

Application of Paramagnetic Lanthanoid Chelating Tags in NMR Spectroscopy and Their Use for the Localization of Ligands Within Biomacromolecules

D Joss, R Vogel, K Zimmermann, and D Häussinger, University of Basel, Basel, Switzerland

© 2020 Elsevier Inc. All rights reserved.

1	Introduction	1
1.1	Paramagnetic Nuclear Magnetic Resonance Spectroscopy	1
2	Lanthanoid Chelating Tags	5
2.1	Introduction and Overview	5
2.2	EDTA-Based Lanthanoid Chelating Tags	5
2.3	DTTA- and DTPA-Based Lanthanoid Chelating Tags	6
2.4	TAHA-Based Lanthanoid Chelating Tags	7
2.5	IDA-, NTA- and Pyridine-Based Lanthanoid Chelating Tags	7
2.6	DOTA-Based Lanthanoid Chelating Tags	8
3	Applications of Lanthanoid Chelating Tags in the Localization of Ligands Within Biomacromolecules	12
3.1	Introduction	12
3.2	Weakly Binding Ligands	12
3.3	Ligands in the Intermediate Exchange Regime	13
3.4	Strongly Binding Ligands	14
4	Localization of Ligands Within Human Carbonic Anhydrase II Using ¹⁹F Pseudocontact Shift Analysis	15
4.1	Human Carbonic Anhydrase and Its Use in the Development of Artificial Metalloenzymes	15
4.2	Localization of Ligands within Human Carbonic Anhydrase II—Introduction and Concept	16
4.3	PCS Analysis and Determination of the Anisotropy Parameters from Multiple Positions Within the Protein	18
4.4	Measurement of ¹⁹ F PCS and Localization of the Ligand Using a Monte-Carlo Approach	19
4.5	Determination of the Ligand Position by Analysis of the Isosurfaces of the Anisotropy Parameters	24
4.6	¹ H- ¹⁹ F-HOESY Experiments for Determination of ¹ H PCS of the Ligand and Protein–Ligand Interactions	24
5	Conclusion and Outlook	26
	Acknowledgment	26
	References	26

1 Introduction

1.1 Paramagnetic Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is ideally suited for the elucidation of structure, dynamics and interactions of biomacromolecules in solution.¹ The development of homo- and heteronuclear experiments in order to establish through-bond^{2–5} and through-space correlations^{6,7} allow for the complete assignment of the protein by experiments linking the amide proton with the backbone and side chain spins from the obtained through-bond correlations. The three-dimensional structure of the protein is then computed using the structural restraints obtained (Fig. 1).^{6,7,10,11} The possibility to tune and adjust the sample conditions, e.g. buffer composition, salt concentration, temperature, additives and pH, allows for realistic experiments that can be adjusted to specific physiological conditions.¹ Furthermore, the dynamics of the protein of interest can be investigated by analysis of relaxation times and heteronuclear Overhauser effects (hetNOEs).^{12,13} Besides the convenient application of advanced NMR techniques to proteins, assignment techniques and further experiments were developed for nucleic acids in order to elucidate their structure, dynamics and protein–nucleic acid complexes.^{14–17} However, the distance limitation of the exploited nuclear Overhauser effect (NOE) of 5–8 Å for structure calculations and unspecific chemical shift perturbation (CSP) experiments, used for the monitoring of interactions of biomacromolecules and ligand-binding, render more advanced methods that yield structural restraints over larger distances highly desirable. Paramagnetic NMR can contribute in this regard by providing long-range, through-space interactions in order to obtain distance and orientational restraints of biomacromolecules and their complexes with other biomacromolecules and ligand molecules.¹⁸

An impressive and early study described the quantitative determination of the conformations of mononucleotides in solution using lanthanoid induced shifts,¹⁹ and in the following, Bertini, Banci, Luchinat et al. exploited the paramagnetism of lanthanoids by substitution of diamagnetic metal ions in metalloproteins and demonstrated thereby thoroughly the application and use of paramagnetic NMR in structural biology.^{20–23} In order to enable the introduction of lanthanoid centers into proteins without native metal site in order to take advantage of the paramagnetism-based structural restraints, lanthanoid binding peptides (LBPs) that imitate the zinc finger and EF hand motif of native metalloproteins were developed^{24–27} and can be added to the N- or C-terminus of the protein of interest. In order to provide more rigid chelators and attach an LCT at any possible position within the scaffold of the biomacromolecule, a large variety of double- and single-armed lanthanoid chelating tags (LCTs) has been developed.^{28–37} LCTs have since then found widespread applications in structural biology and biomolecular NMR, e.g. structural characterization of

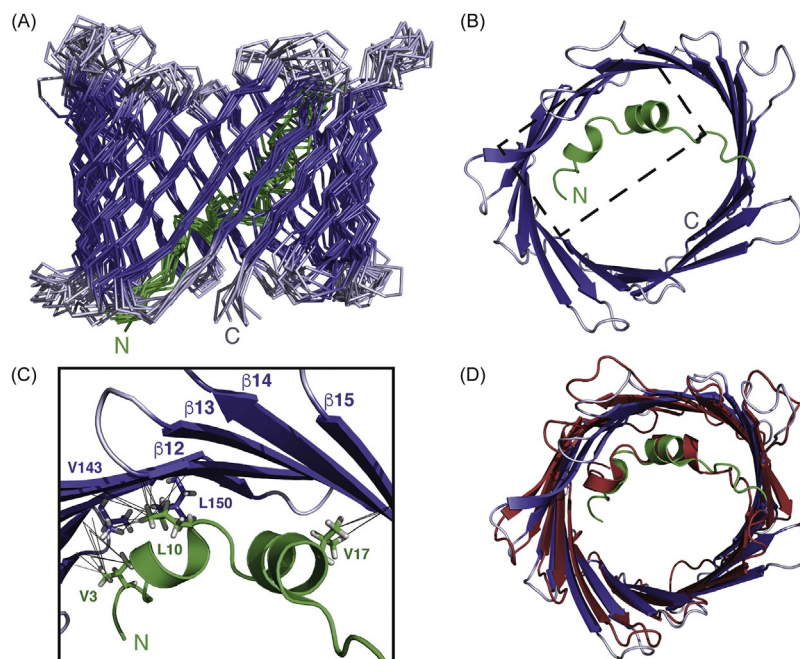


Fig. 1 Solution NMR Structure of hVDAC-1 in LDAO Detergent Micelles refined using NMR restraints. Backbone representation of the 10 conformers with lowest violation of dihedral angles and NOE restraints (A, N-terminus: green, loops: light blue, β -strands: blue). Refined conformer closest to the mean (B, top view). NOE contacts between the N-terminal helices and the wall of the membrane protein (C, NOE contacts: black, N-terminus: green, loops: light blue, β -strands: blue). Overlay of the refined solution NMR structure of hVDAC-1 with the crystal structure of mVDAC-1 (D, PDB 3EMN⁸).⁹ NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture). Wüthrich, K.; *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 3340–3363.

antibiotic drug targets (Fig. 2),³⁹ localization of ligands within proteins,^{40,41} observation of long-range RNA structural information using a paramagnetically tagged reporter protein,⁴² enhancement of signal dispersion in studies of intrinsically disordered proteins,⁴³ structural analysis of biologically highly relevant, large complexes of biomacromolecules⁴⁴ and characterization of the interactions between glycans and proteins.⁴⁵

The introduction of paramagnetic centers in biomacromolecules by either direct complexation by metalloproteins, modification of the protein construct with an LBP or attachment of a small-molecule chelator with appropriate linker moiety leads to sizeable effects visible in NMR experiments that were thoroughly and critically reviewed in multiple reviews.^{18,46–64} The most important effects associated with induced anisotropy caused by paramagnetic centers comprise pseudocontact shifts (PCS),^{28,30,33,34,36,37,65–72} residual dipolar couplings (RDC)^{73–77} and paramagnetic relaxation enhancement (PRE).^{78–84} PCSs and RDCs are exemplified in Fig. 3 by depicting a concrete example of an in-cell NMR study, that delivered *de-novo* derived protein structures from 2D NMR experiments. Besides the most important paramagnetic effects as PCS, RDC and PRE, also contact shifts (CS),^{85–88} residual anisotropic chemical shifts (RACS),^{89,90} dipolar shift anisotropy (DSA) as well as cross-correlation between DSA and dipolar relaxation (DSA/DD)⁹¹ are associated with the incorporation of a paramagnetic center in a target of interest.

Most importantly, the shifts caused by through-space PCS effects provide angular and distance information about the nucleus. Upon application of PCS NMR on proteins, the sequential assignment of backbone amides can be combined with PCS shift data and, therefore, be used in order to obtain structural restraints for each backbone N and H spin. Due to their distance dependence of

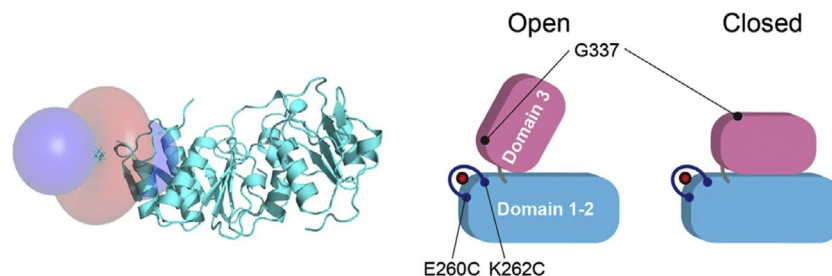


Fig. 2 Caged lanthanoid NMR probe (CLaNP-5)²⁸ attached to a double cysteine mutant of MurD for monitoring of conformational changes. Induced isosurfaces of the Yb³⁺-loaded LCT when ligated to domains 1 and 2 (PCS \pm 0.5 ppm, blue: positive, red: negative, PDB: 3UAG,³⁸ left). Monitoring of the open and closed conformation of domain 3 using PCS restraints (right).³⁹ Ligand-driven conformational changes of MurD visualized by paramagnetic NMR. Saio, T.; Ogura, K.; Kumeta, H.; Kobashigawa, Y.; Shimizu, K.; Yokochi, M.; Kodama, K.; Yamaguchi, H.; Tsujishita, H.; Inagaki, F. *Sci. Rep.* **2015**, *5*, 16685. Nature Publishing Group.

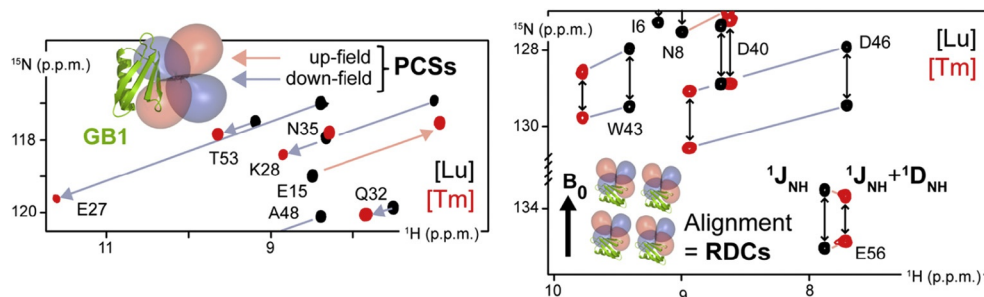


Fig. 3 PCS induced by a Tm-DOTA-M7Py in comparison with its diamagnetic reference Lu-DOTA-M7Py observed on a GB1 E19C construct in ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) experiments (PCS \pm 2.0 ppm, blue: positive, red: negative, *left*). Partial alignment of the labeled protein to the external magnetic field B_0 leading to RDCs observed in ^1H - ^{15}N HSQC IPAP experiments.³³ Reproduced with permission from Müntener, T.; Häussinger, D.; Selenko, P.; Theillet, F.-X. *J. Phys. Chem. Lett.* **2016**, *7*, 2821–2825. Copyright 2016 American Chemical Society.

R^{-3} , PCSs can yield structural restraints over distances of more than 200 Å in favorable cases.¹⁸ In contrast to the through space nature of PCSs, CSs are through-bond effects and only transferred over a few chemical bonds and are, therefore, less interesting for the analysis of biomacromolecules. However, CSs can give valuable insights into ligand structure and conformation, since the observed shifts are dependent on the structure and bonding interactions.^{92,93}

RDCs are especially well suited for structural analysis of large biomacromolecules, since the effect is independent of the distance between the paramagnetic center and the measured nucleus and yields the orientation of the N—H bond vectors in space as structural restraints. RDCs can either be induced by partial alignment of the biomacromolecules by incorporation of a paramagnetic lanthanoid ion with anisotropic electron distribution or by application of an alignment medium as bacteriophages, acrylamide gels or bicelles.^{73,77,94} The RDCs induced by incorporation of paramagnetic centers can be diminished due to internal flexibility of the biomacromolecule, however, this effect can then be further exploited to determine, for example, the stiffness of an α -helix and to draw conclusions about the biological significance of the gained insight.⁷⁷ Combination of LCT-induced PCSs and RDCs was used to determine backbone structures of biomacromolecules inside intact eukaryotic cells, as a unique property of covalently bound high affinity LCTs is their compatibility with *in-cellulo* experiments.^{33,69,95}

PRE is, in a first approximation, an isotropic effect and thus yields only distance and no angular information. Furthermore, the magnitude of the effect decreases with R^{-6} with respect to the distance between the observed spin and the paramagnetic center. However, PREs are not very susceptible to motional averaging, are convenient to induce by commercially available spin labels and are more straightforward to analyze than PCSs and RDCs. Nonetheless, since PCSs provide both distance and angular information, exhibit a distance dependence of R^{-3} and can be observed on any NMR active nuclei, e.g. ^1H , ^{13}C , ^{15}N , ^{19}F , or ^{31}P by one- or multidimensional NMR experiments, PCSs are perfectly well suited for the localization of ligands within biomacromolecules.

The PCS is a dipolar, through-space effect arising from the dipolar interaction between the nuclear spin and, in the case of LCTs, the orbital momentum of the unpaired f-electrons of the paramagnetic lanthanoid center that manifests in NMR experiments by a shift contribution to the NMR resonance of the detected nucleus, depending on the positioning of the nuclear spin with respect to the metal center.⁹⁶ Since the distance between the spin and the paramagnetic metal center is much larger than the ionic radius of the metal center, the point-dipole approximation can be applied that depicts the magnetic dipole moment as located in a single point.⁶² The PCS can then be translated into structural restraints by fitting of the experimental data to a known or calculated structure using software packages like Numbat⁹⁷ or Paramagpy.⁹⁸ Programs capable of PCS analysis minimize the difference between observed and back-calculated PCS in an iterative fashion until the convergence is achieved. The resulting axial and rhombic parameters of the observed anisotropy tensors should be reported in the unique tensor representation (UTR) using the zyz' convention, as only this representation allows for direct comparison with tensors from other LCTs.⁶² In order to visualize the anisotropy parameters induced by the lanthanoid within the biomacromolecule in a suitable fashion, isosurfaces for given ppm values are displayed graphically (Fig. 4).

Thereby, the strength of the shifts in different regions of a biomacromolecule can be directly and visually inspected, e.g. with the program PyMOL. While red parts by definition indicate regions with negative shifts, blue parts constitute regions with positively shifted NMR signals. By the use of PCS analysis, each N and H nucleus can be localized in the three-dimensional space, and therefore valuable structural information can be gained about the structure of the protein of interest. More specifically, a *de-novo* derived structure can be obtained by starting the refinement process from an *in silico* structure, e.g. by use of the Rosetta database. Structures from X-ray crystallography can be refined in solution by PCS and crucial differences between the solution- and solid-state structures, e.g. a different ordering or conformation of more flexible regions within a protein or the alignment of different protein domains, can immediately be detected due to poor fits of the data set obtained via paramagnetic NMR spectroscopy.^{39,99} For nucleic acids, the methodology can be applied e.g. with total correlation spectroscopy (TOCSY) or nuclear Overhauser effect spectroscopy (NOESY) spectra.⁷² The PCS can be calculated by use of the following equation that can be displayed both in polar or Cartesian coordinates⁹⁶:

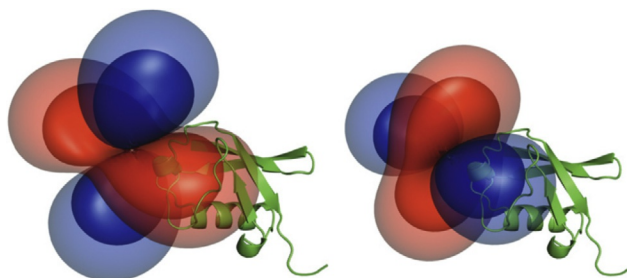


Fig. 4 Isosurfaces induced by the attachment of Ln-P4T-DOTA complexes (Dy^{3+} : left, Tm^{3+} : right) and their relative orientation to ubiquitin S57C (PCS isosurfaces: 1.5 ppm (outer layer), 4.0 ppm (inner layer)).³⁷ Reproduced by permission of The Royal Society of Chemistry from Joss, D.; Häussinger, D. *Chem. Commun.* **2019**, 55, 10543–10546.

$$\text{PCS (polar coordinates)} \delta^{\text{PCS}} = \frac{1}{12\pi r^3} \left[\Delta\chi_{\text{ax}} (3 \cos^2 \theta - 1) + \frac{3}{2} \Delta\chi_{\text{rh}} \sin^2 \theta \cos 2\phi \right] \quad (1)$$

$$\text{PCS (Cartesian coordinates)} \delta^{\text{PCS}}(x, y, z) = \frac{1}{12\pi r^3} \left[\Delta\chi_{\text{ax}} \frac{2z^2 - x^2 - y^2}{r^2} + \frac{3}{2} \Delta\chi_{\text{rh}} \frac{x^2 - y^2}{r^2} \right] \quad (2)$$

For the complete description of the anisotropic part of the magnetic susceptibility that is induced by the lanthanoid ion incorporated in or attached to a biomacromolecule of interest, in total eight parameters are required: axial and rhombic component of the magnetic susceptibility tensor ($\Delta\chi_{\text{ax}}$ and $\Delta\chi_{\text{rh}}$), the coordinates of the paramagnetic center (x, y, z) and the Euler angles describing the orientation with regard to the coordinate frame of the protein (α, β, γ). The anisotropy parameters $\Delta\chi_{\text{ax}}$ and $\Delta\chi_{\text{rh}}$ describe the directional asymmetry of a tensor using the following definitions⁹⁶:

$$\text{Axiality } \Delta\chi_{\text{ax}} = \chi_{zz} - \frac{\chi_{xx} + \chi_{yy}}{2} \quad (3)$$

$$\text{Rhombicity } \Delta\chi_{\text{rh}} = \chi_{xx} - \chi_{yy} \quad (4)$$

As discussed, software packages, e.g. Numbat⁹⁷ or Paramagpy⁹⁸ are used to derive the anisotropy parameters from the experimental data by minimizing the difference between the experimental and back-calculated PCS for a given structure (Fig. 5).

Alternative approaches use a given set of PCS in combination with a known or partially known structure for assignment of methyl groups^{100,101} or the localization of ligand atoms^{40,41,102} within a biomacromolecule of interest.

The aim of this case study is to provide the reader with basic knowledge about LCTs and the application of LCTs for the localization of ligands within biomacromolecules using PCS analysis with a particular focus on the localization of fluorinated ligands within human carbonic anhydrase II (hCA II) using PCS as long-range restraints with a high accuracy over long metal-fluorine distances.

