Magn. Reson., 1, 1–13, 2020
https://doi.org/10.5194/mr-1-1-2020
© Author(s) 2020. This work is distributed under
the Creative Commons Attribution 4.0 License.

Open Access
MAGNETIC
RESONANCE

Phosphoserine for the generation of lanthanide-binding sites on proteins for paramagnetic NMR

Sreelakshmi Mekkattu Tharayil[1,★], **Mithun Chamikara Mahawaththa**[1,★], **Choy-Theng Loh**[1,a], **Ibidolapo Adekoya**[1], **and Gottfried Otting**[1]

[1]ARC Centre of Excellence for Innovations in Peptide and Protein Science, Research School of Chemistry, Australian National University, Canberra ACT 2601, Australia
[a]present address: Hangzhou Wayland Bioscience Co. Ltd, Hangzhou 310030, PR China
★These authors contributed equally to this work.

Correspondence: Gottfried Otting (gottfried.otting@anu.edu.au)

Received: 27 October 2020 – Discussion started: 4 November 2020
Revised: 11 December 2020 – Accepted: 12 December 2020 – Published:

Abstract. TS1Pseudocontact shifts (PCSs) generated by paramagnetic lanthanide ions provide valuable long-range structural information in NMRCE1 spectroscopic analyses of biological macromolecules such as proteins, but labelling proteins site-specifically with a single lanthanide ion remains an ongoing challenge, especially for proteins that are not suitable for ligation with cysteine-reactive lanthanide complexes. We show that a specific lanthanide-binding site can be installed on proteins by incorporation of phosphoserine in conjunction with other negatively charged residues, such as aspartate, glutamate or a second phosphoserine residue. The close proximity of the binding sites to the protein backbone leads to good immobilization of the lanthanide ion, as evidenced by the excellent quality of fits between experimental PCSs and PCSs calculated with a single magnetic susceptibility anisotropy ($\Delta\chi$) tensor. An improved two-plasmid system was designed to enhance the yields of proteins with genetically encoded phosphoserine, and good lanthanide ion affinities were obtained when the side chains of the phosphoserine and aspartate residues are not engaged in salt bridges, although the presence of too many negatively charged residues in close proximity can also lead to unfolding of the protein. In view of the quality of the $\Delta\chi$ tensors that can be obtained from lanthanide-binding sites generated by site-specific incorporation of phosphoserine, this method presents an attractive tool for generating PCSs in stable proteins, particularly as it is independent of cysteine residues.

1 Introduction

Paramagnetic labels offer an attractive tool for the study of protein structure and function, as the magnetic moments of unpaired electrons generate long-range paramagnetic effects in NMR spectra. Among the paramagnetic effects that can be observed in NMR spectra, pseudocontact shifts (PCSs) generated by paramagnetic metal ions stand out for their high information content and ease of observation (Otting, 2008; Parigi and Luchinat, 2018). Specifically, the PCSs provide information about the location of nuclear spins relative to the magnetic susceptibility anisotropy tensor ($\Delta\chi$ tensor) associated with a paramagnetic metal ion, and this information can readily be obtained for nuclear spins as far as 40 Å from the paramagnetic centre (Bertini et al., 2001).

As many lanthanide ions display particularly large $\Delta\chi$ tensors (Bleaney, 1972; Bertini et al., 2001), significant efforts have been made to devise lanthanide complexes for site-specific tagging of proteins (Su and Otting, 2010; Keizers and Ubbink, 2011; Nitsche and Otting, 2017; Joss and Häussinger, 2019; Saio and Ishimori, 2020). In an alternative approach, PCSs can be elicited in proteins by creating binding sites for lanthanides or lanthanide complexes by protein engineering (Yagi et al., 2010; Barthelmes et al., 2011, 2015; Jia et al., 2011).

Please note the remarks at the end of the manuscript.

A common problem of lanthanide tags arises from mobility of the metal-ion complex relative to the target protein. Paramagnetic lanthanide ions always generate paramagnetic relaxation enhancements (PREs) in the protein, which vary relatively little with minor movements of the metal ion. In contrast, PCSs can decrease dramatically if the lanthanide complex reorientates relative to the protein. With a limited degree of tag flexibility, the PCSs may still be explained by a single effective $\Delta\chi$ tensor, although, in principle, a family of $\Delta\chi$ tensors would be required to account for multiple tag conformations (Shishmarev and Otting, 2013). Well-immobilized metal ions thus deliver not only larger PCSs, but also more reliable $\Delta\chi$-tensor fits.

Different strategies have been devised to immobilize lanthanide ions on proteins. Tag motions can be restricted by short tethers and bulky lanthanide complexes to hem in the tag sterically (Nitsche and Otting, 2017). Double-arm tags provide two attachment points (Keizers and Ubbink, 2011), but even these designs have shown signs of tag mobility (Hass et al., 2010). A lanthanide-binding peptide (LBP) engineered into polypeptide loops of protein structures can deliver good metal immobilization but presents a major modification of the target protein (Barthelmes et al., 2011, 2017). Fusions of an LBP combined with disulfide bond formation have also been explored but do not necessarily achieve good immobilization of the lanthanide ion (Saio et al., 2009, 2010, 2011). A successful strategy has been a design where two neighbouring cysteine residues are furnished with metal-chelating tags and a single lanthanide ion is coordinated by both chelating groups (Swarbrick et al., 2011; Welegedara et al., 2017), a design that has also proven successful for Co^{2+} ions (Swarbrick et al., 2016). The most serious drawback of this design is its reliance on cysteine residues, which makes it incompatible with proteins that contain functionally important cysteine residues in their wild-type sequence. In fact, most of the currently available lanthanide tags target cysteines (Su and Otting, 2010; Keizers and Ubbink, 2011; Nitsche and Otting, 2017; Joss and Häussinger, 2019; Saio and Ishimori, 2020), as thiol groups can readily undergo selective chemical reactions. To avoid the mobility of solvent-exposed cysteine side chains, tags have also been designed for attachment to the side chains of aromatic residues, which are more hindered sterically and thus discouraged from populating different rotamers (Loh et al., 2015; Abdelkader et al., 2016), but this approach results in long linkers between the lanthanide ion and the protein, increasing the chances that the lanthanide ion moves and reorientates relative to the protein backbone.

The most elegant strategy for generating a lanthanide-binding site in a protein would be to introduce a lanthanide-binding unnatural amino acid that can be site-specifically incorporated by genetic encoding. This approach would relieve any reliance on cysteine residues. Although systems for genetic encoding have been devised for over 100 different unnatural amino acids, only a few of these can bind metal ions (Dumas et al., 2014), and those that do were found to precipitate proteins upon binding lanthanide ions. For example, protein precipitation has been reported for 2-amino-3-(8-hydroxyquinolin-3-yl) propanoic acid (HQ-Ala; Jones et al., 2009), and we found ourselves incapable of improving on these results. Similarly, bipyridyl-alanine (Bpa) was shown to allow binding of Co^{2+} and the observation of PCSs (Nguyen et al., 2011), but subsequent experiments with Bpa incorporated in different proteins and at different sites showed that this system is also prone to precipitating proteins upon addition of metal ions.

In the present work, we explored the potential of a different unnatural amino acid, phosphoserine (Sep), to create a lanthanide-binding site. Lanthanide ions are known for their affinity with negatively charged oxygens and, with a pK_a TS2 value of 5.6 for the equilibrium between monobasic and dibasic forms (Xie et al., 2005), a phosphoserine residue carries two negative charges under physiological conditions. Phosphorylation of serine residues is a well-known post-translational modification of proteins affected by kinases, but this often is neither quantitative nor easily achievable for specific serine residues. Recently, however, an orthogonal phosphoseryl-tRNA-synthetase/tRNA pair has become available, which allows installation of a Sep residue in response to an amber stop codon (Lee et al., 2013; Pirman et al., 2015; Yang et al., 2016). In the following we show that the system is sufficiently effective at installing two Sep residues in the same protein, exploring the potential to create a lanthanide-binding site using a Sep residue in conjunction with other negatively charged residues, in particular an aspartate or a second Sep residue, and demonstrating the exceptional quality of $\Delta\chi$ tensors that can be obtained with lanthanide ions in these sites.

2 Experimental procedures

2.1 Plasmid preparation for protein expression

The plasmid SepOTSλ, which contains the phosphoseryl-tRNA synthetase/tRNA pair and a suitable EF-Tu mutant for incorporation of Sep in response to an amber stop codon (Pirman et al., 2015), was obtained from Addgene. To create a T7 expression vector that is compatible with SepOTSλ, we subcloned the region containing the T7 promoter, ribosome binding site, multiple cloning site and T7 terminator from pETMCSIII (Neylon et al., 2000) into the plasmid pCDF (Lammers et al., 2014). The gene of interest was inserted into the multiple cloning site and furnished with a C-terminal His_6-tag preceded by a TEV cleavage site. All plasmid constructions were conducted with a QuikChange protocol using mutant T4 DNA polymerase (Qi and Otting, 2019).

