

New Lanthanide chelating tags for PCS NMR spectroscopy with reduction stable, rigid linkers for fast and irreversible conjugation to proteins.

Thomas Müntener, Jérémy Kottelat [†], Annika Huber, Daniel Häussinger *

Department of Chemistry, University of Basel, St. Johannis-Ring 19, 4056 Basel, Switzerland

KEYWORDS: Pseudocontact shift NMR, DOTA, pyridinethiazole, lanthanide chelating tag, thioether formation, reduction stable

ABSTRACT: Lanthanide chelating tags (LCT) have been used with great success for determining structures and interactions of proteins and other biological macromolecules. Recently LCTs have also been used for *in-cell* NMR spectroscopy, but the bottleneck especially for demanding applications like pseudocontact shift (PCS) NMR is the sparse availability of suitable tags that allow for site-selective, rigid, irreversible, fast and quantitative conjugation of chelated paramagnetic lanthanide ions to proteins via reduction stable bonds. We report here several such tags and focus on a new pyridine thiazole derivate of DOTA, that combines high affinity, rigidity and selectivity with unprecedented tagging properties. The conjugation to the cysteine thiol of the protein results in a reductively stable thioether bond and proceeds virtually quantitatively in less than 30 min at 100 μ M protein concentration, ambient temperature and neutral pH. Upon conjugation of the new tag to two single cysteine mutants of ubiquitin and a single cysteine mutant of human carbonic anhydrase type II (30 kDa) only one stereoisomer is formed (square antiprismatic coordination, $\Lambda(\delta\delta\delta\delta)$) and large to very large pseudocontact shifts as well as large residual dipolar couplings (RDC) are observed by NMR spectroscopy. The PCS and RDC show excellent agreement with the solid state structure of the proteins. We believe that the pyridine thiazole moiety reported here has the potential to serve as thiole reactive group in various conjugations applications, furthermore its terbium complex shows strong photoluminescence upon irradiation and may thus serve as a donor group for Förster resonance energy transfer spectroscopy.

INTRODUCTION

NMR has proven to be of great value for the accurate determination of protein structures and monitoring protein – protein¹ and protein – ligand interactions. In particular paramagnetic effects like pseudocontact shifts (PCSs) and residual dipolar couplings (RDCs) induced by lanthanide (III) ions provide extremely valuable long-range and angular restraints. Large PCS and strong paramagnetic alignment can only be obtained by site-specific and rigid incorporation of the paramagnetic source on the protein. Most of the currently used chemically synthesized high-performance lanthanide chelating tags²⁻⁶ (LCTs) are based on the DOTA framework and were extensively reviewed⁷⁻⁸. These tags offer a highly versatile approach to selectively conjugate lanthanide (III) ions to proteins and other biomolecules. In contrast to related applications like paramagnetic relaxation enhancement (PRE), Förster resonance energy transfer (FRET) fluorescence microscopy and others, the conjugation for PCS tags has to fulfill numerous additional conditions. First, a short and rigid linker is beneficial to minimize motional averaging and second, the formation of only one stereoisomer is highly desirable as multiple stereoisomer may lead to peak doubling and complicated spectra.^{3, 9-11} Additionally, more general demands like fast, selective and irreversible reaction with the protein at close to neutral pH, ambient or lower temperature and concentrations in the order of 100 μ M or below have to be met. Conveniently the majority of DOTA-based tags described in literature are conjugated to the protein surface via site-specifically introduced cysteines by the formation of one^{3-6, 12-13} or two disulfide^{2, 11, 14-16} bonds. This approach, however, introduces additional flexibility into the linkage itself and prohibits the use of these tags in strongly reducing environ-

ments like in living cells. Disulfide bridges are readily cleaved by DTT, TCEP or glutathione. A number of tags have been reported recently which can be used for *in-cell* PCS NMR spectroscopy which address the inherent instability of disulfide bonds under reducing conditions by forming non-reducible linkages. Targeting unnatural amino acids¹⁷⁻¹⁸ instead of cysteine residues allows the usage of click-chemistry for the formation of reduction stable triazoles. Solvent exposed cysteine residues can form stable thioether bonds with α -halogen acetamides^{13, 15}, vinyl pyridines¹⁹ and 4-phenyl sulfone pyridines^{5-6, 20}. While the first two approaches result in relatively long and flexible linkers and as a consequence only moderate PCS, the last approach based on phenylsulfones yields excellent PCS tensors but is hampered by sluggish reaction kinetics that require either elevated temperatures or high pH values ≥ 9 during long reaction times and suffers from incomplete tagging.^{5-6, 20} Tagging of proteins under these harsh conditions is not suitable for most more labile proteins and thus has been limited to very small and stable proteins like ubiquitin or GB1. In the current work we have improved the performance of our initial tag DOTA-M7Py by the introduction of fluorine substituents in various positions, yielding significantly faster conjugation rates. In the context of these efforts we explored an activated pyridine-thiazole as a new activator group and accomplished selective and irreversible thioether formation within minutes in diluted protein solution under physiological conditions, which makes this functional group a very promising candidate for the conjugation of small molecules to thiol functionalized proteins.

RESULTS

Synthesis of CF₃ substituted tags and Protein conjugation. Our previously reported Ln-M7PySO₂Ph-DOTA tags⁵ suffered from slow reaction rates towards thiols and, therefore, required elevated temperatures for successful tagging of GB1 mutants. We envisioned the synthesis of CF₃ substituted phenyl leaving groups in order to enhance reactivity towards nucleophilic displacement. Scheme 1 illustrates the synthetic route towards CF₃ substituted analogues of Ln-M7PySO₂Ph-DOTA. Starting from commercially available 4-CF₃ and 3,5-bis-CF₃ substituted thiophenols, we prepared α -bromo methyl pyridine sulfones **1** and **2** by nucleophilic substitution on 4-bromo-2-methyl-pyridine in DMF followed by oxidation of the resulting thiols to sulfones and final NBS bromination using AIBN as radical generating reagent (c.f. Schemes S1 and S2). Attachment of sulfones **1** and **2** to three-fold lactic acid *tert*-butyl ester alkylated M4 cyclen and subsequent acidic deprotection with aqueous hydrochloric acid yielded the free ligands **4** and **5** in good yields. The final metalation step led to diamagnetic (Lu) tags **7** and **9** and paramagnetic (Tm) tags **8** and **10**.

Tagging of the protein G B1 domain (GB1) E42C with the new tags **7** – **10** was performed with a four-fold excess of tag and a protein concentration of 300 μ M in 50 mM phosphate and 0.5 mM TCEP buffer pH 7.0 at room temperature. We found, however, no increased reactivity compared to our previously reported Ln-M7PySO₂Ph-DOTA tags.

