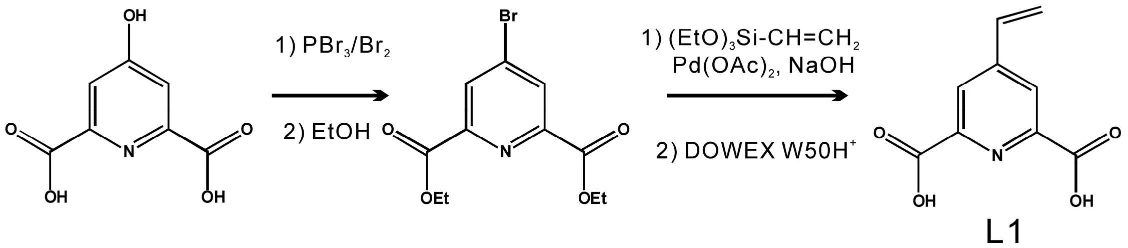
Supporting Information

Thiol-ene reaction: a versatile tool in site-specific labelling proteins with chemically inert tags for paramagnetic NMR

Qing-Feng Li, Yin Yang, Ansis Maleckis, Gottfried Otting, and Xun-Cheng Su

1. Synthesis of L1 and L2

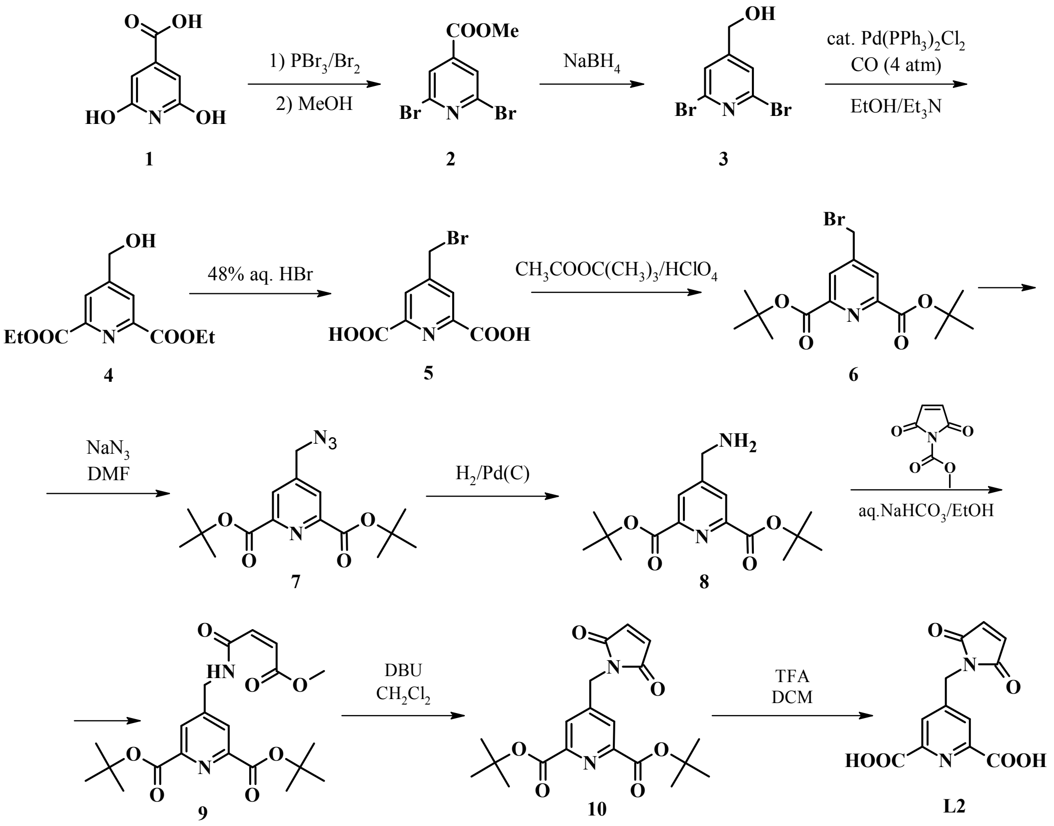
Synthesis of 4-vinylpyridine-2,6-dicarboxylic acid (L1)



Starting from chelidamic acid, **2** were synthesized as reported previously (Schmidt, B.; Ehlert, D. K. *Tetrahedron Lett*. **1998**, *39*, 3999-4002).

Following the cross-coupling protocol reported by Shi et al. (Shi, S.; Zhang, Y. *J. Org. Chem*. **2007**, *72*, 5927-5930), a mixture of 10 mL distilled water, 0.66 g NaOH (16.5 mmol), and 10.0 g PEG 2000 was stirred to homogenous solution, and 1.0 g (3.3 mmol) **2** was added and dissolved. 0.76 g (4mmol) triethoxyvinylsilane and 16 mg (0.07 mmol) $Pd(OAc)_2$ was added and the mixture was stirred at 90 °C for 9 hours and then cooled and filtered. Dowex H^+ ion exchange resin (10.0 g) was added to the filtrate. The solution was filtered after the pH of the suspension decreased to 4. The solution was evaporated under reduced pressure and the solid was suspended in 10 mL ethanol and filtered. 0.5 g white solid **L1** was obtained (yield ~ 80 %). ^{1}H-NMR δ_H ppm (600 MHz, D_2O): 7.95 (s, 2H), 6.82 (dd, *J* = 17.7, 10.9 Hz, 1H), 6.12 (d, *J* = 17.7 Hz, 1H), 5.56 (d, *J* = 10.9 Hz, 1H).

Synthesis of 4-[(2,5-Dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)methyl]pyridine-2,6-dicarboxylic acid (L2)



Compounds **2** and **3** were obtained as described in *J. Org. Chem.* **2006**, *71*, 4696–4699.

<u>4-(Hydroxymethyl)pyridine-2,6-dicarboxylate</u> (**4**)

A medium pressure reaction vessel was charged with 2,6-dibromo-4-(hydroxymethyl)pyridine (**3**) (15.0 g; 56.2 mmol) and $Pd(PPh_3)_2Cl_2$ (2.11 g; 3.01 mmol). The reaction vessel was purged with nitrogen and dry ethanol (150 mL) and triethylamine (19.5 g; 26.9 mL; 193 mmol) were added via canula. The reaction vessel was evacuated and refilled with carbon monoxide and this procedure was repeated three times. Then the reaction mixture was heated to 100 °C under 4 atm of carbon monoxide. Palladium black started to precipitate after approximately 3 hours at 100 °C. The reaction mixture was stirred at 100 °C for another 30 minutes and then cooled to room temperature and filtered. Volatiles were removed under reduced pressure. The residue was dissolved in ethyl acetate (400 mL) and this solution was washed with saturated aq. NH_4Cl solution and with brine, dried over anhydrous sodium sulfate and filtered. Silica gel (approximately 50 g) was added to the filtrate and the solvent was removed under reduced pressure. Column chromatography of the residue (mobile phase: hexane/EtOAc with gradient from 4/1 to 1/1) afforded 12.7 g (89%) of yellowish needles. ^{1}H-NMR δ_H ppm (400 MHz, $CDCl_3$): 8.21 (s, 2H), 4.85 (s, 2H), 4.43 (q, J = 7.0 Hz, 4H), 1.41 (t, J = 7.0 Hz, 6H).

