Design and synthesis of paramagnetic probes for structural biololgy

Liu, W.

Chapter III

The development of new methods for the attachment of caged lanthanoids NMR probes

Manuscript in preparation

Liu, W. -M.; Skinner, S. P.; Timmer, M.; Blok, A.; Filippov, D.; Overhand, M.; Ubbink, M. 'Thioether linkage: A new inert lanthanide-chelating paramagnetic NMR probe'

Abstract

To generate significant paramagnetic effects, site-specific labeling is important. Among all of the possible modification methods, forming disulfide bridges is the most widely used because cysteine residues can be selectively modified in proteins. However, the weak disulfide bridge is easily cleaved by reductants, which restricts its application. Consequently, the development of new methods is important. To increase the stability of the probe-protein linkage, the possibilities of thioether bond formation and click chemistry were investigated. In this chapter, several paramagnetic NMR probes, **CLaNP-6**, **8-12** and **DIBO**, have been designed and synthesized. The experimental results showed that **CLaNP-6** and **CLaNP-8** were highly unstable in aqueous environment, **CLaNP-9** was difficult to purify, and the lanthanoid ions could not be coordinated to **CLaNP-10**. Nevertheless, **CLaNP-9** was still successfully used for tagging two proteins and provided significant PCSs and thus represents a new, more stable double-armed tag. Also, unnatural amino acid tagging was investigated using the artificial amino acid *p*-azido phenylalanine (AzF). AzF was incorporated into a protein and the complementary paramagnetic probes were made. However, the quality of AzF incorporated protein was questionable and the level of tagging was very low.

Introduction

Paramagnetic effects provided by lanthanoid ions, such as pseudocontact shifts (PCSs), residual dipolar couplings (RDCs) and nuclear relaxation enhancements (PREs), have been recognized as powerful tools for the study of biomolecules. To generate a specific paramagnetic effect, the paramagnetic center must be site-specifically and rigidly attached to biomolecules (Chapter 1). In metalloproteins, lanthanoids can be substituted for Ca^{2+} or Mg^{2+} ions in the metal binding sites.[129] However, many proteins do not possess such metal binding sites and in order to overcome this, site-specifically labeling methods for paramagnetic tags have been developed. In recent years, two approaches of site-specific lanthanoid incorporation have been investigated. One of these is the use of genetically engineered lanthanoid binding peptides (LBP) fused to the protein of interest.[159] A drawback of this method is that the engineered LBP region can only be tagged to the N- or C- terminus or into a specific loop region. The other approach is using synthetic Ln-binding probes to introduce the paramagnetic lanthanoids[5] (Chapter 1). So far, most of the synthetic probes are attached to target proteins *via* one or two disulfide bonds. The advantages of this method are that the cysteine residues can be selectively modified in proteins and that they react specifically with thiol-reactive group, avoiding any undesired side-products.[160] However, these bonds are not stable and can be cleaved easily by reductants. As a result, the study of proteins in the presence of a reducing agent such as DTT or TCEP cannot be accomplished by using a paramagnetic tag connected *via* a disulfide bridge.

To increase the stability of the modification, two approaches, thioether bond formation and bioorthogonal chemistry, are employed. Thioether linkers, broadly used for protein modification,[46] are stable in the presence of reducing agents. Several paramagnetic NMR tags were designed for this purpose (Chart 1, **CLaNP-6** and **8-10**). These probes contain functionalities that can react by forming thioether linkages upon reaction with a cysteine residue. The functionalities include an α–ketobromide, a benzylic bromide and a primary bromide. In this chapter, the pitfalls in the synthetic routes towards the reactive **CLaNP-6, 8-10** are described in detail. Nevertheless, **CLaNP-9** was successfully attached to two proteins and significant PCSs were observed. The HSQC spectra of the **CLaNP-9** tagged protein were identical in the presence and absence of DTT or TCEP.

Also, bioorthogonal chemistry is a widely used technique for specific labeling in cellular systems.[48] Among numerous bioorthogonal reactions that have been developed, the introduction of azide groups and their subsequent transformation is the most utilized. Moreover, the azide group can be genetically engineered into proteins at specific sites by using the unnatural amino acid residue *p*-azido phenylalanine (AzF).[161] The azide group of AzF can undergo 1,3-dipolar cycloaddition with a terminal alkyne by using Cu(I) as a catalyst. The reaction product is rigid and only one stereoisomer is generally observed (Chart 2A).[49] The azide group of AzF can also react with cyclooctyne (Chart 2B). In this case, the catalyst Cu(I) is not needed but the cyclolization product has two isomers. Still, a reaction in the absence of the catalyst is advantageous because Cu(I) can denature the protein samples. Thus, AzF represents a potential alternative for Cys in site-specific labeling of proteins with lanthanoid containing probes. We introduced AzF into a test protein, T4 Lysozyme (T4Lys), and synthesized several alkyne containing probes to develop a new approach for the site-specifically tagging of paramagnetic NMR probes. However, the results of labeling were not satisfactory because no paramagnetic effects could be observed.

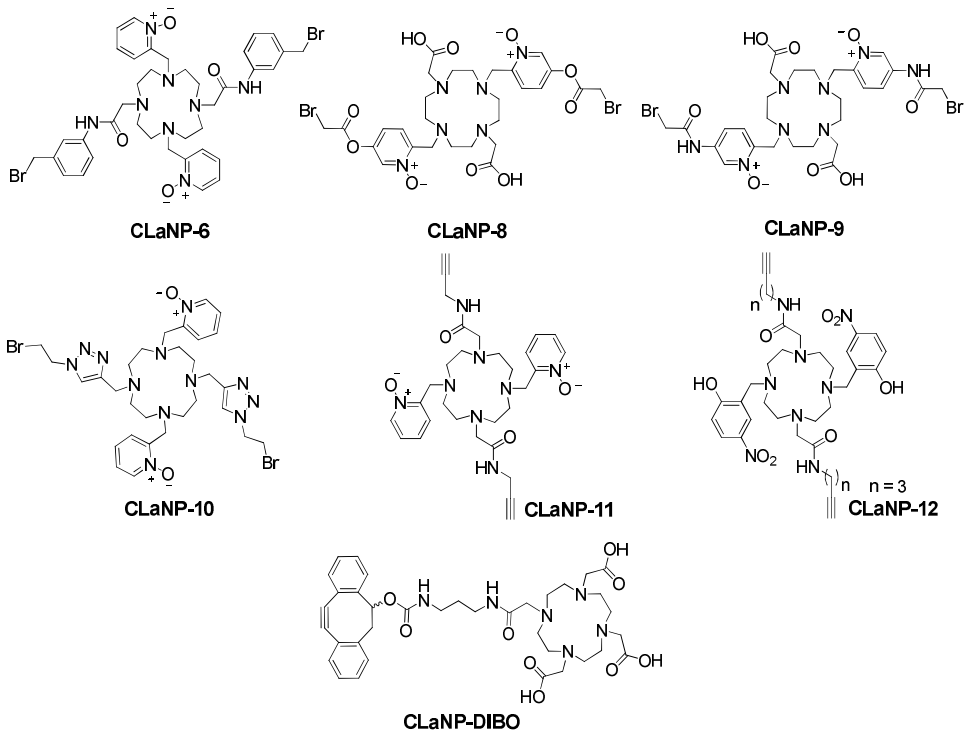


Chart 1. Structures of new CLaNPs, **CLaNP-6, 8-12** and **DIBO.**

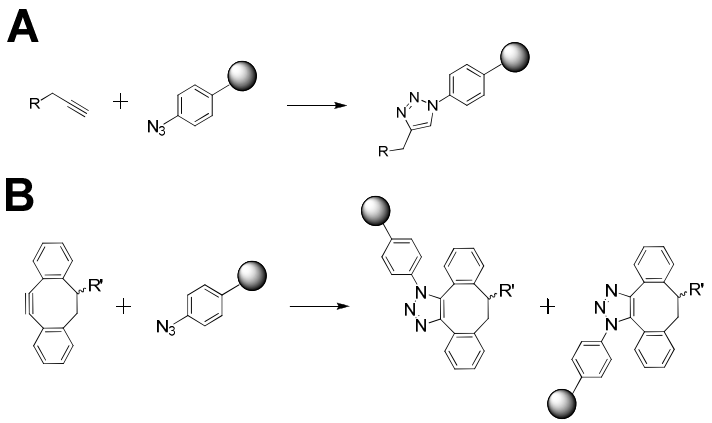


Chart 2. The reaction products of 1,3-dipolar cycloaddition (A) and [3+2] cycloaddition (B).

Results and discussion

Thioether linkage: A new inert lanthanoid-chelating paramagnetic NMR probe

Synthesis of CLaNP-functionalities

Alkyl halides, α-halocarbonyls and electron poor double bonds are typical thiol-reactive functional groups.[46] In order to design an appropriate paramagnetic tag, the labeling ability of acrylic acid, 3-bromo-1-propanol, 3-nitrobenzyl bromide, 4-iodobenzyl bromide, bromoacetic acid, and iodoacetic acid were investigated (Figure 1A). Double cysteine mutants testing proteins (Src SH3 T116C/E117C and T4 Lysozyme N55C/V57C) were first activated by DTT for an hour, and then the reductant was removed by a short gel filtration column. The activated protein was incubated with 10 equiv. of the reactive species at 4 ^{o}C or RT and product formation was monitored by LC/MS (Figure 1). The results showed that under these conditions benzylic bromide is the most reactive group among all of the test species. After 3 h, SH3 T116C/E117C was tagged with 3-nitrobenzyl bromide and 4-iodobenzyl bromide at 4 ^{o}C (Figure 1B and 1C). When the reaction temperature was raised to RT and the reaction time was increased, T4Lys N55C/V57C was also tagged with 2-bromoacetic acid and 2-iodoacetic acid (Figure 1C and 1D). Although α-ketoiodide is more reactive than α-ketobromide, it may also react with methionine and basic residues when cysteine is absent, resulting in non-specific labeling.[162,163] Consequently, benzylic bromide (**CLaNP-6**), α–ketobromide (**CLaNP-8** and **9**) as well as a primary bromide (**CLaNP-10**) were chosen as the target functional groups. Protein expression and purification were kindly performed by Dr. Somireddy V. B.

