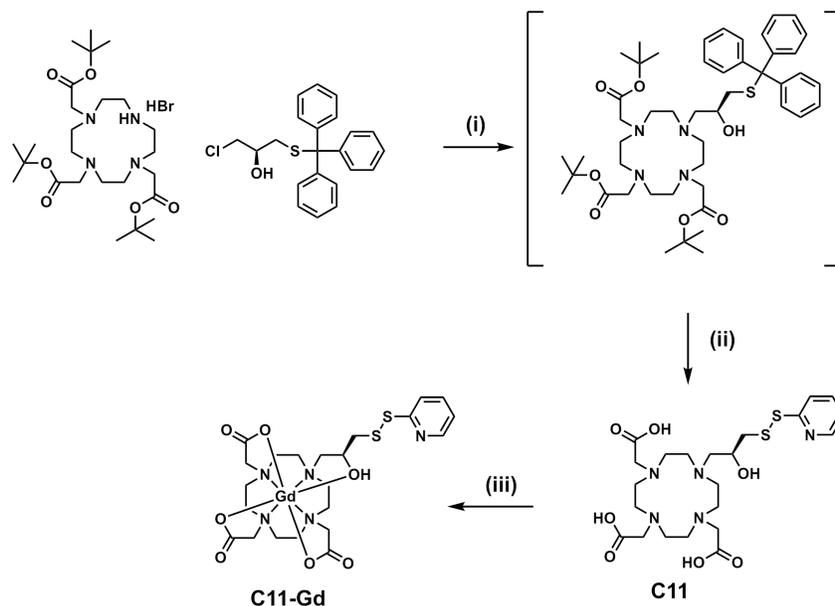


Supporting information

Small neutral Gd³⁺ tags for distance measurements in proteins by double electron–electron resonance experiments

Mithun C. Mahawaththa, Michael D. Lee, Angeliki Giannoulis, Luke A. Adams, Akiva Feintuch, James D. Swarbrick, Bim Graham, Christoph Nitsche, Daniella Goldfarb, Gottfried Otting

Synthesis of C11-Gd



Scheme 1. Synthesis of C11-Gd. Reagents and conditions: (i) K₂CO₃, MeCN, reflux, 3 days; (ii) 2,2'-dipyridyldisulfide, triethylsilane, TFA, CH₂Cl₂, room temperature, 3 days, 20% (over two steps); (iii) GdCl₃, H₂O, pH 5, 80 °C, 2 h, 64%.

Analytical instrumentation used in tag synthesis

All 400 MHz NMR spectra were recorded on a Bruker Avance III 400 MHz NMR spectrometer. Chemical shifts are quoted in units of parts per million (ppm) and were referenced internally to the residual proteo-solvent resonance. Multiplicities and appearances of NMR resonances are abbreviated as: s, singlet; d, doublet; t, triplet; q, quartet; p, pentet; m, multiplet; app, apparent; br, broad. LC-MS data were acquired on an Agilent 1220/6120 LC/MS system, using ChemStation software for instrument control and data analysis. Preparative reverse-phase HPLC was performed on an Agilent 1260 Prep HPLC using an Alltima C8 column (250 mm x 22 mm, 5 micron).

Material

Tri-*tert*-butyl 2,2',2''-(1,4,7,10-tetraazacyclododecane-1,4,7-triyl)triacetate hydrobromide ((^tBu)₃DO3A·HBr) was purchased from Toronto Research Chemicals. (*R*)-1-chloro-3-(tritylthio)propan-2-ol was synthesised following literature procedures.¹

(*S*)-2,2',2''-(10-(2-hydroxy-3-(pyridin-2-yl)disulfaneyl)propyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triyl)triacetic acid (C11)

Potassium carbonate (1.5 g, 10.85 mmol) was added to a solution of (^tBu)₃DO3A·HBr (400 mg, 0.67 mmol) and (*R*)-1-chloro-3-(tritylthio)propan-2-ol (225 mg, 0.75 mmol) in MeCN (15 mL). The mixture was heated to reflux for 3 days. After cooling to room temperature, insoluble salts were removed by filtration and the filtrate was concentrated under reduced pressure to yield a brown oil. The crude material was dissolved in a solution of 2,2'-dipyridyldisulfide (446 mg, 2.03 mmol) and triethylsilane (323 μ L, 2.03 mmol) in CH₂Cl₂ (5 mL), before the dropwise addition of trifluoroacetic acid, TFA (5 mL). The solution was stirred for 3 days at room temperature, after which time LCMS analysis indicated complete *tert*-butyl group deprotection and conversion of the thiotrityl group to a pyridyldisulfide. Volatile solvents were removed by passing a gentle N₂ stream over the open reaction vessel. The resulting residue was taken up in CH₂Cl₂ (10 mL) and washed with a solution of TFA (0.1% v/v) in H₂O (10 mL), followed by H₂O (10 mL). The combined aqueous layers were purified by dry flash column chromatography, using Davisil® P60 C18 (35-70 μ m) silica gel and a gradient from 0-20% MeCN in H₂O with 0.1% (v/v) TFA. Fractions containing pure product were lyophilized to yield **C11** as a beige solid. Yield: 73 mg (20%). ¹H NMR (400 MHz, D₂O) δ 8.46 (m, 1H), 7.91 (m, 1H), 7.86 (d, *J* = 7.99 Hz, 1H), 7.37 (m, 1H), 4.14 (br s, 1H), 3.80–3.55 (m, 7 H), 3.41–2.90 (m, 23 H). ¹³C NMR (101 MHz, D₂O) δ 168.75, 158.00, 148.56, 139.35, 122.38, 66.11, 56.82, 55.81, 54.54, 49.49 (br), 43.36. LC-MS: *m/z* (ESI, 20V) 546.3 [M+H]⁺.

C11-Gd(III) tag

C11 (10 mg, 0.018 mmol) and GdCl₃ (5 mg, 0.20 mmol) were dissolved in H₂O (2 mL) and the pH of the solution was adjusted to ~ 5 by the addition of DIPEA. The solution was heated to 80 °C for 2 h, after which time LC-MS analysis indicated no uncomplexed **C11** remained. After cooling to room temperature, the complex was purified by reverse-phase HPLC (0.1% TFA and 5-95% MeCN over 20 min on a C8 preparative column). Fractions containing pure product were lyophilized to yield **C11-Gd(III)** as a beige solid. Yield: 8 mg (64%). LC-MS: *m/z* (ESI, 20V) 700.2 (complex isotope pattern) [M+H]⁺.

