Abstract: Pseudocontact shifts (PCS) induced by paramagnetic lanthanide ions provide unique long-range structural information in nuclear magnetic resonance (NMR) spectra, but the site-specific attachment of lanthanide tags to proteins remains a challenge. Here we incorporated p-azido-phenylalanine (AzF) site-specifically into the proteins ubiquitin and GB1, and ligated the AzF residue with alkyne derivatives of small nitriirotiacetic acid and iminodiacetic acid tags using the Cu(I)-catalyzed “click” reaction. These tags form lanthanide complexes with no or only a small net charge and produced sizeable PCSs with paramagnetic lanthanide ions in all mutants tested. The PCSs were readily fitted by single magnetic susceptibility anisotropy tensors. Protein precipitation during the click reaction was greatly alleviated by the presence of 150 mM NaCl.

Introduction

Paramagnetic NMR spectroscopy provides valuable information for studies of the structure and dynamics of biomolecules. Among the paramagnetic effects that can be measured by NMR, pseudocontact shifts (PCS) generated by lanthanide ions present particularly useful structural information. PCSs contain information about the spatial location of observed nuclear spins relative to a lanthanide ion, and this information can be obtained for nuclear spins as far as 40 Å from the lanthanide ion. PCSs are easily measured as the difference in chemical shifts between diamagnetic and paramagnetic samples.

In order to obtain PCSs from a lanthanide ion it needs to be attached to the protein in a site-specific manner. Intense efforts over the last decade have produced a range of lanthanide binding tags (LBT) for site-specific covalent attachment. Most LBTs are designed to bind to cysteine residues via formation of disulfide bonds or, less frequently, thioethers. These LBTs rely on the absence of competing cysteine residues, usually requiring the removal of naturally occurring cysteine residues before the introduction of one or two new cysteine residues in the protein by site-directed mutagenesis. In many cases, however, eliminating naturally occurring cysteine residues has proven difficult as alternative amino acids can severely destabilize a protein, especially if the protein contains disulfide bonds. Cysteine residues often also play critical functional roles in enzymes and metalloproteins.

An alternative approach is presented by the fusion of a lanthanide-binding peptide (LBP) to the C- or N-terminus of the target protein but this limits the metal ion position to sites near the N- or C-terminus of the protein. More importantly, it is difficult to produce LBP fusions that immobilize the metal ion position relative to the protein. The importance of immobilization is apparent in the equation governing the PCS of a nuclear spin:

\[ \Delta \rho_{\text{PCS}} = \frac{1}{12\pi r^2} |\Delta \chi| (3\cos^2 \theta - 1) + 1.5\Delta \chi \sin^2 \theta \cos 2\phi \]

where \( \Delta \rho_{\text{PCS}} \) is the PCS (measured in ppm), \( r \) is the distance of the nuclear spin from the metal ion, \( \Delta \chi_{ax} \) and \( \Delta \chi_{rh} \) are the axial and rhombic components of the magnetic susceptibility anisotropy (\( \chi \)) tensor, and \( \theta \) and \( \phi \) are the polar angles describing the position of the nuclear spin with respect to the principal axes of the \( \chi \) tensor.\(^7\) Equation 1 shows that PCSs can be positive or negative. As the \( \chi \) tensor is associated with the metal ion, reorientation movements of the metal complex relative to the target protein easily average the PCSs observed in the protein to zero. More critically, the distance dependence of the PCS implies that, strictly speaking, the PCS data can no longer be interpreted by a single \( \chi \) tensor, if the metal ion undergoes translational movements relative to the protein.\(^6\)

It has been shown that, with prior knowledge of the protein structure, improved immobilization of the lanthanide can be achieved by tying the LBP to the protein not only by an N- or C-terminal fusion but also by an additional disulfide bond. This, however, reintroduces the dependence on single-cysteine mutants. Furthermore, the quality of the paramagnetic NMR spectra depends on the length of the polypeptide that links tag and protein. To circumvent these problems, LBPs have been engineered into solvent-exposed protein loops, but few protein loops are amenable to this approach and the protein structure must be known in detail.\(^6\)

To overcome such limitations, a more versatile strategy has recently been established that relies on the site-specific incorporation of an unnatural amino acid with a uniquely reactive chemical group, allowing subsequent ligation of a tag molecule in a specific “bio-orthogonal” chemical reaction. The incorporation of the unnatural amino acid into a protein can readily be achieved by site-directed mutagenesis to introduce...
an amber stop codon (UAG) at the desired position in the gene of the target protein. Many orthogonal tRNA/aminoacyl-tRNA synthetase pairs have been evolved from Methanocaldococcus jannaschii genes by Schultz and co-workers, which selectively decode the amber stop codon for different unnatural amino acids.\[11]\ In the first demonstration of this strategy for the generation of PCSs, we attached two different cyclohexane-based LBTs (“C3- and C4-tags”) to p-azido-L-phenylalanine (AzF) using an orthogonal tRNA/aminoacyl-tRNA synthetase pair specific for AzF\[12]\ and Cu(I)-catalyzed click chemistry.\[13]\ These tags delivered good PCSs but more often than not resulted in imperfect correlations between experimental and back-calculation PCSs, which can be attributed to variable metal ion positions, owing to multiple tag conformations and a long tether between the lanthanide and the protein backbone.

