Supporting information

Site-specific tagging proteins via a rigid, stable and short thiolether tether for paramagnetic spectroscopic analysis

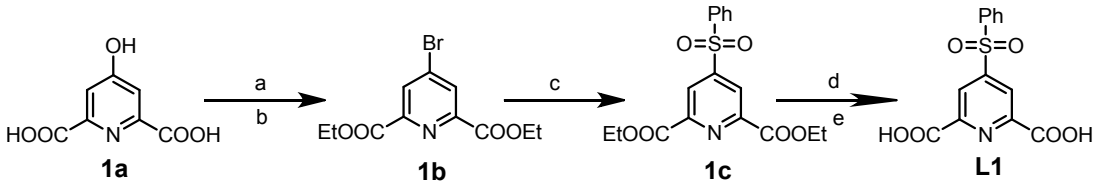
*Yin Yang, Jin-Tao Wang, Ying-Ying Pei, and Xun-Cheng Su**

1. Synthetic procedure of 4-(phenylsulfonyl) and 4-(methylsulfonyl) pyridine derivatives

General information

^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Bruker Avance (400 MHz) spectrometer in the stated solvents unless otherwise mentioned. Chemical shifts were reported in ppm on the δ scale from an internal standard (NMR descriptions: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad). Coupling constants, J, are reported in Hz. Mass spectroscopy was performed in the large facility at State Key Laboratory of Elemento-organic Chemistry, Nankai University, Tianjin. All solvents were commercially available grade. All reactions were carried out under argon atmosphere unless otherwise mentioned.

Scheme S1.



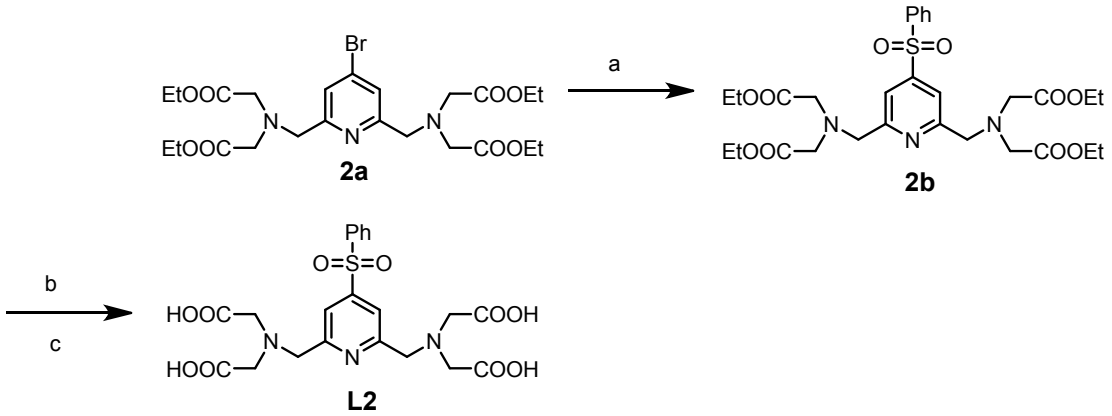
Conditions: (a) PBr_5, 100℃; (b) CH_3CH_2OH, 0℃; (c) $PhSO_2Na$, Bu_4NBr, MeCN, reflux; (d) NaOH, EtOH, H_2O, rt; (e) conc. HCl

Diethyl 4-bromo-2,6-pyridinecarboxylate (**1b**). Staring form 4-hydroxy-2,6-pyridinedicarboxylic acid (**1a**), **1b** was synthesized as previously described[1]. 52.0%。^{1}H-NMR (400 MHz，$CDCl_3$) δ ppm：8.45 (2H，s), 4.52 (4H, q, J = 7.1Hz), 1.48 (6H, t, J = 7.1Hz). ^{13}C-NMR (100 MHz,$CDCl_3$) δ ppm: 163.50, 149.52, 134.89, 131.11, 62.73, 14.14.

Diethyl 4-(phenylsulfonyl)-2,6-pyridinecarboxylate (**1c**). 1.00 g (3.32 mmol) **1b**, 1.11 g (6.64 mmol) sodium phenylsulfinate and 106.24 mg tetrabutylammonium bromide was added into a solution of 30 mL CH_3CN. The resulting solution was heated to reflux with stirring under argon atmosphere for about 36 h, and then the mixture was cooled down to room temperature. The organic solvent was removed under reduced pressure. The resulting solid was mixed with water and extracted with ethyl acetate. The organic layer was washed with brine, dried over sodium sulfate, filtered and concentrated under reduced pressure. The residue was purified by silica gel column chromatography to afford 0.44 g, the title compound **1c** as a white solid (36.7%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ ppm: 8.68 (2H, s), 8.04 (2H, d, J = 7.5Hz), 7.70 (1H, t, J = 7.4Hz), 7.56 (2H, t, J = 7.5Hz), 4.52 (4H, q, J = 7.1Hz), 1.48 (6H, t, J = 7.1Hz). ^{13}C-NMR (100 MHz, $CDCl_3$) δ ppm: 163.25, 152.84, 150.47, 138.84, 134.70, 129.93, 128.44, 124.60, 62.94, 14.19.

4-(Phenylsulfonyl)-pyridine-2,6-dicarboxylic acid (**L1**). 0.44 g 1c (1.21 mmol), 0.19 g NaOH (4.84 mmol), 6.0 mL ethanol and 6.0 mL water was added into a 100 ml-flask and the resulting mixture was stirred at room temperature for 10 h. At 0°C, the pH of the above mixture was adjusted to ~ 2 with concentrated HCl. The white precipitates were filtered, washed with water, and dried. 0.28 g white powder was obtained (75.7%). ^{1}H-NMR (400 MHz, 90%H_2O+10%D_2O, pH＞12) δ ppm: 8.33 (2H, s), 7.96 (2H, d, J = 8.1Hz), 7.66 (1H, t, J = 7.5Hz), 7.56 (2H, t, J = 8.1Hz). ^{13}C-NMR (100 MHz, 90%H_2O + 10%D_2O, pH＞12) δ ppm: 170.61, 154.97, 150.58, 137.53, 135.21, 130.05, 128.07, 121.66.

Scheme S2.



Conditions: (a) $PhSO_2Na$, Bu_4NBr, MeCN, reflux; (b) NaOH, EtOH, H_2O, rt; (c) H^+.

4-Bromo-2,6-bis[N,N'-bis(ethyoxycaronylmethyl)aminomethyl]pyridine (**2a**). Starting from 4-hydroxy-2,6-pyridinedicarboxylic acid, 3a was synthesized as previously described[1].

4-(Phenylsulfonyl)-2,6-bis[N,N'-bis(ethyoxycaronylmethyl)aminomethyl]pyridine (**2b**).

