Supporting information

Protein Conformation by EPR Spectroscopy Using Gadolinium Tags Clicked to Genetically Encoded *p*-Azido-L-Phenylalanine

Elwy H. Abdelkader, Akiva Feintuch, Xuejun Yao, Luke A. Adams, Luigi Aurelio, Bim Graham, Daniella Goldfarb, Gottfried Otting

Protocols for sample preparation

Reagents. *p*-azido-*L*-phenylalanine (AzF) was synthesized as described.[1] Total tRNA containing suppressor tRNA was prepared as described.[2] Other reagents were sourced and prepared as described previously.[3]

Protein synthesis. The wild-type gene (*ybeJ*, NP_415188.1) of the *E. coli* aspartate/glutamate -binding protein (DEBP) without the periplasmic leader sequence was cloned into the pETMCSIII vector[4] with an N-terminal His_6 tag followed by a tobacco etch virus (TEV) protease recognition site (Figure S1). DEBP mutants containing two AzF residues at selected positions were produced, using PCR to insert amber stop codons at the respective sites, followed by a second round of PCR to amplify the DNA with eight-nucleotide overhangs suitable for protein production by cell-free synthesis,[5] and using S30 extract depleted of the release factor RF1.[3,6] The protocol included purified *p*-cyano-*L*-phenylalanyl-tRNA-synthetase[7] to charge optimized suppressor-tRNA,[8] which was provided as a component of total tRNA prepared as described previously.[2]

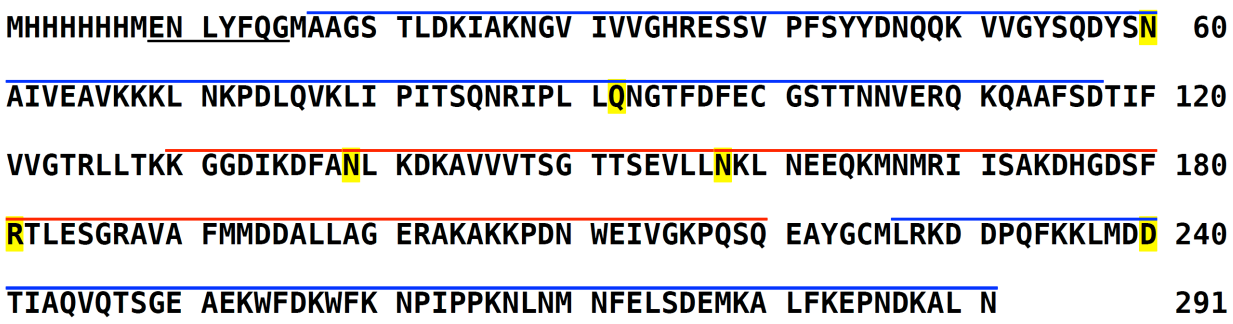

Figure S1. Amino acid sequence of the *E. coli* DEBP construct used. The TEV cleavage site is underlined. Residues of domain 1 are identified by a blue bar above the amino acid sequence and residues of domain 2 by a red bar. The locations of the amber stop mutations introduced for site-specific incorporation of AzF are highlighted in yellow.

15 DEBP double amber mutants were prepared, specifying the incorporation of AzF residues at the following sites: N48/Q80, N48/N127, N48/N146, N48/R169, N48/D228, Q80/N127, Q80/N146, Q80/R169, Q80/D228, N127/N146, N127/R169, N146/R169, N127/D228, N146/D228, R169/D228. The mutation primers are listed in Table S1. Following mutagenesis, double-stranded DNA with 8-nucleotide single-stranded overhangs was produced by PCR as described.[5] Each cell-free reaction was conducted at 30 °C in a dialysis system with 3 mL inner reaction mixture and 30 mL outer buffer. In addition to the usual reagents,[3] the inner buffer contained 0.5 mg/mL total tRNA containing optimized suppressor tRNA[8] and 50 μM purified *p*-cyano-*L*-phenylalanyl-tRNA synthetase.[7] 1 mM AzF was present in both inner and outer buffers.

