

ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

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

Yin Yang[†], Jin-Tao Wang[†], Ying-Ying Pei and Xun-Cheng Su

Received (in XXX, XXX) Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX

DOI: 10.1039/b000000x

Increasing the stability of protein bioconjugates and improving the resolution of protein complexes is important for spectroscopic analysis in structural biology. The reaction of phenylsulfonated pyridine derivative and protein thiol generates a stable, rigid and short thioether tether, which is valuable in high-resolution spectroscopic measurements.

Site-specific labeling proteins with paramagnetic or fluorescent tags provide a powerful technique in determining the structures, interactions and dynamics of proteins and protein-ligand complexes by paramagnetic nuclear magnetic resonance (NMR),¹ electron paramagnetic resonance (EPR),² and fluorescent resonance energy transfer (FRET).³ Many efforts have been made towards high-yield selectivity in site-specific labeling of proteins with a paramagnetic or fluorescent tag including disulfide bond⁴ or thioether bond⁵ formation, or via click-chemistry between non-natural amino acid and functional tags.⁶ The flexibility of the functional tag in protein bioconjugate averages the paramagnetic effects or delocalizes the paramagnetic/fluorescent center, resulting in deteriorated resolution in the structures determined by these spectroscopic methods. Therefore, restriction of the paramagnetic or fluorescent tag is beneficial in improving structural resolution determined by means of paramagnetic NMR, EPR and FRET. Several strategies have been proposed to increase the rigidity of the tag including coordination of protein to the paramagnetic center, increasing the size of tag, and optimizing the ligation site in a protein.⁷ However, an efficient method in tagging a protein is of high importance for high-resolution spectroscopic analysis especially *in situ* / or *in vivo* in structural biology.^{5b,8} We herein present a new way of site-specific tagging proteins with a rigid, stable and short thioether ether via the reaction of phenylsulfonated pyridine derivatives and protein thiol (Scheme 1). The high performance of these new paramagnetic tags in rigidity and stability has been evaluated by paramagnetic NMR spectroscopy.

Phenylsulfone moiety is a valuable building block in organic synthesis, however, it has not been considered as a good leaving group in mild reaction conditions.⁹ To explore the reactivity of 4-phenylsulfonfyl pyridine tags to a free thiol in aqueous solution, we synthesized two lanthanide tags, L1 (4-(phenylsulfonfyl)-pyridine-2,6-dicarboxylic acid, 4PS-DPA) and L2 (4-phenylsulfonfyl-(pyridin-2,6-diyl)bismethylenenitrilo

tetrakis(acetic acid), 4PS-PyMTA) (Fig. 1, details in synthesis of L1 and L2 are provided in the supporting information). L1 and L2 are the phenylsulfonated derivatives of DPA and PyMTA, respectively. We show that phenylsulfonfyl group at the 4-position of pyridine is a thiol-specific reaction reagent in protein bioconjugations via the formation of thioether bond (Scheme 1), with a similar mechanism to a recent report of Juila-Kocienski like reagents with free thiols¹⁰. The reaction of L1 and L-cysteine was first analysed by high resolution NMR spectroscopy in D₂O at pD 7.5 with a proton frequency of 600 MHz. Significant chemical shift perturbations were observed for the resonances of 10 mM L1 after addition of 30 mM L-cysteine. By analysis of 2D NOESY, TOCSY, ¹³C-HSQC and ¹³C-HMBC spectra recorded for the reaction mixture and starting materials, we confirmed thiol ether bond was formed between L1 and L-cysteine, and phenylsulfinate was produced during the reaction (Fig. S1-S5). Notably, the nucleophilic attack to phenyl ring by releasing 4-sulfinatopyridine-2,6-dicarboxylate was not observed. Incubation of 4-(phenylsulfonfyl) pyridine moiety with other amino acids containing potential nucleophiles including hydroxyl and amine groups (up to 48 hours) produced no chemical shift perturbations, suggesting the phenylsulfonated pyridine is highly thiol-specific (Figure S6-S12).