K. Reddy, Ms. Anneloes Blok and Dr. Monika Timmer (Leiden University, Inst. Chemistry).

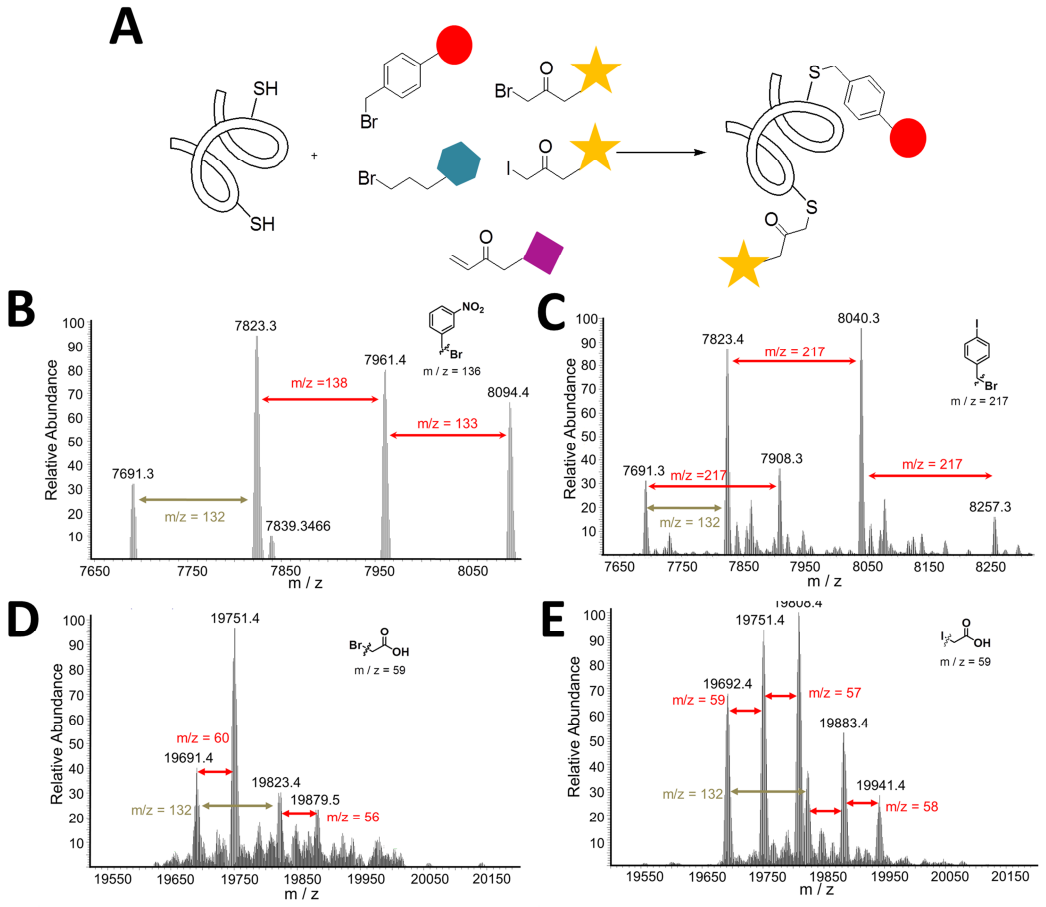
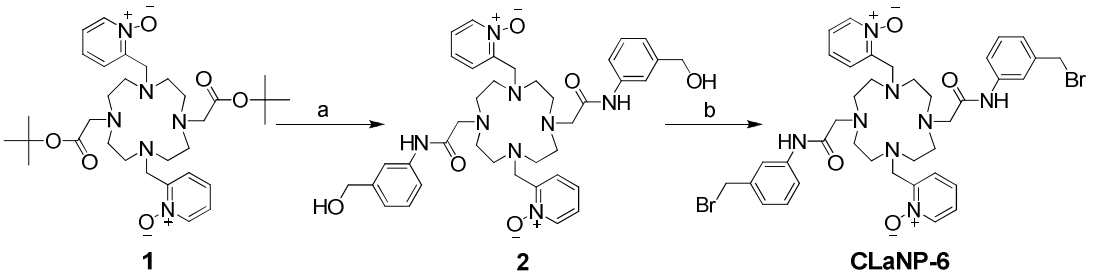


Figure 1. Protein labeling test. (A) In protein labeling experiment, several typical thiol-reactive functional groups were used. The labeling results were checked by LC/MS. For SH3 T116C/E117C, the results showed that the protein could be tagged by 1-(bromomethyl)-3-nitrobenzene (B) and 1-(bromomethyl)-4-iodobenzene (C) at 4 ^{o}C with continuous stirring for 3 h. When the temperature was increased to RT and the reaction time was elongated to 16 h. 2-bromoacetic acid (D) and 2-iodoacetic acid (E) were also attached to another test protein T4Lys N55C/V57C. The mass difference is indicated by a red arrow and the difference between untagged T4Lys with and without the methionine is shown with a dark green arrow.

The synthesis of **CLaNP-6** is depicted in Scheme 1. Removing the protective group of functionalized cyclen **1**[11] and coupling with 3-aminobenzyl alcohol resulted in benzylic alcohol **2** in high yield. The hydroxyl of **2** was converted to the bromide by using PPh_3Br_2 in dry MeCN and product formation was monitored by LC/MS. When the reaction was complete, water was added to quench the reaction. During the reaction **CLaNP-6** was formed, as evidenced by LC/MS, but the compound decomposed after the quenching with water (Figure 2).



Scheme 1. Synthesis of **CLaNP-6**. (a) i) TFA, CH_2Cl_2, RT, 4 h; ii) HATU, 3-aminobenzyl alcohol, TEA, DMF, RT, 16 h; (b) dry MeCN, PPh_3Br_2, 0 ^oC.

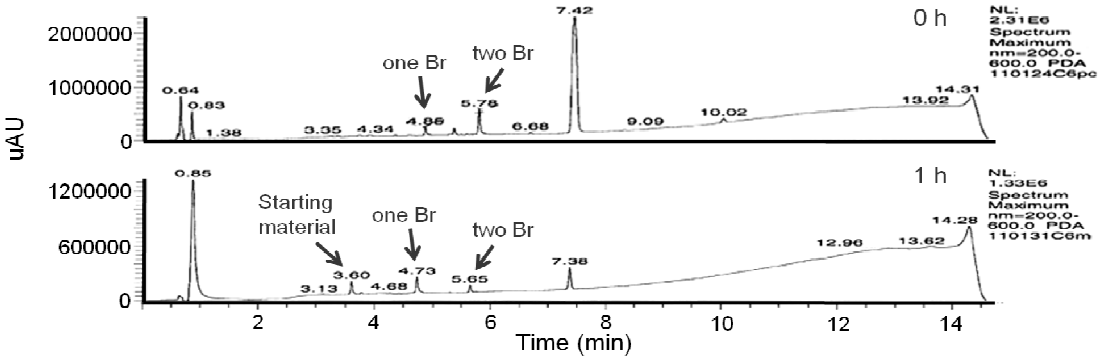
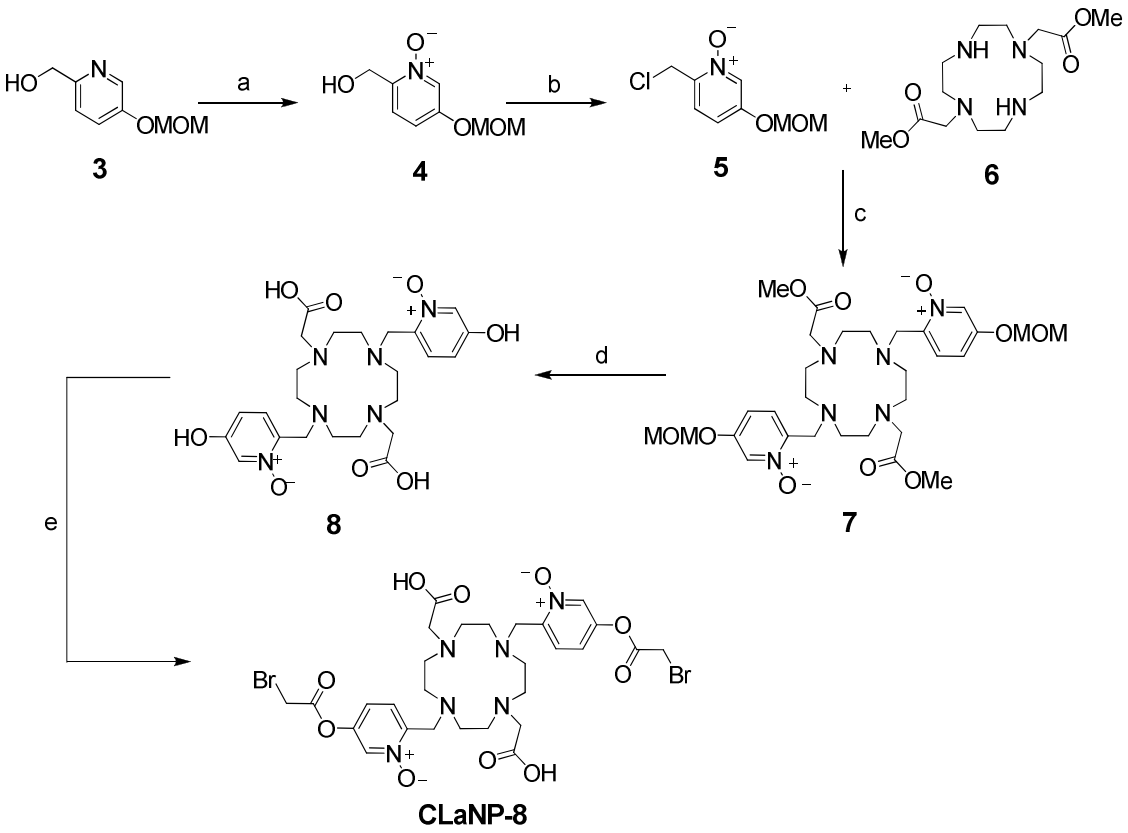


Figure 2. HPLC analysis of the **CLaNP-6**. (A) After quenching by water, a mixture of compounds (**CLaNP-6** with one or two Br) was detected. (B) The same HPLC sample was measured again after 1 h. The results show that **CLaNP-6** was hydrolyzed to the starting material **2**.

The synthetic route of **CLaNP-8** is shown in Scheme 2. The starting material **3**[164] was oxidized with *m*CPBA to produce pyridine-*N*-oxide **4** with a high yield. Substitution of the hydroxyl of **4** using MsCl and LiCl in dry CH_2Cl_2 yielded compound **5**. Functionalized cyclen **6**[42] was treated with **5** and K_2CO_3 at 50 ^oC for 16

h to give compound **7**. Compound **8** was obtained by full deprotection of **7**. Reaction of **8** using bromoacetyl bromide gave the desired **CLaNP-8**. However, also this compound was not to be stable under aqueous conditions (Figure 3).



Scheme 2. Synthesis of **CLaNP-8**. (a) *m*CPBA, $CHCl_3$, RT, 16 h; (b) i) MsCl, dry CH_2Cl_2, TEA, 2 h, 0 ^oC ii) 10 equiv. LiCl, RT, 16 h; (c) K_2CO_3, MeCN, 50 ^oC, 16 h; (d) i) NaOH, 1,4-dioxane, RT, 4 h; ii) TFA, CH_2Cl_2, RT, 4 h; (e) dry MeCN, K_2CO_3, RT, bromoacetyl bromide.

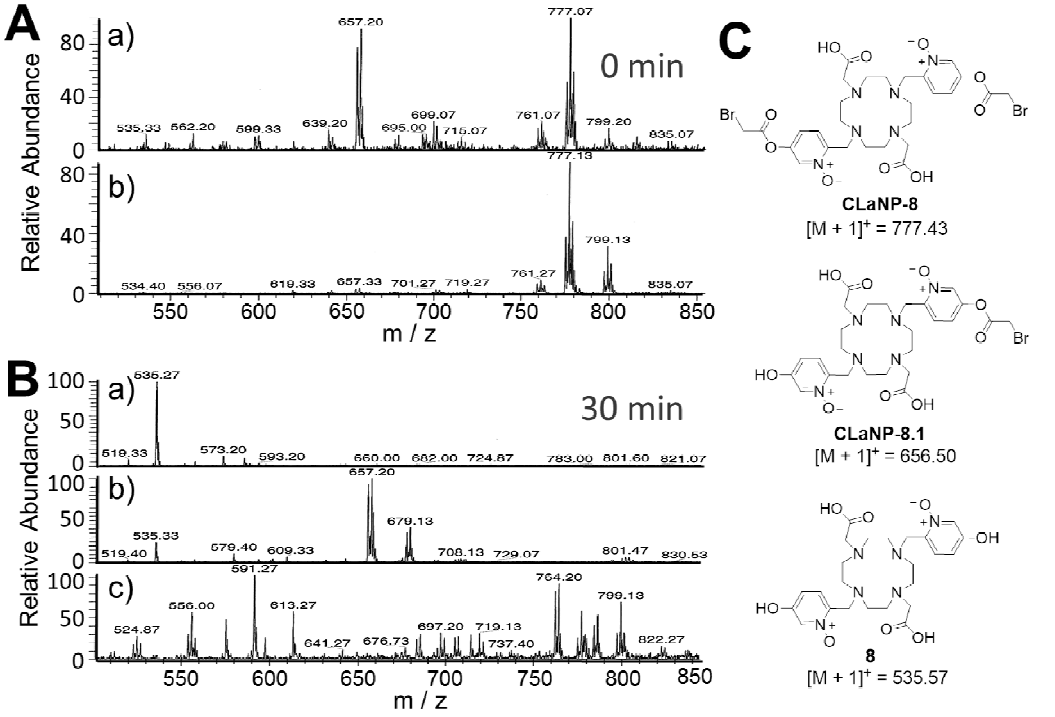
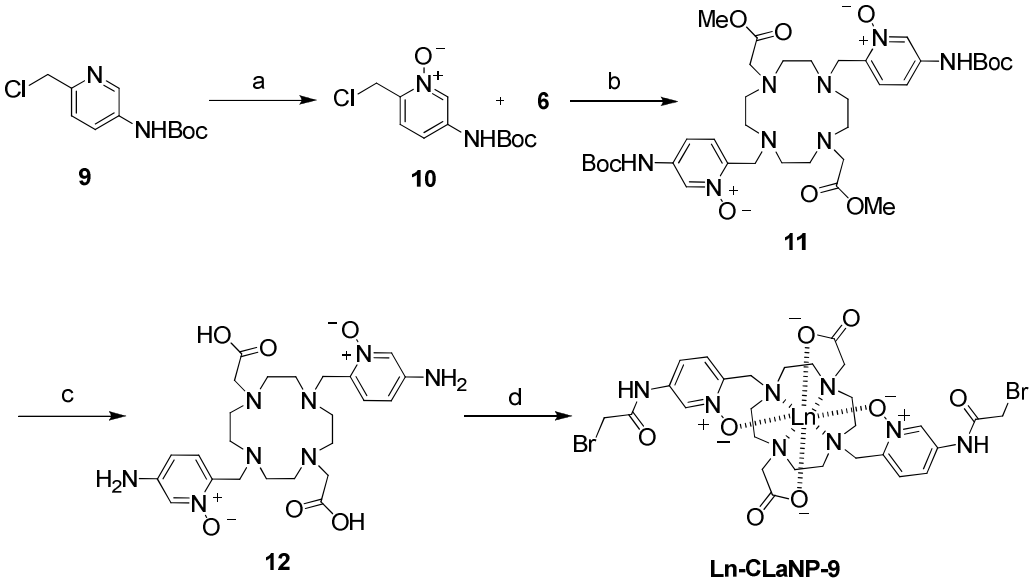