Table S1. Nucleotide sequences of the mutation primers used in this study

Mutation site	Primer Sequence	
N48amber	Forward	5`-TCGCAGGATTACTCC**TAG**GCCATTGTTGAAGCA-3`
	Reverse	5`-TGCTTCAACAATGGC**CTA**GGAGTAATCCTGCGA-3`
Q80amber	Forward	5`-ATTCCACTGCTG**TAG**AACGGCACTTTCGATTTTG-3`
	Reverse	5`-GAAAGTGCCGTT**CTA**CAGCAGTGGAATACG-3`
N127amber	Forward	5`-AAAGATTTTGCC**TAG**CTGAAAGACAAAGCCGTAG-3`
	Reverse	5`-TTTGTCTTTCAG**CTA**GGCAAAATCTTTGATATCGCC-3`
N146amber	Forward	5`-CTGAAGTTTTGCTC**TAG**AAACTGAATGAAGAGCAAAAAATG-3`
	Reverse	5`-TTCATTCAGTTT**CTA**GAGCAAAACTTCAGAGGTAGTGCC-3`
R169amber	Forward	5`-GACTCTTTC**TAG**ACCCTGGAAAGCGGTCG-3`
	Reverse	5`-GCTTTCCAGGGT**CTA**GAAAGAGTCACCGTGATC-3`
D228amber	Forward	5`-AAAAAGCTGATGGAT**TAG**ACCATCGCTCAGGTG-3`
	Reverse	5`-CACCTGAGCGATGGT**CTA**ATCCATCAGCTTTTT-3`

Protein purification. The proteins were purified using a 1 mL Ni-NTA column (GE Healthcare, USA) according to the manufacturer's protocol. Afterwards, the N-terminal His_6 tag was removed by incubation with TEV protease[9] at 4 °C for 16 h in a buffer of 25 mM Tris-HCl, 500 mM sodium chloride, 2 mM 2-mercaptoethanol, pH 8. The cleaved His_6 tag and the TEV protease were separated from DEBP by running the mixture again over a Ni-NTA column. To remove bound glutamate, the protein samples were denatured by the addition of guanidine hydrochloride to a final concentration of 6 M. The solution was left at room temperature for 2 h. Refolding was achieved by dialyzing against click buffer (50 mM sodium phosphate, pH 8, 150 mM sodium chloride) at 4°C, following a published protocol.[10] Finally, the protein samples were concentrated using an Amicon ultrafiltration centrifugal tube with a molecular weight cutoff of 10 kDa. The average yield was about 1 mg of purified protein per mL cell-free reaction mixture (inner buffer). The disulfide bond between Cys88 and Cys213 formed spontaneously during the refolding reaction. Its presence was confirmed by the lack of color formation upon titration with DTNB after denaturation of the protein with 6 M guanidinium hydrochloride. SDS-PAGE analysis of the purified protein samples showed no evidence of truncated product (Figure S2).