Synthesis of M7FPy-DOTA-based tags and Protein conjugation. Our next strategy to increase the reactivity involved the synthesis of a 5-fluoro pyridine-based linker, thus decreasing the electron density of the pendant pyridine ring.

Scheme 1 shows the synthetic route towards M7FPy-DOTA tags **11** and **12**. Starting from commercially available 2-methyl-4-bromo-5-fluoro pyridine we synthesized α -bromo methyl pyridine sulfone **3**. Alkylation and deprotection of three-fold lactic acid *tert*-butyl ester alkylated M4 cyclen yielded the free ligand **6**. We prepared diamagnetic (Lu) tag **11** and paramagnetic (Tm) tag **12** (c.f. Scheme S3).

For tagging reactions, we initially used ubiquitin S57C at a concentration of 100-150 μ M in 10 mM phosphate and 2 mM TCEP buffer pH 7.0 at room temperature. The reaction progress was monitored by ESI-MS and we found a conversion of up to 80% within three hours and >90% in six hours. Close inspection of the ESI-MS spectra revealed partial decomposition of the tag caused by the reaction with TCEP. Therefore, we tagged the protein without additional TCEP. To our surprise we found that under these conditions around 30% of the protein was dimerized after six hours at room temperature. The formation of a non-reducible thioether bond between the tag and the protein releases one molecule of benzenesulfinic acid. Decomposition of the tag with TCEP also leads to the formation of benzenesulfinic acid. We found that benzenesulfinic acid catalyzes the formation of ubiquitin dimers. Best results were obtained by tagging 150 μ M ubiquitin S57C or hCA-II S166C C206S in the presence of 0.45 mM TCEP and using a six-fold excess of the tag. These conditions ensured conversions greater >95% within 16h at room temperature.

We synthesized two additional M7FPy-DOTA-based tags (c.f. Scheme S4) bearing a sulfonamide and a methyl sulfone as leaving group, thus avoiding the release of benzenesulfinic acid. Both tags showed low conversions of cysteine methyl ester under tagging conditions, however, the tag bearing a sulfonamide did not react with ubiquitin S57C while the tag

bearing a methyl sulfone yielded conversions of 5-10% within 24h at room temperature.

Synthesis of M7PyThiazole-based tags and Protein conjugation. A significantly more reactive activator was found in the pyridine thiazole scaffold, increasing the reaction rate of the nucleophilic aromatic substitution by more than an order of magnitude (c.f. Figure 1). Scheme 2 illustrates two synthetic routes towards Ln-M7PyThiazole-SO₂Me-DOTA tags. First, we synthesized the α -bromo methyl pyridinethiazole methyl sulfone **13** from commercially available starting materials. Alkylation of three-fold lactic acid *tert*-butyl ester alkylated M4 cyclen and aqueous acidic deprotection yielded the free ligand **15** in moderate yield. Under the conditions applied, hydrolysis of the methyl sulfone occurred, forming a hydroxy pyridine thiazole. The free ligand **15** was metalated using the corresponding lanthanide triflates in aqueous ammonium acetate (100 mM) at 80°C for 18h in moderate yields. Here we observed the same problem of hydrolyzing the methyl sulfone. In order to avoid this problem, we synthesized the α -bromo methyl pyridinethiazole methyl thioether **14**. The methyl thioether analogue is not susceptible towards nucleophilic substitution by hydroxyl ions. Final alkylation was performed with similar yields and the final acidic deprotection was carried out using non-aqueous conditions. The subsequent metalation of the free ligand **16** with the corresponding lanthanide triflate was higher yielding and the final oxidation could be performed with meta-chloroperbenzoic acid in good yields. This reaction sequence allows for the synthesis of metalated Ln-M7PyThiazole-SO₂Me-DOTA in 50 mg scale starting from 100 mg of three-fold lactic acid *tert*-butyl ester alkylated M4 cyclen corresponding to ~40% overall yield. We prepared diamagnetic (Lu) tag **17** and paramagnetic (Tm, Tb, Dy, Sm) tags **18-20**. For Sm and Lu loaded M7PyThiazole-SO₂Me-

DOTA we recorded ¹H, ¹³C and 2D- HSQC, HMQC, COSY and NOESY spectra and determined the configuration of both complexes to be exclusively $\Lambda(\delta\delta\delta\delta)$ indicating that Dy, Tm and Tb most likely also adopt the $\Lambda(\delta\delta\delta\delta)$ configuration.

We expected these tags to be also susceptible towards decomposition by TCEP. Therefore, we tagged the protein at a concentration of 100-150 μ M in 10 mM phosphate and 0.1 mM TCEP buffer pH 7.0 at room temperature with a four-fold excess of tag. To our delight we found that ubiquitin S57C was tagged to an extent >95% with Ln-M7PyThiazole-SO₂Me-DOTA tags within 15 minutes. To ensure complete tagging the reaction was prolonged for additional four hours. Tagging of ubiquitin S57C for >48h did not result in tagging of undesired amino acid residues i.e. lysine or serine as was checked by 2D-¹H-¹⁵N-HSCQ spectra.

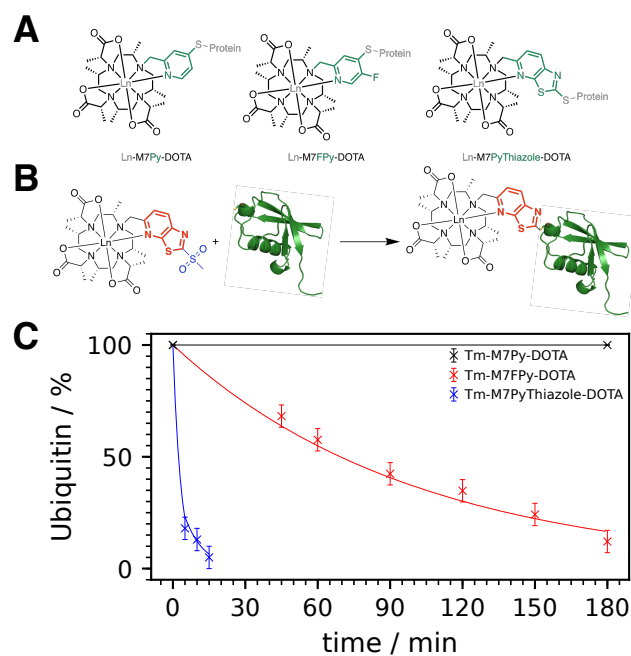
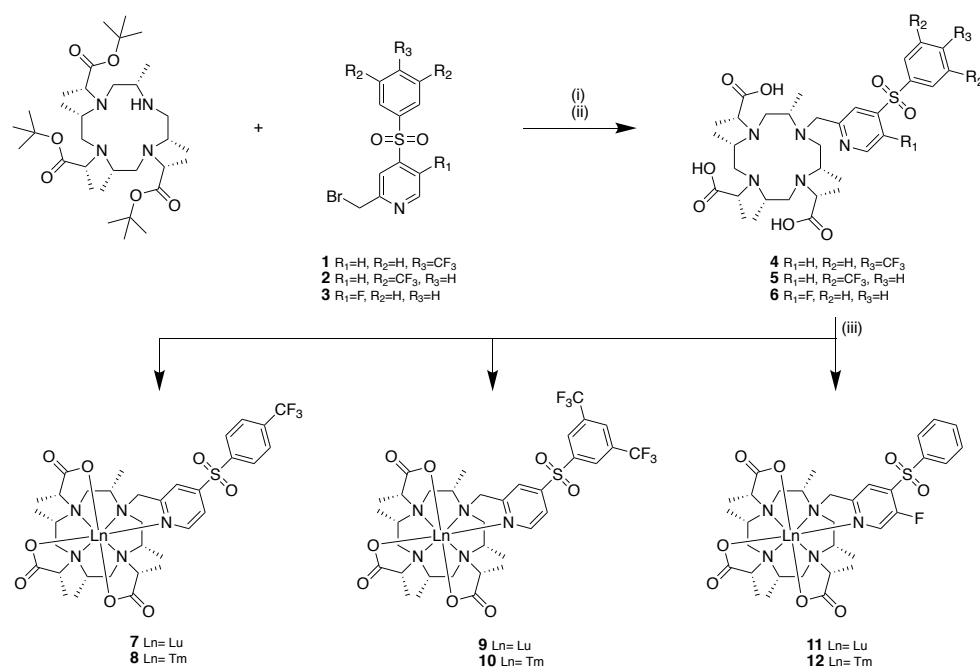