Please note the remarks at the end of the manuscript.

2.2 Protein expression

All proteins were expressed in the BL21ΔserB strain (Park et al., 2011), which lacks phosphoserine phosphatase and thus minimizes the dephosphorylation of phosphoserine to serine. The SepOTSλ and pCDF plasmids were co-transformed into electrocompetent BL21ΔserB cells. In order to minimize usage of amino acids and $^{15}NH_4Cl$, the following top-down expression method was used. Initially, 1 L of cell culture was grown in LB medium with 25 µM spectinomycin and 20 µM kanamycin at 37 °C until the OD_{600} value reached 0.6–0.8. Next, the cells were pelleted and resuspended in 300 mL M9 medium (6 g L^{-1} Na_2HPO_4, 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} NaCl) and supplied with 1 g L^{-1} $^{15}NH_4Cl$ and 1 mM phosphoserine. Subsequently, the cells were incubated for 30 min at 37 °C and induced with IPTG. Protein expression was conducted at 25 °C overnight. Cells were harvested by centrifugation at 5000 *g* for 15 min and lysed by passing twice through a French press (SLM Amicon, USA) at 830 bar. The lysate was then centrifuged at 13 000 *g* for 60 min and the filtered supernatant was loaded onto a 5 mL Ni-NTA column (GE Healthcare, USA) equilibrated with binding buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 5 % glycerol). The protein was eluted with elution buffer (binding buffer containing, in addition, 300 mM imidazole). For the double-amber mutants, the protein was dialysed into TEV protease buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl and 1 mM β-mercaptoethanol) to remove the His_6-tag by digestion with TEV protease overnight at 4 °C. His_6-tagged TEV protease was added in a 0.1 molar ratio. The protease and cleaved His_6-tag were removed by running the sample again over a Ni-NTA column. The resulting protein samples were then treated with 5 mM EDTA to remove any di- or tri-valent metal ion that could have been adsorbed during protein expression and purification. Finally, EDTA was removed by buffer exchange with the NMR buffer (20 mM HEPES-KOH, pH 7.0) using a HiPrep desalting column (GE Healthcare, USA). Mass-spectrometric analysis was conducted using an Elite Hybrid Ion Trap-Orbitrap mass spectrometer (Thermo Scientific, USA) coupled with an UltiMate S4 3000 UHPLC (Thermo Scientific, USA); 7.5 pmol of sample was injected into the mass analyser via an Agilent ZORBAX SB-C3 Rapid Resolution HT Threaded Column (Agilent, USA).

2.3 NMR spectroscopy

All NMR spectra were recorded at 25 °C, using an 800 MHz Bruker Avance NMR spectrometer for all mutants containing a single phosphoserine residue and a 600 MHz Bruker Avance NMR spectrometer for all mutants containing two phosphoserine residues. Samples were prepared in 20 mM HEPES buffer, pH 7.0, in 3 mm NMR tubes; 10 % D_2O was added to provide a lock signal; 0.1–0.5 mM protein samples were used for 2D [^{15}N,^{1}H]-HSQC experiments. Complexes with lanthanides were obtained by titration with 10 mM $LnCl_3$ stock solutions in water. The samples were titrated with paramagnetic lanthanides until the NMR spectra predominantly displayed cross-peaks of the paramagnetic species with minimal interference of cross-peaks from the diamagnetic species. Residual weak cross-peaks from the diamagnetic protein were accepted to avoid over-titration, which could lead to metal binding at additional sites. For the diamagnetic reference with Y^{3+} ions, the reference samples were prepared with the same quantity of YCl_3 stock as the paramagnetic samples.

2.4 PCS measurements and $\Delta\chi$-tensor fitting

PCSs were measured in ppm as the difference in amide proton chemical shift between the paramagnetic and diamagnetic NMR spectrum. PCSs were used to determine the position and orientation of the $\Delta\chi$ tensor of the paramagnetic ions relative to the protein structure. Fitting of $\Delta\chi$ tensors was performed using the program Paramagpy (Orton et al., 2020).

2.5 Isothermal titration calorimetry

Isothermal calorimetric titration experiments were performed using a Nano-ITC low-volume calorimeter (TA Instruments, USA) at 25 °C with stirring at 250 rpm. The protein mutant E18Sep and the titrants $TbCl_3$ and $TmCl_3$ were prepared in the same buffer (20 mM HEPES, pH 7.0) and degassed before use. Data were analysed using the programs NITPIC and SEDPHAT (Keller et al., 2012). The baseline-subtracted power peaks were integrated and the integrated heat values fitted to the single binding site model ($A + B \leftrightarrow AB$, heteroassociation) to obtain the dissociation constant (K_d). The global fitting was done by repeatedly cycling between Marquardt–Levenberg and simplex algorithms in SEDPHAT until modelling parameters converged; 68 % confidence intervals were calculated using the automatic confidence interval search with the projection method using F-statistics in SEDPHAT.

3 Results

3.1 Phosphoserine incorporation

Simultaneous transfection of *E. coli* with the SepOTSλ plasmid and pET vectors containing the genes of proteins targeted for overexpression and Sep incorporation led to slow cell growth and variable colony sizes on plates as described earlier (Pirman et al., 2015). Noting that the SepOTSλ plasmid contains the origin of replication of pUC, which belongs to the same plasmid incompatibility group as pET vectors (Morgan, 2014), we constructed a new expression vector based on pCDF to include T7 promoter, ribosome binding site, multiple cloning site and T7 terminator. This modification restored the usual growth rates of the cells. Proteins con-

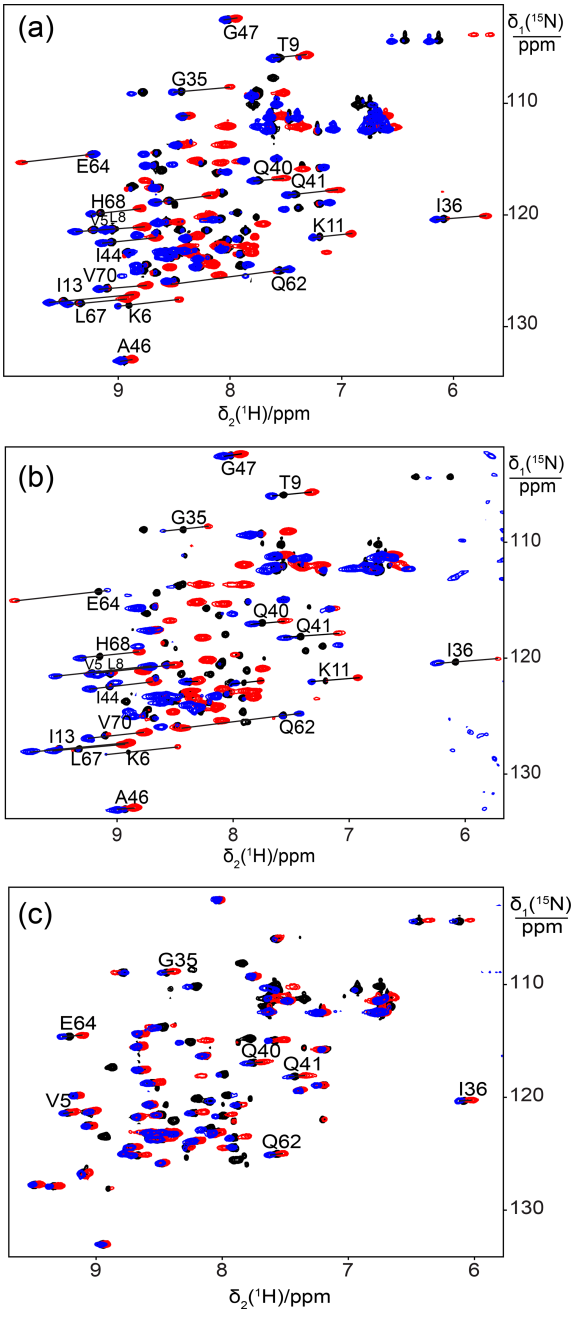

Figure 1. Superimposition of [^{15}N,^{1}H]-HSQC spectra of 0.5 mM solutions of ^{15}N-labelled ubiquitin mutated to generate a lanthanide-binding site at residue 18. Spectra with diamagnetic Y^{3+} are plotted in black and with paramagnetic Tb^{3+} and Tm^{3+} in red and blue, respectively. Lines were drawn to connect some of the cross-peaks belonging to the same residue in the paramagnetic and diamagnetic samples and are labelled with the residue name and sequence number. **(a)** Mutant E18Sep. **(b)** E16Q/E18Sep. **(c)** E18Sep/D21N.