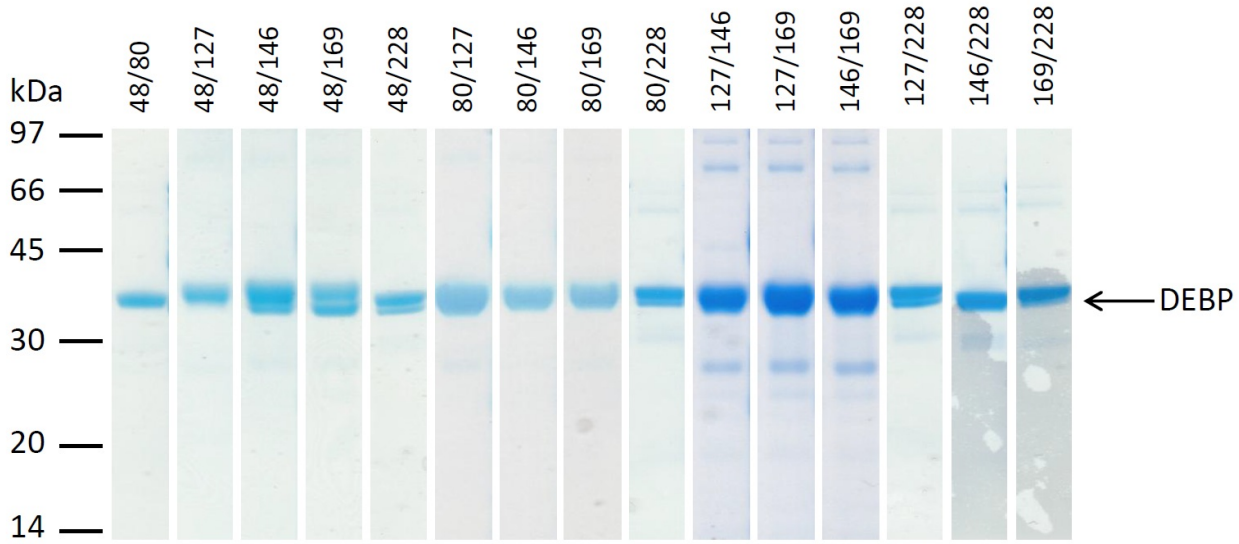

Figure S2. SDS-PAGE gels of the 15 purified DEBP double mutants, showing the absence of truncated product. Staining was with Coomassie blue. All mutants appeared homogeneous when analyzed by mass spectrometry, including those showing two bands in close proximity.

Ligation with C3-Gd^{3+} tag and preparation of EPR samples. For each protein construct, the solution of protein with AzF in "click buffer" (50 mM sodium phosphate, 150 mM sodium chloride, pH 8) was added to the solution of the C3-Gd^{3+} tag, followed by addition of a premixed solution of $CuSO_4$ and BTTAA (copper(I)-binding ligand; BTTAA = 2-[4-({bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino}methyl)-1H-1,2,3-triazol-1-yl]acetic acid),[11] aminoguanidine, glycerol, and finally ascorbic acid to yield a total reaction volume of 0.5 mL.[12] The final concentrations were 0.05 mM protein, 0.5 mM tag, 0.3 mM $CuSO_4$, 1.5 mM BTTAA, 5 mM aminoguanidine, 0.5 mM glycerol, and 5 mM ascorbic acid. The reaction was performed in a glove box under N_2 atmosphere at room temperature with stirring for 16 h. Afterwards, the reactions were terminated by the addition of EDTA to a final concentration of 2 mM, followed by stirring for 30 min in air. The solution was concentrated and exchanged to EPR buffer (20 mM Tris-HCl in D_2O, pH 7.5, uncorrected pH meter reading) using an Amicon ultrafiltration centrifugal tube (10 kDa molecular weight cutoff), and perdeuterated glycerol was added to a final concentration of 20% (v/v) to reach a final protein concentration of 0.1 mM. To form the complex with glutamate, potassium glutamate in EPR buffer was added in 10-fold excess.

Optimization of the click reaction

Assessing chemical reduction of AzF. Analysis of the mass spectra of the purified proteins showed the presence of a minor product (MW - 26 Da, about 15% of the main product in the double mutants), suggesting reduction of azide to amino groups. Dithiothreitol (DTT) is known to reduce aryl azides[13] and 1.7 mM DTT is present in our standard cell-free reaction conditions. Following shortening of the cell-free reaction time from 16 h to 6 h, no evidence for reduction of AzF was found in the mass spectra of purified protein (Figure S3).

