ChemComm

RSC Publishing

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2013, **49**, 9104

Received 1st August 2013,
Accepted 21st August 2013

DOI: 10.1039/c3cc45875j

www.rsc.org/chemcomm

Direct and selective tagging of cysteine residues in peptides and proteins with 4-nitropyridyl lanthanide complexes†

Kathryn L. Gempf, Stephen J. Butler,* Alexander M. Funk and David Parker*

A cysteine-selective tagging method in water is reported, based on rapid displacement of a pyridyl nitro-substituent in simple pyridines and lanthanide complexes. The conjugation reaction creates a short link between the tag and peptide, holding the peptide closer to the Ln^{3+} ion and with reduced flexibility compared to existing methods.

We report the synthesis and application of a new class of lanthanide tags, capable of selectively ligating cysteine-containing peptides and proteins in water *via* a single-bond linkage, potentially aiding the study of these biomolecules by enhancing the shift and relaxation of proximate NMR resonances.

Paramagnetic lanthanide tags aid NMR analysis of biomolecules by providing long-range restraints for structure determination. The ^{1}H NMR spectra of Ln-tagged biomolecules can display signals shifted far from the normal range, allowing clarification of complex spectral data. The large shift and spread of resonances is primarily due to the pseudocontact shift (PCS); a through-space interaction that requires immobilization of the Ln relative to the protein, and decays with distance (r^{-3}) from the Ln centre. In order to maximise the PCS, the distance between the tag and the protein should be minimized (*i.e.* the linker between the two entities must be as short as possible) and the ligand field should be maximised. The pseudocontact shift is directly proportional to the crystal field coefficients, B_0^2 and B_2^2, in Bleaney's theory of magnetic anisotropy.[1]

Lanthanide tags are most commonly introduced into proteins by site-selective attachment to a cysteine residue, owing to the high nucleophilicity of thiols in water. However, only a few synthetic methods for cysteine modification exist; most reported Ln tags for cysteine attachment rely upon disulphide bond formation[2–4] or use maleimide electrophiles.[5,6] Each method has limitations: disulfide bonds are labile under both reducing and basic conditions and the reaction between maleimide and cysteine yields diastereoisomeric bioconjugates, which can complicate NMR analysis. Recently, the thiol-ene reaction was used to create stable Ln tags based on 4-substitutued dipicolinates (DPA).[5] Whilst this approach offers an improvement in stability compared to disulphide bond formation, a shorter, less flexible linkage between the Ln and protein is needed to generate larger PCS values. Lanthanide ions have been incorporated into encoded peptides, based on Ca(II) binding loops, using recombinant protein techniques;[7,8] such a strategy is restricted to proteins with a suitable fusion site.

In this work we report a new conjugation method, applied to a lanthanide complex. Two Ln(III) tags [**Ln.L^n**] (n = 1,2) are presented; each is based on a stable DO3A cyclen ring structure and possess an electron deficient 4-nitropyridine moiety (Fig. 1). The selective displacement of the *p*-nitro substituent by a cysteinyl group of a target protein generates a short single-bond linkage, generating a more conformationally rigid environment for the Ln ion, compared to existing tags. Enhanced PCS values are thereby created. An additional chelation site is offered in the coordinatively unsaturated complex [**Ln.L^2**], permitting ligation of an appropriately positioned anionic amino-acid side chain of a Cys-bound peptide, such as Glu/Asp or a phosphorylated residue (pSer/Thr/Tyr). Such bifunctional tags may further restrict dynamic motion, decreasing chemical exchange broadening in NMR.

