ChemComm

COMMUNICATION

View Article Online
View Journal

Cite this: DOI: 10.1039/c7cc06155b

Received 7th August 2017,
Accepted 5th September 2017

DOI: 10.1039/c7cc06155b

rsc.li/chemcomm

Site-selective tagging of proteins by pnictogen-mediated self-assembly†

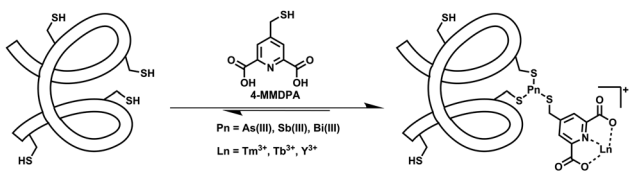
Christoph Nitsche, * Mithun C. Mahawaththa, Walter Becker, ‡ Thomas Huber and Gottfried Otting *

Site-selective chemical protein modification is achieved by self-assembly of a specific di-cysteine motif, trivalent pnictogens (As, Sb or Bi) and an aromatic mercaptomethyl-based probe. The strategy is demonstrated with a quaternary complex involving Zika virus protease and a lanthanide ion, enabling paramagnetic nuclear magnetic resonance spectroscopy and luminescence measurements.

Selective protein tagging strategies are critically important for probing their function, cellular localisation and dynamics. It is difficult, however, to achieve this in a site-specific and bio-orthogonal manner, as is often required to avoid impeding the natural function of the protein in an uncontrolled manner.[1] N- or C-terminal fusions, *e.g.* with green fluorescent protein,[2] tend to alter the biophysical properties of the target protein significantly.[3] Cycloaddition reactions, often referred to as 'click' chemistry, require the selective incorporation of costly unnatural amino acids and depend on the availability of a suitable overexpression system for the protein of interest.[4] Using a single solvent-exposed cysteine residue to attach a tag by formation of a disulfide or thioether linkage usually requires the prior mutation of all other naturally occurring cysteine residues that may compete in the reaction,[5] even though context-dependent reactivity of cysteine residues can sometimes be exploited for site-selective tagging.[6] For many proteins, so many mutations are detrimental to their structure and function.[7]

The present work describes a simple strategy whereby the chemical probe is attached to two cysteine residues by preferential and reversible binding, without interference by a lone cysteine residue. Our method is based on the unique thiophilic characteristics of trivalent pnictogens (As, Sb or Bi). As(III) is well-known to prefer binding to multiple over single thiol groups.[8] For example, biarsenical fluorescent probes, such as FlAsH-EDT$_2$ or ReAsH-EDT$_2$, selectively bind to CCXXCC peptide motifs.[9] Recently, monoarsenical probes that bind to two vicinal cysteines have been used to link protein and peptide chains or detect vicinal dithiol-containing proteins (VDPs) in cells.[10] Unfortunately, the chemical synthesis of the requisite organo-arsenic compounds is challenging and fraught with toxicity concerns, which severely limits broader availability and applicability. While the heavier pnictogens such as antimony or bismuth share the thiophilic character of arsenic and are less toxic, organic antimony and bismuth compounds are even more challenging to synthesise and are unlikely to be sufficiently stable in aqueous solution. Here we report a strategy for pnictogen-mediated tagging of VDPs without requiring the chemical synthesis of organic pnictogen tags, which greatly expands the scope of suitable probes towards readily available thiol compounds.

Our strategy involves the spontaneous attachment of trivalent inorganic arsenic, antimony or bismuth to two vicinal cysteine residues of the protein followed by addition of a thiol-group containing probe to form a covalent thiolate complex (Scheme 1). The ternary complex between VDP, tag and pnictogen selectively self-assembles in aqueous solution. In the present work, we added, in addition, a lanthanide ion to confirm the location of the tag on the protein and explore the chemical stability of the



Scheme 1 Pnictogen-mediated assembly of the lanthanide binding tag 4-MMDPA and a protein that contains solvent-exposed cysteine residues in addition to an engineered binding site with two vicinal thiols.

Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia. E-mail: christoph.nitsche@anu.edu.au, gottfried.otting@anu.edu.au

† Electronic supplementary information (ESI) available: Detailed procedures and conditions for protein expression and purification, sample preparation, ITC, MS and PAGE experiments; details of computational design and molecular modelling; tables of PCSs and fitted $\Delta\chi$ tensor parameters; correlations between back-calculated and measured PCSs; [^{15}N,^{1}H]-HSQC spectra of ZiPro P101C/P131C in presence and absence of various reagents; ITC data; mass spectrometry data. See DOI: 10.1039/c7cc06155b

‡ Present address: Institute of Chemistry, Karl Franzens Universität Graz, 8010 Graz, Austria.

probe by paramagnetic NMR spectroscopy.[11] The approach is demonstrated with the Zika virus NS2B-NS3 protease (ZiPro), which is an established drug target. The wild-type protein contains two solvent-exposed cysteine residues.[12]

Self-assembly of pnictogen thiolates has been described previously for all pnictogens except nitrogen in supramolecular chemistry protocols conducted in organic solvents.[13] The driving force is the formation of labile covalent bonds between sulfur and Pn(III), and additional pnictogen–π interactions have been proposed to stabilise the complex in the case of mercaptomethylaryl-based ligands.[14] Consequently, we chose a simple paramagnetic lanthanide-binding tag with a mercaptomethylaryl moiety, 4-mercaptomethyl-dipicolinic acid (4-MMDPA; Scheme 1) to favour binding of the probe to the pnictogen–VDP assembly and outcompete single cysteine residues and other thiol reagents such as glutathione, which are aliphatic.[15] Suitable sites for pnictogen installation were predicted computationally, starting from the crystal structure of the Zika virus protease (PDB ID: 5LC0)[12] and mutating residues to cysteines, then checking whether standard rotamers could form bonds with an arsenic atom with close to 90° S–As–S bond angle. This identified the double-mutation P101C/P131C as a possible site near the active site of the protease (Fig. 1).

Addition of $NaAsO_2$ in water, or $SbCl_3$ or $BiBr_3$ in DMSO to uniformly ^{15}N-labelled ZiPro P101C/P131C indeed yielded new cross-peaks in the [^{15}N,^{1}H]-HSQC spectrum, indicating slow chemical exchange between bound and free protein, as expected for the formation of covalent-reversible bonds (Fig. S4, ESI†).

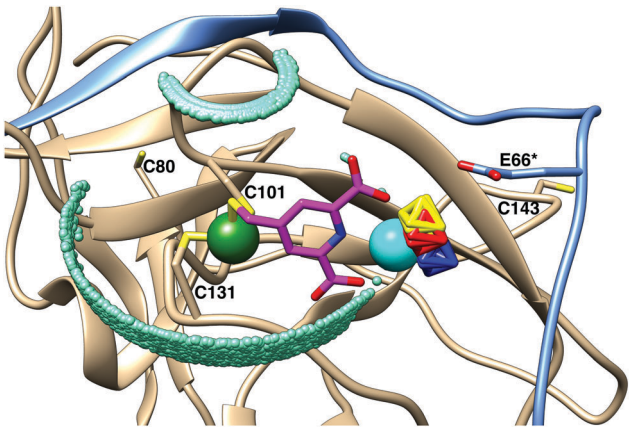


Fig. 1 Model of the self-assembled complex between ZiPro P101C/P131C, a trivalent pnictogen (green), 4-MMDPA (magenta) and lanthanide (cyan), highlighting bonds with nitrogen in blue, oxygen in red, and sulfur in yellow. Small cyan balls identify all sterically allowed metal positions determined by modelling the protein–pnictogen–4-MMDPA–lanthanide complex with variable dihedral angles of the 4-MMDPA moiety. The lanthanide coordinates found by NMR are shown as octahedrons for the complexes assembled with arsenic (yellow), antimony (red), and bismuth (blue). These metal positions were determined using pseudocontact shifts measured with Tm^{3+} and Tb^{3+} to fit magnetic susceptibility anisotropy ($\Delta\chi$) tensors to the crystal structure (PDB ID: 5LC0)[12] of the protein (Table S2, ESI†). NS3 is shown in beige and the NS2B cofactor in light blue. The carboxyl group of Glu 66* (where the star identifies residues of NS2B) is near the metal site and may assist in positioning the lanthanide ion. Cys 80 and Cys 143 are part of the wild-type amino acid sequence of ZiPro.

Isothermal titration calorimetry (ITC) with As(III) indicated a dissociation constant K_d of about 260 μM (Fig. S6, ESI†).