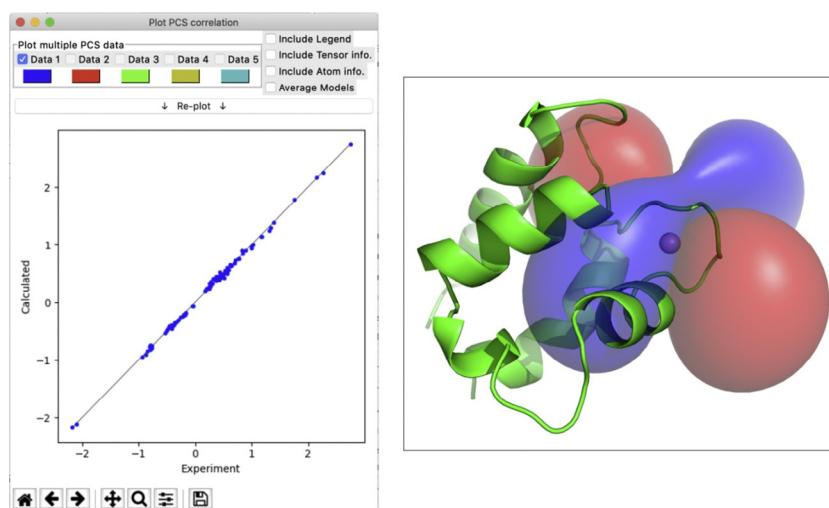


Fig. 5 View on a part of the interface of the novel software package Paramagpy for PCS fitting. Correlation plot of experimental and calculated PCS (left) and isosurfaces of Er^{3+} incorporated in calbindin D9k (PyMOL, right).⁹⁸ Paramagpy: Software for Fitting Magnetic Susceptibility Tensors Using Paramagnetic Effects Measured in NMR Spectra. Orton, H. W.; Huber, T.; Otting, G. *Magn. Reson. Discuss.* **2019**, 1–18. Copernicus publications

2 Lanthanoid Chelating Tags

2.1 Introduction and Overview

In order to exploit the paramagnetic effects induced by lanthanoids in biomacromolecules, the most important and basic consideration is the incorporation of the lanthanoid center into the biological target of interest. As demonstrated by Bertini et al., the most simple and convenient starting point is a protein that possesses a native metal binding site, e.g. calbindin D9k. Due to the similar oxophilicity and ionic radius of the members of the lanthanoid series when compared with Ca^{2+} ions,¹⁰³ paramagnetic lanthanoid ions can often be incorporated in calcium binding sites. Given that the protein exhibits only one calcium binding site or only one calcium binding site that is amenable for substitution with a lanthanoid ion, the induced paramagnetic effects can then be used as structural restraints for the characterization of the protein scaffold.¹⁰⁴ Furthermore, the fixed protein coordination site can also be used for a study of the induced anisotropy parameters by the full series of lanthanoids.^{47,105} However, it is important to note that the displayed anisotropy of lanthanoid centers can depend strongly on the type of donor atoms and the coordination polyhedron of the lanthanoid complex.^{106–110} While the direct substitution of calcium by lanthanoid ions is highly interesting for any protein containing a suitable metal binding site, the incorporation of lanthanoid ions into other proteins has to rely on external tags that are fused to the protein during expression or by attachment to specific residues on the surface of the protein.

LBP mimics the native metal binding site of metalloproteins and can conveniently be inserted into the sequence or fused to the termini of the protein.^{25–27,111,112} In order to enhance the rigidity of the lanthanoid binding site delivered by the LBP with respect to the protein scaffold, two-point anchored LBPs, i.e. an LBP linked to the protein through one peptide fusion and one disulfide bond, present an appealing improvement.²⁷ Although this modification always leads to complete ligation yields for the attachment of the tag to the protein, significant drawbacks are associated with LBPs, i.e. the incorporation of the tag is mostly performed at the termini of the sequence since otherwise a significant distortion of the protein scaffold is possible, the additional residues lead to more crowded protein NMR spectra and the rigidity as well as properties of the chelator cannot be that specifically tuned as e.g. with chelators based on ethylenediaminetetraacetic acid (EDTA) or 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA) scaffolds. To avoid these drawbacks, LBPs that circumvent this severe restriction by anchoring the LBP via cysteine residues to the protein were proposed.^{113,114}

In order to enable the tuning of the induced paramagnetic effects and the incorporation of different linker moieties and ligand donor atoms, the attachment of common metal chelators as EDTA,^{115,116} diethylenetriaminepentaacetic acid (DTPA),¹¹⁷ or DOTA^{28,29} and their site-specific attachment to single cysteine residues proved to be highly successful (Fig. 6). Cysteine residues are specifically well-suited to serve as attachment point of an LCT, since they can be introduced on the surface of the protein by mutation by replacing a serine for a cysteine residue without distortion of the local structure. Most importantly, the nucleophilicity of cysteines, due to the significant fraction of deprotonated cysteine thiols around the physiological pH, renders them very attractive for ligation to an LCT with an electrophilic attachment point. Furthermore, cysteines display a low abundance of only 1.4% in protein sequences (average of proteins contained in the UniProtKB/Swiss-Prot data bank).¹¹⁸

Synthetic LCTs can be designed in virtually any substitution and donor atom pattern¹⁸ and allow a high variety of thiol-reactive moieties to be incorporated, i.e. activated mixed disulfides,^{29,34,36} activators with pyridinesulfone leaving groups,^{33,69,119} pyridine-thiazole moieties,^{35,37} bioorthogonal activators,^{67,120} or two-point anchoring disulfides that strongly restrict the flexibility of the LCT on the surface of the protein.^{28,121–124}

However, a serine to cysteine mutation can be problematic for proteins that show instability upon displaying cysteine residues on their surface or for proteins that display multiple cysteine residues, which are required in order to fulfill their biological function, on their surface. In the future, bio-orthogonal conjugation to artificial amino acids that can be incorporated into the genetic code, are expected to play a key role in the incorporation of LCTs into proteins and paramagnetic NMR spectroscopy of biomacromolecules.

2.2 EDTA-Based Lanthanoid Chelating Tags

In order to evolve the methodology of inducing paramagnetic effects by lanthanoid-loaded peptide motifs further to a more versatile chelator that is independent of the process of protein expression, suitable synthetic chelators were required and auspicious

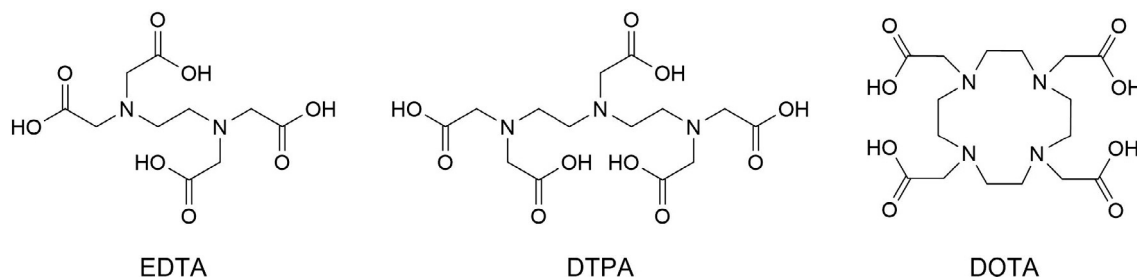


Fig. 6 Scaffolds of chelators used for the synthesis of LCTs.

candidates were EDTA and DTPA scaffolds. The development of LCTs started therefore with the synthesis of EDTA-derived LCTs that were then conjugated to a cysteine residue of a protein in order to observe structural restraints on covalently tagged proteins.

The development of an EDTA-derived LCT and its application for paramagnetic protein NMR was first reported by Dvoretzky et al.¹²⁵ The Yb³⁺-loaded chelator equipped with a pyridine-2-thiol activated linker moiety was attached to barnase H102C and the obtained structural restraints, i.e. PCS and RDC, were fitted to the X-ray structure of wild-type barnase.

Two years later, an enantiomerically pure EDTA-based LCT with a disulfide linker was proposed by Ikegami et al.¹¹⁵ Upon conjugation of the LCT to trigger factor, a molecular chaperone associated with the ribosome, RDCs of up to 8 Hz were observed at 800 MHz proton frequency. However, upon equipping the EDTA-based LCT with a linker, and coordination of the metal that leads to a fixed geometry, a nitrogen atom within the chelator becomes a stereogenic center and two diastereomeric species are formed. The two diastereomeric species lead, when attached to the protein of interest, to a second set of peaks in ¹H-¹⁵N HSQC spectra. A significant improvement was then achieved by Leonov et al., who attached the linker of the LCT to the carbon backbone of the EDTA ligand instead of the nitrogen atom.¹²⁶ Thereby, each of the two nitrogen atoms is substituted by two identical groups and despite the fact, that still two enantiomers of the LCT can be formed, in favorable cases only one set of signals was observed in ¹H-¹⁵N HSQC experiments.

In 2006, two novel EDTA chelators were reported by Haberz et al. that lead to linearly independent anisotropy parameters when compared to previously synthesized EDTA-tags.¹²⁷ Interestingly, the LCTs were applied for the solution structure refinement of trigger factor and the authors demonstrated that the affinity of the newly synthesized LCTs for lanthanoids is higher than the one of the native metal binding site in calmodulin. By the design and synthesis of EDTA-derived LCTs, it was successfully demonstrated that the methodology of the LBPs can be successfully transferred to the use of synthetic metal chelators. The important advantage of the newly synthesized LCTs is that they can be attached at multiple sites within a protein. However, the observed anisotropy parameters were only of modest size and for some of the EDTA-based tags peak doubling was observed that severely complicates assignment of the spectra. Therefore, other metal chelators were tested in order to serve as scaffolds for LCTs, as e.g. diethylene-triaminetetraacetic acid (DTTA) and DTPA ligands.

2.3 DTTA- and DTPA-Based Lanthanoid Chelating Tags

In order to improve the initially developed EDTA tags, DTTA- and DTPA-derived LCTs were synthesized and applied in paramagnetic NMR of biomacromolecules.

As a first example, Prudêncio et al. proposed in 2004 a DTPA-based LCT with two linker moieties that was termed CLaNP and is attached to suitable double-cysteine mutants of the protein of interest.¹¹⁷ However, the authors detected up to five signals per residue in ¹H-¹⁵N HSQC spectra. In Ln-DTPA-bisamide complexes, in theory up to eight different stereoisomers can be formed, although it is likely that the occurrence of some species is promoted by interaction of the chelator with the surface of the protein.^{62,128} Although the occurrence of multiple species was rather discouraging, observable PCSs were detected even on nuclei further away than 40 Å, a result that confirmed that valuable long-range restraints can be obtained by using synthetic paramagnetic metal tags. Upon investigation of DTPA-based LCTs that lead to multiple detected species, DTTA-derived LCTs were envisioned to be synthesized.

Jiang et al. developed bio-orthogonally linked diethylene-triamine-tetraacetate propyl-1-yne (DTTA-C3-yne) and diethylene-triamine-tetraacetate butyl-1-yne (DTTA-C4-yne) LCTs that are attached to the protein of interest by the use of a copper-catalyzed azide alkyne click reaction of an alkyne moiety on the LCT with a *p*-azido-L-phenylalanine (AzF) residue of the protein (Fig. 7).¹²⁰ Thereby, the authors demonstrated that large anisotropy parameters can be obtained and RDCs up to 8 Hz were detected using Tb³⁺ and Tm³⁺ loaded LCTs on ubiquitin and enzyme II B. Based on intermolecular PCS measured for a UBA1-ubiquitin complex, the authors reported a refined model on the interactions between the two proteins.

Two further DTTA-derived LCTs equipped with pyridinesulfone linker moieties were reported in 2017 by Chen et al.¹²⁹ Both the 4PS-PyDTTA and the methyl-substituted analogue 4PS-6 M-PyDTTA were tested on ubiquitin and SrtA constructs and yielded

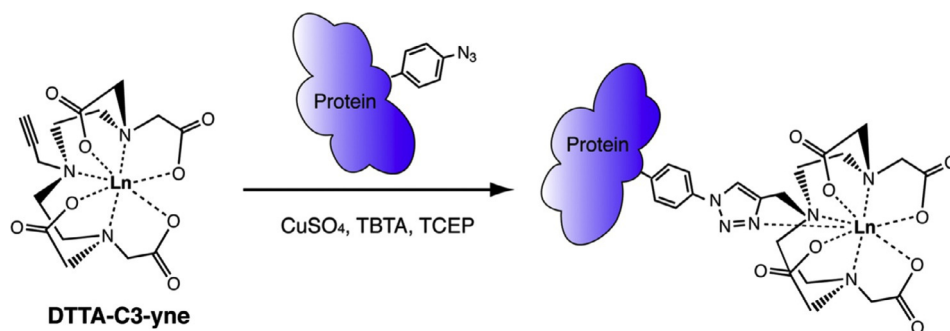


Fig. 7 Copper catalyzed Huisgen azide alkyne cycloaddition of DTTA-C3-yne to a 4-azidophenylalanine residue of the target protein. Remarkably, the triazole ring forms during the ligation reaction a further coordinating nitrogen site and thereby creates a very short linkage from the LCT to the protein of interest.¹²⁰ Lanthanoid tagging via an unnatural amino acid for protein structure characterization. Jiang, W.-X.; Gu, X.-H.; Dong, X.; Tang, C. *J. Biomol. NMR* **2017**, *67*, 273–282. Springer.

large PCSs and RDCs. In line with the results of Haberz et al. for the EDTA-derived LCT,¹²⁷ the authors of the present study observed a high stability of the nonmethylated lanthanoid complex, i.e. the affinity of the LCT to the lanthanoid ion exceeds the one of the native calcium binding site in SrtA. Interestingly, 4PS-6 M-PyDTTA was more susceptible to the addition of EDTA than its nonmethylated analogue.¹²⁹ However, the methyl substitution leads to a significant increase in the anisotropy parameters (4PS-6 M-PyDTTA: $\Delta\chi_{ax} = 11.0 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = 3.5 \times 10^{-32} \text{ m}^3$ vs. 4PS-PyDTTA: $\Delta\chi_{ax} = 7.6 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = 3.0 \times 10^{-32} \text{ m}^3$ for Tm^{3+} on SrtA D82C), which suggests that the translational and rotational averaging of the paramagnetic effects by reorientation of the LCT on the surface of the protein is suppressed.

While the DTPA-derived LCTs can lead in theory to observation of up to eight signal sets for each amide in ^1H - ^{15}N HSQC spectra and in practice up to five different signal sets were reported, the DTTA-based LCTs equipped with bio-orthogonal and pyridine-sulfone linker provide one set of signals and yield sizeable anisotropy parameters. Nonetheless, further improvement of the anisotropy parameters and increase of stability under all buffer conditions remains an interesting target.

2.4 TAHA-Based Lanthanoid Chelating Tags

In order to investigate the use of further chelators for the use as LCT in paramagnetic NMR spectroscopy, Peters et al. demonstrated in 2011 that cysteinyl-phenyl-triaminohexaacetate (Cys-Ph-TAHA) can be used as suitable scaffold for the design and synthesis of an LCT (Fig. 8).³⁰ Advantageously, Cys-Ph-TAHA is highly symmetric, leads therefore to only one observed set of signals in ^1H - ^{15}N HSQC spectra and shows thermal stability over a wide range of temperature (278–315 K). Furthermore, a metal affinity in the femtomolar range is provided and PCSs up to 2 ppm and RDCs up to 18 Hz are induced by the novel LCT. The authors of the study conjugated Cys-Ph-TAHA to a 90 kDa complex of Lac repressor protein, deoxyribonucleic acid (DNA) and isopropyl β -D-1-thiogalactopyranoside (IPTG). A minor disadvantage of the presented LCT is its pH-sensitivity that excludes the use of the tag for studies at pH values below 4.0, e.g. in order to monitor the unfolding process of a protein of interest under acidic conditions.

2.5 IDA-, NTA- and Pyridine-Based Lanthanoid Chelating Tags

Simple, acyclic chelators captivate the interest of researchers due to their cheap and easily accessible precursors. Although it is clear that the induced anisotropy parameters are expected to be smaller than with extremely rigid, sterically crowded, macrocyclic LCTs, simple acyclic chelators can constitute an interesting alternative to more advanced LCTs for applications that involve only small- and medium-sized proteins. Small LCTs are conveniently built from picolinic acid derivatives or iminodiacetic acid (IDA) and nitrilotriacetic acid (NTA)-scaffolds that are closely related to their larger analogues DTTA and DTPA.¹⁸

Su et al. developed a 4-mercaptomethyl-dipicolinic acid (4MMDPA) LCT that yields PCS up to 2 ppm for the Yb^{3+} -complex when conjugated to the N-terminal domain of arginine repressor protein (ArgN).⁶⁵ Upon attachment of the LCT to a cysteine residue in position 68 of ArgN by using Ellman's reagent, the LCT is further restricted in its mobility relative to the protein surface via the coordination of a nearby carboxylate of the protein scaffold to the lanthanoid. Closely related to the LCT presented by Su et al.,⁶⁵ Man et al. reported a 3-mercapto-2,6-pyridinedicarboxylic acid (3MDPA) LCT that bears a shorter linker in the 3-position of the pyridine instead of the 4-position.¹³⁰ Upon conjugation of the LCT to ArgN and T4 lysozyme C54T C97A Q69C, the LCT was shown to yield for such a small-molecule chelator sizeable anisotropy parameters for the Tm^{3+} loaded construct. As in the case of 4MMDPA, the LCT is suspected to interact with a carboxylic acid group of the Glu21 sidechain of the protein. Jia et al. extended the repertoire of dipicolinic acid derived LCTs by the design and synthesis of 4MDPA that includes a short linker in the 4-position of the pyridine scaffold.¹³¹ While the tag offers quantitative tagging yields, the obtained anisotropy parameters still offer the potential of significant advances. In order to enable the measurement of PCS and RDC effect under reductive conditions as e.g. found while performing in-cell NMR studies, it would be highly desirable to develop a reduction-stable linker moiety based on the dipicolinic acid LCTs.

Li et al. developed therefore in 2011 a dipicolinic acid-based LCT, which bears a vinylpyridine moiety that reacts in a thiol-ene reaction with the cysteine thiol of the protein, and was demonstrated to be applicable on both ArgN as well as ubiquitin (Fig. 9).¹³² The protein-LCT conjugate was shown to be stable both in dithiothreitol (DTT) and 3,3',3''-phosphanetriyltripropanoic acid (TCEP) due to the reductively stable thioether that is formed during the tagging reaction. The LCT induces sizeable effects both on ArgN and ubiquitin constructs.

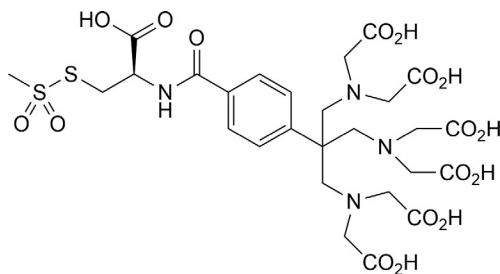


Fig. 8 Chemical structure of the Cys-Ph-TAHA LCT³⁰ synthesized in a total yield of 28% over seven steps.