Scheme 1 Site-specific tagging proteins in a short, rigid and stable manner via thiol-ether bond formation between 4-(phenylsulfonfyl) pyridine moiety and protein thiol.

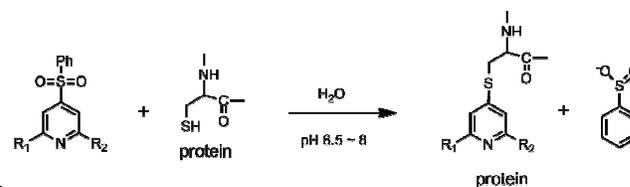


Fig. 1 Lanthanide binding tags of pyridine derivatives containing a phenylsulfone or vinyl group.

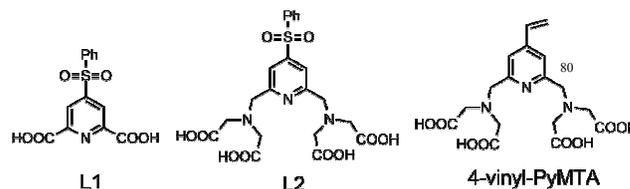
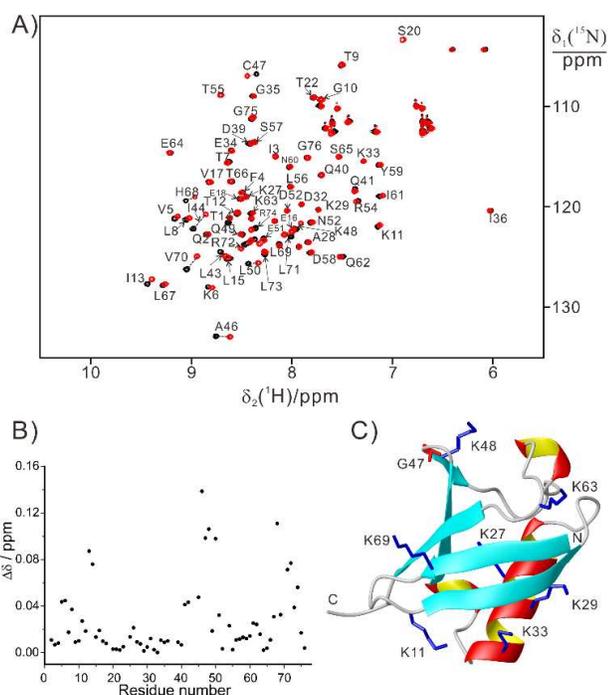
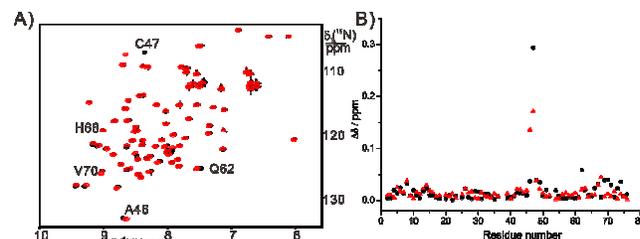


Fig. 2 A) Superimposition of ^{15}N -HSQC spectra of a 0.1 mM solution of uniformly ^{15}N -labeled G47C ubiquitin (black) and G47C-L1 conjugate (red). B) Chemical-shift differences between G47C and G47C-L1 conjugate, calculated as $\delta = ((\Delta\delta_{\text{H}})^2 + (\Delta\delta_{\text{N}}/10)^2)^{1/2}$. C) Ribbon drawing of human ubiquitin (PDB code:1UBI)¹¹. The mutation site of G47 was highlighted in red and side-chains of lysine residues in blue. The spectra were recorded at 298 K in 20 mM MES buffer, pH 6.4.