Subsequent NMR-titrations of the pnictogen–protease complexes with 4-MMDPA also indicated slow chemical exchange for the association and dissociation of the ternary complex (data not shown). Mass spectrometry confirmed the formation of the ZiPro–As(III)–4-MMDPA assembly in equilibrium with free ZiPro (Fig. S8, ESI†). ITC-titration of a 1 : 1 ZiPro–As(III) complex with 4-MMDPA indicated a K_d value of approximately 170 μM (Fig. S7, ESI†). Despite the fairly large dissociation constants, the kinetics of the ternary self-assembly are slow. Whereas each titration step in the ITC experiment took about 1 minute for thermal re-equilibration when forming the ZiPro–As(III) complex (Fig. S6, ESI†), about 10 minute intervals were needed for re-equilibration upon stepwise addition of 4-MMDPA (Fig. S7, ESI†). Similarly slow equilibration times (2.5 minutes) have been reported for the organoarsenical fluorescent probe FAsH, which similarly binds to two vicinal cysteine residues.[10b] In the case of bismuth, the samples showed pronounced precipitation during longer NMR measurements at 25 °C, presumably by the formation of insoluble Bi_2O_3. A dimeric coordination complex of Bi(III) with dipicolinic acid has been described[16] and a corresponding complex with 4-MMDPA may assist the disassembly of the complex in the absence of lanthanide.

Addition of lanthanide ions produced the quaternary complex of ZiPro P101C/P131C, pnictogen, 4-MMDPA and lanthanide, with the paramagnetic lanthanide ions Tm^{3+} and Tb^{3+} generating pseudocontact shifts (PCS) of opposite sign in the [^{15}N,^{1}H]-HSQC spectrum (Fig. 2 and Fig. S2, S3, ESI†). The chemical exchange between the paramagnetic quaternary complex and the dissociated diamagnetic species was slow on the NMR time scale for all three pnictogens, with the equilibrium favouring the fully assembled complex for arsenic and, even more so, antimony (Fig. 2B). The latter assembly showed the highest degree of association in our experiments. In contrast, the complex with bismuth appeared to be more dissociated, but it is unclear how far this was a consequence of limited solubility or lower affinity. To determine the metal position, we measured the PCSs of well-resolved [^{15}N,^{1}H]-HSQC cross-peaks of 29 residues in the protease generated by Tm^{3+} and Tb^{3+} (Table S1, ESI†) and fitted $\Delta\chi$ tensors for all pnictogen complexes to the crystal structure of the protein (Table S2, ESI†).[17] Excellent correlations between experimental and back-calculated PCSs were obtained with quality factors between 0.05 and 0.06 (Fig. S1 and Table S2, ESI†). This indicates that the binding of pnictogen, tag or lanthanide caused no global distortion of the protein structure. The magnitudes of the axial $\Delta\chi$ tensor components (up to 20×10^{-32} m^3) were significantly larger than those observed previously for 4-MMDPA attached to cysteine or other DPA derivatives that form shorter linkers with cysteine residues, such as 4-mercapto-DPA (4-MDPA) or 3-mercapto-DPA (3-MDPA),[15,18] indicating good immobilisation of the lanthanide ions. Fig. 1 shows the positions of the metal ions obtained by the $\Delta\chi$ tensor determinations. Comprehensive modelling of the range of tag conformations compatible with the pnictogen coordination by Cys 101 and Cys 131 indicated two arcs of localisation spaces, where the lanthanide ion could, in principle, be located (Fig. 1).