NMR and LC-MS data

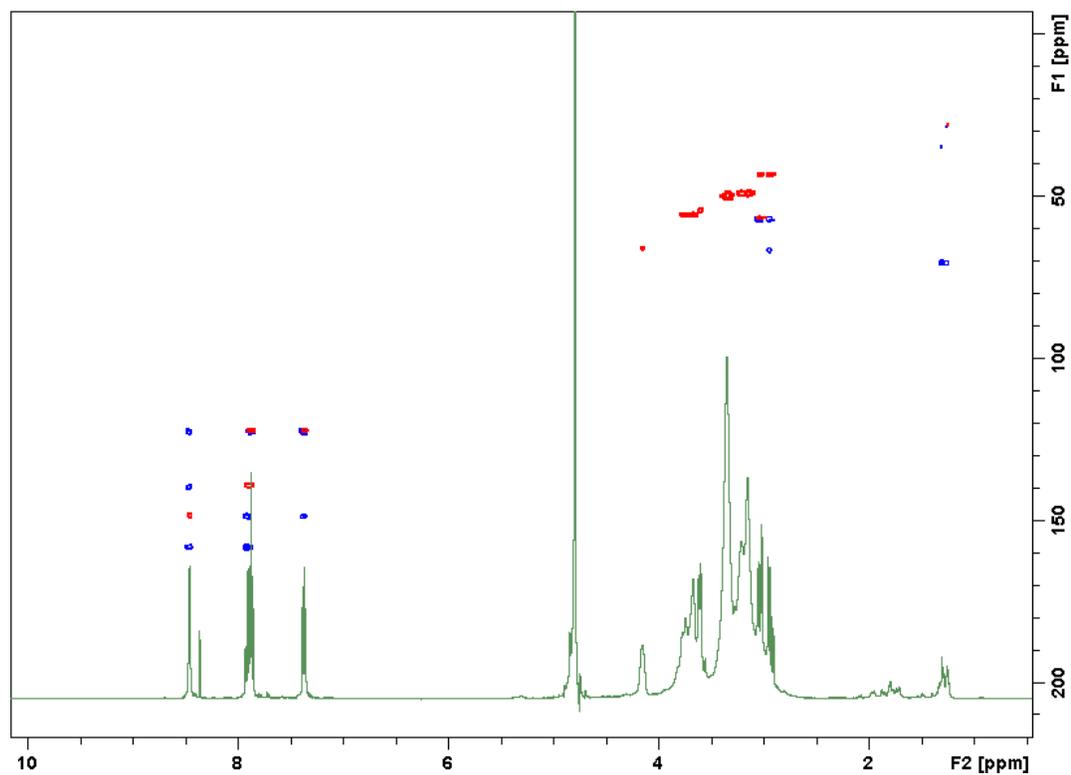


Figure S1. Superimposition of ^1H -NMR (green), ^{13}C -HSQC (red) and ^{13}C -HMBC (blue) spectra of **C11** in D_2O at pH 3.

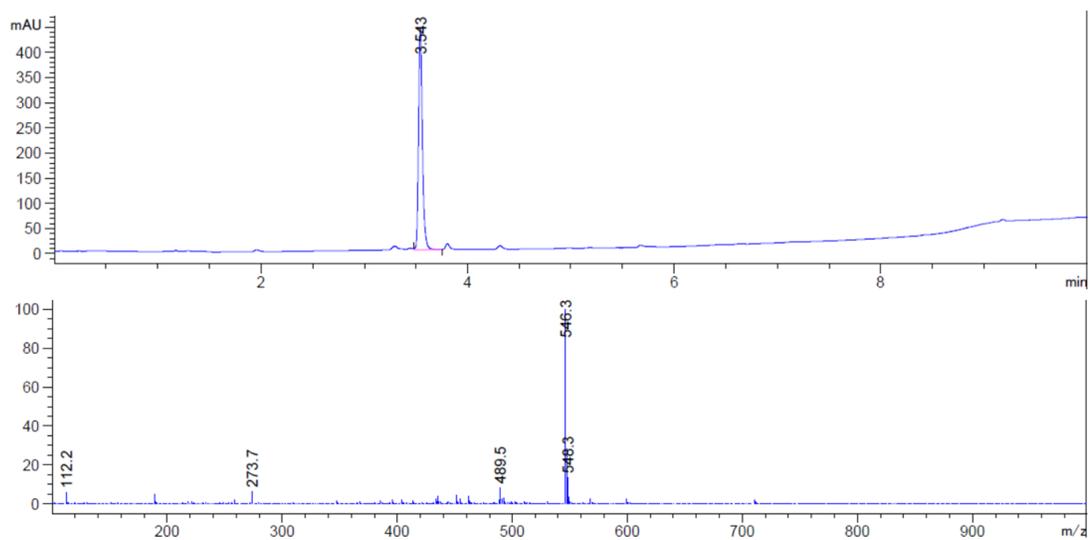


Figure S2. LC-MS UV trace at 254 nm (upper panel) and positive mass spectrum (lower panel) of **C11**.

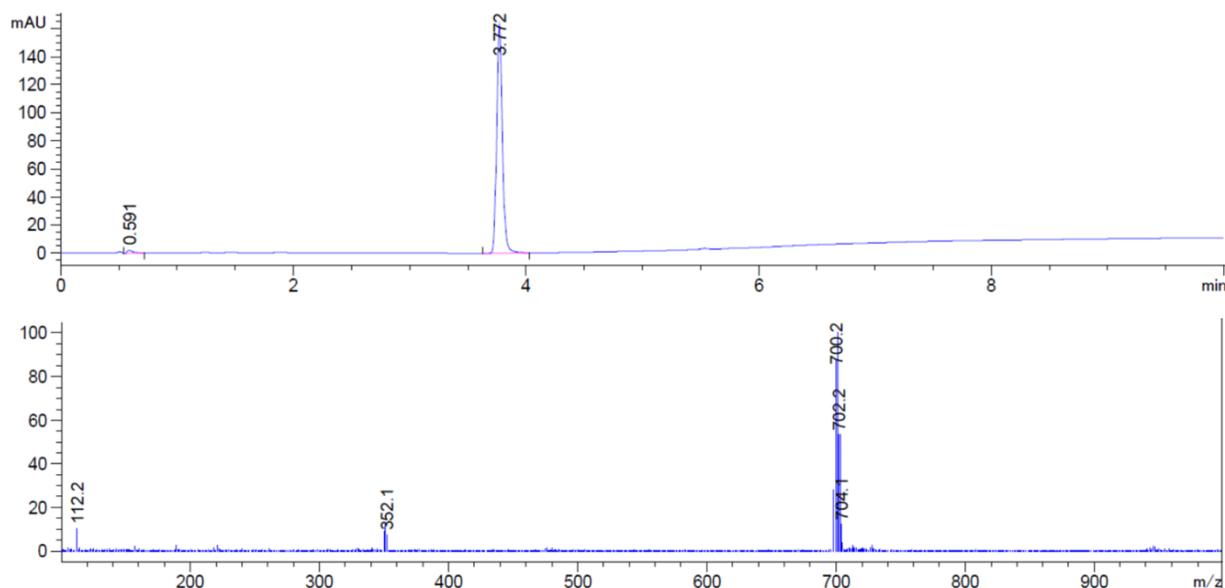


Figure S3. LC-MS UV trace at 254 nm (upper panel) and positive mass spectrum (lower panel) of **C11** loaded with gadolinium.

EPR measurements and analysis

Simulations of ED-EPR spectra of DEBP labeled with C3 or propargyl-DO3A

The ED-EPR spectra were simulated using the ‘pepper’ function in the program EasySpin² and taking into account the distribution of the axial (D) and rhombic (E) parameters of the ZFS as suggested by Raitsimring *et al.*^{3,4} The distribution over D is given by two Gaussian functions centered at $-D$ and $+D$ with equal widths σ_D and equal weights. The probability distribution of E/D is given by

$$P(E/D) = (E/D) - 2(E/D)^2 \quad (1)$$

The frequency was 94.9 GHz and the temperature was 10 K. To obtain good fits we had to introduce a linewidth (lw). Table S1 shows the parameters used in the simulations of Figures S4 and S8. The simulations did not take into account the underestimation of the amplitude of the broad background arising from an adjustment of the nominal $\pi/2$ and π pulses to the central transition, which has the highest transition probability. This may be one of the reasons for not obtaining better fits.