0.80g (1.43 mmol) **2a**, 0.48g sodium phenylsulfinate (2.86 mmol) and tetrabutylammonium bromide 45.76mg was added to the solution of 30 mL acetonitrile, and the above mixture was heated to reflux for 12 h under argon atmosphere till the reaction was completed. The solvent was removed under reduced pressure. The resulting solid was mixed with water and then extracted with ethyl acetate. The organic layer was washed with brine, dried over sodium sulfate, and filtered. The solvent was removed, resulting 0.62 g the title compound **2b** as yellow oil (69.7%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ ppm: 8.04 (4H, d, J = 5.4Hz), 7.63 (1H, t, J = 7.1Hz), 7.56 (2H, t, J = 7.5Hz), 4.18 (8H, q, J = 7.1Hz), 4.12 (4H,s), 3.60 (8H, s), 1.28 (12H, t, J = 7.1Hz). ^{13}C-NMR (100 MHz, $CDCl_3$) δ ppm: 170.93, 161.33, 151.24, 140.25, 133.77, 129.43, 128.26, 118.13, 60.65, 59.72, 55.11, 14.24.

4-Phenylsulfonyl-(pyridin-2,6-diyl)bismethylenenitrilo tetrakis(acetic acid) (**L2**). 0.62g **2b** (1.00mmol) was dissolved in 10 mL ethanol and then 0.33g NaOH (8.30 mmol) in 10 ml water was added. The above mixture was stirred at room temperature for 5 h. Dowex H^+ ion exchange resin was then added to the solution till the pH was decreased to ~3. The solution was filtered and the filtrate was evaporated under reduced pressure. The solid was suspended in 20 mL acetone and filtered, resulting 0.41g yellow power (80.6%). ^{1}H-NMR (400 MHz, 90%H_2O + 10%D_2O, pH>12) δ ppm: 7.93 (2H, d, J = 7.7Hz), 7.83(2H, s), 7.66(1H, t, J = 7.4Hz), 7.56 (2H, t, J = 7.7Hz), 3.84 (4H, s), 3.07 (8H, s). ^{13}C-NMR (100 MHz, 90%H_2O + 10%D_2O, pH＞12) δ ppm: 179.16, 161.16, 149.88, 137.90, 134.95, 129.98, 127.83, 118.70, 58.98, 57.97. MS-ESI ([M–1], cal.508.1): 508.2.

2. Reaction of 4-(phenylsulfonyl)-pyridine-2,6-dicarboxylic acid 1 with L-cysteine in aqueous solution and structural characterization of the ligation products

General procedure for the conjugation reaction: 10 mM sulfonated pyridine derivative **1** was prepared in D_2O and the pD was adjusted to 7.5. The 1D ^{1}H-NMR, 2D ^{13}C-HSQC and ^{13}C-HMBC were recorded at 298 K with a proton frequency of 600 MHz. For comparison, the NMR spectra were recorded for 10 mM sodium phenylsulfinate. A mixture of 10 mM **1** with 30 mM L-cysteine was incubated in D_2O at pD 7.5 for 2 h at 298 K and NMR spectra were then recorded. The final product was characterized by NMR spectra, resulting released phenylsulfinate during the reaction of **1** and L-cysteine. The reaction demonstrated that pyridyl phenylsulfone is reactive to free thiols.

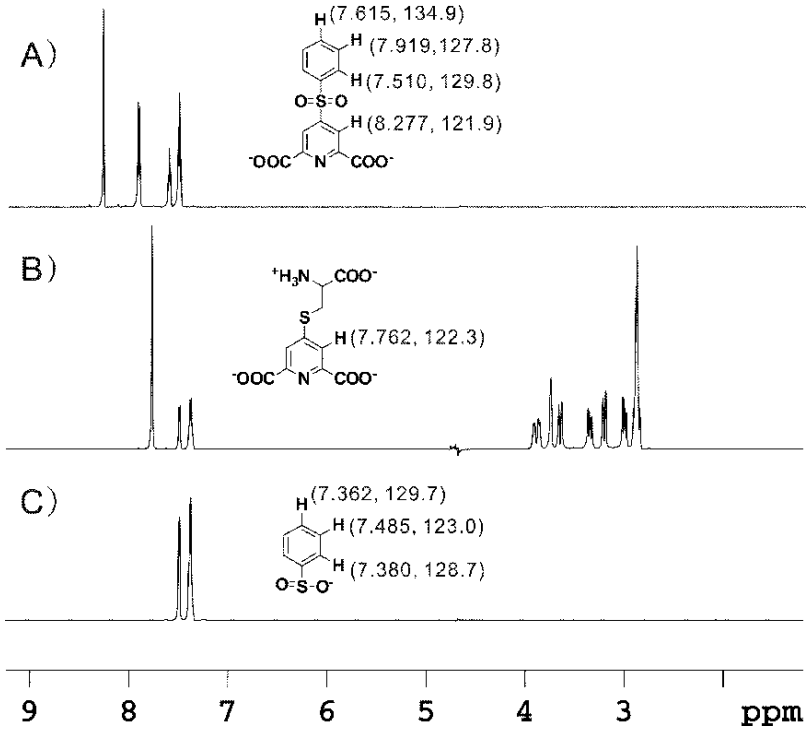


Fig. S1 1D ^{1}H-NMR spectra recorded in monitoring the reaction of tag **1** and L-cysteine. A) 10 mM **1**. B) Incubation the mixture of 10 mM **1** and 30 mM L-cysteine at room temperature after 2 h. C) 10 mM sodium sulfinate. The spectra were recorded in D_2O at pD 7.5 and 298 K with a ^{1}H NMR frequency of 600 MHz and the chemical shifts of aromatic resonances (^{1}H and ^{13}C) were depicted in the structures of molecules.

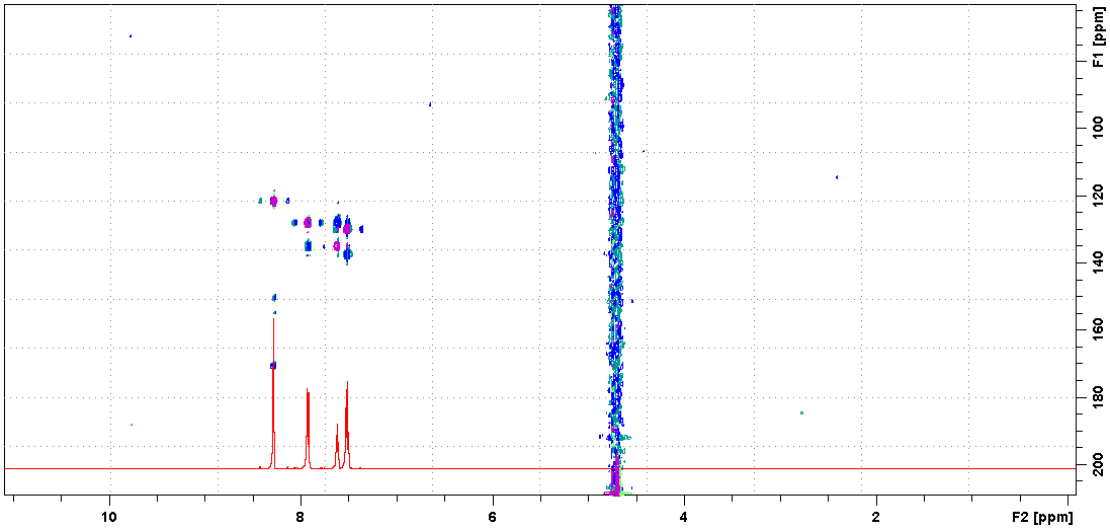


Fig. S2 Superimposition NMR spectra of ^{13}C-HSQC (magenta); ^{13}C-HMBC (blue and green) and 1D proton (red) recorded for the solution of 10 mM **1** in in D_2O at pD 7.5 and 298 K with a proton frequency of 600 MHz.