In the present work, we explored a new set of LBTs for ligation to AzF in a Cu(I)-catalyzed click reaction. To achieve rigid positioning of the lanthanide ion relative to the protein, these tags retain free coordination sites on the lanthanide ion to allow additional coordination by a side-chain carboxyl group of a natural amino acid (glutamate or aspartate) of the protein. The concept of recruiting a side-chain carboxyl group for additional coordination to a lanthanide has previously been used successfully with iminodiacetic acid and nitritotriacetic acid tags with thiol groups (IDA-SH and NTA-SH, respectively) attached to cysteine residues in α-helices.\[14,30,32,34\] Here we used IDA and NTA tags with alkyne functional groups for ligation with AzF (Scheme 1). In the presence of a Cu(I) catalyst, the alkyne group of the tags undergoes a 1,3-dipolar cycloaddition reaction with the AzF residue in the protein to form a triazole-LBT. Apart from the C'-C' bond of the AzF residue, bond rotations in the tether of the reaction product change the position of the lanthanide only to a limited degree, despite large changes in its orientation. Therefore, the new tags perform remarkably well in generating PCSs that can be accurately interpreted by a single effective tensor, even without immobilization by additional coordination to a side-chain carboxyl group of the protein.

In addition, the present work uncovered a beneficial effect of NaCl in preventing protein precipitation during the click reaction.

**Results**

**Synthesis of the alkyne-NTA and alkyne-IDA tags:** The alkyne tags have been described previously in the context of other applications, and were prepared according to the reported literature methods with only minor variations.\[18\]

**Synthesis of proteins with AzF:** Experiments were performed with two different AzF mutants of uniformly $^{15}$N-labelled human ubiquitin (at positions 18 and 66) and three different mutants of uniformly $^{15}$N-labelled GB1 (at positions 6, 21 and 44), resulting in the mutants Ubi-E18AzF, Ubi-T66AzF, GB1-16AzF, GB1-V21AzF and GB1-T44AzF, respectively. All mutants were prepared **in vivo** following transformation of *E. coli* BLP21(DE3) with a T7 vector containing the gene of the target protein and the pEVOL vector to express the orthogonal tRNA/aminoacyl-tRNA synthetase pair for incorporation of AzF.

**Tagging of proteins:** The ligation reaction between tag and AzF was performed at room temperature with 10-fold excess of alkyne tag with the help of copper sulfate, BTTAA, sodium ascorbate, glycerol and aminoacuanidine. BTTAA is a Cu(I)-binding ligand that has been shown to enhance the yields of Cu(I)-catalyzed click reactions.\[13,16\] Aminoacuanidine has been reported to prevent protein degradation by byproducts of ascorbate oxidation.\[15,17\] Judged by the cross-peak intensities observed in $^{15}$N-HSQC spectra of the proteins recorded in the presence of paramagnetic lanthanides, the yields of tagged protein varied between 54 and 78%, with an average of 64%. Mass spectrometric analysis indicated that complete ligation yields were prevented by partial reduction of the azido group of the AzF residue to an amino group during protein expression. The degree of reduction varied strongly between different AzF sites (Figure S2).

While the click reaction proceeded without precipitation in the case of the ubiquitin samples, the reaction conditions resulted in partial precipitation of the GB1 mutants. We noted that the precipitation could be suppressed by the presence of 150 mM NaCl without any negative impact on the ligation yields. Consequently, all GB1 samples were ligated in the presence of 150 mM NaCl.

To explore the general validity of the salt effect, we also investigated AzF mutants of ERP29,\[18\] PpiB,\[16\] the intracellular domain of the p75 neurotrophin receptor (p75ICD)\[20\] and the
West Nile virus NS2B-NS3 protease (WNVpro) with respect to precipitation during the ligation reaction. All four proteins contained a C-terminal His$_6$-tag to facilitate the separation of full-length protein containing AzF from truncation products arising from recognition of the amber stop codon by release factor 1. The mutants were ERp29 S114AzF, PpiB H147AzF, p75ICD C416AzF and WNVpro H87AzF. All of them precipitated during the ligation reaction in the absence of salt. Like for GB1, salt suppressed precipitation of the PpiB and ERp29 mutants. For the p75ICD and WNVpro mutants, however, quantitative precipitation was observed even in the presence of salt. To probe whether the His$_6$-tag contributed to the precipitation, a construct of the WNVpro mutant was prepared, which allowed removal of the C-terminal His$_6$-tag by cleavage with TEV protease. The resulting protein still precipitated during the ligation reaction. To probe whether the cysteine residue in WNVpro (Cys 78) could cause the precipitation, we made the double-mutant H87AzF/C78V. This construct still precipitated. Removing, in addition, the C-terminal His$_6$-tag with TEV protease, however, greatly alleviated precipitation in the presence of 150 mM NaCl. These results point to His$_6$-tags as well as cysteine residues as triggers of precipitation, possibly due to coordination to copper. Nonetheless, salt suppressed precipitation of PpiB and ERp29 even though these proteins contained His$_6$-tags and cysteine residues.

