Supporting Information

Choy Theng Loh, Kiyoshi Ozawa, Kellie L. Tuck, Nicholas Barlow, Thomas Huber, Gottfried Otting, and Bim Graham

Lanthanide tags for site-specific ligation to an unnatural amino acid and generation of pseudocontact shifts in proteins

Contents

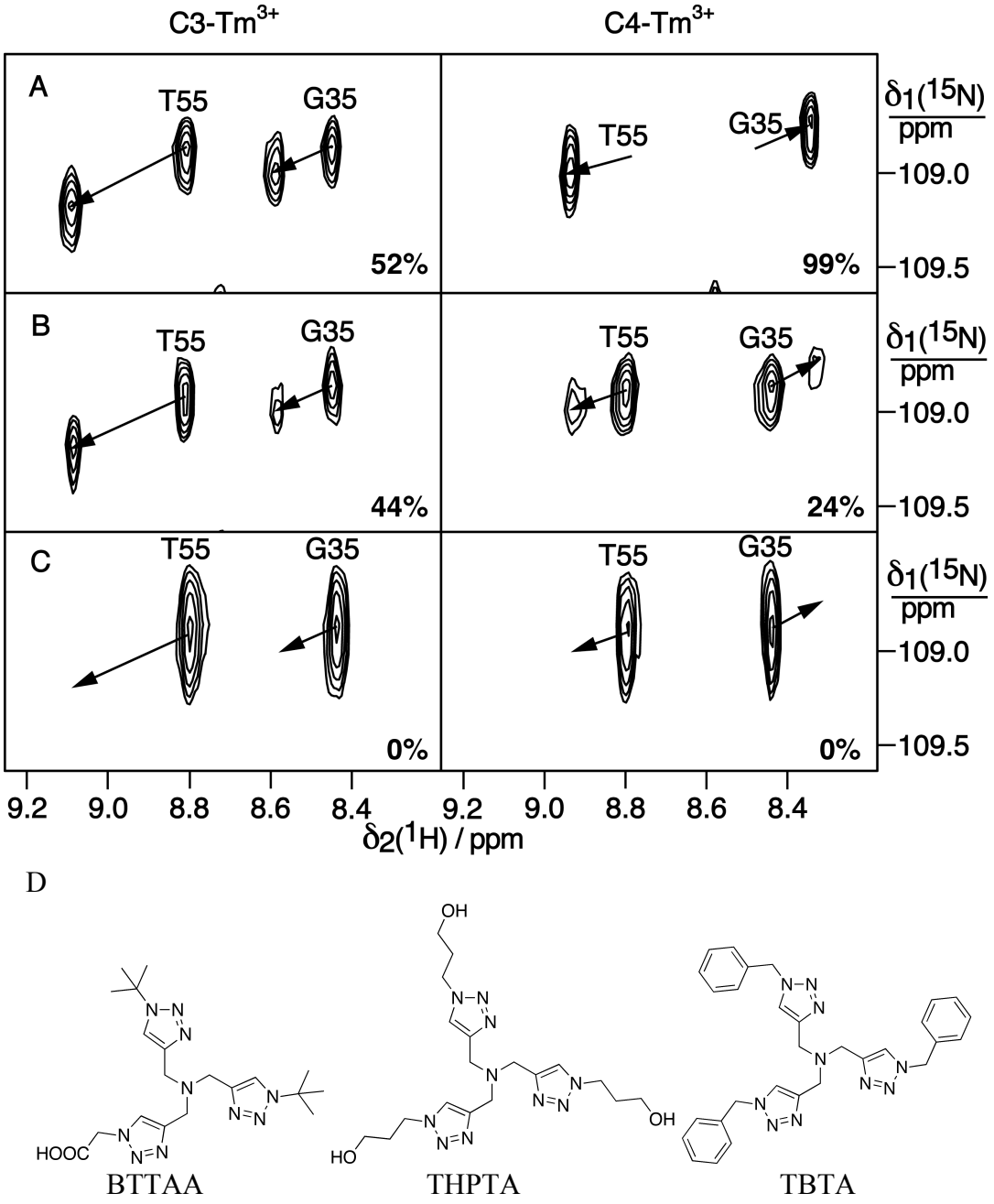

Figure S1. Comparison of the ligation yields obtained by click chemistry using different copper-complexing ligands. (A)-(C) Selected spectral region of ^{15}N-HSQC spectra recorded of 0.1 mM solutions of uniformly ^{15}N-labelled ubiquitin Glu18AzF (without His_6 tag) ligated with C3-Tm^{3+} (left panel) and C4-Tm^{3+} (right panel). The spectra were recorded at 25 °C on a 600 MHz NMR spectrometer, using 50 mM HEPES buffer, pH 7.0. The spectra were plotted using the same contour levels. The pseudocontact shifts of Gly35 and Thr55 are identified by arrows pointing from the cross-peaks of unligated protein to the corresponding cross-peaks of the ligated protein. The click reactions were performed using the same conditions for all samples (see Experimental Section), except for the presence of (A) 1 mM BTTAA, (B) 1 mM THPTA, or (C) 1 mM TBTA. The reaction yields are indicated in each panel. (D) Chemical structure of the BTTAA, THPTA, and TBTA ligands.

