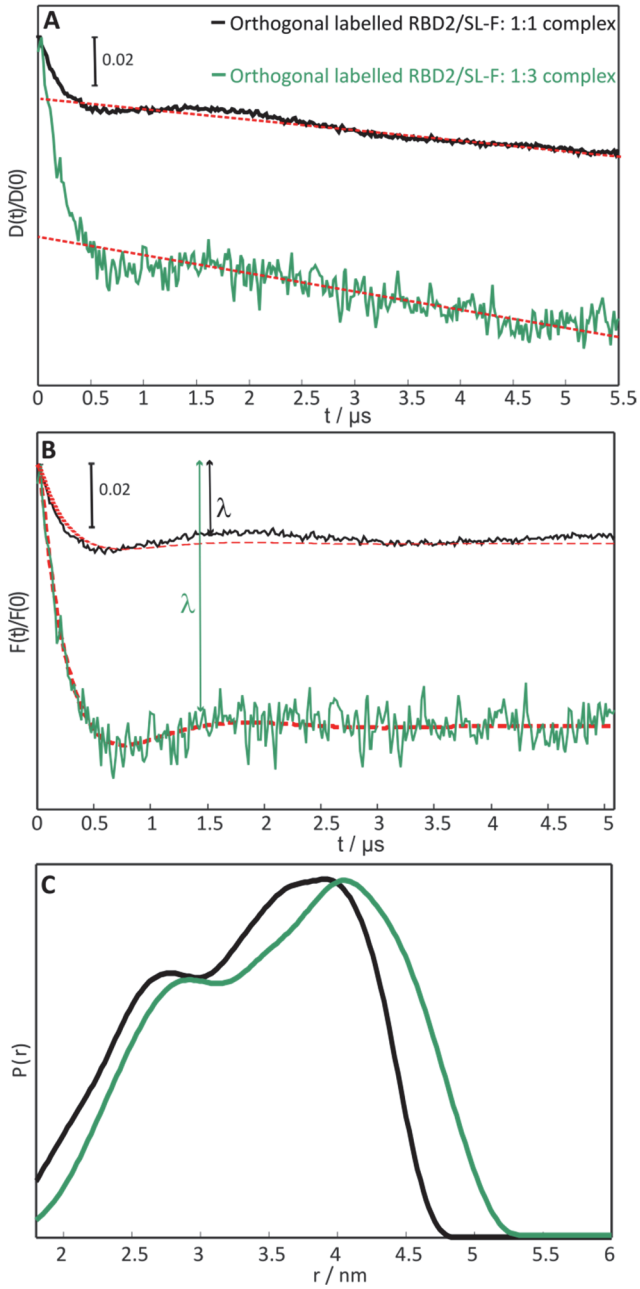

Figure S23 Orthogonal DEER of the complex RBD2/SL-F. Positions S205C and S240C of the protein are labelled with Gd(III)-DOTA and position U15 in the RNA with IAP. RBD/RNA complexes were formed in a 1:1 ratio (black) and in a 1:3 ratio (green). Experimental data including background correction (red dashed lines, A) and background-corrected data including form factor fitting (red dashed lines, B) show a significant increase of the modulation depth for the 1:3 ratio where the probability of complex formation between RBD2 and SL-F is higher. The distance distributions (C) are similar, as expected, and fit to the simulation (see Fig. 9).

4.4 Relaxation measurements T_1 and T_2 of RBDs

T_1 relaxation

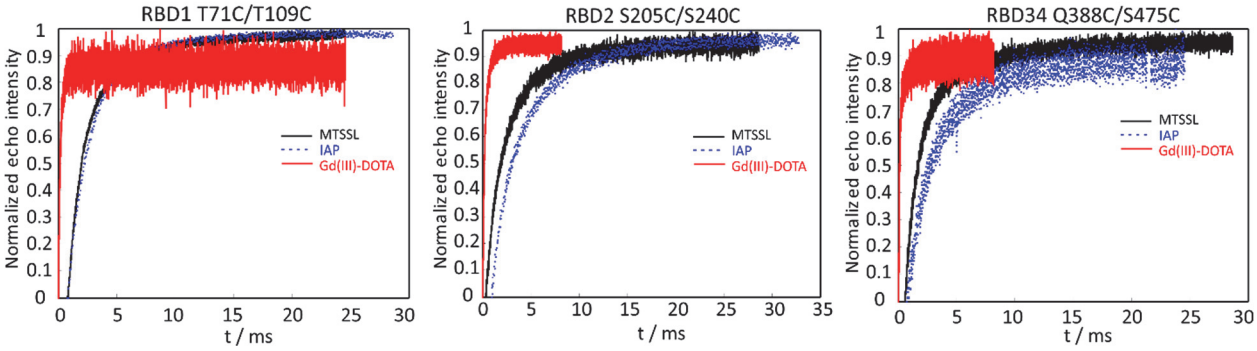

Figure S24 T_1 relaxation measurements of MTSSL- (black), IAP- (blue) and Gd(III)-DOTA- (red) labelled RBD1, RBD2 and RBD34. Since Gd(III) represents a high-spin system, T_1 is significantly shorter compared to nitroxide radicals. All samples were measured in D_2O/d8-glycerol (1:1) (Sigma Aldrich). The inversion recovery experiments were carried out at 50 K for nitroxide labelled RBDs and at 10K for Gd(III)-DOTA labelled RBDs.