4-(Bromomethyl)pyridine-2,6-dicarboxylic acid (5)

4-(Hydroxymethyl)pyridine-2,6-dicarboxylate (**4**) (910 mg; 3.59 mmol) was heated in hydrobromic acid (48%; 4 mL) at 100 °C for 4 hours. The reaction mixture was poured into water (15 ml) and neutralized with solid $NaHCO_3$ to pH 2. The resulting suspension was extracted with EtOAc (5 × 25 mL). The combined organic phases were washed with brine (2 × 30 mL), dried over anhydrous Na_2SO_4, filtered, and evaporated under reduced pressure to yield the product (762 mg; 82%) as a yellowish crystalline solid. This product was sufficiently pure for further transformations. Analytically pure compound was obtained after crystallization from EtOAc. ^{1}H-NMR δ_H ppm (300 MHz, DMSO-d_6): 8.26 (s, 2H), 4.84 (s, 2H). MS (ES-) $[M-H]^+$: calculated $C_8H_5{}^{79}BrNO_4$ 257.9, found 258.2.

di-*tert*-Butyl 4-(bromomethyl)pyridine-2,6-dicarboxylate (6)

4-(Bromomethyl) pyridine-2,6-dicarboxylic acid (**5**) (640 mg; 2.46 mmol) was suspended in *tert*-butyl acetate (10.0 mL; 8.90 g; 76.6 mmol) and the round bottom reaction flask was tightly sealed. Aq. $HClO_4$ (70%; 387 μl; 4.50 mmol) was then added via a septum and the reaction mixture was stirred at room temperature for 6 hours. The resulting homogenous solution was slowly poured into saturated aq. sodium bicarbonate solution. The mixture was extracted with ethyl acetate (3 × 20 mL) and the combined organic phases were washed with brine, dried over anhydrous sodium sulfate, filtered, and evaporated under reduced pressure. Column chromatography of the residue (silica gel, mobile phase: EtOAc/hexane with gradient from 1/10 to 1/2) yielded the product (819 mg; 89%) as a yellowish solid. ^{1}H-NMR δ_H ppm (300 MHz, $CDCl_3$): 8.17 (s, 2H), 4.48 (s, 2H), and 1.64 (s, 18H).

di-*tert*-Butyl 4-(azidomethyl)pyridine-2,6-dicarboxylate (**7**)

A solution of di-*tert*-butyl 4-(bromomethyl)pyridine-2,6-dicarboxylate (**6**) (800 mg; 2.15 mmol) and sodium azide (260 mg; 4.00 mmol) in DMF (7 mL) was stirred at room temperature for 16 hours. The reaction mixture was poured into water and the resulting suspension was extracted with ethyl acetate (3 × 20 mL). The combined organic phases were washed with brine, dried over anhydrous sodium sulfate, filtered, and evaporated under reduced pressure. Column chromatography of the residue (silica gel, mobile phase: EtOAc/hexane with gradient from 1/10 to 1/2) yielded the product (665 mg; 93%) as a yellowish solid. ^{1}H-NMR δ_H ppm (300 MHz, $CDCl_3$): 8.12 (s, 2H), 4.56 (s, 2H), and 1.64 (s, 18H).

di-*tert*-Butyl 4-(aminomethyl)pyridine-2,6-dicarboxylate (**8**)

Palladium on carbon (10%; 100 mg) was added to the solution of di-*tert*-butyl 4-(azidomethyl)pyridine-2,6-dicarboxylate (**7**) (600 mg; 1.79 mmol) in THF (20 mL). Hydrogen was passed through the suspension at room temperature. When there was no starting material left (TLC control; about 1 hour), the reaction mixture was filtered through a pad of celite. Solvent was removed under reduced pressure to yield product (524 mg; 95%) as a yellowish viscous oil which was used without further purification. ^{1}H-NMR δ_H ppm (300 MHz, DMSO-d_6): 8.30 (s, 2H), 4.22 (s, 2H), and 1.56 (s, 18H).

di-*tert*-Butyl- 4-({[(2*Z*)-4-methoxy-4-oxobut-2-enoyl]amino}methyl)pyridine-2,6-dicarboxylate (**9**)

di-*tert*-Butyl 4-(aminomethyl)pyridine-2,6-dicarboxylate (**8**) (180 mg; 0.584 mmol) and *N*-(methoxycarbonyl)-maleimide (116 mg; 0.750 mmol) were dissolved in EtOAc (4 mL) and this solution was stirred at room temperature for 4 hours. The precipitates were collected by filtration, washed with EtOAc (2 × 2 mL), and dried under reduced pressure. Yield: 209 mg (77%). *N*-(methoxycarbonyl)maleimide was obtained as described in the Supporting Information of *J. Am. Chem. Soc.* **2008**, *130*, 11288–11289. ^{1}H-NMR δ_H ppm (300 MHz, $CDCl_3$): 8.93 (t, *J* = 5.9Hz, 1H), 8.03 (s, 2H), 6.46 (d, *J* = 11.9 Hz, 1H), 6.22 (d, *J* = 11.9 Hz, 1H), 4.46 (d, *J* = 5.9 Hz, 2H), 3.60 (s, 3H), and 1.55 (s, 18H).

di-*tert*-Butyl 4-[(2,5-Dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)methyl]pyridine-2,6-dicarboxylate (**10**)

Amide (**9**) (200 mg; 0.432 mmol) was suspended in CH_2Cl_2 (10 mL) and DBU (~15 μL; 0.1 mmol) was added. The reaction mixture was stirred at room temperature for about 30 minutes until no starting material was left (TLC control). The resulting pink solution was diluted with EtOAc (30 mL), washed with saturated aq. NH_4Cl and brine, dried over anhydrous Na_2SO_4, filtered, and evaporated under reduced pressure. Column chromatography of the residue (silica gel, mobile phase: EtOAc/hexane with gradient from 1/1 to 1/0) afforded the product (130 mg; 78%) in the form of white crystals. ^{1}H-NMR δ_H ppm (300 MHz, $CDCl_3$): 8.04 (s, 2H), 6.81 (s, 2H), 4.78 (s, 2H), and 1.63 (s, 18H). MS (Magnet EI+) $[M]^+$: calculated $C_{20}H_{24}N_2O_6$ 388.2, found 388.2.

4-[(2,5-Dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)methyl]pyridine-2,6-dicarboxylic acid (**L2**)

Trifluoroacetic acid (3 mL) was added to the solution of di-*tert*-butyl 4-[(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)methyl]pyridine-2,6-dicarboxylate (**10**) (120 mg; 0.309 mmol) in CH_2Cl_2 (6 mL). The reaction mixture was stirred at room temperature for 3 hours. Next, the volatiles were removed under reduced pressure. The residue was suspended in EtOAc (5 mL), stirred for 5 minutes, and the solvent was evaporated under reduced pressure. The remaining white crystals were thoroughly washed with EtOAc (3 × 2 mL) and dried under reduced pressure. Yield: 64 mg (75%). ^{1}H-NMR δ_H ppm (300 MHz, DMSO-d_6): 8.05 (s, 2H), 7.12 (s, 2H), and 4.79 (s, 2H). HRMS (TOF ES+) $[M+Na]^+$: calculated $C_{12}H_8N_2O_6Na$ 299.0280, found 299.0280.

2. Construct design and expression of ubiquitin T22C

The single-point mutant T22C of human ubiquitin was prepared in a pET3a vector by PCR-mediated site-directed mutagenesis using the

primers 5' GAGGTTGAGCCCAGTGACTGCATCGAGAATGTCAAGGCAAA 3'

and 5' CATTCTCGATGCAGTCACTGGGCTCAACCTCGAGGGTGA 3'. The coding region in all plasmids was confirmed by DNA sequencing. The correct plasmid was transformed into the *E. coli* strain Rosetta(DE3) (Novagen).