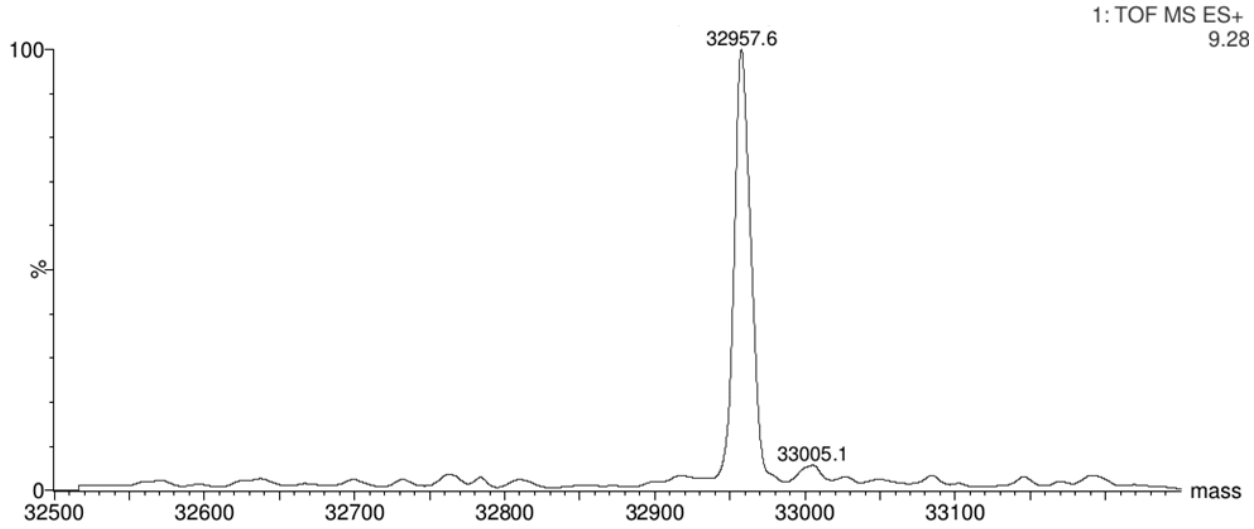

Figure S3. ESI mass spectrum of DEBP N48AzF/Q80AzF after 6 h of cell-free reaction (expected MW 32960.0, observed MW 32957.6).