The high reactivity and chemoselectivity of 4-phenylsulfonyl pyridine moiety to free thiols as demonstrated above suggests the 4-phenylsulfonate pyridine can be a versatile building block in protein modifications. As the chemical environment for individual residues in structured biomolecules may affect the chemical reactivity and selectivity in protein bioconjugations, we thus undertook the selectivity assay of microenvironment effects on the reactions of phenylsulfonated pyridine derivatives towards protein thiols. The chemical reaction process was monitored by ^{15}N -HSQC spectra for a mixture of 1.0 mM ^{15}N -labelled G47C mutant of human ubiquitin and 5.0 mM L1 in 20 mM Tris buffer at pH 7.6 and 298 K. The reaction was completed within 10 hours. The chemical shift perturbations showed that the residues with large chemical shift changes were those vicinal to the ligation site of G47C, and no chemical modifications of Met 1, K6, K11, K27, K29, K33, K48 and K63 residing in different structural environments were determined in ^{15}N -HSQC spectrum (Figure 2). This was further demonstrated by addition of Gd^{3+} ion into the solution of G47C-L1 conjugate, which was purified after a small desalting PD10 column. Gd^{3+} is the strongest paramagnetic metal ion, and it generates very strong paramagnetic effects (PREs) on the vicinal nuclear spins closed to its binding area (Fig. S13).^{1ac} These results clearly show that phenylsulfonated pyridine moiety can be used as a thiol-specific reaction reagent in site-specific labelling of proteins.

Fig. 3 A) Superimposition of ^{15}N -HSQC spectra of a 0.1 mM solution of uniformly ^{15}N -labeled G47C/R72A/R74A ubiquitin (black) and G47C/R72A/R74A-L2 conjugate (red). The NMR spectra were recorded in the same condition as in Fig 2C. B) Chemical-shift differences between G47C/R72A/R74A and its L2 conjugate (black sphere), and protein-L2 and its Y^{3+} complex (triangle), respectively, calculated as $\delta = ((\Delta\delta_{\text{H}})^2 + (\Delta\delta_{\text{N}}/10)^2)^{1/2}$.



Since L1 has only three coordinating atoms for lanthanide ion, its protein conjugate requires the coordination of protein to immobilize the lanthanide ion. L2, a ligand with seven coordinating sites and very high-binding-affinity for lanthanide ion, was explored in the subsequent study. Quantitative ligation of G47C/R72A/R74A triple mutant of human ubiquitin with L2 was achieved by incubating the protein with a five-fold excess of L2 at room temperature and pH 7.6 for 16 hours.

Low-molecular weight compounds were subsequently separated from the ligated product by ion-exchange chromatography (yield > 85%). Small chemical shift perturbations close to G47C were observed after ligation with L2 (Fig. 3). The chemical shift mapping indicates only G47C was modified in the protein-L2 conjugate (also in Fig. S14). The resulting ligation product formally resembles the previously described ubiquitin G47C-4vinylPyMTA complex,^{5b} except that two methylene groups are absent. Titration of G47C/R72A/R74A-L2 with paramagnetic lanthanide ions produced very larger chemical shift changes in the ^{15}N -HSQC spectra (Fig. 4 and Fig. S15) in comparison with those obtained with G47C-4vinylPyMTA.^{5b} Despite the very short thioether tether in the protein construct, only a single paramagnetic species was observed as the molar ratio of lanthanide to protein was increased from 0.2 to 1.2. Large PCSs of up to -0.6 ppm were observed even in the presence of the weakly paramagnetic lanthanide Yb^{3+} . Four sets of PCSs were measured for Tb^{3+} , Dy^{3+} , Tm^{3+} , and Yb^{3+} , and the respective $\Delta\chi$ -tensor parameters were determined using the program Numbat¹² to fit the crystal structure of ubiquitin¹¹ (Table 1). Good correlations between the experimental and backcalculated PCSs were determined (Fig. S16). The metal position was found at a distance of 8.66 Å to carbon atom C^α of G47 (Fig. 5).