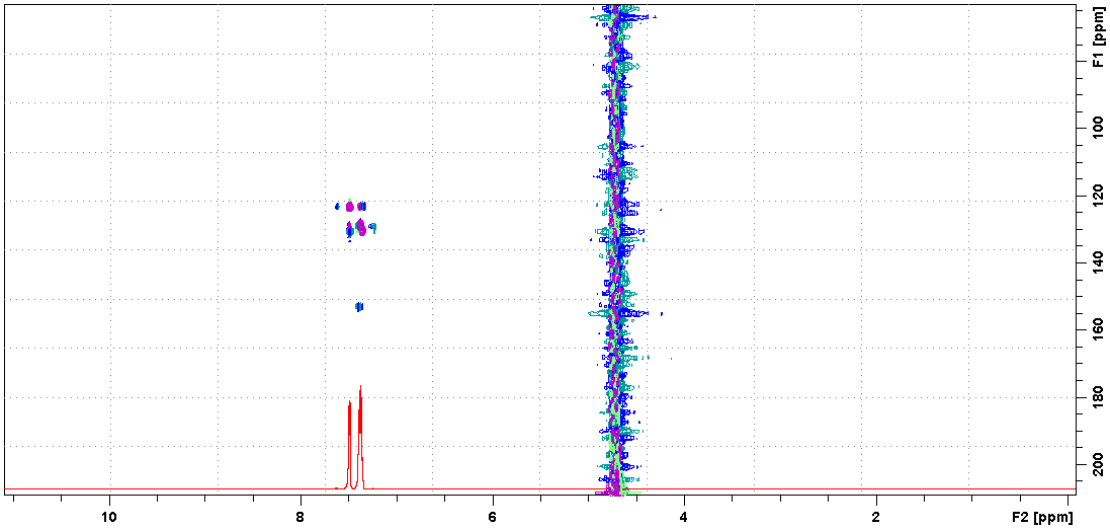


Fig. S3 Superimposition NMR spectra of ^{13}C-HSQC (magenta); ^{13}C-HMBC (blue and green) and 1D proton (red) recorded for 10 mM sodium phenylsulfinate in D_2O at pD 7.5 and 298 K with a proton frequency of 600 MHz.

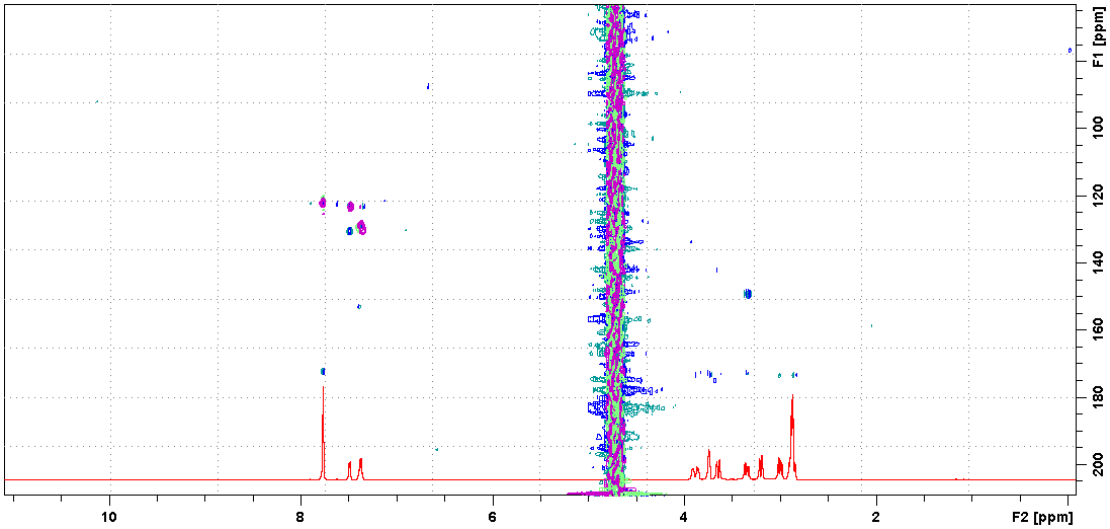


Fig. S4 Superimposition NMR spectra of ^{13}C-HSQC (magenta); ^{13}C-HMBC (blue and green) and 1D proton (red) for the solution of 10 mM **1** after incubation with 30 mM L-cysteine for 2 h. The NMR spectra were recorded in D_2O at pD 7.5 and 298 K with a proton frequency of 600 MHz.