T_2 relaxation of all RBDs

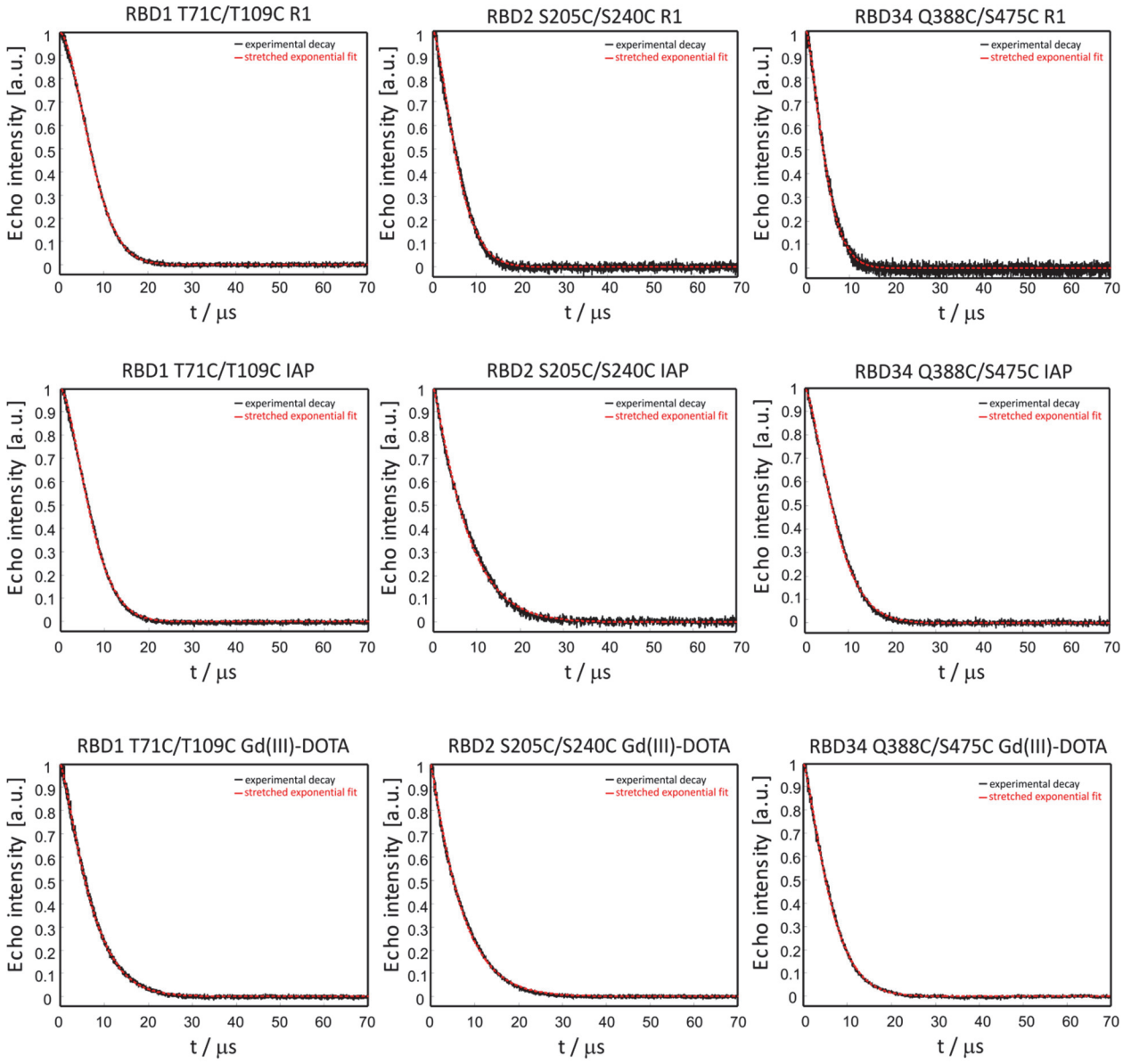

Figure S25 T_2 relaxation measurements of MTSSL-, IAP- and Gd(III)-DOTA-labelled RBD1 (left), RBD2 (centre) and RBD34 (right) including stretched exponential fits (red). All samples were measured in D_2O/d8-glycerol (1:1) (Sigma Aldrich) and the Hahn-echo experiments were carried out at 50 K for nitroxide-labelled RBDs and at 10 K for Gd(III)-DOTA labelled RBDs. As it can be seen in Table 2, T_2 relaxation of the measured samples were comparable.

T_1 and T_2 relaxation of orthogonal-labelled RBD/RNA complexes

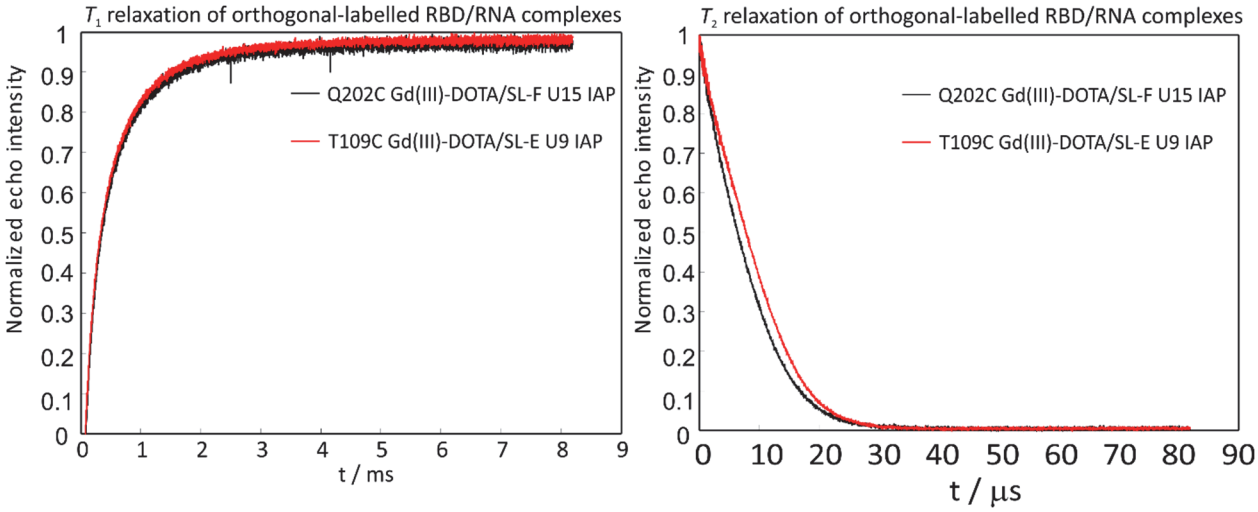

Figure S26 T_1 and T_2 relaxation measurements of orthogonal-labelled RBD/RNA complexes having one Gd(III)-DOTA spin label on the protein and one IAP on the RNA SLs. All samples were measured in D_2O/d8-glycerol (1:1) (Sigma Aldrich) and both experiments, the inversion recovery and the Hahn-echo experiment, were carried out at 10 K. The overlaid spectra show similar relaxation behaviour for T_1 and T_2 of the respective samples (see also Table 2).