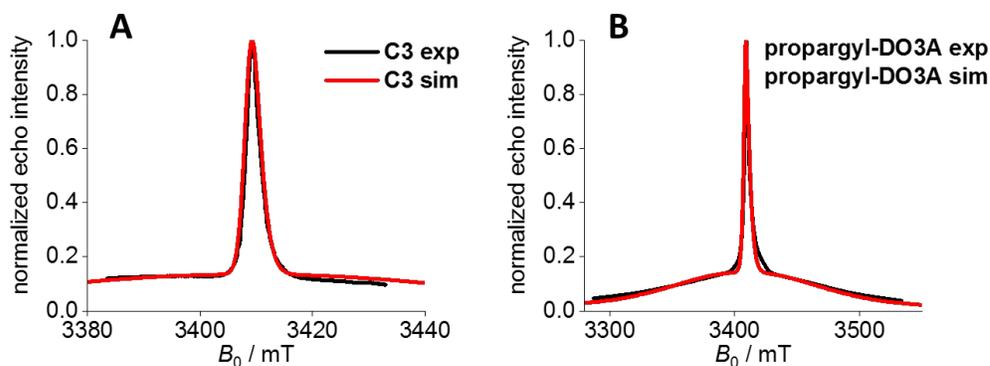


Figure S4. Simulations of the ED-EPR spectra of DEBP and comparison with experimental spectra. (A) N48/R169 mutant labelled with the **C3** tag. (B) N48/R169 mutant labelled with **propargyl-DO3A**.

Table S1. Summary of the parameters used to simulate the spectra presented in Figures S1 and S5.

Tag	D/MHz	linewidth/mT
C3	850	2
propargyl-DO3A	1250	1
C11^a	1400	1
C9^b	800	3

^a Different weights of the two Gaussians were required for a reasonable fit. The relative weights for the $-D$ and D parameters were 100:1.

^b The D value of **C9** is probably overestimated, as an overly large linewidth had to be added to reproduce the broad signal due to the transitions other than the central one.

Echo decays of DEBP

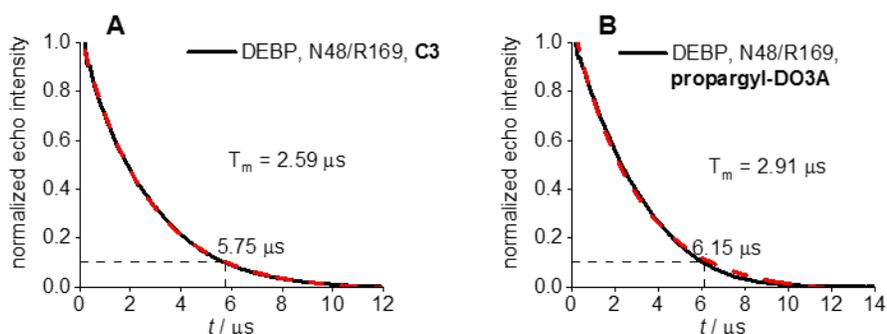


Figure S5. Echo decay data measured at W-band and 10 K. The T_m values obtained by fitting a mono-exponential decay (shown as a red dashed line) are given as well as the values at which the echo intensity has decayed to 10% of its initial intensity. (A) DEBP mutant N48/R169 labelled with the C3 tag. (B) DEBP mutant N48/R169 labelled with propargyl-DO3A.

ED-EPR spectra of the Zika virus NS2B-NS3 protease mutants and simulations

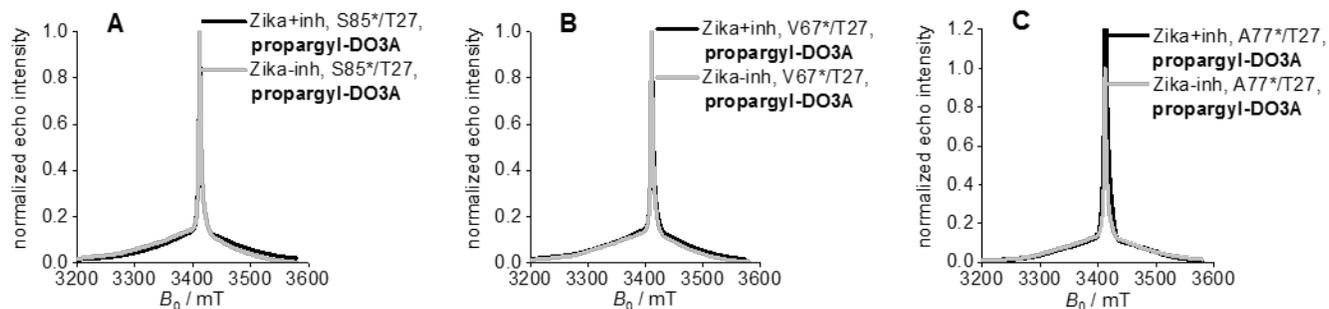


Figure S6. ED-EPR spectra of the Zika virus NS2B-NS3 protease mutants tagged with propargyl-DO3A measured at W-band at 10 K. Spectra recorded with and without the inhibitor cn-716 are shown in black and grey, respectively. (A) Mutant S85*AzF/T27AzF. (B) Mutant V67*AzF/T27AzF. (C) Mutant A77*AzF/T27AzF.