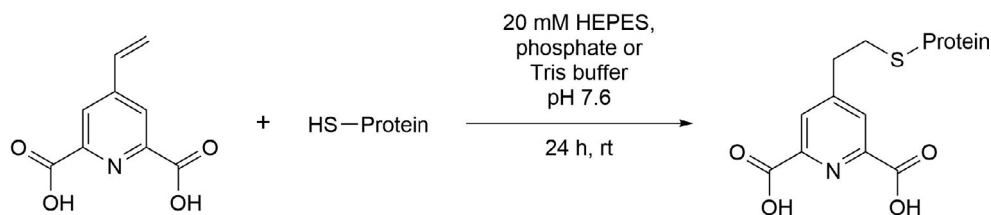


Fig. 9 Nucleophilic addition of the cysteine residue of a protein via thiol-ene reaction to the LCT reported by Li et al.¹³² in aqueous solution in order to create a reduction-stable linkage.

Since the interaction of the small molecule chelators with sidechain carboxylates can directly be used in order to immobilize the LCT on the surface, Swarbrick et al. developed a surprisingly simple LCT based on IDA.¹³³ Upon conjugation to the protein on a residue in an α -helix via a disulfide-bond, the LCT was shown to interact with an Asp residue in the $i + 4$ position by its free coordination sites. The synthesis of the LCT shows an overall yield of 45% for three steps, a result that demonstrates the straightforward availability of this LCT. However, the presented LCT is not generally applicable, since a basic residue in the $i + 4$ position, which is either present in the native state of the protein or has to be introduced by mutation, is required. Following the study with IDA-SH tags, Swarbrick et al. demonstrated that instead of IDA-based tags, also NTA-derived LCTs can be used for the generation of structural restraints using immobilization by a nearby carboxylate or with a second NTA-LCT in the $i + 4$ position resulting in an octadentate coordination motif.¹³⁴ In order to obtain large structural restraints by additional coordination to the IDA-SH and NTA-SH LCTs, Yagi et al. determined in a comprehensive study the optimal position of coordinating carboxylate groups of the sidechain in relation to the tag position.¹³⁵ While IDA-SH tags require additional coordination in order to yield large PCS, NTA-SH LCTs can also yield large paramagnetic effects independent of additional coordination by carboxylate groups of the protein.

Interestingly, Loh et al. incorporated an alkyne moiety in both NTA- and IDA-derived LCTs, giving rise to the possibility of bio-orthogonal tagging reactions, i.e. a copper catalyzed click reaction between the LCT equipped with an alkyne and an azide functional group of an AzF residue incorporated in the protein.¹³⁶ Thereby, the authors obtained large anisotropy parameters on a GB1 V21 construct using the alkyne-NTA LCT ($\Delta\chi_{\text{ax}} = 16.3 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 7.8 \times 10^{-32} \text{ m}^3$ for Tm^{3+}) and demonstrated the applicability of the novel LCT by conjugation to AzF-containing constructs of PpiB and ERp29. LCTs that offer the possibility for bio-orthogonal tagging reactions are generally applicable, even to proteins, that show structural distortions upon introduction of an additional cysteine or lose their biological activity.

Taking advantage both from vinylpyridine LCTs as well as IDA- and NTA-derived LCTs led to the synthesis of the 4-vinyl(pyridine-2,6-diyl)bismethylenitrilo tetrakis(acetic acid) (4VPyMTA) LCT by Yang et al.¹³⁷ Most likely due to the increased flexibility of the chelator and its pendant arms as well as the flexible linker moiety that can lead to averaging of paramagnetic effects, the newly developed LCT displays rather small anisotropy parameters. However, in contrast to DTPA-based LCTs, the presented LCT shows only one set of signals and furthermore, the protein-tag conjugate is stable under reductive conditions and the lanthanoid complex displays a higher stability than EDTA-lanthanoid chelates.

In order to synthesize a pyridine-derived LCT that provides more rigid side-arms, Huang et al. developed the 4'-mercapto-2,2':6,2''-terpyridine-6,6''-dicarboxylic acid (4MTDA) LCT.¹³⁸ The reported LCT provides very large anisotropy parameters of $\Delta\chi_{\text{ax}} = 38.9 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 5.2 \times 10^{-32} \text{ m}^3$ for the Dy^{3+} -loaded chelator on an ubiquitin A28C construct. The large observed PCSs and RDCs up to 13 Hz can be attributed to additional immobilization of the chelator on the surface of the protein by the carboxylate bearing sidechain of the residue E24. However, a significant drawback of the reported 4MTDA LCT is the ligation yield of only 50–70%, which is lower than for the previously described 3MDPA and 4MMDPA LCTs.

In order to provide LCTs that form a reduction-stable thioether bond upon conjugation to the protein, Yang et al. developed the (4-(phenylsulfonyl)-pyridine-2,6-dicarboxylic acid (4PS-DPA) and 4-phenylsulfonyl-(pyridine-2,6-diyl)bismethylenitrilo tetrakis(acetic acid) (4PS-PyMTA)) LCTs.¹¹⁹ The reported LCTs, which are closely related to the 4VPyMTA LCT published by Yang et al.¹³⁷ but incorporate a significantly shorter and more rigid linker moiety, yield large anisotropy parameters of $\Delta\chi_{\text{ax}} = -24.5 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = -7.1 \times 10^{-32} \text{ m}^3$ for Dy^{3+} on a ubiquitin G47C R72A R74A construct.

2.6 DOTA-Based Lanthanoid Chelating Tags

Due to the excellent chelation of lanthanoids by DOTA-derived chelators (equilibrium constant for the formation of the Gd^{3+} -DOTA complex = 10^{25} M^{-1} at pH 7)¹³⁹ that constitutes the main reason for its extensive application in magnetic resonance imaging (MRI) as well as radiopharmacy,^{140,141} the macrocyclic DOTA framework presents an optimal scaffold for the introduction of various substituents, donor atoms and linker moieties. Furthermore, by three-fold installation of a negatively charged donor atom, e.g. three lactic acid side-arms, and a noncharged linker donor atom, neutral lanthanoid DOTA complexes can be conveniently synthesized and the chance for Coulomb interactions of the LCT with the protein is successfully minimized. Due to the high affinity towards lanthanoid ions in combination with the possibility to rationally design the LCTs by introduction of suitable side arms, the DOTA ligand displays very favorable properties for the synthesis of high-performance LCTs.

Since DOTA-based LCTs can, based on two conformations of the ring scaffold ($\delta\delta\delta\delta$ and $\lambda\lambda\lambda\lambda$) and two conformations of the pendant arms of the chelator (Δ and Λ), form up to four different stereoisomers in solution ($\Delta(\delta\delta\delta\delta)$, $\Delta(\lambda\lambda\lambda\lambda)$, $\Lambda(\delta\delta\delta\delta)$, $\Lambda(\lambda\lambda\lambda\lambda)$) (Fig. 10), methyl substituents were introduced in order to restrict the conformational flexibility of the ligand.

Häussinger et al. reported in 2009 an eight-fold methyl-substituted LCT (Ln-DOTA-M8-(8S)-SSPy) with a single-armed linker that leads to very large PCSs over 5 ppm and RDCs over 20 Hz, properties that were unprecedented for single-armed LCTs.²⁹ The methyl substituents on the basic ring scaffold occupy in an equatorial-upper position^{128,142} upon coordination of the metal in order to minimize steric clashes and provide a suitable cavity for the lanthanoid ion. As shown by the authors, the introduction of methyl-substituents leads to a substantial increase of the conformational stability of the ligand that can be demonstrated in variable temperature spectra. Furthermore, the LCT can be used under extreme buffer conditions as e.g. strongly acidic pH-values, thereby enabling to monitor the unfolding of a protein.²⁹ The Tm³⁺ loaded LCT reported by Häussinger et al. adopts in solution a twisted square antiprismatic $\Delta(\delta\delta\delta\delta)$ conformation with negligible amounts of the square antiprismatic $\Lambda(\delta\delta\delta\delta)$ isomer (9% based on HPLC detection), while the Dy³⁺ complex, that adopts a square antiprismatic $\Lambda(\delta\delta\delta\delta)$ geometry as major conformation, contains a more substantial amount of the minor conformational isomer (27%).¹⁴²

As reported by Joss et al., the occurrence of the minor conformational isomer can be suppressed by an inversion of the stereochemistry of the side arms that results in exclusive presence of the $\Lambda(\delta\delta\delta\delta)$ isomer in the strongly paramagnetic Tb³⁺, Dy³⁺ and Tm³⁺ complexes as well as in the diamagnetic lutetium complex.³⁴ Subsequently, it was demonstrated by Opina et al. for the Ln-DOTA-M8-(4R4S) complexes, that the observation of the occurrence of only one conformational species, i.e. $\Lambda(\delta\delta\delta\delta)$, holds true for the whole lanthanoid series (the radioactive promethium was not investigated).¹⁴³ Surprisingly, the fitted anisotropy parameters on ubiquitin and human carbonic anhydrase II constructs both yield that the observed paramagnetic susceptibility of Ln-DOTA-M8-(4R4S)-SSPy³⁴ is significantly different from the previously published Ln-DOTA-M8-(8S)-SSPy.²⁹ Since both LCTs are structurally identical apart from the stereochemistry of the side arms and the associated switch in conformational geometry, it can be concluded that the coordination polyhedron and ligand-metal interactions can lead to a significant change in the induced anisotropy as suggested earlier for a range of coordination polyhedra by Mironov et al.¹⁰⁶

In order to expand the repertoire of DOTA-derived tags and enhance the obtained shifts with DOTA-M8-SSPy LCTs, Joss et al. introduced isopropyl substituents on the basic ring scaffold (Ln-P4M4-DOTA) and evaluated the performance of the novel chelator on ubiquitin and human carbonic anhydrase II constructs.^{36,144} The Ln-P4M4-DOTA LCT induces large PCS of up to 6.5 ppm on ubiquitin S57C and shows significantly increased anisotropy parameters when compared to its predecessor Ln-DOTA-M8-(4R4S)-SSPy. A comparison of the different properties of the newly synthesized paramagnetic complexes with Ln-DOTA-M8-(4R4S)-SSPy showed that the shift range (measured in one-dimensional ¹H spectra), position of donor atoms (monitored by DFT calculations) and translational motion of the LCT on the surface of the protein (modeled with a program developed by Suturina et al.¹⁴⁵) are highly similar for both LCTs. Therefore, it can be concluded that the reason for the strongly enhanced shifts induced by Tm- and

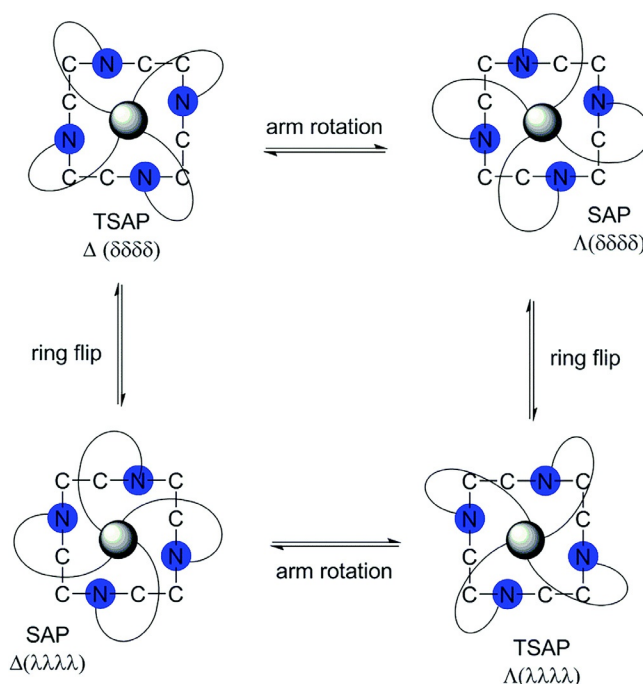


Fig. 10 Schematic representation of the different conformations of a DOTA-type ligand upon metal coordination and the paths of conversion between the conformers.¹⁴² Reproduced by permission of The Royal Society of Chemistry from Opina, A. C. L.; Strickland, M.; Lee, Y.-S.; Tjandra, N.; Byrd, R. A.; Swenson, R. E.; Vasalatiy, O. *Dalton Trans.* **2016**, 45, 4673–4687.

Dy-P4M4-DOTA when compared to Tm- and Dy-DOTA-M8-(4R4S)-SSPy can be attributed to decreased rotational freedom of the LCT relative to the protein due to the bulkier isopropyl substituents of the basic ring scaffold.

Since it would be highly desirable to apply rigidified DOTA chelators under reductive conditions as e.g. used for in-cell NMR, Müntener et al. developed a rigidified DOTA chelate bearing a pyridine-derived linker moiety (Ln-DOTA-(3R4S)-M7Py) that forms upon conjugation of the LCT to the protein a rigid and stable thioether bond.³³ Since the LCT showed only limited reactivity, an improved version with a fluorine substituent in *ortho* position of the leaving group was subsequently developed.³⁵

The electron-withdrawing substituent allows to stabilize the Meisenheimer-complex intermediate¹⁴⁶ which involves delocalization of negative charge within the aromatic ring during the S_NAr tagging reaction. However, the anisotropy parameters are partially averaged due to the suboptimal placement of the leaving group in the *para* position of the pyridine linker moiety. Therefore, a novel pyridinethiazole linker was developed that aligns the main axis of the induced anisotropy with the C_{pyr}-S_{Cys} bond and, thereby, leads to a favorable averaging of the induced anisotropy and hence to strongly enhanced PCSs and RDCs (Fig. 11).³⁵

In order to further enhance the obtained anisotropy parameters with the pyridinethiazole linker, the methyl substituents on the macrocyclic ring scaffold were replaced by isopropyl substituents to yield Ln-P4T-DOTA (Fig. 12).³⁷ Ln-P4T-DOTA induces very large PCS and was benchmarked on ubiquitin S57C, ubiquitin K48C and ¹⁵N selectively leucine labeled hCA II S166C constructs (anisotropy parameters: $\Delta\chi_{ax} = 44.3 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = 4.8 \times 10^{-32} \text{ m}^3$ for Tm³⁺ on a hCA II S166C construct). Due to its reduction-stable linker, the LCT could be applied for in-cell NMR studies.³⁷

In order to reduce motional freedom of the LCT on the surface of the protein, a two-point anchored LCT named CLaNP-3 was reported in 2007 by Vlasie et al.¹⁴⁷ In contrast to single-armed LCTs, two-point anchored LCTs yield in general strongly enhanced anisotropy parameters, since they are significantly more tightly attached to the surface of the protein. However, the expression of protein constructs that bear two cysteine residues in close proximity and provide thereby the attachment site of the double-armed LCT is more challenging than the expression of single cysteine mutants, since protein misfolding after expression can more easily

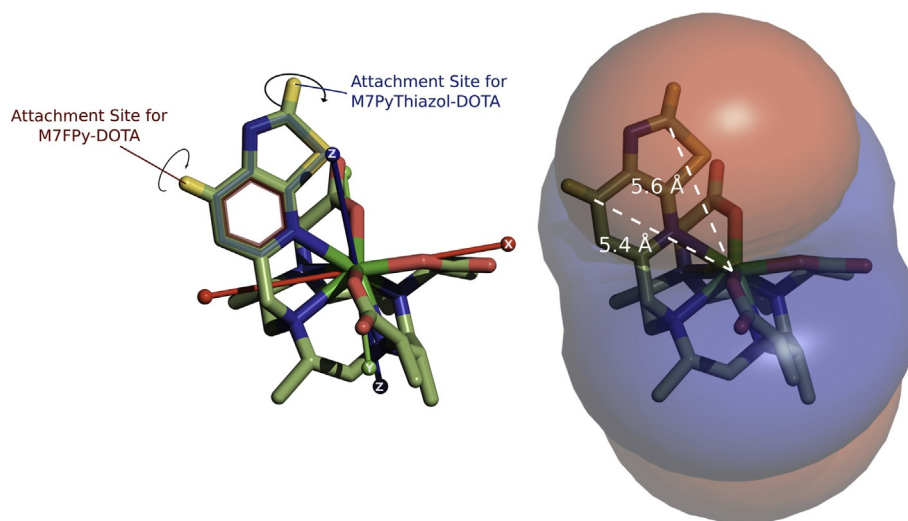


Fig. 11 Difference in the attachment point between Ln-M7FPy-DOTA and Ln-M7PyThiazol-DOTA. The more favorable alignment of the C_{pyr}-S_{Cys} linkage of Ln-M7PyThiazol-DOTA leads to a reduced averaging of the induced anisotropy and thus to larger paramagnetic effects observed on the protein.³⁵ Reproduced with permission from Müntener, T.; Kottelat, J.; Huber, A.; Häussinger, D. *Bioconjug. Chem.* **2018**, *29*, 3344–3351. Copyright 2016 American Chemical Society.

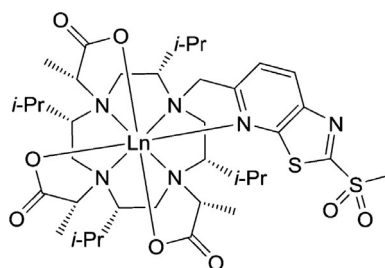


Fig. 12 Chemical structure of the $\Lambda(\delta\delta\delta\delta)$ conformer of Ln-P4T-DOTA.³⁷ The lanthanoid ion is coordinated by three carboxylate and one pyridine donors and the sterically demanding substituents rigidify the DOTA scaffold of the ligand. The methylsulfonyl on the pyridinethiazole moiety serves as leaving group upon reaction of the LCT with a cysteine residue of the protein.

occur than in proteins with cysteines that are located farther from each other. Using CLaNP-3 loaded with Yb^{3+} , PCSs for residues up to 35 Å away from the metal center were detected. Although the LCT presented by Vlasie et al. shows multiple conformations in solution that are in exchange with each other, in ^1H - ^{15}N HSQC experiments only one set of signals for most of the residues is detected, leading to the successful determination of one set of anisotropy parameters.

After the initial assessment of two-armed DOTA LCTs by Vlasie et al.,¹⁴⁷ Keizer et al. synthesized and reported the double-armed, DOTA-based CLaNP-5 LCT.^{28,148} The CLaNP-5 LCT presented by Keizers et al. yields large anisotropy parameters of $\Delta\chi_{\text{ax}} = 55.3 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 6.9 \times 10^{-32} \text{ m}^3$ for Tm^{3+} and $\Delta\chi_{\text{ax}} = 9.0 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 2.3 \times 10^{-32} \text{ m}^3$ for Yb^{3+} on a Paz E51C E54C construct.

Since the CLaNP-5 LCT displays a relatively high, three-fold positive charge, Liu et al. developed the one-fold positively charged CLaNP-7 LCT that is equipped with two nitro-substituted phenol side arms instead of the two pyridine N-oxide moieties found in CLaNP-5.¹²¹ Upon coordination of the metal, the two phenol groups are deprotonated to give the corresponding phenoxides that partially compensate the three-fold positive charge of the lanthanoid ion. Interestingly, the anisotropy parameters were found to be pH-dependent when a histidine residue is located in the immediate vicinity, a behavior that could originate in an apically coordinated water/hydroxide that can interact with the histidine sidechain.