Figure 1. Superimposition of $^{15}$N-HSQC spectra of 0.1 mM solutions of uniformly $^{15}$N-labelled AzF mutants of ubiquitin and GB1 ligated with different NTA and IDA tags by Cu(I)-catalyzed click chemistry. Spectra of complexes with diamagnetic Y$^{3+}$ are in black, and with paramagnetic Tb$^{3+}$ and Tm$^{3+}$ in blue and red, respectively. All spectra were measured in 50 mM HEPES buffer, pH 7.0, at 25 °C on a 600 MHz NMR spectrometer equipped with a cryoprobe. Selected pairs of cross-peaks in the diamagnetic and paramagnetic samples are connected by lines and labelled with their resonance assignments. A) Ubi-E18AzF ligated with alkyne-NTA tag, B) Ubi-T66AzF C) GB1-V21AzF ligated with IDA1 tag, D) GB1-T44AzF ligated with alkyne-IDA2 tag.
PCS measurements: Pronounced PCSs were observed when Tm$^{3+}$ or Tb$^{3+}$ were added to the tagged AzF-proteins. Figure 1A and B show the results for two of the ubiquitin mutants, Ubi-E18AzF and Ubi-T66AzF, following ligation with the alkyne-NTA tag and loading with either diamagnetic Y$^{3+}$ or paramagnetic Tm$^{3+}$ or Tb$^{3+}$. The examples shown in Figure 1C and D are of two of the GB1 mutants, GB1-V21AzF and GB1-T44AzF, with alkyne-IDAl and alkyne-IDA2 tag, respectively. Single resonances were observed for each backbone amide in the paramagnetic states, indicating that the tags did not produce diastereomers with the protein,[34,22] although the coordination of the lanthanide ion is, in principle, chiral. This observation is in agreement with previous results on PCSs from NTA- and IDA-tags tied to the protein via disulfide bonds.[36,34,14] The apparent absence of diastereomers can be attributed to rapid interconversion of any different enantiomeric forms of the lanthanide complexes.

For most amides, the PCSs generated by Tm$^{3+}$ and Tb$^{3+}$ shifted the $^{15}$N-HSQC cross-peaks in opposite directions, facilitating the assignment of at least a set of cross-peaks in the paramagnetic samples. Further cross-peak assignments in the paramagnetic states used the initial set of PCSs to fit $\Delta\chi$ tensors to the respective protein crystal structures and predict the PCSs of the remaining amide protons, leading to iterative assignment of the amide cross-peaks in the paramagnetic states and refined $\Delta\chi$ tensors. PCSs were measured for amide proton chemical shifts only, avoiding the significant residual anisotropic chemical shift effects associated with $^{15}$N nuclei.[23]

Identification of the metal position: While $\Delta\chi$ tensor fits readily deliver the metal ion coordinates that best explain the PCSs observed in the protein, mobility of the metal complex relative to the protein can result in chemically unrealistic metal ion coordinates when a fit with a single effective $\Delta\chi$ tensor is attempted. The magnitude of the fitted tensor is in turn sensitive to the metal ion position. For example, if the paramagnetic centre is positioned far from the protein, the fit of the experimental PCSs requires a correspondingly large $\Delta\chi$ tensor.[19] To restrain the tensor magnitudes, we restrained the metal ion coordinates to positions compatible with the covalent structure of the tag. The best fit was identified as the fit that best explained the PCSs using a large set of metal ion positions modelled by a rotamer library approach. The rotamer library was established by systematically changing the dihedral angles of the rotatable bonds of the tag identified in Scheme 1. $\Delta\chi$ tensors were fitted to each model for which the tag rotamer was stERICally compatible with the protein structure.

Table 1 shows the $\Delta\chi$ tensor parameters obtained with alkyne-NTA tags loaded with Tb$^{3+}$ and Tm$^{3+}$ for two different ubiquitin AzF mutants and three different GB1 AzF mutants. The axial tensor values were generally larger than those observed previously for proteins with thiol-NTA tags.[34,14] The large tensor magnitudes suggest that either the conformation of the AzF-alkyne-NTA conjugate is relatively rigid or the vicinity of the triazole increases the magnitude of the $\Delta\chi$ tensor.

### Table 1. $\Delta\chi$ tensor parameters of AzF mutants of ubiquitin and GB1 with alkyne-NTA tag loaded with Tb$^{3+}$ and Tm$^{3+}$

<table>
<thead>
<tr>
<th>Protein</th>
<th>$\Delta\chi_{ax}$</th>
<th>$\Delta\chi_{rh}$</th>
<th>Q</th>
<th>dA$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ubi-E18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>-17.6 (0.5)</td>
<td>-7.4 (0.2)</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>8.5 (0.2)</td>
<td>0.9 (0.1)</td>
<td>0.11</td>
<td>1.2</td>
</tr>
<tr>
<td>Ubi-T66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>-12.4 (0.1)</td>
<td>-2.7 (0.1)</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>8.7 (0.1)</td>
<td>2.6 (0.2)</td>
<td>0.03</td>
<td>1.0</td>
</tr>
<tr>
<td>GB1-16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>7.2 (0.4)</td>
<td>3.7 (0.3)</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>-5.8 (0.3)</td>
<td>-1.9 (0.2)</td>
<td>0.24</td>
<td>4.4</td>
</tr>
<tr>
<td>GB1-V21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>-20.5 (0.8)</td>
<td>-10.9 (0.3)</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>16.3 (1.7)</td>
<td>7.8 (1.2)</td>
<td>0.15</td>
<td>4.2</td>
</tr>
<tr>
<td>GB1-T44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>-22.7 (5.5)</td>
<td>-11.3 (3.2)</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>14.3 (1.3)</td>
<td>3.3 (1.0)</td>
<td>0.18</td>
<td>102</td>
</tr>
</tbody>
</table>