As show in Table 1, the $\Delta\chi$ -tensors determined in G47C/R72A/R74A-L2 are significantly larger than those obtained with G47C-4vinylPyMTA.^{5b} It indicates that the short thioether tether in protein-L2 conjugate is sufficiently rigid to restrict the paramagnetic tag. Both L2 and 4vinylPyMTA in reaction with protein generate a highly stable thioether tether, but L2, 4PS-PyMTA, is a more rigid paramagnetic tag. The magnetic susceptibility anisotropies

generated by protein-L2 complex are comparable with a number of rigid lanthanide binding tags like 4MMDPA^{7a}, TAHA^{7e}, LBT^{7fghk} and 4MTDA⁷ⁱ, but are generally smaller than DOTA-tags^{5d, 7bedj} that are also varied to the ligation site. The reaction of phenylsulfonated pyridine derivative and protein has several advantageous features: generation of a rigid and stable thioether tether between protein and functional tag, which is resistant to the reducing reagents, is simple; and it doesn't need to optimize the ligation site in protein. Moreover, 4PS-PyMTA is relatively small but has high binding affinity for lanthanide ion. As PyMTA has higher affinity for lanthanide ion than EDTA, the lanthanide complex of protein-L2 conjugate can be readily prepared by simply mixing the modified protein and lanthanide ion. Excess of lanthanide ion can be removed by addition of EDTA.

Fig. 4 Superimposition of ¹⁵N-HSQC spectra of uniformly ¹⁵N-labeled G47C/R72A/R74A ubiquitin derivatized with L2 at G47C in the presence of diamagnetic Y³⁺ (red) and paramagnetic Dy³⁺ or Tm³⁺ (black), respectively. The molar ratio of lanthanide to protein was about 1.2:1. The spectra were recorded at 298K and pH 6.4 with a ¹H NMR frequency of 600 MHz in 20 mM MES buffer.

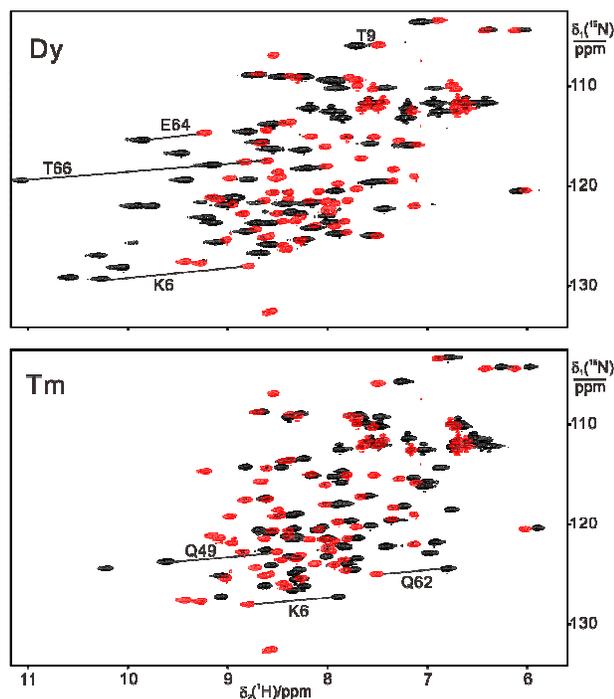


Table 1. $\Delta\chi$ -tensor parameters of ubiquitin G47C/R72A/R74A-L2 complexed with Tb³⁺, Dy³⁺, Tm³⁺, and Yb³⁺, respectively.^[a]

	Tb ³⁺	Dy ³⁺	Tm ³⁺	Yb ³⁺
$\Delta\chi_{ax}$	-21.5 (3.4)	-24.5 (4.3)	16.6 (1.7)	6.7 (0.9)
$\Delta\chi_{rh}$	-5.1 (0.9)	-7.1 (2.3)	3.9(1.1)	1.4 (0.4)
α	91.4	91.2	91.4	91.6
β	122.2	121.7	123.2	121.2
γ	113.9	109.9	116.8	112.8

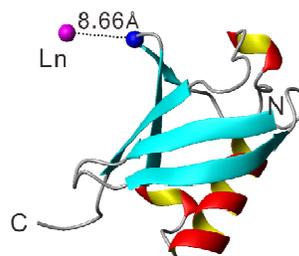
^[a] The tensor parameters are in units of 10^{-32} m^3 . The tensor parameters determined for the corresponding lanthanide complexes of ubiquitin G47C-4vinylPyMTA^{5b} are shown in brackets for comparison.