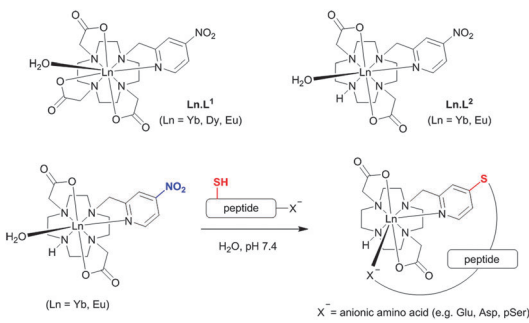

Fig. 1 (upper) Structures of lanthanide tags, **Ln.L$^{1–2}$**, and (lower) scheme showing the tagging reaction between **Ln.L^2** and a Cys-bound peptide, with additional ligation of an anionic amino-acid side chain. The conjugation reaction proceeds in the absence of the lanthanide ion.

Department of Chemistry, Durham University, South Road, Durham, DH1 3LE, UK. E-mail: s.j.butler@dur.ac.uk, david.parker@dur.ac.uk; Fax: +44 (0)191-3844737; Tel: +44 (0)191-3342033

† Electronic supplementary information (ESI) available: Details of complex synthesis and characterization are available. See DOI: 10.1039/c3cc45875j

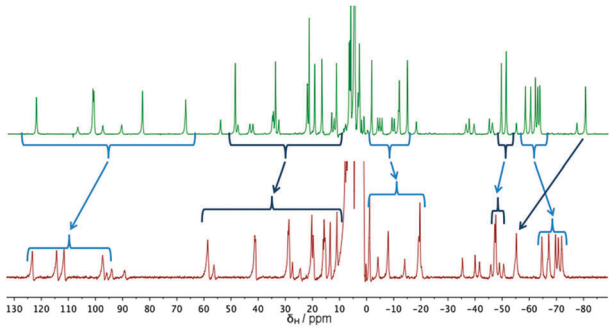

Fig. 2 ^{1}H NMR spectra of: (upper) [**Yb.L^1**], and (lower) [**Yb.L^1**] with a bound Boc-Cys-OMe residue (pD 6, 400 MHz).

The lanthanide tags **L^1** and **L^2** were synthesized using established methodology (ESI†) and the corresponding complexes of Yb, Dy, and Eu isolated as trifluoroacetate salts. Ytterbium was selected for its relatively narrow shifted signals, associated with a Bleaney coefficient, C_J of +39.2 and a modest magnetic moment of 4.3 BM. For comparison, Dy was also chosen (C_J = −181; μ_{eff} = 10.4 BM).[5c] The Eu complexes were prepared to allow monitoring of the change in the emission spectral fingerprints.

The ^{1}H NMR spectrum of [**Yb.L^1**] was fully assigned (Fig. 2, top) and revealed the presence of two main isomers in 1 : 3 ratio. Structurally similar complexes of [Ln(DOTA)]$^-$ are well known to exist as two diastereoisomers, in square antiprismatic (SAP) and twisted square antiprismatic (TSAP) forms.[9] Resonances for the major conformer were assigned by considering the polar coordinates of complex nuclei and their observed PCS. The four signals shifted to highest frequency correspond to the ring axial protons closest to the Ln^{3+} ion. An average axial ring proton shift of +93.4 ppm was observed; such a value strongly indicates that the major isomer is a SAP form.[10] The other set of axial protons were characterised by a shift to lower frequency (average δ_{Hax} = −61.0 ppm). Conversely, the eight ring equatorial protons occupy positions further from the Ln and are depicted as resonances within the 50 to 2.5 ppm range. The acetate CH_2 protons occupy different positions, giving rise to six signals at low frequency (−1.6, −11.7, −14.7, −49.4, −51.2, and −80.4 ppm). The CH_2 protons of the pyridyl arm appear to be strongly shifted (101.3 and −62.8 ppm), consistent with their close proximity to the Ln, whereas the relatively distant pyridyl protons give rise to modestly shifted signals at similar frequencies (6.8, 6.5 and 6.2 ppm). For the heptadentate complex [**Yb.L^2**], the ^{1}H NMR spectral width corresponds well with that of [**Yb.L^1**] (ESI†). However, the spectrum is rather broad, due to the presence of chemical exchange between various conformers.