Figure 3. The reaction progress of **CLaNP-8** was monitored by LC/MS. (A) The reaction was quenched by water and the sample was immediately analyzed by mass spectrometry. The result showed that **CLaNP-8.1** (a) and **CLaNP-8** (b) were present. (B) After 30 min, the same sample was analyzed again and the results showed that **CLaNP-8** was hydrolyzed to compound **8** (a) and **CLaNP-8.1** (b) by water. Also, the final product was difficult to observe (c). The expected weights of **CLaNP-8**, **CLaNP-8.1** and **8** are presented in (C).

According to the results of LC/MS, the hydrolysis occurred at the active ester bond. Therefore, the ester was replaced with an amide to increase the stability of desired product, **CLaNP-9**, and the synthesis is shown in Scheme 3. The compound **9**[165,166] was synthesized *via* following a previously published procedure and then oxidized with *m*CPBA, yielding **10**. Functionalized cyclen **6**[42] was treated with **10** in MeCN at 50 ^{o}C and was converted into **11**. Removal of the protective groups in **11** followed by HPLC-purification yielded compound **12**. Subsequently, **12** was dissolved in dry MeCN, treated with bromoacetyl bromide and purified by HPLC to give **CLaNP-9** in very low yield (3 %), which was caused by the reactive carboxylic acid and *N*-oxide. Moreover, the final product was hydrolyzed after freeze drying.

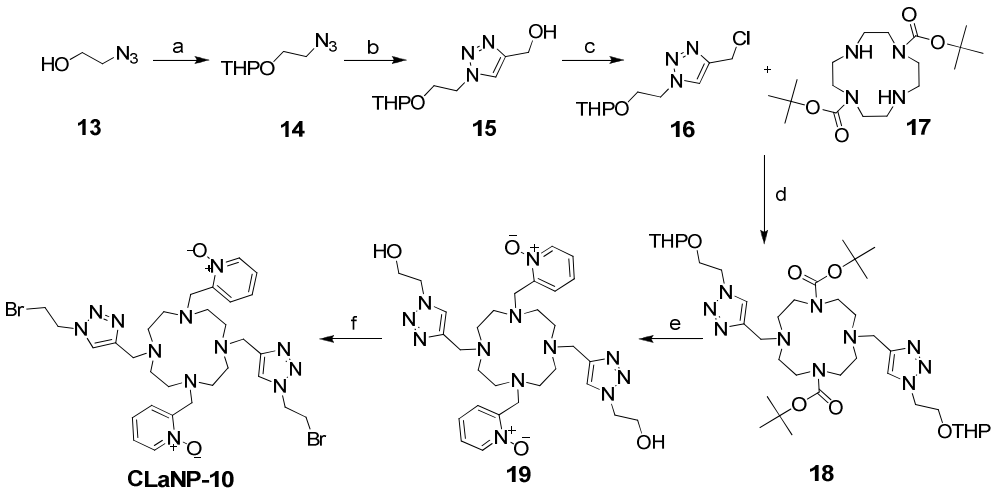
Hence, compound **12** was coordinated to $Ln(OAc)_3$ to "block" the reactivity of the acids and *N*-oxides, and was subsequently treated with bromoacetyl bromide. Indeed, the reaction became cleaner and complete conversion to **Ln-CLaNP-9** was observed. The solution mixture was purified over a short Prevail™ C18 column (Alltech® Etract-Clean™) to remove the impurities and the product was directly used for protein labeling (see probe attachment section).



Scheme 3. Synthesis of **CLaNP-9**. (a) *m*CPBA, $CHCl_3$, RT, 4 h; (b) K_2CO_3, MeCN, 50 °C, 2 days; (c) i) TFA, CH_2Cl_2, RT, 4 h; ii) NaOH, 1,4-dioxane, RT, 4 h; (d) i) $Ln(OAc)_3$, DMF, RT; ii) $NaHCO_3$, dry MeCN, bromoacetyl bromide, RT, 30 min.

Compared to benzylic bromides and α-ketobromides, primary bromides are stable and can easily be purified. However, these compounds did not react with Cys residues during the labeling test. Nevertheless, the synthesis of **CLaNP-10** was undertaken and is shown in Scheme 4. The hydroxyl group of known compound **13**[167] was protected using a THP group and clicked with propargyl alcohol, yielding **15**. The alcohol of **15** was activated with MsCl and converted to chloride **16**, after which **16** was treated with protected cyclen **17**.[42] The reaction was heated under reflux for 16 h and compound **18** was obtained. The protecting groups of **18** were removed by TFA and the crude compound was reacted with 2-(chloromethyl) pyridine-*N*-oxide to give **19**. Compound **19** was converted into **CLaNP-10** by using PPh_3Br_2 and purified by HPLC. However, **CLaNP-10** did not readily chelate

lanthanoid ions, which might be caused by the steric hindrance by the four aromatic groups. When the temperature was raised, to speed-up the complex formation process, the compound started to decompose.



Scheme 4. Synthesis of **CLaNP-10**. (a) DHP, PTTs, CH_2Cl_2, RT, 16 h; (b) propargyl alcohol, ascorbate acid, $CuSO_4$, THF/H_2O solvent mixture, RT, 16 h; (c) i) MsCl, TEA, CH_2Cl_2, 0 ^oC, 1 h; ii) 10 equniv. LiCl, acetone, RT, 16 h; (d) K_2CO_3, MeCN, reflux, 16 h; (e) i) TFA, CH_2Cl_2, RT, 4 h; ii) 2-(chloromethyl) pyridine-*N*-oxide, K_2CO_3, MeCN, reflux, 16 h; (f) dry MeCN, PPh_3Br_2, at 0 ^oC to RT, 16 h.

Probe attachment

^{15}N-enriched T4Lys N55C/V57C was activated with DTT, washed, and directly incubated with 10 equiv. of Ln-**CLaNP-9** at RT for 6 hours, after which the tagged protein was purified using an ion-exchange column. The [^{15}N-^{1}H]-HSQC spectra of tagged T4Lys were recorded in the presence of three lanthanoid ions (Lu^{3+}, Yb^{3+} and Tm^{3+}). The paramagnetic lanthanoid ions (Yb^{3+} and Tm^{3+}) caused changes in the resonance frequencies of the observed nuclei, which were identified as PCSs (Figure 4). The spectra of tagged T4Lys were identical in the presence and absence of DTT, indicating that the probe was stable under reductive conditions. Moreover, the spectra were the same after keeping the protein samples at 4 ^oC for 2 weeks. The copper protein Paz mutant E51C/E54C was the second tested protein. Instead of DTT, TCEP was used to activate the protein because the copper ion is removed from the protein by a high concentration of DTT. The [^{15}N-^{1}H]-HSQC spectra of tagged

Cu Paz also exhibited single sets of PCSs. Both experiments indicated that the probe was attached rigidly to the protein and existed in one dominant conformation. To demonstrate the efficacy of this new probe, the Ln-**CLaNP-9** tagged Cu(II)-Paz was reduced to Cu(I)-Paz by using TCEP. As expected, the tagged probe was stable under reducing conditions and significant paramagnetic effects were observed (Figure 5). The $\Delta\chi$-tensors values for Tm- and Yb-**CLaNP-9**, and **CLaNP-5** for comparison, are reported in Table 1 and the back-calculated PCSs are plotted versus the observed ones in Figures 6. The yield of labeling after purification was 50 ~ 80%, which estimated from the intensity of diamagnetic peaks in the spectra of samples with paramagnetic tags. The LC/MS results show that the diamagnetic peaks come from unlabeled protein and not one-arm attached tag. A simple explanation is that the probe concentration was insufficient to yield complete labeling. Theoretically, the probe was in 10-fold excess over the protein, but the accuracy of this number is questionable, due to the purification step. In the future, 20 or 30 excess fold of probe and longer incubation times could be tried to increase the labeling yield.

Table 1. PCSs-based $\Delta\chi$-tensors of CLaNP-5[11] and CLaNP-9[a,b]

Protein	probe	Ln	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	Restraints	Q
Paz Cu(I) E51C/E54C	CLaNP-5	Yb	9.4 ± 0.2	1.9 ± 0.4	93	0.04
		Tm	55.5 ± 0.8	10 ± 1	89	0.03
Paz Cu(II) E51C/E54C	CLaNP-9	Yb	7.5± 0.4[c]	2.3 ± 0.4[c]	66	0.04
		Tm	39.5± 1.1[c]	10.6 ± 1.5[c]	41	0.02
T4Lys N55/V57C	CLaNP-9	Yb	6.3 ± 0.8[c]	2.2 ± 0.8[c]	88	0.04
		Tm	32.5 ± 2.3	15.6 ± 1.5	62	0.03

[a] The unit of axial and rhombic components ($\Delta\chi_{ax}$ and $\Delta\chi_{rh}$) is 10^{-32} m^3.

[b] The program Numbat[168] was used to calculate the $\Delta\chi$-tensors.

[c] The error is calculated by randomly removing 10% of the data and repeating the $\Delta\chi$-tensor fit 100 times.

The $\Delta\chi_{ax}$ of Tm^{3+} and Yb^{3+} of **CLaNP-9** are smaller than those of **CLaNP-5**. For **CLaNP-9**, the functionalized pyridine-*N*-oxide was the site for attachment to the protein, which may cause a larger mobility of the probe relative to the protein. In the case of T4Lys, the mutant site was at the end of the elongated shape of the protein

and therefore only part of amide proton resonances experienced a significant PCS, which can explain the larger error in $\Delta\chi_{ax}$ as compared to Paz.