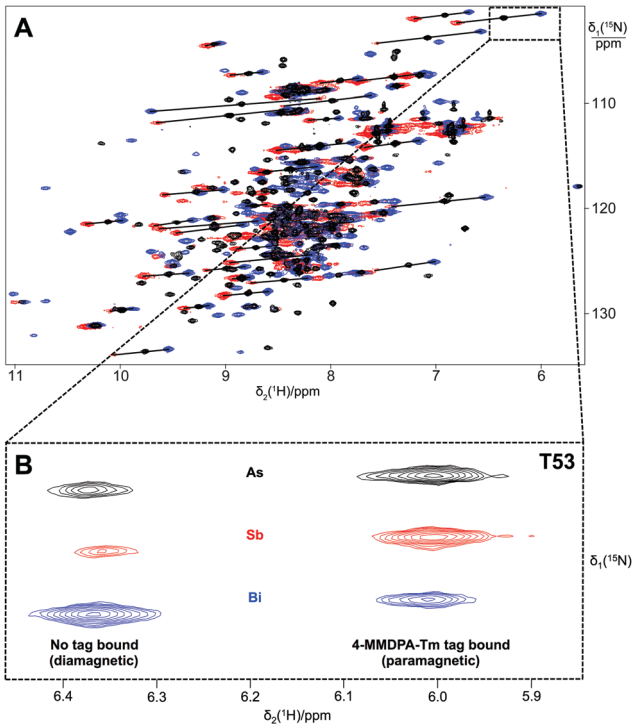


Fig. 2 Superimposition of [^{15}N,^{1}H]-HSQC spectra of 0.3 mM solutions of ZiPro P101C/P131C in the presence of 0.3 mM $SbCl_3$, 4-MMDPA and paramagnetic Tm^{3+} (blue), Tb^{3+} (red) or diamagnetic Y^{3+} (black). The spectra were recorded at 25 °C in 20 mM MES pH 6.5, 150 mM NaCl, 1 mM TCEP on a Bruker 800 MHz NMR spectrometer. (A) Overview of the spectra. Black lines connect corresponding cross-peaks observed with diamagnetic and paramagnetic lanthanides. (B) Selected spectral region showing the cross-peak of the backbone amide of Thr 53 in the complexes assembled with $NaAsO_2$ (black), $SbCl_3$ (red) or $BiBr_3$ (blue), in the presence of 0.3 mM 4-MMDPA and Tm^{3+}. The persistence of diamagnetic cross-peaks illustrates the equilibrium between protein with and without probe molecules, which is most strongly shifted towards the complex for the sample assembled with antimony. The spectra were shifted in the δ_1 dimension for improved visibility.

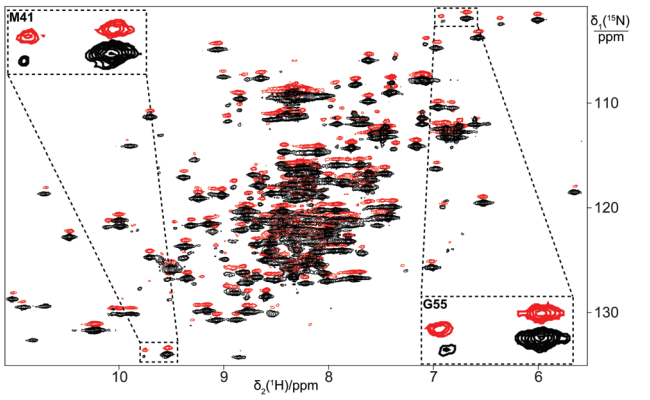


Fig. 3 Superimposition of [^{15}N,^{1}H]-HSQC spectra of 0.3 mM solutions of ZiPro P101C/P131C, $SbCl_3$, 4-MMDPA and paramagnetic Tm^{3+}, in the absence (black) and presence (red) of 1 mM glutathione (GSH). The cross-peaks of Met 41 and Gly 55 illustrate the persistence of the paramagnetic peaks in the presence of excess GSH. For improved visual representation, the black spectrum was shifted in the δ_1 dimension by 0.6 ppm.

Proximity of the metal to the side chain carboxyl group of Glu 66* suggests that coordination to this side chain helps immobilise the lanthanide ion.

Initially, we performed these studies with a ZiPro mutant, where the solvent-exposed residues Cys 80 and Cys 143 of the wild-type protein were substituted by serine to avoid any potential competition for pnictogen binding. This proved to be unnecessary, as the [^{15}N,^{1}H]-HSQC cross-peaks of Cys 80 and Cys 143 showed no chemical shift perturbations in the presence of equimolar trivalent antimony (Fig. S5A, ESI†). Furthermore, the quaternary complex with Sb(III) proved to be inert towards additional thiol compounds, as demonstrated by [^{15}N,^{1}H]-HSQC spectra recorded in the presence of glutathione (GSH). Again, the [^{15}N,^{1}H]-HSQC cross-peaks of Cys 80 and Cys 143 remained unperturbed (Fig. S5B, ESI†), while GSH competed with 4-MMDPA for binding to the pnictogen, illustrating the site-selectivity of the approach for different probe molecules (Fig. 3). Fig. 3 also shows that the 4-MMDPA complex clearly is more stable, which may in part be explained by a pnictogen–π interaction with the mercaptomethylpyridine group. In agreement with this result, a mass spectrum displayed only a small peak for the ternary complex of ZiPro, As(III) and GSH (data not shown).