The crystal field parameter, B_0^2 that is directly proportional to the PCS of resonances, was calculated for [**Eu.L^1**/**L^2**] by analysis of the ΔJ = 1 splitting in the europium emission spectra. Relatively high values of −420 and −427 cm^{-1} were found for the Eu complexes of **L^1** and **L^2** respectively, compared to +100 cm^{-1} for the Eu(III) complexes of picolinate ligands.[5]

The addition of Boc-Cys-OMe, to complexes [**Ln.L$^{1-2}$**] (Ln = Yb, Dy, Eu) resulted in rapid and efficient displacement of the nitro substituent in DMF within 1 h at 20 °C. When Boc-Cys-OMe was added to [**Dy.L^1**] in the presence of a 20-fold excess of Boc-Lys-OH, complete thiol-selectivity was observed and the reaction was over within 15 min. Each displacement reaction was monitored by LCMS and UV-vis spectroscopy, revealing a shift in the main π–π* transition from 300 nm for [**Ln.L$^{1-2}$**] to 280 nm for the Cys-bound complexes.

Each tagging reaction gave rise to clear and distinct changes in the ^{1}H NMR spectra. Representative spectra for [**Yb.L^1**] before and after tagging with the Cys residue (Fig. 2) indicated the presence of two isomeric species for [**Yb.L^1**-Cys]. For the major isomer, clear shifts in the distribution of shifted resonances were identified. The four cyclen axial protons closest to the Ln have an average shift of 109.4 ppm, suggesting that the major isomer of the Cys-bound tag adopts a SAP conformation.[10] The appearance of multiple, overlapping resonances in the 10 to −1 ppm range correspond to the more distant protons; namely the pyridyl protons and the fifteen protons of the protected Cys residue.

To demonstrate the versatility of this thiol-specific reaction under ambient aqueous conditions, the selective ligation of the single surface-exposed Cys residue (Cys-58) in bovine serum albumin (BSA) was accomplished following incubation with [**Dy.L^1**] in water for 1 h. Analysis by MALDI-TOF MS revealed species at 66 456 and 67 059 Da, corresponding to native BSA and the **Dy.L^1**-derivatised BSA, respectively. To identify the exact site of ligation, a trypsin digestion of both the native protein and **Dy.L^1**-derivatised BSA was undertaken. Peptide mass fingerprinting revealed a shift in the mass of the peptide fragment Gly45-Lys65 (which contains Cys58) from m/z = 2435.2 to 3032.0 after modification, confirming unequivocally the Cys ligation of a single molecule of [**Dy.L^1**] to the BSA sequence.

Lanthanide complexes of **L^2** offer an additional binding site for the chelation of an anionic residue (*e.g.* Glu, pSer) of a Cys-bound peptide or protein. Such a process was envisaged to immobilize the complex further and limit dynamic motion (Fig. 1). Intermolecular oxyanion binding was demonstrated following addition of a ten-fold excess of acetate and phosphate to [**Yb.L^2**]. Significant shifts in the ^{1}H NMR spectra were observed consistent with displacement of the bound water molecules by the added anions, associated with the variation in the polarizability of the oxygen ligand donors.[11] Additionally, Eu emission spectra of [**Eu.L^2**] were recorded in the presence of a 100-fold excess of NH_4OAc and Na_2HPO_4 (Fig. 3). Fingerprint spectra for each anion adduct displayed characteristic changes in spectral form, associated with alteration of the Eu coordination environment.[12]

Next, [**Ln.L^2**] (Ln = Yb, Eu) was reacted with three short pentapeptides Cys-Phe-DPro-Gly-Glu-$CONH_2$ (**1**), Cys-Phe-DPro-Gly-pSer-$CONH_2$ (**2**), and Cys-Phe-DPro-Gly-Ser-$CONH_2$ (**3**). Each peptide contains an N-terminal Cys residue for thiol ligation; peptides **1** and **2** possess a C-terminal Glu and pSer residue respectively, for chelation to the Ln metal. The model peptide **3** has a non-phosphorylated Ser residue. In addition, each peptide contains the β-turn inducing sequence DPro-Gly, which may assist in binding of the Glu/pSer side chain to the Ln by bringing the two components together.[13]