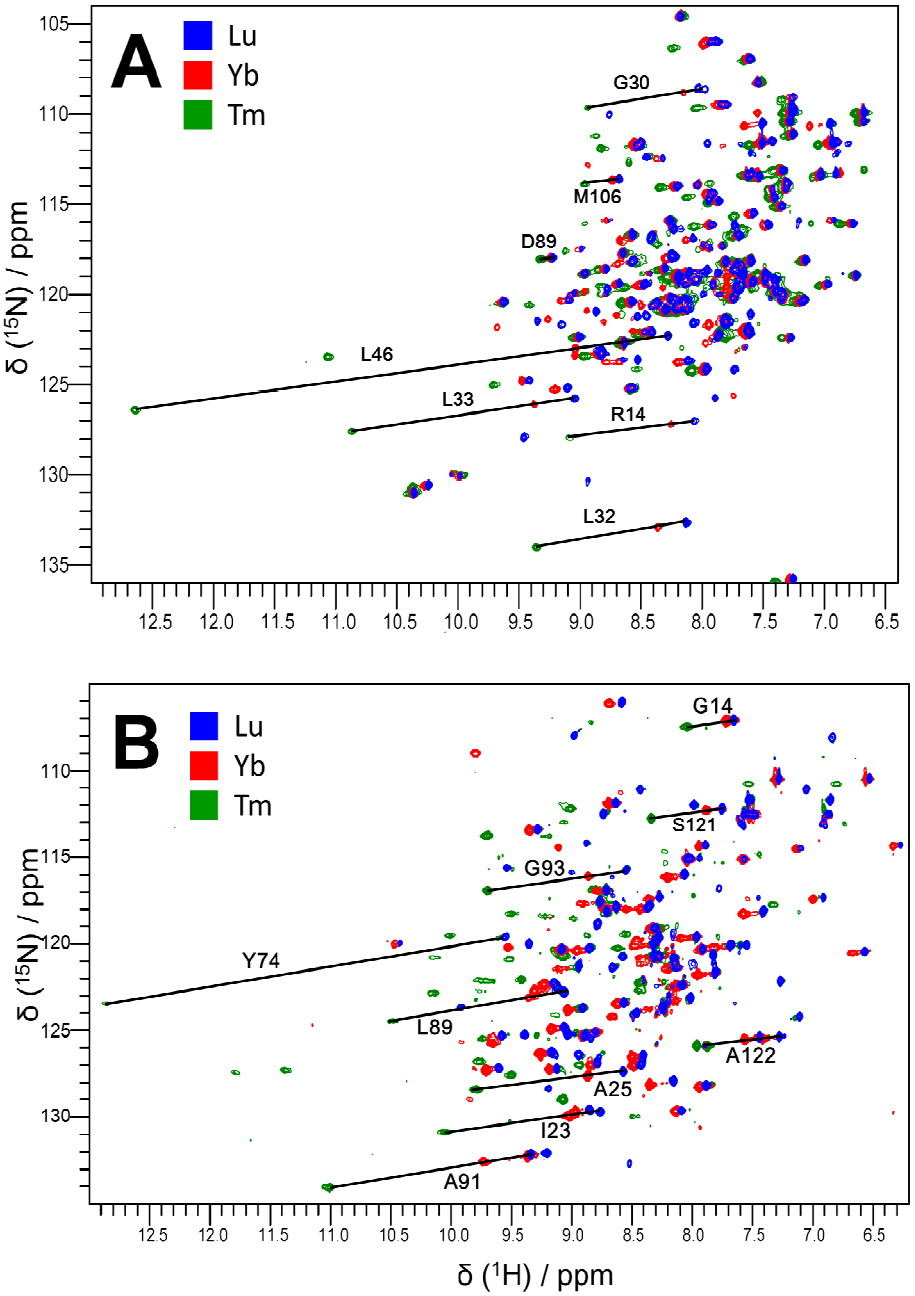


Figure 4. Overlay of [^{15}N, ^{1}H]-HSQC spectra of T4Lys N55C/V57C (A) and Paz-Cu(II) E51C/E54C(B) with **CLaNP-9**-Lu (blue), Yb (red) and Tm (green) attached. Several PCSs are indicated with solid lines.

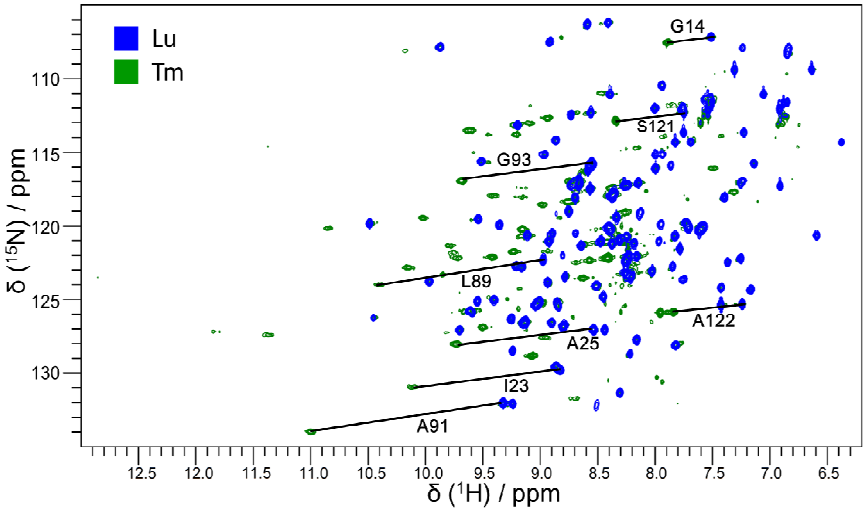


Figure 5. Overlay of [^{15}N, ^{1}H]-HSQC spectra of Paz-Cu(I) E51C/E54C (A) with **CLaNP-9**-Lu (blue) and Tm (green) attached. Several PCSs are indicated with solid lines.

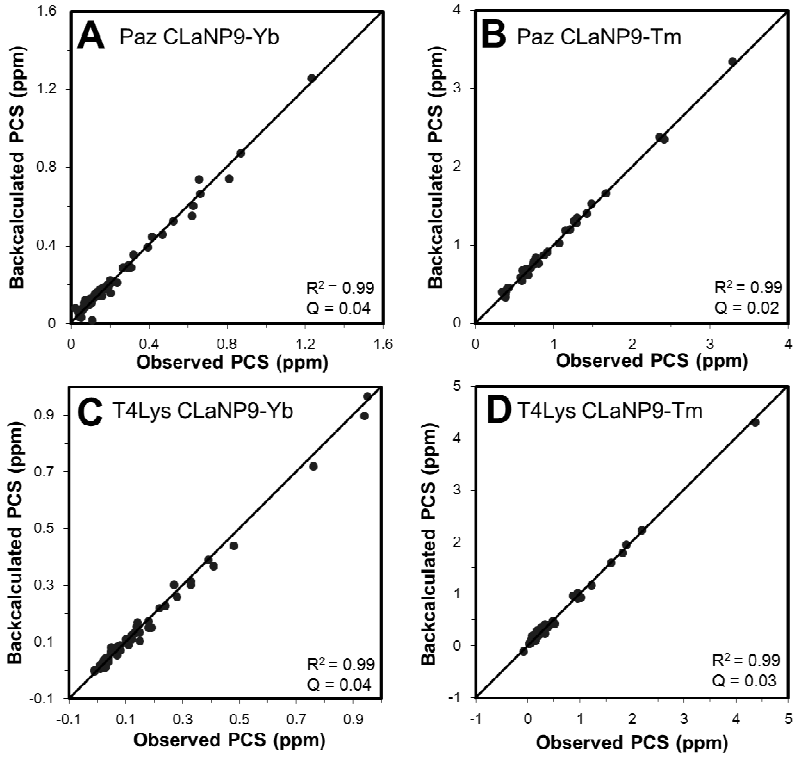
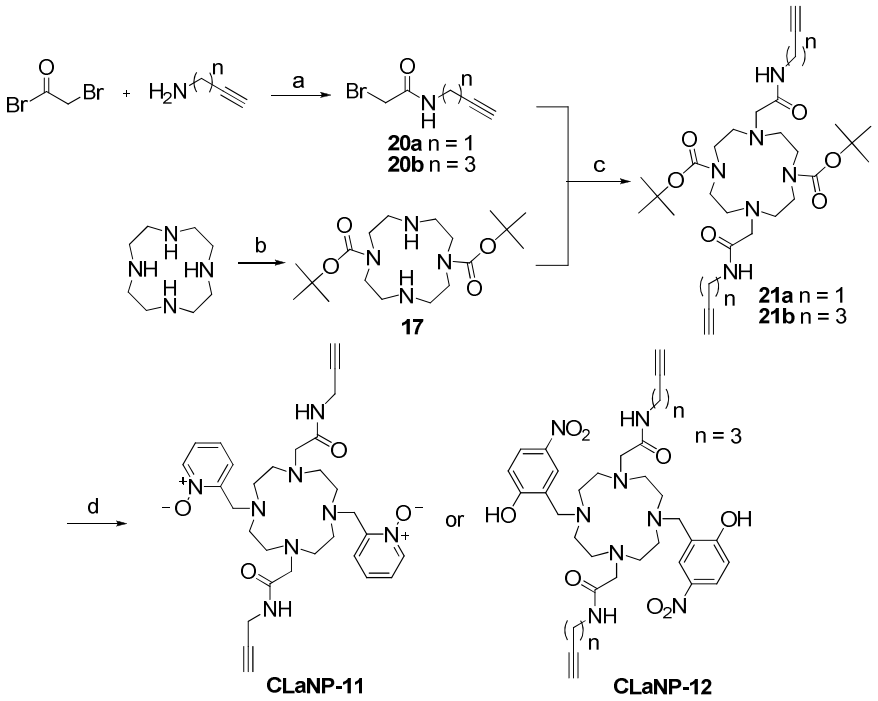


Figure 6. Experimentally observed amide proton PCSs of Paz Cu(II) E51C/E54C (A, B) and T4Lys N55C/V57C (C, D) Ln-**CLaNP-9** plotted against the back-calculated PCSs. The solid lines represent the perfect correlation (y=x).

Bioorthogonal reactions: A potential method for site-specifically attaching lanthanoid probes

Synthesis of CLaNP-click

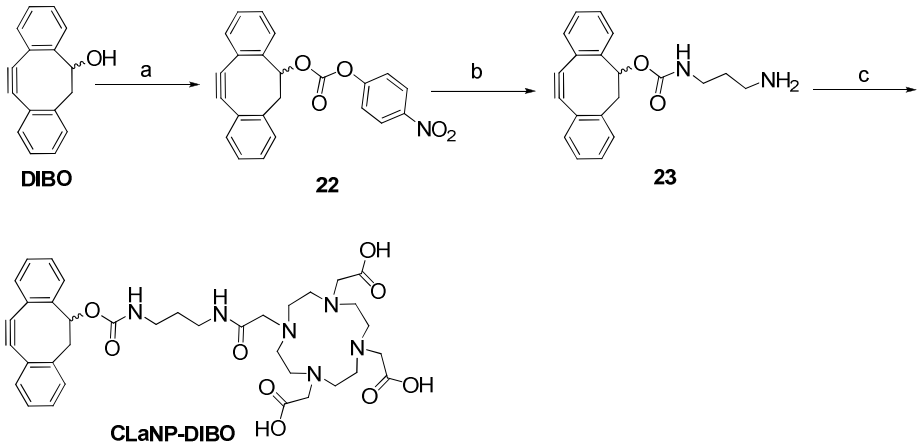
The structures of the CLaNP-click series are based on two successful probes, **CLaNP-5** and **CLaNP-7** (Chapter 1).[11,42] In the design of the bioorthogonal probes, the two reactive groups on these CLaNPs are replaced with terminal alkynes to allow a bioorthogonal reaction. **CLaNP-11** and **CLaNP-12** were synthesized as depicted in Scheme 5. Compound **20a**, **20b** and **17** were prepared using previously published methods.[42,169] To obtain **21a** and **21b**, **17** was functionalized with alkyl functional arms (**20a** or **20b**). Removal of the Boc protective group in **21a** and **21b**, alkylation with 2-(chloromethyl)pyridine-*N*-oxide or 2-(bromomethyl)-4-nitrophenol, and purification by HPLC yielded the designed products.