Table 2. Comparison of alignment tensor and $\Delta\chi$ -tensor parameters of ubiquitin G47C/R72A/R74A-L2 complexed with Tb³⁺, Dy³⁺, Tm³⁺, and Yb³⁺, respectively.^a

	$A_{ax} (\times 10^4)^b$	$A_{rh} (\times 10^4)^b$	$\Delta\chi_{ax} (\times 10^4)^c$	$\Delta\chi_{rh} (\times 10^4)^c$
Tb ³⁺	1.4	0.2	-5.6	-1.3
Dy ³⁺	1.6	0.2	-6.4	-1.9
Tm ³⁺	-1.2	-0.3	4.3	1.1
Yb ³⁺	-0.4	-0.1	1.7	0.4

^a All RDCs for protein backbone amide group were measured at 25 °C and 600 MHz ¹H NMR frequency. ^b Alignment tensors determined by RDCs of backbone amides in regularly structural segments. ^c A_{ax} and A_{rh} were obtained from $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ values in Table 1 by using equation (1).

Fig. 5 Ribbon drawing of ubiquitin showing the metal position (magenta sphere) determined by fitting the PCSs of the ubiquitin G47C/R72A/R74A-L2-lanthanide complexes to the ubiquitin structure (PDB ID: 1UBI),¹¹ and the distance to the C^α atom (blue sphere) of G47.



Using Module program,¹³ the alignment tensors were determined by simulation the backbone amide RDCs of regularly structured residues to the protein structure (PDB code: 1UBI). Notably, the alignment tensors determined by RDCs are significantly smaller than those back-calculated from PCSs based on $\Delta\chi$ -tensor using equation (1),

$$A_{ax,rh} = \frac{B_0^2}{15kT\mu_0} \Delta\chi_{ax,rh} \quad (1)$$

where $A_{ax, rh}$ are the axial and rhombic components of the alignment tensor, $\Delta\chi_{ax, rh}$ the axial and rhombic components of magnetic susceptibility anisotropy tensor, respectively (Table 2 and Fig. S17). The different sensitivities of PCSs and RDCs to protein local mobility, first reported in the K79A mutant of cytochrome c in an alkaline condition,¹⁴ have been found in many lanthanide loaded proteins.⁷ In the case of ubiquitin G47C-L2 conjugate, the much smaller RDCs are probably not only caused by the local mobility of protein. The residual mobility of the tag and non-specific transient contact of the tag with protein surface are also likely to average RDCs. It should be pointed out that the hydrophobic patch formed by residues Leu8, Ile44 and Val70 in ubiquitin dominates most interactions of ubiquitin with its binding partners.¹⁵ Since G47C sits close to the hydrophobic hot spot, any non-specific contact of L2 with these hydrophobic residues will decrease the RDCs greatly. The less sensitivity of PCSs to protein local mobility and probably also residual fluctuations of tag than RDCs offers additional advantages in

structural biology by paramagnetic NMR spectroscopy. The PCS defined rigidity of protein-L2 conjugate provides the possibility of high-resolution distance measurements by double electron-electron resonance (DEER) and FRET methods, which can be extended to high-molecule-weight systems.

In conclusion, we report a new way of generating a stable, short, and rigid thioether bridge between the target protein and functional tag, which is also resistant to reducing reagent, for paramagnetic NMR spectroscopy analysis. 4PS-PyMTA is a small and high-affinity lanthanide binding tag, and the protein ligation reaction is simple and in high yield. As pyridine is capable of binding lanthanide and transition metal ions, the 4-phenylsulfonated pyridine derivatives generates a shortest and stable tether in the protein conjugates for paramagnetic NMR and also for DEER and FRET analysis. The sulfonated pyridine moiety can also be engineered into the established paramagnetic ion binding tags including DOTA-derivatives or fluorescent tags (manuscript in preparation). With the high demand of simplicity and chemoselectivity in protein bioconjugation, and rigidity and stability in protein-tag conjugate for the study of DEER, FRET and paramagnetic NMR, we anticipate the phenylsulfonated pyridine moiety will find broad applications in structural biology and cell biology.