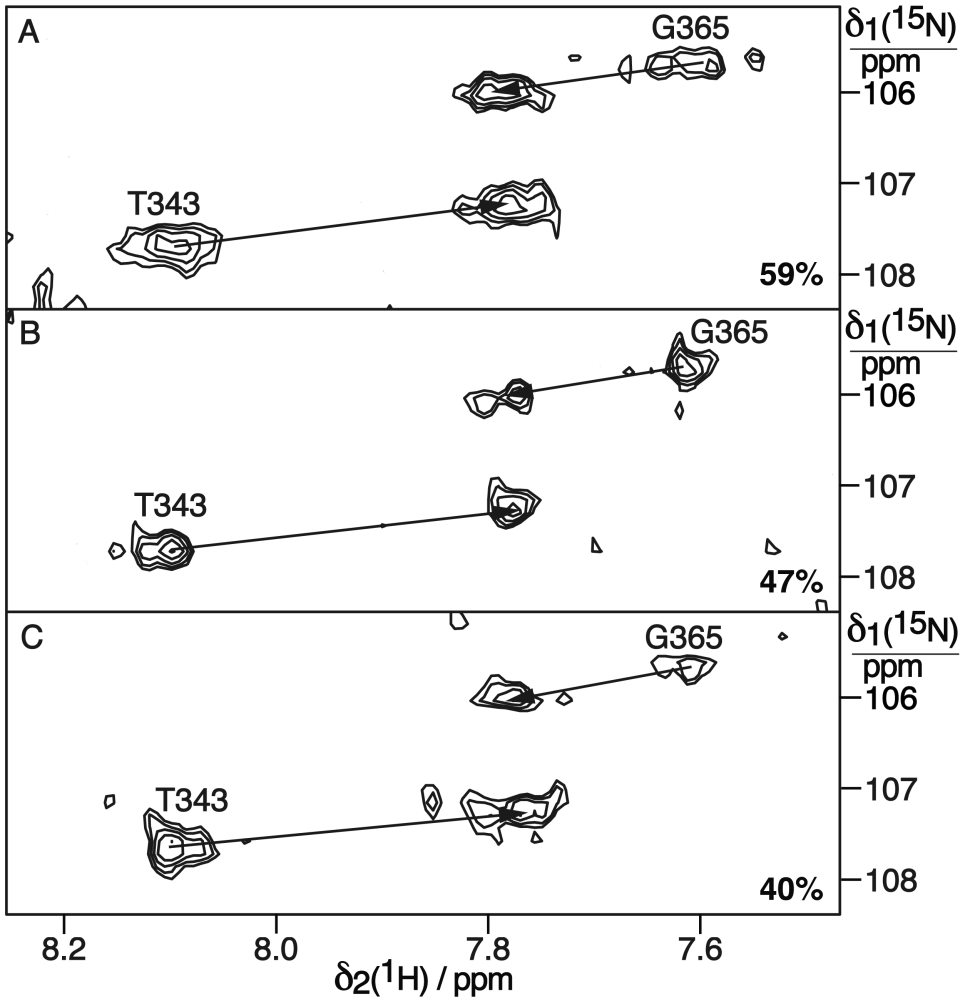

Figure S2. Effect of aminoguanidine and glycerol on the click ligation yields of the Lys350AzF mutant of uniformly ^{15}N-labelled intracellular domain of the rat p75 neurotrophin receptor with the C3-Tm^{3+} tag. ^{15}N-HSQC spectra were recorded at 25 °C of about 0.1 mM protein solutions in 20 mM MES buffer (pH 6.5) with 1 mM DTT, using a 600 MHz NMR spectrometer. All spectra are plotted with the same contour levels. The pseudocontact shifts of the cross-peaks of Thr343 and Gly365 are indicated by arrows pointing from the cross-peaks of the unligated protein to the corresponding peaks of the ligated protein. The percentage of successfully ligated protein as reflected by the relative cross-peak intensities is reported in each panel. Standard click reaction conditions were used (0.05 mM protein, 0.5 mM tag, 0.2 mM $CuSO_4$, 1 mM BTTAA, 5 mM sodium ascorbate, 50 mM HEPES buffer, pH 7.5), except for the addition of (A) 5 mM aminoguanidine and 0.5 mM glycerol, (B) 5 mM aminoguanidine only, and (C) no additives beyond the standard reagents.

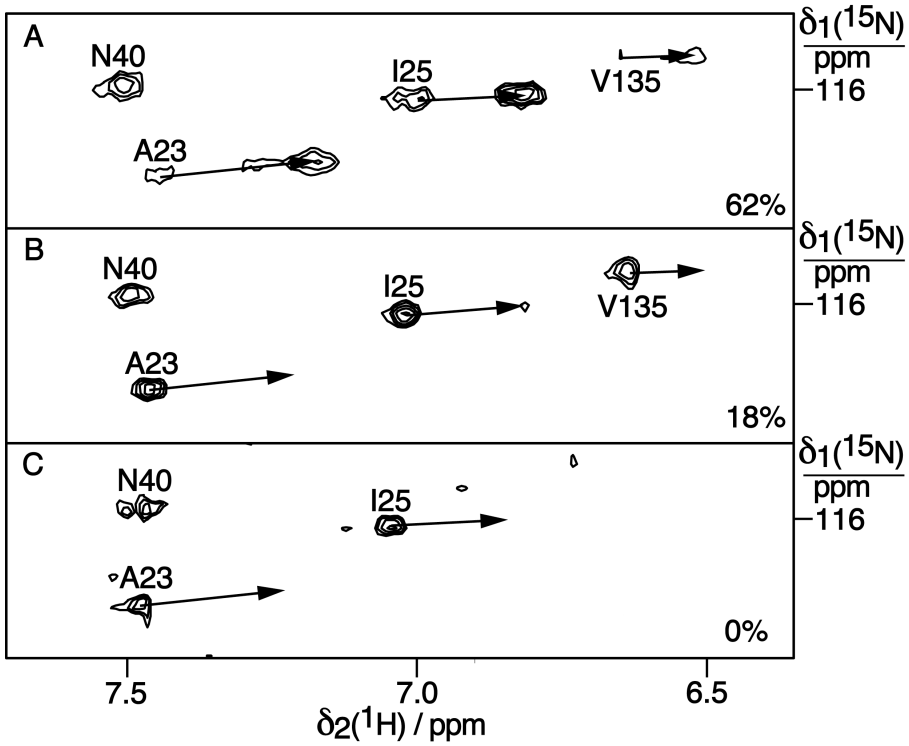

Figure S3. Same as Figure S2, except for the Ser99AzF mutant of uniformly ^{15}N-labelled *S. aureus* sortase A with the C3-Tm^{3+} tag. The pseudocontact shifts of Ala23, Ile25, Asn40, and Val135 are indicated by arrows. The percentage of successfully ligated protein as reflected by the relative cross-peak intensities is reported in each panel. Standard click reaction conditions were used, except for the addition of (A) 5 mM aminoguanidine and 0.5 mM glycerol, (B) 5 mM aminoguanidine only, and (C) no additives beyond the standard reagents.