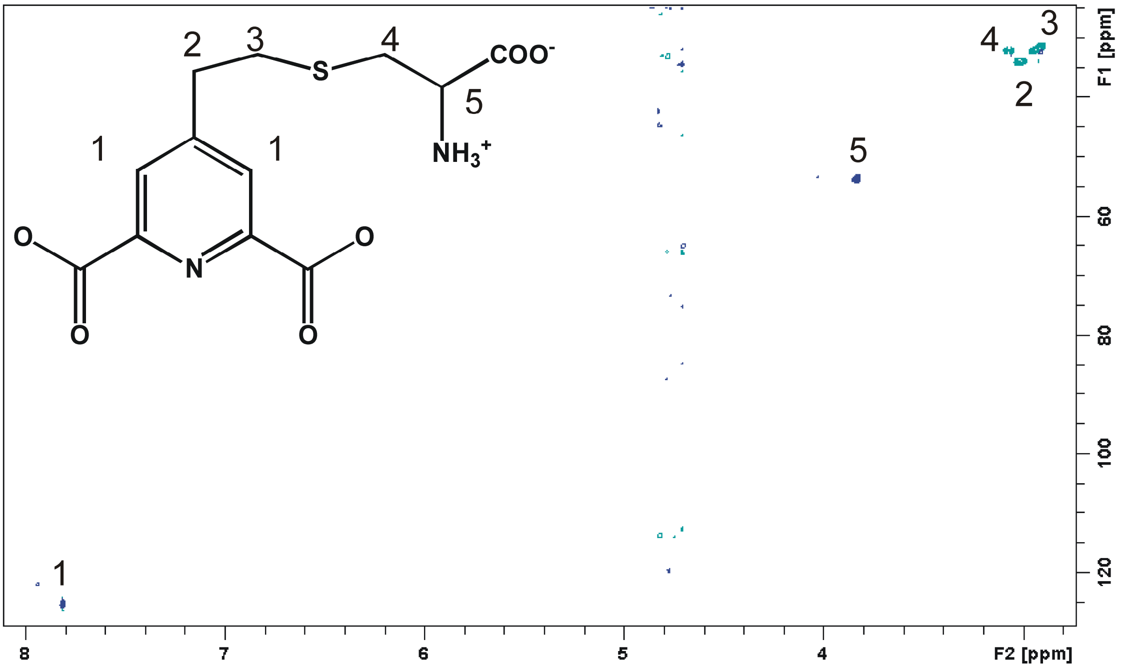
Protein was prepared by growing the cells with the high-density methods reported by Marley et al. (Marley, J.; Lu, M.; Bracken, C. *J. Biomol NMR* **2001**, *20*, 71-75). The mutant ubiquitin was purified from the soluble fraction of the lysed cells by ammonium sulfate precipitation, followed by chromatography on DEAE columns (GE Healthcare Biosciences) and G50 (GE Healthcare Biosciences). Typically, 40 mg of protein was obtained from 250 mL media.

3. Reaction of L1 and L2 with cysteine and proteins

The rate of the Michael reaction of L1 with cysteine was monitored by NMR at room temperature between pH 6.5 to 8. Titration of the solution of L1 with *L*-cysteine showed that the reaction was complete within one hour. Product formation was confirmed by HMBC and ^{13}C-HSQC spectra (Figure S1). No ligation of L1 to the amino group of L-cysteine was observed in the presence of excess L1 (up to four equivalents). The reaction of L1 and L2 with proteins was slower than the reaction with free cysteine but incubation of eight equivalents L1 or L2 with the proteins at pH 7.6 in 20 mM HEPES, phosphate or Tris buffer at room temperature for 24 hours resulted in the ligation product with high yields (final >80%). The ligation product was purified by FPLC using ion exchange or gel filtration. ^{15}N-HSQC spectra showed no evidence for reactions of the amino groups of protein lysine side chains with L1 or L2.

4. Ligation protocol

4.0 mL of a 1.0 mM solution of ^{15}N-labeled protein were diluted into 5 mL 20 mM Tris at pH 7.6 and eight equivalents of vinyl DPA tag (100 mM stock solution) were added to the protein solution. The pH of the mixture was adjusted to 7.6 using 1.0 M NaOH and the solution was incubated at room temperature for about 16 to 48 hours. Subsequently, the solution was concentrated to 1.0 mL (optionally in the presence of 1.0 mM TCEP). The ligation product was purified through either ion exchange or gel filtration, depending on the property of protein. The overall yield of purified ligation product was usually above 80%.



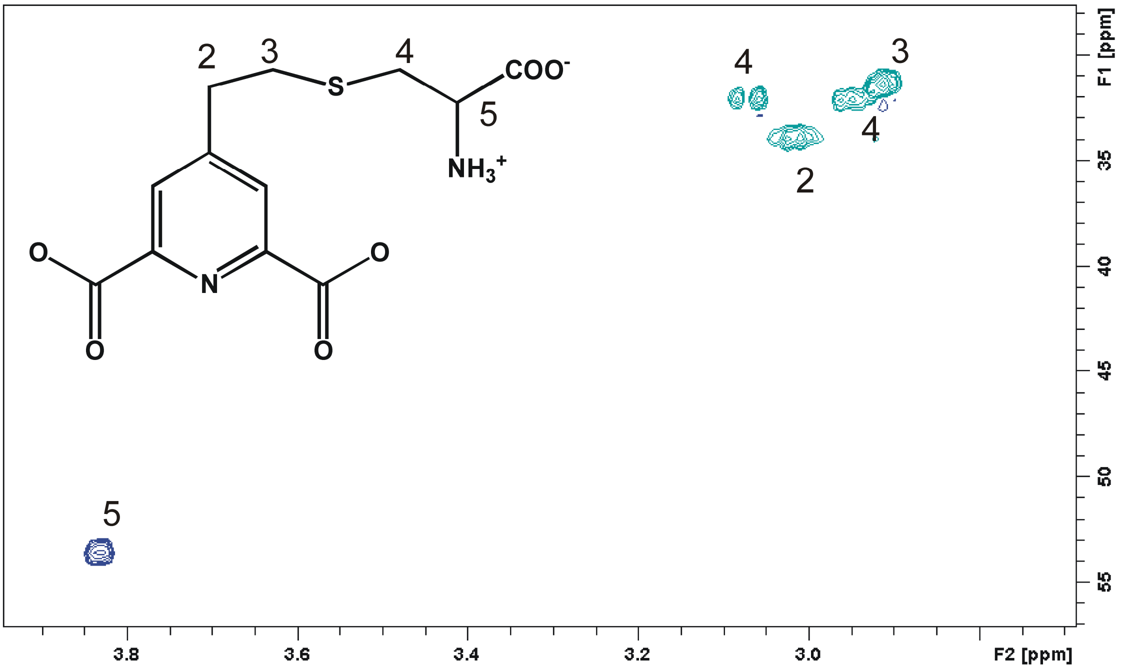


Figure S1. Overview and selected spectral region of a ^{13}C-HSQC spectrum recorded at pH 7.5 of the mixture of 4-vinyl DPA and *L*-cysteine after incubation for 1 hour. This spectrum and all other spectra in the Supporting Information were recorded in 90%H_2O/10% D_2O at 25 °C, using a ^{1}H NMR frequency of 600 MHz.