Scheme 5. Synthesis of CLaNP-11 & 12. (a) $NaHCO_3$, CH_2Cl_2, RT, 16 h; (b) BocOSu, dry $CHCl_3$, RT, 1 day; (c) **20a** or **20b** (2.5 equiv.), MeCN, K_2CO_3, 0 ^oC to RT, 16 h.; (d) i) TFA, CH_2Cl_2, RT, 4 h; ii) 2-(chloromethyl)pyridine-*N*-oxide or 2-(bromomethyl)-4-nitrophenol, MeCN, 80 ^oC, 16 h.

For the ring strain-promoted [3+2] cycloaddition, **CLaNP-DIBO** was synthesized (Scheme 6). Compound **DIBO**[170] was activated by *p*-nitrophenyl

chloroformate, then reacted with 1,3-diaminopropane yielding **23**.[171,172] The final product was synthesized through EDC coupling of the free amine of **23** with a single carboxylic acid group of a DOTA ring yielding the desired compound.[173]



Scheme 6. Synthesis of CLaNP-DIBO. (a) *p*-nitrophenyl chloroformate, CH_2Cl_2, pyridine, RT, 16 h; (b) 1,3-diaminopropane, dry CH_2Cl_2, TEA, RT, 1 h; (c) DOTA, DIPEA, EDC, water/MeOH, pH = 4.5-8.0, RT, 1 h, yield 20% (over 3 steps).

Protein labeling study

Copper-catalyzed azide-alkyne 1,3-diploar cycloaddition (CuAAC)

Site-specific introduction of unnatural amino acids (uaa) into a protein system *in vivo* was first reported by Prof. P. Schultz and co-workers.[174] This method uses the stop codon, UAG, as the codon of uaa. In *Eschirichia coli* the genes of complementary tRNA, the gene for a dedicated AzF-tRNA synthetase and the target gene with the UAG codon(s) introduced at the site(s) for AzF incorporation were expressed together. To demonstrate the utility of this new method, two AzF mutants of T4Lys were designed, N68AzF/D72AzF and N55AzF/V57AzF. However, the protein yield was low due to the presence of two UAG codons close together in the sequence. Therefore, single mutant proteins were also produced and used as test proteins. Protein expression and purification were performed by Mrs. Anneloes Blok (Leiden University, Inst. Chemistry).

In order to reduce the catalyst Cu(II) to Cu(I) and to stabilize the Cu(I), a reductor and a metal stabilizer (TBTA) are employed, respectively. In the labeling experiment, the probes were added in five-fold excess and the reaction was quenched

by EDTA after 30 min. This procedure was performed because it was found that Cu(I)-induced protein precipitation increases with time. Firstly, the amounts of $CuSO_4$ and TBTA were optimized. SDS-PA gel was used to detect the tagged protein. A weak second band indicative of the tagged protein was present at the highest concentration of $CuSO_4$ (Figure 7A). Tris(2-carboxyethyl)phosphine (TCEP), sodium ascorbate (NaAsc) and dithiothreitol (DTT) are commonly used to reduce Cu(II). TCEP gave the best results in our system (Figure 7B). The commercially available fluorescent probe, Tamra, was also used as a control to confirm that the employed reaction conditions allowed tagging of the protein. In the Tamra test, a second protein band was present in SDS-PA gels and showed intensive fluorescence, even though only a small portion of the protein was labeled (Figure 7C). The labeling yield increased with a higher concentration of TCEP. Using these conditions, T4Lys D72AzF was incubated with the Yb-**CLaNP-11** and excess probe was removed using a HiTrapSP column.

[^{15}N, ^{1}H]-HSQC spectra taken of the T4Lys D72AzF before and after the reaction are shown in Figure 8. Unexpectedly, the HSQC spectra before and after reaction are identical except for few residues. These residues are within 10 Å of the C-terminus, with has an additional His tag. The His tag could chelated Cu(II) resulting PREs. No PCSs were observed, indicating that the labeling reaction was unsuccessful.

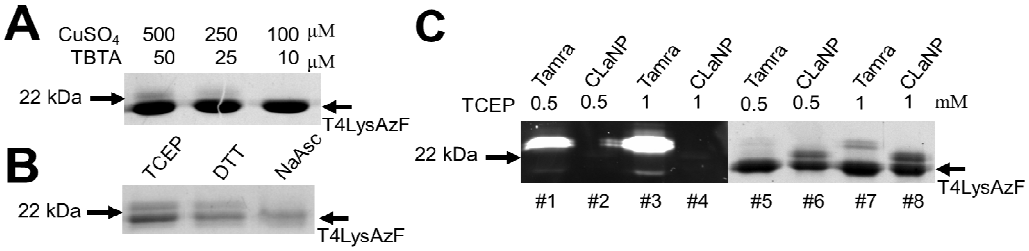


Figure 7. Optimization of the click reaction. (A) Influence of the amount of $CuSO_4$ used during the click reaction. (B) The effect of different reductors. (C) Influence of the amount of TCEP on the click labeling of Tamra and CLaNP to T4LysAzF. The reaction mixture contained $CuSO_4$ (500 μM), TBTA (50 μM), CLaNP or Tamra (250 μM), and T4Lys D72AzF (50 μM). Lanes 1-4 show the fluorescence and 5-8 the Coomassie stain of the SDS-PA gel.

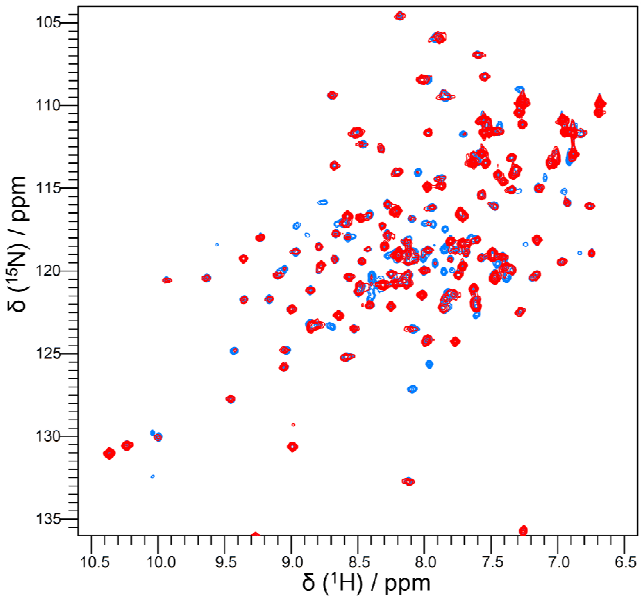


Figure 8. Overlay of [^{15}N, ^{1}H]-HSQC spectra of T4Lys D72AzF before (blue) and after the reaction with Yb-CLaNP-11 (red). No PCS are observed.

The low level of labeling may be caused by the short alkyl arms in **CLaNP-11**, and therefore, **CLaNP-12** was used for testing. **CLaNP-12** has an additional absorption band at 390 nm, which simplifies the detection of tagged protein compared with **CLaNP-11**. However, the poorly aqueous solubility of **CLaNP-12** was a serious problem and caused a low labeling yield. It was not pursued further.

Copper free click chemistry

For protein labeling, the yield of the tagging reaction could also be site-dependent. In T4Lys D72AzF a lysine and an arginine are close to the AzF, which is on an α-helix. The long side-chains of Lys and Arg could increase the steric hindrance and negatively influence the labeling efficiency. Moreover, the charges of those amino acids could repel the positive **CLaN-11** complexes. For these reasons, another AzF mutant, T4Lys N55AzF, was selected as a model protein for **CLaNP-DIBO** labeling. To test the reactivity of **CLaNP-DIBO**, it was reacted with free AzF and product formation was monitored by LC/MS. The sample was incubated at 4 ^{o}C with continuous stirring, and **CLaNP-DIBO** was completely converted into the cycloaddition products and both isomers (Chart 2B) were present in a 1:1 ratio. Upon

linking **CLaNP-DIBO** to a protein, two products are expected and, therefore, two sets of PCS, which is not desired. However, **CLaNP-DIBO** can be loaded with Gd^{3+}, to obtain a PRE probe. The presence of diastereoisomers will not complicate the spectrum in that case. Following the same protocol, T4Lys N55AzF was incubated with the Ln-**CLaNP-DIBO** complexes and the labeling experiment was checked by SDS-PA gel (Figure 9). The gel suggests that a fraction of the protein is labeled with the probe. The excess probe was removed and the protein sample was concentrated and analyzed without further purification. The HSQC spectra of T4Lys N55AzF reacted with Lu^{3+} and Gd^{3+} complexes are shown in Figure 10. The peak intensity ratios of Gd^{3+}/Lu^{3+} samples is close to 1 (data not shown), indicating that the Gd^{3+}-containing probe was not attached onto the protein in sufficient quantity to cause substantial PRE.

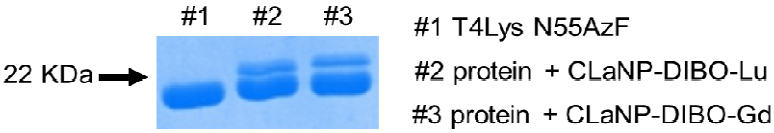


Figure 9. Detail of the SDS-PA gel showing the results of the CLaNP-DIBO labeling experiment.

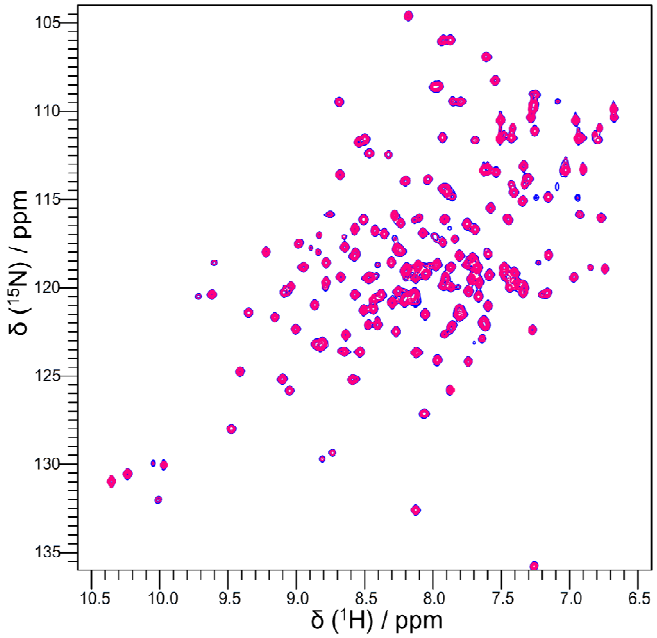


Figure 10. Overlay of [^{15}N, ^{1}H]-HSQC spectra of T4Lys N55AzF reacted with Lu-CLaNP-DIBO (blue) and Gd-CLaNP-DIBO (red).

The application of copper-catalysed 1,3-dipolar azide-alkyne cycloaddition (CuAAC) to obtain a modified protein amenable to NMR spectroscopy is a challenge, as compared to fluorescent studies, because high tagging yields are required. In Figure 7, the SDS-PA gel showed an intense fluorescent band indicating the tagged protein, yet, only a small fraction of the protein was actually labeled. To deduce the reason for the low labeling yield, the quality of T4LysAzF protein was assessed by mass spectroscopy. The MS results demonstrate that incorporation of the unnatural amino acid is 100%, but the results always had a mass of 28 Da less than the theoretical value indicating the loss of two nitrogens from the azide group. The degradation of AzF might be caused by the chemical reactivity and photoinstability of the aryl-azide group, which could have decomposed either during the MS analysis or the protein synthesis and purification.[175] Hence, the quality of protein samples could not be accurately assessed. The NMR spectra showed that there were no paramagnetic effects, even though the SDS-PA gel showed that a small fraction of the protein was modified. A similar result was observed in the **CLaNP-DIBO** labeling experiments.