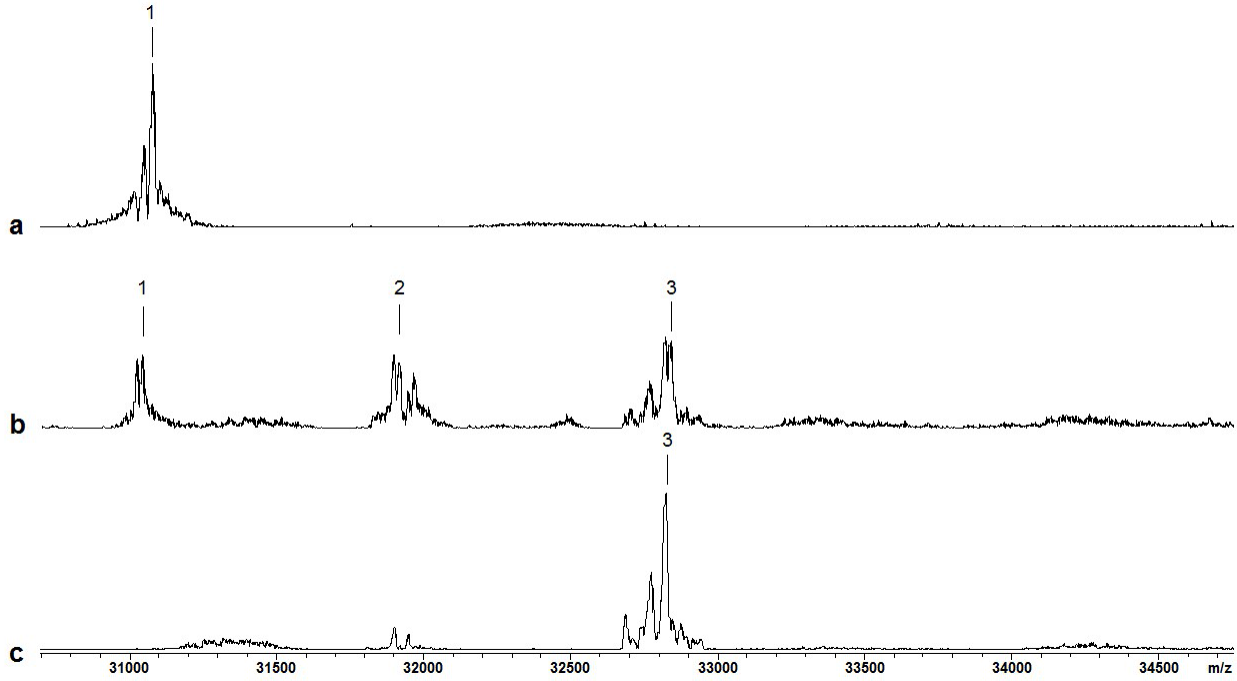

Figure S4. ESI mass spectra of DEBP Q80AzF/N127AzF following click reaction with the C3-Gd^{3+} tag in the presence of different amounts of sodium chloride. Peaks 1–3 correspond to DEBP without tag, with one tag, and with two tags, respectively. (1: expected MW 31080.0, observed MW 31077.5; 2: expected MW 31917.1, observed MW 31916.4; 3: expected MW 32823.16 (+3 Na^+), observed MW 32822.3) (a) Protein prior to the click reaction. (b) Protein obtained following click reaction without added sodium chloride. (c) Product obtained in the presence of 150 mM sodium chloride.

EPR spectroscopy

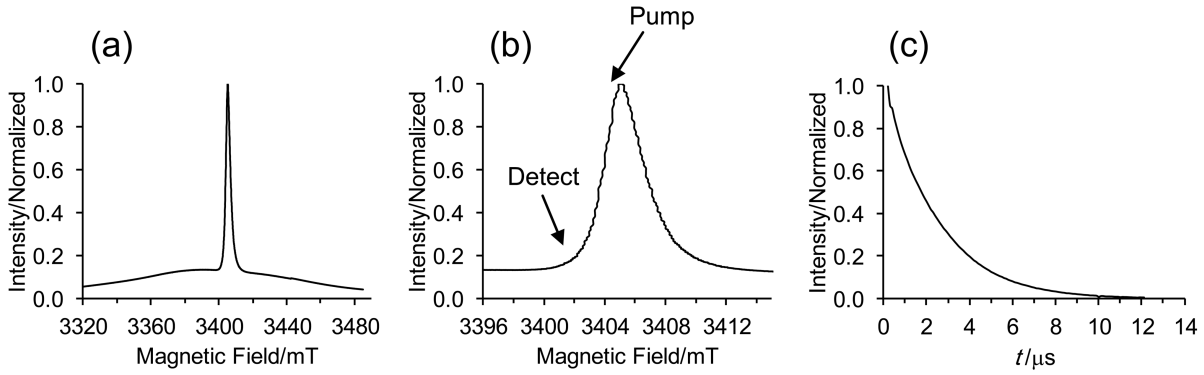

Figure S5. EPR characteristics of the AzF-C3-Gd^{3+} tag. (a) Representative echo-detected EPR spectrum of the DEBP mutant Q80/R169 at 94.9 GHz. The narrow line is due to the |-1/2>→|1/2> transition and the broad background is due to all other transitions. The width of the central narrow line (about 3 mT) is somewhat larger than for the chemically similar C1 (about 1.5 mT) or DOTA tags.[14,15] (b) Central part of the EPR spectrum, indicating the positioning of the DEER pulses. (c) Measurement of phase memory time, T_m, performed at the position of the detection pulse.

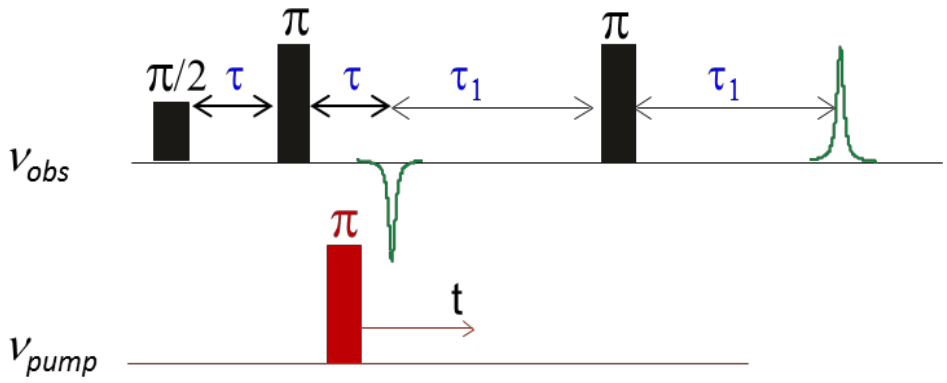

Figure S6. The four pulse DEER pulse sequence.[16] In the DEER experiment the echo intensity is measured as a function of the time *t*, while all other time intervals are kept constant.