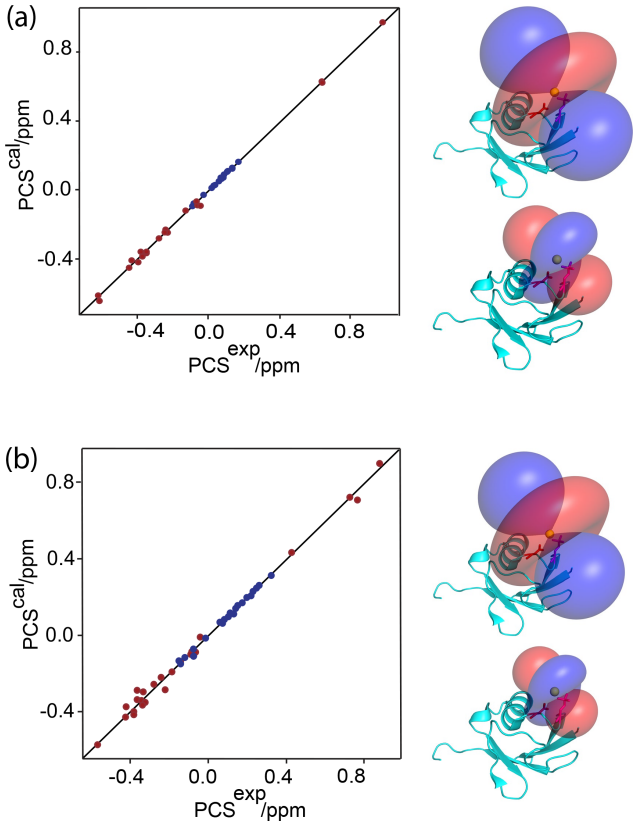

Figure 2. Correlation between back-calculated and experimental PCSs, and lanthanide locations on the ubiquitin mutants **(a)** E18Sep and **(b)** E16Q/E18Sep. Left panel: PCS data obtained with Tb^{3+} and Tm^{3+} plotted in red and blue, respectively. Right panel: blue and red PCS isosurfaces plotted on the protein structure and indicating PCSs of ±1 ppm, respectively. The isosurfaces illustrate the $\Delta\chi$ tensors obtained with Tb^{3+} (upper structure) and Tm^{3+} (lower structure). The side chains of E16 and the phosphoserine residue in position 18 are shown in a stick representation.

taining phosphoserine were expressed from a two-plasmid system containing SepOTSλ and a modified pCDF vector in BL21ΔserB. Expression yields of up to 3 mg purified protein per litre of growth medium were obtained.

3.2 Single phosphoserine residues for lanthanide binding

We used the proteins ubiquitin and GB1 to test whether a single phosphoserine residue is sufficient to create a lanthanide-binding site. We hypothesized that a phosphoserine residue assisted by an additional carboxyl CE2 group from a glutamate or aspartate residue (in the following referred to as "helper residue") could potentially be sufficient to generate a tridentate complex with a lanthanide ion, positioning the metal ion close to the protein backbone and compensating its positive charge. In the first example, we made the mutant E18Sep of ubiquitin, where Glu16 and Asp21 could

Please note the remarks at the end of the manuscript.

act as potential helper residues. Subsequent titration with Tb^{3+} ions succeeded in generating PCSs of up to almost 1 ppm (Table S1). The paramagnetic peaks appeared at chemical shifts different from the diamagnetic parent peaks, indicating slow exchange between lanthanide-bound and free protein. Isothermal calorimetric experiments with Tb^{3+} and Tm^{3+} ions indicated dissociation constants of about 30–50 µM (Fig. S1).

Figure 1a shows the PCSs observed with Tb^{3+} and Tm^{3+} ions. Using the NMR ensemble structure of ubiquitin (PDB ID: 2KOX; Fenwick et al., 2011) and the measured PCSs, the metal position was determined by fitting the $\Delta\chi$ tensor using the program Paramagpy (Table 1; Orton et al., 2020). The correlation between back-calculated and experimental PCSs was excellent (Fig. 2a, Table S1), resulting in Q factors lower than 0.04. This indicated that the Sep residue and lanthanide complex did not alter the structure of the protein. Furthermore, the tensor fit positioned the lanthanide ion between the phosphoserine residue and D21, suggesting that D21 acts as a helper residue rather than E16. To verify this result, we prepared the two ubiquitin mutants E18Sep/E16Q and E18Sep/D21N. As expected, the former delivered similar PCSs (Fig. 1b, Table S1), a similarly good $\Delta\chi$-tensor fit (Table 1) and a similar metal position, whereas the latter showed only very small chemical shift changes upon titration with lanthanides, indicating a faster exchange (Fig. 1c). The paramagnetic centre identified by the fits placed the lanthanide ions between the aspartate and Sep residues as expected (Fig. 2b).

3.3 Phosphoserine and aspartate for introducing a lanthanide-binding site into GB1

The scheme of combining a phosphoserine with an aspartate helper residue to create a lanthanide-binding site was also successful with the GB1 mutant K10D/T11Sep, where Tb^{3+} ions generated PCSs as large as 0.55 ppm (Fig. 3a, Table S2) and, as for the ubiquitin mutants, the lanthanide complex was in slow exchange with the free protein. The $\Delta\chi$-tensor fit produced an excellent correlation between back-calculated and experimental PCSs with a Q factor of 0.01, indicating good immobilization of the lanthanide ion (Fig. 3c, Table S2). The best fit of the $\Delta\chi$ tensor positioned the lanthanide between the phosphoserine and aspartic acid residues as expected (Fig. 3e).

3.4 Double-phosphoserine motifs in GB1

Next we assessed the possibility of generating a lanthanide-binding motif by the introduction of two phosphoserine residues. For comparison with the GB1 mutant K10D/T11Sep, a double-amber mutant of GB1 was made to replace both K10 and T11 with phosphoserine. The protein was obtained in good yield (1.5 mg from 1 Ll of cell culture) despite the presence of two amber stop codons. Successful double-amber suppression was confirmed by mass spectrometry (Fig. S2a). Following titration with Tb^{3+} ions, we observed PCSs up to 1 ppm (Fig. 3a, Table S2). The $\Delta\chi$-tensor fit indicated that the lanthanide ion binds between the Sep residues in positions 10 and 11 as expected and the agreement between back-calculated and experimental PCSs was excellent (Fig. 3d and f). The very low Q factor associated with the $\Delta\chi$-tensor fit (Table 2) demonstrates that the PCSs are adequately explained by a single $\Delta\chi$ tensor, indicating CE3 the absence of averaging between different tensors arising from translational movements of the paramagnetic centre.

In previous work, we reported that two nitrilotriacetic acid (NTA) tags attached to cysteine residues in positions i and $i+4$ of an α-helix yielded larger PCSs with lanthanides than a single NTA tag combined with an acidic helper residue (Swarbrick et al., 2011). In view of this result, we also attempted to position two phosphoserine residues in positions i and $i+4$ of the α-helix of GB1. About 1 mg of GB1 A24Sep/K28Sep was obtained from 300 mL cell culture, and the successful and complete incorporation of two Sep residues was confirmed by mass spectrometry (Fig. S2b).

Following titration with Tb^{3+} ions, PCSs of up to 3 ppm were observed (Table S2). Figure 4a shows the PCSs observed with Tb^{3+} and Tm^{3+} ions following titration to lanthanide : protein ratios at which the diamagnetic peaks had completely or mostly disappeared. Excess lanthanide ions resulted in significant peak broadening, indicating weak binding of the excess lanthanide ions to less specific sites. The $\Delta\chi$-tensor fits to the crystal structure of GB1 revealed relatively large $\Delta\chi$ tensors and a small Q factor (Fig. 4b and Table 2), indicating good immobilization of the lanthanide ion. The paramagnetic centre identified by the fits placed the lanthanide ions between the two Sep residues as expected (Fig. 4c).

3.5 Double-phosphoserine incorporation into other proteins

To test the broader validity of double-phosphoserine motifs as lanthanide-binding sites, we generated double-amber mutants for double-phosphoserine incorporation in 16 different sites in four different proteins (Fig. S3). The double-amber mutations were designed to position two phosphoserine residues in α-helices (positions i and $i+4$), loops (positions i and $i+2$), and β-strands (positions i and $i+2$, as well as two positions located in parallel β-strands). Among the constructs made of GB1, ubiquitin, *E. coli* PpiB, Zika virus NS2B-NS3 protease, and the N-terminal ATP-binding domain of *Plasmodium falciparum* Hsp90 (Hsp90-N), in vivo expression attempts produced protein only for two of the constructs, namely Hsp90-N S36Sep/D40Sep (where the phosphoserine residues are in an α-helix) and ubiquitin T66/H68 (where the phosphoserine residues are in a β-strand). All the other constructs failed to produce protein.

Please note the remarks at the end of the manuscript.