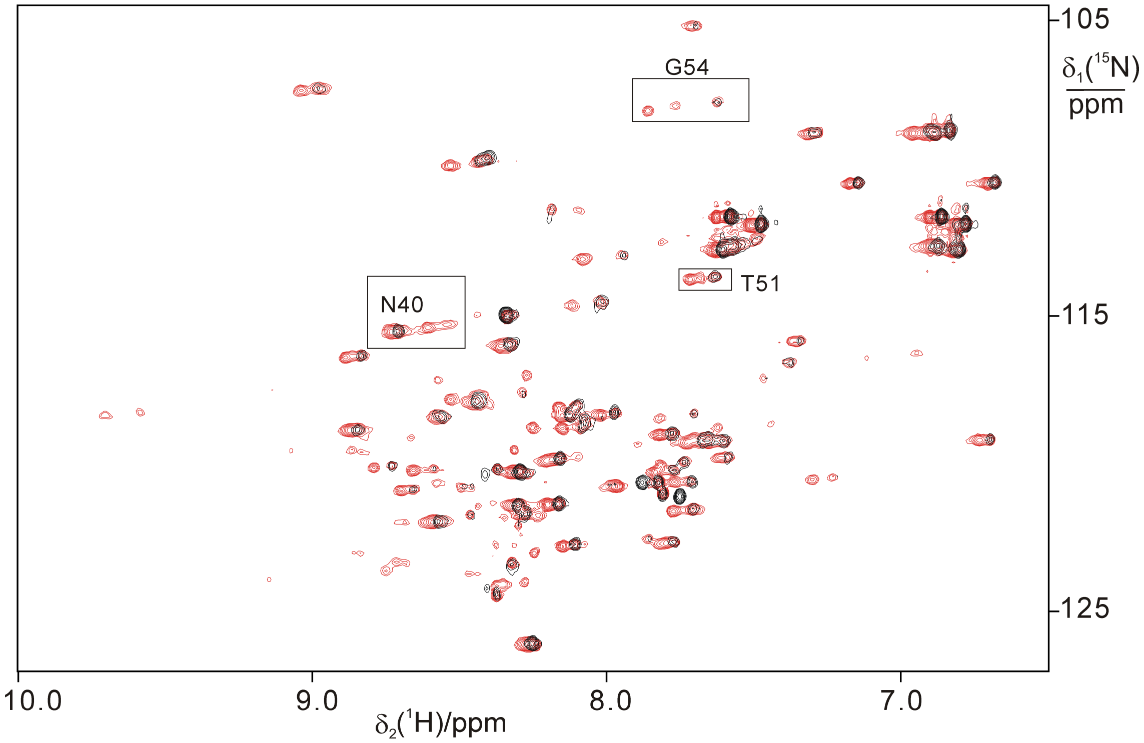


Figure S2. Superimposition of ^{15}N-HSQC spectra of uniformly ^{15}N-labeled ArgN derivatized with L2 at Cys68 in the presence of a 1:1 mixture of Y^{3+} and Yb^{3+} (red) and in the presence of Y^{3+} alone (black). The ratio of lanthanides to protein was about 1.5:1. The spectra were recorded at pH 6.5 in 20 mM MES buffer. The boxes highlight examples of peak doubling arising from the formation of diastereomers.

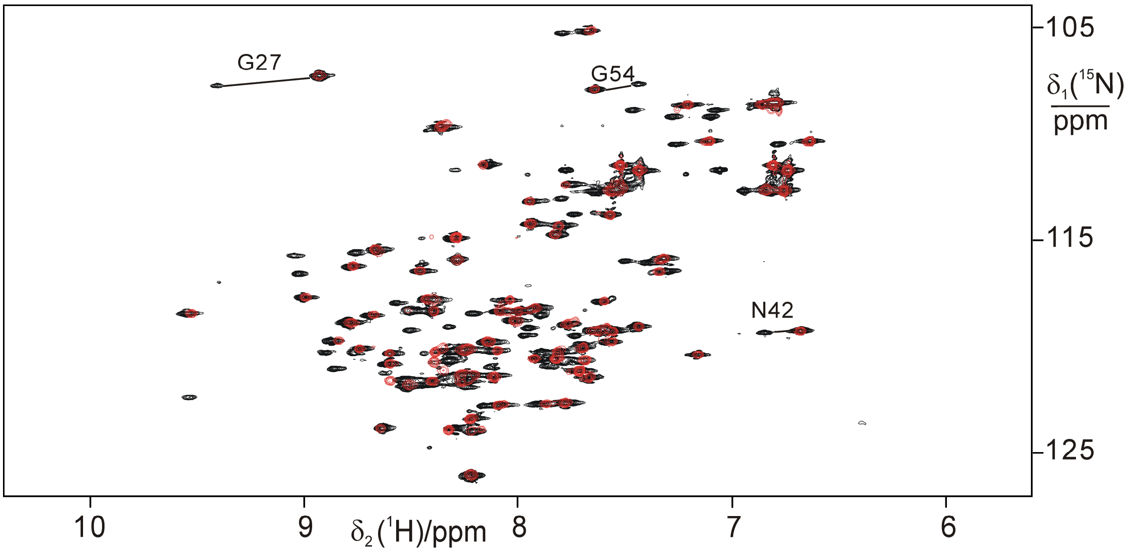


Figure S3. Superimposition of ^{15}N-HSQC spectra of uniformly ^{15}N-labeled ArgN derivatized with L1 at Cys68 in the presence of a 1:1 mixture of Y^{3+} and Tm^{3+} (black) and in the presence of Y^{3+} (red). The ratio of lanthanide to protein was about 1.2:1. The spectra were recorded at pH 6.5 in 20 mM MES.

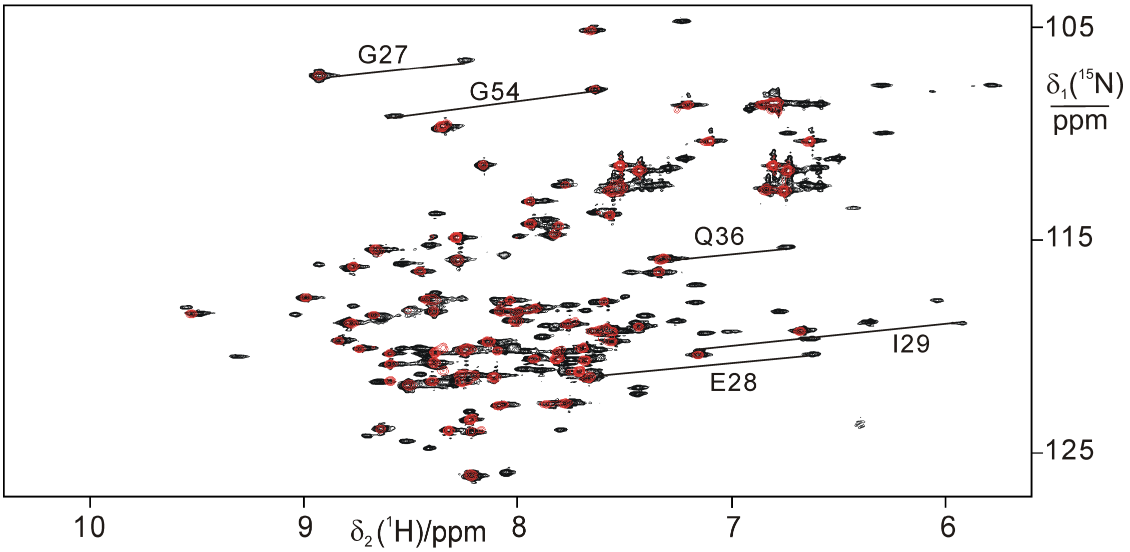


Figure S4. Same as Figure S3, except that Dy^{3+} was present instead of Tm^{3+}.