Figure 1. (A) Molecular formula of the three different LCTs investigated and their abbreviations. (B) Tagging reaction between Ln-M7PyThiazole-SO₂Me-DOTA and ubiquitin S57C forming a reduction stable thioether bond. (C) Reaction kinetics measured by HPLC-ESI-MS for tagging of ubiquitin S57C with

M7Py-DOTA derivatives (black), M7FPy-DOTA (red) and M7PyThiazole-DOTA (blue) in 10 mM phosphate and 0.45 μ M TCEP buffer pH 7.0. Protein concentrations were adjusted to 100-150 μ M.

Scheme 1. Synthesis of pyridine-based M7-DOTA LCTs^a



^a Reaction conditions: (i) K_2CO_3 , MeCN, RT, 18h; (ii) 1M HCl aq., MeCN, 80 °C, 4 h; (iii) $Ln(OTf)_3$, 100 mM NH_4OAc aq., 80 °C, 18h.

During tagging no formation of the hydrolyzed tag was observed indicating increased stability at pH 7.0 and room temperature. Nevertheless, we detected small quantities of the decomposition product originating from the reaction between the tag and TCEP. However, tagging under the conditions applied is orders of magnitude faster than the reaction with TCEP at room temperature and therefore negligible. Figure 1 shows the reaction kinetics of the tagging reaction between Tm loaded tags **8**, **10**, **12** and ubiquitin S57C measured by ESI-MS. For further kinetics studies we synthesized a tetrapeptide mimicking the tagging site of ubiquitin S57C (Leu-Cys-Asp) with a C-terminal tryptophan for increased UV absorbance. Table 1 shows the conversion of the tetrapeptide using a four-fold excess of Dy-M7PyThiazole-SO₂Me-DOTA under various conditions.

Table 1. Reaction kinetics of Dy-M7PyThiazole-DOTA

[TP] ^a / μ M	pH	T / K	Conversion % ^b	
			5 min	30 min
100	7.0	295	>99	-
100	7.0	278	76	>99
10	7.0	295	56	79 ^c
100	6.0	295	11	36
100	5.0	295	<1	2

^a TP: tetrapeptide (Leu-Cys-Asp-Trp), ^b Conversion determined by analytical HPLC-ESI-MS, ^c 84% conversion after 60 min.

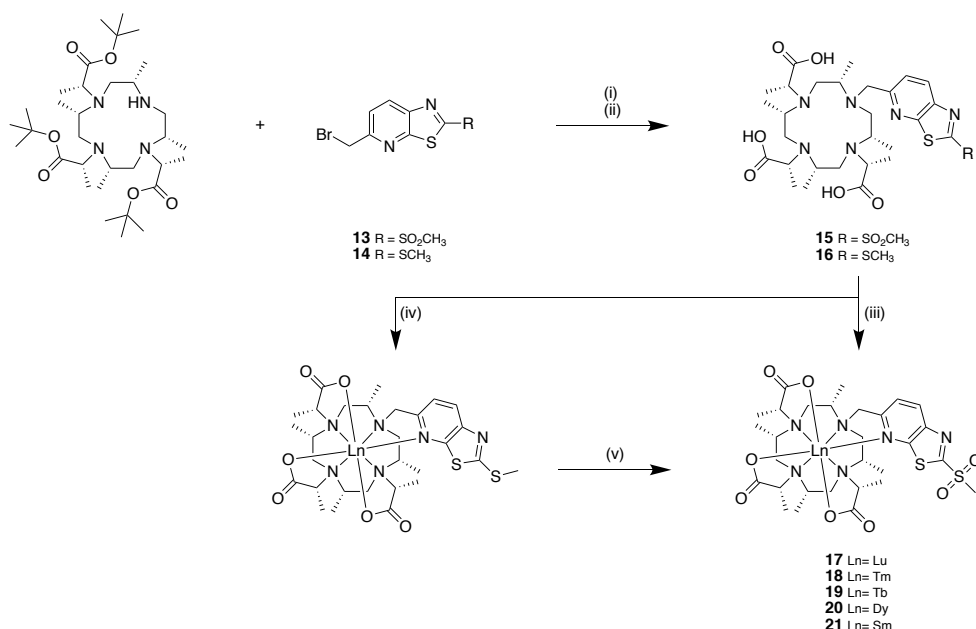
Dy-M7PyThiazole-SO₂Me-DOTA showed an extremely high reactivity at pH 7.0 offering a conversion >99% within five

minutes at 295 K and within 30 minutes at 273 K. Tagging even proceeded at peptide concentrations as low as 10 μ M or 100 μ M at pH 6.0. At pH 5.0 the desired product was still detectable by ESI-MS. Despite the high reactivity of the tag no formation of undesired *N*-terminal conjugation was observed.

PCS and RDC measurements and $\Delta\chi$ Determination. We prepared diamagnetic (Lu) and paramagnetic (Tm, Tb and Dy) M7PyThiazole-DOTA tagged uniformly 15 N labelled ubiquitin S57C and K48C samples, diamagnetic (Lu) and paramagnetic (Dy) M7PyThiazole-DOTA tagged selectively 15 N leucine labelled human carbonic anhydrase II (hCA-II) S166C C206S samples and diamagnetic (Lu) and paramagnetic (Tm) M7FPy-DOTA tagged uniformly 15 N labelled ubiquitin S57C samples.