Table 1. $\Delta\chi$-tensor parameters of the ubiquitin mutants E18Sep, E16Q/E18Sep, and T22Sep/N25D/K29Q and the GB1 mutant K10D/T11Sep complexed with Tb^{3+} and Tm^{3+} ions[a]. TS4

Protein	N[b]	$\Delta\chi^{c}_{ax}(10^{-32}\,m^3)$	$\Delta\chi^{c}_{rh}(10^{-32}\,m^3)$	x (Å)	y (Å)	z (Å)	$\alpha(°)$	$\beta(°)$	$\gamma(°)$	Q[d]
Ubiquitin E18Sep (Tb^{3+})	20	17.1 (0.6)	2.8 (0.3)	10.095	−1.846	−11.711	170	138	50	0.03
Ubiquitin E18Sep (Tm^{3+})	27	−2.7 (0.1)	−1.0 (0.1)	9.463	−0.674	−12.207	168	129	49	0.03
Ubiquitin E16Q/E18Sep (Tb^{3+})	27	15.9 (0.6)	3.4 (0.8)	9.695	−1.754	−11.833	162	135	37	0.04
Ubiquitin E16Q/E18Sep (Tm^{3+})	28	−4.5 (0.1)	−2.1 (0.1)	9.441	−1.902	−11.918	164	131	59	0.03
GB1 K10D/T11Sep (Tb^{3+})	26	7.3 (0.1)	1.6 (0.1)	3.513	14.367	0.093	35	116	174	0.01
Ubi. T22Sep/N25D/K29Q (Tb^{3+})	20	3.5 (0.1)	1.3 (0.1)	5.505	1.144	−8.867	150	104	9	0.03

[a] The $\Delta\chi$-tensor fits used PCSs measured with Tb^{3+} and Tm^{3+}, using Y^{3+} as the diamagnetic reference. The metal coordinates and tensor parameters for the ubiquitin and GB1 mutants are reported relative to the NMR ensemble structure of ubiquitin (PDB ID: 2KOX; Fenwick et al., 2011) and the crystal structure of GB1 (PDB ID: 1PGA; Gallagher et al., 1994), respectively. [b] N: number of PCSs used in the fit. [c] Uncertainties (in brackets) were determined from fits obtained by randomly omitting 10 % of the PCS data. [d] The quality factor was calculated as the root-mean-square deviation between experimental and back-calculated PCSs divided by the root-mean square of the experimental PCSs.

Table 2. $\Delta\chi$-tensor parameters of the GB1 mutants K10Sep/T11Sep and A24Sep/K28Sep[a].

Mutant	N	$\Delta\chi_{ax}(10^{-32}\,m^3)$	$\Delta\chi_{rh}(10^{-32}\,m^3)$	x (Å)	y (Å)	z (Å)	$\alpha(°)$	$\beta(°)$	$\gamma(°)$	Q
K10Sep/T11Sep (Tb^{3+})	31	−14.5 (0.1)	−3.2 (0.1)	27.455	13.449	12.675	88	13	155	0.01
A24Sep/K28Sep (Tb^{3+})	34	34.7 (0.6)	5.3 (0.1)	17.628	34.049	21.869	178	46	69	0.02
A24Sep/K28Sep (Tm^{3+})	31	−15.5 (0.4)	−2. (0.1)	17.666	34.141	21.937	178	46	47	0.03

[a] The $\Delta\chi$-tensor fits used the crystal structure 1PGA (Gallagher et al., 1994) and the PCSs measured with Tb^{3+} (or Tm^{3+}) and Y^{3+}. See footnotes b–d of Table 1 for further details.

Disappointingly, neither Hsp90-N S36Sep/D40Sep nor ubiquitin T66Sep/H68Sep displayed any PCSs upon titration with paramagnetic lanthanides.

The difficulties in expressing most of the double-phosphoserine mutants were not due to expression into insoluble inclusion bodies, as we did not find the proteins in the insoluble fraction after cell lysis. As the read-through efficiency of amber stop codons has been reported to depend on neighbouring nucleotides (Pott et al., 2014), we tested the incorporation of Boc-lysine (BoK) to produce ubiquitin A28BoK/D32BoK, *E. coli* PpiB K25BoK/D29BoK and GB1 T51BoK/T53BoK, using a previously published pyrrolysyl-tRNA synthetase/tRNA pair (Bryson et al., 2017). All these proteins were expressed successfully (Fig. S4), demonstrating that the difficulty in expressing these mutants with two phosphoserine residues arises not simply from the difficulty in reading through two amber stop codons in the same gene. These observations suggest that too many negatively charged amino acids located in close proximity interfere with protein folding, making the protein prone to proteolytic degradation during overexpression. Likewise, the ubiquitin mutant A28Sep/D32Sep could not be overexpressed, whereas the single mutant A28Sep was produced in good yield. Unfortunately, ubiquitin A28Sep did not display PCSs following titration with $TbCl_3$ (data not shown).

3.6 Lanthanide binding by three amino acid side chains

The high failure rate of double-phosphoserine incorporation prompted us to carefully assess the two GB1 double-Sep mutants that did express and deliver PCSs. Notably, both constructs feature an additional glutamate residue near the lanthanide-binding site, which could potentially assist with the binding of the lanthanide. Specifically, Glu26 is near the lanthanide-binding site of GB1 A24Sep/K28Sep (Fig. 5a), and the side chain of Glu56 is near the loop region harbouring the K10Sep/T11Sep mutations and could point towards the two phosphoserines in the loop (Fig. 5b). Indeed, the lanthanide positions determined by the $\Delta\chi$-tensor fits are not simply between the two phosphoserine side chains, but potentially also within reach of the side-chain carboxyl groups of the nearby glutamate residues. The excellent Q factors associated with the $\Delta\chi$-tensor fits (Table 2) suggest that the metal positions are reliable. Notably, none of the other double-phosphoserine mutants investigated (Fig. S3) provided the possibility of additional lanthanide coordination by a negatively charged helper residue. To test the functional importance of Glu26, we produced the GB1 A24Sep/K28Sep/E26N triple mutant and probed for lanthanide binding. Indeed, this mutant produced no PCSs upon titration with $TbCl_3$.

3.7 Effect of salt bridges

In wild-type proteins, most aspartate and glutamate residues are located sufficiently close to positively charged side chains that they can engage in salt bridges. This raises the question whether such salt bridges can affect the lanthanide-binding affinity of sites constructed with negatively charged residues by compensating some of the negative charge. For