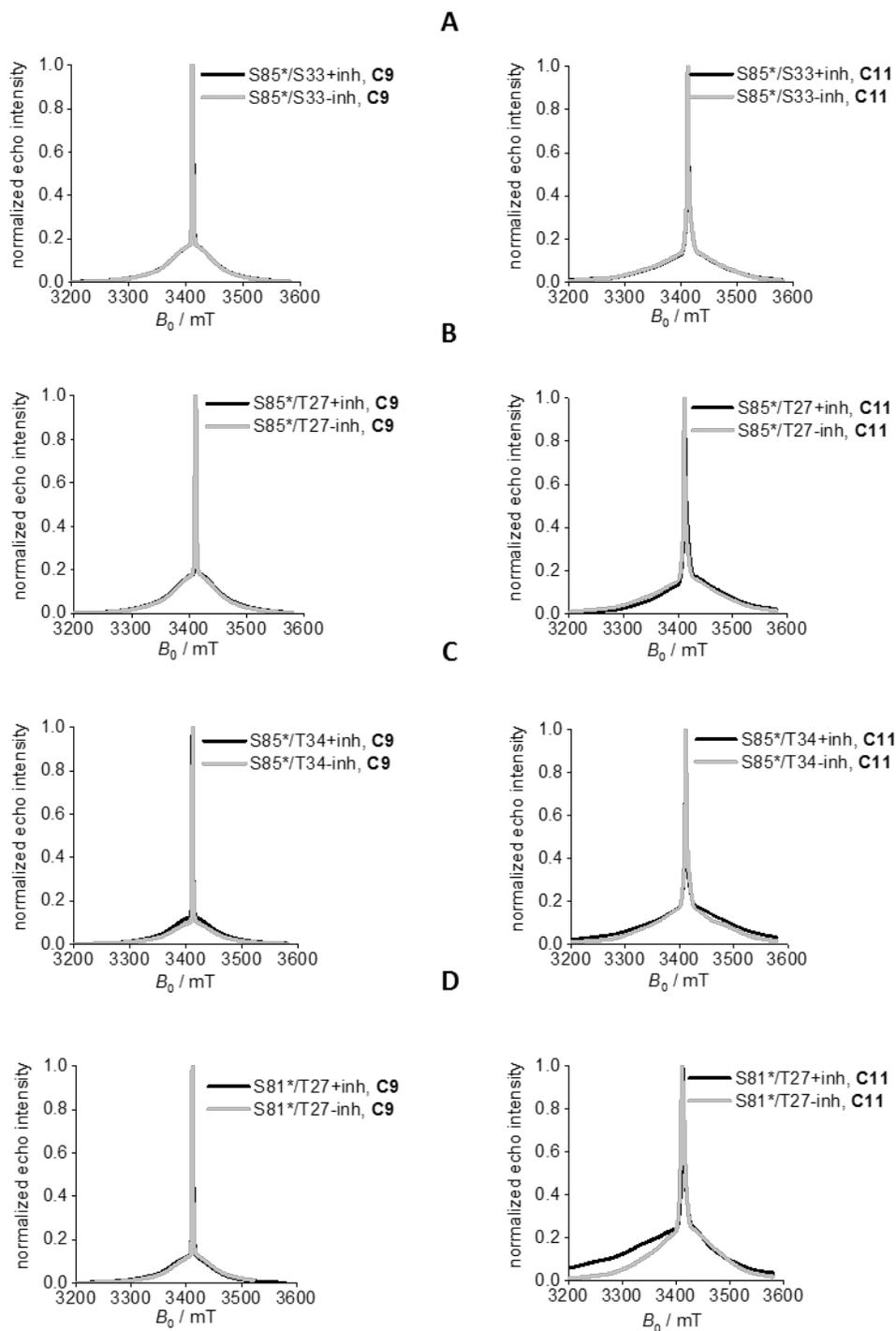


Figure S7. ED-EPR spectra of the Zika virus NS2B-NS3 protease mutants tagged with **C9** (left panels) or **C11** (right panels) measured at W-band at 10 K. Spectra recorded with and without the inhibitor cn-

716 are shown in black and grey, respectively. (A) Mutant S85*C/S33C. (B) Mutant S85*C/T27C. (C) Mutant S85*C/T34C. (D) Mutant S81*C/T27C.

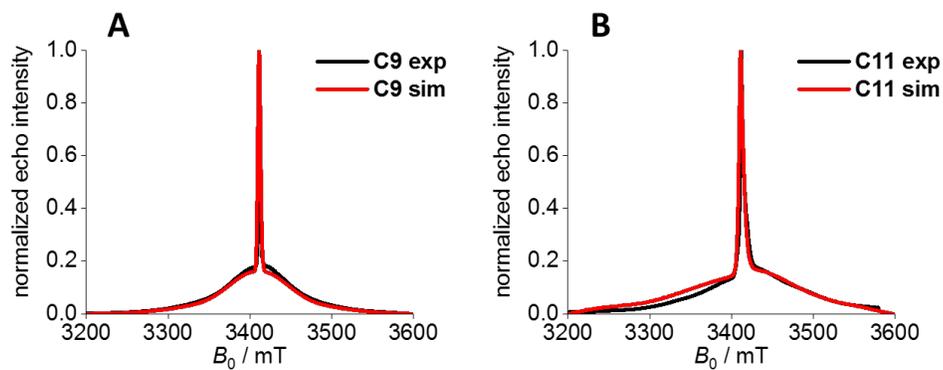


Figure S8. Simulations of the ED-EPR spectra of Zika virus NS2B-NS3 protease mutant S85*C/T27C mutant and comparison with experimental spectra. (A) Labelled with the C9 tag. (B) Labelled with the C11 tag.