Subsequently, Liu et al. developed the one-fold positively charged CLaNP-9 LCT that bears two carboxylate side arms.¹²² The presented LCT yields large PCS and anisotropy parameters on Paz Cu^{II} E51C E54C as well as T4 Lysozyme N55C V57C ($\Delta\chi_{\text{ax}} = 39.5 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 10.6 \times 10^{-32} \text{ m}^3$ for Tm^{3+} on the Paz construct and $\Delta\chi_{\text{ax}} = 32.5 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 15.6 \times 10^{-32} \text{ m}^3$ for Tm^{3+} on the T4 Lysozyme construct). Although the Yb -CLaNP-9 complex exhibits larger shifts in its one-dimensional ^1H spectrum, it displays smaller anisotropy parameters upon attachment to the protein when compared to CLaNP-5. This result suggests that the CLaNP-9 LCT shows a slightly enhanced mobility on the surface of the protein, which can be attributed to the linker that is elongated by ~ 2 Å when compared to CLaNP-5.

In order to decrease the conformational flexibility of the DOTA scaffolds, Graham et al. reported a DOTA-amide LCT that displays sterically demanding side arms that restrict the flexibility of the basic ring scaffold and side arms of the LCT and thus leads to only one set of signals in ^1H - ^{15}N HSQC experiments.³¹ It is noteworthy, that the shift range in one-dimensional ^1H spectra of the ytterbium complex amounts only to 170.5 ppm, while the LCT induces very large anisotropy parameters of $\Delta\chi_{\text{ax}} = 37.0 \times 10^{-32} \text{ m}^3$, $\Delta\chi_{\text{rh}} = 12.0 \times 10^{-32} \text{ m}^3$, for $\text{Ln} = \text{Tm}^{3+}$ on an ArgN protein construct. This result indicates that the bulky side arms do not only lead to a decrease in flexibility of the macrocyclic scaffold, but also efficiently suppress rotational averaging of the paramagnetic effects. Besides the introduction of more steric bulk, also shortening of the linker moiety can lead to significantly enhanced paramagnetic effects observed on the protein.

Lee et al. reported compact, hydrophilic DOTA-based LCTs bearing methyl substituents on the side arms and a short linker moiety resulting in the formation of a disulfide linkage to the protein.⁶⁸ The LCT presented by Lee et al. yields large observed anisotropy parameters on ubiquitin A28C ($\Delta\chi_{\text{ax}} = -19.4 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = -7.8 \times 10^{-32} \text{ m}^3$ for Tm^{3+}) and HPPK S112C C80A ($\Delta\chi_{\text{ax}} = 54.5 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 12.5 \times 10^{-32} \text{ m}^3$ for Tm^{3+}). In order to exploit the advantages of two-point anchored LCTs, Lee et al. converted both single-armed LCTs from the reported study,⁶⁸ i.e. the LCTs with a disulfide linker similar to Häussinger et al.²⁹ as well as the single-armed LCTs with a shorter disulfide linker, into two-point anchored LCTs that yield very large anisotropy parameters for the Tm^{3+} -loaded LCT on the tested ubiquitin E24C A28C and HPPK K76C C80 constructs (Fig. 13).^{123,124}

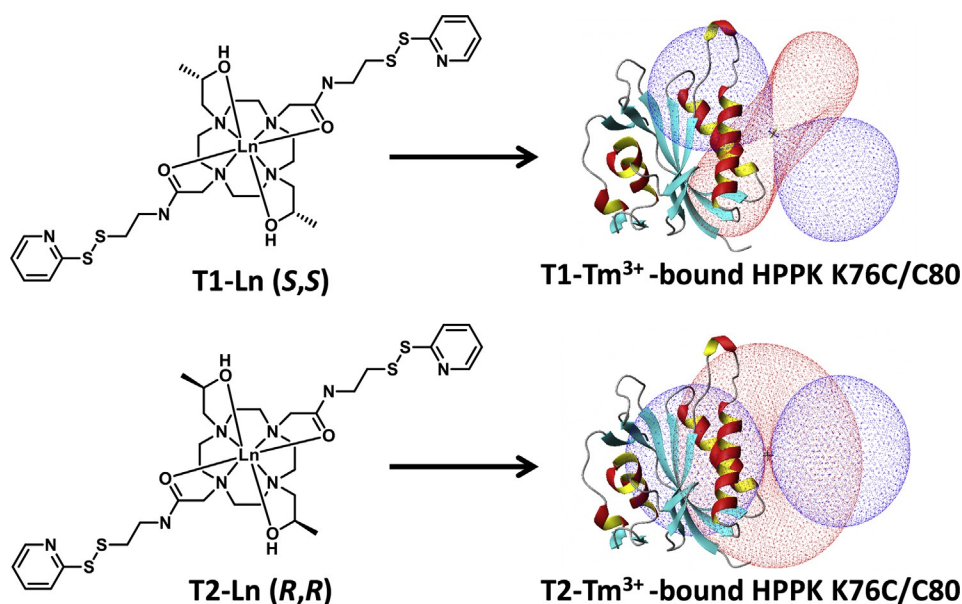


Fig. 13 Chemical structures of the Ln-T1 and Ln-T2 and isosurfaces induced by the Tm^{3+} -loaded LCTs attached to a HPPK K76C C80 construct.¹²³ Reproduced by permission of The Royal Society of Chemistry from Lee, M. D.; Dennis, M. L.; Swarbrick, J. D.; Graham, B. *Chem. Commun.* **2016**, 52, 7954–7957.

Remarkably, the magnitude of the induced PCS by the two-point anchored LCTs with a longer linker is significantly pH-dependent, i.e. the induced shifts at pH 6.5 are 40% smaller than the ones observed at pH 8.0.¹²³ This feature could be attributed to the hydroxyl donor groups, since for the related single-armed Ln-DOTA-M8-(8S)-SSPy LCTs equipped with carboxy groups, the pH dependence of the induced anisotropy was found to be not significant by Strickland et al. in the pH range of 5.0–7.4.⁷⁰ However, the pH sensitivity of the reported LCTs by Lee et al.¹²³ was not found for the LCTs with hydroxyl donor groups and a shortened linker,¹²⁴ therefore other reasons than the hydroxyl groups may be involved in the observed pH-dependence.

In order to provide reductively-stable linkages that can be applied, e.g. in in-cell NMR, the attachment of the LCT to a AzF residue incorporated in the protein via copper catalyzed azide-alkyne click reaction provides a very valuable solution.⁶⁷ Loh et al. reported two different DOTA-based LCTs with different linker lengths. For the shorter alkyne linker presented in the study, the triazole-moiety formed upon reaction of the alkyne group on the LCT with the azide side chain directly serves as donor to the lanthanoid ion after successful conjugation reaction. Both LCTs induce sizeable anisotropy parameters on a ubiquitin Glu18AzF construct with Tm³⁺ or Tb³⁺-loaded complexes. Importantly, the used protein must tolerate free copper ions used for the catalysis of the conjugation reaction. Therefore, protein constructs with His₆ tags are usually not suitable for ligation reactions with the described methodology.

A further example of an LCT yielding a reduction-stable linkage upon conjugation constitutes the iodoacetamide LCT Ln-M8-CAM-I reported by Hikone et al., which incorporates a carbamidemethyl linker with an iodide leaving group on an acetyl amide moiety that reacts as electrophile with the nucleophilic thiol group of cysteine residues of the protein.⁹⁵ In order to demonstrate the suitability of the LCT derived from Ln-DOTA-M8-(8S)-SSPy²⁹ for in-cell NMR studies, Hikone et al. ligated the newly developed LCT to ubiquitin S57C and introduced the tagged constructs successfully into HeLa cells via electroporation. The derived anisotropy parameters ($\Delta\chi_{ax} = 4.0 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = 1.8 \times 10^{-32} \text{ m}^3$ for Dy³⁺ at 310 K) demonstrate the suitability of the reported construct for in-cell NMR, however, the magnitude of the anisotropy parameters is significantly smaller than for Ln-DOTA-M8-(8S)-SSPy ($\Delta\chi_{ax} = 19 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = 11 \times 10^{-32} \text{ m}^3$ for Dy³⁺ at 313 K), which can be attributed to the significantly longer and therefore more flexible iodoacetamide linker.^{29,95}

Yang et al. presented in 2016 an LCT that yields extremely large anisotropy parameters and PCSs on an ubiquitin G47C construct ($\Delta\chi_{ax} = -84.3 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = -17.2 \times 10^{-32} \text{ m}^3$ for Dy³⁺ and $\Delta\chi_{ax} = 65.2 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = 31.8 \times 10^{-32} \text{ m}^3$ for Tm³⁺).³² The large PCSs are highly desirable for evaluation of challenging applications involving large biomacromolecules, however, due to the nonstabilized Meisenheimer complex¹⁴⁶ during tagging reaction of the cysteine with the pyridinesulfone moiety, the LCT requires severely basic tagging conditions, i.e. pH 9.0–9.3 over a time of 16–18 h. For a successful application of the strongly paramagnetic LCT, only proteins withstanding these harsh conditions can be investigated.

In order to attach LCTs to other biomacromolecules than proteins, Wu et al. developed a bromoacetamide linker that was demonstrated to be able to ligate to DNA incorporating a phosphorothioate modification.⁷² However, conjugation to an internal phosphorothioate in DNA generates an additional stereocenter, i.e. a stereogenic phosphorus atom, in racemic form and hence to two sets of diastereotopic NMR signals. In the reported case, the two diastereomers can be successfully separated using conventional reverse-phase HPLC. PCSs were extracted from 2D NOESY spectra and fitted to a structural model of the DNA double-helix based on fiber diffraction data in order to give the first anisotropy parameters that were obtained by using an LCT with a covalent linkage to DNA.

The discussed examples in this section highlight the outstanding properties of DOTA-type ligands for the development of tuneable, strongly paramagnetic LCTs with a diverse range of linker and activator moieties and their subsequent application in structural studies of protein–ligand complexes.

3 Applications of Lanthanoid Chelating Tags in the Localization of Ligands Within Biomacromolecules

3.1 Introduction

The localization of ligands within biomacromolecules is crucial for the understanding of the cellular machinery and the underlying processes. Furthermore, the positioning and orientation of a ligand within a drug target is an important prerequisite for rational design of drug molecules. In contrast to X-ray crystallography, which provides in principle a precise tool for the determination of ligand-protein complexes, the use of NMR spectroscopy allows for monitoring of biomacromolecule-ligand interactions under physiologically relevant conditions. While NOESY experiments are optimally suited for monitoring distances below 5–8 Å and require a full NMR assignment obtained from multidimensional experiments by using isotopically labeled biopolymers, CSP can result in misleading conclusions due to conformational changes or unexpected intermolecular interactions of the biomacromolecules. The exploitation of PCS for the characterization of biomacromolecule-ligand complexes can cover large distances beyond 100 Å with high accuracy. Furthermore, if the anisotropy parameters from multiple sites have been determined, the accordingly tagged protein constructs can be used with a large number of ligands, thus providing a valuable tool for screening of potential drug molecules.

3.2 Weakly Binding Ligands

In order to determine the position of a weakly binding ligand within a protein that is in fast exchange, the ligand shifts can be conveniently obtained from one-dimensional NMR experiments. Since the ligand can be employed in excess, the signal discrimination between ligand resonances and protein peaks is ensured by the intensity and the PCS can be extracted from the observed

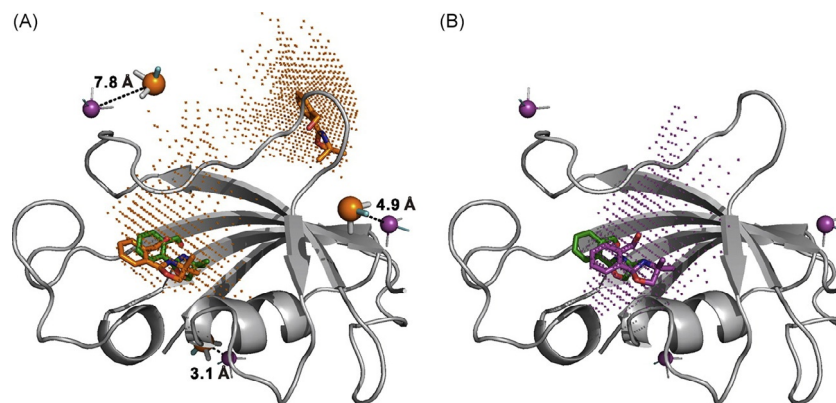


Fig. 14 Localization of ligands within a protein scaffold (predicted metal positions: magenta, experimentally determined metal positions: orange, NOE-based ligand positions: green, ligand positions determined by PCS: orange, ligand position determined by use of predicted anisotropy parameters: magenta). Fitted ligand position based on experimental data is in accordance to NOE position, with simultaneous occurrence of a “ghost” site during the fitting process (A). Fitted ligand position based on predicted anisotropy parameters (B).⁴⁰ Reproduced with permission from Guan, J.-Y.; Keizers, P. H. J.; Liu, W.-M.; Löhr, F.; Skinner, S. P.; Heeneman, E. A.; Schwalbe, H.; Ubbink, M.; Siegal, G. *J. Am. Chem. Soc.* **2013**, *135*, 5859–5868. Copyright 2013 American Chemical Society.

shift for the fast exchanging bound and unbound fraction of the ligand provided the affinity is known, and the PCS is large compared to the CSP. Saio et al. determined the positioning of low affinity ligands within the growth factor receptor-bound protein 2 Src homology domain by use of PCS analysis.¹⁰² The paramagnetic shifts were induced by a lanthanoid-binding peptide tag attached via two linkages on the N-terminus and one cysteine mutant in the protein. Upon determination of the anisotropy parameters, the weakly binding pYTN tripeptide ligand and the tagged protein construct were combined and the PCS were extracted from the determined averaged shift of the unbound and the bound state using the binding affinity of the tripeptide to the protein obtained from a ^1H - ^{15}N HSQC-based titration.

The combination of the structural restraints obtained by PCS analysis and docking simulations yielded then the position of the ligand within the biomacromolecule. In 2013, Guan et al. presented the localization of ligands using PCS analysis using the rigidly, two-point anchored LCT CLaNP-5 attached to three different sites of the protein.⁴⁰ Interestingly, two-point anchored LCTs as CLaNP-5 that are strongly immobilized on the surface of the protein, yield reproducible and predictable anisotropy parameters (Fig. 14). This leads to the convenient feature that the predicted $\Delta\chi$ parameters can be directly used for a coarse grain localization of the ligand within the biomacromolecule (RMSD of 4.4–4.7 Å relative to NOE structure). Hence, for an initial assessment or discrimination of two already established binding sites in a protein, no isotopically labeled protein or backbone amide assignment is required and the only input required is the PCS from one-dimensional ^1H experiments. The authors reported that upon the paramagnetic labeling of the protein on three different sites and fitting of the anisotropy parameters based on backbone amide NMR shifts, the ligand can be localized and positioned over a distance of 15–25 Å with an average root-mean-square deviation (RMSD) of 2.8 Å when compared to the NOE structure. In contrast to the attachment of the paramagnetic tag to the biomacromolecule of interest, as demonstrated by Brath et al., the LCT can also be directly attached to the weakly binding ligand.¹⁴⁹ Thereby, the authors identified the interaction site of calmodulin with a low-affinity anesthetic agent. However, although the direct attachment of the LCT to the ligand of interest has the advantage that dissolution of hydrophobic ligands is enhanced in aqueous solution and engineering of protein cysteine mutants is omitted, the approach reported by Brath et al. is less suitable for the screening of a large number of different ligands, since the LCT would have to be synthetically coupled to each of the applied ligands.

3.3 Ligands in the Intermediate Exchange Regime

Ligands in the intermediate exchange regime can pose significant challenges for PCS NMR approaches, since their resonances are often broadened beyond detection and therefore no structural restraints can be measured. An elegant tool to circumvent this difficulty is the use of the chemical saturation exchange transfer methodology (CEST) as proposed by Gao et al.¹⁵⁰ The authors of the study suggested to detect the PCS on the spectrum of the free ligand by scanning through the ^{19}F saturation frequencies of the bound state. Since intensity dips can be observed at the frequencies of the signals of the bound ligand, a fluorine spectrum of the bound ligand can thereby be collected. The authors of the study then used the PCS recorded by use of their ^{19}F CEST methodology in combination with PCS-driven docking programs as HADDOCK or XPLOR-NIH containing the PARArestraints module in order to localize mono- and difluorinated inhibitors in the BRM bromodomain.

In order to extend the presented method by Gao et al.¹⁵⁰ for the localization of ligands with micromolar affinities and NMR resonances in the intermediate exchange regime that do not incorporate fluorine, Xu et al. developed a new method that is suitable for localization of ligands with micromolar affinities.¹⁵¹ Since ^1H CEST usually is not applicable in a clean fashion, since it can suffer severely from NOE interferences, Xu et al. proposed the use of relaxation dispersion (Fig. 15). By application of their novel methodology, the authors obtained PCS restraints and used them successfully as input restraints in docking simulations of the ligand within the protein.

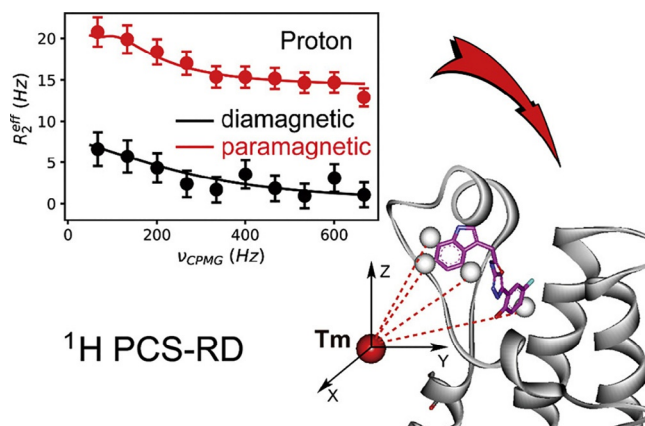


Fig. 15 Extraction of PCS measured on ligands in the intermediate exchange regime by use of relaxation dispersion.¹⁵¹ Reproduced with permission from Xu, D.; Li, B.; Gao, J.; Liu, Z.; Niu, X.; Nshogoza, G.; Zhang, J.; Wu, J.; Su, X.-C.; He, W.; Ma, R.; Yang, D.; Ruan, K. J. *Phys. Chem. Lett.* **2018**, *9*, 3361–3367. Copyright 2018 American Chemical Society.

3.4 Strongly Binding Ligands

In the localization of strong ligands, usually large PCS can be detected, since the observed signal originates exclusively from the bound state that experiences the full PCS and not an average shift of the bound and unbound state as in the case of the weakly binding ligands in fast exchange. However, the signal discrimination of the bound, strongly binding ligand against the signals of the protein, which is present in equimolar amounts within the sample, bears a significant challenge. Besides the equimolar ratio and therefore similar shift intensities, tightly binding ligands exhibit the same rotational correlation time as the signals of the protein and are therefore broadened in similar extent. Isotopic labeling of the ligand, e.g. with ^{13}C atoms, is certainly achievable, however, for a large number of ligands that have to be screened against a protein of interest this cannot be considered as a convenient approach. Saio et al. achieved the required signal discrimination by deuteration of the protein and measurement of one-dimensional ^1H spectra.¹⁰² Thereby, the authors were able to localize a high-affinity inhibitor of the Src homology 2 domain of the growth factor receptor-bound protein 2 by use of the observed PCS and docking simulations over a distance of $\sim 10 \text{ \AA}$.