Table S1. Pseudocontact shifts measured for backbone amide protons of ubiquitin Glu18AzF linked by cycloaddition with C3 and C4 tags, and of ubiquitin Thr66AzF with the C3 tag. Tags were loaded with Tb^{3+} and Tm^{3+}. [a]

		Glu18AzF				Thr66AzF	
		C3		C4		C3	
		Tb^{3+}	Tm^{3+}	Tb^{3+}	Tm^{3+}	Tb^{3+}	Tm^{3+}
2	Gln	0.99	-0.56	-0.45	0.44		
3	Ile	-0.09	0.05	-0.18	0.22		
4	Phe	-0.08	0.06	-0.12	0.13	-0.69	0.42
5	Val	-0.18	0.12	-0.05	0.05		-0.76
6	Lys	-0.15	0.10	-0.04	0.04		
7	Thr	-0.10	0.08	-0.01	0.01		-1.07
8	Leu	-0.11	0.07				-0.50
10	Gly	-0.06	0.05	-0.01			
11	Lys	-0.06	0.05	-0.01			
12	Thr	-0.14	0.03	-0.02			
13	Ile	-0.12	0.10	-0.01	0.01		
14	Thr	-0.09	0.06		-0.02		-0.82
15	Leu	-0.20	0.24	-0.06	0.09	0.49	0.12
16	Glu		0.55	-0.33	0.38		
17	Val	-0.03	-0.03	-0.09	0.10		
18	Glu			-0.78	0.88		
20	Ser	-1.42	0.75	-0.69	0.73	-0.58	0.35
21	Asp	-1.51	0.92			-0.50	0.39
22	Thr	-1.24	0.83	-0.65	0.67	-0.31	0.19
23	Ile	-0.62	0.39	-0.14	0.17	-0.30	0.17
25	Asn	-0.97	0.61	-0.34	0.31	-0.22	0.12
26	Val	-1.10	0.68				
27	Lys	-0.76	0.50	-0.10	0.13	-0.17	0.09
28	Ala					-0.15	
29	Lys	-1.21	0.76	-0.01	0.09	-0.08	
30	Ile	-0.80	0.56	0.07	0.02		-0.08
31	Gln			0.02	-0.02	0.06	-0.05
32	Asp	-0.60	0.37	0.28	-0.23	0.13	-0.12
33	Lys	-0.47	0.30	0.20	-0.18	0.29	-0.22
34	Glu	-0.30	0.20	0.10	-0.10	0.38	-0.29
35	Gly	-0.25	0.15	0.11	-0.10	0.30	-0.22
36	Ile	-0.23	0.16	0.08	-0.08	0.24	-0.19
39	Asp	-0.23	0.14	0.11	-0.10		
40	Gln	-0.21	0.13	0.07	-0.07	0.04	-0.07
41	Gln	-0.24	0.15	0.05	-0.04		-0.07
42	Arg	-0.21	0.13	0.01	-0.01		-0.13
43	Leu	-0.27	0.15	-0.02	0.02	-0.09	0.05
44	Ile	-0.21	0.13	-0.04	0.04	-0.24	0.01
45	Phe	-0.18	0.09	-0.05	0.06	-0.34	0.16

46	Ala	-0.12	0.04	-0.05	0.06		
47	Gly	-0.09	0.07	-0.04	0.04	-0.21	0.12
48	Lys	-0.13	0.07	-0.04	0.04	-0.20	0.23
49	Gln	-0.14	0.08	-0.02	0.03		0.06
50	Leu	-0.23	0.14	-0.04	0.04	-0.22	0.09
51	Glu	-0.26	0.17	-0.03	0.05	-0.14	0.11
52	Asp	-0.27	0.18	0.02		-0.14	0.06
54	Arg	-0.36	0.23	-0.02	0.03	-0.23	0.12
55	Thr	-0.40	0.27	-0.09	0.12	-0.32	0.20
56	Leu	-0.77	0.45	-0.29	0.32	-0.54	0.32
57	Ser	-0.57	0.32	-0.26	0.29	-0.60	0.36
58	Asp	-0.38	0.23	-0.14	0.17	-0.45	0.27
59	Tyr	-0.30	0.18	-0.13	0.15	-0.50	0.30
60	Asn	-0.22	0.13	-0.13	0.14	-0.54	0.34
61	Ile	-0.26	0.13	-0.17	0.19	-0.86	0.52
62	Gln	-0.08	0.03	-0.17	0.19		
63	Lys	0.15	-0.17	-0.28	0.30		
64	Glu	0.25	-0.19	-0.23	0.24		
65	Ser	0.04	-0.05	-0.17	0.19		
66	Thr			-0.37	0.51		
67	Leu	-0.17	0.10	-0.08	0.09		0.01
68	His	-0.18	0.11	-0.05	0.05		0.09
69	Leu	-0.10	0.06	0.02	-0.02	0.04	
70	Val	-0.18	0.11				
71	Leu	-0.12	0.08	0.01	-0.01	0.05	-0.15
72	Arg	-0.10	0.13	0.14	-0.11	0.05	-0.10
73	Leu	-0.17	0.11	-0.02	0.02	-0.01	-0.08
74	Arg	-0.09	0.06				
75	Gly	0.01	0.01			-0.10	0.13
76	Gly	-0.07	0.04	0.02	-0.02		

[a] Pseudocontact shifts were measured as the chemical shift observed for the complex with paramagnetic lanthanide minus the chemical shift observed for the complex with diamagnetic Y^{3+}. The uncertainties in PCSs were estimated to be ±0.02 ppm. ^{15}N-HSQC spectra were recorded at 25 ^{o}C of 0.1 mM protein solutions in a buffer of 50 mM HEPES at pH 7.0, using a Bruker 600 MHz NMR spectrometer equipped with a cryoprobe. Residues 74-76 are disordered in the NMR structure 1D3Z (Cornilescu et al., 1998) and therefore were not used in the $\Delta\chi$ tensor fits.