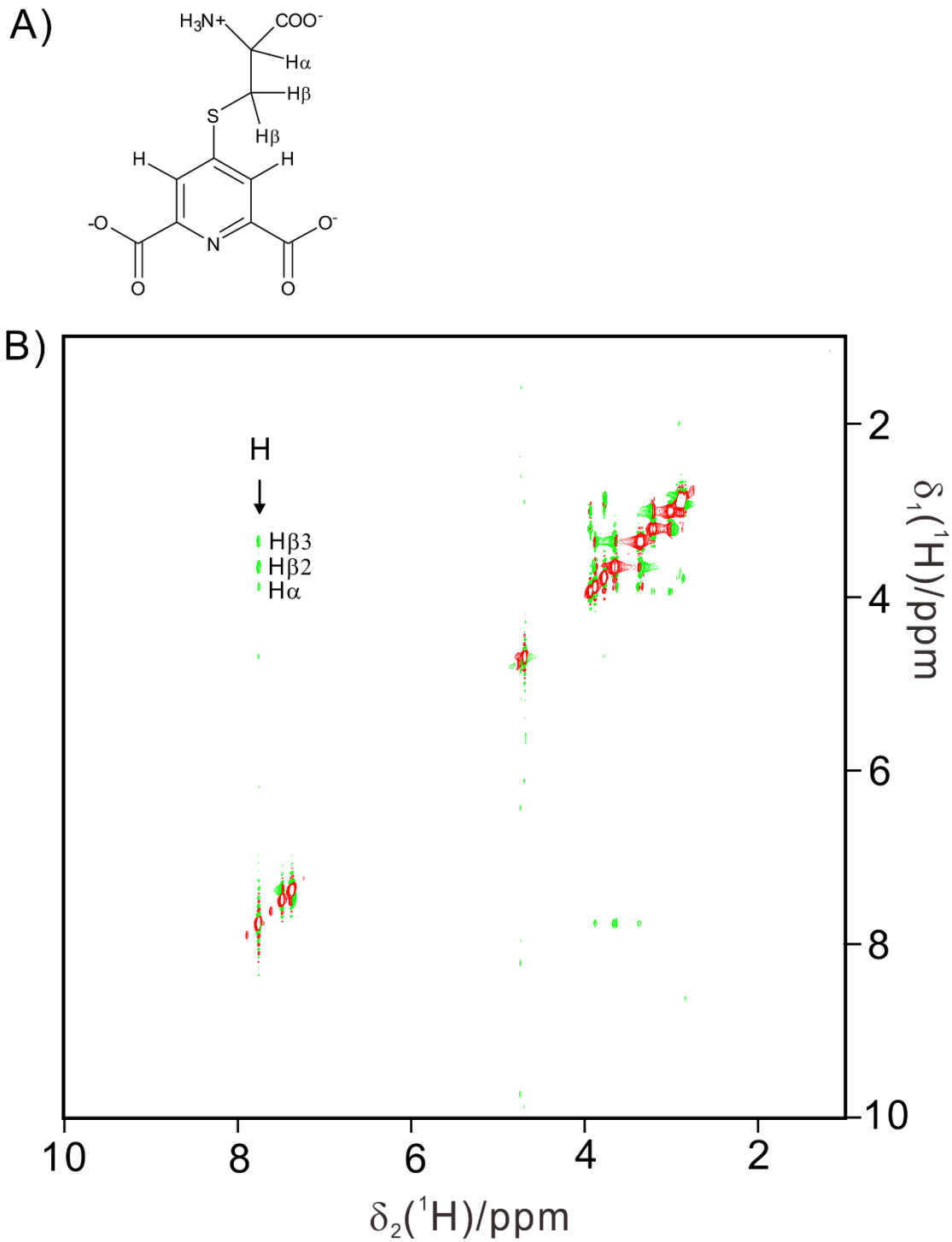


Fig. S5 (A) Structure of the ligation product from the reaction of **1** and L-cysteine. (B) 2D ^{1}H-^{1}H NOESY recorded for the mixture of 10 mM **1** after incubation with 30 mM L-cysteine for 5 hours. The NMR spectra were recorded in D_2O at pD 7.5 and 298 K with a proton frequency of 600 MHz and 300 ms mixing time was applied.

3. Phenylsulfonated pyridine derivative 1 in reaction with other amino acids

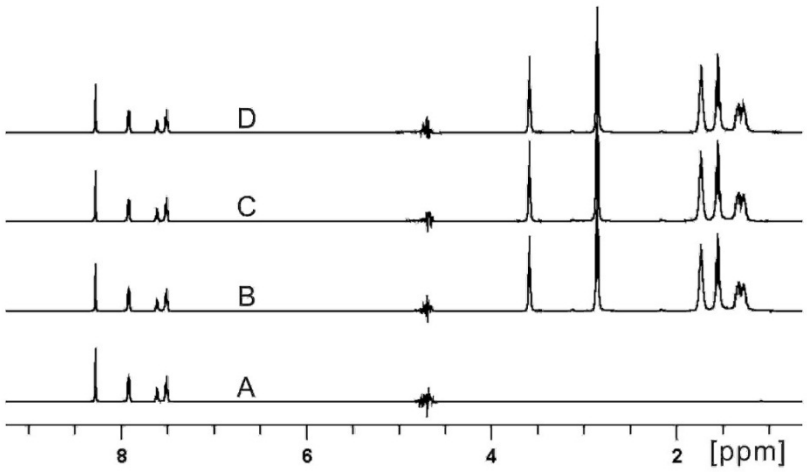


Figure S6. Reaction of **1** and L-lysine in 20 mM phosphate buffer at pH 7.5, which was monitored by 1D ^{1}H NMR spectroscopy. A) 1.0 mM **1**; B) 1.0 mM **1** and 6 mM L-lysine were mixed and incubation for 5 min; C) incubation of the mixture B at room temperature for 24 h; D) as in C for 48 h. The NMR spectra were recorded at 298 K with a proton frequency of 600 MHz.

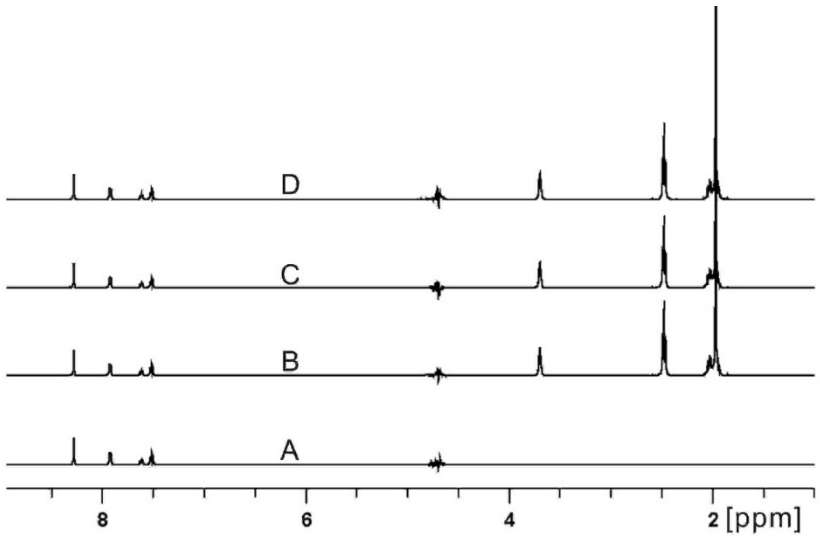


Fig. S7 Reaction of **1** and L-methionine in 20 mM phosphate buffer at pH 7.5, which was monitored by 1D ^{1}H NMR spectroscopy. A) 1.0 mM **1**; B) 1.0 mM **1** and 6 mM L-methionine were mixed and incubation for 5 min; C) incubation of the mixture B at room temperature for 24 h; D) as in C for 48 h. The NMR spectra were recorded at 298 K with a proton frequency of 600 MHz.

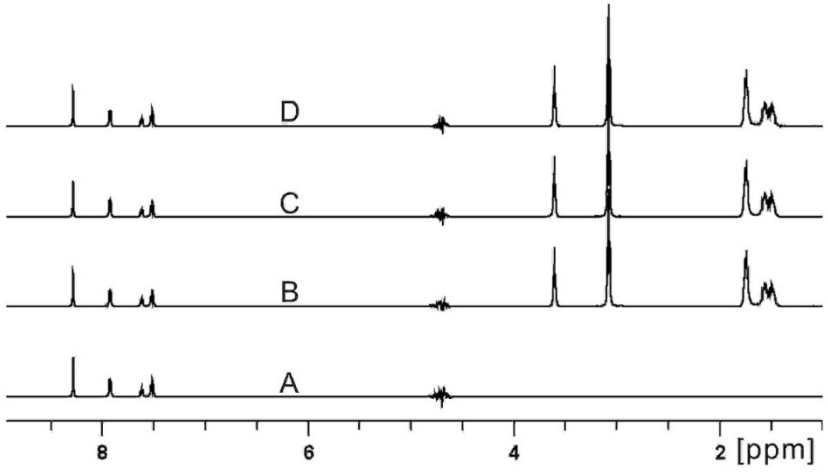


Fig. S8 Reaction of **1** and L-arginine in 20 mM phosphate buffer at pH 7.5, which was monitored by 1D ^{1}H NMR spectroscopy analysis. A) 1.0 mM **1**; B) 1.0 mM **1** and 6 mM L-arginine were mixed and incubation for 5 min; C) incubation of the mixture B at room temperature for 24 h; D) as in C for 48 h. The NMR spectra were recorded at 298 K with a proton frequency of 600 MHz.