*a* The axial and rhombic components of the $\Delta\chi$ tensors are given in $10^{-24}$ m$^2$ and the Euler angles in degrees, using the xyz convention and unique tensor representation.[34] The metal coordinates and tensor parameters for the ubiquitin and GB1 mutants are reported relative to the first conformer of the NMR structure of ubiquitin (PDB ID 1D32)[21] and relative to the crystal structure of GB1 (PDB ID 1PGA)[24], respectively. Fits were performed using a rotamer library to identify all conformationally possible metal ion positions. In a second step, $\Delta\chi$ tensors were fitted using the PCS data measured with Tb$^{3+}$ and Tm$^{3+}$. Standard deviations (in brackets) were determined from fits obtained by using the same metal ion position while randomly omitting 10% of the PCS data. Quality factors were calculated as the root-mean-square deviation between experimental and back-calculated PCSs divided by the root-mean-square of the experimental PCSs.

*b* The sites of AzF incorporation are indicated.

*c* The distance $d$ reports the shift in the metal ion position between the full $\Delta\chi$ tensor fit, obtained by including the metal ion coordinates among the parameters to be fitted, and the $\Delta\chi$ tensor fit that restrained the metal coordinates to those found in the rotamer library as described in footnote a.

The V21AzF and T44AzF mutants of GB1 produced $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ components of the same sign as the ubiquitin mutants above, albeit with greater magnitude. The larger quality factors indicate that the tensor fits are not as good for these mutants, suggesting greater metal ion movements. For the I6AzF mutant, the quality factors are also quite large and the tensors show inverted signs, suggesting even greater metal ion movements.
The same three mutants of GB1 were also ligated with the alkyne-IDA1 and alkyne-IDA2 tags. Among the tensor fits (Table 2), only the T44AzF mutant with IDA1 tag combined small quality factors for both lanthanides with $\Delta \chi$ tensor magnitudes that are in the range of literature values for IDA tags (about $20 \times 10^{-32}$ m$^3$ for Tb$^{3+}$ and $10 \times 10^{-32}$ m$^3$ for Tm$^{3+}$).\cite{3q,14,27}

The reliability of the metal ion position found by the rotamer library approach was assessed by comparison with the metal ion coordinates found by a full $\Delta \chi$ tensor fit that included the metal ion coordinates as fitting parameters. A small shift in the best-fitting metal ion coordinates between both fitting approaches indicates a reliable metal ion position. In the case of Ubi-E18AzF and Ubi-T66AzF ligated with the alkyne-NTA tag, the metal ion positions determined by the full $\Delta \chi$ tensor fit were within 1.2 Å from the positions determined by the rotamer library approach, confirming the metal ion locations. In contrast, all GB1 mutants with alkyne-NTA tag displayed metal shifts of at least 4.2 Å between the positionally restrained and unrestrained fitting approaches. Among the GB1 mutants with alkyne-IDA1 and alkyne-IDA2 tags, metal ion shifts < 2.5 Å were obtained only for the mutants I6AzF and T44AzF.

Table 2. $\Delta \chi$ tensor parameters of AzF mutants of GB1 with alkyne-IDA1 and alkyne-IDA2 tags loaded with Tb$^{3+}$ and Tm$^{3+}$.

<table>
<thead>
<tr>
<th></th>
<th>$\Delta \chi_{ax}$</th>
<th>$\Delta \chi_{rh}$</th>
<th>$Q$</th>
<th>$d/Å$</th>
<th>$\Delta \chi_{ax}$</th>
<th>$\Delta \chi_{rh}$</th>
<th>$Q$</th>
<th>$d/Å$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>alkyne-IDA1 tag</td>
<td></td>
<td></td>
<td></td>
<td>alkyne-IDA2 tag</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB1-I6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>-24.5 (2.2)</td>
<td>-11.0 (1.6)</td>
<td>0.12</td>
<td></td>
<td>-24.5 (1.5)</td>
<td>-8.5 (0.8)</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>4.0 (0.1)</td>
<td>0.9 (0.1)</td>
<td>0.12</td>
<td>1.0</td>
<td>-10.3 (0.3)</td>
<td>-5.8 (0.4)</td>
<td>0.16</td>
<td>0.9</td>
</tr>
<tr>
<td>GB1-V21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>-31.5 (1.3)</td>
<td>-12.4 (0.3)</td>
<td>0.03</td>
<td></td>
<td>-46.2 (5.4)</td>
<td>-24.1 (3.8)</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>17.7 (0.4)</td>
<td>9.4 (0.2)</td>
<td>0.08</td>
<td>11</td>
<td>-13.0 (1.2)</td>
<td>-7.8 (0.8)</td>
<td>0.11</td>
<td>18</td>
</tr>
<tr>
<td>GB1-T44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb$^{3+}$</td>
<td>16.9 (0.2)</td>
<td>1.5 (0.6)</td>
<td>0.07</td>
<td></td>
<td>12.9 (0.1)</td>
<td>3.5 (0.6)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Tm$^{3+}$</td>
<td>-12.8 (0.1)</td>
<td>-4.1 (0.6)</td>
<td>0.06</td>
<td>2.5</td>
<td>-4.3 (3.5)</td>
<td>-2.7 (1.9)</td>
<td>0.13</td>
<td>3.8</td>
</tr>
</tbody>
</table>

* See footnotes of Table 1.