We recorded 2D ^1H - ^{15}N HSQC and 2D ^1H - ^{15}N IPAP HSQC spectra at 600 MHz and 298 K and detected large pseudocontact shifts reaching up to 10 ppm and strong RDCs (amplitudes reaching 32 Hz). On ubiquitin S57C samples we detected 61, 54, 51 and 67 NH amide resonances out of a total of 72 for Tm-, Tb-, Dy-M7PyThiazole-DOTA and Tm-M7FPy-DOTA, respectively. Similarly, we detected 54, 54 and 53 resonances for Tm-, Tb and Dy-M7PyThiazole-DOTA on ubiquitin K48C. 23 out of 26 leucine resonances were detected for selectively ^{15}N leucine labelled hCA-II S166C C206S conjugated to Dy-PyThiazole-DOTA. Resonances in close proximity to the metal centers were broadened beyond detection due to the strong PRE-effect. Tm-M7FPy-DOTA induced predominately negative shifts on ubiquitin S57C whereas shifts induced by Tm-M7PyThiazole-DOTA were mostly positive. This originates from a different relative orientation of the tag with respect to the protein surface due to the different attachment site.

Scheme 2. Synthesis of Ln-M7PyThiazole-DOTA.^a



^a Reaction conditions: (i) K₂CO₃, MeCN, RT, 18h; (ii) 1M HCl aq., MeCN, 65 °C, 7h or TFA / CH₂Cl₂, RT, 36h; (iii) Ln(OTf)₃, 100 mM NH₄OAc aq., 80 °C, 18h; (iv) mCPBA, CH₂Cl₂, 5 °C → RT, 18h; (v) Ln(OTf)₃, 100 mM NH₄OAc aq., 80 °C, 18h.

We observed that M7PyThiazole-DOTA tags loaded with Tm are shifting in general to opposite directions compared to their Dy and Tb analogues and shifts measured for Dy were usually larger than those obtained for Tb. These circumstances allow for a simplified resonance assignment due to the fact that on superimposed spectra cross peaks belonging to one residue lie on a straight line with the diamagnetic reference peak being somewhere in-between (Figure 2B and S4 to S5). We assigned all cross peaks and fitted the components of the anisotropy of the magnetic susceptibility using the program Numbat²¹ and using only residues residing in stable secondary structure elements. The metal position was simultaneously fitted for samples tagged with Tm-, Tb- and Dy-PyThiazole-DOTA **18**, **19** and **20**. The pseudocontact shifts were fitted against the X-ray structure of the corresponding protein (ubiquitin 1UBI, hCA-II 3KS3). Table 2 shows the fitted $\Delta\chi$

parameters. Tensor parameters for Tm-M7FPy-DOTA on ubiquitin S57C are significantly smaller than those obtained for Tm-M7PyThiazole-DOTA. DFT geometry optimized structures of the corresponding Lutetium tags **11** and **17** revealed almost identical coordination geometries around the Lu centers with an RMSD of the eight coordinating atoms of 0.03 Å (c. f. Figure S18). Therefore, the significant difference in tensor magnitudes is very unlikely to be caused by a distortion of the coordination sphere. We attribute the increased performance to a partially hindered rotation around the newly formed C-S bond between the pyridine thiazole ring and the cysteine sulfur for proteins tagged with M7PyThiazole-based tags (c.f. Figure S16). Specifically, for proteins tagged with M7FPy-based tags the angle between the formed C(pyridine) – S(cysteine) bond and the metal - N(pyridine) bond is 180°, which means that a rotation around this C-S bond is almost

parallel to the principal axis of inertia, allowing for unhindered rotation. As a consequence, this rotation leads to rotational averaging and reduces the observed tensor in tags based on M7FPy. In marked contrast, the aforementioned angle is 100.1° in the M7PyThiazole case. Therefore, rotation around the C-S bond results in a much larger radius of gyration and presumably causes steric clashes of the tag with the protein and, thus, results in a restricted rotation and as a consequence in less averaging of the tensor anisotropy parameters.

In general, we found excellent agreement between the experimental PCS and the back-calculated PCS which is well represented in the small Q factors. The fitted tensor parameters for Dy-M7PyThiazole-DOTA on ubiquitin S57C are significantly larger than those obtained for the closely related Dy-M8-DOTA tag reaching magnitudes of $58.0 \times 10^{-32} \text{ m}^3$ for $\Delta\chi_{\text{ax}}$ and $24.6 \times 10^{-32} \text{ m}^3$ for $\Delta\chi_{\text{rh}}$. In agreement with the large PCS we also measured large RDC and calculated the tensor parameters using the web tool FANTEN²² and the metal position from the PCS fitting (c.f. Table S6).

Photoluminescence of Tb-M7PyThiazole-SO₂Me-DOTA.

During our synthesis of Ln-M7PyThiazole-SO₂Me-DOTA we realized a strong emission of in particular Tb loaded complexes under UV irradiation. We recorded UV/Vis absorption and fluorescence emission spectra (see Figure S17) and found an absorption maximum at 300 nm and typical Tb centered narrow emission bands at 487, 547, 587 and 622 nm resulting in a yellow emission color. These properties allow for potential

application in FRET spectroscopy.

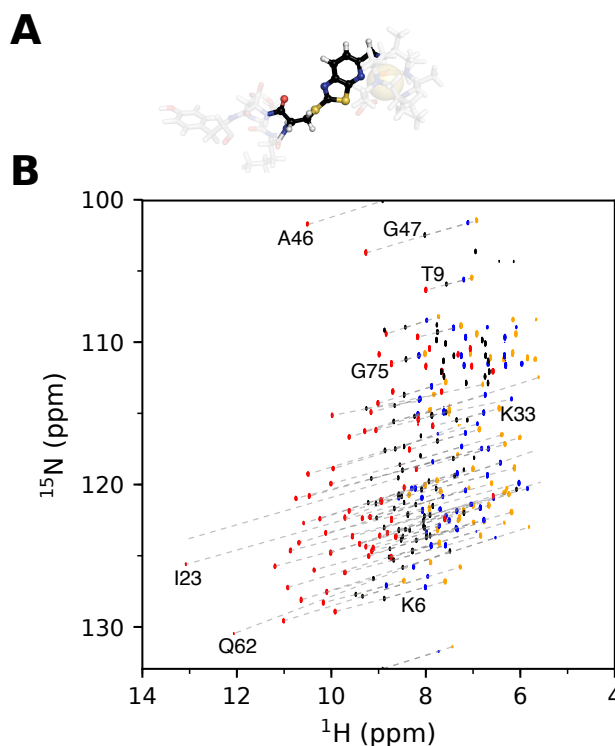


Figure 2. (A) Conjugation of Ln-M7PyThiazole-DOTA to ubiquitin S57C (only a small excerpt is displayed) via formation of a thioether bond between the Cys57 and the pyridine thiazole moiety. (B) ^1H - ^{15}N HSQC spectra of ubiquitin S57C tagged with Ln-M7PyThiazole-DOTA, diamagnetic Lu (black) and paramagnetic Tm (red), Tb (blue) and Dy (orange). All spectra were recorded at 298 K and 14.1 T in 10 mM phosphate buffer pH 6.0. Water signal removed for clarity.