Echo decay data for all Zika virus NS2B-NS3 protease samples

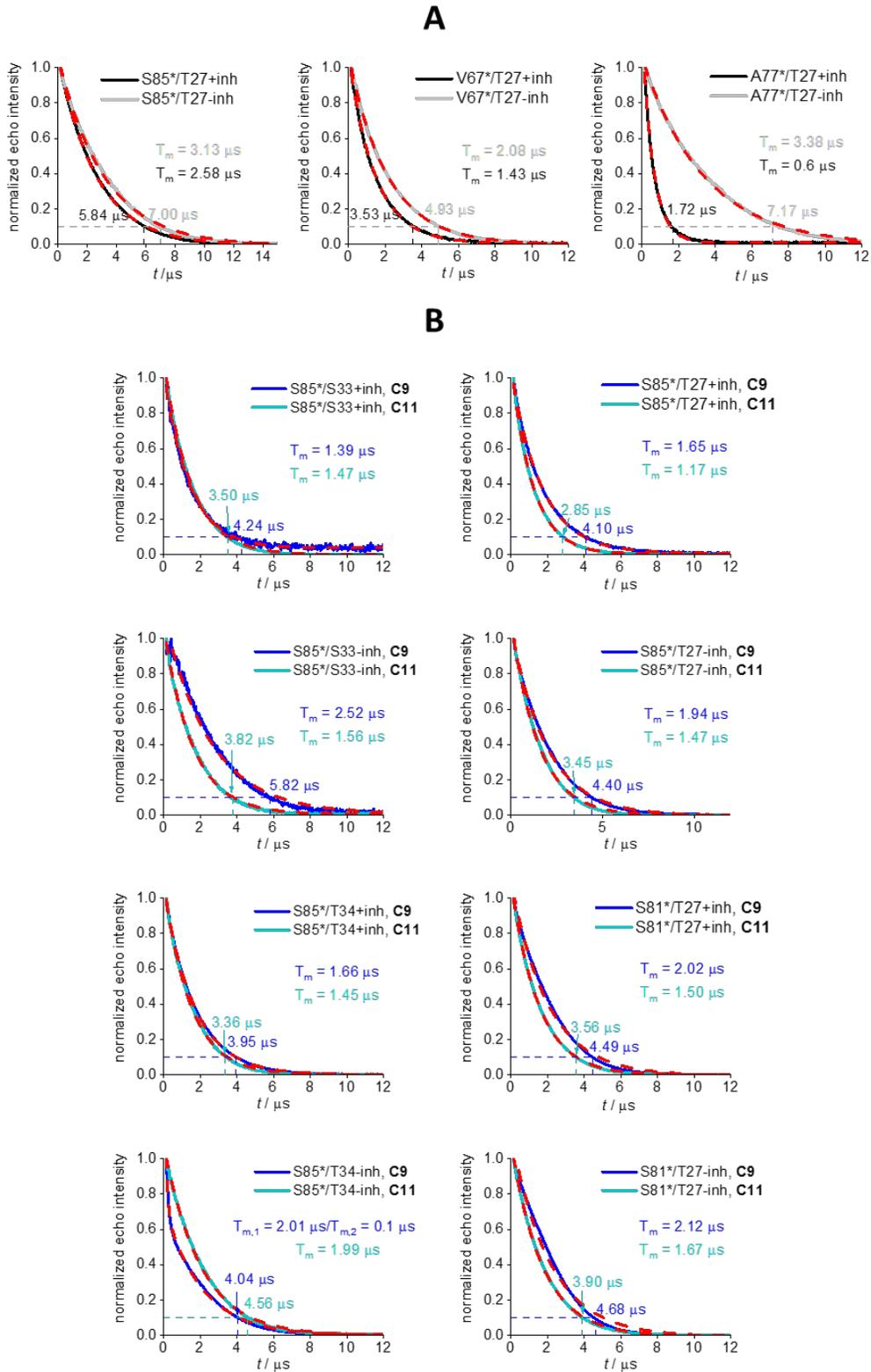
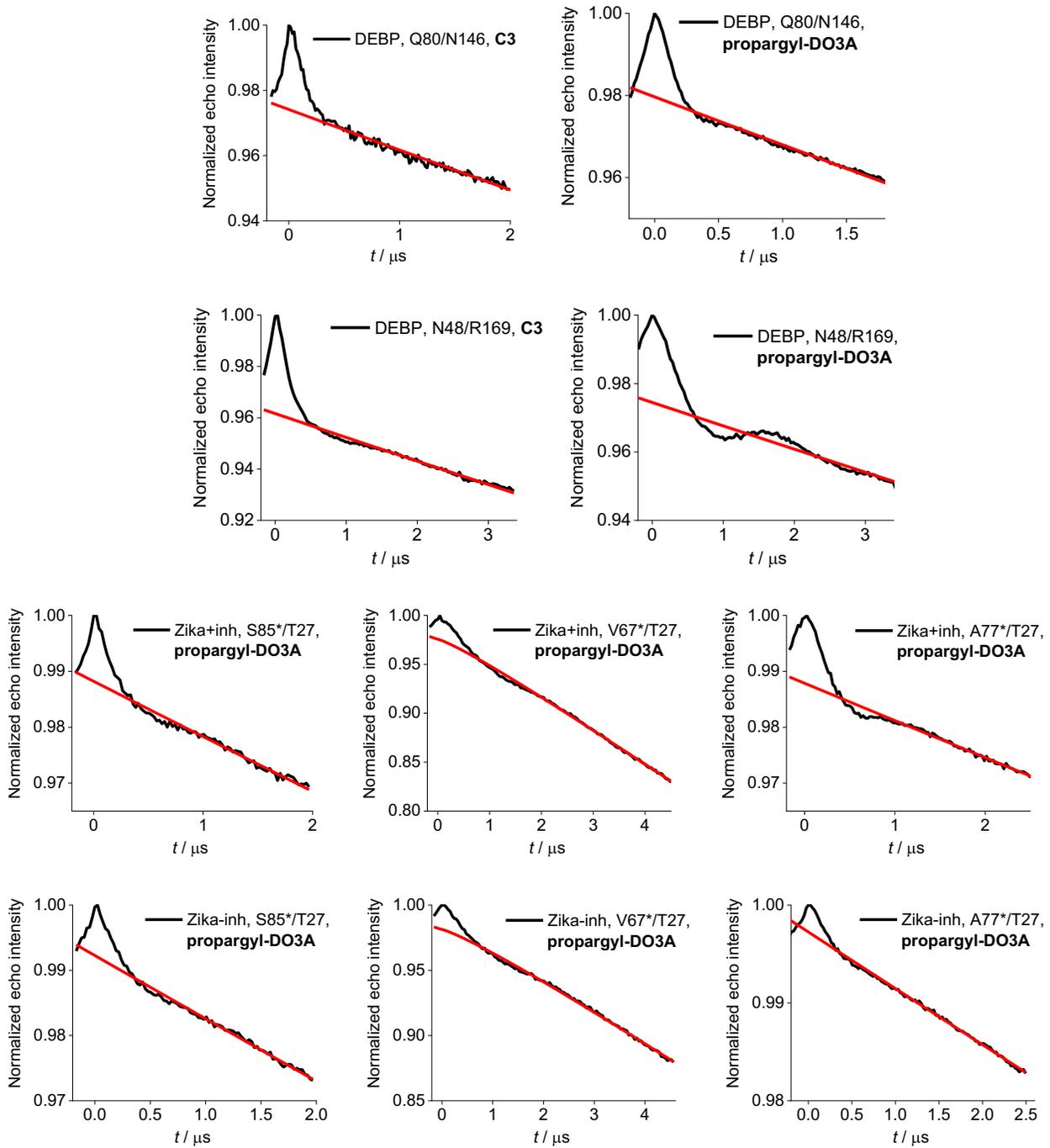
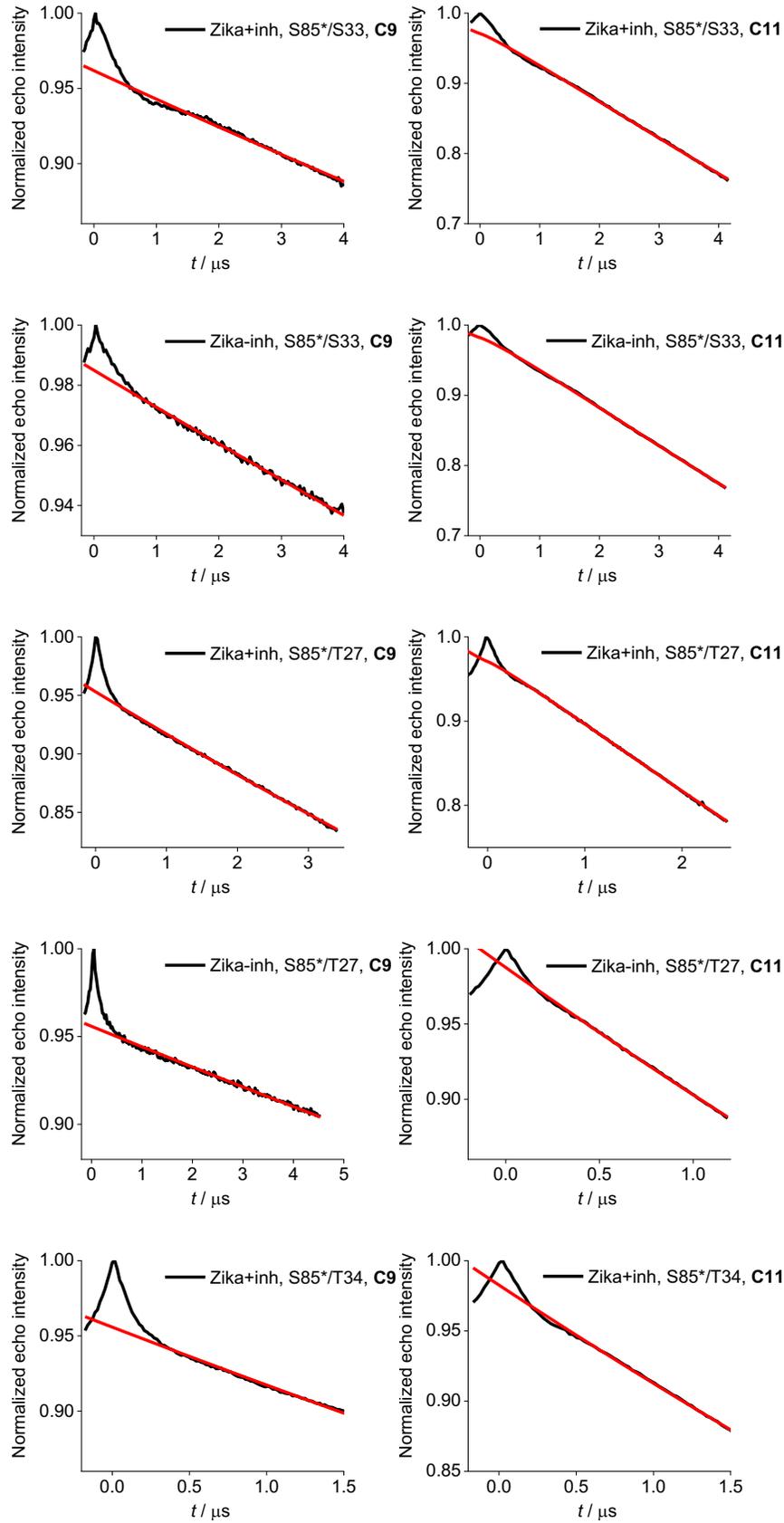