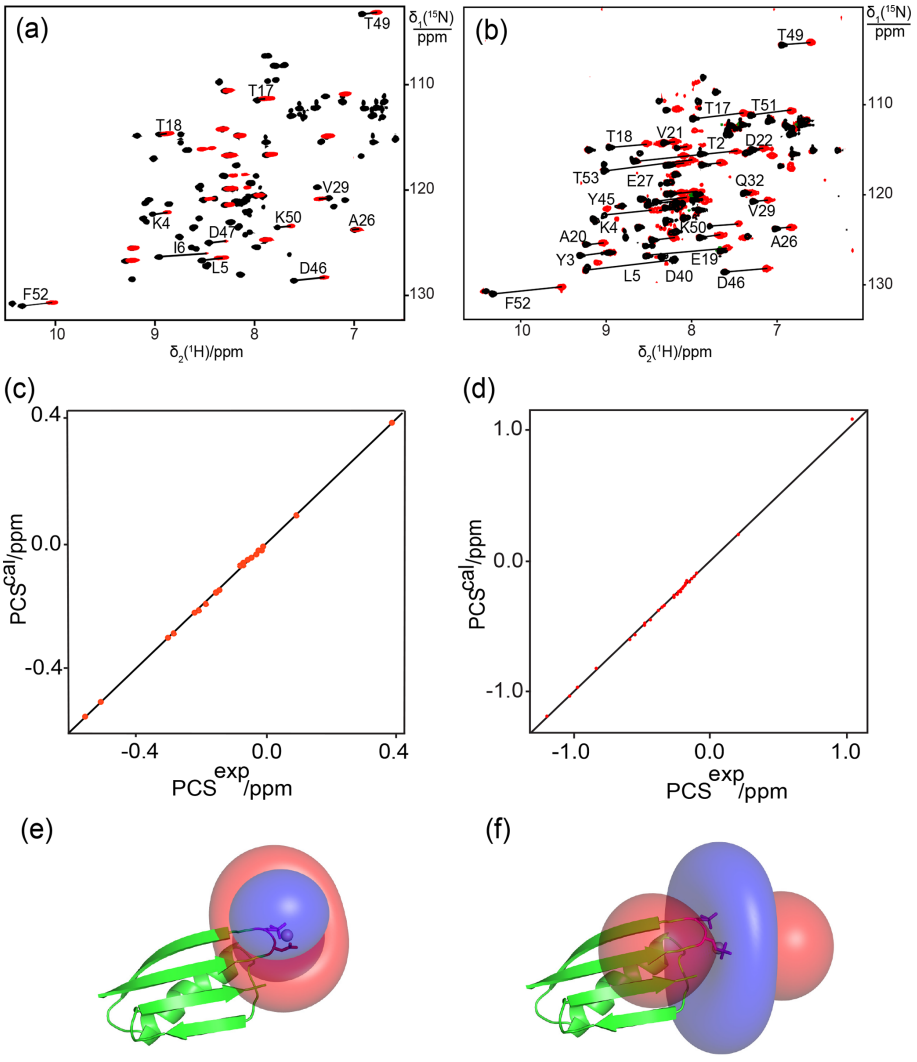

Figure 3. Close agreement between experimental and back-calculated PCSs of amide protons in the protein GB1 obtained with lanthanide-binding sites generated with one or two phosphoserine residues. **(a)** and **(b)**: superimposition of [^{15}N,^{1}H]-HSQC spectra of 0.3 mM solutions of GB1 K10D/T11Sep and GB1 K10Sep/T11Sep, respectively. The spectra were recorded in the presence of Tb^{3+} (red) or Y^{3+} (black). Lines connect cross-peaks belonging to the same residue in the paramagnetic and diamagnetic samples. **(c)** and **(d)**: correlation between back-calculated and experimental PCSs for GB1 K10D/T11Sep and GB1 K10Sep/T11Sep, respectively. **(e)** and **(f)**: location of the Tb^{3+} ion on the GB1 mutants K10D/T11Sep and K10Sep/T11Sep, respectively, and PCS isosurfaces plotted on the structure of GB1. Blue and red isosurfaces indicate PCSs of ± 1 ppm, respectively.

example, the ubiquitin mutant T22Sep/N25D features a lysine residue (K29) in the α-helix harbouring D25, with the potential to form a salt bridge (Fig. 6a). To test the effect of this interaction, we replaced K29 by glutamine in the mutant T22Sep/N25D/K29Q. Indeed, while the mutant T22Sep/N25D displayed only very small PCSs with Tb^{3+} ions if any (Fig. 7a), the mutant T22Sep/N25D/K29Q displayed PCSs up to 0.3 ppm (Fig. 7b, Table S1). Using the NMR ensemble structure of ubiquitin (PDB ID: 2KOX) and the measured PCSs, we determined the metal position in the triple mutant by fitting the $\Delta\chi$ tensor. The correlation between back-calculated and experimental PCSs was excellent, resulting in a Q factor of 0.03 (Fig. 7c, Table 1).

Similarly, fitting of a $\Delta\chi$ tensor to the small PCSs observed for the ubiquitin mutant Q2D/E64Sep (Fig. S5), which has a lysine residue in position 63, suggested metal coordinates far from the protein, which is a hallmark of a variable metal position (Shishmarev and Otting, 2013). Unfortunately, the attempt to remove the potential salt bridge between K63 and the Sep residue in position 64 in the triple mutant Q2D/K63Q/E64Sep resulted in a construct that failed to express.

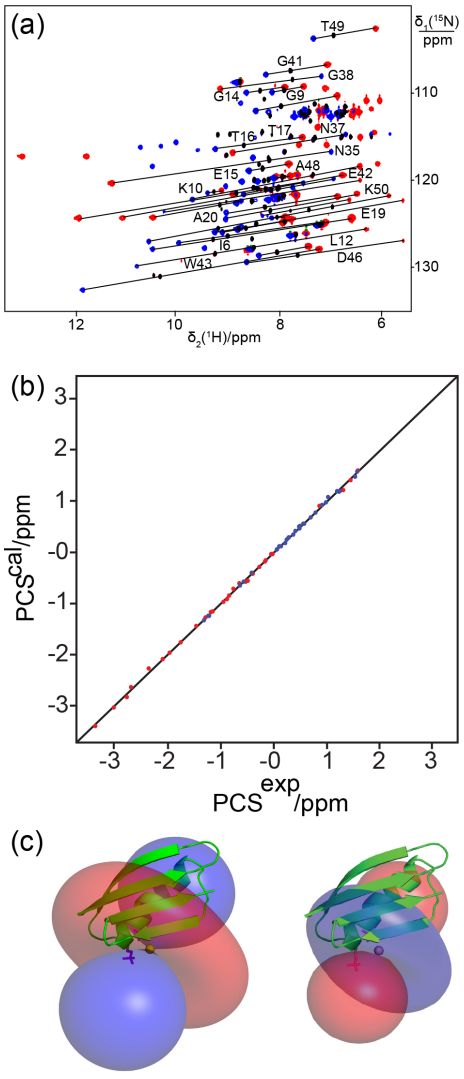

Figure 4. The double-phosphoserine mutant GB1 A24Sep/K28Sep generates high-quality PCSs. **(a)** Superimposition of [^{15}N,^{1}H]-HSQC spectra of 0.3 mM solutions of GB1 A24Sep/K28Sep in the presence of Tb^{3+} (red cross-peaks), Tm^{3+} (blue cross-peaks), or Y^{3+} (black cross-peaks). Lines were drawn to connect selected corresponding cross-peaks observed with diamagnetic and paramagnetic metal ions. **(b)** Correlation between back-calculated and experimental PCSs. **(c)** Blue and red isosurfaces indicating PCSs of ±1 ppm, respectively, as determined by the $\Delta\chi$ tensors of Tb^{3+} (left) and Tm^{3+} (right). The side chains of Sep residues modelled at positions 24 and 28 are highlighted by a stick representation.

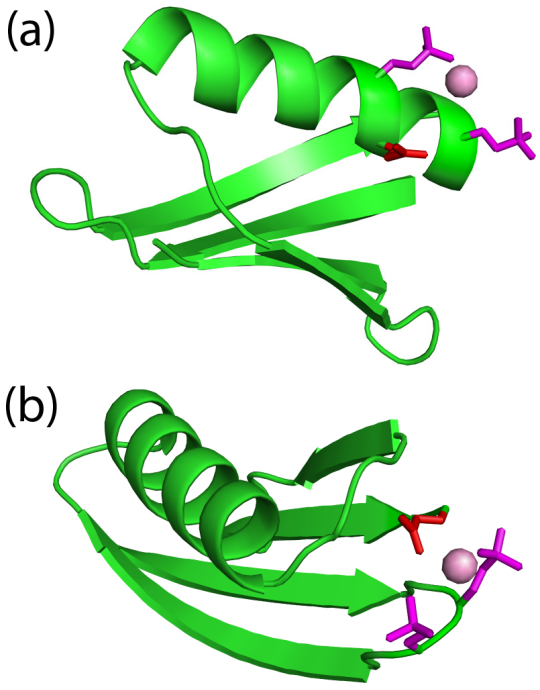

Figure 5. An additional glutamate residue acts as a helper residue to bind a lanthanide ion in double-phosphoserine mutants of GB1. **(a)** GB1 A24Sep/K28Sep with the side chains of the phosphoserine residues (purple) and Glu26 (red) modelled to indicate their possible proximity to a lanthanide ion (purple ball). A side-chain oxygen of Glu26 (in the conformation of the crystal structure 1PGA; Gallagher et al., 1994) is within 4.9 Å of the metal ion. **(b)** GB1 K10Sep/T11Sep showing the phosphoserine side chains in purple and Glu56 in red. Using the side-chain conformation of Glu56 in the crystal structure, a carboxylate oxygen is within 3.6 Å of the metal ion.

Attempts to express the ubiquitin mutant R54Sep and the GB1 mutant K50Sep failed. We speculate that this may be due to the destabilizing effect associated with the disruption of salt bridges involving these sites (Fig. S6a). Conversely, the ubiquitin mutant T55Sep and the GB1 mutant A24Sep expressed in high yield but did not display PCSs upon titration with paramagnetic lanthanides. The structure of ubiquitin indicates that a Sep residue in position 55 could form a salt bridge with R54, and the structure of GB1 suggests that a Sep residue in position 24 could form a salt bridge with K28 (Fig. S6b). These results suggest that the expression even of highly stable proteins like ubiquitin and GB1 can be affected by the presence of too many charges in close proximity, while compensating the negative charge density by salt bridges affects lanthanide binding.