Recently, site-specifically labeling *via* a UAA was demonstrated by a similar approach.[50] A terminal alkyne functionalized DOTA-base probe was attached to a protein using CuAAc. These researchers faced similar problems to obtain high reaction yields. For example, the protein precipitated during the reaction and a His tag poisoned the Cu(I) yielding lower labeling efficiency.[176] Instead of TBTA, a new Cu(I)-stabilizing ligand, 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), was employed. The new stabilizing ligand performs significantly better and the solubility is higher than commercially available TBTA.[177] Instead of TCEP, NaAsc was employed to reduce the Cu(II). Glycerol and aminoguanidine were also added into the reaction mixture because they prevent protein aggregation caused by the oxidized NaAsc.[176,178] Additionally, the reaction was performed in a glovebox under N_2 with continuous stirring for 16 h at RT. Although the alkyne functionalized probe was successfully attached to protein, the complicated protocol limits the application of this method to very stable proteins. Moreover, the application of CuAAC appears to require optimized reaction conditions for each protein.[179] Consequently, a general protocol is difficult to develop.

Conclusion

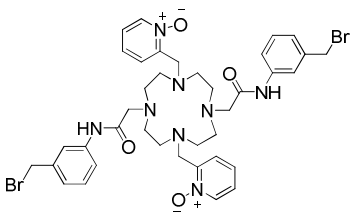
A new thioether paramagnetic NMR probe, **CLaNP-9**, has been reported. It provides stable tagged products under reducing conditions creating the possibility of studying reductive proteins and produces large single PCSs. The charge of **CLaNP-9** complexes is the same as that of **CLaNP-7** (Chapter 2), but is far more soluble under aqueous conditions. The $\Delta\chi$-tensors for Yb^{3+} and Tm^{3+} were also defined using two different model proteins. It is therefore an attractive probe for paramagnetic NMR.

AzF containing T4Lys was successfully produced, however, the reactivity of protein is questionable in the light of the results of MS. Azide reactive probes have been produced and ligated with AzF functionalized protein. However, the attachment experiments showed that it is very difficult and impractical to obtain yields sufficient for NMR spectroscopy.

Experimental Section

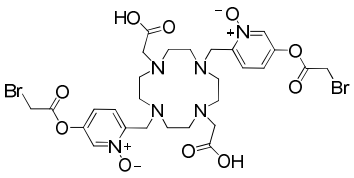
General Procedures:

Compound **1**[11], **3**[164], **6**[42], **9**[165,166] and **13**[167], **17**[42], **20a**, **22b**[169], and **DIBO**[170] were prepared according to methods published in the literature. All other chemicals were used as purchased without further purification. TLC-analysis was conducted on DC-alufolien (Merck, Kieselgel60, F254) with detection by UV-absorption (254 nm). Flash chromatography was performed on Screening Devices silica gel 60 (0.04-0.063 mm). A Biocad Vision HPLC (PerSeptive Biosystems, inc.) and an Äkta Basic FPLC (GE Healthcare Inc.) were used for purification. Semipreparative and preparative reversed phase C18 columns were obtained from Phenomenex (Torrance, CA). Superdex 75, CM sepharose and HiTrap SP columns were obtained from GE Healthcare. NMR spectra were recorded on a Bruker AV-400 (400/100 MHz) and Bruker Avance-III 600 (600/150 MHz) spectrometer equipped with a TCI-Z-GRAD cryoprobe. A LCQ LCMS system and a Finnigan LTQ Orbitrap system were used for HRMS and protein conjugation analysis. FTIR was performed on a Perkin-Elmer (Shelton, CT) Paragon 1000 FTIR spectrometer. Melting points were obtained using a SMP3 scientific melting apparatus (Stuart, Bibby Sterlin Ltd.)



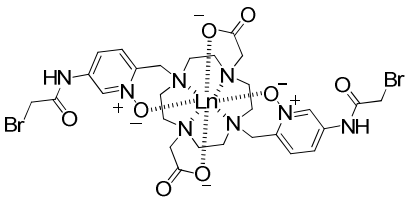
Towards CLaNP-6

The compound **1**[11] (1 g, 1.6 mmol) was dissolved in a mixture of TFA/CH_2Cl_2 (3/1, v/v, 8 mL) . After 4 h stirring at RT, the reaction mixture was concentrated by coevaporating with toluene. The crude residue obtained was redissolved in 10 mL DMF and triethylamine was added until the pH was neutral. Subsequently, HATU (1.3 g, 3.4 mmol) and 3-aminobenzyl alcohol (0.4 g, 3.4 mmol) were added and stirring was continued for 16 h at RT. Concentration of the reaction mixture, redissolving the residue in mini-Q water, filtering out the solids and concentration of the aqueous layer resulted in crude compound **2**, after which crude **2** (90 mg, 0.13 mmol) was dissolved in 0.5 mL dry MeCN and added into a 2.5 mL dry MeCN solution containing PPh_3Br_2 (531 mg, 1.3 mmol) under argon. The reaction mixture was stirred at 0 ^oC for 1 h and was quenched with 0.5 mL demi water. During the reaction, the product formation was monitored by LC/MS. The result showed that the **CLaNP-6** decomposed after quenching with water.



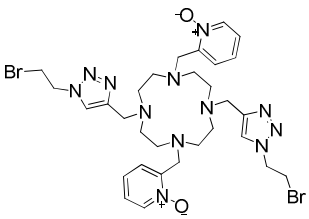
Towards CLaNP-8

Compound **5** (398 mg, 1.92 mmol), **6** (276 mg, 0.87 mmol) and K_2CO_3 (361 mg, 2.61 mmol) were dissolved in 6 mL MeCN, heated to 50 ^oC and stirred for 16 h. Filtering out the excess K_2CO_3 and concentration of the reaction mixture resulted in crude **7**. Redissolving crude **7** in a mixture of 1,4-dioxane (final concentration is 15 mM) and NaOH (final concentration 0.4 M) and stirring for 3 h at RT yielded compound **7**. Acid-resin was added until the pH ~ 7, and this was then removed by filtration. The filtrate was concentrated *in vacuo* and redissolved in a TFA (6 mL) and CH_2Cl_2 (2 mL) solution mixture. Stirring for 4 h, the reaction mixture was coevaporated with toluene twice. The crude compound **8** (100 mg, 0.19 mmol) obtained and K_2CO_3 (65.5 mg, 0.47 mmol) were dissolved in 1 mL dry MeCN. Bromoacetyl bromide (94.8 mg, 0.47 mmol) was added into the reaction mixture with continuous stirring at RT. Product formation was monitored by LC/MS. When all of the starting material was converted to the desired **CLaNP-8**, the reaction was quenched by adding water, to hydrolyze the excess-reagent(s), which resulted in degradation of the target compound.



Ln-CLaNP-9

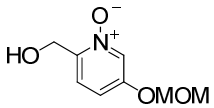
To a solution of **12** (100 mg, 0.187 mmol) in 100 μL DMF 1.1 equiv. $Lu(OAc)_3$ was added. The solution mixture was stirred at RT 16 h and the formation of lanthanoid complex was checked by LC/MS. The previously mentioned solution mixture was added into 6.5 mL dry MeCN. Subsequently, $NaHCO_3$ (15 equiv.) and bromoacetyl bromide (15 equiv.) were added to the reaction mixture. The reaction mixture was stirred at RT for 30 mins and monitored by LC/MS. When the starting material was completely converted to product, the reaction was quenched by few drops of water and run through a Prevail™ C18 column (Alltech® Etract-Clean™) with a 10-50% acetonitrile gradient. The acetonitrile was removed *in vacuo* and the remained aqueous solution was directly used for protein labeling experiment. HR-MS m/z: 947.0408 $[M]^+$, calcd $[C_{28}H_{36}Br_2LuN_8O_8]$: 947.0407.



CLaNP-10

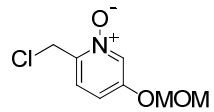
Compound **16** (1.6 g, 6.4 mmol), **17** (1.1 g, 2.9 mmol) and K_2CO_3 (1 g, 7.3 mmol) were dissolved in 30 mL MeCN and the solution mixture was stirred at 50 ^oC for 2 days. The excess K_2CO_3 was removed by filtration and the filtrate was concentrated under reduced pressure. The obtained crude compound **18** was redissolved in a solution mixture containing 6 mL TFA and 2 mL CH_2Cl_2. After 4 h stirring at RT, the TFA was removed by coevaporation with toluene (twice), and then dissolved in 30 mL MeCN. 2-(chloromethyl) pyridine-*N*-oxide (2.3 g, 16 mmol) and K_2CO_3 (2.2 g, 16 mmol) were added to the previously described solution with continuous stirring at 80 ^oC for 1 day. The mixture was filtered and concentrated *in vacuo* yielding crude **19**. Crude compound **19** (100 mg) was dissolved in 2 mL dry MeCN and PPh_3Br_2 (20 equiv.) was added at 0 ^oC. The reaction mixture was allowed to warm up from 0 ^oC to RT with continuous stirring for 16 h, and then which was extracted with CH_2Cl_2. Removing the organic layer, the water layer was purified by HPLC (0.1% TFA and a 10-30% acetonitrile gradient on C18 preparative column) yielding **CLaNP-10** (6 % yield over 3 steps). ^1H NMR (600 MHz, D_2O, 323 K): δ = 8.77 (d, 2H, *J* = 6.6 Hz), 8.54 (s, 2H), 8.21 (d, 2H, *J* = 7.8 Hz), 8.05 (t, 2H, *J* = 7.8 Hz), 7.90 (t, 2H, *J* = 7.2 Hz), 5.14 (t, 4H, *J* = 6 Hz), 4.54 (s, 4H), 4.39 (s, 4H), 4.14

(s, 4H, *J* = 6 Hz), 3.87-3.56 (br, 8H), 3.50 (br, 8H). ^{13}C NMR (150 MHz, D_2O, 323 K): δ = 144.72, 140.35, 135.67, 131.05, 128.96, 127.56, 127.14, 51.89, 51.55, 49.53, 47.96, 47.28, 30.25. HR-MS m/z: 834.2380 $[M]^+$, calcd $[C_{36}H_{42}Br_2N_{12}O_2]$: 834.1894.



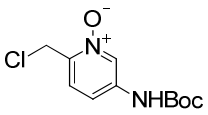
(5-methoxymethoxy-2-hydroxymethyl)pyridine-*N*-oxide (4)

To a solution of **3** (900 mg, 5.3 mmol) in 25 mL chloroform *m*CPBA (1.6 g, 6.4 mmol) was added and the reaction mixture was stirred at RT for 16 h. The mixture was quenched with saturated aqueous $Na_2S_2O_3$, extracted with CH_2Cl_2, dried over $MgSO_4$ and concentrated *in vacuo*. Purification by silica column chromatography afforded **4** (689 mg, 70%) as a white solid. R_f = 0.16 (5% methanol in ethyl acetate). m.p. = 81-82 ^{o}C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.11 (d, 1H, *J* = 2.4 Hz), 7.19 (d, 1H, *J* = 8.8 Hz), 7.03 (dd, 1H, *J* = 8.8 Hz, *J* = 2.4 Hz), 5.14 (s, 2H), 4.71 (s, 2H), 3.44 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 154.83, 143.54, 129.79, 124.33, 115.86, 94.96, 60.95, 56.44. FTIR: 3140.2, 3056.3, 2920.3, 1620.3, 1555.6, 1501.6, 1393.6, 1269.2, 1157.3, 1066.7, 1003.0, 917.2, 825.6 cm^{-1}. HR-MS m/z: 186.0765 $[M+H]^+$, calcd $[C_8H_{12}NO_4]$: 186.0761.