This work was supported by 973 program (2013CB910200) and the National Science Foundation of China (21473095, 21273121 and 21121002).

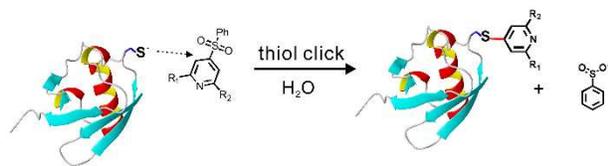
Notes and references

State-Key Laboratory of Elemento-organic Chemistry, Innovation Center of Chemical Science and Engineering (Tianjin), Nankai University, Tianjin, China 300071, E-mail: xunchengsu@nankai.edu.cn

[†]YY and JTW contributed equally in this work.

[†] Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. /

- [1] a) I. Bertini, C. Luchinat and G. Parigi, *Prog. NMR Spectr.*, 2002, **40**, 249-273; b) G. Otting, *J. Biomol. NMR*, 2008, **42**, 1-9; c) G. M. Clore and J. Iwahara, *Chem. Rev.*, 2009, **109**, 4108-4139; d) J. Koehler and J. Meiler, *Prog. NMR Spectr.*, 2011, **59**, 360-389; e) P. H. Keizers and M. Ubbink, *Prog. NMR Spectr.*, 2011, **58**, 88-96; f) C. Göbl, T. Madl, B. Simon and M. Sattler, *Prog. NMR Spectr.*, 2014, **80**, 26-63; g) I. Bertini, C. Luchinat, G. Parigi and P. Pierattelli, *Dalton Trans.*, 2008, 3782-3790; i) Y. Yang, C. J. Chen and X. C. Su, *Chin. J. Magn. Reson.*, 2014, **31**, 155-171.
- [2] a) D. Goldfarb, *Phys. Chem. Chem. Phys.*, 2014, **16**, 9685-9699; b) W. L. Hubbell, C. J. López, C. Altenbach and Z. Yang, *Curr. Opin. Struct. Biol.*, 2013, **23**, 725-733; c) G. Jeschke, *Annu. Rev. Phys. Chem.*, 2012, **63**, 419-446; d) E. J. Hustedt and A. H. Beth, *Annu. Rev. Biophys. Biomol. Struct.*, 1999, **28**, 129-153; e) M. Ji, S. Ruthstein and S. Saxena, *Acc. Chem. Res.*, 2014, **18**, 688-695.
- [3] a) K. N. Allen and B. Imperiali, *Curr. Opin. Chem. Biol.*, 2010, **14**, 247-254; b) P. R. Selvin, *Methods Enzymol.*, 2002, **246**, 300-334; c) D. J. Posson, *Neuron*, 2008, **59**, 98-109.
- [4] a) G. E. Fanucci and D. S. Cafiso, *Curr. Opin. Struct. Biol.*, 2006, **16**, 644-653; b) F. Rodriguez-Castañeda, P. Haberz, A. Leonov and C. Griesinger, *Magn. Reson. Chem.*, 2006, **44**, S10-16; c) X. C. Su and G. Otting, *J. Biomol. NMR*, 2010, **46**, 101-112.
- [5] a) Q. F. Li, Y. Yang, A. Maleckis, G. Otting and X. C. Su, *Chem. Comm.*, 2012, **48**, 2704-2706; b) Y. Yang, Q. F. Li, C. Cao, F. Huang and X. C. Su, *Chem. Eur. J.*, 2013, **19**, 1097-1103; c) K. L. Gempf, S. J. Butler, A. M. Funk and D. Parker, *Chem. Comm.*, 2013, **49**, 9104-9106; d) W. M. Liu, S. P. Skinner, M. Timmer, A. Blok, M. A. Hass, D. V. Filippov, M. Overhand and M. Ubbink, *Chem. Eur. J.*, 2014, **20**, 6256-6258.
- [6] C. T. Loh, K. Ozawa, K. L. Tuck, N. Barlow, T. Huber, G. Otting and B. Graham, *Bioconjug. Chem.*, 2013, **24**, 260-268.