The addition of 1 equiv. of each peptide to [**Ln.L^2**] (Ln = Yb, Eu) in water resulted in good conversions (>95%) to the tagged peptides, as judged by analytical RP-HPLC (see ESI†). The Eu emission spectrum of [**Eu.L^2**-peptide **1**] was almost identical to

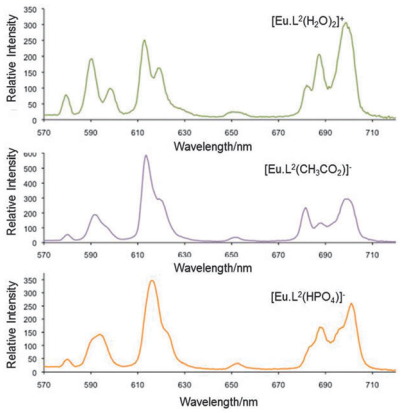

Fig. 3 Emission spectra of: (upper) [**Eu.L^2**]; (centre) [**Eu.L^2**] in 1 mM NH_4OAc solution and; (lower) [**Eu.L^2**] in 1 mM Na_2HPO_4 solution (293 K, H_2O, pH 6, 10 μM complex).

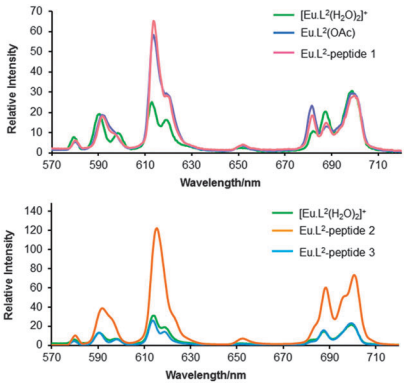

Fig. 4 (upper) Emission spectra of: [**Eu.L^2**] in water (green); [**Eu.L^2**] in 1 mM NH_4OAc solution (blue) and; [**Eu.L^2**-peptide **1**] (pink) in water. (lower) Overlaid emission spectra of: [**Eu.L^2**] in water (green); [**Eu.L^2**-peptide **2**] (orange) and; [**Eu.L^2**-peptide **3**] (blue) in water.

the emission fingerprint spectrum recorded for [**Eu.L^2**] in acetate solution (Fig. 4), and differs clearly from that of [**Eu.L^2**] in water alone, indicating that the terminal Glu residue is directly interacting with the Eu metal. Likewise, the Eu emission spectral form of [**Eu.L^2**-peptide **2**] was very similar to that recorded for [**Eu.L^2**] in phosphate solution, consistent with ligation of the terminal pSer residue of peptide **2** to the Eu centre. Additionally, the ^{31}P NMR spectrum of **Yb.L^2**-peptide **2** showed two signals at 3.4 and −50.6 ppm in the ratio 2 : 1, indicating that two species are present in solution. The signal at 3.4 ppm was assigned to the species in which the pSer moiety was unbound (free peptide δ_p = 0.2 ppm). The broad signal at −50.6 ppm corresponds to the pSer-bound species, confirmed by relaxation studies (ESI†), that revealed the enhanced rate of relaxation of this signal as a function of magnetic field (R_1 = 84, 114 and 143 s^{-1} at 11.7, 14.1 and 16.5 T, 295 K). Such a rate variation accords with an Yb-P separation, r, of 3.50(0.07) Å, consistent with structural data for related Yb-phosphinate complexes, where r = 3.39(0.02) Å.[14] Finally, the europium emission spectrum of [**Eu.L^2**-peptide **3**] was identical to that of [**Eu.L^2**] in water, indicating that in this case binding of the terminal residue to the lanthanide ion was not occurring.