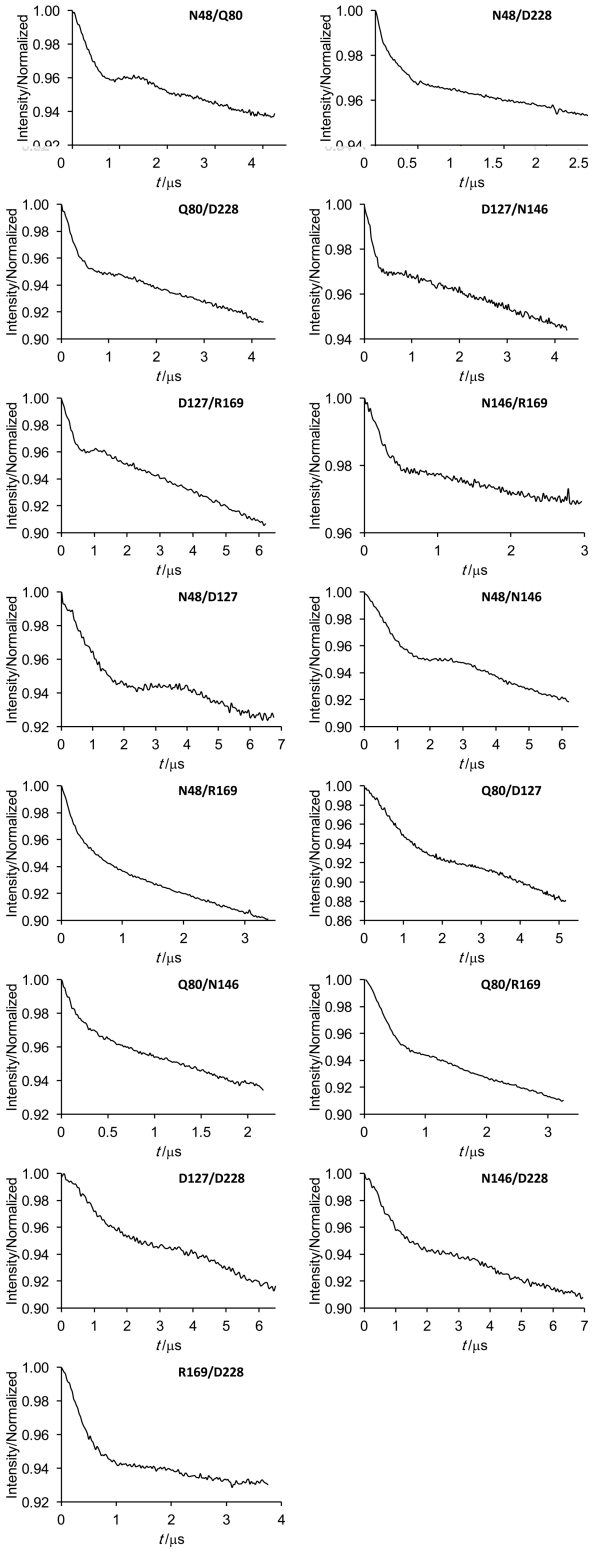

Figure S7. Raw data from DEER measurements for 15 different double AzF mutants of DEBP ligated with C3-Gd^{3+} tags. For all samples an exponential background correction was used, but because of the low modulation depth, the actual correction corresponded to a linear function.