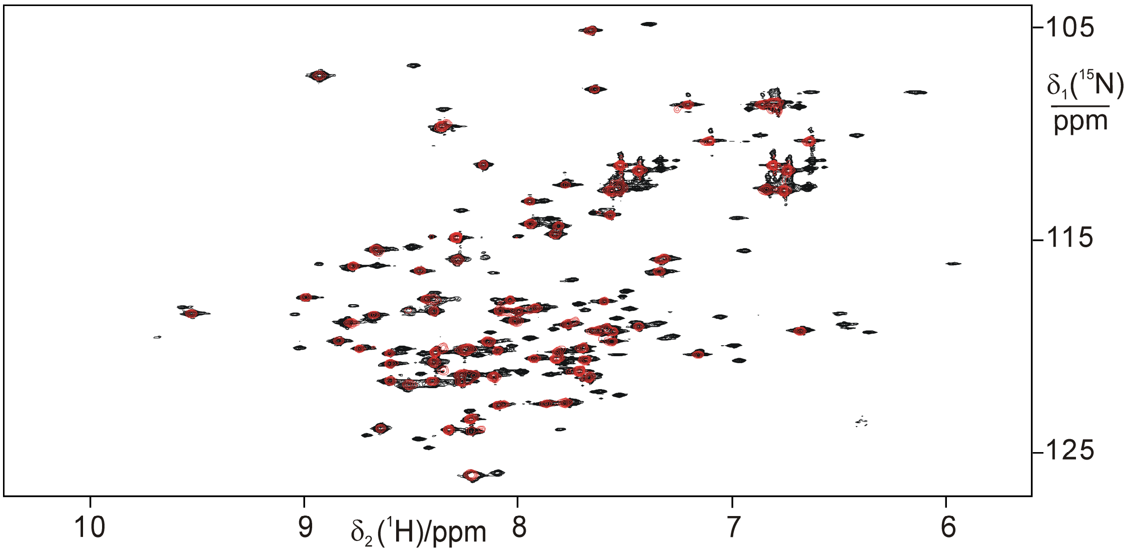


Figure S5. Same as Figure S3, except that Tb^{3+} was present instead of Tm^{3+}.

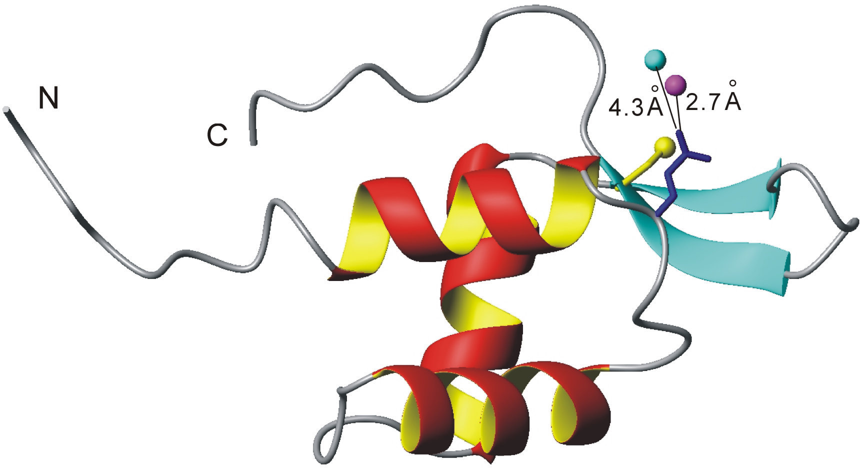


Figure S6. Comparison of the metal position determined for the ArgN-4MMDPA-Ln complex (with the Ln^{3+} ion shown in magenta) and for the ArgN-L1-Ln complex (Ln^{3+} in cyan). The side chains of E21 and C68 are colored in blue and yellow, respectively, with the sulfur of C68 shown as a ball.

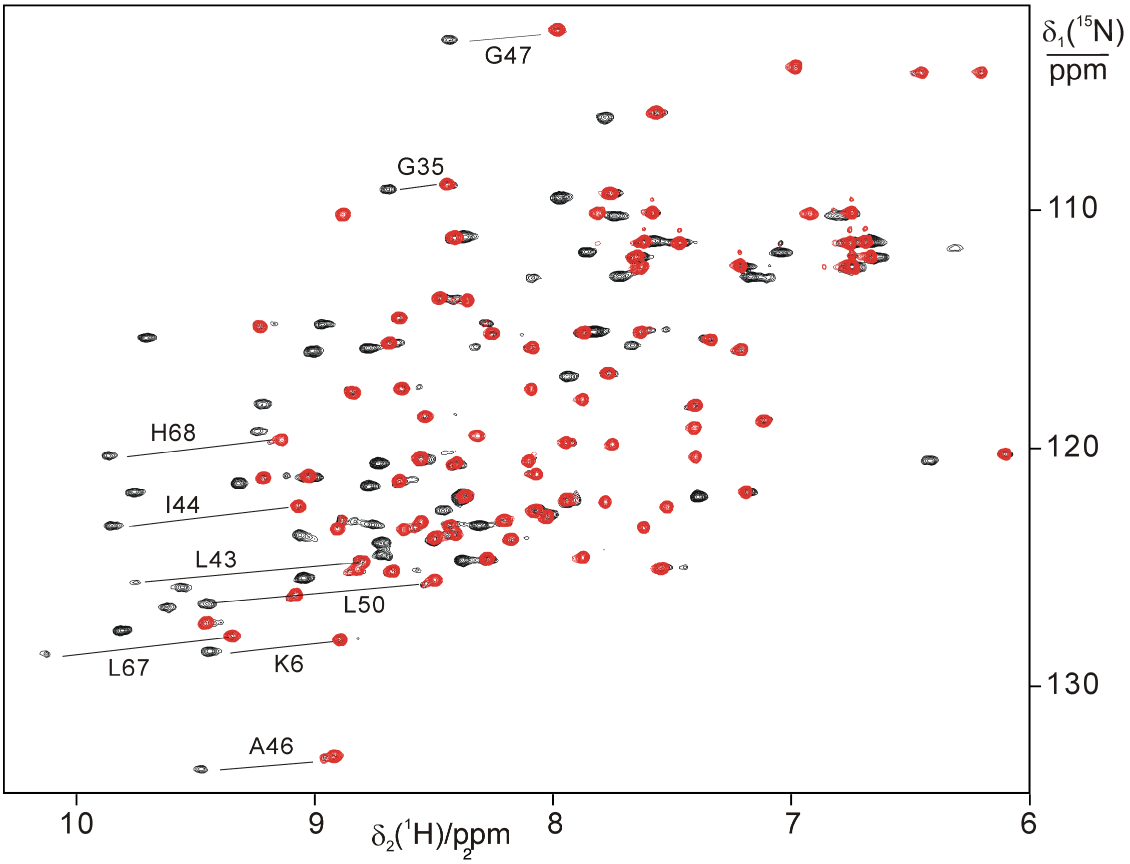


Figure S7. Superimposition of ^{15}N-HSQC spectra of a 0.08 mM solution of uniformly ^{15}N-labeled ubiquitin T22C derivatized with L1 in the presence of a 1:1 mixture of Y^{3+} and Tm^{3+} (black) and in the presence of Y^{3+} (red). The ratio of lanthanides to protein was about 1.2:1. The spectra were recorded at pH 6.4 in 20 mM MES buffer.