Table 2. Magnetic Susceptibility Tensor Parameters ^a for Ubiquitin ^b and hCA-II ^c.

Protein	DOTA-tag	Ln ³⁺	$\Delta\chi_{ax} / 10^{32} \text{ m}^3$	$\Delta\chi_{rh} / 10^{32} \text{ m}^3$	$\alpha / ^\circ$	$\beta / ^\circ$	$\gamma / ^\circ$	Q ^d
Ubi S57C	M7FPy	Tm ^e	17.8 (1.5)	9.8 (1.0)	2.8 (8.3)	154.3 (7.0)	27.0 (6.4)	0.02
Ubi S57C	M7PyThiazole	Tm ^f	35.4 (0.6)	11.3 (0.7)	50.3 (1.6)	44.0 (0.9)	17.9 (3.0)	0.06
Ubi S57C	M7PyThiazole	Dy ^f	41.9 (0.8)	26.1 (0.5)	145.9 (1.3)	84.3 (1.1)	138.4 (2.6)	0.05
Ubi S57C	M7PyThiazole	Tb ^f	30.8 (1.1)	19.6 (0.7)	146.5 (0.9)	84.5 (1.3)	132.0 (4.1)	0.05
Ubi K48C	M7PyThiazole	Tm ^g	47.4 (1.7)	8.3 (1.3)	88.3 (0.5)	81.5 (1.1)	143.9 (2.1)	0.04
Ubi K48C	M7PyThiazole	Dy ^g	58.0 (2.0)	24.6 (1.0)	8.5 (0.5)	71.2 (0.8)	178.2 (2.5)	0.04
Ubi K48C	M7PyThiazole	Tb ^g	36.0 (1.3)	21.8 (0.9)	7.1 (0.4)	70.7 (1.1)	5.0 (1.1)	0.07
hCA-II S166C	M7FPy	Tm ^h	-18.4 (0.6)	-4.4 (0.8)	73.5 (2.1)	128.2 (1.8)	120.6 (1.2)	0.13
hCA-II S166C	M7PyThiazole	Dy ⁱ	44.0 (1.0)	24.7 (1.5)	163.1 (5.0)	51.7 (1.6)	98.3 (7.3)	0.07

^a All tensor parameters (errors) are represented in the universal tensor representation (UTR). Tensor parameters (errors) were calculated using the program Numbat²¹ and the metal position was simultaneously fitted for all different lanthanide ions. ^b Only amino acids residing in a stable secondary structure were used (2-7, 12-16, 23-34, 38-40, 41-45, 48-49, 56-59 and 66-71). Metal coordinates were given with respect to the ubiquitin X-ray structure (PDB code: 1UBI). ^c Selectively ¹⁵N leucine labelled human carbonic anhydrase (hCA) II S166C C206S. Metal coordinates were given with respect to the hCA-II X-ray structure (PDB code: 3KS3). ^d Q factors are calculated as root-mean-square deviation between measured and predicted PCS divided by the root-mean-square of the measured PCS. ^e Metal position / Å: x = 20.7 (0.5), y = 11.4 (0.2), z = 12.7 (0.5). ^f Metal position / Å: x = 20.5 (0.2), y = 15.1 (0.1), z = 6.3 (0.2). ^g Metal position / Å: x = 19.0 (0.2), y = 18.6 (0.2), z = 22.8 (0.1). ^h Metal position / Å: x = -16.7 (0.3), y = -0.7 (0.7), z = -9.8 (0.3). ⁱ Metal position / Å: x = -16.7 (0.3), y = -0.7 (0.7), z = -9.8 (0.3).

DISCUSSION

Conjugation of lanthanide chelating tags to proteins via a reductively stable covalent bond is one key prerequisite for any *in-cell* NMR experiments. We have recently shown⁵ that lanthanide complexes of M7PySO₂Ph-DOTA fulfil this condition and were suitable for the determination of the structure of GB1 inside intact eukaryotic cells. More general applications of this tag are, however, hampered by the sluggish reac-

tion of the phenyl-sulfone leaving group of this tag with cysteine side chain thiol at ambient temperature and neutral pH value.

In a first approach the properties of the leaving group were altered by the introduction of one CF₃-group in the *para* position (tag **4**) or by two CF₃-groups in *meta* position (tag **5**) but in both cases no significant acceleration of the tagging speed was observed. This finding indicates that the rate determining step is not the cleavage of the pyridine – sulfone bond but

rather the formation of the Meisenheimer-type complex, as has been postulated before.²³⁻²⁴ As a consequence, the synthetically more demanding introduction of a fluorine substituent in the 5-position of the pyridine ring was performed and the resulting tag **6** (c.f. Schemes 1 and S3) was obtained in good yields. To our delight, the new tag turned out to have much superior reactivity and conjugation to 100 μ M ubiquitin S57C was achieved at r.t. and neutral pH to greater than 90% within six hours (see Figure 1), thus corroborating the hypothesis on the rate determining step. Unexpectedly, we noticed that **6** slowly reacts also with the phosphorous donor of TCEP, whereas no reaction with nitrogen donors, neither in organic solvents nor in aqueous solution was observed. Performing the tagging reaction without the reducing agent TCEP revealed a further complication of Ln-**6** tags, which is, however, to be expected for all tags that use phenyl-sulfonyl leaving groups: upon conjugation and also upon reaction with TCEP, the liberated benzene sulfinic acid is capable of catalyzing the disulfide formation between two cysteine side chains, thus leading to an oxidized ubiquitin dimer. Obviously, this reaction is in competition with the aqueous hydrolysis of benzene sulfinic acid to form phenol and SO₂ and was, therefore, not noticeable in the very slowly reacting parent tag M7PySO₂Ph-DOTA. Careful monitoring of the concentrations and ratios of tag, TCEP and protein allowed to achieve product formation >95% within 16 hours at ambient temperature and neutral pH. Efforts to replace the phenyl on the sulfone with either an amino or a methyl group that should undergo faster hydrolysis of the leaving group compared to benzene sulfinic acid, yielded the corresponding compounds (c.f. SI, compounds **S9** and **S12**) but here again the conjugation reaction speed was compromised, rendering these derivatives unsuitable as LCTs.