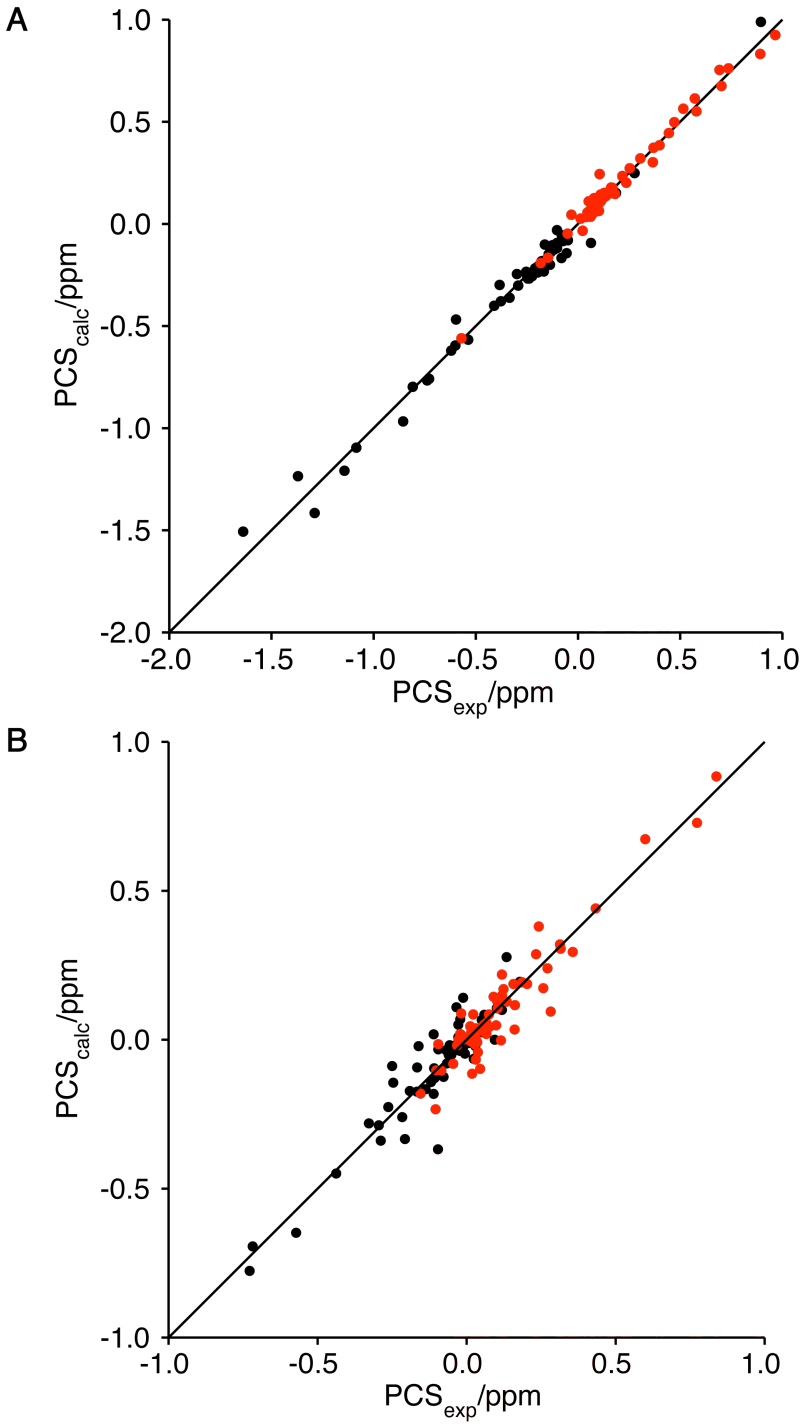

Figure S4. Correlation between back-calculated and experimental pseudocontact shifts of ubiquitin Glu18AzF ligated with (A) the C3 and (B) the C4 tags by the Cu(I)-catalyzed click reaction. Black and red points mark the PCSs obtained with tags loaded with Tb^{3+} and Tm^{3+}, respectively.