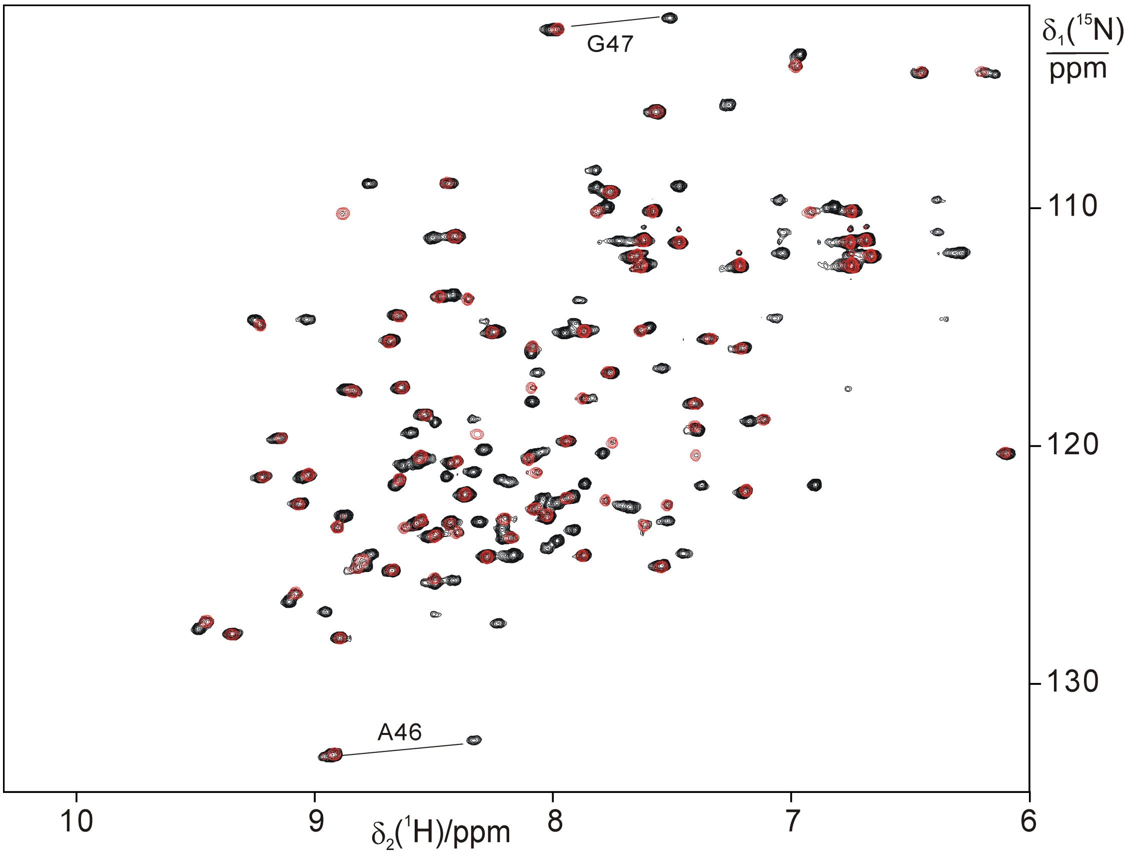
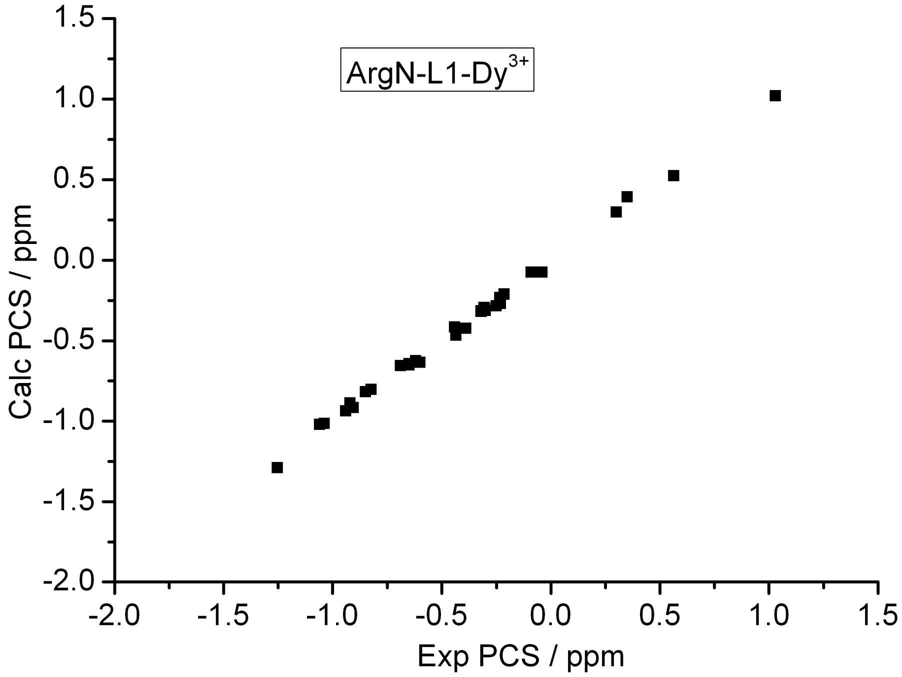
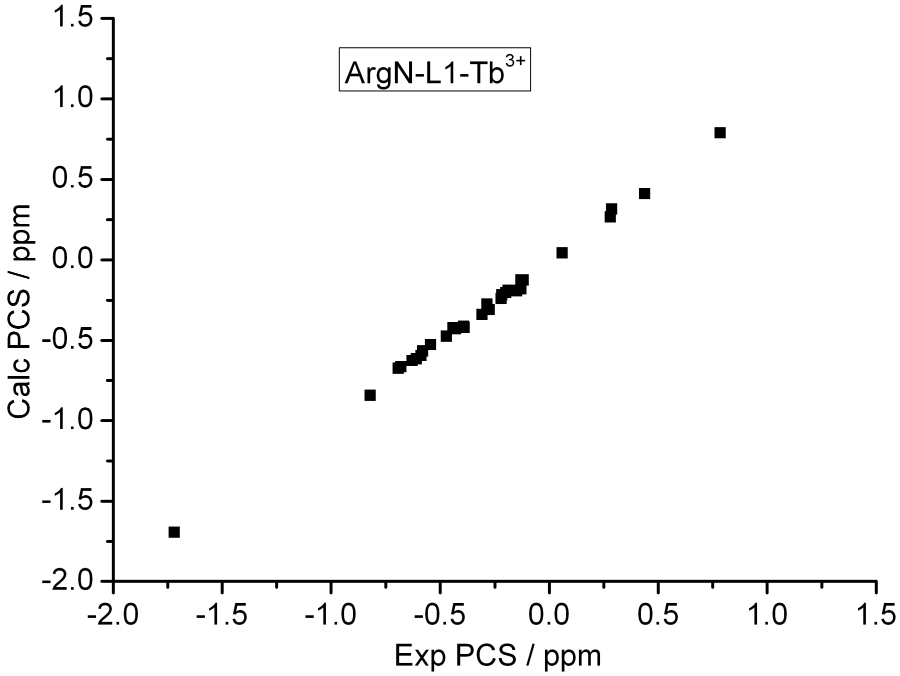
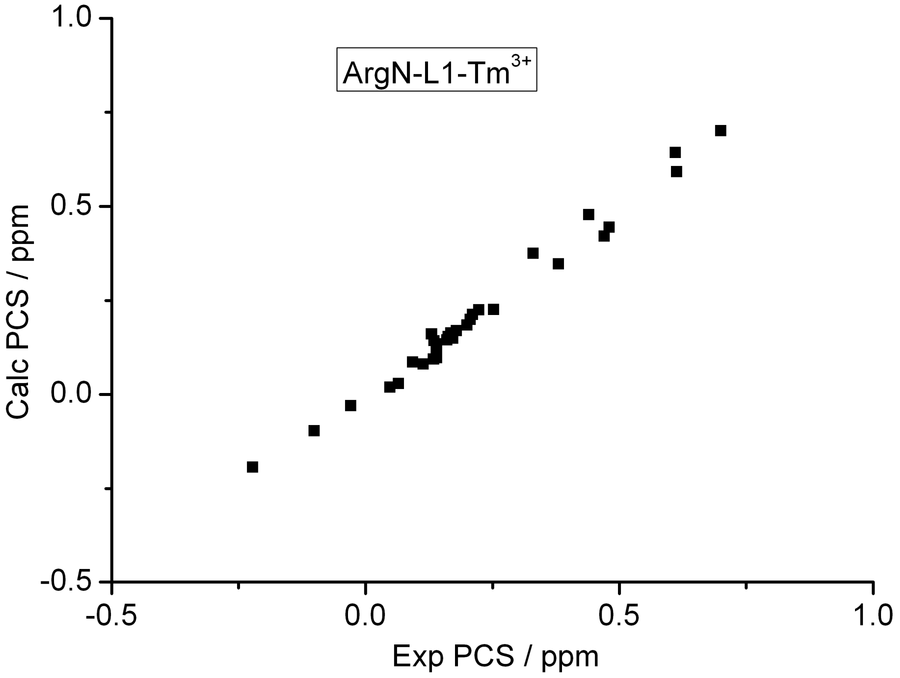
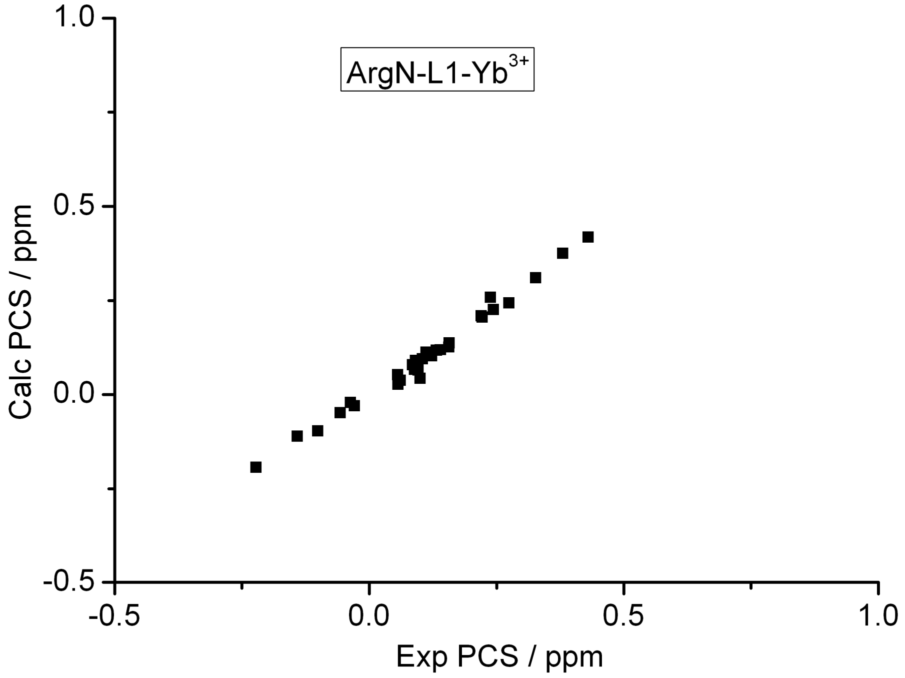