Remarkably, when we probed the stability of the quaternary complex between ZiPro, pnictogen, 4-MMDPA and lanthanide by native polyacrylamide electrophoresis (PAGE) and terbium luminescence, the bismuth-mediated complex was found to survive intact. In contrast, the complexes with As(III) and Sb(III) decomposed, indicating a slower dissociation rate of the bismuth complex (Fig. 4).

In summary, a VDP containing a pair of cysteine residues with appropriate geometry presents opportunities for pnictogen-mediated self-assembled tagging with mercaptomethylaryl containing molecules. The assembly from simple inorganic pnictogen compounds presents a facile way of tagging, which is hardly perturbed by the presence of single cysteine residues and other thiol-group containing biomolecules. Interestingly, the assembly appears to be more favoured with antimony than with arsenic (Fig. 2B). The present work establishes a strategy to dispense with the difficult synthesis of toxic organoarsenical compounds and to expand the spectrum of pnictogen-mediated protein modifications from arsenic to less toxic antimony and non-toxic bismuth. The assembly is orthogonal to conventional tagging reactions of cysteine residues, can be performed in aqueous solution at neutral pH and produces remarkably stable complexes. The thermodynamic stability of the tag assembly presented here may not be sufficient for in-cell experiments as achieved, *e.g.*, by the organoarsenical FAsH tag, which can be used in submicromolar concentrations.[10b] Nonetheless, pnictogen-mediated self-assembly opens a path to site-specific reversible tagging of sensitive proteins under very mild conditions for use with a broad range of commercially available probes.

C. N. thanks the Alexander von Humboldt Foundation for a Feodor Lynen fellowship. Financial support by the Australian Research Council is gratefully acknowledged.