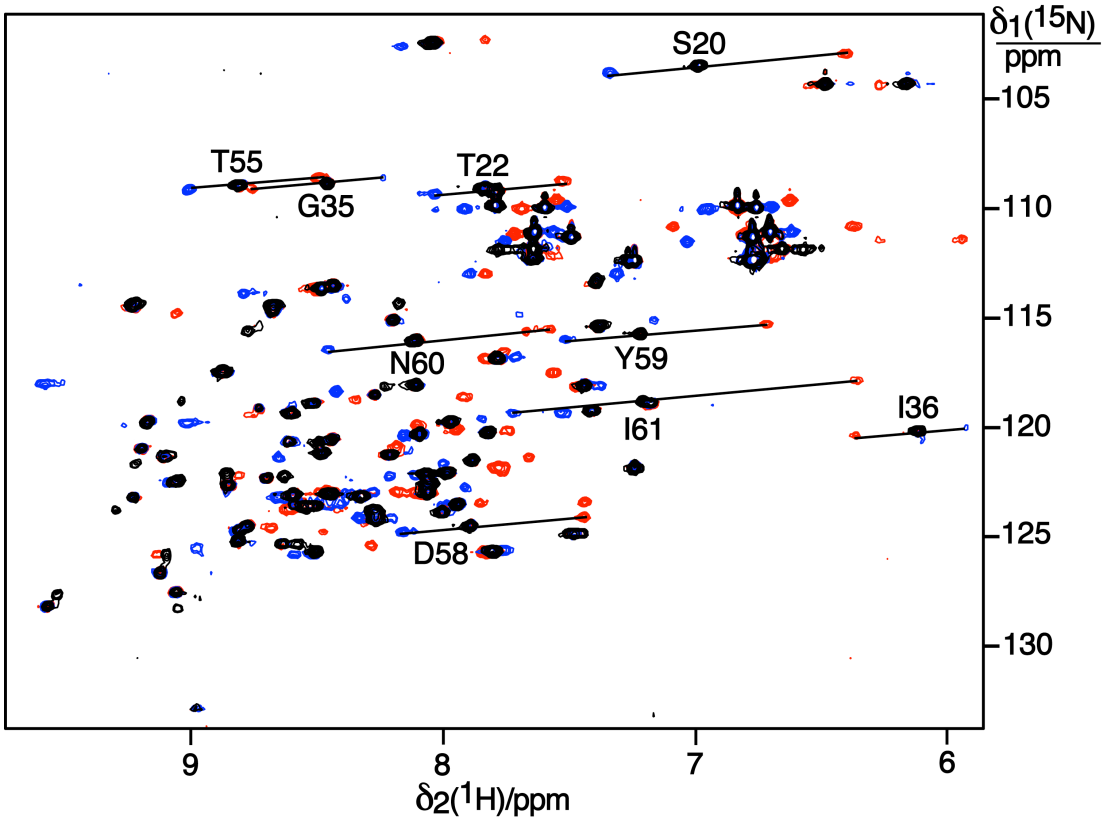

Figure S5. Superimposition of ^{15}N-HSQC spectra of 0.1 mM solutions of uniformly ^{15}N-labelled ubiquitin with the C3 tag ligated to an AzF residue in position 66. The spectrum with the tag loaded with Y^{3+} is shown in black, Tb^{3+} in red, and Tm^{3+} in blue. All spectra were recorded in 50 mM HEPES, pH 7.0. Selected diamagnetic cross-peaks are labelled with their resonance assignments and connected by lines with their paramagnetic partners.

Table S2. Δχ-tensor parameters of the ubiquitin mutant Thr66AzF with the C3 tag, determined by the rotamer library approach of PyParaTools [a,b]

	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	*Q*	*X*	*y*	*z*	α	β	Γ
C3-Tb	20.0 (0.5)	9.2 (0.3)	0.19	38.368	-83.909	11.752	76 (1)	32 (1)	165 (1)
C3-Tm	-12.7 (0.3)	-5.3 (0.2)	0.22	38.368	-83.909	11.752	71 (2)	27 (1)	168 (2)

[a] The axial and rhombic components of the Δχ tensors are given in 10^{-32} m^3 and the Euler angles in degrees, using the zyz convention and unique tensor representation (Schmitz et al., 2008). As in Table 1 of the main text, the metal coordinates and tensor parameters are reported relative to the first conformer of the NMR structure of ubiquitin (PDB ID 1D3Z; Cornilescu et al., 1998). Fits were performed using a rotamer library to identify all conformationally possible metal positions. In a second step, Δχ tensors were fitted to every metal position, simultaneously using the PCS data of Tm^{3+} and Tb^{3+}. The table displays the overall best fits. Standard deviations (in brackets) were determined from fits obtained by using the same metal position while randomly omitting 10% of the PCS data. Quality factors were calculated as the root-mean-square deviation between experimental and back-calculated PCSs divided by the root-mean-square of the experimental PCSs.

[b] Δχ tensor fits were also performed using Numbat (Schmitz et al., 2008) without any restraints on the covalent structures of the tags. The resulting quality factors were 0.18 for C3-Tb and 0.37 for C3-Tm, and the metal shifted by 2.5 Å from the coordinates of Table S2.

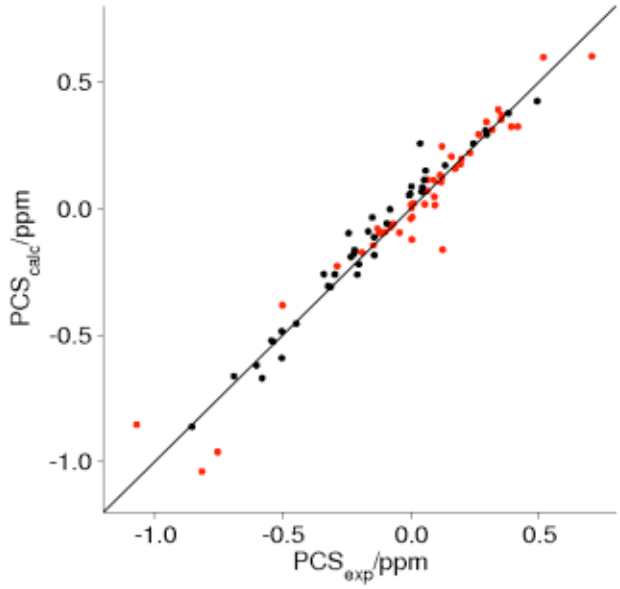

Figure S6. Correlation between back-calculated and experimental pseudocontact shifts of ubiquitin Thr66AzF ligated with the C3 tag. Black and red points mark the PCSs obtained with tags loaded with Tb^{3+} and Tm^{3+}, respectively.