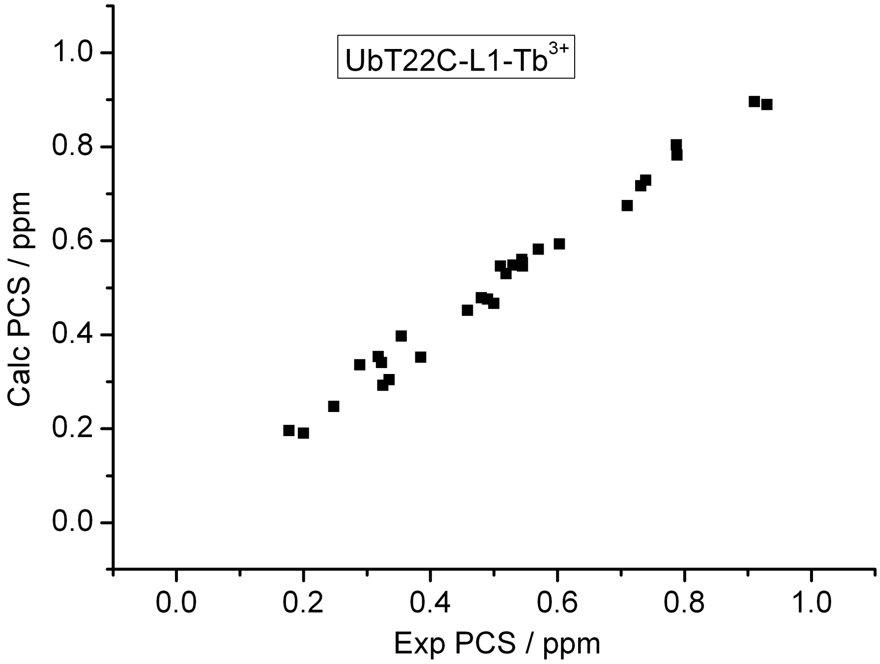
Figure S8. Same as Figure S7, except that Tm^{3+} was substituted by Tb^{3+}.

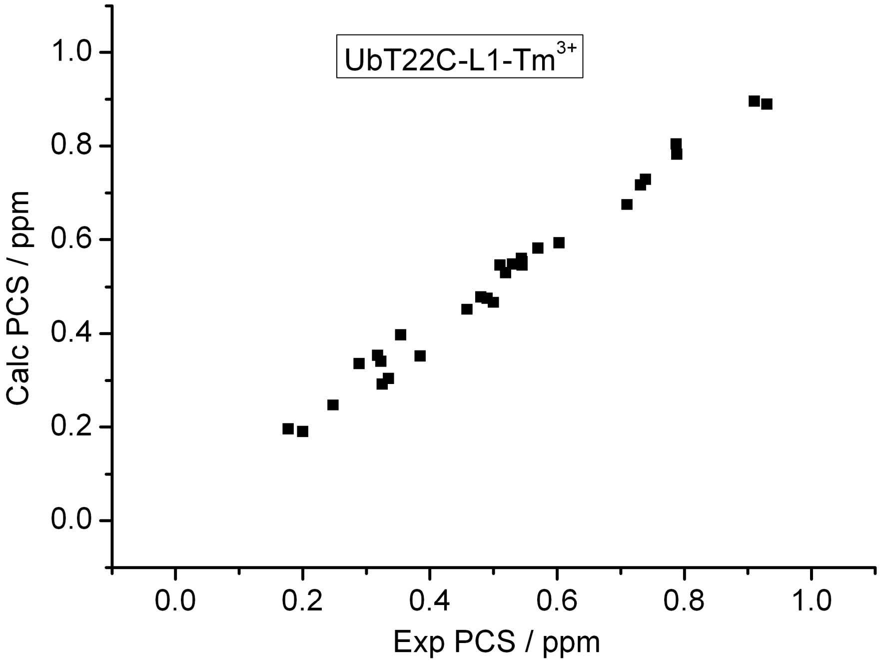

ArgN-L1-Dy^{3+}
Calc PCS / ppm
Exp PCS / ppm
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5


ArgN-L1-Tb^{3+}
Calc PCS / ppm
Exp PCS / ppm
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5