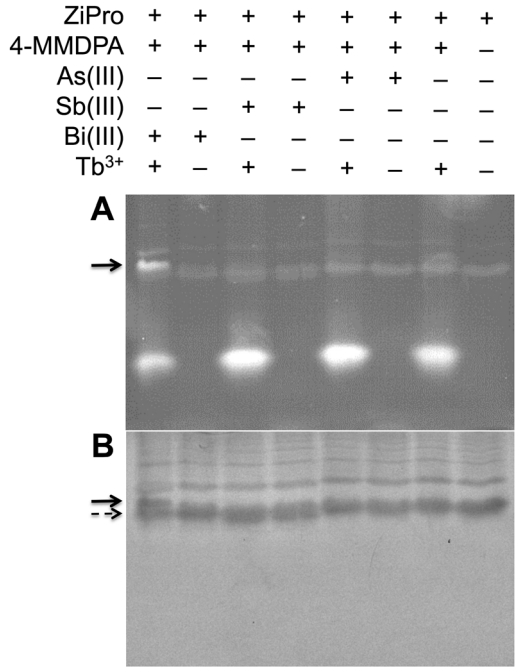


Fig. 4 Native PAGE for ZiPro C80S/C143S/P101C/P131C in the presence and absence of 4-MMDPA, trivalent pnictogens, and Tb^{3+}. (A) Terbium luminescence at 254 nm excitation wave length. (B) Coomassie blue stain. The dashed arrow identifies the band of native ZiPro without tag. The bold arrow identifies the band of the quaternary complex of ZiPro with trivalent bismuth, 4-MMDPA and terbium.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 (*a*) E. M. Sletten and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2009, **48**, 6974; (*b*) D. M. Patterson, L. A. Nazarova and J. A. Prescher, *ACS Chem. Biol.*, 2014, **9**, 592; (*c*) Y. Takaoka, A. Ojida and I. Hamachi, *Angew. Chem., Int. Ed.*, 2013, **52**, 4088; (*d*) K. Lang and J. W. Chin, *ACS Chem. Biol.*, 2014, **9**, 16; (*e*) C. Jing and V. W. Cornish, *Acc. Chem. Res.*, 2011, **44**, 784.
2 (*a*) E. A. Rodriguez, R. E. Campbell, J. Y. Lin, M. Z. Lin, A. Miyawaki, A. E. Palmer, X. Shu, J. Zhang and R. Y. Tsien, *Trends Biochem. Sci.*, 2017, **42**, 111; (*b*) V. Sample, R. H. Newman and J. Zhang, *Chem. Soc. Rev.*, 2009, **38**, 2852.
3 A. Brieger, G. Plotz, I. Hinrichsen, S. Passmann, R. Adam and S. Zeuzem, *PLoS One*, 2012, **7**, e31863.
4 (*a*) K. Lang and J. W. Chin, *Chem. Rev.*, 2014, **114**, 4764; (*b*) C. S. McKay and M. G. Finn, *Cell Chem. Biol.*, 2014, **21**, 1075.
5 K. Wang, H. Peng and B. Wang, *J. Cell. Biochem.*, 2014, **115**, 1007.
6 (*a*) F. H. Ma, J. L. Chen, Q. F. Li, H. H. Zuo, F. Huang and X. C. Su, *Chem. – Asian J.*, 2014, **9**, 1808; (*b*) J.-L. Chen, X. Wang, F. Yang, C. Cao, G. Otting and X.-C. Su, *Angew. Chem., Int. Ed.*, 2016, **55**, 13744.
7 (*a*) R. A. Studer, B. H. Dessailly and C. A. Orengo, *Biochem. J.*, 2013, **449**, 581; (*b*) B. Reva, Y. Antipin and C. Sander, *Nucleic Acids Res.*, 2011, **39**, e118.
8 (*a*) S. Shen, X. F. Li, W. R. Cullen, M. Weinfeld and X. C. Le, *Chem. Rev.*, 2013, **113**, 7769; (*b*) B. Chen, Q. Liu, A. Popowich, S. Shen, X. Yan, Q. Zhang, X.-F. Li, M. Weinfeld, W. R. Cullen and X. C. Le, *Metallomics*, 2015, **7**, 39.
9 (*a*) B. A. Griffin, S. R. Adams and R. Y. Tsien, *Science*, 1998, **281**, 269; (*b*) S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis and R. Y. Tsien, *J. Am. Chem. Soc.*, 2002, **124**, 6063.
10 (*a*) P. Wilson, A. Anastasaki, M. R. Owen, K. Kempe, D. M. Haddleton, S. K. Mann, A. P. Johnston, J. F. Quinn, M. R. Whittaker, P. J. Hogg and T. P. Davis, *J. Am. Chem. Soc.*, 2015, **137**, 4215; (*b*) Y. Wang, X.-F. Yang, Y. Zhong, X. Gong, Z. Li and H. Li, *Chem. Sci.*, 2016, **7**, 518; (*c*) C. Huang, T. Jia, M. Tang, Q. Yin, W. Zhu, C. Zhang, Y. Yang, N. Jia, Y. Xu and X. Qian, *J. Am. Chem. Soc.*, 2014, **136**, 14237.
11 C. Nitsche and G. Otting, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2017, **98–99**, 20.
12 J. Lei, G. Hansen, C. Nitsche, C. D. Klein, L. Zhang and R. Hilgenfeld, *Science*, 2016, **353**, 503.
13 (*a*) V. M. Cangelosi, T. G. Carter, J. L. Crossland, L. N. Zakharov and D. W. Johnson, *Inorg. Chem.*, 2010, **49**, 9985; (*b*) M. S. Collins, R. Y. Choi, L. N. Zakharov, L. A. Watson, B. P. Hay and D. W. Johnson, *Chem. Sci.*, 2015, **6**, 2444.
14 (*a*) V. M. Cangelosi, M. A. Pitt, W. J. Vickaryous, C. A. Allen, L. N. Zakharov and D. W. Johnson, *Cryst. Growth Des.*, 2010, **10**, 3531; (*b*) M. M. Watt, M. S. Collins and D. W. Johnson, *Acc. Chem. Res.*, 2013, **46**, 955.
15 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.
16 V. Stavila, R. L. Davidovich, A. Gulea and K. H. Whitmire, *Coord. Chem. Rev.*, 2006, **250**, 2782.
17 C. Schmitz, M. J. Stanton-Cook, X.-C. Su, G. Otting and T. Huber, *J. Biomol. NMR*, 2008, **41**, 179.
18 (*a*) X. Jia, A. Maleckis, T. Huber and G. Otting, *Chem. – Eur. J.*, 2011, **17**, 6830; (*b*) B. Man, X. C. Su, H. Liang, S. Simonsen, T. Huber, B. A. Messerle and G. Otting, *Chem. – Eur. J.*, 2010, **16**, 3827.