In order to further improve our tags, we investigated a pyridine thiazole moiety as the mercaptan reactive site. Benzene thiazole methyl sulfone has been described as moderately reactive towards cysteine (45% conversion within 5 min at 20 mM thiol concentration and 1.2 fold excess of sulfone)²⁵ but no data are available for the pyridine analogues. According to the results we obtained in the first section on the fluorinated pyridine phenyl sulfones, we expected the pyridine thiazole to better stabilize the negative charge of the Meisenheimer intermediate during the nucleophilic attack of the thiol. Lanthanide coordination to the pyridine nitrogen donor should further enhance the reactivity. Pyridine thiazoles have found application²⁶ in medicinal chemistry²⁷ (e.g. as ulcer inhibitors²⁸) and as agrochemicals²⁹⁻³⁰ like fungicides and pesticides, but to the best of our knowledge, have never been described as suitable for protein conjugation via thiol nucleophilic aromatic substitution. The observed acceleration of the tagging speed for ubiquitin S57C by more than an order of magnitude compared to the M7FPy-DOTA tags **11** and **12** (Figure 1) is, therefore, an important finding. In order to explore the limits of our new tag we performed the tagging reaction also at 4°C, at pH 6.0 and 5.0, and reduced the thiol concentration to 10 μ M as shown in Table 1. In order to be able to monitor the reaction more easily by HPLC-ESI/MS, we analyzed the reactivity towards the tetrapeptide Leu-Cys-Asp-Trp instead of ubiquitin for the experiments shown in Table 1. It is clearly demonstrated that even for less favorable conditions, the tagging reaction proceeds relatively fast. The typical setup for sensitive proteins (4°C, 100 μ M, pH 7.0) still shows quantitative conversion within 30 min. These very promising results fulfil all the necessary requirements like thiol-selective, fast, irreversible and virtually quantitative tagging. The fact, that there is only one rotatable bond and no consid-

erably deformable dihedral angle between the cysteine sulfur and the lanthanide, increased our expectations on the performance as PCS tag. The anisotropy parameters of the new tag **15** were determined for two single cysteine mutants of ubiquitin (S57C and K48C) as well as for the more demanding S166C mutant of the 30 kDa protein human carbonic anhydrase type II (hCA II). To facilitate the assignment on hCA II³¹, we used a selectively ¹⁵N-Leu labelled construct, which allowed for fast assignment of the 26 Leu residues in the paramagnetically shifted spectra. Indeed, we found that the anisotropy parameters for tag **15** are unusually large, in the case of thulium they are between two and three times the values found for the Tm complex of **6** (c.f. Table 2). For the dysprosium complex of **15** tagged to ubi-K48C, $\Delta\chi_{\text{ax}}$ of $58.0 (2.0) \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}}$ of $24.6 (1.0) \times 10^{-32} \text{ m}^3$ were determined, which is amongst the largest values ever reported for LCTs. The PCS observed for both new tags **6** and **15** are up to 10 ppm for the Tm complexes in ubiquitin, whereas for the Dy derivatives residues that are close to the metal center tend to be bleached beyond detection by paramagnetic relaxation enhancement. For all protein constructs investigated, virtually every single spin experiences significant PCS (c.f. Figures 1 and S4 to S7, Tables S3 and S5), that give excellent agreement when fitted to the X-ray crystal structures (Figure S1). Using the tensor parameters for Dy **15** tagged to ubi-K48C, calculations yield a PCS larger than 0.03 ppm for a residue in 100 Å distance and 0.005 ppm in 180 Å if it is in a favorable orientation. This will allow for the monitoring of relevant biological interactions even in considerable distance to the tagging site. It is interesting to note, that the relative orientation of the tensor to the protein is more favorable to cause large PCS in the case of pyridine thiazole tag **15** (c.f. Figure S9) which can be explained by the almost orthogonal orientation of the C-S bond

vector in **15** compared to **6** (Figure S16) whereas the attachment points are almost equally distant from the metal center.

The RDC for ubiquitin K48C and S57C caused by the Dy tag of **15** at a field of 14.1 T (600 MHz) exceed 30 Hz (see Table S4) and are in excellent agreement with the NH vector orientations reported in the X-ray structure (see Figures S2 and S3). The size of the RDC alignment tensors (Table S6) is within error margins identical to the tensor parameters obtained from PCS, for the ubiquitin K48C and the hCA II construct, which indicates that the tag undergoes only very subtle motions relative to the protein. This is further corroborated by an analysis of the probability distribution in space of the metal according to the method of Suturina and Kuprov³² (see Figure S10). In contrast, the tagging site in ubiquitin S57C seems to allow for slightly larger pendent motions. In addition, it is noteworthy, that for all tags described in this work only one set of signals was found for the tags themselves (¹H-NMR spectra see Figures S10 to S13) as well as for all HSQC spectra recorded. This is in marked contrast to the presence of a second conformation for the medium sized lanthanides like Dy with the parent DOTA-M8 tag.^{3,33} While early lanthanides adopt SAP geometry in DOTA-M8 complexes, the late lanthanides prefer TSAP conformation.³³ We assume for all lanthanide complexes reported here a $\Lambda(\delta\delta\delta\delta)$ or SAP geometry, as was confirmed by the analysis of the NOESY spectra for the diamagnetic compounds as well as a DFT calculation (see Figure S15).

A further interesting property of the pyridine thiazole tag is the pronounced photoluminescence upon irradiation at 300 nm, that was observed for the Tb complex **19** (Figure S17) which makes this tag a potential donor for FRET applications and is currently studied in more detail in our group.

In conclusion, we have synthesized two new high affinity LCT that show highly selective, fast, quantitative and irreversible tagging towards cysteines. The attachment is unusually rigid and evokes very large PCS and RDC and in addition yields only the SAP stereoisomer for all lanthanides studied. The new pyridine thiazole methyl sulfone activator is particularly suitable for protein tagging at moderate protein concentrations and may be an interesting approach for a wide range of protein conjugation applications.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Synthetic procedures, methods and materials, supporting figures and tables, complete lists of PCS and RDC. (44 pages, PDF)

AUTHOR INFORMATION

Corresponding Author

* E-mail: daniel.haessinger@unibas.ch

ORCID

Daniel Häussinger: 0000-0002-4798-0072

Thomas Müntener: 0000-0001-5508-1117

Present Address

†(J.K.) School of Engineering and Architecture of Fribourg,
Boulevard de Pérolles 80, 1705 Fribourg, Switzerland.

Funding Sources

This work was supported by the Fondation Claude et Giuliana, Vaduz, Liechtenstein and the Department of Chemistry, University of Basel, Switzerland.

ACKNOWLEDGMENT

We thank C. E. Housecroft and E. C. Constable for helpful discussions and R. A. Byrd for providing a chemical.