ArgN-L1-Tm^{3+}
Calc PCS / ppm
Exp PCS / ppm
1.0
0.5
0.0
-0.5
-0.5
0.0
0.5
1.0


ArgN-L1-Yb^{3+}
Calc PCS / ppm
Exp PCS / ppm
1.0
0.5
0.0
-0.5
-0.5
0.0
0.5
1.0


UbT22C-L1-Tb3+
Calc PCS / ppm
Exp PCS / ppm
0.0
0.2
0.4
0.6
0.8
1.0


UbT22C-L1-Tm3+
Calc PCS / ppm
Exp PCS / ppm
0.0
0.2
0.4
0.6
0.8
1.0

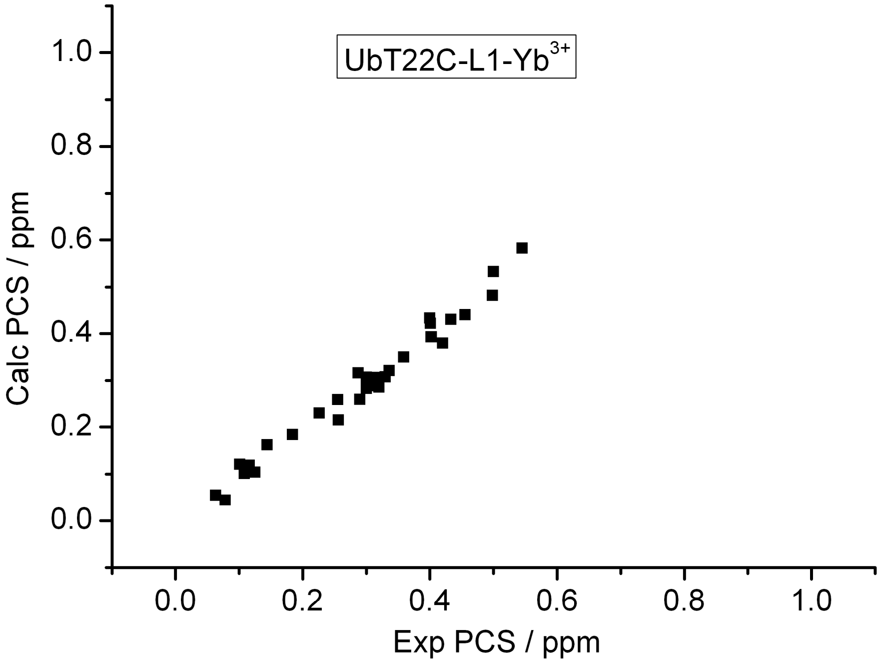
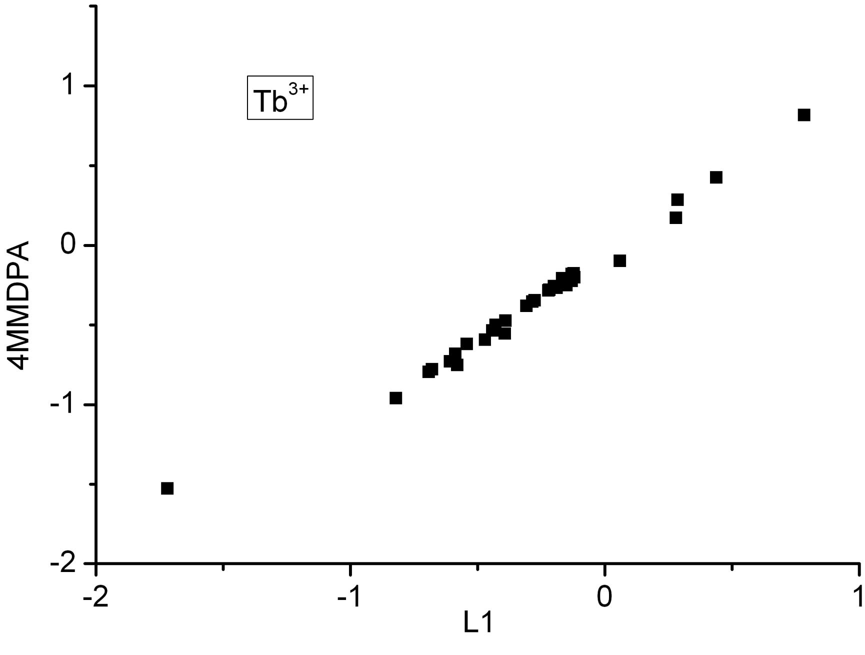


Figure S9. PCSs back-calculated with the program Numbat (Calc PCS) plotted against the experimentally measured PCSs (Exp PCS).



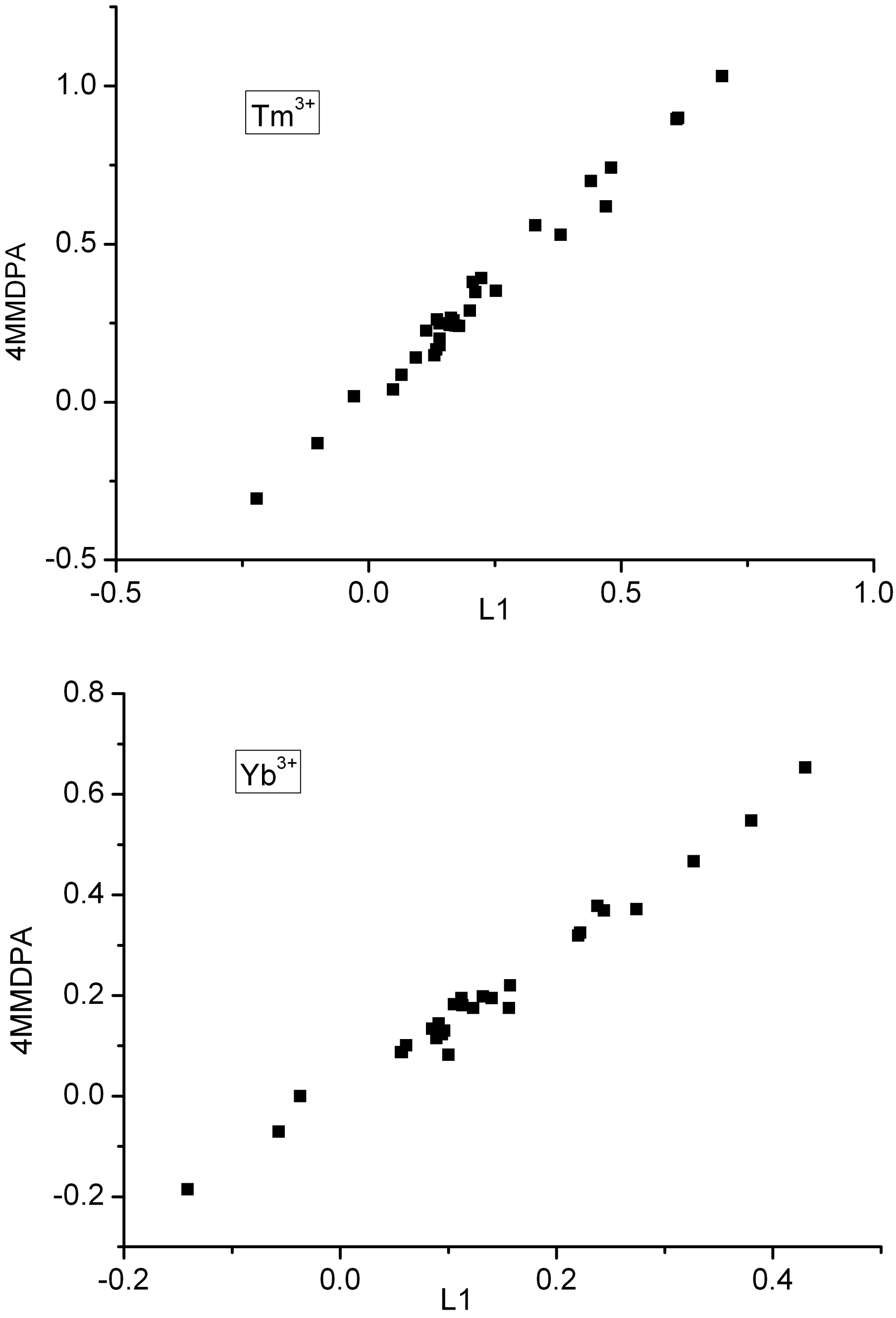


Figure S10. Correlations of the PCSs measured of ArgN-L1 and ArgN-4MMDPA loaded with different lanthanides.