An approach for overcoming the issue of signal discrimination of strongly binding ligands was proposed by Chen et al., who used a *tert*-butyl group, which displays a strong signal with a specific NMR shift, as reporter group for detection of the signal of the bound ligand (Fig. 16).¹⁵²

The researchers found that broadening and simultaneous shift of the signals due to the paramagnetic tags can lead to ambiguous assignments caused by insufficient signal discrimination when compared to the protein. However, the *tert*-butyl group incorporated in the ligand produces very intense NOESY correlations even when the protein is hardly detectable. Thereby, the signals of the ligand can be unambiguously assigned and discriminated towards the protein background. Since the used approach omits the need for perdeuteration of the protein samples and spectra can be acquired at very low concentration, which is important for screening of a large number of ligands, the methodology reported by Chen et al. is highly suitable for all ligands containing *tert*-butyl groups. In order to unambiguously localize fluorine-containing ligands over large distances, Zimmermann et al. developed an approach based on the acquisition of one-dimensional ^{19}F spectra with paramagnetic tags attached to different mutation sites on the protein scaffold.⁴¹ This methodology will be described in the next section, which covers the enzyme human carbonic anhydrase II and the localization of ligands within the protein by PCS, and lies at the heart of the present case study.

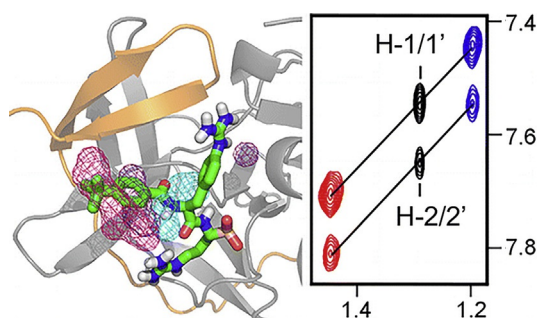


Fig. 16 Strong correlation peaks of the *tert*-butyl group of the applied, strongly binding ligand (right). Structural models of the protein–ligand complex of DENpro and ligand 1. The fitted regions with a deviation less than 0.02 ppm from the experimental PCSs are indicated with color coding: *tert*-butyl group (magenta), aromatic protons (purple and cyan).¹⁵² Reproduced with permission from Chen, W.-N.; Nitsche, C.; Pilla, K. B.; Graham, B.; Huber, T.; Klein, C. D.; Otting, G. *J. Am. Chem. Soc.* **2016**, *138*, 4539–4546. Copyright 2016 American Chemical Society.

4 Localization of Ligands Within Human Carbonic Anhydrase II Using ^{19}F Pseudocontact Shift Analysis

4.1 Human Carbonic Anhydrase and Its Use in the Development of Artificial Metalloenzymes

Human carbonic anhydrase II (hCA II) is a medium-sized, monomeric, globular protein with a molecular weight of 29 kDa. hCA II catalyzes the reversible hydration of CO_2 to bicarbonate and displays a catalytic turnover rate of $k_{\text{cat}} = 1.4 \times 10^6 \text{ s}^{-1}$ and an efficiency of $k_{\text{cat}}/K_{\text{M}}$ of $1.5 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$, which rank among the highest known for enzymatic reactions and can be considered to be close to the limit imposed to the reaction by diffusion.^{153–156} The reaction is best described by a two-step mechanism.¹⁵⁶ First, CO_2 is directly attacked by the zinc bound hydroxide ion to form a metal bound bicarbonate, which is then replaced by a water molecule in order to release the bicarbonate anion from the zinc metal center. In the second step, a proton from the zinc bound water molecule is transferred, using a histidine residue as proton shuttle, to the aqueous solvent surroundings in order to regenerate the zinc bound hydroxide. The central zinc(II) ion of hCA II is coordinated by three histidine residues and one hydroxide ion in a distorted tetrahedral geometry in order to form together the active site. The active site itself is located at the bottom of a conical shaped cavity, which is about 15 Å deep and 15 Å wide at the entrance. The cavity leading to the active site comprises a hydrophilic and a hydrophobic side.¹⁵⁷ The Zn(II) ion in combination with the hydrophobic side is the binding site for virtually all hCA II inhibitors.¹⁵⁶ An interesting structural feature displayed by hCA II is the knot-like topology, i.e. if one were to pull at both ends of the peptide chain, a knot would form (Fig. 17).¹⁵⁷

In order to use hCA II as a drug target, high affinity inhibitors of hCA II have been developed.¹⁵⁶ Strongly binding inhibitors of hCA II are based on an aromatic sulfonamide functional group within the ligand.^{158,159} The strong binding affinity originates in the ability of the sulfonamide anion ($\text{RS}(\text{O})_2\text{NH}^-$) to strongly coordinate the zinc ion.¹⁶⁰ The properties of the inhibitor can be tailored by adjusting the aromatic ring and adding different functional groups to the aromatic scaffold.¹⁵⁶ Importantly, the inhibition of hCA II allows the control of different diseases, e.g. glaucoma, a disease that results in nerve damage followed by a loss of the patient's vision. hCA II plays an important role in the control of the intraocular pressure. The conversion of CO_2 into bicarbonate catalyzed by hCA II in the ciliary body causes the uptake of sodium ions and thus also water into the eye. The sodium ions together with water, other electrolytes and components as amino acids and glutathione form the aqueous humor. Increased intraocular pressure is one of the most important risk factors for glaucoma. The inhibition of hCA II suppresses the formation of bicarbonate, which results in less uptake of sodium ions and water and thus leads to a reduction of the intraocular pressure.¹⁶¹

The large binding pocket of hCA II in combination with strongly binding and readily derivatizable inhibitors rendered hCA II likely to be an optimal scaffold for artificial metalloenzymes. Additionally, hCA II is easily overexpressed and is therefore highly suitable for an optimization by directed evolution in order to improve the performance of the organometallic protein construct. By today, directed evolution has become a highly advanced technique that can be readily used in order to optimize the catalytic properties of computationally designed enzyme folds or of proteins with promiscuous activity that are auspicious catalysts for other than only their natively catalyzed reaction.^{162–164} Transition metal complexes can be anchored to hCA II by tethering a bidentate ligand to an aryl-sulfonamide inhibitor.¹⁶⁵ Iridium-based piano stool complexes incorporated into hCA II were then shown to be able to catalyze an asymmetric transfer hydrogenation of imines.¹⁶⁶ Subsequently, the catalytic performance was improved by computational design, which led both to an increase of the enantioselectivity (70–96% ee) as well as the turnover number (9–59) (Fig. 18).¹⁶⁷

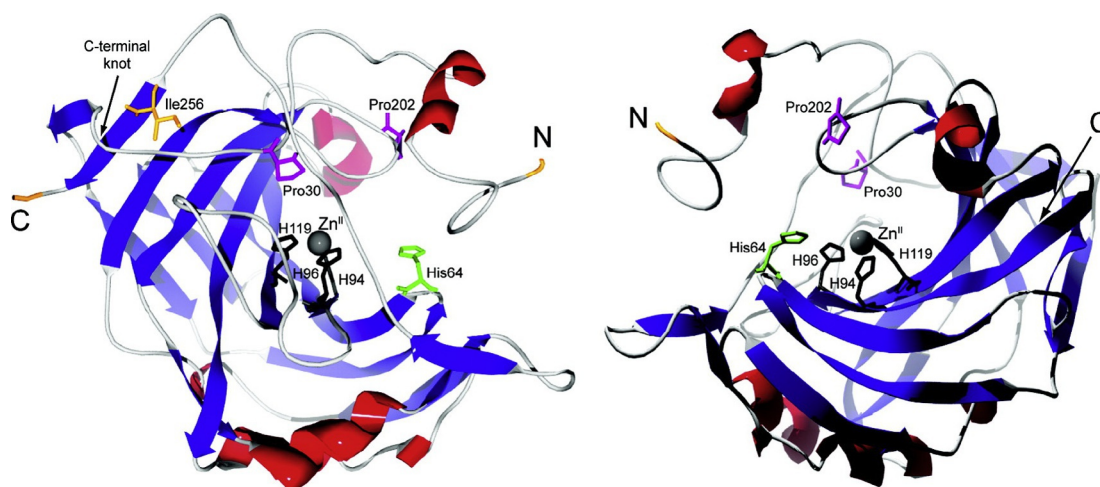


Fig. 17 View on the fold of hCA II displayed as ribbon structure (α -helices: red, β -sheets: blue). The active site Zn(II) ion is coordinated by three histidine residues, and lies at the center of the enzyme. Furthermore, the critical residues for the initiation of folding of hCA II are marked within the displayed structure.¹⁵⁶ Reproduced with permission from Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudixsen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108*, 946–1051. Copyright 2016 American Chemical Society.

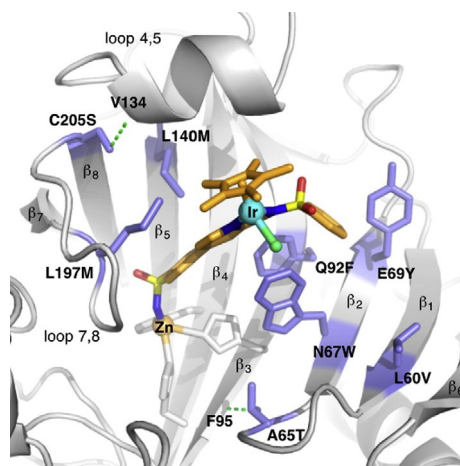


Fig. 18 View on the active site of an artificial metalloenzyme incorporating an sulfonamide iridium piano stool complex for efficient, asymmetric catalysis in the synthesis of (*S*)-salsolidine, an inhibitor of monoamine oxidase. The H-bonds introduced using computational design are marked with dashed lines.¹⁶⁷ Reproduced with permission from Heinisch, T.; Pellizzoni, M.; Dürrenberger, M.; Tinberg, C. E.; Köhler, V.; Klehr, J.; Häussinger, D.; Baker, D.; Ward, T. R. *J. Am. Chem. Soc.* **2015**, *137*, 10414–10419. Copyright 2016 American Chemical Society.

In a following study, *in vivo* transfer hydrogenation enabled by iridium catalysts incorporated into the hCA II scaffold was investigated.¹⁶⁸ Furthermore, computational studies elucidating the amino acid residues, which contribute significantly to the binding of the ligands,¹⁵⁹ as well as comparing both point-charge and atomic multipole-scored methods for studying the binding of sulfonamide ligands to hCA II were reported.¹⁶⁹ One of the most promising fields of applications for metalloenzymes is their ability to activate drugs at a specific target location *in vivo*. This could be used to specifically target cancer cells, since cancer cells overexpress carbonic anhydrase IX in order to arrange with acidosis due to their strongly accelerated and anaerobic metabolism.¹⁷⁰ In this context, the ring-closing metathesis (RCM) constitutes a promising reaction as it allows the bioorthogonal conversion of acyclic diolefins to cyclic products. By use of hCA-based artificial metalloenzymes, RCM can be performed by anchoring a second-generation Hoveyda-Grubbs catalyst to the active site of the protein using an arylsulfonamide part on the catalyst.¹⁷¹

The same properties that render hCA II as an interesting host for metalloenzymes also result in an excellent suitability as test protein for biophysical studies of protein–ligand binding.¹⁵⁶ The convenient derivatization of the inhibitors in combination with reported X-ray structures of protein–ligand complexes pose hCA II as an excellent choice in order to develop new methods of ligand localization. In order to localize ligands for hCA II within the protein scaffold in solution, a method based on paramagnetic NMR spectroscopy was developed.⁴¹ More specifically, ¹⁹F detected, lanthanoid-induced PCSs served as structural restraints in order to localize the ligands of interest within the protein scaffold in solution resulting in an accuracy of up to 0.8 Å over a distance of 22–38 Å.

4.2 Localization of Ligands within Human Carbonic Anhydrase II—Introduction and Concept

Localization of ligands within proteins in solution is crucial to the rational design of drugs and investigation of biomedical pathways within the cellular machinery. In order to study the structure of ligand–protein complexes that are not straightforward amenable to X-ray crystallography¹⁷² as well as cryo-electron microscopy¹⁷³ and in order to get insights into the dynamics of the complexes, NMR spectroscopy and in particular paramagnetic NMR spectroscopy constitutes a highly promising method.¹⁸ As discussed in the previous section, the localization of firmly binding ligands within proteins in solution using NMR poses significant challenges as signal discrimination and the short distance range covered by conventional NMR spectroscopy. Given the three-dimensional structural restraints and distance dependency of R^{-3} displayed by PCSs, they deliver an optimal tool for the calculation of the ligand position within the protein based on long-range restraints and can thereby overcome the distance limitation of NOE experiments and unspecific results by CSP. Furthermore, during the screening process of the ligands, the protein is not required to be labeled, since only the ligand PCS is used, detected from one-dimensional spectra. In order to achieve signal discrimination between the signals of a strongly binding ligand and the protein, different approaches were already implemented. Saio et al. used deuterated protein samples in order to conveniently detect the signals of the ligand.¹⁰² In order to remove the need for protein labeling, Chen et al. proposed to use a *tert*-butyl substituent as suitable reporter group, since in NOESY spectra exceedingly strong correlations are detected for the *tert*-butyl group, even when the protein correlations are barely visible.¹⁵² Since the percentage of drugs incorporating fluorine amounted to 25% in 2014¹⁷⁴ and the fluorine detection removes the need for protein labeling while giving an astonishingly high sensitivity due to the favorable gyromagnetic ratio and the 100% natural abundance of the NMR active isotope, we envisioned in our study to introduce a fluorine-detected localization of ligands within proteins using PCSs as structural restraints.⁴¹ In order to test our approach, we incorporated three different, strongly binding sulfonamide ligands within hCA II (Fig. 19).

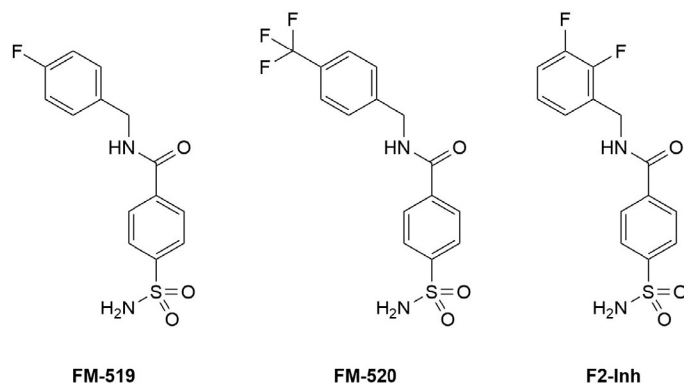


Fig. 19 hCA II inhibitors for the evaluation of the ^{19}F -PCS-based approach for localization of ligands within proteins.

For the unambiguous localization of a ligand within a protein scaffold in three-dimensional space, PCSs from different tagging sites are needed. Alternatively, a smaller number of tagging sites can be used in combination with different LCTs that yield independent anisotropy tensors. By using multiple tagging sites, first, the anisotropy parameters for each tagging site are fitted by using the backbone amides. Upon detection of the PCSs on the ligand signals, the values are compared with the back-calculated positions for a given PCS derived from the fitted anisotropy parameters depicted as isosurfaces around the paramagnetic center.

Intersection of two isosurfaces from two different positions yields a curve that includes the possible ligand positions (Fig. 20). By use of a third isosurface, the curve gets intersected at two different points that depict the remaining two possible positions for the ligand position. Intersection with a fourth isosurface leads then to one point in space as unique solution for the fluorine atom. However, since monitoring the ligand position from three different tagging sites results in two possible locations and often one of these possibilities is located completely outside the protein scaffold or in very dense regions in the interior of the protein, in practice usually a fourth position is not required to determine unambiguously the position of the ligand within the protein.

In order to obtain different tagging sites for attachment of the LCTs, five serine to cysteine mutations were introduced on the surface of the protein: S50C, S166C, S173C, S217C and S220C. Furthermore, different labeling schemes were introduced, each with its own purpose: uniformly ^{15}N labeled hCA II for the determination of the anisotropy parameters from ^1H - ^{15}N HSQC spectra and selectively ^{15}N leucine labeled hCA II in order to serve as starting point for the fitting process due to crowded paramagnetic spectra. The 26 leucine residues in hCA II show a favorable distribution in the primary sequence, amount to 10% abundance and the chemical shift dispersion in ^1H - ^{15}N HSQC experiments is reasonable as expected for a well-folded, globular protein. Upon tagging with a diamagnetic LCT, a small number of residues close to the tagging site are shifted. In order to contribute for the occurrence of these shifts, we attached, besides the paramagnetic Tm^{3+} -loaded LCT, also the diamagnetic Lu^{3+} -loaded LCT to the protein in order to acquire reference spectra. Furthermore, in order to provide complete evidence for the assignment of the shifted residues

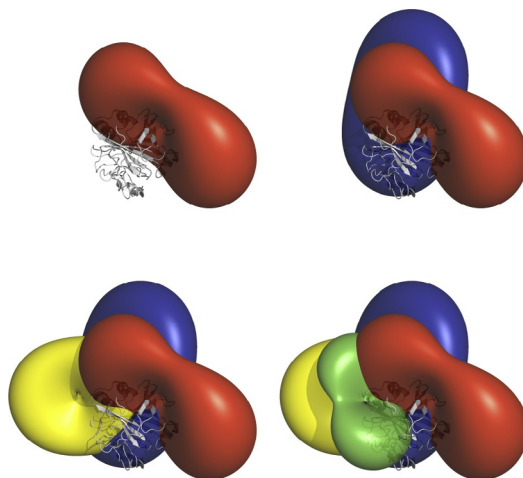


Fig. 20 Intersection of different isosurfaces for the localization of a specific nucleus within a protein scaffold using experimental and back-calculated PCS: A single isosurface plotted besides hCA II (*top left*), intersection of two isosurfaces originating from PCS measured from two different sites within the protein and their curve of intersection (*top right*), intersection of three isosurfaces resulting in two intersection points (*bottom left*, first intersection point observable in the center of the protein scaffold, the second on the backside of the protein), intersection of four isosurfaces resulting in one and thus unambiguous intersection point (*bottom right*).⁴¹ Reproduced by permission of The Royal Society of Chemistry from Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.

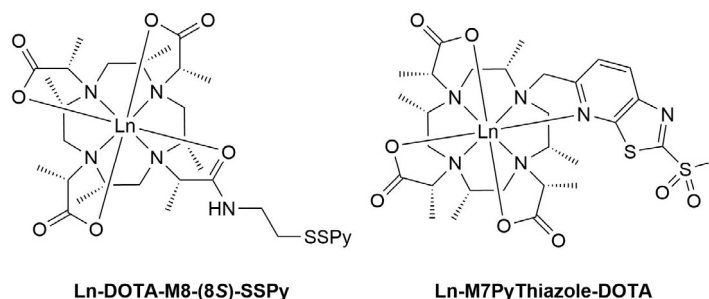


Fig. 21 Strongly paramagnetic LCTs applied in the localization of ligands within hCA II. Ln-DOTA-M8-(8S)-SSPy in $\Delta(\delta\delta\delta\delta)$ and Ln-M7-(3*R*,4*S*)-PyThiazole-DOTA in $\Lambda(\delta\delta\delta\delta)$ conformation (see also Section 2.6).^{29,35}

upon tagging reaction, we produced a triple labeled mutant that was then used for backbone assignment with and without attached Lu^{3+} -LCT.