5 Representation of the Gd(III)-DOTA spin clouds and comparison with the distance distributions

Table S4. Comparison of peak maxima and standard deviations of experimental obtained and simulated distance distributions and representation of the spin cloud orientations.

Samples	Spin labels	orientation of spin clouds	$r_{max.}^{A}$ / nm	$r_{max.}^{B}$ / nm	Δ	$s(r)^{A}$ / nm	$s(r)^{B}$ / nm	Δ
RBD1								
T71C/T109C	Gd(III)-DOTA		3.83	3.25	0.58	0.86	0.41	0.45
	MTSSL		3.22	3.45	0.23	0.31	0.35	0.04
RBD2								
S205C/S240C	Gd(III)-DOTA		2.86	2.40	0.46	0.58	0.42	0.16
	MTSSL		2.00	2.40	0.40	0.47	0.34	0.13
RBD34								
Q388C/S475C	Gd(III)-DOTA		5.03	4.60	0.38	0.62	0.40	0.22
	MTSSL		3.86	4.10	0.24	0.41	0.30	0.11
RBD1/SL-E								
T109C/U11	Gd(III)-DOTA / IAP		5.10	4.40	0.70	0.70	0.46	0.24
RBD2/SL-F								
Q202C/U15	Gd(III)-DOTA / IAP		3.31	3.10	0.21	0.52	0.55	0.03
S205C/U15	Gd(III)-DOTA / IAP		3.24	3.15	0.09	0.40	0.51	0.11
S240C/U15	Gd(III)-DOTA / IAP		4.46	4.75	0.29	0.44	0.44	0.00

[A] Experimental data. [B] Simulated data.

SDSL attachment and spatial distributions of the spin density was simulated with software package MMM.[10] When comparing experimental and simulated distance distributions for Gd(III)-labelled RBDs, shifts in the peak maxima of about 0.4 to 0.7 nm could be seen (Table 4). In contrast, nitroxide-nitroxide DEER delivered always distance distributions with peak maxima that differ by only about 0.2-0.4 nm from the simulation which is in line with previously studies.[11] However, the shifts of the peak maxima and the differences in the width for Gd(III)-DOTA labelled RBDs were different for the individual RBDs as well as for the RBD/RNA complexes. Table S4 represents a comparison of the Gd(III)-DOTA results to MTSSL-labelled RBDs since this spin label showed better agreement with the simulations than Gd(III)-DTPA.

In general, the spatial spin cloud has rather a mushroom-like shape than a spherically symmetry. Thus, we created 'mushroom'-models for the spatial spin clouds to abstract the shape. The mushroom-like shape is shown by an oblate ellipsoid and a vector (see Fig. S.27, 28), both computed from the coordinates of the spin centres of all rotamers and their populations. This covers the N-O bond in case of nitroxides and the Gd^{3+}-ion in case of Gd(III) spin labels. The cap is the ellipsoid representation of a tensor (3x3 matrix), where distance vectors, determined by the difference of spin centre coordinates and the centre coordinate of the cap, are considered. To quantify the flatness of the cap, the oblaticity is defined using $(\sigma_{22}\sigma_{33})^{1/2}/\ \sigma_{11}$, where σ represents the eigenvalues of the tensor. In addition, the contraction of the standard deviation tensor due to clashes is quantified by the factor $\chi = (\sigma_{11}\ \sigma_{22}\ \sigma_{33})/\ (\sigma_{11,0}\ \sigma_{22,0}\ \sigma_{33,0})$.

In Figure S27 and S28, the 'mushroom'-models are shown for the free label (green ellipsoid for the cap and green line for the vector between spin centre and Cα atom) and for the restrained label considering clashes with the biomolecule (brown ellipsoid for the cap and red line for the vector between spin centre and Cα atom). The angle θ represents the deviation of the free label and the clashed-restrained label. All determined contraction-, oblaticity- and θ-values for the labelling sites in RBD1, RBD2 (Fig. S27) and RBD34 (Fig. S28) are shown in Table S5.