Figure S9. Echo decay data measured at W-band at 10 K at the maximum of the ED-EPR spectra of the Zika virus protease mutants. The T_m values are given, as well as the times at which the echo intensity had decayed to 10% of its initial intensity, according to the exponential decay obtained by fitting (monoexponential for all mutants except for the sample “S85*/T34-inh, **C9**”, where a biexponential decay function was used). The fits are shown by red dashed lines. The times of echo intensity decay to 10%, $\tau_{10\%}$, provide estimates of the duration of evolution times at which DEER experiments can be conducted. For the mutant “S81*/T27-inh, **C9**”, we could not obtain a good fit and the T_m value given here was obtained by fitting a monoexponential decay function. (A) Mutants (left to right): S85*AzF/T27AzF, V67*AzF/T27AzF, A77*AzF/T27AzF. Data recorded with and without the inhibitor cn-716 are shown in black and grey, respectively. (B) Double-cysteine mutants. The mutation sites and presence or absence of inhibitor are specified in each panel. Dark blue: data recorded with the **C9** tag. Light blue: data recorded with the **C11** tag.

Primary DEER data





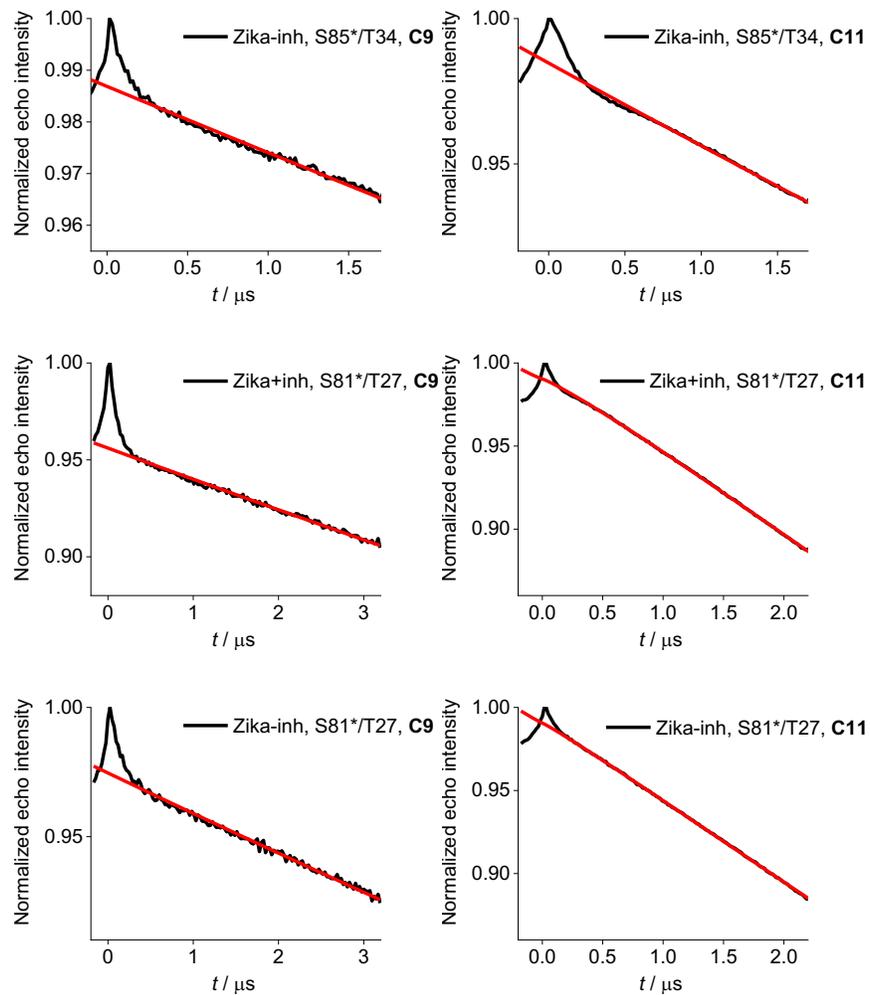
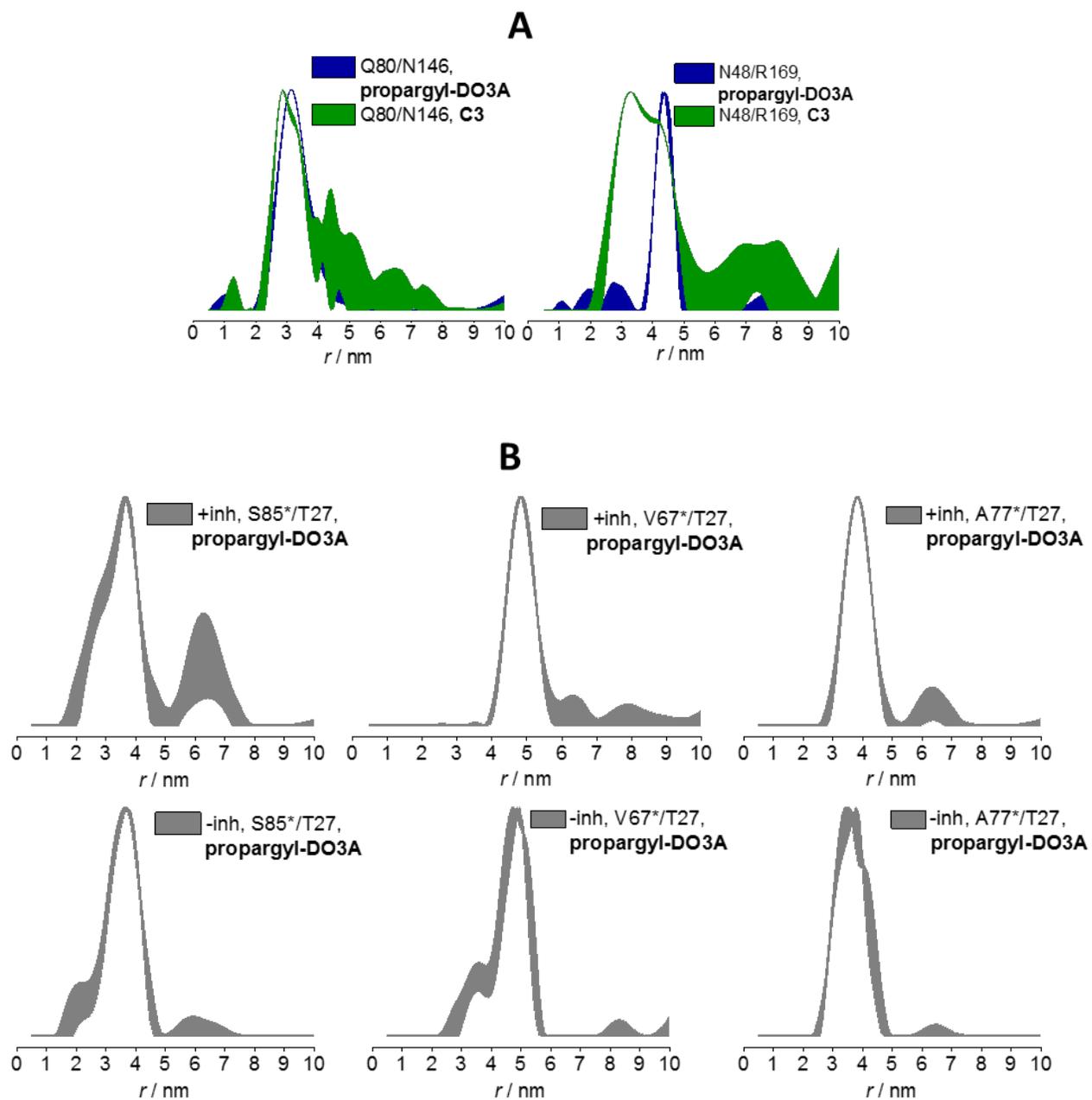


Figure S10. Primary DEER data (black) and fitted background decays (red). The protein, mutant, presence of inhibitor and tag are indicated in each panel.