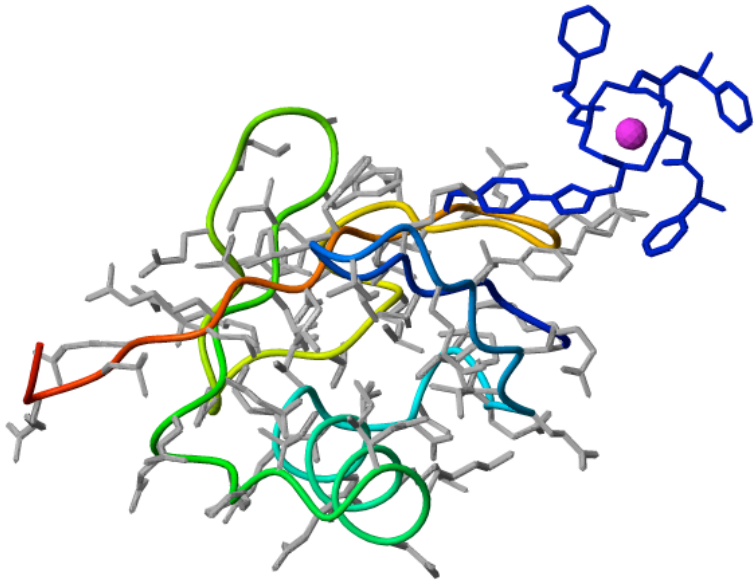

Figure S7. Structure of ubiquitin Thr66AzF with C3 tag. The first conformer of the NMR structure 1D3Z (Cornilescu et al., 1998) is displayed as a ribbon drawing with the heavy atoms of the amino acid side chains in grey. The tag is shown in blue with the lanthanide as a magenta sphere. The position and orientation of the tag corresponds to the tag conformation that delivered the best fits of the $\Delta\chi$ tensor to the experimental PCS.

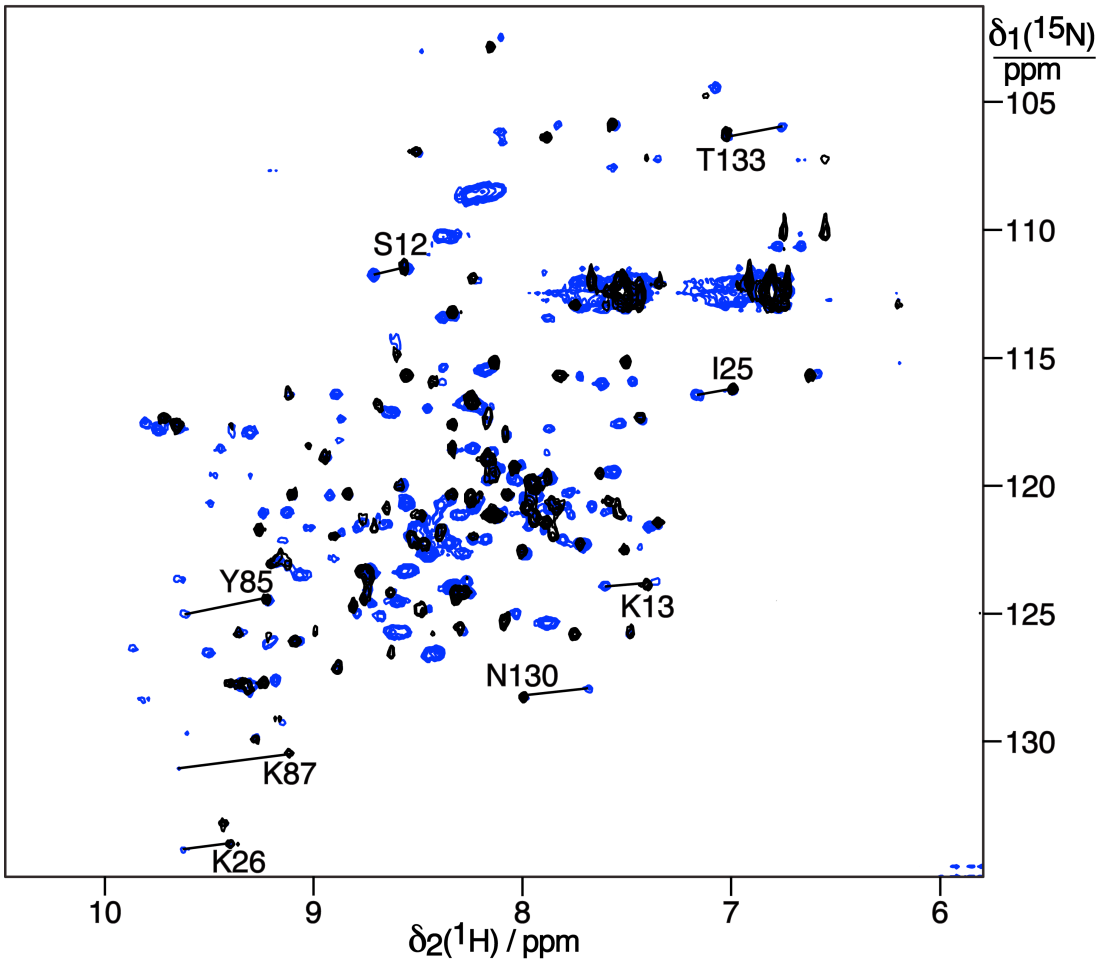

Figure S8. Superimposition of ^{15}N-HSQC spectra of a 0.1 mM solution of uniformly ^{15}N-labelled *S. aureus* sortase A without tag (black peaks) or with C3-Tm^{3+} (blue peaks) ligated to an AzF residue at position 55. Both spectra were recorded in 20 mM MES buffer (pH 6.5) with 1 mM DTT. Selected diamagnetic cross-peaks are labelled with their resonance assignments and connected by lines with their paramagnetic partners.

Table S3. Pseudocontact shifts measured for backbone amide protons of *S. aureus* sortase A Gln55AzF with the C3 tag loaded with Tm^{3+}. [a]

Residue		Tm^{3+}	Residue		Tm^{3+}
12	Ser	0.16	69	Asn	0.06
13	Lys	0.21	70	Tyr	0.10
15	Ala	0.46	71	Gln	0.16
16	Gly	0.54	72	Phe	0.28
18	Ile	0.45	73	Thr	0.17
19	Glu	0.30	74	Asn	0.10
20	Ile	0.27	84	Val	0.28
22	Asp	0.10	85	Tyr	0.41
23	Ala	0.12	86	Phe	0.52
24	Asp	0.11	87	Lys	0.54
25	Ile	0.18	92	Thr	0.35
26	Lys	0.24	94	Lys	0.45
29	Val	0.46	95	Tyr	0.40
30	Tyr	0.49	128	Asp	-0.11
32	Gly	0.71	130	Asn	-0.30
39	Leu	0.22	133	Thr	-0.26
40	Asn	0.14	134	Gly	-0.33
41	Arg	0.16	135	Val	-0.41
42	Gly	0.28	142	Phe	-0.59
61	Gly	0.39	143	Val	-0.87
62	His	0.43			

[a] Pseudocontact shifts were measured as the chemical shift observed for the complex with paramagnetic Tm^{3+} minus the chemical shift observed for the sample without tag. The uncertainties in PCSs were estimated to be ±0.02 ppm. The residue numbering refers to the NMR structure 1IJA (Ilangovan et al., 2001) in which residues 100-119 are disordered and therefore were omitted from the $\Delta\chi$ tensor fits.