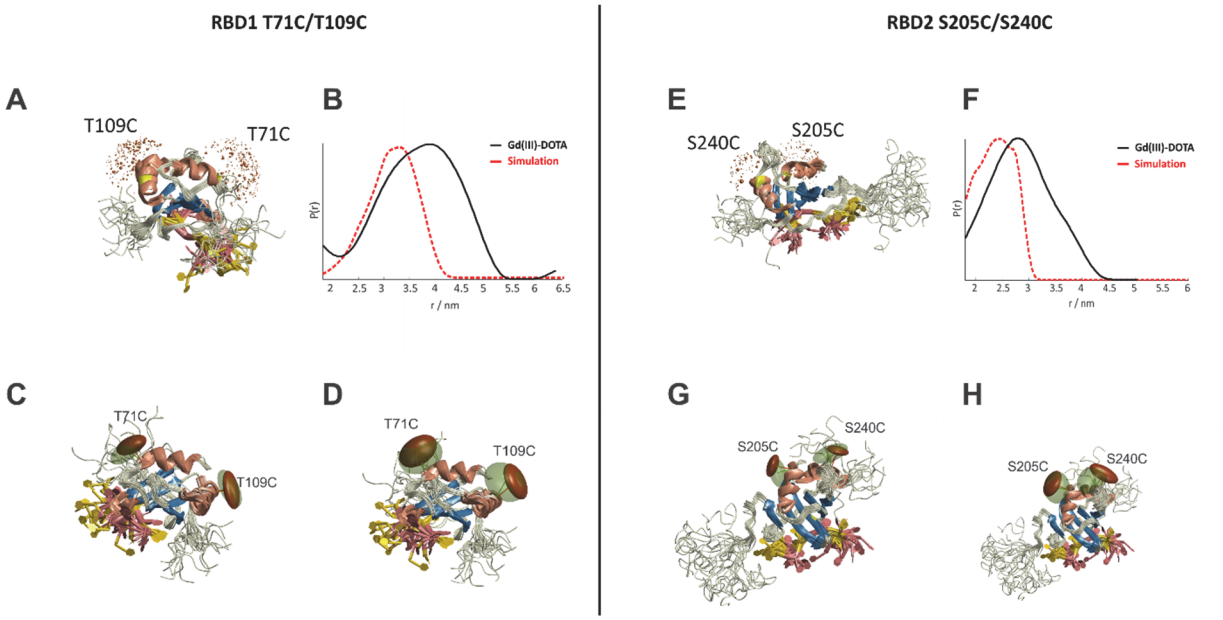

Figure S27 Spin label conformations and spatial spin clouds of MTSSL and Gd(III)-DOTA in RBD1 and RBD2. A) Spin label conformation of RBD1 T71C/T109C Gd(III)-DOTA; B) Distance distribution of RBD1 T71C/T109C Gd(III)-DOTA; C) 'Mushroom model' for MTSSL spin clouds in RBD1 T71C/T109C; D) 'Mushroom model' for Gd(III)-DOTA spin clouds in RBD1 T71C/T109C; E) Spin label conformation of RBD2 S205C/S240C Gd(III)-DOTA; F) Distance distribution of RBD2 S205C/S240C Gd(III)-DOTA; G) 'Mushroom model' for MTSSL spin clouds in RBD2 S205C/S240C; H) 'Mushroom model' for Gd(III)-DOTA spin clouds in RBD2 S205C/S240C.

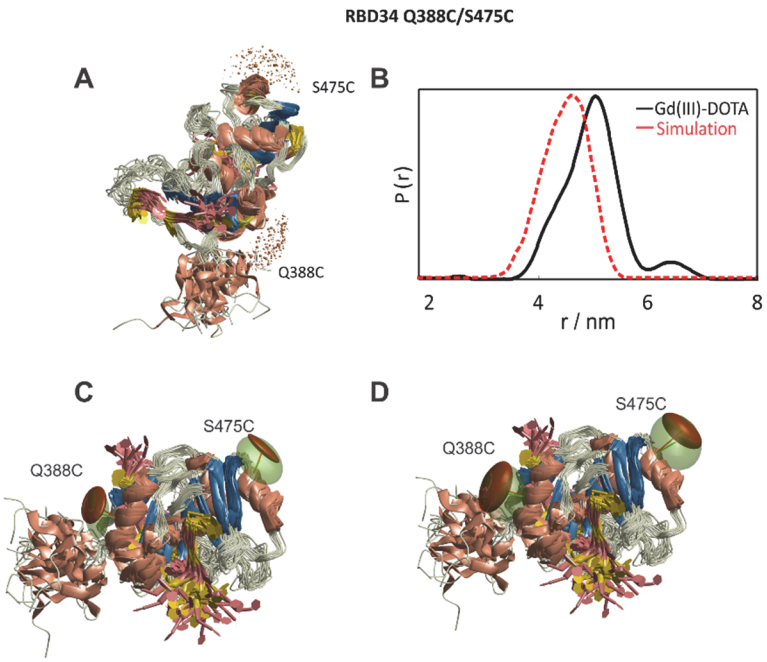

Figure S28 Spin label conformations and spatial spin clouds of MTSSL and Gd(III)-DOTA in RBD34. A) Spin label conformation of RBD34 Q388C/S475C Gd(III)-DOTA; B) Distance distribution of RBD34 Q388C/S475C Gd(III)-DOTA; C) 'Mushroom model' for MTSSL spin clouds; D) 'Mushroom model' for Gd(III)-DOTA spin clouds.

Table S5. Comparison of spatial spin clouds ('mushroom'- models) in terms of contraction, oblaticity and θ at different labelling positions in the individual RBDs.

Position	Label	Contraction	Oblaticity*	θ [°]	Rotamers+ [min. | max.]	Rotamers++ [average]	partition fct.	ζ^{+++}
RBD1 (2AD9)								
T71C	MTSSL	0.778	2.174	15.9	94 | 121	113 +/- 7	1.326 +/- 0.222	0.53
	Gd(III)-DOTA	0.694	1.618	18.3	197 | 308	272 +/- 27	1.186 +/- 0.147	0.42
T109C	MTSSL	0.560	2.659	11.2	68 | 87	79 +/- 5	1.328 +/-0.073	0.37
	Gd(III)-DOTA	0.381	2.380	17.3	171 | 227	209 +/- 14	0.922 +/- 0.105	0.32
RBD2 (2ADB)								
S205C	MTSSL	0.469	2.610	11.5	31 | 61	49 +/- 8	0.739 +/- 0.311	0.23
	Gd(III)-DOTA	0.383	2.342	11.7	121 | 193	162 +/- 17	0.907 +/- 0.138	0.25
S240C	MTSSL	0.292	3.025	7.6	23 | 48	36 +/- 5	1.463 +/- 0.251	0.17
	Gd(III)-DOTA	0.283	1.605	23.6	61 | 99	76 +/- 10	0.430 +/- 0.096	0.12
RBD34 (2ADC)								
Q388C	MTSSL	0.334	3.208	19.6	13 | 87	69 +/- 19	1.111 +/- 0.270	0.32
	Gd(III)-DOTA	0.417	2.078	7.3	32 | 268	207 +/- 62	0.888 +/- 0.178	0.32
S475C	MTSSL	0.206	3.718	25.1	29 | 66	54 +/- 9	0.949 +/- 0.458	0.25
	Gd(III)-DOTA	0.369	2.102	9.5	120 | 201	163 +/- 24	0.959 +/- 0.159	0.25

* The oplaticity for the free label is given by 2.056 (MTSSL) and 1.274 (Gd(III)-DOTA). [+]Minimum and maximum number of rotamers selected from the NMR ensemble containing 20 structures. [++]Averaged number of rotamers over the whole NMR ensemble containing 20 structures. [+++]Ratio between simulated attached averaged rotamers and rotamer numbers of the free labels (216 for MTSSL and 648 for Gd(III)-DOTA).