Figure 2. Models of AzF mutants of Ubi-E18AzF and Ubi-T66AzF ligated with alkyne-NTA tag. The models display the first conformer of the NMR structures of human ubiquitin (PDB ID 1D3Z)\cite{25} after modelling the tag on the protein and performing a comprehensive search of the rotamer space to find the best $D_c$ tensor compatible with any of the tag models. The tags and nearby glutamate residues are shown in a stick representation coloured with the colour of the backbone segment of their location in the amino acid sequence, with nitrogen atoms in dark blue, oxygens in red and the lanthanide ion displayed as a magenta sphere. A) Ubi-E18AzF. The metal ion is within 4.8 Å of the nearest carboxyl oxygen of Glu16. B) Ubi-T66AzF. The metal ion is within 1.7 Å of the nearest carboxyl oxygen of Glu64.

Immobilization of the metal ion: Figure 2 shows that the $\Delta \chi$ tensor fits located the lanthanide ions as close as 4.8 Å and 1.7 Å from the side-chain carboxyl groups of Glu16 on Ubi-E18AzF and Glu64 on Ubi-T66AzF, respectively. This suggests that these glutamate side chains occupy some of the remaining coordination sites of the lanthanide ion, assisting its
immobilization. Unique locations of the metal ion are also supported by small values of the quality factors, indicating the validity of fitting the PCSs with a single $\Delta \chi$ tensor. Measurements of the residual dipolar couplings (RDC) between amide nitrogens and amide protons on a 800 MHz NMR spectrometer, however, indicated that the lanthanide ions retained residual mobility: the RDC values ranged from -5.1 to 4.1 Hz for Ubi-E18AzF and -4.1 to 4.2 Hz for Ubi-T66AzF ligated with alkyne-NTA-Tm$^{3+}$ tag and thus are smaller than the maximal RDC of about 6.5 Hz expected for $\Delta \chi_{\text{fit}} = -12 \times 10^{-32}$ m$^3$. Therefore, additional interactions of the lanthanide ion with side-chain carboxyl groups of the protein may be an important aid in immobilizing the lanthanide ion but they are not sufficient to achieve complete immobilization. Notably, however, modelling indicated that the Ubi-T66AzF mutant could accommodate the tag in two different $\chi_\alpha$ rotamers, but the $\chi_\beta$ rotamer found experimentally corresponds to the one that placed the lanthanide in proximity of a negatively charged side chain.

Compared with the ubiquitin mutants containing AzF with alkyne-NTA tag, the corresponding GB1 constructs displayed greater variability in tensor signs (GB1-I6AzF; Table 1), generally larger quality factors, and less conserved lanthanide ion positions, when full $\Delta \chi$ tensor fits were compared with rotamer library-restricted fits (Table 1). This indicates that the metal ion was not well immobilized in the GB1 mutants with alkyne-NTA tag. Indeed, the best $\Delta \chi$ tensor fits obtained in the rotamer-library approach did not place the lanthanide ion sufficiently close to any of the carboxyl groups in the protein for direct coordination.

Table 2 reports the $\Delta \chi$ tensor fits for the same GB1 AzF mutants, but with alkyne-IDA1 and alkyne-IDA2 tags. With two exceptions, no carboxyl group of the protein was near the lanthanide ion positions determined by the $\Delta \chi$ tensor fits. The exceptions are the mutants I6AzF and T44AzF with alkyne-IDA1 tag, with distances from the nearest carboxyl oxygen of the protein of 5 Å from Glu15 and 4.6 Å from Glu42, respectively, based on the crystal structure 1PGA[28]. Potentially, this suggests some immobilization of the tag due to electrostatic attraction with the positively charged IDA-lanthanide complex. GB1-I6AzF with alkyne-IDA1 tag, however, produced $\Delta \chi$ tensors of Tb$^{3+}$ and Tm$^{3+}$ with unexpected relative magnitudes and the quality factors were not particularly small (Table 2), suggesting that the metal may not be well immobilized after all. For all other constructs with alkyne-IDA1 or alkyne-IDA2 tag, no protein carboxyl group seems capable of stably coordinating the lanthanide ion and significant translational movements of the lanthanide are indicated by unexpected tensor magnitudes or signs, large quality factors, or metal ion positions strongly depending on whether the $\Delta \chi$ tensor fit was restrained by a rotamer library or not.

Most importantly, every single AzF mutant of ubiquitin and GB1 prepared produced a good, or at least acceptable, correlation between back-calculated and experimental PCSs, in particular for the samples ligated with the alkyne-IDA1 tag (Figures 3 and S1). Therefore, the PCSs measured with alkyne-NTA and alkyne-IDA tags ligated to AzF residues are highly suitable for generating structural restraints.

**Figure 3.** Correlations between back-calculated and experimental PCSs of A) Ubi-E18AzF ligated with alkyne-NTA tag, B) Ubi-T66AzF ligated with alkyne-NTA, C) GB1-V21AzF ligated with IDA1 tag, D) GB1-T44AzF ligated with alkyne-IDA2 tag. The data from complexes with Tb$^{3+}$ and Tm$^{3+}$ are shown in black and red, respectively. The fits were performed using the rotamer library approach to restrain the metal ion positions to sites that are compatible with the covalent structures of the tags.