- [7] a) X. C. Su, B. Man, S. Beeren, H. Liang, S. Simonsen, C. Schmitz, T. Huber, B. A. Messerle and G. Otting, *J. Am. Chem. Soc.*, 2008, **130**, 10486-10487; b) P. H. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2008, **130**, 14802-14812; c) D. Häussinger, J. R. Huang and S. Grzesiek, *J. Am. Chem. Soc.*, 2009, **131**, 14761-14767; d) B. Graham, C. T. Loh, J. D. Swarbrick, P. Ung, J. Shin, H. Yagi, X. Jia, S. Chhabra, N. Barlow, G. Pintacuda, T. Huber and G. Otting, *Bioconjug. Chem.*, 2011, **22**, 2118-2125; e) F. Peters, M. Maestre-Martinez, A. Leonov, L. Kovačič, S. Becker, R. Boelens and C. Griesinger, *J. Biomol. NMR*, 2011, **51**, 329-337; f) L. J. Martin, M. J. Hähnke, M. Nitz, J. Wöhnert, N. R. Silvaggi, K. N. Allen, H. Schwalbe and B. Imperiali, *J. Am. Chem. Soc.*, 2007, **129**, 7106-7113; g) K. Barthelme, A. M. Reynolds, E. Peisach, H. R. Jonker, N. J. DeNunzio, A. N. Allen, B. Imperiali and H. Schwalbe, *J. Am. Chem. Soc.*, 2011, **133**, 808-819; h) T. Saio, K. Ogura, M. Yokochi, Y. Kobashigawa and F. Inagaki, *J. Biomol. NMR*, 2009, **44**, 157-166; i) F. Huang, Y. Y. Pei, H. H. Zuo, J. L. Chen, Y. Yang and X. C. Su, *Eur. J. Chem.*, 2013, **19**, 17141-17149; j) L. Cerofolini, G. B. Fields, M. Fragai, C. F. Geraldes, C. Luchinat, G. Parigi, E. Ravera, D. I. Svergun and J. M. Teixeira, *J. Biol. Chem.*, 2013, **288**, 30659-30671; k) X. C. Su, K. McAndrew, T. Huber and G. Otting, *J. Am. Chem. Soc.*, 2008, **130**, 1681-1687.
- [8] A. Martorana, G. Bellapadrona, A. Feintuch, E. Di Gregorio, S. Aime and D. Goldfarb, *J. Am. Chem. Soc.*, 2014, **136**, 13458-13465.
- [9] A. El-Awa, M. N. Noshi, X. M. du Jourdin and P. L. Fuchs, *Chem. Rev.*, 2009, **109**, 2315-2349.
- [10] a) D. Zhang, N. O. Devarie-Baez, Q. Li, J. R. Lancaster Jr. and M. Xian, *Org. Lett.*, 2012, **14**, 3396-3399; b) N. Toda, S. Asano and C. F. Barbas III, *Angew. Chem. Int. Ed. Engl.*, 2013, **52**, 12592-12596.
- [11] R. Ramage, J. Green, T. W. Muir, O. M. Ogunjobi, S. Love and K. Shaw, *Biochem. J.*, 1994, **299**, 151-158.
- [12] C. Schmitz, M. J. Stanton-Cook, X. C. Su, G. Otting and T. Huber, *J. Biomol. NMR*, 2008, **41**, 179-189.
- [13] P. Dosset, J. C. Hus, D. Marion and M. Blackledge, *J. Biomol. NMR*, 2001, **20**, 223-231.
- [14] M. Assfalg, I. Bertini, P. Turano, A. G. Mauk, J. R. Winkler and H. B. Gray, *Biophys. J.*, 2003, **84**, 3917-3923.
- [15] I. Dikic, S. Wakatsuki and K. J. Walker, *Nat. Rev. Mol. Cell. Biol.*, 2009, **10**, 659-671.



The reaction of phenylsulfonated pyridine derivative and protein thiol is suitable for high-resolution spectroscopic analysis by generation of a rigid, stable and short thioether tether.

⁵ TOC