Validated distance distributions



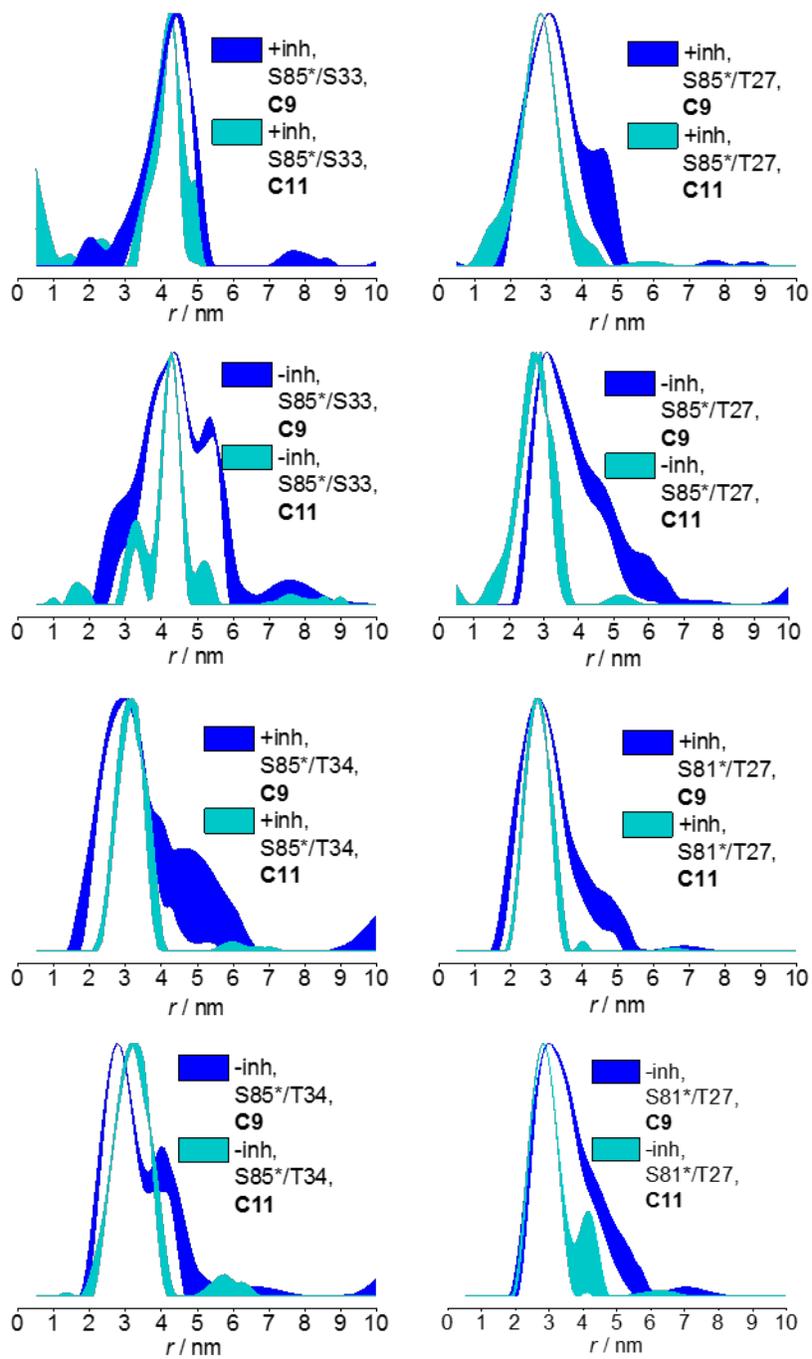


Figure S11. Distance distributions for all mutants including confidence intervals. The mutant, presence of inhibitor and tag are indicated in each panel. (A) DEBP mutants. (B) Zika virus NS2B-NS3 mutants.

DEER with dual-mode cavity

When performing Gd(III)–Gd(III) DEER distance measurements, one needs to consider the pseudo-secular terms of the dipolar Hamiltonian in order to reliably extract the distance distribution(s).⁵ Neglecting the pseudo-secular terms has been found to be reasonable for distances above 3.4 nm,⁶ but not necessarily for shorter Gd(III)–Gd(III) distances. In the latter case, processing of the time domain data using the ‘common practice’ DeerAnalysis software can cause artificial broadening of the distance distributions, as the software utilises a kernel function that assumes the weak coupling approximation to be valid. The effect is more pronounced for short Gd(III)–Gd(III) distances and small zero-field splittings (i.e. a narrow central EPR line corresponding to the $|-1/2 \rightarrow +1/2\rangle$ transition of Gd(III)). It has been shown that this artificial broadening can be overcome experimentally by using large probe-pump offsets that increase the contribution from higher order transitions. In the present work, this was achieved by performing the DEER experiment using a dual cavity probe,⁷ which allows the use of a large probe-pump offset (Figure S9). This enabled us to test whether the broad distance distribution, which we had measured for the S85*C/T27C mutant labelled with the **C9** tag and in the absence of inhibitor, originates from protein/tag flexibility or artificial broadening due to the neglect of pseudo-secular terms in the analysis of data recorded with the standard probe-pump offset of 100 MHz. Equally broad distance distributions were obtained in both setups, indicating intrinsic flexibility of the protein and/or tags.

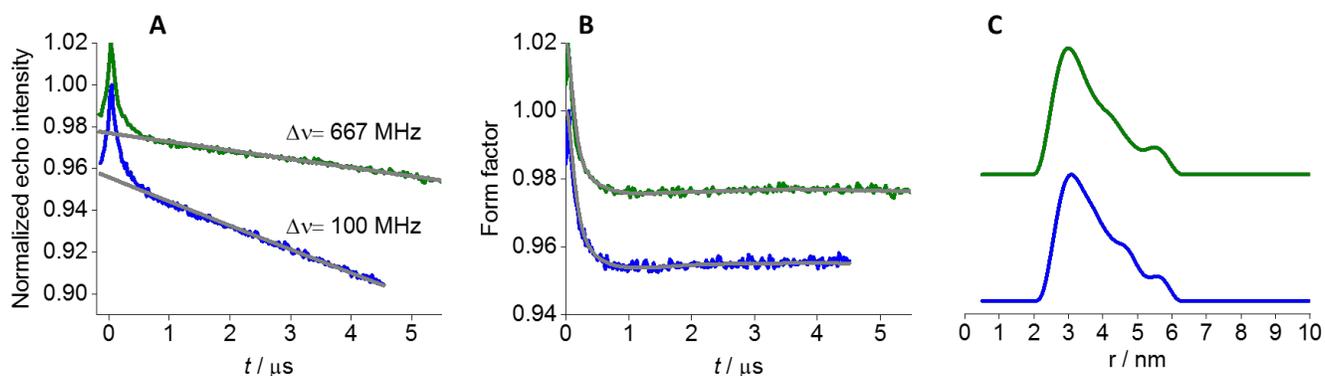


Figure S12. Comparison of the DEER traces of the Zika virus NS2B-NS3 protease mutant S85*C/T27C labeled with **C9** and without inhibitor recorded with two values of the $\Delta\nu$ parameter as noted in the figure. (A) Primary DEER data with the fitted background decay. (B) Corresponding form factor with the fitted data (in grey) obtained with the distance distributions shown in (C).