**Discussion**

**Protein tagging using AzF:** The overwhelming majority of LBTs produced to date for site-specific attachment to proteins make use of one or two cysteine residues as the chemically most reactive site in a target protein.[11,28] Tagging of proteins containing cysteine residues, however, requires independence of sulfur chemistry. The present work demonstrates that alkyne-NTA and alkyne-IDA tags can be ligated with p-azidophenylalanine in a Cu(I)-catalysed cycloaddition reaction to produce the smallest LBTs yet available that use an unnatural amino acid for site-specific incorporation. Despite the fairly long tether created in the reaction between protein backbone and paramagnetic centre, good correlations between back-calculated and experimental PCSs were obtained, even without the additional immobilization that could be gained by coordination of the lanthanide ion by a carboxyl group of an amino-acid side chain in the protein. This suggests that the generation of PCSs by these tags is robust, making them suitable for a wide range of proteins.
Substantial translational movements of the metal ion tend to be valid. Investigated in the present work, these predictions a single average averaging, out small changes in produced by the phenylene and triazole rings, is sensitive to metal ion position, which is located at the end of the long tether dihedral angle is centre at a chemically plausible sites. For tags ligated to p-azido-phenylalanine can also be incorporated into proteins by an orthogonal [30 mRNA/mRNA-synthetase system. This amino acid would be inert towards intracellular reducing conditions, and an NTA or IDA tag with an azide group could yield the same final product as the react. It has been shown that p-ethylphénylalanine can also be incorporated into proteins by an orthogonal [31 mRNA/mRNA-synthetase system. This amino acid would be inert towards intracellular reducing conditions, and an NTA or IDA tag with an azide group could yield the same final product as the react. Unfortunately, this process of the orthogonal systems of Scheme 1. Unfortunately, the synthesis of NTA and IDA tags with azide groups would be much more demanding.

Representing PCSs by a single effective $\Delta \chi$ tensor: Excessive mobility of a lanthanide-binding tag can compromise its utility. Ultimately, to produce useful structural restraints, paramagnetic LBTs must generate sizeable PCSs, and back-calculated and observed PCSs must correlate closely. So long as the fitted effective $\Delta \chi$ tensor explains the data well, it is thus of lesser importance whether the fit places the paramagnetic centre at a chemically plausible location or not. For tags ligated to p-azido-phenylalanine by cyclodaidation, the most critical dihedral angle is the $\chi_1$ angle of the A2F residue, because the metal ion position, which is located at the end of the long tether produced by the phenylene and triazole rings, is sensitive to small changes in $\chi_1$ angle. Rotations around bonds near the outer end of the tag, i.e. around the bonds following the triazole ring, reduce the magnitude of the PCSs in the protein by averaging, without greatly compromising their representation by a single average $\Delta \chi$ tensor.

If the three-dimensional structure of the protein is known, the potential of an aromatic amino acid to populate one or more different $\chi_1$ rotamers can be estimated using the mutation tool in PyMOL. In the case of the ubiquitin and GB1 mutants investigated in the present work, these predictions appeared to be valid.

Assessing the mobility of lanthanide tags: Several criteria can be used as evidence for mobile lanthanide tags. (i) Substantial translational movements of the metal ion tend to generate large PCSs in an extended area near the tagging site, causing unconstrained $\Delta \chi$ tensor fits to position the paramagnetic centre far from the protein. In this situation, a very different metal ion position results from an unrestrained fit and a fit that restricts the metal ion position to chemically plausible sites. (ii) Large quality factors associated with any $\Delta \chi$ tensor fit where the metal ion is restricted to a chemically plausible position indicate difficulties to fit the data with a single effective $\Delta \chi$ tensor. (iii) The alignment tensor due to magnetic alignment is proportional to the $\Delta \chi$ tensor. Therefore, if the largest RDC measured is much smaller than expected from the $\Delta \chi_{\text{obs}}$ value, this is evidence for movement of the metal ion, although the magnitude of the effect depends strongly on the details of the metal ion movement. (iv) Very small $\Delta \chi$ tensors indicate substantial averaging between different tensor orientations. The experimental results obtained in the present work show that mobility is most clearly revealed by a combination of the criteria above and that alkyne-NTA and alkyne-IDA tags clicked onto A2F deliver acceptable but incomplete immobilization of the lanthanide ion.

Simultaneous lanthanide ion complexation by two different coordinating moieties: Previous work showed that two NTA-SH tags ligated to a pair of cysteine residues in neighboring turns of an $\alpha$-helix confers excellent immobilization of a lanthanide ion that simultaneously binds to both NTA moieties. Similarly, an IDA-SH tag ligated to a cysteine in position i of an $\alpha$-helix can cooperate with an aspartic acid residue in position i+4 to immobilize a lanthanide ion and lanthanide complexes with dipicolinic acid tags have also been noted to co-opt side-chain carboxyl groups for additional coordination. By the same mechanism, the ubiquitin mutant T66A2F with alkyne-NTA tag seems to capture the lanthanide ion by the NTA moiety and, simultaneously, by the side-chain carboxyl group of Glu64. Unfortunately, this situation proved to be the exception rather than the rule. In almost all other examples of the present work, the tags positioned the lanthanide ions too far from protein carboxyl groups for direct coordination, indicating that the remaining coordination sites of the lanthanides were filled by hydration water even when model building indicated the possibility of direct coordination by a protein carboxyl group. This result agrees with the observation that salt bridges in proteins rarely lead to stable direct associations of carboxyl and amino groups. Nonetheless, the metal ion positions seemed to be biased towards side-chain carboxyl groups, suggesting that transient complex formation or longer-range electrostatic interactions help to confine the metal ion position.