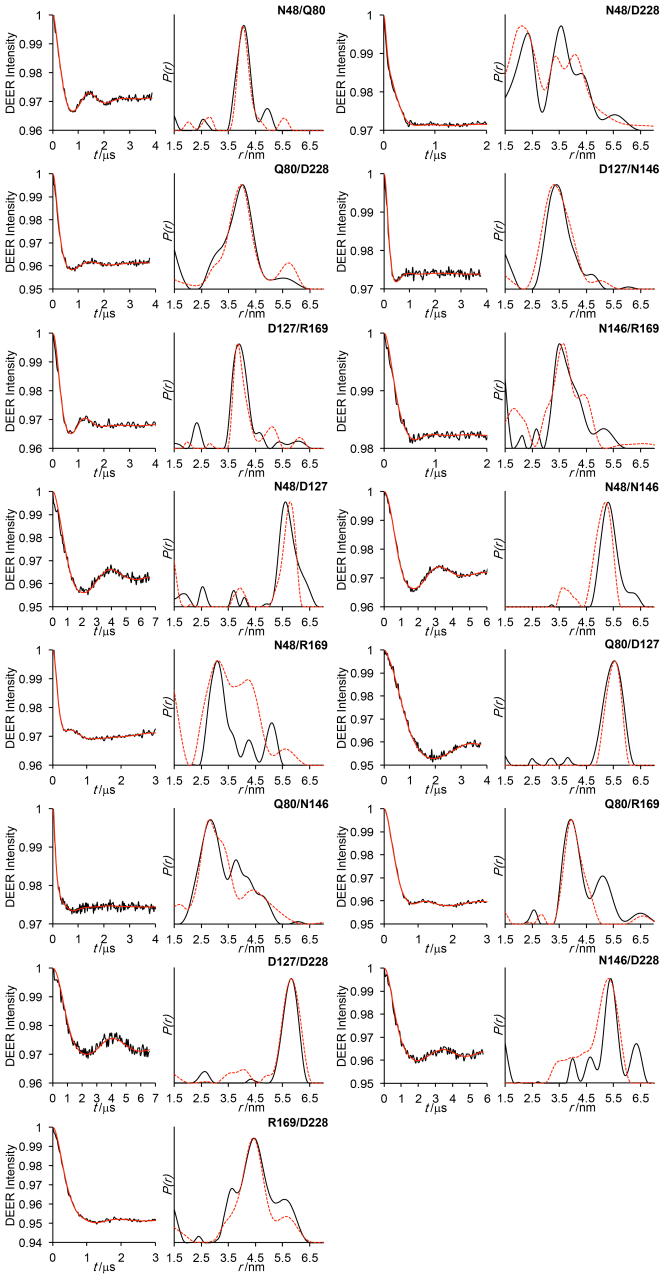

Figure S8. Distributions of inter-gadolinium distances determined by Tikhonov regularization from DEER measurements of double AzF mutants of DEBP ligated with C3-Gd^{3+} tags after dialysis under denaturing conditions to remove bound glutamate. The experimental conditions and processing parameters were the same as for the closed conformation in the presence of glutamate (Fig. 2). Left panels: Normalized DEER traces after background subtraction, showing the fits obtained by Tikhonov regularization in red. Right panels: corresponding distance distributions obtained by the Tikhonov regularization. For comparison, red dotted lines trace the distance distributions measured in the presence of glutamate (Fig. 2). The top three rows show the intra-domain distance measurements.

Tag conformation

The AzF-C3 residue has several rotatable bonds. The coordinates of the C3 tag were modelled using the enantiomeric mirror image of the crystal coordinates of the Gd^{3+} complex of the DOTA tetramide ligand featuring 1-phenylethyl groups, DOTAMPh (CSD accession code EQOZUF[17]). The dihedral angles between the triazole and cyclen rings were taken from the crystal coordinates of a cyclen compound with a methylene triazole pendant (CSD accession code ACOHIL[18]). In this structure, the Gd^{3+} ion is almost in the same plane as the atoms of the triazole ring. The dihedral angle χ_6 between the triazole and phenylene rings (Figure 1 of the main text) was assumed to be either 0 or 180 degrees to maintain a conjugated double-bond system. In the Cambridge Structure Database, all compounds with this structural motif contain additional substituents in the ortho positions that cause loss of co-planarity unless both rings simultaneously coordinate a common metal ion. Assuming $\chi_6 = 130^o$ for all AzF-C3 residues except for residue 169 made the correlation between experimental and predicted distances only slightly worse compared with the assumption of $\chi_6 = 180^o$.