Table S5 shows the number of rotamers for MTSSL and Gd(III)-DOTA of all labelling sites in the different RBD samples. Note, that the structure with the smallest rotamer number of the 20-structure NMR ensemble as well as the one with the largest rotamer number was taken to show the variation. In addition, the average of all rotamer numbers of the whole ensemble was determined. The same procedure was done for the partition function. It is clearly visible that Gd(III)-DOTA has a larger number of rotamers but also a larger difference between minimum and maximum value. Further, we computed the ratio between the averaged simulated values and the rotamer numbers of the free labels MTSSL (216 rotamers) and Gd(III)-DOTA (648 rotamers). Unfortunately, both methods, the spatial spin cloud and the rotamer analysis, did not provide any correlation to the deviation in the distance distributions of Gd(III)-DOTA spin labels explaining that the main contribution to these deviations is not a systematic error due to insufficient quality of the rotamer library. However, a distinct fact is that Gd(III)-DOTA shows in all cases a lower oblaticity, which is comprehensible, because of its larger volume the spin label has a higher tendency to a spherical shape than MTSSL. A small deviation in the angles of the spatial spin clouds in RBD1 and RBD2 Gd(III)-DOTA samples compared to MTSSL can be seen but since the opposite is the case for RBD34, we cannot correlate this to the deviation in the distance distributions of the RBD1 and RBD2 Gd(III)-DOTA samples. Furthermore, we performed rotamer analysis of all used maleimido-linker spin labels (MAP, Gd(III)-DOTA and Gd(III)-DTPA) to confirm that the label size itself does not restrict the labelling attachment and hence does not have an influence on the labelling efficiency. This is actually in line with our experiments (see Table 4 and Table S6). This analysis, in combination with the respective intra-domain distance distributions (Fig. 5 and 7, Fig. S17-20) for MAP, Gd(III)-DOTA and Gd(III)-DTPA, showed that the use of the larger maleimido linker (compared to the MTS linker) can be neglected as potential source for these deviations since data derived from MAP spin label agree very accurately with the prediction. This strongly underlines our expectations.
However, the uncertainties might rather occur from different sources, as for example, in prediction the interaction of the label with the local environment or experimental Gd(III)-Gd(III) DEER data with lower quality as a result of the lower modulation depth. Latter one could be removed by using either ultra wide band DEER and RIDME techniques.[12-14]

Table S6. Rotamer analysis of nitroxide and Gd(III)-based spin labels with a maleimido group as linker type.

Position	Label	Rotamers [min. | max.]+	Rotamers++ [average]	partition fct.	ζ^{+++}
RBD1 (2AD9)					
T71C	MAP	58 | 83	76 +/- 6	1.420 +/- 0.100	0.7
	Gd(III)-DOTA	197 | 308	272 +/- 27	1.186 +/- 0.147	0.42
	Gd(III)-DTPA	354 | 605	519 +/- 55	0.826 +/- 0.145	0.21
T109C	MAP	51 | 61	58 +/- 3	1.122 +/- 0.123	0.54
	Gd(III)-DOTA	171 | 227	209 +/- 14	0.922 +/- 0.105	0.32
	Gd(III)-DTPA	328 | 432	382 +/- 31	0.917 +/- 0.125	0.16
RBD2 (2ADB)					
S205C	MAP	28 | 47	38 +/- 5	0.933 +/- 0.207	0.35
	Gd(III)-DOTA	121 | 193	162 +/- 17	0.907 +/- 0.138	0.25
	Gd(III)-DTPA	233 | 390	301 +/- 45	0.518 +/- 0.186	0.12
S240C	MAP	13 | 30	24 +/- 4	0.324 +/- 0.068	0.22

	Gd(III)-DOTA	61 | 99	76 +/- 10	0.430 +/- 0.096	0.12
	Gd(III)-DTPA	76 | 220	155 +/-28	0.377 +/- 0.089	0.06
RBD34 (2ADC)					
Q388C	MAP	8 | 72	55 +/- 16	1.285 +/- 0.372	0.51
	Gd(III)-DOTA	32 | 268	207 +/- 62	0.888 +/- 0.178	0.32
	Gd(III)-DTPA	42 | 419	302 +/- 101	0.714 +/- 0.250	0.12
S475C	MAP	25 | 51	41 +/- 7	1.200 +/- 0.311	0.38
	Gd(III)-DOTA	120 | 201	163 +/- 24	0.959 +/- 0.159	0.25
	Gd(III)-DTPA	164 | 355	255 +/- 61	0.483 +/- 0.224	0.11

[+]Minimum and maximum number of rotamers selected from the NMR ensemble containing 20 structures. [++]Averaged number of rotamers over the whole NMR ensemble containing 20 structures. [+++]Ratio between simulated attached averaged rotamers and rotamer numbers of the free labels (108 for MAP, 648 for Gd(III)-DOTA and 2430 for Gd(III)-DTPA).