ABBREVIATIONS

DOTA, 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid; DTT, dithiothreitol; FRET, Förster resonance energy transfer; GB1, B1 domain of the protein G; hCA II, human carbonic anhydrase type II; LCT, lanthanide chelating tag; PCS, pseudocontact shift; RDC, residual dipolar coupling; SAP, square antiprism; TCEP, tris(2-carboxyethyl)phosphine; TSAP, twisted square antiprism;

REFERENCES

1. Hass, M. A.; Ubbink, M., Structure determination of protein-protein complexes with long-range anisotropic paramagnetic NMR restraints. *Curr Opin Struct Biol* **2014**, *24*, 45-53.
2. Keizers, P. H. J.; Saragliadis, A.; Hiruma, Y.; Overhand, M.; Ubbink, M., Design, synthesis, and evaluation of a lanthanide chelating protein probe: CLaNP-5 yields predictable paramagnetic effects independent of environment. *J. Am. Chem. Soc.* **2008**, *130* (44), 14802-14812.

3. Häussinger, D.; Huang, J. R.; Grzesiek, S., DOTA-M8: An extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. *J. Am. Chem. Soc.* **2009**, *131* (41), 14761-14767.
4. Graham, B.; Loh, C. T.; Swarbrick, J. D.; Ung, P.; Shin, J.; Yagi, H.; Jia, X.; Chhabra, S.; Barlow, N.; Pintacuda, G.; et al., DOTA-amide lanthanide tag for reliable generation of pseudocontact shifts in protein NMR spectra. *Bioconjugate Chem.* **2011**, *22* (10), 2118-2125.
5. Müntener, T.; Häussinger, D.; Selenko, P.; Theillet, F. X., In-Cell Protein Structures from 2D NMR Experiments. *J. Phys. Chem. Lett.* **2016**, *7* (14), 2821-2825.
6. Yang, F.; Wang, X.; Pan, B. B.; Su, X. C., Single-armed phenylsulfonated pyridine derivative of DOTA is rigid and stable paramagnetic tag in protein analysis. *Chem Commun (Camb)* **2016**, *52* (77), 11535-8.
7. Nitsche, C.; Otting, G., Pseudocontact shifts in biomolecular NMR using paramagnetic metal tags. *Prog Nucl Magn Reson Spectrosc* **2017**, *98-99*, 20-49.
8. Liu, W. M.; Overhand, M.; Ubbink, M., The application of paramagnetic lanthanoid ions in NMR spectroscopy on proteins. *Coord. Chem. Rev.* **2014**, *273-274*, 2-12.
9. Ikegami, T.; Verdier, L.; Sakhaii, P.; Grimme, S.; Pescatore, B.; Saxena, K.; Fiebig, K. M.; Griesinger, C., Novel techniques for weak alignment of proteins in solution using chemical tags coordinating lanthanide ions. *J. Biomol. NMR* **2004**, *29* (3), 339-349.
10. Pintacuda, G.; Moshref, A.; Leonchiks, A.; Sharipo, A.; Otting, G., Site-specific labelling with a metal chelator for protein-structure refinement. *J. Biomol. NMR* **2004**, *29* (3), 351-361.
11. Vlasie, M. D.; Comuzzi, C.; van den Nieuwendijk, A. M.; Prudêncio, M.; Overhand, M.; Ubbink, M., Long-range-distance NMR effects in a protein labeled with a lanthanide-DOTA chelate. *Chemistry* **2007**, *13* (6), 1715-1723.
12. Lee, M. D.; Loh, C. T.; Shin, J.; Chhabra, S.; Dennis, M. L.; Otting, G.; Swarbrick, J. D.; Graham, B., Compact, hydrophilic, lanthanide-binding tags for paramagnetic NMR spectroscopy. *Chem. Sci.* **2015**, *6* (4), 2614-2624.
13. Hikone, Y.; Hirai, G.; Mishima, M.; Inomata, K.; Ikeya, T.; Arai, S.; Shirakawa, M.; Sodeoka, M.; Ito, Y., A new carbamidemethyl-linked lanthanoid chelating tag for PCS NMR spectroscopy of proteins in living HeLa cells. *J. Biomol. NMR* **2016**, 1-12.
14. Liu, W. M.; Keizers, P. H.; Hass, M. A.; Blok, A.; Timmer, M.; Sarris, A. J.; Overhand, M.; Ubbink, M., A pH-sensitive, colorful, lanthanide-chelating paramagnetic NMR probe. *J Am Chem Soc* **2012**, *134* (41), 17306-13.
15. Liu, W. M.; Skinner, S. P.; Timmer, M.; Blok, A.; Hass, M. A.; Filippov, D. V.; Overhand, M.; Ubbink, M., A two-armed lanthanoid-chelating paramagnetic NMR probe linked to proteins via thioether linkages. *Chemistry* **2014**, *20* (21), 6256-6258.
16. Lee, M. D.; Dennis, M. L.; Swarbrick, J. D.; Graham, B., Enantiomeric two-armed lanthanide-binding tags for complementary effects in paramagnetic NMR