Table S2. Experimental and modelled Gd^{3+}-Gd^{3+} distances

Mutant[a]	$d_{experimental}$/nm[b]	$d_{predicted}$/nm[c]	$d_{predicted}$/nm[d]
N48/Q80	4.07	3.96	3.68
N48/D228	2.11	2.06	2.00
Q80/D228	3.97	3.85	3.51
D127/N146	3.31	3.32	3.09
D127/R169	3.86	3.71	3.59
N146/R169	3.64	3.59	3.46
N48/D127	5.82	5.77	5.50
N48/N146	5.19	5.37	5.34
N48/R169	3.00	3.09	2.88
Q80/D127	5.55	5.52	5.09
Q80/N146	2.79	2.98	3.38
Q80/R169	3.93	4.04	4.02
D127/D228	5.82	5.82	5.46
N146/D228	5.34	5.44	5.66
R169/D228	4.42	4.36	4.12

[a] The first six mutants report intra-domain distances.

[b] The experimental distances correspond to the peak maxima of the distance distributions in Figure 2.

[c] Prediction for all χ_6 angles set to 180°.

[d] Prediction for all χ_6 angles set to 0°.

References

1 R. Seyer, A. Aumelas, M. Tence, J. Marie, J.-C. Bonnafus, S. Jard and B. Castro, *Int. J. Peptide Protein Res*., 1989, **34**, 234–245.

2 K. Ozawa, K. V. Loscha, K. V. Kuppan, C. T. Loh, N. E. Dixon and G. Otting, *Biochem. Biophys. Res. Commun*., 2012, **418**, 652–656.

3 M. Apponyi, K. Ozawa, N. E. Dixon and G. Otting, *Methods in Molecular Biology*, 2008, **426**, Structural proteomics: high-throughput methods; B. Kobe, M. Guss and T. Huber Eds., Humana Press, Totowa, USA, pp. 257–268.

4 C. Neylon, S. E. Brown, A. V. Kralicek, C. S. Miles, C. A. Love and N. E. Dixon, Biochemistry, 2000**, 39**, 11989–11999.

5 P. S. C. Wu, K. Ozawa, S. P. Lim, S. G. Vasudevan, N. E. Dixon and G. Otting, *Angew. Chemie Int. Ed.*, 2007**, 46**, 3356–3358.

6 K. V. Loscha, A. J. Herlt, R. Qi, T. Huber, K. Ozawa and G. Otting, *Angew. Chem. Int. Ed.*, 2012**, 51**, 2243–2246.

7 D. D. Young, T. S. Young, M. Jahnz, I. Ahmad, G. Spraggon and P. G. Schultz, Biochemistry, 2011, **50**, 1894–1900.

8 T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *J. Mol. Biol.*, 2010, **395**, 361–374.

9 L. D. Cabrita, D. Gilis, A. L. Robertson, Y. Dehouck, M. Rooman and S. P. Bottomley, *Protein Sci.*, 2007, **16**, 2360–2367.

10 Y. Hu, C.-P. Fan, G. Fu, D. Zhu, Q. Jin and D.-C. Wang, *J. Mol. Biol.*, 2008, **382**, 99–111.

11 C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. Soriano del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu and P. Wu, *Angew. Chemie Int. Ed.*, 2011, **50**, 8051–8056.

12 C. T. Loh, K. Ozawa, K. L. Tuck, N. Barlow, T. Huber, G. Otting and B. Graham, *Bioconjugate Chem.*, 2013, **24**, 260–268.

13 J. V. Staros, H. Bayley, D. N. Standring and J. R. Knowles, *Biochem. Biophys. Res. Commun.***,** 1978, **80**, 568–572.

14 H. Yagi, D. Banerjee, B. Graham, T. Huber, D. Goldfarb and G. Otting, *J. Am. Chem. Soc*., 2011, **133**, 10418–10421.

15 E. Matalon, T. Huber, G. Hagelueken, B. Graham, V. Frydman, A. Feintuch, G. Otting and D. Goldfarb, *Angew. Chemie Int. Ed*., 2013, **52**, 11831–11834.

16 M. Pannier, S. Veit, A. Godt, G. Jeschke and H. W. Spiess, *J. Magn. Reson.*, 2000, **142**, 331-340.

17 D. Parker, H. Puschmann, A. S. Batsanov and K. Senanayake, *Inorg. Chem*., 2003, **42**, 8646–8651.

18 H. Ito, H. Tsukube and S. Shinoda, *Chem. Commun*., 2012, **48**, 10954–10956.