In conclusion, the Ln complexes reported here function as efficient, selective and chemically stable tags for cysteine-containing peptides and proteins. Moreover, we have devised a new method of protein attachment,[17] based on displacement of the nitro group in a *p*-substituted pyridine moiety that creates a short link between the tag and protein, compared to previous methods with electron poor alkene or maleimide electrophiles. The tagging reaction is thiol-specific and proceeds rapidly in water at room temperature, demonstrated by the facile ligation of the single, exposed Cys residue of BSA. Using [**Ln.L^2**], tagged peptides can be further constrained by coordination of proximal anionic amino acid residues (Glu and pSer) to the Ln centre.

These examples highlight the advantages of the nucleophilic substitution of an activated aromatic nitro group over existing protocols for Cys binding. The range of applications could embrace introduction of a heavy atom label to aid crystallographic analysis, may include examples in NMR analysis created by the introduction of one (or more) Ln ions of differing shift and relaxation ability and could extend to protein labelling with radioisotopes or luminescent species, *e.g.* a Eu(III) complex analogue of the recent, very bright series of complexes.[12,15,16]

We thank Dr Jackie Mosely for analysis of the MALDI MS data and the ERC for support (FCC 266804).

Notes and references

1 B. Bleaney, *J. Magn. Reson.*, 1972, **8**, 91.
2 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.
3 P. H. J. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2008, **130**, 14802.
4 D. Häussinger, J.-r. Huang and S. Grzesiek, *J. Am. Chem. Soc.*, 2009, **131**, 14761.
5 (*a*) Q.-F. Li, Y. Yang, A. Maleckis, G. Otting and X.-C. Su, *Chem. Commun.*, 2012, **48**, 2704; (*b*) For the original application to protein linkage: J. R. Morphy, D. Parker, R. Kataky, M. A. W. Eaton, A. T. Millican, R. Alexander, A. Harrison and C. Walker, *J. Chem. Soc., Perkin Trans. 2*, 1990, 573(*c*) A. M. Funk, P. Fries, P. Harvey, A. M. Kenwright and D. Parker, *J. Phys. Chem. A*, 2013, **117**, 905.
6 J. Chen and P. R. Selvin, *Bioconjugate Chem.*, 1999, **10**, 311.
7 J. Wöhnert, K. J. Franz, M. Nitz, B. Imperiali and H. Schwalbe, *J. Am. Chem. Soc.*, 2003, **125**, 13338.
8 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.
9 S. Aime, M. Botta and G. Ermondi, *Inorg. Chem.*, 1992, **31**, 4291.
10 R. S. Dickins, D. Parker, J. I. Bruce and D. J. Tozer, *Dalton Trans.*, 2003, 1264; D. Parker, *Chem. Soc. Rev.*, 2004, **33**, 156–165.
11 S. J. Butler and D. Parker, *Chem. Soc. Rev.*, 2013, **42**, 1652–1666.
12 S. J. Butler, B. K. McMahon, R. Pal, D. Parker and J. W. Walton, *Chem.–Eur. J.*, 2013, **19**, 9511.
13 H. E. Stanger and S. H. Gellman, *J. Am. Chem. Soc.*, 1998, **120**, 4236.
14 S. Aime, A. S. Batsanov, A. Beeby, M. Botta, R. S. Dickins, S. Faulkner, C. E. Foster, J. AS. K. Howard, J. M. Moloney, T. J. Norman, D. Parker and J. A. G. Williams, *J. Chem. Soc., Dalton Trans.*, 1997, 3623.
15 J. W. Walton, M. Soulie, A. Bourdolle, S. J. Butler, M. Delbianco, L. Lamarque, C. Andraud, O. Maury, B. K. McMahon, R. Pal, D. Parker, H. Puschmann and J. Zwier, *Chem. Commun.*, 2013, **49**, 1600.
16 B. K. McMahon, R. Pal and D. Parker, *Chem. Commun.*, 2013, **49**, 5763.
17 Reactions of 4-nitro-2-hydroxymethyl pyridine with cysteine or glutathione (pH 7, 20 °C, 40 min) went to completion, suggesting a role in tagging an unbound ligand for antibody/peptide radiolabelling.