Table S4. Δχ-tensor parameters of the Gln55AzF mutant of *S. aureus* sortase A with the C3 tag loaded with Tm^{3+}. [a,b]

	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	Q	x	y	z	α	β	Γ
C3-Tm	13.3 (0.2)	3.7 (0.3)	0.18	0.942	-4.242	13.668	43 (1)	143 (1)	62 (5)

[a] The axial and rhombic components of the Δχ tensors are given in 10^{-32} m^3 and the Euler angles in degrees, using the zyz convention and unique tensor representation (Schmitz et al., 2008). The metal coordinates and tensor parameters are reported relative to the first conformer of the NMR structure of *S. aureus* sortase A (PDB ID 1IJA; Ilangovan et al., 2001). Fits were performed using PyParaTools with a rotamer library to identify all conformationally possible metal positions, as described in the main text for the AzF mutants of human ubiquitin. The table displays the parameters of the overall best fit.

[b] Δχ tensor fits were also performed using Numbat (Schmitz et al., 2008) without any restraints on the covalent structures of the tags. The resulting quality factors were 0.17 and the metal shifted by no more than 1.4 Å from the coordinates reported in the table.

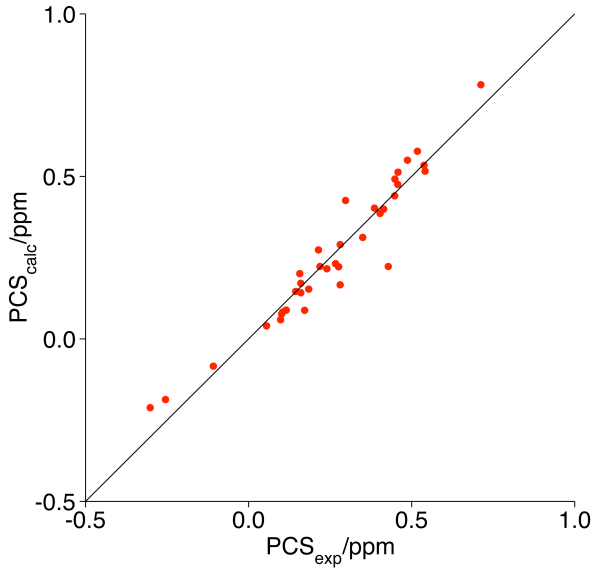

Figure S9. Correlation between back-calculated and experimental pseudocontact shifts of *S. aureus* sortase A Gln55AzF ligated with the C3 tag loaded with Tm^{3+}.