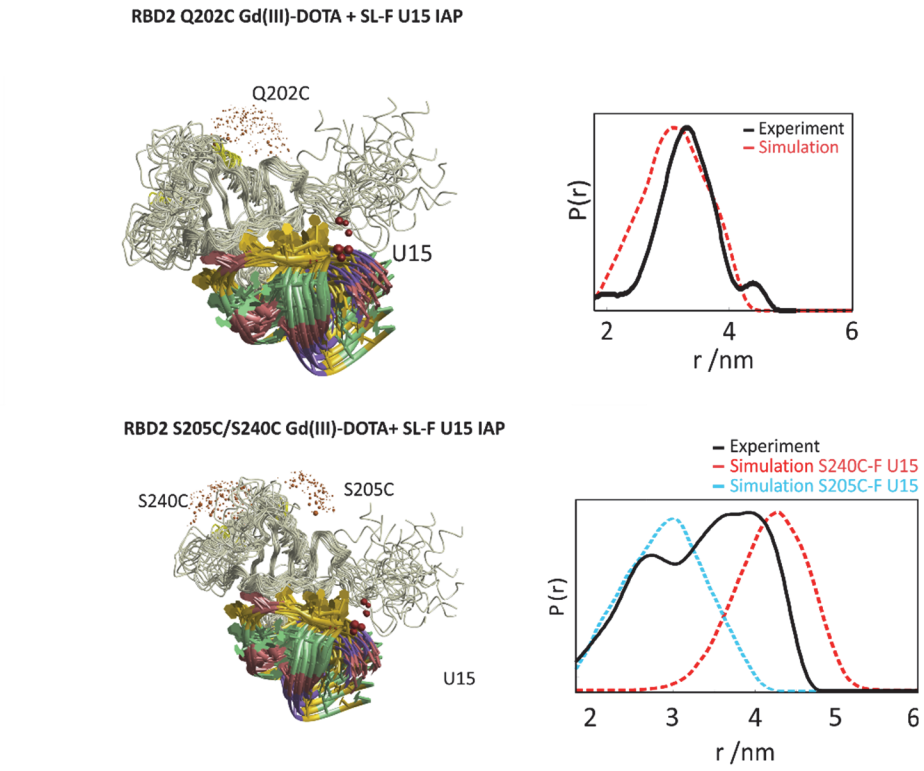

Figure S29 Orthogonal-labelled RBD2/SL-F complex with Q202C (Gd(III)-DOTA) and SL-F U15 (IAP) is shown above and the RBD2/SL-F complex with S205C/S240C (Gd(III)-DOTA) and SL-F U15 (IAP) is shown below. In the first case, the spin clouds are pointing almost perpendicular to each other which could cause a slight shift to longer distances, but the width agrees very well with the simulation. In RBD2 S205C/S240C Gd(III)-DOTA/SL-F U15 IAP, the spin clouds, representing the shorter protein-RNA distance (S205C-U15), are pointing more or less in the same direction leading to a very good agreement between simulation and experiment. Although for the longer distance (S240C-U15) both clouds are nearly parallel oriented, a slight shift of the peak maximum is again observed.