- spectroscopy. *Chem Commun (Camb)* **2016**, 52 (51), 7954-7.
17. Loh, C. T.; Ozawa, K.; Tuck, K. L.; Barlow, N.; Huber, T.; Otting, G.; Graham, B., Lanthanide tags for site-specific ligation to an unnatural amino acid and generation of pseudocontact shifts in proteins. *Bioconjugate Chem.* **2013**, 24 (2), 260-268.
18. Loh, C. T.; Graham, B.; Abdelkader, E. H.; Tuck, K. L.; Otting, G., Generation of pseudocontact shifts in proteins with lanthanides using small "clickable" nitrilotriacetic acid and iminodiacetic acid tags. *Chemistry* **2015**, 21 (13), 5084-5092.
19. Yang, Y.; Li, Q. F.; Cao, C.; Huang, F.; Su, X. C., Site-specific labeling of proteins with a chemically stable, high-affinity tag for protein study. *Chem. Eur. J.* **2013**, 19 (3), 1097-1103.
20. Pan, B. B.; Yang, F.; Ye, Y.; Wu, Q.; Li, C.; Huber, T.; Su, X. C., 3D structure determination of a protein in living cells using paramagnetic NMR spectroscopy. *Chem Commun (Camb)* **2016**, 52 (67), 10237-40.
21. Schmitz, C.; Stanton-Cook, M. J.; Su, X. C.; Otting, G.; Huber, T., Numbat: an interactive software tool for fitting Deltachi-tensors to molecular coordinates using pseudocontact shifts. *J Biomol NMR* **2008**, 41 (3), 179-89.
22. Rinaldelli, M.; Carlon, A.; Ravera, E.; Parigi, G.; Luchinat, C., FANTEN: a new web-based interface for the analysis of magnetic anisotropy-induced NMR data. *J Biomol NMR* **2015**, 61 (1), 21-34.
23. Bunnett, J. F.; Garbisch, E. W.; Pruitt, K. M., The "Element Effect" as a Criterion of Mechanism in Activated Aromatic Nucleophilic Substitution Reactions^{1,2}. *J. Am. Chem. Soc.* **1957**, 79 (2), 385-391.
24. Gandler, J. R.; Setiarahardjo, I. U.; Tufon, C.; Chen, C., Nucleophilic reactivity: nucleophilic aromatic substitution reactions of 2,4-dinitrochlorobenzene and picryl chloride in aqueous and methanol solutions. *The Journal of Organic Chemistry* **1992**, 57 (15), 4169-4173.
25. Toda, N.; Asano, S.; Barbas, C. F., 3rd, Rapid, stable, chemoselective labeling of thiols with Julia-Kocienski-like reagents: a serum-stable alternative to maleimide-based protein conjugation. *Angew Chem Int Ed Engl* **2013**, 52 (48), 12592-6.
26. Bourdais, J.; Abenheim, D.; Sabourault, B.; Lorre, A., Polycyclic thiazoles. I. Anionic nucleophilic substitution of 2-methylsulfonyl derivatives of thiazolo[5,4-b]pyridine and benzothiazole. *J. Heterocycl. Chem.* **1976**, 13 (3), 491-6.
27. Crawford, J. J.; Dossetter, A. G.; Finlayson, J. E.; Heron, N. M. Cyanocyclopropylcarboxamides as cathepsin inhibitors and their preparation and use in the treatment of diseases. WO2009001127A1, 2008.
28. Katano, K.; Tomomoto, T.; Ogino, H.; Yamazaki, N.; Hirano, F.; Yuda, Y.; Konno, F.; Nishio, M.; Machinami, T.; et, a. Preparation of thiazolopyridines as ulcer inhibitors. EP405976A1, 1991.
29. Wagner, O.; Wetterich, F.; Eicken, K.; Rack, M.; Hamprecht, G.; Lamm, G.; Speakman, J.-B.; Lorenz, G.; Ammermann, E.; Strathmann, S. Preparation of

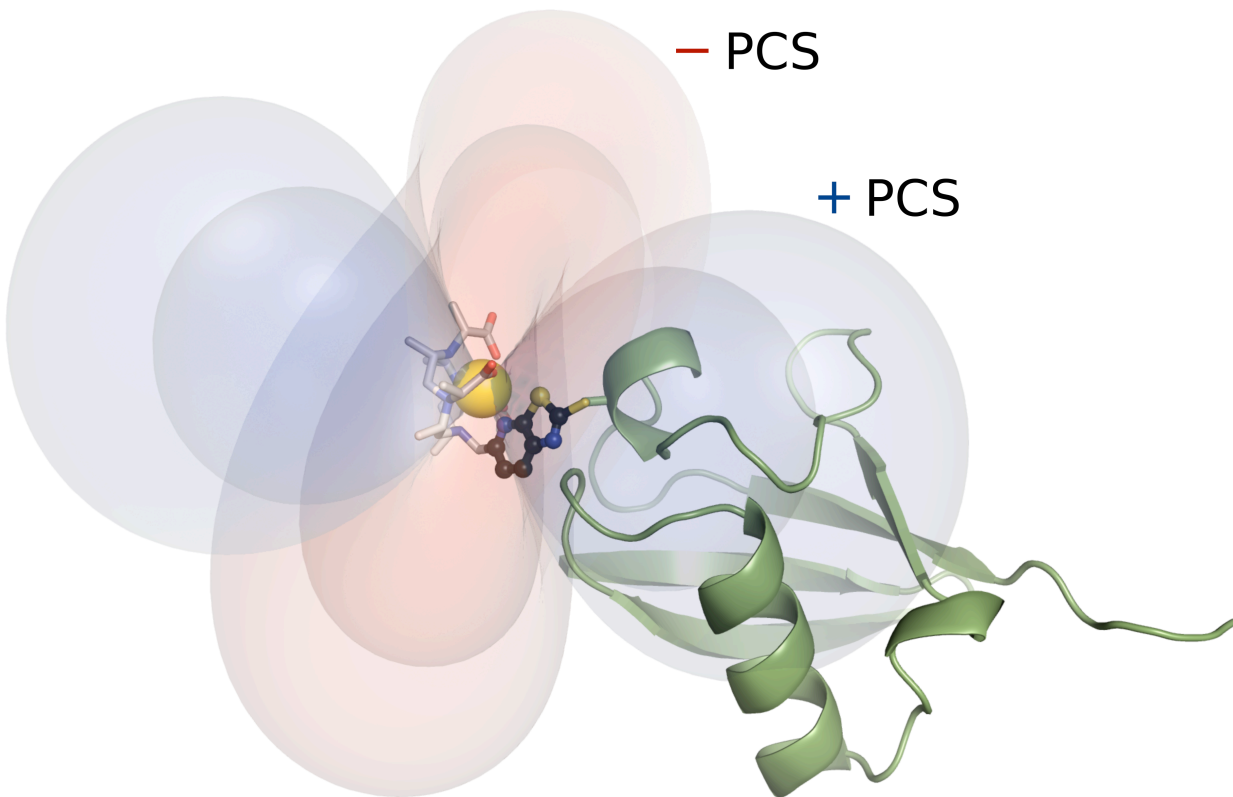
heterocyclisulfonylalkyl benzoates and analogs as agrochemical fungicides. WO9708147A1, 1997.

30. Gayral, P.; Bourdais, J.; Lorre, A.; Abenhaim, D.; Dusset, F.; Pommies, M.; Fouret, G., Synthesis and study on antiparasitic and molluscicidal activities of polycyclic pyridine derivatives. I. Comparison of 2-substituted derivatives of thiazolo[5,4-b]pyridine and benzothiazole. *Eur. J. Med. Chem. - Chim. Ther.* **1978**, *13* (2), 171-5.

31. Varghese, S.; Halling, P. J.; Haussinger, D.; Wimperis, S., Two-dimensional (1)H and (1)H-detected NMR study of a heterogeneous biocatalyst using fast MAS at high magnetic fields. *Solid State Nucl Magn Reson* **2018**, *92*, 7-11.

32. Suturina, E. A.; Kuprov, I., Pseudocontact shifts from mobile spin labels. *Phys Chem Chem Phys* **2016**, *18* (38), 26412-26422.

33. Opina, A. C. L.; Strickland, M.; Lee, Y. S.; Tjandra, N.; Byrd, R. A.; Swenson, R. E.; Vasalatiy, O., Analysis of the isomer ratios of polymethylated-DOTA complexes and the implications on protein structural studies. *Dalton T* **2016**, *45* (11), 4673-4687.



Tm-M7PyThiazole-DOTA — Ubi S57C