While NTA tags bind lanthanides very tightly, the affinity to IDA tags is considerably lower. In the present work, we extensively explored IDA tags to increase the chances of additional coordination by carboxyl groups of the protein, considering that the net charge of lanthanide-NTA complexes is zero whereas the complex with an IDA tag retains a positive net charge. The experimental results showed, however, that the IDA tags resulted in no better immobilization of the lanthanide ion. In particular, the alkyne-IDA2A tag was designed with a flexible amide bond to facilitate the reorientation of the IDA group.
relative to the tether, in case this would promote coordination to a carboxyl group of the protein. The results, however, did not show any evidence for easier capture of the lanthanide ion by protein carboxyl groups. Instead, the alkyn-IDA2 tag produced overall smaller PCSs in the NMR spectra (Figure S1), indicating greater mobility of the lanthanide. This corroborates our previous results with IDA-SH tags, which showed that small deviations from ideal geometry upset additional coordination by protein carboxyl groups, even if they are in close proximity. In the light of these results, we prefer the alkyn-NTA tag over the alkyn-IDA tags for its better binding affinity.

**PCSs from tags attached by azide-alkyne cycloaddition:**

Most important and initially unexpected was the observation that paramagnetic lanthanides associated with the alkyn-NTA and alkyn-IDA tags ligated to AzF readily confer sizeable PCSs to the proteins they are attached to, irrespective of the presence of additional protein carboxyl groups to assist the coordination of the lanthanide ion. This is a remarkable result, as immobilization of tags attached via a single disulfide bond usually relies on short tethers and/or bulky LBTs such as DOTA derivatives. Tags based on phenylalanine have the special advantage that the \( \chi_1 \) and \( \chi_2 \) angles of phenylalanine can be predicted with good reliability, as the sterically allowed ranges of both dihedral angles are significantly limited by the bulkiness of the phenyl ring. Protein crystal structures and theoretical predictions show that the \( \chi_2 \) angle of solvent-exposed phenylalanines usually is close to 90° and the preference in \( \chi_2 \) angle helps restricting the \( \chi_1 \) angle to populate a single rotamer. In addition, the tether produced in the cycloaddition reaction is remarkably rigid, as the phenyl and triazole rings can be assumed to be co-planar due to double-bond conjugation and changes in \( \chi_2 \) angle reorient the triazole ring without relocating it (Scheme 1). A 180° flip around either the bond defined by the \( \chi_2 \) angle or the bond between phenyl and triazole rings would displace the metal ion linked to the triazole ring, but this displacement is no more than ± 5.3 Å in the case of the alkyn-NTA tag, ± 3.5 Å in the case of the alkyn-IDA1 tag and ± 6.3 Å in the case of the alkyn-IDA2 tag.

In the case of the alkyn-IDA1 tag, coordination of the lanthanide ion by the nearest nitrogen of the triazole ring would further restrict the movements of the lanthanide ion. For comparison, the established NTA-SH and IDA-SH tags can displace the metal ion by up to ± 4.5 Å. Another advantage of the AzF-triazole conjugate lies in its capability to position the metal ion at a greater distance from the protein surface than the disulfide bridge formed by the NTA-SH and IDA-SH tags. This alleviates the impact of translational metal ion movements on the approximation made when fitting the PCSs of the protein by a single effective \( \Delta \chi \) tensor.

**Conclusions**

In conclusion, the alkyn-NTA and alkyn-IDA tags provide excellent tools for generating pronounced PCSs in proteins with paramagnetic lanthanide ions. These tags are attractive because they can be attached to the unnatural amino acid AzF, which can be incorporated at any site using a well-established orthogonal rRNA/aminoacyl-rRNA synthetase system, also in the presence of cysteine residues. Although the metal ion is not completely immobilized relative to the protein, sizeable PCSs are obtained routinely and, in general, these can be fitted acceptably well by a single effective \( \Delta \chi \) tensor. Prior detailed knowledge of the protein structure is not required, as incorporation of an AzF residue into a regular secondary structure element at a site predicted to be solvent exposed stands a good chance of producing a tagged residue that populates a single \( \chi_1 \) rotamer and thus delivers useful PCSs. In situations where the lanthanide complex displays large translational motions, a local \( \Delta \chi \) tensor fit to a selected region of the protein is still likely to provide useful structural data. The ligation product is chemically stable under all biologically relevant conditions. In combination, these properties make the alkyn-NTA and alkyn-IDA tags exceptionally versatile and highly attractive for structural studies. Compared to the established C3 and C4 tags, they are considerably smaller and less highly charged, thus minimizing the impact on protein structure and function. In addition, they allow the production of samples with different lanthanides by straightforward exchange of the metal ions rather than requiring new protein for each lanthanide.

**Experimental Section**

**Tag syntheses:** The tags were prepared according to published procedures. Analytical data were in agreement with those reported previously.