3.8 Protein unfolding due to charge repulsion

Our failure to produce most of the proteins designed with two phosphoserine residues in close proximity led us to hypothesize that low expression yields could in part be caused by unfolding due to electrostatic repulsion, which would increase susceptibility to proteolytic degradation during expression in *E. coli*. Supporting evidence came from two observations. First, the GB1 mutant K4D/I6Sep displayed an NMR spectrum characteristic of an unfolded protein (Fig. S7). In wild-type GB1, Glu15 is in close proximity to Lys4 (Fig. 6c). By disrupting this salt bridge, the mutant K4D/I6Sep contains several uncompensated negative charges in close proximity

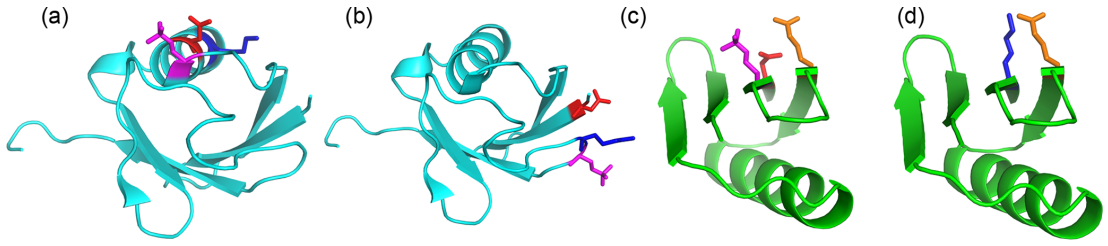

Figure 6. Single-site phosphoserine mutants tested for the effect of positively charged residues nearby. The locations of selected residues are highlighted by displaying side-chain atoms, with phosphoserine in magenta, aspartate in red, glutamate in orange and lysine residues in blue. Side-chain conformations are those of the crystal structure, except for phosphoserine, which was modelled. **(a)** Ubiquitin T22Sep/N25D. A lysine residue (K29) is located next to the engineered lanthanide-binding site in the same α-helix, where it can form a salt bridge with residue 25. **(b)** Ubiquitin Q2D/E64Sep. There is a lysine residue in position i-1 of the Sep residue. **(c)** Wild-type GB1 showing the salt bridge between K4 and E15. **(d)** GB1 K4D/I6Sep. Introduction of the aspartate and Sep residue resulted in denaturation of the protein (Fig. S7).

(Fig. 6d). Alternatively, Glu15 could also form a salt bridge with Lys13. Therefore we attempted to reduce the number of negative charges by producing the mutant K4D/I6Sep/E15Q. Unfortunately, this mutant failed to express.

The second piece of evidence of charge-driven unfolding came from phosphoserine mutants of Hsp90-N. Although the wild-type protein can be produced in good yield, the single-phosphoserine mutants K70Q/T71Sep, K70Q/N72Sep, K70Q/N72Sep/D69, N72Sep, Q54D/S57Sep and R98D/S99Sep (Fig. S8a) failed to express, and the mutants D88Sep/N91D, E162D/T163Sep and K160Q/E162D/T163Sep (Fig. S8b) were produced only in very low yields. Only the mutant N91Sep (Fig. S8c) expressed in sufficient yield for isotope labelling. Its [^{15}N,^{1}H]-HSQC spectrum showed evidence of partial unfolding, as the signals of many amides vanished, while new peaks appeared at chemical shifts characteristic of unfolded proteins. Assignment of the well-resolved cross-peaks by comparison with the wild-type protein showed that the β-sheet of Hsp90-N was conserved in the N91Sep mutant, whereas no evidence was found of structural conservation of the protein region near residue 91 (Fig. S9). Notably, Hsp90-N is a protein of limited stability that is prone to precipitation and degradation within a couple of days.

4 Discussion

The present study shows the potential of phosphoserine to generate lanthanide-binding sites on proteins. Using phosphoserine to construct lanthanide-binding sites in proteins is uniquely attractive for multiple reasons. (i) Systems are available to genetically encode phosphoserine as an unnatural amino acid for site-specific insertion into polypeptide chains (Pirman et al., 2015). This provides facile access to the requisite protein mutants. The main alternative way, in which lanthanide ions can be attached to an unnatural amino acid, relies on copper-catalysed click chemistry of alkyne tags with a site-specifically introduced *p*-azidophenylalanine residue (Loh et al., 2013, 2015). In our hands, about half of the proteins have proven to precipitate quantitatively when exposed to the copper catalyst. (ii) Phosphoserine allows us to construct the lanthanide-binding site without the need for post-translational modification by a lanthanide-binding chemical tag. Without the need for chemical modification, the approach is independent of the presence or absence of cysteine residues or whether the target protein tolerates the chemicals needed for specific tagging. (iii) The side chain of phosphoserine is relatively short, leading to a lanthanide position close to the protein backbone. This makes it easier to predict the position of the lanthanide ion relative to the protein. While a single phosphoserine residue is not sufficient to bind a lanthanide ion with high affinity, this study shows that a nearby aspartate residue can assist in forming a good lanthanide-binding site, with the lanthanide ion coordinated by both the phosphoserine and aspartate residues. This delivers a better localization of the lanthanide ion than most of the chemical tags designed for binding to cysteine residues and, hence, $\Delta\chi$-tensor fits with very small Q factors can be obtained. The small size of the Q factors also indicates that the introduction of a phosphoserine residue does not induce any significant conformational changes in the target protein. High-quality $\Delta\chi$-tensor fits open the door for exploiting PCSs as accurate long-range restraints in structural biology.

Exceptionally low Q factors were obtained for a lanthanide-binding site in GB1, which was made of two phosphoserine residues in positions i and $i+4$ of the α-helix together with Glu15. The site also generated relatively large $\Delta\chi$ tensors, indicating excellent immobilization of the metal ion relative to the protein (Shishmarev and Otting, 2013) as well as full conservation of the 3D structure of the protein. Two phosphoserine residues in a loop region of GB1 also produced a very small Q factor. It was disappointing, however, that attempts to produce other proteins with two phosphoserine residues met with a high failure rate. This may be explained by a failure to fold due too many negatively charged residues located in close proximity (Baneyx and Mu-

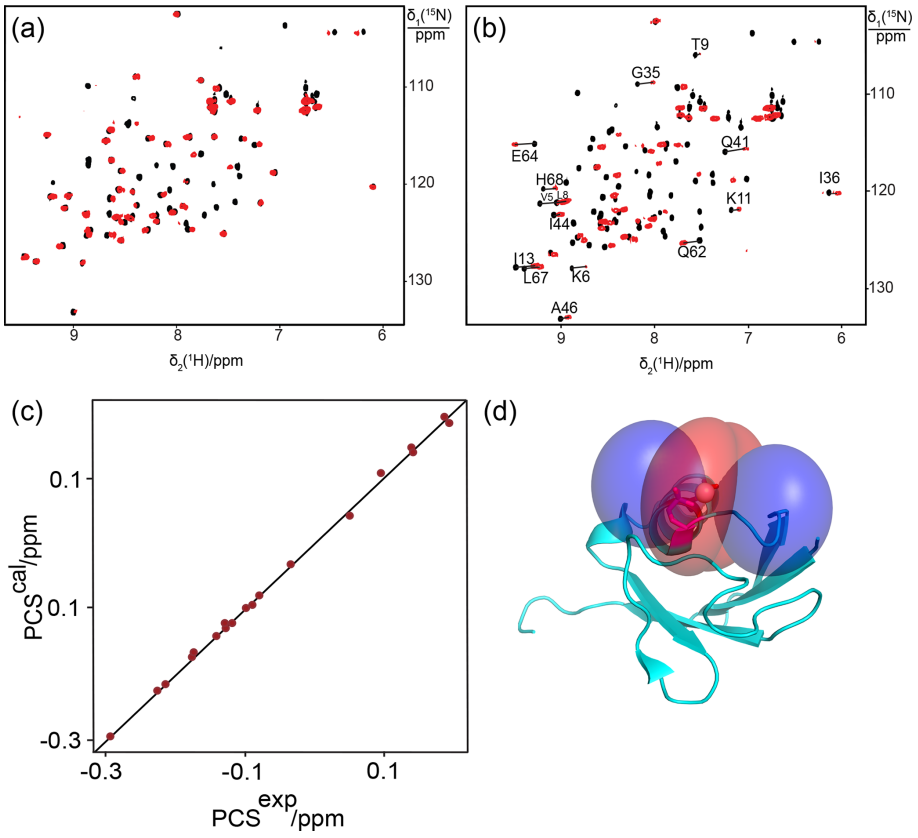

Figure 7. Breaking a salt bridge in ubiquitin T22Sep/N25D generates a specific lanthanide-binding site. **(a)** Superimposition of [^{15}N,^{1}H]-HSQC spectra of 0.3 mM solutions of ubiquitin T22Sep/N25D recorded in the presence of Tb^{3+} (red) or Y^{3+} (black). **(b)** Same as **(a)** but for ubiquitin T22Sep/N25D/K29Q. Lines connect cross-peaks belonging to the same residue in the paramagnetic and diamagnetic samples. **(c)** Correlation between back-calculated and experimental PCSs for ubiquitin T22Sep/N25D/K29Q with Tb^{3+}. **(d)** Blue and red PCS isosurfaces indicating PCSs of ±1 ppm, respectively. The side chains of D25 and the phosphoserine residue are highlighted by a stick representation.

jacic, 2004), resulting in degradation of the proteins during expression.

We succeeded in producing double-phosphoserine mutants of only two proteins other than GB1. These were the Hsp90-N mutant S36Sep/D40Sep and the ubiquitin mutant T66Sep/H68Sep. Both expressed in good yield but failed to produce PCSs with lanthanides. Furthermore, the absence of paramagnetic relaxation enhancements upon titration with lanthanides indicated the failure to bind. Inspection of the 3D structures of these proteins indicated that nearby residues with positively charged side chains were in positions capable of at least partially compensating the negative charges of the phosphoserine residues. The fact that the ubiquitin mutant T22Sep/N25D/K29Q produced much better PCSs than the mutant T22Sep/N25D (Fig. 7) illustrates the potentially detrimental effect of salt bridges on lanthanide binding.