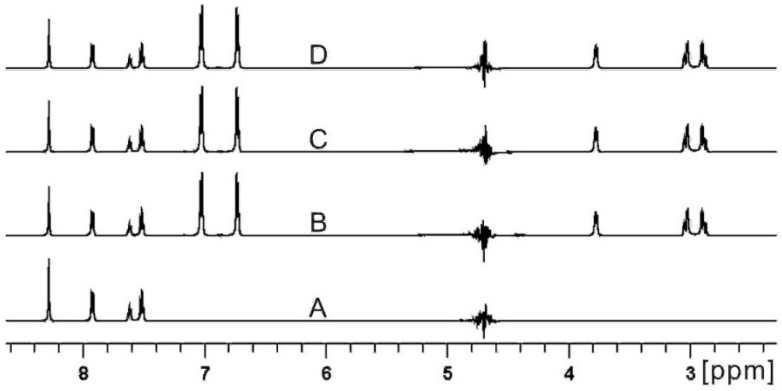


Fig. S9 Reaction of **1** and L-tyrosine in 20 mM phosphate buffer at pH 7.5, which was monitored by 1D ^{1}H NMR spectroscopy. A) 1.0 mM **1**; B) 1.0 mM **1** and 6 mM L-tyrosine were mixed and incubation for 5 min; C) incubation of the mixture B at room temperature for 24 h; D) as in C for 48 h. The NMR spectra were recorded at 298 K with a proton frequency of 600 MHz.

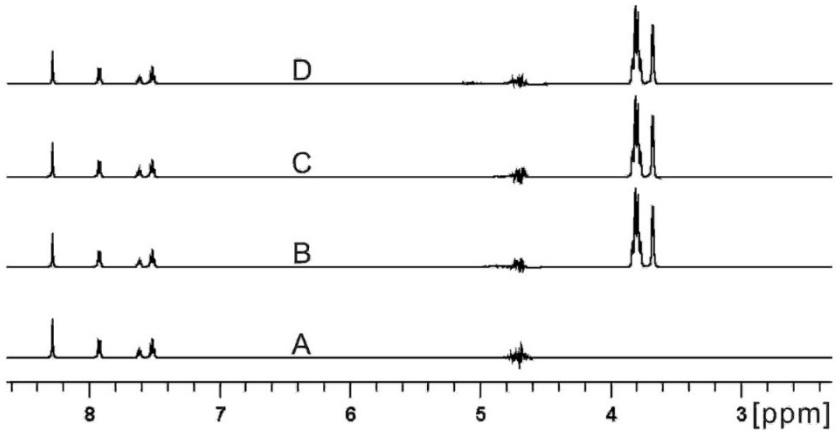


Fig. S10 Reaction of **1** and L-serine in 20 mM phosphate buffer at pH 7.5, which was monitored by 1D ^{1}H NMR spectroscopy analysis. A) 1.0 mM **1**; B) 1.0 mM **1** and 6 mM L-serine were mixed and incubation for 5 min; C) incubation of the mixture B at room temperature for 24 h; D) as in C for 48 h. The NMR spectra were recorded at 298 K with a proton frequency of 600 MHz.

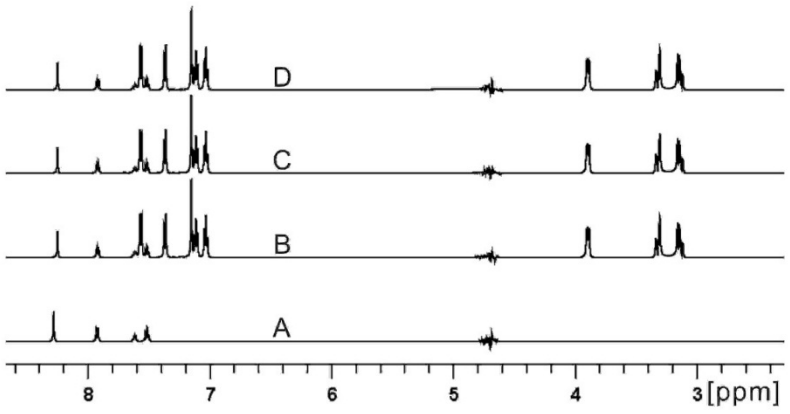


Fig. S11 Reaction of **1** and L-tryptophan in 20 mM phosphate buffer at pH 7.5, which was monitored by 1D ^{1}H NMR spectroscopy. A) 1.0 mM **1**; B) 1.0 mM **1** and 6 mM L-tryptophan were mixed and incubation for 5 min; C) incubation of the mixture B at room temperature for 24 h; D) as in C for 48 h. The NMR spectra were recorded at 298 K with a proton frequency of 600 MHz.

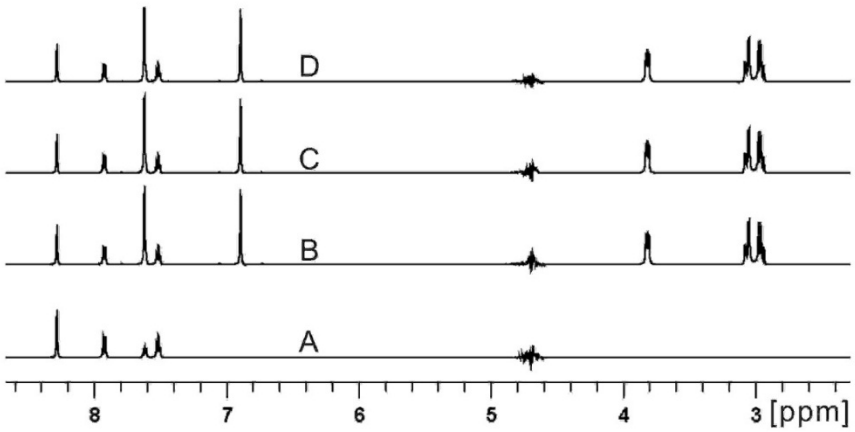


Fig. S12 Reaction of **1** and L-histidine in 20 mM phosphate buffer at pH 7.5, which was monitored by 1D ^{1}H NMR spectroscopy. A) 1.0 mm **1**; B) 1.0 mM **1** and 6 mM L-histidine were mixed and incubation for 5 min; C) incubation of the mixture B at room temperature for 24 h; D) as in C for 48 h. The NMR spectra were recorded at 298 K with a proton frequency of 600 MHz.