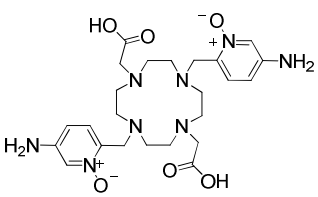
2-(chloromethyl)-5-(methoxymethoxy)pyridine-*N*-oxide (5)

Compound **4** (100 mg, 0.54 mmol) was dissolved in 2 mL dry CH_2Cl_2 and cooled with an ice bath. To this cooled solution MsCl (74.5 mg, 0.65 mmol) and TEA (84.0 mg, 0.65 mmol) were added and the reaction was stirred for 1 h at 0 ^{o}C. Lithium chloride (226.8 mg, 5.4 mmol) was added into the solution with continuous stirring at RT for 16 h. The white solid and the excess lithium chloride were filtered off, the filtrate was diluted with CH_2Cl_2 and extracted with brine. The organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. Purification by silica column chromatography afforded **5** (109.7 mg, 91%) as light yellow solid. R_f = 0.41 (5% methanol in ethyl acetate). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.12 (d, 1H, *J* = 2 Hz), 7.38 (d, 1H, *J* = 8.8 Hz), 6.99 (dd, 1H, *J* = 8.8 Hz, *J* = 2 Hz), 5.12 (s, 2H), 4.72 (s, 2H), 3.41 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 155.11, 140.84, 129.66, 125.44, 114.75, 94.79, 56.37, 39.26. FTIR: 3038.9, 2966.6, 1622.2, 1557.6, 1501.6, 1432.2, 1398.4, 1304.9, 1259.6, 1183.4, 1148.7, 1015.6, 992.4, 915.3 cm^{-1}. HR-MS m/z: 204.1384 $[M+H]^+$, calcd $[C_8H_{11}ClNO_3]$: 204.0422.



5-(tert-butoxycarbonylamino)-2-(chloromethyl)pyridine-*N*-oxide (10)

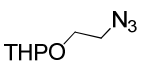
To a solution of **9** (140 mg, 0.57 mmol) in 6 mL chloroform was added *m*CPBA (169.3 mg, 0.69 mmol). After 4 h stirring at room temperature, the reaction mixture was diluted with 10 mL chloroform and quenched with $Na_2S_2O_3$. The crude reaction mixture was washed with brine and the organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. Silica column chromatography afforded title compound **10** (90%, 132.3 mg, 0.51 mmol) as a white solid. The compound decomposed above 194 ^oC. R_f = 0.2 (EtOAc/PET : 2/1 (v/v)). ^1H NMR (400 MHz, d^6-DMSO): δ = 9.87 (s, 1H), 8.50 (s, 1H), 7.58 (d, 1H, *J* = 8.8 Hz), 7.38 (d, 1H, *J* = 8.8 Hz), 4.79 (s, 2H), 1.47 (s, 9H). ^{13}C NMR (100 MHz, d^6-DMSO): δ = 152.3, 139.8, 138.9, 129.0, 126.8, 114.8, 80.5, 39.5, 27.9. FTIR: 3146.4, 3029.3, 2977.3, 1725.4, 1573.9, 1509.4, 1382.0, 1252.8, 1172.8, 1054.1, 988.6, 871.9 cm^{-1}. HR-MS m/z: 259.1905 $[M+H]^+$, calcd $[C_{11}H_{16}ClN_2O_3]$: 259.0844.



6,6'-(4,10-bis(carboxymethyl)-1,4,7,10-tetraazacyclo-dodecane-1,7-diyl)bis(methylene)bis(3-aminopyridine-*N*-oxide) (12)

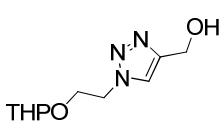
Compound **10** (200 mg, 0.39 mmol), **6** (219 mg, 0.85 mmol) and $NaHCO_3$ (71 mg, 0.85 mmol) were dissolved in 4 mL DMF. The solution mixture was heated to 50 ^oC with continuous stirring for 2 days and product formation was monitored by LC/MS. When the reaction was complete, the excess $NaHCO_3$ was filtered off and the reaction mixture was concentrated *in vacuo* without further purification giving crude compound **11**. Crude compound **11** was taken up in a TFA/CH_2Cl_2 (3/1, v/v, 2 mL) solution mixture. After 4 h stirring at RT, the TFA was removed by coevaporating with toluene (twice) and the crude mixture was redissolved in a 1,4-dioxane and NaOH solution mixture (The final concentration of starting material was 15 mM and 0.4 M for NaOH). The reaction was stirred for 4 h at RT, after which it was neutralized by acid resin, followed by filtrating off the resin, concentration, and purification by HPLC (0.1% TFA and a 0-10% acetonitrile gradient on C18 preparative column) yielding **12** (60%, over 3 steps). ^1H NMR (600 MHz, D_2O, 323 K): δ = 8.24 (d, 2H, *J* = 2.4 Hz), 7.81 (d, 2H, *J* = 9 Hz), 7.36 (dd, 2H, *J* = 9 Hz, *J* = 2.4 Hz), 4.91 (s, 4H), 3.68 (br, 8H), 3.60 (s, 4H), 3.48 (br, 8H). ^{13}C NMR (150 MHz, D_2O, 323 K): δ = 174.7, 163.8, 149.8,

131.4, 129.0, 128.2, 118.6, 116.7, 54.9, 54.1, 52.5, 49.6. FTIR: 3344.7, 3203.9, 3065.9, 1666.5, 1611.6, 1425.5, 1169.9, 1126.5, 1085.9, 978.9, 829.4 cm^{-1}. HR-MS m/z: 532.4383 $[M]^+$, calcd $[C_{24}H_{36}N_8O_6]$: 532.2752.



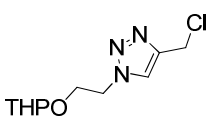
2-(2-azidoethoxy)-tetrahydro-2H-pyran (14)

To a solution of **13** (1 g, 11.5 mmol) in 38 mL dry CH_2Cl_2 DHP (1.3 mL, 13.8 mmol) and PTTs (288 mg, 1.15 mmol) were added. After 16 h stirring at RT, the reaction mixture was washed with $NaHCO_3$, dried by $MgSO_4$, concentrated *in vacuo* and purified by column chromatography yielding compound **14** (95%, 1.9 g). R_f = 0.38 (EtOAc/PET = 1:1 (v/v)). ^1H NMR (400 MHz, $CDCl_3$): δ = 4.64 (t, 1H), 3.92-3.82 (m, 2H), 3.59-3.50 (m, 2H), 3.42-3.33 (m, 2H), 1.86-1.58 (m, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 98.8, 66.4, 62.0, 50.9, 30.4, 25.4, 19.1. FTIR: 2941.4, 2870.0, 2098.5, 1440.8, 1282.7, 1122.6, 1033.8, 977.9, 871.8, 813.9 cm^{-1}. HR-MS m/z: 172.0969 $[M+H]^+$, calcd $[C_7H_{14}N_3O_2]$: 172.1081.



(1-(2-(tetrahydro-2H-pyran-2-yloxy)ethyl)-1H-1,2,3-triazol-4-yl)methanol (15)

Compound **14** (856 mg, 5 mmol) was dissolved in a solvent mixture of THF (7 mL) and H_2O (3 mL). Propargyl alcohol (321 µL, 5.5 mmol), 100 µL of 500 mM $CuSO_4$ (aq.) and 1 mL of 1 M ascorbate acid (aq.) were added to the solution mixture. The reaction mixture was stirred at RT for 16 h, and then it was extracted with ethyl acetate. The organic layer was dried using $MgSO_4$, concentrated *in vacuo* and purified by column chromatography giving product **15** (96%, 1.09 g). R_f = 0.15 (Ethyl acetate). ^1H NMR (400 MHz, $CDCl_3$): δ = 7.65 (s, 1H), 4.69 (s, 2H), 4.55-4.43 (m, 3H), 4.30 (s, 1H), 4.04-3.99 (m, 1H), 3.73-3.67 (m, 1H), 3.62-3.56 (m, 1H), 3.41-3.36 (m, 1H), 1.72-1.39 (m, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 147.8, 122.9, 98.9, 65.7, 62.3, 55.9, 50.4, 30.3, 25.1, 19.3. FTIR: 3375.4, 2941.4, 2870.0, 1734.0, 1440.8, 1348.2, 1122.6, 1031.9, 869.9, 812.0 cm^{-1}. HR-MS m/z: 227.1278 $[M+H]^+$, calcd $[C_{10}H_{17}N_3O_3]$: 227.1264.



4-(chloromethyl)-1-(2-(tetrahydro-2H-pyran-2-yloxy)ethyl)-1H-1,2,3-triazole (16)

To a dry solution of **15** (9.58 g, 42.2 mmol) in 125 mL CH_2Cl_2 at 0 ^oC MsCl (3.94 mL, 50.6 mmol) and triethylamine (7.2 mL, 50.6 mmol) were added

with continuous stirring for 3 h. The reaction was extracted with brine, dried using $MgSO_4$ and concentrated *in vacuo*. The crude mixture was dissolved in 20 mL dry acetone and lithium chloride (10 equiv.) was added. The reaction was stirred 16 h at RT. Concentration *in vacuo*, redissolving in CH_2Cl_2, extraction with brine, drying ($MgSO_4$), concentrating, and purifying by silica chromatography gave product **16** (85%, 8.81 g). R_f = 0.22 (EtOAc/PET = 1:1 (v/v)). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.73 (s, 1H), 4.69 (s, 2H), 4.59-4.51 (m, 2H), 4.09-4.05 (m, 1H), 3.76-3.73 (m, 1H), 3.59-3.56 (m, 1H), 3.44-3.42 (m, 1H), 1.76-1.64 (m, 2H), 1.56-1.53 (m, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 144.6, 123.9, 99.0, 65.6, 62.4, 50.6, 36.3, 30.4, 25.2, 19.3. FTIR: 2943.4, 2870.1, 1734.0, 1352.1, 1228.7, 1122.6, 1033.8, 964.4, 869.9, 813.9, 729.1 cm^{-1}. HR-MS m/z: 246.0782 $[M+H]^+$, calcd $[C_{10}H_{17}ClN_3O_2]$: 246.1004.

General method for protein labeling experiment

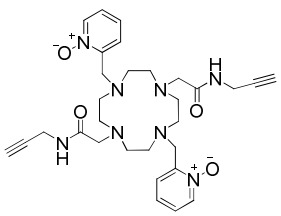
Both of the ^{15}N enriched Paz and T4Lys were kindly performed by Mrs. Anneloes Blok and Dr. Monika Timmer (Leiden University, Inst. Chemistry). To attach Ln-**CLaNP-9** to T4Lys N55C/V57C and Paz E51C/E54C, protein sample (1 mL, 150-300 μM) was treated with DTT (final concentration 5 mM) at 0 ^{o}C for 1 h to remove possible dimers. The reaction mixture was loaded on a PD-10 column (GE Healthcare) pre-equilibrated with labeling buffer (20 mM Tris, 150 mM NaCl and pH 8.0) to remove DTT. To avoid any reoxidation by air, the buffer was degassed and the PD-10 column kept under an argon atmosphere. To the eluted protein eight equivalents Ln-**CLaNP-9** were added. For T4Lys N55C/V57C, the labeling reaction was stirred at RT for 6 h, after which the tagged protein was purified on a HiTrap-SP column. The mass of the resulting ^{15}N T4Lys Lu-**CLaNP-9** (20478 ± 2 Da) agreed with the expected mass of 20478 Da assuming 98% ^{15}N enrichment. In the case of Paz E51C/E54C, the protein was activated with TCEP, after which the probe attached Paz sample was concentrated to 500 μL, $K_3[Fe(CN_6)]$ was added until 1 mM, and purified over a Superdex 75 gel filtration column. The mass of ^{15}N Paz Yb-**CLaNP-9** (14410 ± 2 Da) also agreed with expected mass 14411 Da. The yield of labeling after purification was 50 ~ 80%, which estimated from the intensity of diamagnetic peaks in the spectra of samples with paramagnetic tags.

PCS analysis

The PCS analysis, Δχ-tensor calculation and Q factor are based on published method (Chapter 2). Instead of NIH-XPLOR, a program NUMBAT[168] was used to calculated the Δχ-tensor. The structures of Paz and T4Lys were taken from PDB entries 1PY0 and 3DKE and hydrogens were added. The variation in the Ln positions was calculated by randomly removing 10% PCS data and repeating the Δχ-tensor fit 100 times.