Upon a successful assignment of the backbone of hCA II, the anisotropy parameters induced by the LCTs ligated to different attachment sites can be determined by fitting the amide signals detected in ^1H - ^{15}N HSQC experiments to the X-ray structure of hCA II. Since the uniformly ^{15}N labeled protein displays crowded spectra, selectively ^{15}N leucine labeled mutants are used to provide a reasonable starting point for the fitting routine. Full assignment of the PCS and fitting to the X-ray structure leads then to the derived anisotropy parameters. After fitting of the anisotropy parameters, the protein is incubated with the ligand of interest and excess ligand is subsequently removed by simple centrifugation. By acquiring one-dimensional ^{19}F experiments, the PCS of the fluorine atoms incorporated in the ligand can be measured. Back-calculation of the expected values in different positions in the protein by use of the previously fitted anisotropy parameters from multiple tagging sites leads then straightforwardly to the positioning of the ligand within the protein. Since large PCS decrease significantly the error and lead to more precise structural restraints, an LCT is required that induces large PCS and is suitable for exploration of the hCA II scaffold that is characterized by a diameter of 40–56 Å and a molecular weight of 30 kDa.¹⁵⁶ We used the eight-fold methyl-substituted Ln-DOTA-M8-(8S)-SSPy²⁹ and the recently published Ln-M7PyThiazole-DOTA,³⁵ which incorporates a short and rigid pyridinethiazole-derived linker moiety (Fig. 21).

4.3 PCS Analysis and Determination of the Anisotropy Parameters from Multiple Positions Within the Protein

The PCS and the corresponding anisotropy parameters are generated by the attachment of an LCT to the protein scaffold on multiple positions. In order to generate valuable structural restraints, a few criteria have to be met: (i) The LCT needs to be attached in a stable and rigid position relative to the protein. (ii) The global fold of the protein must not be compromised by the attachment of the LCT. (iii) The attachment sites should be distributed all over the protein in order to achieve a reasonable orthogonality of the different induced anisotropy parameters and the corresponding isosurfaces. In order to meet all of the mentioned criteria, five solvent-exposed serine residues in various positions of the protein located in stable secondary structure elements were mutated in order to produce five different cysteine mutants (Fig. 22). The native cysteine 206 present in human carbonic anhydrase II was mutated vice versa into a serine residue.

The PCS methodology requires a backbone assignment of the protein, in the best case a backbone assignment of the diamagnetically tagged protein in order to avoid ambiguities caused by shifted residues close to the tagging site, and therefore a backbone assignment of uniformly ^2H ^{13}C ^{15}N labeled hCA II S50C Lu-DOTA-M8 was performed. With the assigned amide resonances for the diamagnetic spectra in hand, the shifted residues due to PCSs are assigned. Based on the experimentally observed PCS on the amide resonances from multiple tagging sites and the structure of the protein, the anisotropy parameters induced by the LCT for each tagging site can then be fitted. However, due to the crowded ^1H - ^{15}N HSQC spectra of uniformly ^{15}N labeled hCA II, the direct assignment of all PCS can be cumbersome (Fig. 23).

In order to reduce the chance of obtaining an erroneous PCS assignment, multiple approaches were developed: (i) selectively labeled proteins, (ii) variable temperature NMR or (iii) use of different lanthanoids.

Selectively labeled proteins, e.g. selectively ^{15}N leucine labeled proteins as used in this study, decrease drastically the complexity of the ^1H - ^{15}N HSQC spectra (Fig. 24). The assignment of the shifted resonances in the spectra with reduced complexity that are therefore more amenable to PCS analysis leads then to an initial set of anisotropy parameters. The back-calculation of the remaining unassigned PCS based on the initially derived anisotropy parameters then strongly facilitates the complete assignment.

Measurement of PCS at variable temperatures produces a series of differently shifted signals, since PCSs are temperature dependent and, in addition, an increase in temperature leads to enhanced movements of the LCT on the surface of the protein and flexibility of the protein scaffold, which results in decreased PCS. The thereby produced pattern of signals significantly simplifies the assignment due to the “linkages” provided between the diamagnetic shift and the paramagnetic shift at room temperature (Fig. 25).

A very similar approach involves the application of different lanthanoid ions. The major determinants of the strength of the induced PCS are the intrinsic anisotropy of a lanthanoid within a given ligand-field/chelator and its motion on the surface of the protein. Since the motion can be expected to be similar for all different lanthanoids, except for very rare cases (when a larger

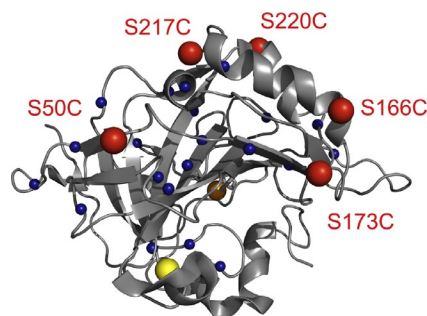


Fig. 22 Cysteine mutations introduced in the hCA II samples visualized in the X-ray structure of hCA II (PDB 3KS3,¹⁷⁵ introduced cysteine mutations: red, native cysteine 206 (mutated to serine): yellow, leucine residues: blue, zinc(II) ion: brown).⁴¹ Reproduced by permission of The Royal Society of Chemistry from Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.

lanthanoid ion would enable coordination of a protein sidechain on the ninth coordination site and a smaller lanthanoid would not show this effect), the complexation of different lanthanoids and attachment of the different LCTs to the protein leads to a pattern of signals similar to the measurement at various temperatures. The pattern of the shifts induced by application of different lanthanoid ions can then be used for a convenient assignment of the PCSs (Fig. 26).

The selective labeling approach is usually for severely crowded ^1H - ^{15}N HSQC spectra the more convenient choice, since the variable temperature approach and application of a series of lanthanoids even increase the complexity of the overlaid spectra in strict technical terms.

Therefore, selectively ^{15}N leucine labeled hCA II mutants were expressed and led in combination with the assignment of uniformly labeled hCA II tagged with Lu-DOTA-M8 to an initial fitting of the acquired PCS. Based on the leucine derived tensor parameters, the PCS of the remaining residues of the uniformly ^{15}N labeled hCA II were back-calculated and the experimentally measured PCS fitted to the structure (PDB 3KS3)¹⁷⁵ using the software Numbat.⁹⁷ A full description of the fitting routines can be found in the supporting information of the original publication.⁴¹

The complete results of all hCA II constructs tagged with Tm-DOTA-M8 are depicted in Table 1. Importantly, in order to obtain an estimation of the error included in the fitted anisotropy parameters, two different Monte-Carlo approaches can be applied. While the first one uses a structure variation during the fitting process, the second one uses only subsets of PCS to account for the possibility that erroneous assignments are hidden due to the good quality of the rest of the fitted resonances. Subsequently, two different criteria can be used to assess the quality of the fits. First, the location of the metal center obtained in a fit without used restraints in a reasonable positioning to the cysteine residue and in a distance of 6–8 Å to the γ -oxygen confirms the metal position found by the fitting routine. Second, the Q-factors give an impression about the quality of an obtained fit. The Q-factors for the used LCTs^{29,35} and hCA II constructs were found to be excellent (0.035–0.106). The complete fits of all PCS observed with the Tm-DOTA-M8-SSPy²⁹ and Tm-Thiazole³⁵ LCT constitutes the basis for an accurate back-calculation/prediction of the ^{19}F PCS observed on the ligands and thus their localization within the protein. For the Tm tagged S166C mutant, significantly larger PCS and anisotropy parameters were found than for the other constructs. This result can be assigned to a different environment of the tagging site that causes a different flexibility of the LCT on the protein.¹⁴⁵

With the full set of fitted PCS from various attachment sites in hand, the localization of the ligands can be tested by incubation of the tagged hCA II constructs with the three different ligands. The high-affinity phenyl-sulfonamide ligand molecules (FM-519, FM-510 and F2-Inh) were added to the buffered protein sample as a solution in DMSO in small excess (1.1 eq.) and the incubated protein was then washed several times with buffer in order to remove any excess ligand. The protein was then concentrated and subjected to ^1H - ^{15}N HSQC experiments that confirmed the complete loading and the structural integrity of the protein.

4.4 Measurement of ^{19}F PCS and Localization of the Ligand Using a Monte-Carlo Approach

^{19}F PCS were observed in simple one-dimensional ^{19}F experiments for all ligand-containing protein constructs (Fig. 27). Due to the strongly paramagnetic LCTs, sizeable PCS of up to 0.4 ppm were detected, ensuring that only a minimal error is included in the PCS that serves as raw data input for the fitting process of the ligand position. The linewidths of the fluorine signals in the range of 20–70 Hz indicate clearly residual mobility of the phenyl ring of the ligand that points out of the enzyme's pocket, since comparable strongly binding carazolol ligands showed linewidths of 220–230 Hz in a study of Eddy et al.¹⁷⁶ and the Debye-Waller factors of the PDB 1G54 agree with this observation.¹⁷⁷ As for the amide N—H spin pairs, the largest PCSs were observed for the Tm-tagged hCA II S166C constructs, most likely due to the suppressed flexibility of the LCT attached to this mutant when compared to the other attachment sites.¹⁴⁵

By using the obtained anisotropy parameters from the fitting to the backbone amides of the hCA II constructs, the expected ^{19}F PCS can be back-calculated and an estimate of the position of the ligand can be achieved. The expected PCS for given anisotropy parameters in relation to the distance and angular position of the fluorine atom of the ligand can be calculated using the following equation:

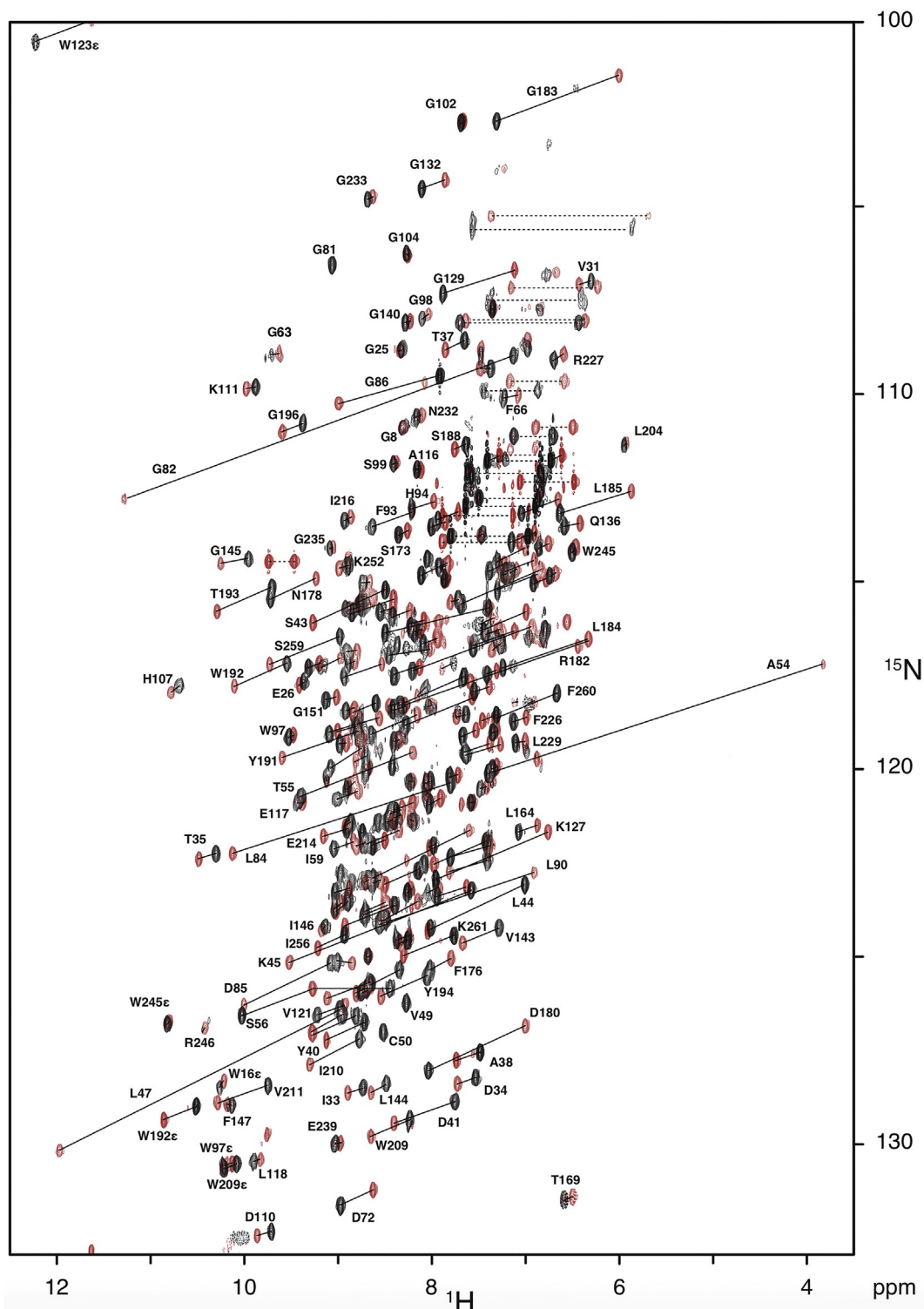


Fig. 23 Overlay of the ^1H - ^{15}N HSQC spectra of uniformly ^{15}N labeled hCA II S50C Lu-DOTA-M8-SSPy (*black*) and hCA II S50C Tm-DOTA-M8-SSPy (*red*) recorded at 600 MHz proton frequency (Aliased signals with negative intensity are shown with dashed contour lines, PCS are indicated with solid and NH_2 groups with dashed lines, water signal was omitted for clarity). Due to the highly crowded overlay of the spectra, an ambiguous result could be obtained upon direct assignment of all PCS.⁴¹ Reproduced by permission of The Royal Society of Chemistry from Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.

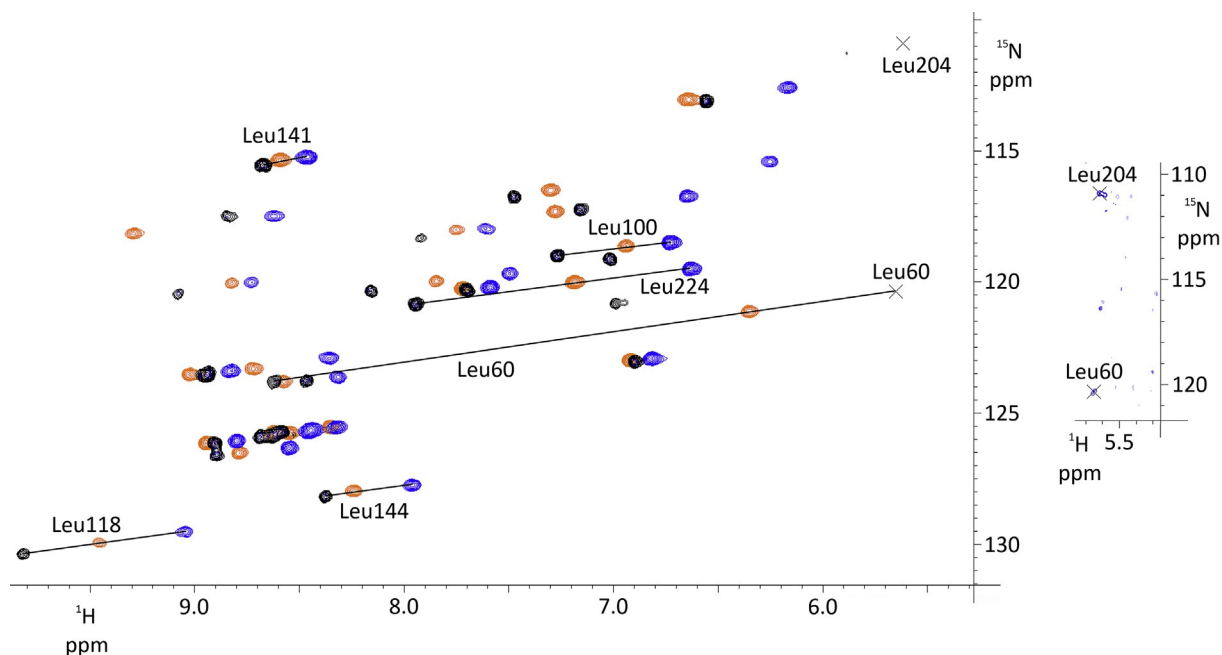


Fig. 24 Overlay of ^1H - ^{15}N HSQC spectra of selectively ^{15}N leucine labeled hCA II S166C labeled with Tm-DOTA-M8-(4*R*,4*S*)-SSPy³⁴ (orange), Tm-P4M4-DOTA (blue) and Lu-P4M4-DOTA (black) measured at 600 MHz proton frequency (left). Remarkably, due to the low spectral complexity, the PCSs can be readily assigned and leucine-derived anisotropy parameters can be fitted. Presence of peaks affected in intensity due to the applied water suppression method is demonstrated in an extra panel (right).³⁶

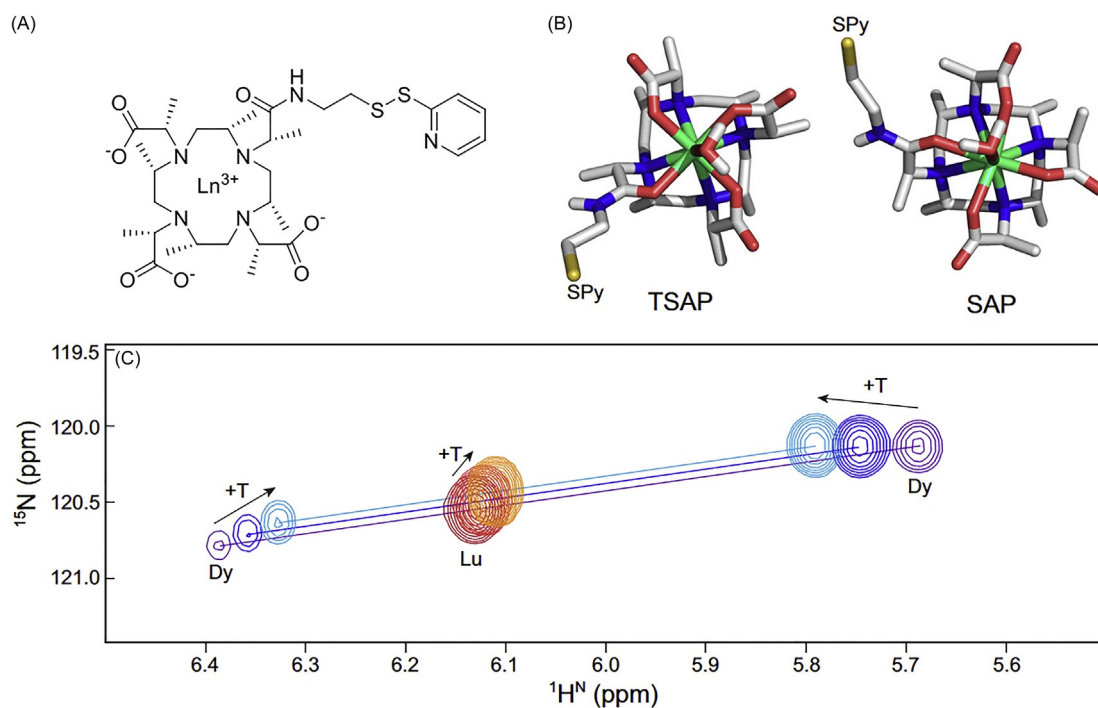


Fig. 25 Influence of different temperatures on shifts induced by the Dy-DOTA-M8-(8*S*)-SSPy LCT²⁹ on an ubiquitin S57C construct (structure of the LCT depicted in part A).⁷⁰ The major and minor species originate in two conformational isomers (B) that interconvert on a slow timescale (twisted square antiprism (TSAP) and square antiprism (SAP), see also Section 2.6). The observed temperature dependence of the induced PCS, i.e. the movement of the paramagnetic shift to the diamagnetic reference peak upon increase of the measuring temperature, can be used in order to unambiguously assign PCSs in a crowded spectral region of a ^1H - ^{15}N HSQC experiment (C, 288 K: purple, 298 K: blue, 308 K: cyan).⁷⁰ Strickland, M.; Schwieters, C. D.; Göbl, C.; Opina, A. C. L.; Strub, M.-P.; Swenson, R. E.; Vasalatiy, O.; Tjandra, N., *J. Biomol. NMR* **2016**, *66*, 125–139.