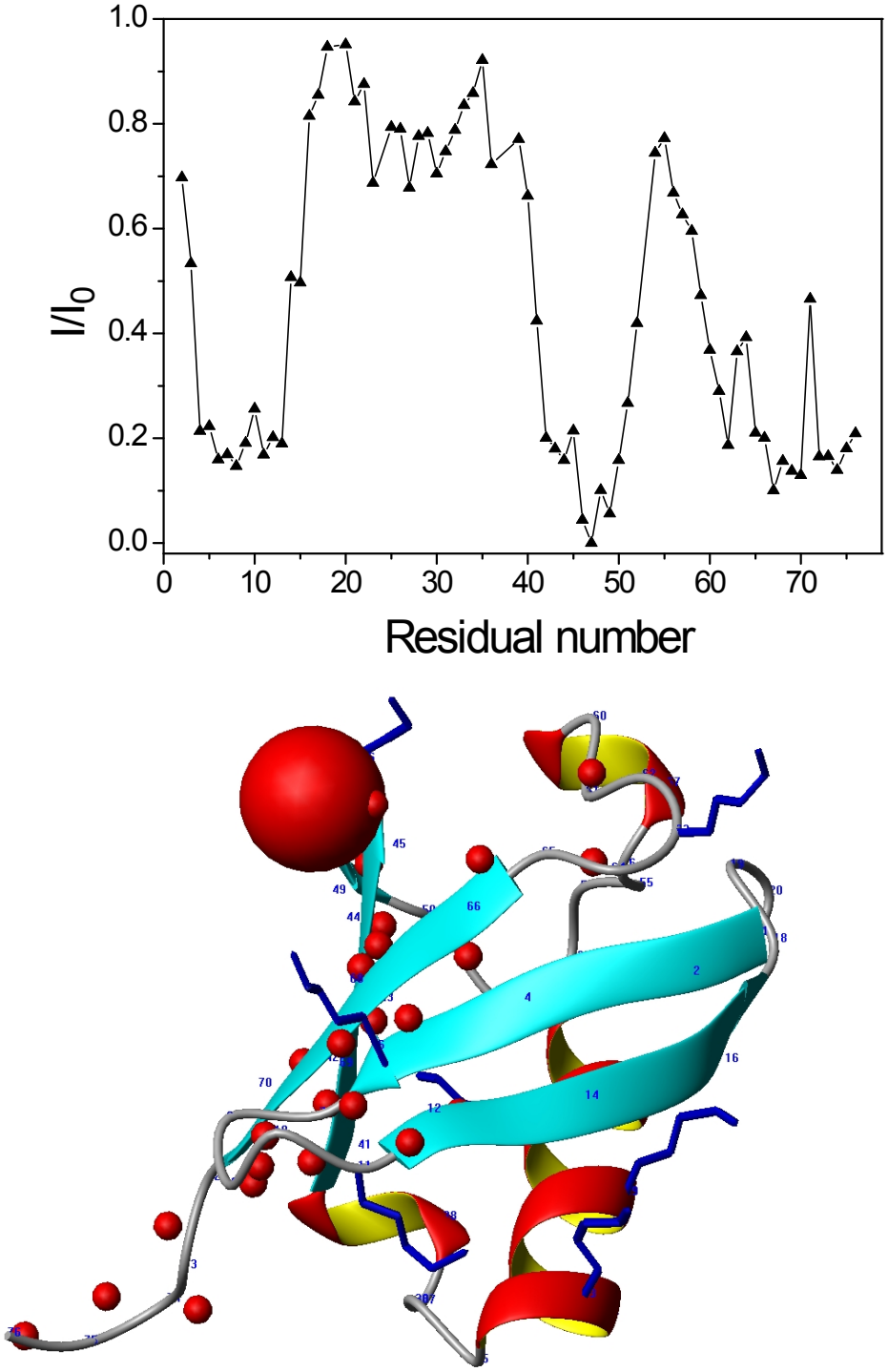


Fig. S13 The relative peak-intensity ratios (I_{para}/I_{dia}) in the ^{15}N-HSQC spectra of a 0.10 mM solution of ^{15}N-labeled G47C-L1 in the presence (I_{para}) or absence (I_{dia}) of one equivalent of Gd^{3+} (top panel). Ribbon drawing of the ubiquitin structure, of which the backbone amide protons with

$I_{para}/I_{dia} < 0.2$ were shown in read spheres and side-chains of Lys in blue sticks (bottom). Notably, high PREs are observed in the hydrophobic patch of ubiquitin, which plays a central role in the function of protein. Significant PREs are also found on the C-terminal residues similar to our previous observations on the interaction of ubiquitin and $[Gd(DPA)_3]^{3-}$,[2] which are not clearly understood.

4 Protein expression and purifications

All oligonucleotides used for mutagenesis were from Promega (China, Beijing). The plasmid for expression of the target protein was constructed using pET3a vector for expression under control of the T7 promoter. *E coli* BL21 (Rosetta) strain was used for protein expression. $^{15}NH_4Cl$ was purchased from Aldrich-Sigma ISO-TECH.

The protein samples were made by overexpression of the target protein using the high-cell-density method described previously.[3] The bacterial cells were grown at 37 °C in 1 L Luria broth rich medium shaken at 220 rpm. Once the optical cell density at 600 nm (OD_{600}) reached about 0.6, the cell culture was gently spun down at 3000 rpm for 5 min, and the cells were washed with MilliQ water and then resuspended in 250 mL of minimal M9 medium with $^{15}NH_4Cl$ as the only nitrogen source. The cells were allowed to recover by incubation at 37 °C for 20 min. Protein expression was subsequently induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a concentration of 0.5 mM. After about overnight incubation, the cells were harvested and stored at -20 °C. The cells were thawed and then lysated by sonication and purified by FPLC through anion/ion exchange column (GE Healthcare Biosciences) followed by G50 (GE Healthcare Biosciences) gel filtration. 20 mg of protein was usually obtained from 250 mL media.

4. General procedure for protein bioconjugation

0.5 mL of a 2.0 mM solution of ^{15}N-labeled protein in presence of 1.0 mM TCEP was diluted into 2 mL 20 mM Tris at pH 7.6 and five equivalents of phenylsulfonated pyridine derivative tag (50 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 2 to 16 h. Subsequently, the solution was concentrated to 1.0 mL. The ligation product was purified through either ion/anion exchange column or gel filtration, depending on the pI of protein. The overall yield of purified ligation product was usually above 75%. The ^{15}N-HSQC spectra were recorded in 20 mM MES at pH 6.5 and 298K with a proton frequency of 600 MHz equipped with QCI-cryoprobe.