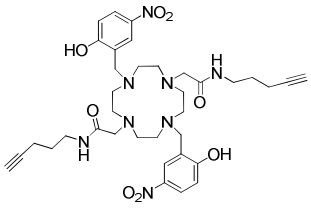
General protocol for CLaNP-click series

To a solution of **17** (100 mg, 0.31 mmol) in 5 mL MeCN **20a** or **20b** (2.5 equiv.) and K_2CO_3 (106.9 mg, 0.77 mmol) were added with continuous stirring at RT for 16 h. The excess K_2CO_3 was filtered off and the filtrate was concentrated *in vacuo* without any purification. The concentrated crude compounds were re-dissolved in 10 mL MeCN. To synthesize **CLaNP-11** and **CLaNP-12**, K_2CO_3 (2.5 equiv.), 2-(chloromethyl)pyridine-N-oxide and 2-(bromomethyl)-4-nitrophenol (2.5 equiv.), respectively, were added and the reaction mixture was stirred at 80 ^oC. After 16 h stirring, the reaction was allowed to cool to RT, the excess K_2CO_3 was filtered off, concentrated, and purified by HPLC.



CLaNP-11

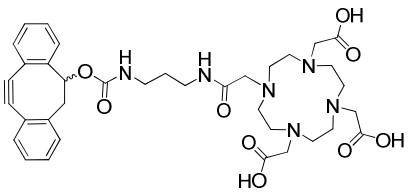
Purification by HPLC (0.1% TFA and a 5-30% acetonitrile gradient on C18 preparative column) yielded **CLaNP-11** (41.6 %). ^1H NMR (600 MHz, D_2O, 323 K): δ = 8.75 (d, 2H, *J* = 6.6 Hz), 8.13 (d, 2H, *J* = 7.8 Hz), 8.06 (t, 2H, *J* = 7.8 Hz), 8.00 (t, 2H, *J* = 7.8 Hz). 4.99 (s, 4H), 4.35 (s, 4H), 3.65-3.62 (br, 8H), 3.54-3.47 (br, 12H), 2.85 (t, 2H, *J* = 2.4 Hz). ^{13}C NMR (150 MHz, D_2O, 323 K): δ = 170.29, 139.76, 132.18, 129.93, 128.44, 117.16, 115.22, 79.70, 71.65, 56.77, 53.80, 50.40, 49.02, 28.69. FTIR: 3556.7, 3408.2, 3244.3, 2985.8, 2125.1, 1654.8, 1124.5, 829.4 cm^{-1}.



CLaNP-12

Final product was purified by HPLC (0.1% TFA and a 0-50% acetonitrile gradient on C18 preparative column). ^1H NMR (600 MHz, d^6-DMSO, 323 K): δ = 8.50 (s, 2H),

8.21-8.19 (m, 4H), 7.12 (d, 2H, *J* = 6 Hz), 4.35 (s, 4H), 3.67 (br, 2H), 3.26-3.30 (m, 16H), 3.09 (s, 8H), 2.67 (t, 2H, *J* = 2.4 Hz), 2.19 (t, 4H, *J* = 7.2 Hz), 1.66-1.62 (m, 4H). ^{13}C NMR (150 MHz, d^6-DMSO, 323 K): δ = 169.58, 163.81, 139.24, 129.31, 127.02, 116.31, 83.67, 71.02, 54.33, 51.14, 49.65, 48.97, 37.77, 27.72, 15.29. FTIR: 3304.1, 3115.0, 2958.8, 2112.0, 1647.2, 1338.6, 1197.8, 1080.1, 717.5 cm^{-1}. HR-MS m/z: 721.3672 $[M+H]^+$, calcd $[C_{36}H_{49}N_8O_8]^+$: 721.3668.



CLaNP-DIBO

Pyridine (0.36 mL, 4.55 mmol) and *p*-nitrophenyl chloroformate (362 mg, 7.28 mmol) were added into a CH_2Cl_2 (27 mL) solution containing **DIBO** (200 mg, 0.91 mmol). The reaction was stirred 16 h, and then quenched with brine (3 mL). The solution mixture was extracted with CH_2Cl_2 and the organic layer was washed with brine, dried by $MgSO_4$, concentrated under reduced pressure to provide crude **22**. To the solution of **22** in 5 mL CH_2Cl_2 was added dry TEA and 1,3-diaminopropane and the reaction was stirred at room temperature. After being stirred for 30 min, the reaction was diluted with CH_2Cl_2 (10 mL), washed with 1 M NaOH (10 mL x 2) and brine (10 mL). The organic phase was dried ($MgSO_4$), filtered, and evaporated. The crude residue was dissolved in MeOH (2.3 mL) and added into a water solution mixture (6.9 mL) containing DOTA (232 mg, 0.91 mmol) and HOBT (61 mg, 0.91 mmol). The reaction mixture was cooled on ice and pH was adjusted to 4.5-5.0 using N,N-diisopropylethylamine. EDC (130 mg) in water (2.3 mL) was added slowly into the mixture which was then stirred with cooling. After a 30 min stirring, the pH was elevated to 8 by adding *N*,*N*-diisopropylethylamine and the solution was stirred for a further 30 min at room temperature. The reaction was monitored by LC/MS and the title compound was purified by HPLC (0.1% TFA and a 35-55% acetonitrile gradient on C18 preparative column). ^{1}H NMR (600 MHz, D_2O, 323 K): δ = 7.76 (d, 1H, *J* = 7.8 Hz), 7.73-7.64 (m, 7H), 5.58 (m, 1H), 4.07-3.96 (m, 8H), 3.55-3.47 (m, 24H). ^{13}C NMR (150 MHz, D_2O, 323 K): δ = 158.7, 152.49, 150.26, 133.34, 132.95, 131.43, 129.88, 128.86, 127.74, 127.70, 126.62, 125.03, 124.30, 121.76, 114.21, 111.14, 78.46, 70.90, 56.36, 55.43, 51.13, 50.83, 46.79, 39.4, 38.1. FTIR: 3352.3, 3068.7, 2856.6, 1674.2, 1199.2, 1134.1 cm^{-1}. HR-MS m/z: 707.2844 $[M+H]^+$, calcd $[C_{36}H_{47}N_6O_9]^+$: 707.3399.

Protein production and purification

The production and purification of the ^{15}N enriched bacteriophage T4 Lysozyme (T4Lys) single AzF mutant, N55AzF and D72AzF, double Cys mutant N55C/V57C and the ^{15}N enriched double Cys mutant Paz E51C/E54C were performed by Ms. Anneloes Blok and Dr. Monika Timmer (Leiden University, Inst. Chemistry). The expression and purification of Paz E51C/E54C were followed by published protocol (Chapter 2).

Double Cys mutation T4Lys N55C/V57C were prepared by the WHOPS method[180] using the expression plasmid pET28-T4Lys as a template. The oligonucleotides 5'-GGTCGTAATGCCTGTGGTTGCATTACCAAAGATGAAG CCG-3' and its complement were used as the forward and reverse primers, respectively. *Escherichia coli* BL21 (DE3) was used to express the mutant gene. The transformed cells were incubated overnight in 10 mL ^{15}N rich M9 medium with kanamycin (final concentration 50 μg/mL) at 37 °C, 250 rpm. Overnight pre-cultures were diluted 100 times and incubated at 37 ^{o}C, 250 rpm. When the OD_{600} value was around 0.8, cultures were induced by adding 1 mM IPTG. The temperature was reduced to 30 ^{o}C and cells were harvested 18 h after induction by centrifugation. Cells were lysed with a French pressure cell, and the lysates cleared by centrifugation (45 minutes, 2,500 rpm). Cleared lysate was loaded on a 5 mL HisTrap column (equilibrated in 20 mM Tris pH 7.5, 500 mM NaCl) and eluted with a 0 to 500 mM linear imidazole gradient. After a Ni^{2+}-column, EDTA (1 mM) was added to the concentrated protein and the protein was purified over a Superose 12 column (equilibrated in 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM DTT).

In the case of AzF mutant, single AzF mutations T4Lys N55AzF and T4Lys D72AzF were prepared by the WHOPS method[180] using the expression plasmid of wt T4Lys as a template. The oligonucleotides 5'-GGTCGTAATGCCTAGGGT GTGATTACCAAAGATGAGGCCG-3' and its complement were used as the forward and reverse primers, respectively, for N55AzF. Similarly, the oligonucleotides 5'-CTGTTTAATCAGGATGTTTAGGCAGCCGTTCGTGGAAT TCTGC-3' and its complement were used as the forward and reverse primers, respectively, for D72AzF. pEVOL-AzF (coding for the AzF-tRNA transferase and AzF-tRNA, CAMr) were kindly provided by Prof. Peter Schultz, Scripps Research Institute (California) and was first transformed to *Escherichia coli* BL21 (DE3) and cells were made chemically competent. Then, the pEVOL containing BL21 (DE3)

were transformed with the KANr plasmid coding for the T4Lys AMBER mutant. The doubly transformed cells were grown overnight in 10 mL ^{15}N rich M9 Medium with kanamycin (final concentration is 50 μg/mL) and chloramphenicol (final concentration is 34 μg/mL). Overnight pre-cultures were diluted 100 times, grown for approximately 3 h at 37 ^{o}C, 250 rpm, after which the temperature was reduced to 30 ^{o}C. When the OD_{600} value was around 1.4, cultures were induced by adding 1 mM IPTG, 0.02% (w/v) arabinose and 1, 2 or 5 mM Paz. Cells were harvested 24 h after induction by centrifugation, resuspended in 20 mM Tris pH7.5, 500 mM NaCl). Cells were lysed by French press, cleared by centrifugation (45 minutes, 9,000 rpm). Cleared lysate was loaded on a 5 mL HisTrap column (equilibrated in 20 mM Tris pH 7.5, 500 mM NaCl) and eluted with a 0 to 500 mM linear imidazole gradient. The final yields were 96 mg/L and 112 mg/L of N55AzF and D72AzF culture, respectively.

Protein labeling experiment

The CuAAC reaction was performed in 20 mM HEPES buffer containing 150 mM NaCl (pH = 7.4). A solution mixture of T4LysAzF (50 μM) and paramagnetic NMR probe (250 μM) was treated with 1 mM TCEP, and then 500 μM of $CuSO_4$ and 50 μM of TBTA in 1:4 DMSO/tBuOH were added. After incubation for 30 min at room temperature, the reaction was stopped by adding 1 μL of 1 M EDTA. All the reactions were checked by SDS-PA gel and excess probe was removed by gel filtration chromatography.

For the ring strain-promoted [3+2] cycloaddition, CLaNP-DIBO coordinated to Gd^{3+} was prepared in DMF and was added to a protein solution, which was in 20 mM HEPES buffer containing 150 mM NaCl (pH = 7.4). The final concentration of the protein was 50 μM and the probe was in five time excess. After 16 h stirring at 4 ^{o}C, the reaction mixture was concentrated to 1 mL and applied to a PD10 column to remove the unreacted probe. The filtrate was concentrated without further purification and prepared for NMR spectroscopy.

NMR spectroscopy on proteins

The NMR samples of Paz Ln-**CLaNP-9** (80-150 μM) were prepared in 20 mM sodium phosphate, 150 mM NaCl buffer and 6% (v/v) D_2O at pH = 7. The Cu(II) was reduced by 5 equiv. TCEP. All T4Lys samples (100-200 μM) contained 20 mM

HEPES, 150 mM NaCl buffer and 6% (v/v) D_2O at pH 7.2. All [^{15}N, ^{1}H]-HSQC were recorded at 298 K on a Bruker Avance III 600 MHz spectrometer. Data were processed with NMRPipe and analyzed with CCPNMR Analysis version 2.1. Assignments of the resonances were based on previous work[181] and kindly provided by Mr. Simon P. Skinner (Leiden University, Inst. Chemistry).