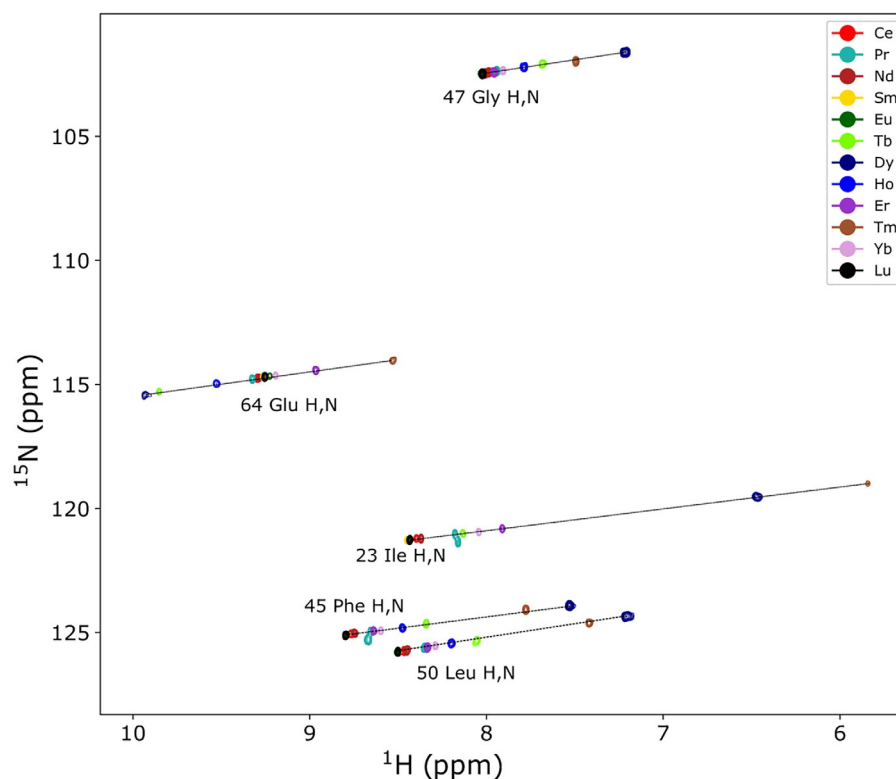


Fig. 26 Spectral overlay of a selection of peaks in ^1H - ^{15}N HSQC experiments of Ln-DOTA-M7FPy,³⁵ loaded with various lanthanoid ions, attached to an ubiquitin S57C construct. The measurement of multiple lanthanoid ions allows to critically assess the PCS assignment in crowded regions of a ^1H - ^{15}N HSQC spectrum by comparing the induced patterns of shifts.

Table 1 Anisotropy parameters for the hCA II constructs tagged with Tm-DOTA-M8-(8S)-SSPy.

Parameter	Unit	S50C Tm-DOTA	S166C Tm-DOTA	S217C Tm-DOTA	S220C Tm-DOTA	S166C Tm-Thiazole
No. of PCS	–	366	397	364	366	44
$\Delta\chi_{\text{ax}}$	$[10^{-32} \text{ m}^3]$	21.6 ± 1.2	38.5 ± 2.0	25.7 ± 1.0	23.6 ± 0.9	34.7 ± 0.6
$\Delta\chi_{\text{rh}}$	$[10^{-32} \text{ m}^3]$	8.5 ± 0.7	8.0 ± 1.0	13.2 ± 0.6	4.3 ± 0.3	13.3 ± 1.1
x	$[\text{\AA}]$	-27.8 ± 0.3	-16.3 ± 0.4	-24.9 ± 0.2	-13.1 ± 0.3	-11.8 ± 0.5
y	$[\text{\AA}]$	13.7 ± 0.3	-3.6 ± 0.4	-17.7 ± 0.3	-26.7 ± 0.3	-1.7 ± 0.1
z	$[\text{\AA}]$	18.1 ± 0.3	-11.2 ± 0.4	19.6 ± 0.2	3.2 ± 0.2	-11.0 ± 0.2
α	$[\text{\textcircled{C}}]$	104.1 ± 1.6	52.2 ± 1.8	143.7 ± 0.8	14.9 ± 1.4	119.4 ± 4.4
β	$[\text{\textcircled{C}}]$	142.3 ± 1.1	123.6 ± 1.4	70.9 ± 0.5	153.6 ± 0.6	162.1 ± 0.5
γ	$[\text{\textcircled{C}}]$	116.2 ± 1.7	140.3 ± 5.6	125.5 ± 1.0	1.0 ± 2.6	44.8 ± 3.7
Q		0.072	0.037	0.064	0.106	0.061
Monte-Carlo structure variation with $\sigma = 0.5 \text{ \AA}$.						$\sigma = 0.05 \text{ \AA}$
$\Delta\chi_{\text{ax}}$	$[10^{-32} \text{ m}^3]$	21.1 ± 0.9	37.4 ± 0.9	25.5 ± 1.1	23.0 ± 2.1	34.3 ± 0.8
$\Delta\chi_{\text{rh}}$	$[10^{-32} \text{ m}^3]$	8.5 ± 0.6	7.8 ± 0.6	13.1 ± 0.6	4.4 ± 0.8	12.7 ± 1.6
x	$[\text{\AA}]$	-27.5 ± 0.3	-16.2 ± 0.2	-24.8 ± 0.2	-13.0 ± 0.6	-12.0 ± 0.7
y	$[\text{\AA}]$	13.6 ± 0.2	-3.6 ± 0.2	-17.5 ± 0.3	-26.4 ± 0.6	-1.8 ± 0.2
z	$[\text{\AA}]$	18.2 ± 0.2	-11.0 ± 0.2	19.6 ± 0.2	3.2 ± 0.5	-10.8 ± 0.4
α	$[\text{\textcircled{C}}]$	104.0 ± 1.8	52.7 ± 1.2	143.3 ± 0.8	16.2 ± 3.5	121.6 ± 6.8
β	$[\text{\textcircled{C}}]$	141.8 ± 0.8	123.1 ± 0.7	71.2 ± 0.7	153.7 ± 1.5	161.6 ± 1.1
γ	$[\text{\textcircled{C}}]$	115.9 ± 1.7	141.1 ± 2.7	125.1 ± 1.1	3.9 ± 6.6	45.1 ± 5.2
Q		0.071	0.035	0.068	0.106	0.059

Monte-Carlo protocol where random subsets consisting of 20% of the available PCS were used.

Reproduced by permission of The Royal Society of Chemistry from Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.

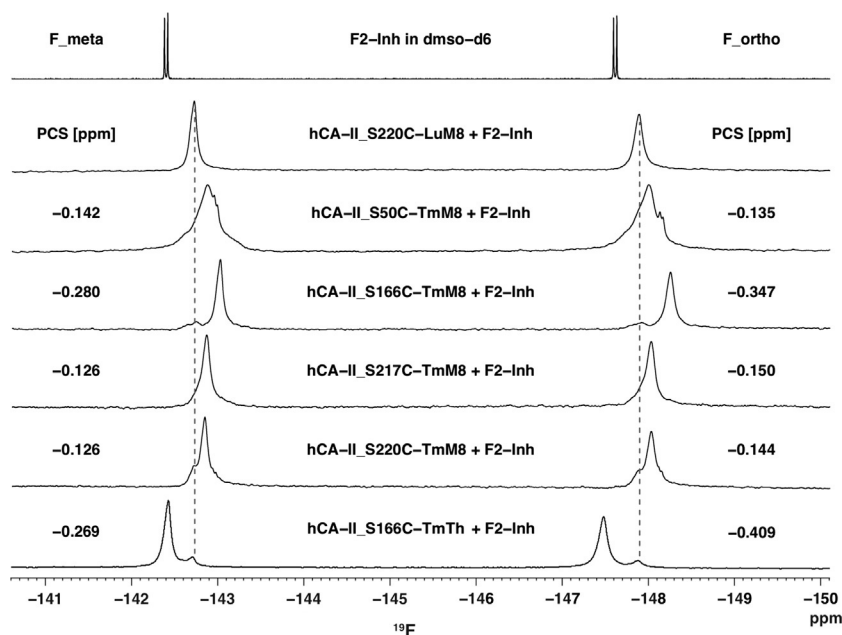


Fig. 27 One-dimensional ^{19}F NMR spectra of the difluorinated ligand F2-Inh incorporated into hCA II constructs that were labeled at different sites within the protein. The observed peaks of the ligand incorporated in the different hCA II constructs are significantly broadened when compared with the unbound ligand due to the higher rotational correlation time. Since the different hCA II constructs labeled with a diamagnetic LCT did not exhibit any differences, a hCA II S220C construct labeled with Lu-DOTA-M8-(8S)-SSPy was used as reference. The one-dimensional ^{19}F NMR spectra were referenced using an internal signal of trifluoroacetate at -79.0 ppm. Reproduced by permission of The Royal Society of Chemistry from Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.

$$\delta^{PCS} = \frac{1}{12\pi r^3} \left[\Delta\chi_{ax} (3 \cos^2 \theta - 1) + \frac{3}{2} \Delta\chi_{rh} \sin^2 \theta \cos 2\phi \right]$$

The position of the observed spin characterized by the distance to the metal center r as well as the azimuth (ϕ) and zenith (θ) angles leads then to the calculated PCS by taking into account the axial and rhombic anisotropy ($\Delta\chi_{ax}$ and $\Delta\chi_{rh}$) parameters.

In order to calculate the position of the fluorine atom(s) of the chosen sulfonamide ligands, the fitted anisotropy tensors based on the PCS of the amide N—H spin pairs can be used to back-calculate the position of the ligand by taking as input the ^{19}F PCS measured on the ligand while being complexed with the tagged protein. In order to obtain the position in an unambiguous fashion by using least-squares fitting, four different anisotropy tensors were used in the study. Furthermore, in order to further enhance the accuracy, a fifth anisotropy tensor with an additional orthogonal isosurface can be incorporated during the least-squares minimization. The square residual that is minimized during the process is given by the following definition:

$$s(x, y, z) = \sum_{i=1}^n \left(\delta_i^{PCS}(x'_i, y'_i, z'_i) - \delta_i^{PCS}(\text{experiment}) \right)^2$$

The definition is composed by the sum of the deviation between the experimentally measured PCS and the back-calculated PCS, whereas i is the corresponding tensor index and $\delta_i^{PCS}(\text{experiment})$ is the experimental PCS. The back-calculated PCS is given by the semi-empirical formula of Kurland and McGarvey¹⁷⁸ for the PCS in Cartesian coordinates that incorporates the positioning of the spin (x, y, z) , the distance from the spin to the paramagnetic center as well as the anisotropy parameters $\Delta\chi_{ax,i}$ and $\Delta\chi_{rh,i}$ which give the axiality and rhombicity of the earlier determined anisotropy tensor.

$$\delta_i^{PCS}(x'_i, y'_i, z'_i) = \frac{1}{12\pi r_i^3} \left[\Delta\chi_{ax,i} \frac{2(z'_i)^2 - (x'_i)^2 - (y'_i)^2}{r_i^2} + \frac{3}{2} \Delta\chi_{rh,i} \frac{(x'_i)^2 - (y'_i)^2}{r_i^2} \right]$$

The difference between back-calculated and measured PCS, i.e. the function $s(x, y, z)$, leads then straightforward to the position of the fluorine nuclei of the tested ligand. However, one has further to account for the uncertainties of the used anisotropy tensors. Therefore, a Monte-Carlo protocol was implemented in the algorithm that introduces errors based on a random seed and performs 10,000 iterations in order to have a statistically valid sampling number.¹⁷⁹ The introduced uncertainties are based on the errors of the tensor values determined during the fitting process using Numbat.⁹⁷

By application of the described protocol in our model system, the tested inhibitors were successfully localized using PCS restraints with an accuracy of up to 0.8 \AA . Furthermore, the inhibitors were localized over a through-space distance of $22\text{--}38 \text{ \AA}$, a previously unprecedented distance range. Earlier studies reported localizations of ligands over distances of $10\text{--}25 \text{ \AA}$,^{40,102,151} results

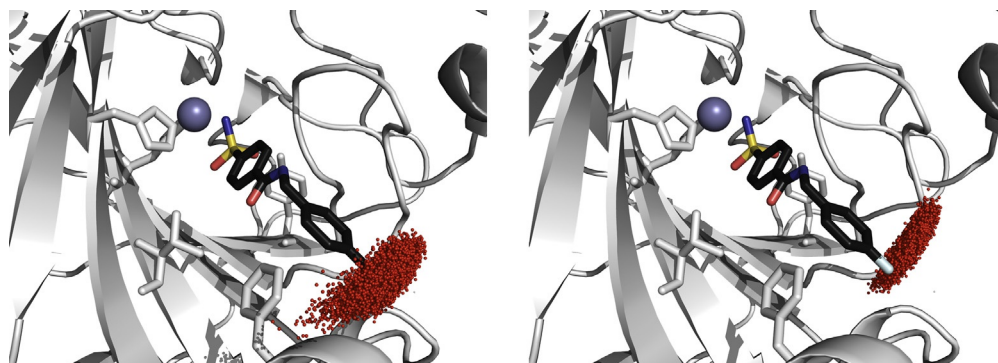


Fig. 28 Point cloud obtained in the Monte-Carlo calculation of the position of the fluorine atoms of FM-520 (*left*) and FM-519 (*right*). Light blue sticks indicate fluorine atoms, while the blue sphere represents the zinc(II) ion (modified PDB 1G54).^{41,177} Reproduced by permission of The Royal Society of Chemistry from Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.

that were already considered as highly promising given that for NOESY a distance limitation of $\sim 5\text{--}8\text{ \AA}$ applies and a complete sidechain assignment is required. The key factor for the distance range that was monitored in the present study lies in the use of the strongly paramagnetic LCTs that lead to large PCS detected for the ligand atoms. With the newest generations of strongly paramagnetic LCTs available today, distances over 200 \AA can be sampled in the optimal case, i.e. given a favorable alignment of the induced anisotropy with the protein.¹⁸ For proteins with a mainly globular fold, the newest generation of LCTs would even allow to localize ligands in proteins over 500 kDa , which includes most proteins suitable for NMR spectroscopy by today.

The determined position for the ligands FM-519 and FM-520 shows deviations of 3.3 and 0.8 \AA when compared to the X-ray structure of hCA II incorporating a pentafluorobenzyl-substituted ligand (PDB 1G54)¹⁷⁷ (Fig. 28). Interestingly, the found deviations of the position found for the ligand F2-Inh by PCS and in the X-ray structure (PDB 1G52)¹⁷⁷ deviate more strongly, although the intramolecular distance between the two fluorine atoms in *ortho*- and *meta* position of the phenyl ring of the ligand is very accurately reproduced (determined as 3.5 \AA compared to 2.8 \AA in the X-ray structure (modified PDB 1G52)¹⁷⁷). However, based on our results and MP2 calculations that investigate the interactions of differently fluorinated benzyl rings with another nonfluorinated benzyl ring (Motif 1b on page 2 of Ref. ¹⁸⁰), it was proposed that the outer aromatic ring of the ligand aligns in solution different from the conformation found in the crystal structure due to the interactions with the neighboring phenyl substituent of Phe-131. By taking into account the possible structural changes of the outer phenyl ring of the ligand in solution, the positional deviations coincided within 1.6 \AA with the fluorine in *meta* position of the difluorophenyl substituent and 2.6 \AA for the fluorine in *ortho* position.

4.5 Determination of the Ligand Position by Analysis of the Isosurfaces of the Anisotropy Parameters

In principle, the position of the investigated ligand can very simply be obtained by using only the intersection of three back-calculated isosurfaces for a specific PCS based on the respective previously determined anisotropy tensors. This approach wins clearly over the Monte-Carlo method, which incorporates more than three anisotropy tensors, by its pure straightforwardness. However, by taking in account only three anisotropy tensors, the result could be corroborated due to an unfavorable alignment and thus not sufficient orthogonality of the used isosurfaces. In order to investigate the described approach and quantify its deviations and pitfalls, the position was determined for all three applied ligands in the study with all possible combinations of tensor isosurfaces. The orthogonality of tensors was determined by the intersection angle of the normal vectors at the intersection point.

The obtained results clearly show that the closer the intersection angles match 90 degrees, the more precise the position can be determined. In order to reflect the found intersection angles and the suitability of a given combination of isosurfaces, an angle score was defined by calculating the sum of the deviations of the intersection angles to 90 degrees and dividing the result by 3. Therefore, an angle score of 0 degree states perfect orthogonality of the used isosurfaces, while 90 degrees indicates parallel isosurfaces. We found that the combination of isosurfaces with an angle score below 30 degrees closely match the positions determined using the Monte-Carlo protocol (average deviation for combinations of isosurfaces with an angle score below 30 degrees of all four ligands: 2.3 \AA) (Fig. 29). For combinations exhibiting an angle score of over 40 degrees, the deviations amount to 7 \AA or more. Therefore, it can be concluded that the use of only three different isosurfaces can be used to localize the ligand given that the used isosurfaces show a favorable angle score. Even for combinations of isosurfaces with an unfavorable angle score, the achieved accuracy would still allow a coarse-grained localization of the ligand and thus enable to discriminate two different, distant binding sites for the ligands within the investigated protein.

4.6 ^1H - ^{19}F -HOESY Experiments for Determination of ^1H PCS of the Ligand and Protein–Ligand Interactions

Since it is impossible to discriminate the proton signals of the bound ligand and the protein host present in equimolar ratios within the sample, ^1H - ^{19}F heteronuclear Overhauser effect spectroscopy (HOESY) experiments were measured in order to determine the ^1H



Fig. 29 Points of intersection of isosurfaces induced by LCTs on different tagging sites (ligand FM-520, center of gravity of the CF₃ fluorine atoms: cyan (modified PDB 1G54),¹⁷⁷ fluorine atoms: light blue sticks, position obtained by least square minimization using all four tensors: gold, positions obtained by intersection of three isosurfaces in all possible combinations with an angle score below 30: orange, positions obtained by intersection of three isosurfaces in all possible combinations with an angle score above 30: red).⁴¹ Reproduced by permission of The Royal Society of Chemistry from Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.