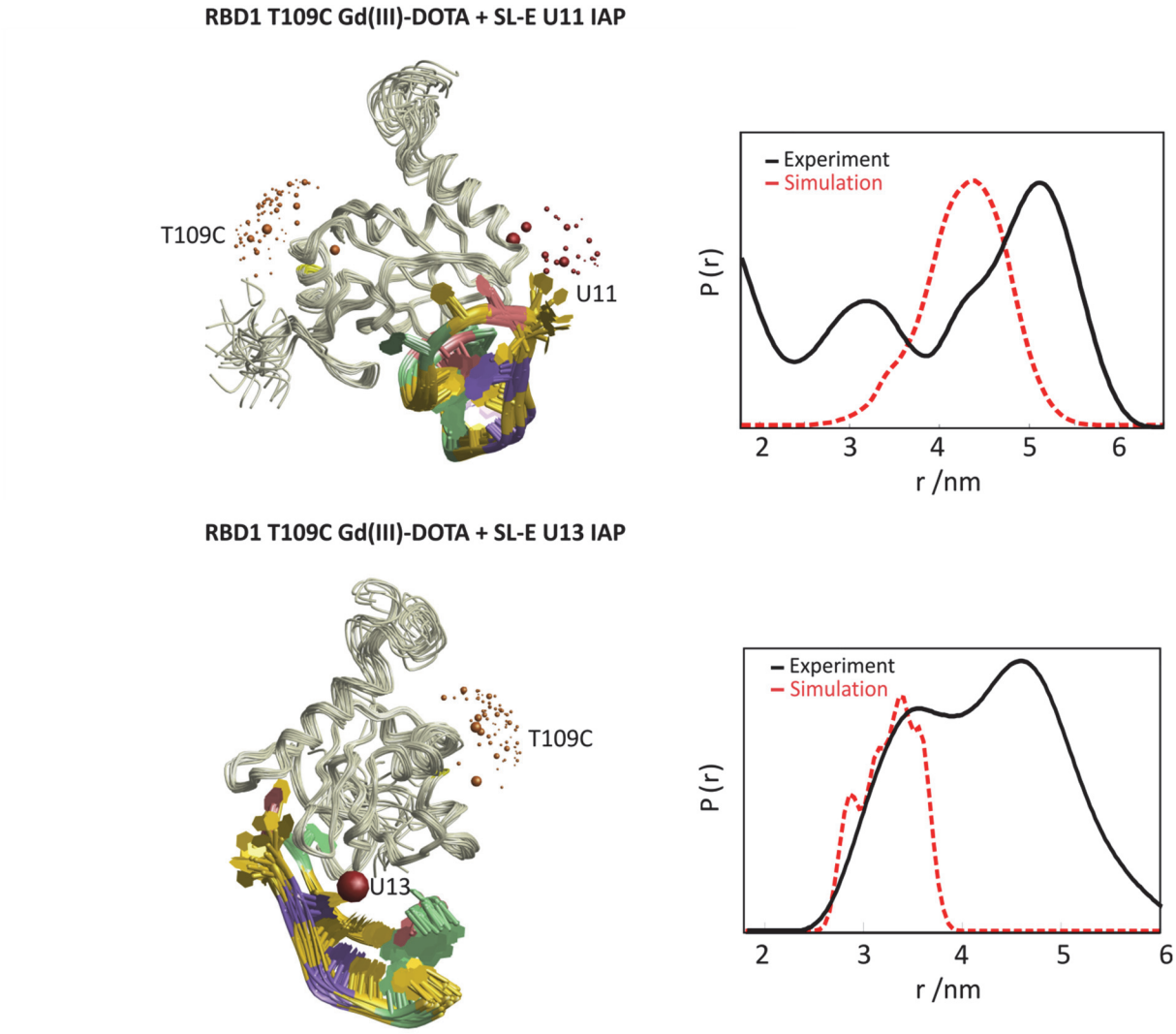

Figure S30 Orthogonal-labelled RBD1/SL-E complexes models with T109C (Gd(III)-DOTA) and SL-E U11 (IAP), respectively U13 (IAP). In the upper case, both spin label clouds are pointing almost perpendicular to each other which leads to a slight shift of 0.7 nm of the main peak to higher distances compared to the simulation (red dashed). A shorter protein-RNA distance at 3 nm is also visible, which indicates bimodality. In the case shown in the lower panel, the 4-thiouridine in the stem is labelled which might lead to conformational changes in the RNA stem and thus might be the reason for the obtained bimodal distance distribution. The shorter protein-RNA distance agrees very well with the predicted distance distribution (red dashed). In this model both spin clouds are pointing more or less in the same direction.

6 Influence of structural elements on the distance distributions in RBD1/SL-E complex

Besides the obtained biomodality in RBD1(T109C)/SL-E U9 and RBD1(T109C)/SL-E U13, where at least one peak of the distance distributions match very accurately with the respective simulation, the situation is different for RBD1(T109C)/SL-E U11. In this particular complex, one can recognize a more spread protein-RNA distance between the labelling sites. Since RBD1 seems to form the third α-helix only when it is bound to structured RNA, we took into account that this helix might have an influence on the spatial spin label cloud. Indeed, by creating a model of RBD1 (XAD9), based on the solution-NMR structure without an additional helix (2AD9), bound to RNA SL-E we could detect influences of the helix to the conformational space of the spin label (Fig. S31). However, this influence does not alter neither the rotamer analysis (Table S7) nor the simulated distance distribution when both models are compared. The fact, that the labelling position U11 in the RNA stem loop is not experimentally restrained and thus shows varied conformations in the ensemble of the CLIR-MS/MS model[15] (Fig. S31), seems to be the most probable reason for the deviation in the simulation and the experimental data. When comparing all single models of the ensembles (Fig. S31) one can recognize that some models (~ 20 %) match with the shorter experimental distance peak around 3 nm. Further, nearly half of the models match quite good with the most prominent experimental distance distribution around 4.5 to 5 nm with peak maxima differences of about 0.5 nm, which is on the edge of our acceptable accuracy range. The rest of these 20 models, which incorporates nearly a third, represents a distance distribution peak between both experimental distributions, leading to a shift of the ensemble distance distribution in a range of 4 to 4.5 nm. This analysis supports our assumption that the disagreement of simulation and experimental data seems to occur from uncertainties in the RNA structure, especially for nucleotide U11.