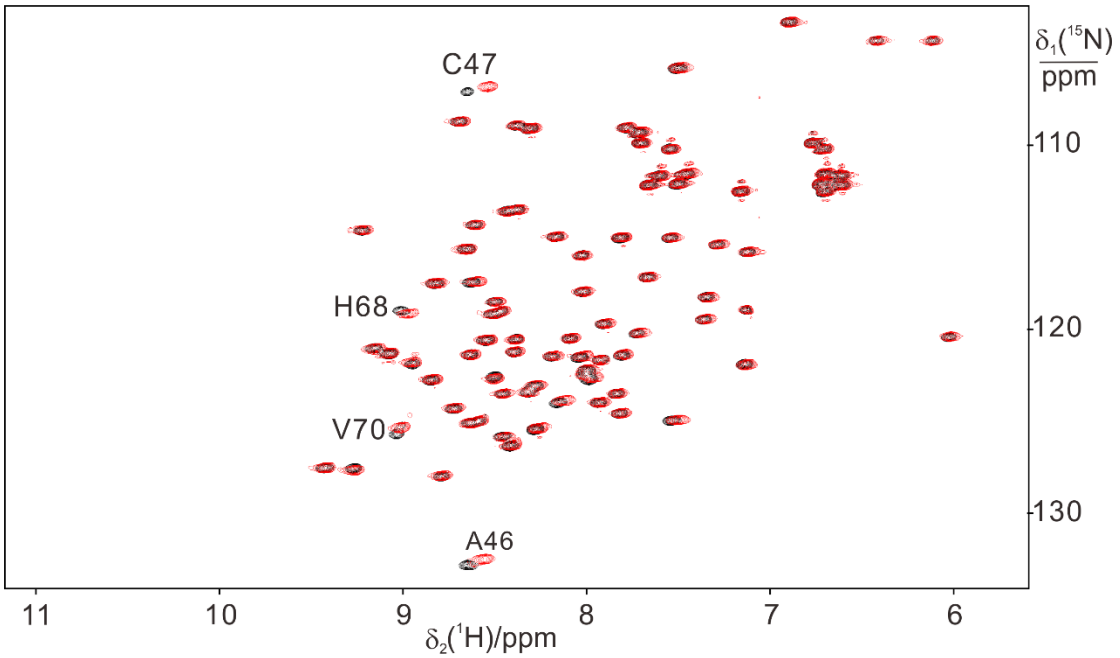
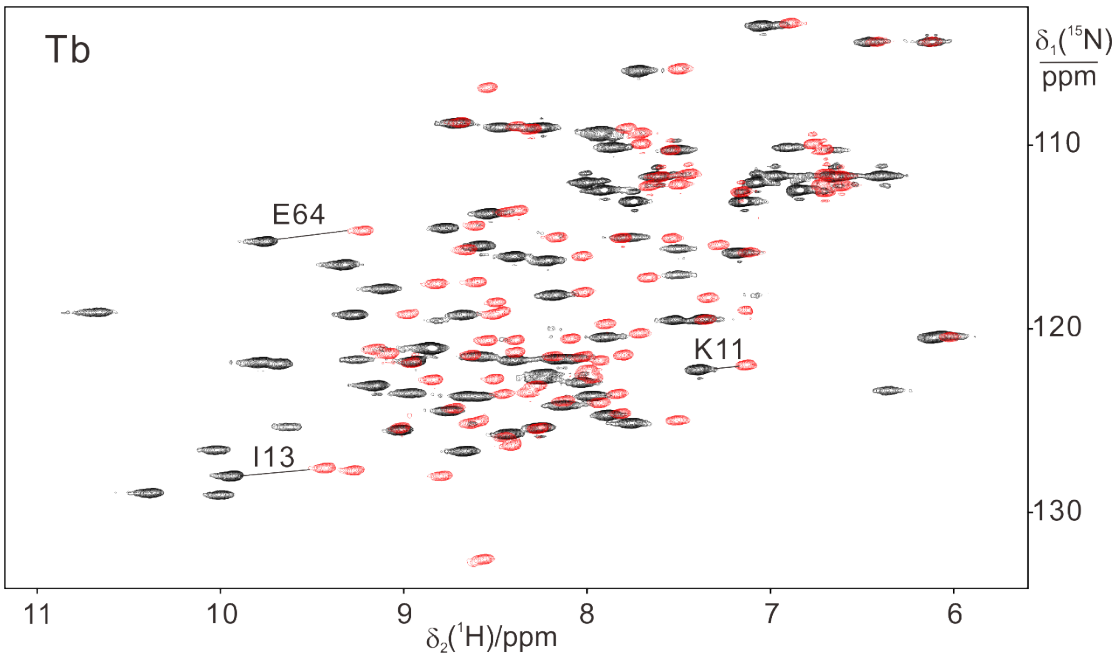


Fig. S14 Superimposition of ^{15}N-HSQC spectra of ^{15}N-labeled 0.1 mM ubiquitin G47C/R72A/R74A-L2 in the absence (black) and presence of 1.2 equivalents Y^{3+} (red). The spectra were recorded at 298 K in 20 mM MES buffer, pH 6.4



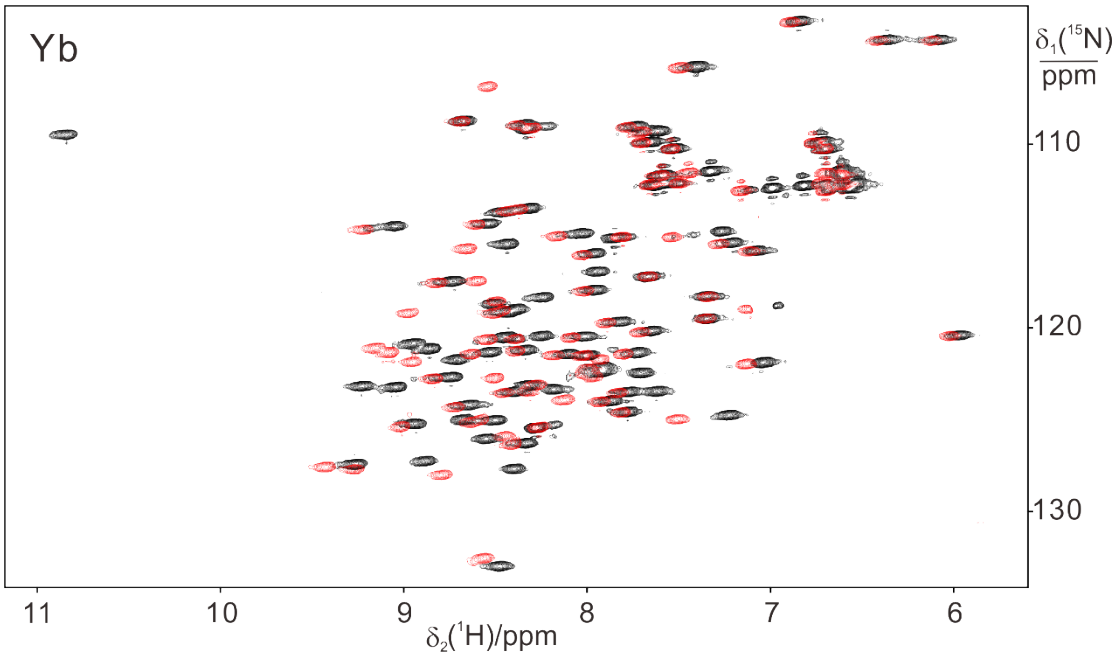


Fig. S15 Superimposition of ^{15}N-HSQC spectra of ^{15}N-labeled 0.1 mM ubiquitin G47C/R72A/R74A-L2 in the presence one equivalent Y^{3+} (red) and 1.2 equivalents paramagnetic lanthanides (black) Tb^{3+} or Yb^{3+}. The spectra were recorded at 298 K in 20 mM MES buffer, pH 6.4.