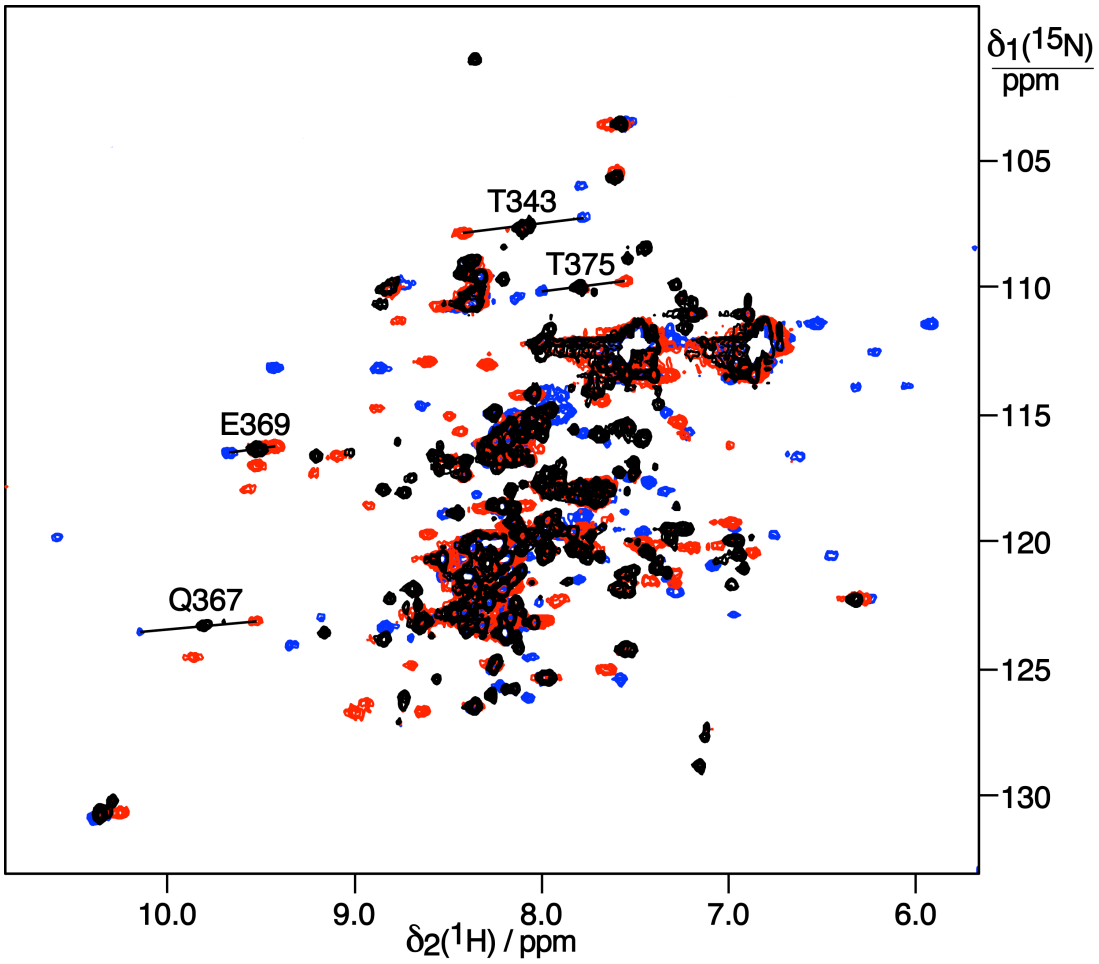

Figure S10. Superimposition of ^{15}N-HSQC spectra of 0.1 mM solutions of the uniformly ^{15}N-labelled intracellular domain of the Lys350AzF mutant of p75 neurotrophin receptor, ligated with diamagnetic C3-Y^{3+} (black), or paramagnetic C3-Tb^{3+} (red) or C3-Tm^{3+} (blue). All spectra were recorded at 25 °C in 20 mM MES buffer (pH 6.5) with 1 mM DTT on a 600 MHz NMR spectrometer. Selected diamagnetic cross-peaks are labelled with their resonance assignments and connected by lines with their paramagnetic partners. No attempt was made to determine the $\Delta\chi$ tensor for this sample, as only a small number of cross-peaks were resolved in the 2D NMR spectrum.

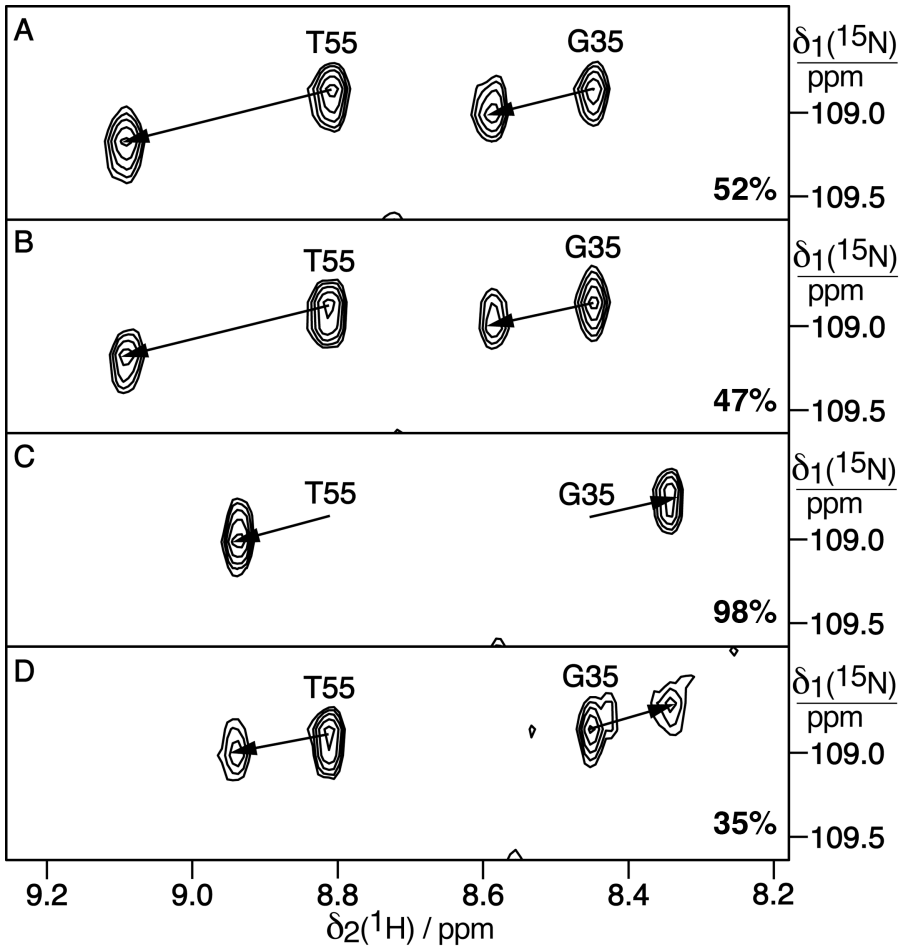

Figure S11. Selected spectral region of the ^{15}N-HSQC spectra of uniformly ^{15}N-labelled ubiquitin Glu18AzF with and without C-terminal His_6 tag. The same ligation conditions were used with all samples, using 0.05 mM protein, 0.5 mM tag, 0.2 mM $CuSO_4$, 1 mM BTTAA, and 5 mM sodium ascorbate. The percentage of successfully ligated protein is reported in each panel. The spectra were recorded at 25 °C on a 600 MHz NMR spectrometer in 50 mM HEPES buffer, pH 7.0, and plotted using the same contour levels. The pseudocontact shifts of the amide cross-peaks of Gly35 and Thr55 are indicated by arrows pointing from the cross-peaks of the unligated protein to the cross-peaks of the ligated protein. (A) Without His_6 tag, using C3-Tm^{3+}. (B) With His_6 tag, using C3-Tm^{3+}. (C) Without His_6 tag, using C4-Tm^{3+}. (D) With His_6 tag, using C4-Tm^{3+}.

REFERENCES

Cornilescu, G., Marquardt, L. J., Ottiger, M., and Bax, A. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc. 120*, 6836–6837.

Ilangovan, U., Ton-That, H., Iwahara, J., Schneewind, O., and Clubb, R. T. (2001) Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc. Nat. Acad. Sci. 98*, 6056–6061.

Schmitz, C., Stanton-Cook, M. J., Su, X.-C., Otting, G., and Huber, T. (2008) Numbat: an interactive software tool for fitting $\Delta\chi$-tensors to molecular coordinates using pseudocontact shifts. *J. Biomol. NMR 41*, 179–189.