**Protein preparation:** The AzF mutants of human ubiquitin and GB1 were expressed in vivo by double transformation of E. coli BL21(DE3) with pETMCSI 15 and pEVOL vectors. The genes of ubiquitin and GB1 were encoded on the pETMCSI vector, with the sites of AzF incorporation defined by amber stop codons to replace Glu18 or Thr66 in ubiquitin, or ile6, Val21 or Thr44 in GB1. The pEVOL plasmid encodes the p-azido-L-phenylalanyl-tRNA synthetase (AzF-RS) from M. jannaschii 15a,36 Following transformation into E. coli BL21(DE3), the cells were grown at 37 °C in the presence of 100 mg/L ampicillin and 33 mg/L chloramphenicol. 10 mL of an overnight culture was used to inoculate 1 L Luria-Bertani medium supplemented with 0.02% arabinose. After growing to an OD600 of 1.5-2.0, the cultures were changed to 500 mL minimal media for 15N-labelling, containing 0.5 g/L 15N-ammonium chloride, 0.02% arabinose and 1 mM p-azido-L-phenylalanine. Following incubation at 37 °C for another 1-2 h, overexpression was induced with 1 mM isopropyl-\( \beta \)-D-thiogalactopyranoside (IPTG). The cultures were harvested by centrifugation after expression for about 16 h at 25 °C.

Except for Ubi-E18AzF, all protein constructs contained a C-terminal His\(_6\) tag. These were purified by resuspending the cell pellet in buffer A (20 mM HEPES buffer, pH 7.5, 150 mM sodium chloride, 5% glycerol, 20 mM imidazole) followed by lysis using a French press at 12,000 psi. The cell lysates were centrifuged for 1 h at 34,000 g. The supernatant was loaded onto a 5 mL Ni-NTA column (GE Healthcare, USA) and the protein eluted with buffer B (same as buffer A but containing 500 mM imidazole). The fractions were analyzed by 15% SDS-PAGE with Coomassie brilliant blue staining. Fractions containing...
protein were pooled and dialyzed against buffer C (50 mM HEPES, pH 7.5). Finally, the samples were concentrated using an Amicon ultrafiltration centrifugal tube with a molecular weight cutoff of 3 kDa.

The Ubi-E18AzF sample was prepared in the same way, except that the cell pellets were resuspended in buffer D (20 mM MES, pH 6.8) prior to lysis, and the supernatant from the cell lysate was loaded onto an SP Sepharose column (30 mL) equilibrated with buffer D. The column was washed with buffer D to remove unbound proteins and the bound proteins were eluted with a linearly increasing gradient of sodium chloride (0–1 M).

Ligation of proteins with alkyne-NTA and alkyne-IDA tags: The click reactions between tag and AzF were performed in 0.8 mL total reaction volume in the presence of 0.05 mM AzF-protein in 50 mM HEPES buffer (pH 7.5), 0.5 mM alkyne-NTA tag, 0.2 mM copper(II) sulfate, 1 mM BTTAA (BTTAA = 2-[4-[[bis(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino]methyl]-1H-1,2,3-triazol-1-yl]acetic acid)[5], 5 mM sodium ascorbate, 0.1 mM glycerol and 5 mM aminoguanidine. After overnight incubation at room temperature, 5 mM EDTA were added to facilitate the removal of copper. Following 30 min incubation at room temperature, the solution was concentrated and exchanged with NMR buffer (50 mM HEPES buffer, pH 7.0) using an Amicon ultrafiltration centrifugal tube.

NMR measurements: All NMR measurements were performed at 25 °C using a Bruker 600 MHz NMR spectrometer equipped with a TCI cryoprobe. Complexes with lanthanides or Y[41] were formed by titration with the metal trichlorides. PCSs were measured of amide protons as the [25] reference structures chosen for the initial fits of the covalent structures of AzF ligated with the 〈6H〉 tensor. The covalent structures of AzF ligated with the CH triazole and the CH angle of AzF (60°, 120°, and 180°), the dihedral angle of the conjugated bond between the phenyl and triazole groups (0° and 180°), the dihedral angle and the dihedral angle between the triazole and the CH group (any rotation angle), and the CH bond following the CH group (60°, 120°, and 180°). An uncertainty range of ±10° was allowed for all dihedral bond variations. Conformations with steric clashes between protein and tag were eliminated. The Δχ tensors were fitted to each of the remaining conformers using the experimental PCSs and the final Δχ tensor was read from the conformer with the best least-squares fit to the experimental PCS data.[24] Error estimates of the Δχ tensors and metal ion position were obtained by a Monte Carlo protocol that used a random selection of 90% of the experimental PCS data, and by comparison with the metal ion positions obtained by the program Numbat that does not restrain the metal ion positions.[24]

Acknowledgements

pEVOL was a kind gift from Prof. Peter G. Schultz (The Scripps Research Institute, La Jolla, USA). Financial support by the Australian Research Council is gratefully acknowledged.

Keywords: click chemistry • lanthanide-binding tags • NMR spectroscopy • p-azido-L-phenylalanine • pseudocontact shifts
Stable and versatile: Three different lanthanide-binding tags have been explored that can be attached to p-azidophenyl-L-alanine in a click reaction. Loaded with lanthanides, the tags reliably induced pseudocontact shifts in proteins and the data were readily fitted by a single magnetic susceptibility anisotropy tensor. They present broadly applicable tools for protein structure analysis by paramagnetic NMR spectroscopy.

Choy-Theng Loh, Bim Graham, Elwy H. Abdelkader, Kelly L. Tuck, Gottfried Otting

Page No. – Page No.
Generation of Pseudocontact Shifts in Proteins with Lanthanides Using Small "Clickable" Nitrilotriacetic Acid and Iminodiacetic Acid Tags