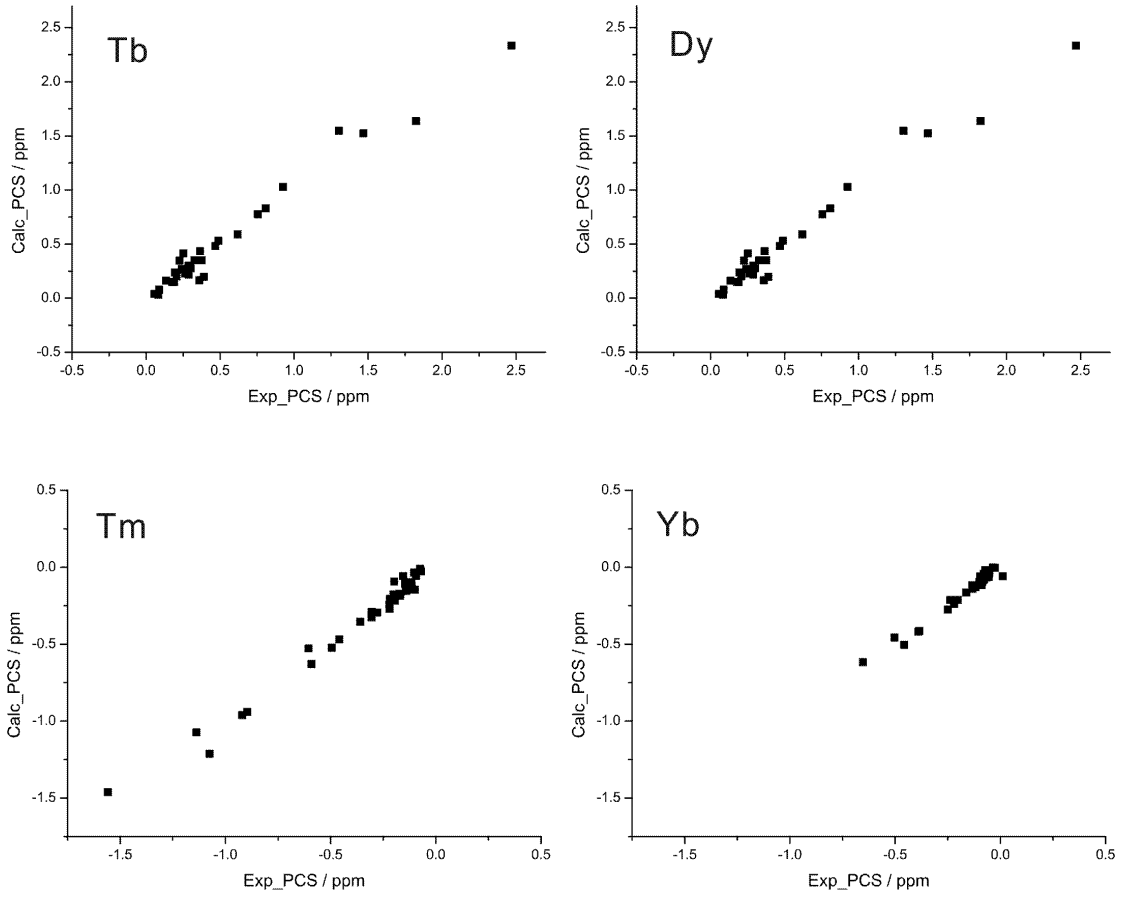


Fig. S16 Correlations between the back-calculated PCS values (Calc_PCS) and the experimentally measured PCS values (Exp_PCS) for a mixture of 0.10 mM G47C/R72A/R74A ubiquitin-L2 complexed with 1.2 equivalents of paramagnetic lanthanide ion.

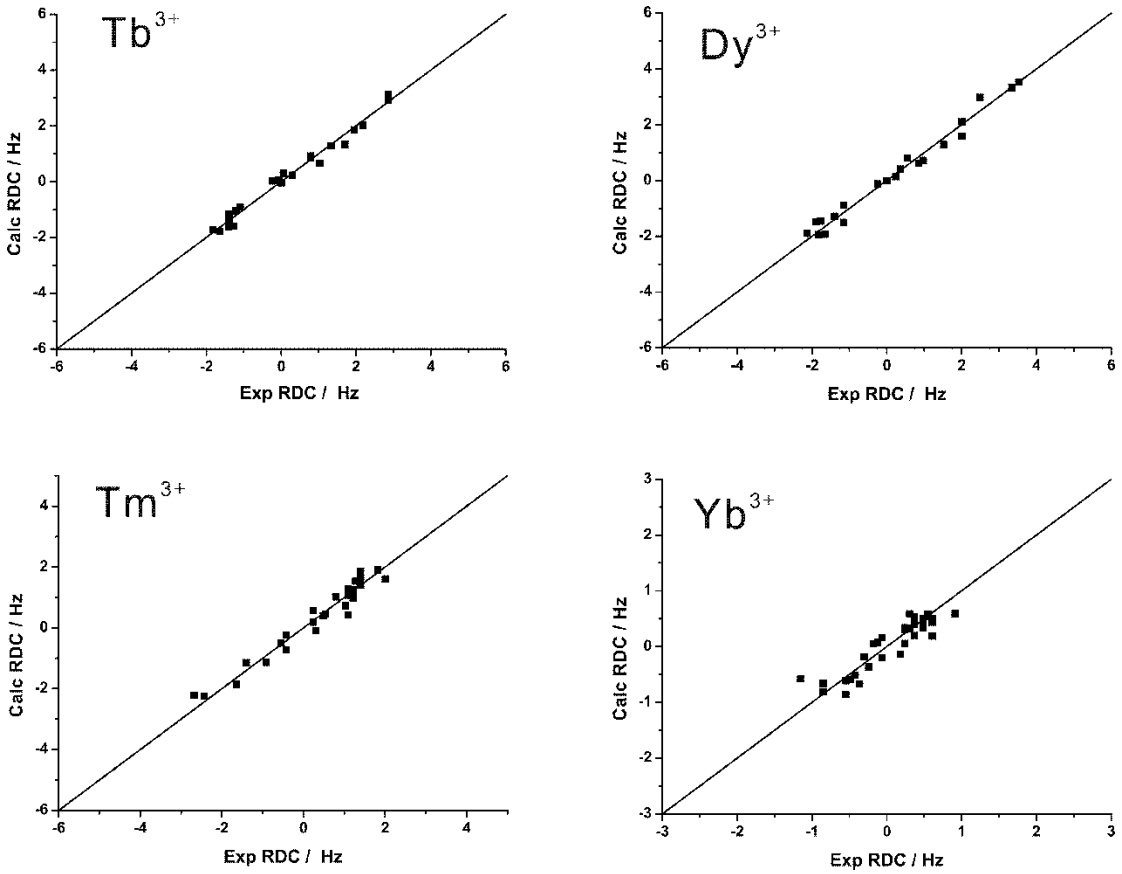


Fig. S17. Correlations of the back-calculated RDC (Calc RDC) plotted against the experimentally measured RDC (Exp RDC) for the G47C/R72A/R74A ubiquitin-L2 conjugate in complex with paramagnetic lanthanide ion, of which the Q-factors were determined.

References:

1. Takalo, H.; Kankare, J., *Acta Chemi. Scand. Ser. B* 1987, **41**, 219-221.

2. Wei, Z.; Yang, Y.; Li, Q. F.; Huang, F.; Zuo, H. H.; Su, X. C., *Chem. Eur. J.* 2013, **19**, 5758-5764.

3. Cao, C.; Chen, J. L., Yang, Y., Huang, F.; Otting, G., Su, X. C., *J. Biomol. NMR* 2014, **59**, 251-161.