The $\Delta\chi$ tensors obtained with Tm^{3+} instead of Tb^{3+} ions were unexpectedly low for the single-Sep mutants but not for the GB1 mutant A24Sep/K28Sep. We observed previously that the ratio between the $\Delta\chi_{ax}$ components of these two ions can vary between different tags and even for the same tag at different sites of a protein (Loh et al., 2015). These differences are not an artifact of fitting the tensors for Tm^{3+} and Tb^{3+} independently, as, with the exception of ubiquitin E18Sep, the fits converged to very similar metal positions (Tables 1 and 2). We do not understand the origin of different magnitudes of χ-tensor anisotropies for Tm^{3+} and Tb^{3+} ions. In addition, much larger $\Delta\chi$ tensors have been reported for sterically rigid cyclen CE4 tags (Joss and Häussinger, 2019), suggesting that a rigid ligand field promotes large $\Delta\chi$ tensors.

Using tags preloaded with metal ions saves subsequent titration of the protein samples after tagging. In our titration experiments, using calculated amounts of lanthanide ions did not always deliver the optimal ratio of metal to protein, resulting in over- or under-titration. The most successful approach to establishing the paramagnetic complexes was operational, namely by titrating until the original signal of the protein had vanished or, at least, substantially decreased (to avoid over-titration and potential binding to other

Please note the remarks at the end of the manuscript.

sites). Under-titration was not problematic, as the paramagnetic peaks are well resolved due to PCSs. Incomplete saturation with diamagnetic yttrium would mainly affect signals of amide protons near the metal-binding site, which are unobservable in the paramagnetic samples due to PREs.

In summary, when designing lanthanide-binding sites with phosphoserine residues, a single phosphoserine residue in combination with an aspartate can deliver binding affinities in the micromolar range, but positively charged side chains near the designed lanthanide-binding site can compromise its ability to bind lanthanides. At the same time, the difficulty in producing proteins that contain many negatively charged residues in close proximity points to the importance of salt bridges to ensure the structural integrity of proteins.

5 Conclusions

The present study demonstrates, for the first time, that a lanthanide-binding motif can be introduced into a protein via genetically encoded unnatural amino acids without further chemical modification. It is particularly promising that the lanthanide-binding motif can be generated in either an α-helix or a loop region by a single phosphoserine residue combined with an aspartate, provided these residues are not engaged in salt bridges. While two phosphoserine residues potentially bind lanthanide ions even more strongly, too many negatively charged residues in close proximity tend to severely affect the in vivo expression yields as well as the folding of the target protein. For proteins, where lanthanide-binding sites can successfully be installed with the help of phosphoserine residues, however, $\Delta\chi$ tensors of extraordinary quality can be obtained.

Code and data availability. The NMR spectra are available at https://doi.org/10.25911/5fc5bd5f0f872 (Mekkattu Tharayil et al., 2020).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/mr-1-1-2020-supplement.

Author contributions. GO initiated the project and edited the final version of the manuscript. SMT performed most experiments and wrote the first version of the manuscript. MCM produced and analysed all protein mutants with two Sep residues. CTL established the phosphoserine incorporation protocols and performed the first successful PCS measurements with lanthanide ions. IA produced wild-type and mutant samples of Hsp90-N, provided NMR resonance assignments and analysed partially unfolded mutants.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue "Robert Kaptein Festschrift". It is not associated with a conference. TS5

Acknowledgements. TS6

Financial support. This research has been supported by the Australian Research Council (grant no. FL170100019) and the Centre of Excellence for Innovations in Peptide and Protein Science, Australian Research Council (grant no. CE200100012). TS7

Review statement. This paper was edited by Rolf Boelens and reviewed by Claudio Luchinat and Marcellus Ubbink.

References

Abdelkader, E. H., Yao, X., Feintuch, A., Adams, L. A., Aurelio, L., Graham, B., Goldfarb, D., and Otting, G.: Pulse EPR-enabled interpretation of scarce pseudocontact shifts induced by lanthanide binding tags, J. Biomol. NMR, 64, 39–51, https://doi.org/10.1007/s10858-015-0003-z, 2016.

Baneyx, F. and Mujacic, M.: Recombinant protein folding and misfolding in *Escherichia coli*, Nat. Biotechnol., 22, 1399–1408, https://doi.org/10.1038/nbt1029, 2004.

Barthelmes, D., Gränz, M., Barthelmes, K., Allen, K. N., Imperiali, B., Prisner, T., and Schwalbe, H.: Encoded loop-lanthanide-binding tags for long-range distance measurements in proteins by NMR and EPR spectroscopy, J. Biomol. NMR, 63, 275–282, https://doi.org/10.1007/s10858-015-9984-x, 2015.

Barthelmes, D., Barthelmes, K., Schnorr, K., Jonker, H. R. A., Bodmer, B., Allen, K. N., Imperiali, B., and Schwalbe, H.: Conformational dynamics and alignment properties of loop lanthanide-binding-tags (LBTs) studied in interleukin-1β, J. Biomol. NMR, 68, 187–194, https://doi.org/10.1007/s10858-017-0118-5, 2017.

Barthelmes, K., Reynolds, A. M., Peisach, E., Jonker, H. R. A., DeNunzio, N. J., Allen, K. N., Imperiali, B., and Schwalbe, H.: Engineering encodable lanthanide-binding tags into loop regions of proteins, J. Am. Chem. Soc., 133, 808–819, https://doi.org/10.1021/ja104983t, 2011.

Bertini, I., Janik, M. B. L., Lee, Y.-M., Luchinat, C., and Rosato, A.: Magnetic susceptibility tensor anisotropies for a lanthanide ion series in a fixed protein matrix, J. Am. Chem. Soc., 123, 4181–4188, https://doi.org/10.1021/ja0028626, 2001.

Bleaney, B.: Nuclear magnetic resonance shifts in solution due to lanthanide ions, J. Magn. Reson., 8, 91–100, https://doi.org/10.1016/0022-2364(72)90027-3, 1972.

Bryson, D. I., Fan, C., Guo, L.-T., Miller, C., Söll, D., and Liu, D. R.: Continuous directed evolution of aminoacyl-tRNA synthetases, Nat. Chem. Biol., 13, 1253–1260, https://doi.org/10.1038/nchembio.2474, 2017.

Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G.: Designing logical codon reassignment – Expanding the chemistry in biology, Chem. Sci., 6, 50–69, https://doi.org/10.1039/C4SC01534G, 2014.

Please note the remarks at the end of the manuscript.

Fenwick, R. B., Esteban-Martín, S., Richter, B., Lee, D., Walter, K. F. A., Milovanovic, D., Becker, S., Lakomek, N. A., Griesinger, C., and Salvatella, X.: Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition, J. Am. Chem. Soc., 133, 10336–10339, https://doi.org/10.1021/ja200461n, 2011.

Gallagher, T., Alexander, P., Bryan, P., and Gilliland, G. L.: Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR, Biochemistry, 33, 4721–4729, https://doi.org/10.1021/bi00181a032, 1994.

Hass, M. A. S., Keizers, P. H. J., Blok, A., Hiruma, Y., and Ubbink, M.: Validation of a lanthanide tag for the analysis of protein dynamics by paramagnetic NMR spectroscopy, J. Am. Chem. Soc., 132, 9952–9953, https://doi.org/10.1021/ja909508r, 2010.

Jia, X., Yagi, H., Su, X.-C., Stanton-Cook, M., Huber, T., and Otting, G.: Engineering $[Ln(DPA)_3]^{3-}$ binding sites in proteins: a widely applicable method for tagging proteins with lanthanide ions, J. Biomol. NMR, 50, 411, https://doi.org/10.1007/s10858-011-9529-x, 2011.

Jones, D. H., Cellitti, S. E., Hao, X., Zhang, Q., Jahnz, M., Summerer, D., Schultz, P. G., Uno, T., and Geierstanger, B. H.: Site-specific labeling of proteins with NMR-active unnatural amino acids, J. Biomol. NMR, 46, 89, https://doi.org/10.1007/s10858-009-9365-4, 2009.

Joss, D. and Häussinger, D.: Design and applications of lanthanide chelating tags for pseudocontact shift NMR spectroscopy with biomacromolecules, Prog. Nucl. Mag. Res. Sp., 114–115, 284–312, https://doi.org/10.1016/j.pnmrs.2019.08.002, 2019.

Keizers, P. H. J. and Ubbink, M.: Paramagnetic tagging for protein structure and dynamics analysis, Prog. Nucl. Mag. Res. Sp., 58, 88–96, https://doi.org/10.1016/j.pnmrs.2010.08.001, 2011.

Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C. A., and Schuck, P.: High-precision isothermal titration calorimetry with automated peak-shape analysis, Anal. Chem., 84, 5066–5073, https://doi.org/10.1021/ac3007522, 2012.

Lammers, C., Hahn, L. E., and Neumann, H.: Optimized plasmid systems for the incorporation of multiple different unnatural amino acids by evolved orthogonal ribosomes, Chem Bio Chem, 15, 1800–1804, https://doi.org/10.1002/cbic.201402033, 2014.

Lee, S., Oh, S., Yang, A., Kim, J., Söll, D., Lee, D., and Park, H.-S.: A facile strategy for selective incorporation of phosphoserine into histones, Angew. Chem. Int. Edit., 52, 5771–5775, https://doi.org/10.1002/anie.201300531, 2013.

Loh, C. T., Ozawa, K., Tuck, K. L., Barlow, N., Huber, T., Otting, G., and Graham, B.: Lanthanide tags for site-specific ligation to an unnatural amino acid and generation of pseudocontact shifts in proteins, Bioconjugate Chem., 24, 260–268, https://doi.org/10.1021/bc300631z, 2013.

Loh, C.-T., Graham, B., Abdelkader, E. H., Tuck, K. L., and Otting, G.: Generation of pseudocontact shifts in proteins with lanthanides using small "clickable" nitrilotriacetic acid and iminodiacetic acid tags, Chem. Eur. J., 21, 5084–5092, https://doi.org/10.1002/chem.201406274, 2015.

Mekkattu Tharayil, S., Mahawaththa, M. C., Adekoya, I., Otting, G.: NMR spectra, Australian National University, https://doi.org/10.25911/5fc5bd5f0f872, 2020. TS8

Morgan, K.: Plasmids 101: origin of replication, available at: https://blog.addgene.org/plasmid-101-origin-of-replication (last access: 27 October 2020), 2014.

Neylon, C., Brown, S. E., Kralicek, A. V., Miles, C. S., Love, C. A., and Dixon, N. E.: Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: A model derived from DNA-Binding studies of mutant proteins by surface plasmon resonance, Biochemistry, 39, 11989–11999, https://doi.org/10.1021/bi001174w, 2000.

Nguyen, T. H. D., Ozawa, K., Stanton-Cook, M., Barrow, R., Huber, T., and Otting, G.: Generation of pseudocontact shifts in protein NMR spectra with a genetically encoded cobalt(II)-binding amino acid, Angew. Chem. Int. Ed., 50, 692–694, https://doi.org/10.1002/anie.201005672, 2011.

Nitsche, C. and Otting, G.: Pseudocontact shifts in biomolecular NMR using paramagnetic metal tags, Prog. Nucl. Mag. Res. Sp., 98–99, 20–49, https://doi.org/10.1016/j.pnmrs.2016.11.001, 2017.

Orton, H. W., Huber, T., and Otting, G.: Paramagpy: software for fitting magnetic susceptibility tensors using paramagnetic effects measured in NMR spectra, Magn. Reson., 1, 1–12, doi:https://doi.org/10.5194/mr-1-1-2020, 2020.

Otting, G.: Prospects for lanthanides in structural biology by NMR, J. Biomol. NMR, 42, 1–9, https://doi.org/10.1007/s10858-008-9256-0, 2008.

Parigi, G. and Luchinat, C.: NMR consequences of the nucleus-electron spin interactions, in: Paramagnetism in Experimental Biomolecular NMR, edited by: Luchinat, C., Parigi, G., and Ravera, E., RSC Publishing, Cambridge, United Kingdom, 1–41, https://doi.org/10.1039/9781788013291-00001, 2018.

Park, H.-S., Hohn, M. J., Umehara, T., Guo, L.-T., Osborne, E. M., Benner, J., Noren, C. J., Rinehart, J., and Söll, D.: Expanding the genetic code of *Escherichia coli* with phosphoserine, Science, 333, 1151–1154, https://doi.org/10.1126/science.1207203, 2011.

Pirman, N. L., Barber, K. W., Aerni, H. R., Ma, N. J., Haimovich, A. D., Rogulina, S., Isaacs, F. J., and Rinehart, J.: A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation, Nat. Commun., 6, 8130, https://doi.org/10.1038/ncomms9130, 2015.

Pott, M., Schmidt, M. J., and Summerer, D.: Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids, ACS Chem. Biol., 9, 2815–2822, https://doi.org/10.1021/cb5006273, 2014.

Qi, R. and Otting, G.: Mutant T4 DNA polymerase for easy cloning and mutagenesis, Plos One, 14, e0211065, https://doi.org/10.1371/journal.pone.0211065, 2019.

Saio, T. and Ishimori, K.: Accelerating structural life science by paramagnetic lanthanide probe methods, BBA-Gen. Subjects, 1864, 129332, https://doi.org/10.1016/j.bbagen.2019.03.018, 2020.

Saio, T., Ogura, K., Yokochi, M., Kobashigawa, Y., and Inagaki, F.: Two-point anchoring of a lanthanide-binding peptide to a target protein enhances the paramagnetic anisotropic effect, J. Biomol. NMR, 44, 157–166, https://doi.org/10.1007/s10858-009-9325-z, 2009.

Saio, T., Yokochi, M., Kumeta, H., and Inagaki, F.: PCS-based structure determination of protein–protein complexes, J. Biomol. NMR, 46, 271–280, https://doi.org/10.1007/s10858-010-9401-4, 2010.

Please note the remarks at the end of the manuscript.

Saio, T., Ogura, K., Shimizu, K., Yokochi, M., Burke, T. R., and Inagaki, F.: An NMR strategy for fragment-based ligand screening utilizing a paramagnetic lanthanide probe, J. Biomol. NMR, 51, 395, https://doi.org/10.1007/s10858-011-9566-5, 2011.

Shishmarev, D. and Otting, G.: How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study, J. Biomol. NMR, 56, 203–216, https://doi.org/10.1007/s10858-013-9738-6, 2013.

Su, X.-C. and Otting, G.: Paramagnetic labelling of proteins and oligonucleotides for NMR, J. Biomol. NMR, 46, 101–112, https://doi.org/10.1007/s10858-009-9331-1, 2010.

Swarbrick, J. D., Ung, P., Su, X.-C., Maleckis, A., Chhabra, S., Huber, T., Otting, G., and Graham, B.: Engineering of a bis-chelator motif into a protein α-helix for rigid lanthanide binding and paramagnetic NMR spectroscopy, Chem. Commun., 47, 7368–7370, https://doi.org/10.1039/C1CC11893E, 2011.

Swarbrick, J. D., Ung, P., Dennis, M. L., Lee, M. D., Chhabra, S., and Graham, B.: Installation of a rigid EDTA-like motif into a protein α-helix for paramagnetic NMR spectroscopy with cobalt(II) ions, Chem. Eur. J., 22, 1228–1232, https://doi.org/10.1002/chem.201503139, 2016.

Welegedara, A. P., Yang, Y., Lee, M. D., Swarbrick, J. D., Huber, T., Graham, B., Goldfarb, D., and Otting, G.: Double-arm lanthanide tags deliver narrow Gd^{3+}–Gd^{3+} distance distributions in double electron–electron resonance (DEER) measurements, Chemistry, 23, 11694–11702, https://doi.org/10.1002/chem.201702521, 2017.

Xie, Y., Jiang, Y., and Ben-Amotz, D.: Detection of amino acid and peptide phosphate protonation using Raman spectroscopy, Anal. Biochem., 343, 223–230, https://doi.org/10.1016/j.ab.2005.05.038, 2005.

Yagi, H., Loscha, K. V., Su, X.-C., Stanton-Cook, M., Huber, T., and Otting, G.: Tunable paramagnetic relaxation enhancements by $[Gd(DPA)_3]^{3-}$ for protein structure analysis, J. Biomol. NMR, 47, 143–153, https://doi.org/10.1007/s10858-010-9416-x, 2010.

Yang, A., Ha, S., Ahn, J., Kim, R., Kim, S., Lee, Y., Kim, J., Söll, D., Lee, H.-Y., and Park, H.-S.: A chemical biology route to site-specific authentic protein modifications, Science, 354, 623–626, https://doi.org/10.1126/science.aah4428, 2016.

Remarks from the language copy-editor

CE1 Please define “NMR” here, on first mention in the main body of the text, and in the main title.
CE2 Please confirm.
CE3 Please confirm.
CE4 Please confirm.

Remarks from the typesetter

TS1 Please send a new supplement as a *.pdf without the title, authors, correspondence author, etc. as we will generate a supplement title page during publication (with a citation including the DOI), which will contain this information.
TS2 What does the subscript "a" stand for? Please define.
TS3 Please provide short title.
TS4 Please define subscripts "ax" and "rh".
TS5 Please confirm.
TS6 Would you like to add an acknowledgement?
TS7 Please note that the funding information has been added to this paper. Please check if it is correct.
TS8 Please confirm added reference taken from the Assets.