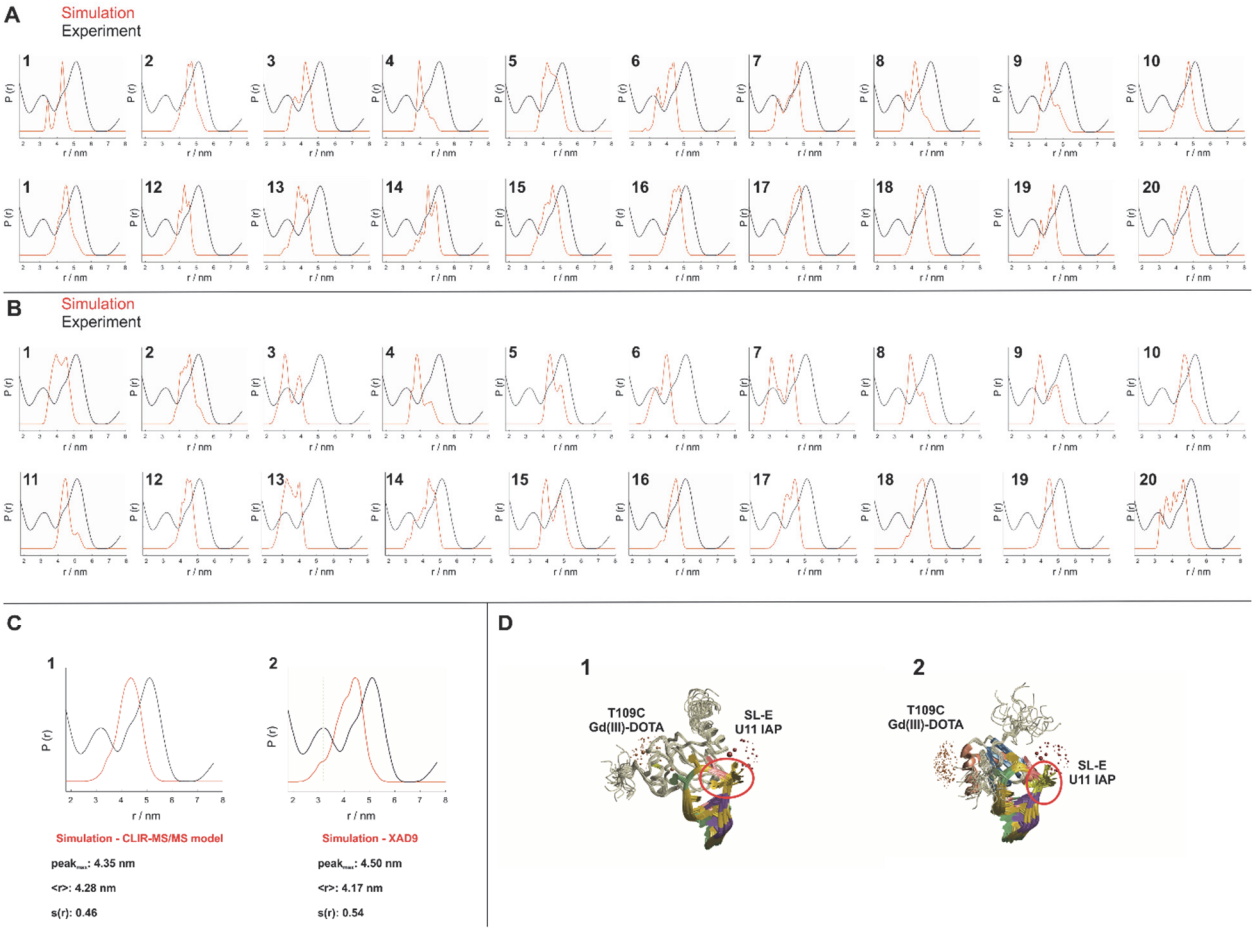

Figure S 31. Comparison of experimental and simulated distance distribution based on the 20 single models of the structural ensemble of the CLIR-MS/MS model[15] (including the third α-helix) and the XAD9 model (without additional α-helix). A) Distance distribution of T109C (Gd(III)-

DOTA) and SL-E U11 (IAP) compared to all single models of the CLIR-MS/MS model. B) Distance distribution of T109C (Gd(III)-DOTA) and SL-E U11 (IAP) compared to all single models of the XAD9 model. C) Experimental distance distribution compared to the whole ensemble of the CLIR-MS/MS model (1) and the created XAD9 model (2). D) Spatial spin label cloud of T109C Gd(III)-DOTA and SL-E U11 IAP of the CLIR-MS/MS model (1) and the XAD9 model (2). The third α-helix seems to alter the conformational distribution of the spin label around position U11 in the RNA. Further, the varied conformations of U11 is visible which occurs because U11 is experimentally not restrained.

Table S 7. Comparison of rotamer analysis for of both RBD1/SL-E models (CLIR-MS/MS model and XAD9 model).

Model	Position	Rotamers	partition fct.	Model	Position	Rotamer	partition fct.
CLIR-MS/MS	*SL-E U11 IAP*	25	0.65385	*XAD9*	*SL-E U11 IAP*	32	0.68665
		28	0.83145			34	0.70178
		23	0.63829			26	0.78954
		31	0.51155			31	0.55709
		39	0.51393			37	0.42772
		44	0.48651			41	1.43728
		34	0.51609			43	0.68947
		43	0.61433			44	0.52867
		37	0.9693			43	0.62629
		39	0.90079			43	0.53788
		45	0.67113			46	1.18373
		48	0.64675			50	1.10253
		47	1.14017			48	0.6769
		40	0.82448			45	1.01035
		49	1.02459			49	1.00182
		50	1.01839			51	1.01069
		50	1.02003			50	1.01763
		51	1.02688			50	1.01881
		38	1.08939			37	1.25063
		36	0.51413			48	1.23964
	$Rotamers_{Min}$	23	0.487			26	0.428
	$Rotamers_{Max}$	51	1.140			51	1.437
	$Rotamers_{Ave}$	40 +/- 8	0.781			42 +/- 7	0.875

7 References

1. F. C. Oberstrass, S. D. Auweter, M. Erat, Y. Hargous, A. Henning, P. Wenter, L. Reymond, B. Amir-Ahmady, S. Pitsch, D. L. Black and F. H. Allain, *Science*, 2005, 309, 2054-2057.
2. P. J. Simpson, T. P. Monie, A. Szendroi, N. Davydova, J. K. Tyzack, M. R. Conte, C. M. Read, P. D. Cary, D. I. Svergun, P. V. Konarev, S. Curry and S. Matthews, *Structure*, 2004, 12, 1631-1643.
3. A. Joshi, M. B. Coelho, O. Kotik-Kogan, P. J. Simpson, S. J. Matthews, C. W. Smith and S. Curry, *Structure*, 2011, 19, 1816-1825.
4. P. Kafasla, N. Morgner, T. A. Poyry, S. Curry, C. V. Robinson and R. J. Jackson, *Mol Cell*, 2009, 34, 556-568.
5. P. Kafasla, H. Lin, S. Curry and R. J. Jackson, *RNA*, 2011, 17, 1120-1131.
6. E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch, *Nucleic Acids Research*, 2003, 31, 3784-3788.
7. H. S. McHaourab, M. A. Lietzow, K. Hideg and W. L. Hubbell, *Biochemistry*, 1996, 35, 7692-7704.
8. M. Parisien and F. Major, *Nature*, 2008, 452, 51-55.
9. R. Lorenz, S. H. Bernhart, C. H. Z. Siederdissen, H. Tafer, C. Flamm, P. F. Stadler and I. L. Hofacker, *Algorithm Mol Biol*, 2011, 6.
10. Y. Polyhach, E. Bordignon and G. Jeschke, *Phys Chem Chem Phys*, 2011, 13, 2356-2366.
11. G. Jeschke, *Prog Nucl Magn Reson Spectrosc*, 2013, 72, 42-60.
12. D. Akhmetzyanov, H. Y. Ching, V. Denysenkov, P. Demay-Drouhard, H. C. Bertrand, L. C. Tabares, C. Policar, T. F. Prisner and S. Un, *Phys Chem Chem Phys*, 2016, 18, 30857-30866.
13. S. Razzaghi, M. Qi, A. I. Nalepa, A. Godt, G. Jeschke, A. Savitsky and M. Yulikov, *J Phys Chem Lett*, 2014, 5, 3970-3975.
14. A. Doll, M. Qi, N. Wili, S. Pribitzer, A. Godt and G. Jeschke, *J Magn Reson*, 2015, 259, 153-162.
15. G. Dorn, A. Leitner, J. Boudet, S. Campagne, C. von Schroetter, A. Moursy, R. Aebersold and F. H. Allain, *Nat Methods*, 2017, DOI: 10.1038/nmeth.4235.