Modelling of DEER distance distributions

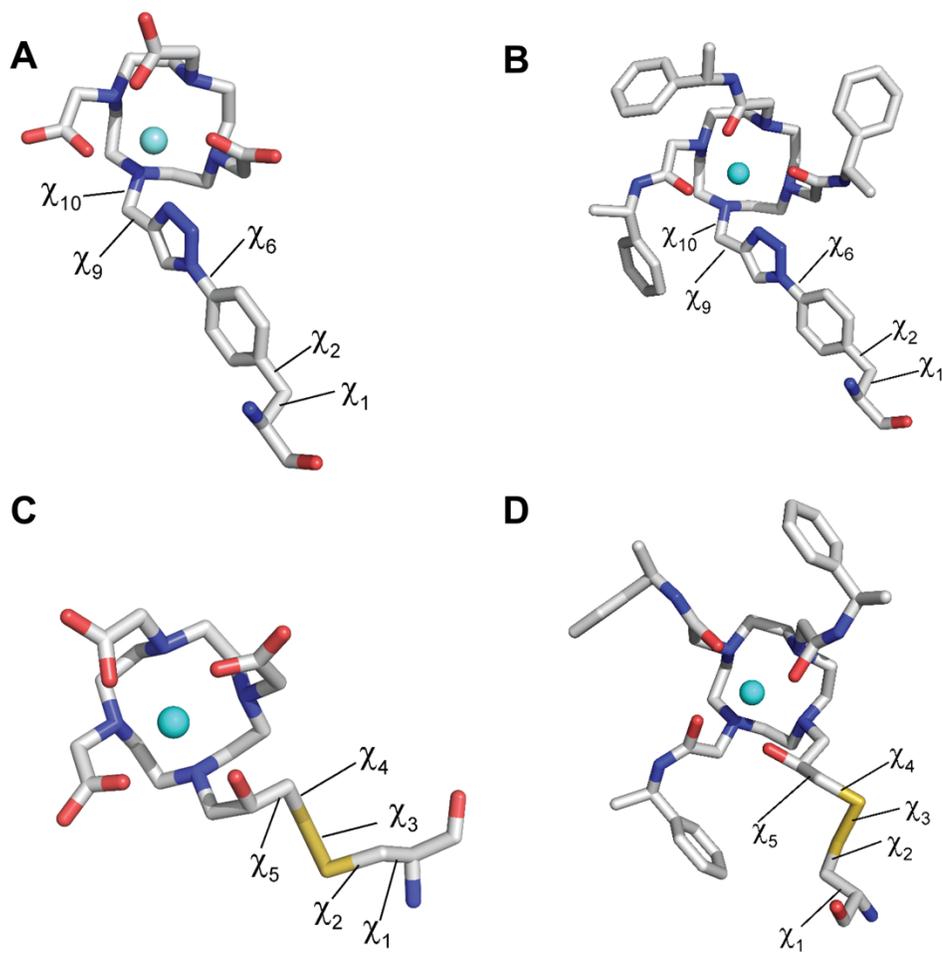


Figure S13. Dihedral angles varied to predict distance distributions with the program PyParaTools⁸ as described previously.⁹ The figure depicts the amino acid residues with tags attached. The rotamer libraries were established by varying the dihedral angles in random combinations, allowing the dihedral angles to vary either completely randomly or within ranges of $\pm 10^\circ$. Conformations generating steric clashes with the protein were excluded. (A) AzF residue with **propargyl-DO3A** tag. χ_1 and χ_2 angles were selected by using those conformations, which were identified by the mutation tool of the program PyMOL to have minimal steric clashes. For most of the different mutation sites in the protein, the χ_1 angles were centered about -175° , -80° , 60° and 175° , and the χ_2 angles about 20° , 85° and 100° . The χ_6 angle was centered about 0° or 180° , χ_9 was completely randomized and the χ_{10} angle was restricted to -60° , 60° and 180° . The triazole ring is known to coordinate the lanthanide ion, but the coordination is not very stable.¹⁰ (B) AzF residue with **C3** tag. Same dihedral angle variations as in (A). The pendant arms of the cyclen complex were modelled according to the crystal structure of the DOTA-tetraamide Gd^{3+} complex with 1-phenylethyl amine (DOTAMPh, CSD accession code EQOZUF¹¹). (C) Cysteine residue with **C11** tag. χ_1 : completely random. χ_2 : $-50^\circ - -70^\circ$, $50^\circ - 70^\circ$, $170^\circ - 190^\circ$. χ_3 : $-80^\circ - -100^\circ$, $80^\circ - 100^\circ$. χ_4 and χ_5 : same dihedral angle ranges as for χ_2 . (D) Cysteine residue with **C9** tag. Same dihedral angle variations as in (C).

References

- 1 M. D. Lee, C.-T. Loh, J. Shin, S. Chhabra, M. L. Dennis, G. Otting, J. D. Swarbrick and B. Graham, *Chem. Sci.*, 2015, **6**, 2614–2624.
- 2 S. Stoll and A. Schweiger, *J. Magn. Reson.*, 2006, **178**, 42–55.
- 3 A. M. Raitsimring, A. V. Astashkin, O. G. Poluektov and P. Caravan, *Appl. Magn. Reson.*, 2005, **28**, 281–295.
- 4 A. M. Raitsimring, A. V. Astashkin and P. Caravan, Highfrequency EPR and ENDOR characterization of MRI contrast agents, in *Biological Magnetic Resonance*, ed. L. Berliner and G. Hanson, Springer, New York, 2009, pp. 581–621.
- 5 N. Manukovsky, A. Feintuch, I. Kuprov and D. Goldfarb, *J. Chem. Phys.*, 2017, **147**, 044201.
- 6 A. Dalaloyan, M. Qi, S. Ruthstein, S. Vega, A. Godt, A. Feintuch and D. Goldfarb, *Phys. Chem. Chem. Phys.*, 2015, **17**, 18464–18476.
- 7 M. R. Cohen, V. Frydman, P. Milko, M.A. Iron, E.H. Abdelkader, M. D. Lee, J. D. Swarbrick, A. Raitsimring, G. Otting, B. Graham, A. Feintuch and D. Goldfarb, *Phys. Chem. Chem. Phys.*, 2016, **18**, 12847–12859.
- 8 M. Stanton-Cook, X.-C. Su, G. Otting and T. Huber, <http://comp-bio.anu.edu.au/mscook/PPT/>, accessed February 14, 2018.
- 9 A. Potapov, H. Yagi, T. Huber, S. Jergic, N. E. Dixon, G. Otting and D. Goldfarb, *J. Am. Chem. Soc.*, 2010, **132**, 9040–9048.
- 10 M. Tropiano, C. J. Record, E. Morris, H. S. Rai, C. Allain and S. Faulkner, *Organometallics*, 2012, **31**, 5673–5676.
- 11 D. Parker, H. Puschmann, A. S. Batsanov and K. Senanayake, *Inorg. Chem.*, 2003, **42**, 8646–8651.