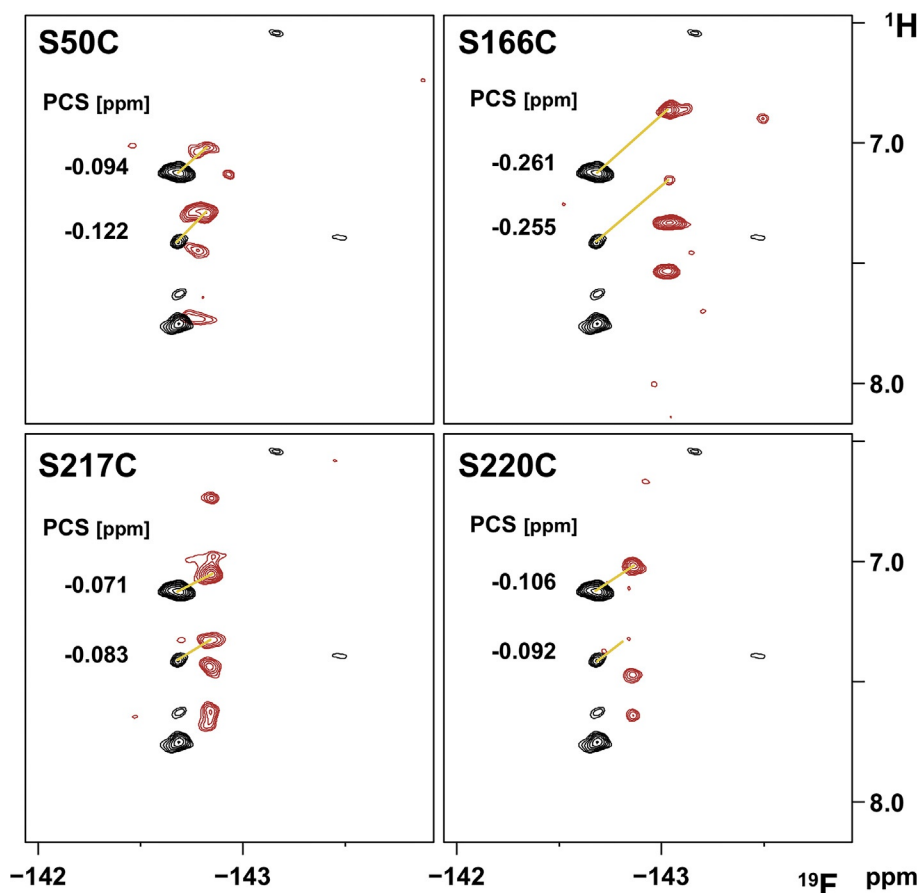


Fig. 30 HOESY spectra of hCA II constructs incorporating the ligand F2-Inh. The correlation peaks represent through-space interactions of ¹⁹F with ¹H. Due to the narrow sweep width, the signals for aliphatic protons appear at their aliased frequency.¹⁸¹ Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D., *Chem. Sci.* **2019**, *10*, 5064–5072.

PCS of proton atoms of the bound ligand (F2-Inh), which are close in space to the fluorine atoms. The experiments revealed four correlation peaks with variable intensities (Fig. 30). Due to the small sweep width, the peaks appear at their aliased frequency. In order to analyze their original NMR shift, spectra with a larger spectral window were recorded. The obtained results show that three peaks originate from correlation of fluorine with sidechain methyl groups, since their proton frequency was found in the aliphatic region between 0.4 and 0.75 ppm. The strongest ¹H-¹⁹F HOESY correlation peak was considered to constitute the cross peak of the proton in *para*-position of the fluorinated phenyl ring of the ligand with the fluorine in *ortho*- and *meta* position.

Upon determination of the ^1H PCS, the aromatic proton atom of F2-Inh in *para* position was localized within the protein by fitting of its experimentally determined PCS to the back-calculated PCS based on the anisotropy tensors derived from the H—N signals. The positional deviation of the proton, based on the PCS approach, amounted to only 1.4 Å when compared with the X-ray structure of hCA II incorporating the difluorinated ligand F2-Inh (PDB 1G52)¹⁷⁷. Interactions of the benzylic fluorine atoms with side-chain methyl groups are clearly observed and the resonance at 0.41 ppm was, based on its observed PCS, attributed to interactions with the methyl groups of the residue Val-135.

5 Conclusion and Outlook

To conclude, the strongly paramagnetic LCTs that are available today yield accurate restraints for the structural elucidation of biomacromolecules in solution, i.e. *in vitro* as well as *in cellulo*. PCS analysis offers the possibility to investigate the three-dimensional structure, conformations and dynamics of biomacromolecules. Furthermore, the site of ligand-binding and the structure of the protein–ligand complex can be resolved by the structural restraints obtained by application of an LCT.

As demonstrated by various research groups, nuclei as ^1H , ^{13}C and ^{19}F in combination with different NMR approaches, e.g. conventional ^1H or ^{19}F one-dimensional spectra, ^{19}F CEST and ^1H relaxation dispersion, can be applied in order to achieve a reasonable signal discrimination of the ligand signals from the protein. The chosen approach depends on the measured nucleus as well as on the affinity of the ligand, i.e. ligands that show weak, intermediate or strong binding. The distance range that can be covered according to the reported literature (10–38 Å) renders this unique method highly suitable for biologically relevant applications. This can be achieved by the application of covalently binding and strongly paramagnetic LCTs to the protein that generate large PCSs and provide therefore accurate structural restraints for the measured nuclei.

By use of human carbonic anhydrase II and incorporated sulfonamide ligands, which interact with nM to pM K_d -values with the zinc metal center, it was shown that it is possible to unambiguously localize a ligand bound ^{19}F atom within a protein approaching an accuracy of 0.8 Å. Solely based on the fluorine PCS caused by LCTs over a distance of 22–38 Å by measuring 4–5 one-dimensional ^{19}F spectra per ligand, one unique solution for the position of the fluorine atoms was obtained. Besides Monte-Carlo calculations with 4 or 5 different anisotropy tensors, graphical analysis using only three isosurfaces shows a high accuracy as long as sufficient orthogonality of the isosurfaces is confirmed by application of an angle score. When only combinations are taken into account showing an angle score below 30 degrees, the obtained accuracy allows a clear localization of the fluorine atom and the position matches the one determined in Monte-Carlo protocols using four tensors. By omitting the need for solvent suppression, chemical modification of the ligand and extensive measurement times during the screening process, the method constitutes a fast, reliable and convenient approach to screen a high number of fluorine-containing ligands for a specific protein.

The use of LCTs for the characterization of protein–ligand complexes in solution can deliver accurate structural restraints over a large distance and thus holds great promise for future applications in structural biology and rational drug design. With the newest generation of LCTs in hand, the localization of ligands by PCS can further be extended to proteins with molecular weights significantly larger than 30 kDa and large, multimeric protein complexes.

Acknowledgment

The Chemistry Department of the University of Basel and the Swiss National Science Foundation grant 200021_130263 are acknowledged for financial support. Biological structures were generated using the open source software PyMOL (<http://www.pymol.org/>).

References

1. Wüthrich, K. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 3340–3363.
2. Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366–12371.
3. Salzmänn, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13585–13590.
4. Salzmänn, M.; Wider, G.; Pervushin, K.; Senn, H.; Wüthrich, K. *J. Am. Chem. Soc.* **1999**, *121*, 844–848.
5. Salzmänn, M.; Wider, G.; Pervushin, K.; Wüthrich, K. *J. Biomol. NMR* **1999**, *15*, 181–184.
6. Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
7. Wider, G.; Macura, S.; Kumar, A.; Ernst, R. R.; Wüthrich, K. *J. Magn. Reson.* **1984**, *56*, 207–234.
8. Ujwal, R.; Cascio, D.; Colletier, J.-P.; Faham, S.; Zhang, J.; Toro, L.; Ping, P.; Abramson, J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17742.
9. Böhm, R.; Amodeo, G. F.; Murlidaran, S.; Chavali, S.; Wagner, G.; Winterhalter, M.; Brannigan, G.; Hiller, S. *Structure* **2019**. <https://doi.org/10.1016/j.str.2019.11.015>.
10. Wüthrich, K. *Nat. Struct. Biol.* **2001**, *8*, 923–925.
11. Marion, D. *Mol. Cell. Proteomics* **2013**, *12*, 3006.
12. Zhu, G.; Xia, Y.; Nicholson, L. K.; Sze, K. H. *J. Magn. Reson.* **2000**, *143*, 423–426.
13. Renner, C.; Schleicher, M.; Moroder, L.; Holak, T. A. *J. Biomol. NMR* **2002**, *23*, 23–33.
14. Wijmenga, S. S.; van Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287–387.
15. Jamin, N.; Toma, F. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *38*, 83–114.
16. Fürtig, B.; Richter, C.; Wöhnert, J.; Schwalbe, H. *ChemBioChem* **2003**, *4*, 936–962.
17. Al-Hashimi, H. M. *J. Magn. Reson.* **2013**, *237*, 191–204.

18. Joss, D.; Häussinger, D. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *114–115*, 284–312.
19. Barry, C. D.; North, A. C. T.; Glasel, J. A.; Williams, R. J. P.; Xavier, A. V. *Nature* **1971**, *232*, 236–245.
20. Bertini, I.; Turano, P.; Vila, A. J. *Chem. Rev.* **1993**, *93*, 2833–2932.
21. Banci, L.; Bertini, I.; Bren, K. L.; Cremonini, M. A.; Gray, H. B.; Luchinat, C.; Turano, P. *J. Biol. Inorg. Chem.* **1996**, *1*, 117–126.
22. Banci, L.; Bertini, I.; Savellini, G. G.; Romagnoli, A.; Turano, P.; Cremonini, M. A.; Luchinat, C.; Gray, H. B. *Proteins Struct. Funct. Bioinf.* **1997**, *29*, 68–76.
23. Bertini, I.; Luchinat, C.; Parigi, G.; Pierattelli, R. *ChemBioChem* **2005**, *6*, 1536–1549.
24. Wöhnert, J.; Franz, K. J.; Nitz, M.; Imperiali, B.; Schwalbe, H. *J. Am. Chem. Soc.* **2003**, *125*, 13338–13339.
25. Franz, K. J.; Nitz, M.; Imperiali, B. *ChemBioChem* **2003**, *4*, 265–271.
26. Nitz, M.; Franz, K. J.; Maglathlin, R. L.; Imperiali, B. *ChemBioChem* **2003**, *4*, 272–276.
27. Saio, T.; Ogura, K.; Yokochi, M.; Kobashigawa, Y.; Inagaki, F. *J. Biomol. NMR* **2009**, *44*, 157–166.
28. Keizers, P. H.; Saragliadis, A.; Hiruma, Y.; Overhand, M.; Ubbink, M. *J. Am. Chem. Soc.* **2008**, *130*, 14802–14812.
29. Häussinger, D.; Huang, J.-R.; Grzesiek, S. *J. Am. Chem. Soc.* **2009**, *131*, 14761–14767.
30. Peters, F.; Maestre-Martinez, M.; Leonov, A.; Kovacic, L.; Becker, S.; Boelens, R.; Griesinger, C. *J. Biomol. NMR* **2011**, *51*, 329–337.
31. Graham, B.; Loh, C. T.; Swarbrick, J. D.; Ung, P.; Shin, J.; Yagi, H.; Jia, X.; Chhabra, S.; Barlow, N.; Pintacuda, G.; Huber, T.; Otting, G. *Bioconjug. Chem.* **2011**, *22*, 2118–2125.
32. Yang, F.; Wang, X.; Pan, B.-B.; Su, X.-C. *Chem. Commun.* **2016**, *52*, 11535–11538.
33. Müntener, T.; Häussinger, D.; Selenko, P.; Theillet, F.-X. *J. Phys. Chem. Lett.* **2016**, *7*, 2821–2825.
34. Joss, D.; Walliser, R. M.; Zimmermann, K.; Häussinger, D. *J. Biomol. NMR* **2018**, *29*, 34–38.
35. Müntener, T.; Kottelat, J.; Huber, A.; Häussinger, D. *Bioconjug. Chem.* **2018**, *29*, 3344–3351.
36. Joss, D.; Bertrams, M.-S.; Häussinger, D. *Chem. Eur. J.* **2019**, *25*, 11910–11917.
37. Joss, D.; Häussinger, D. *Chem. Commun.* **2019**, *55*, 10543–10546.
38. Bertrand, J. A.; Auger, G.; Martin, L.; Fanchon, E.; Blanot, D.; Le Beller, D.; van Heijenoort, J.; Dideberg, O. *J. Mol. Biol.* **1999**, *289*, 579–590.
39. Saio, T.; Ogura, K.; Kumeta, H.; Kobashigawa, Y.; Shimizu, K.; Yokochi, M.; Kodama, K.; Yamaguchi, H.; Tsujishita, H.; Inagaki, F. *Sci. Rep.* **2015**, *5*, 16685.
40. Guan, J.-Y.; Keizers, P. H. J.; Liu, W.-M.; Löhr, F.; Skinner, S. P.; Heeneman, E. A.; Schwalbe, H.; Ubbink, M.; Siegal, G. *J. Am. Chem. Soc.* **2013**, *135*, 5859–5868.
41. Zimmermann, K.; Joss, D.; Müntener, T.; Nogueira, E. S.; Schäfer, M.; Knörr, L.; Monnard, F. W.; Häussinger, D. *Chem. Sci.* **2019**, *10*, 5064–5072.
42. Strickland, M.; Catazarro, J.; Rajasekaran, R.; Strub, M.-P.; O'Hern, C.; Bermejo, G. A.; Summers, M. F.; Marchant, J.; Tjandra, N. *J. Am. Chem. Soc.* **2019**, *141*, 1430–1434.
43. Göbl, C.; Resch, M.; Strickland, M.; Hartmüller, C.; Viertler, M.; Tjandra, N.; Madl, T. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 14847–14851.
44. Brewer, K. D.; Bacaj, T.; Cavalli, A.; Camilloni, C.; Swarbrick, J. D.; Liu, J.; Zhou, A.; Zhou, P.; Barlow, N.; Xu, J.; Seven, A. B.; Prinslow, E. A.; Voleti, R.; Häussinger, D.; Bonvin, A. M. J. J.; Tomchick, D. R.; Vendruscolo, M.; Graham, B.; Südhof, T. C.; Rizo, J. *Nat. Struct. Mol. Biol.* **2015**, *22*, 555–564.
45. Canales, A.; Boos, I.; Perkams, L.; Karst, L.; Lubert, T.; Karagiannis, T.; Domínguez, G.; Cañada, F. J.; Pérez-Castells, J.; Häussinger, D.; Unverzagt, C.; Jiménez-Barbero, J. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 14987–14991.
46. Rodríguez-Castañeda, F.; Haberz, P.; Leonov, A.; Griesinger, C. *Magn. Reson. Chem.* **2006**, *44*, S10–S16.
47. Pintacuda, G.; John, M.; Su, X.-C.; Otting, G. *Acc. Chem. Res.* **2007**, *40*, 206–212.
48. John, M.; Otting, G. *ChemPhysChem* **2007**, *8*, 2309–2313.
49. Otting, G. *J. Biomol. NMR* **2008**, *42*, 1–9.
50. Otting, G. *Annu. Rev. Biophys.* **2010**, *39*, 387–405.
51. Su, X.-C.; Otting, G. *J. Biomol. NMR* **2010**, *46*, 101–112.
52. Allen, K. N.; Imperiali, B. *Curr. Opin. Chem. Biol.* **2010**, *14*, 247–254.
53. Koehler, J.; Meiler, J. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 360–389.
54. Keizers, P. H. J.; Ubbink, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *58*, 88–96.
55. Bhaumik, A.; Luchinat, C.; Parigi, G.; Ravera, E.; Rinaldelli, M. *CrystEngComm* **2013**, *15*, 8639–8656.
56. Hass, M. A. S.; Ubbink, M. *Curr. Opin. Struct. Biol.* **2014**, *24*, 45–53.
57. Prestegard, J. H.; Agard, D. A.; Moremen, K. W.; Lavery, L. A.; Morris, L. C.; Pederson, K. *J. Magn. Reson.* **2014**, *241*, 32–40.
58. Göbl, C.; Madl, T.; Simon, B.; Sattler, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *80*, 26–63.
59. Liu, W.-M.; Overhand, M.; Ubbink, M. *Coord. Chem. Rev.* **2014**, *273–274*, 2–12.
60. Kato, K.; Yamaguchi, T. *Glycoconjugate J.* **2015**, *32*, 505–513.
61. Carlon, A.; Ravera, E.; Andralojć, W.; Parigi, G.; Murshudov, G. N.; Luchinat, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *92–93*, 54–70.
62. Nitsche, C.; Otting, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *98–99*, 20–49.
63. Su, X.-C.; Chen, J.-L. *Acc. Chem. Res.* **2019**, *52*, 1675–1686.
64. Saio, T.; Ishimori, K. *Biochim. Biophys. Acta. Gen. Subj.* **2020**. <https://doi.org/10.1016/j.bbagen.2019.03.018>. in press.
65. Su, X.-C.; Man, B.; Beeren, S.; Liang, H.; Simonsen, S.; Schmitz, C.; Huber, T.; Messerle, B. A.; Otting, G. *J. Am. Chem. Soc.* **2008**, *130*, 10486–10487.
66. Su, X.-C.; Liang, H.; Loscha, K. V.; Otting, G. *J. Am. Chem. Soc.* **2009**, *131*, 10352–10353.
67. Loh, C. T.; Ozawa, K.; Tuck, K. L.; Barlow, N.; Huber, T.; Otting, G.; Graham, B. *Bioconjug. Chem.* **2013**, *24*, 260–268.
68. Lee, M. D.; Loh, C. T.; Shin, J.; Chhabra, S.; Dennis, M. L.; Otting, G.; Swarbrick, J. D.; Graham, B. *Chem. Sci.* **2015**, *6*, 2614–2624.
69. Pan, B.-B.; Yang, F.; Ye, Y.; Wu, Q.; Li, C.; Huber, T.; Su, X.-C. *Chem. Commun.* **2016**, *52*, 10237–10240.
70. Strickland, M.; Schwieters, C. D.; Göbl, C.; Opina, A. C. L.; Strub, M.-P.; Swenson, R. E.; Vasalatiy, O.; Tjandra, N. *J. Biomol. NMR* **2016**, *66*, 125–139.
71. Pearce, B. J. G.; Jabar, S.; Loh, C.-T.; Szabo, M.; Graham, B.; Otting, G. *J. Biomol. NMR* **2017**, *68*, 19–32.
72. Wu, Z.; Lee, M. D.; Carruthers, T. J.; Szabo, M.; Dennis, M. L.; Swarbrick, J. D.; Graham, B.; Otting, G. *Bioconjug. Chem.* **2017**, *28*, 1741–1748.
73. Barbieri, R.; Bertini, I.; Cavallaro, G.; Lee, Y.-M.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **2002**, *124*, 5581–5587.
74. Kamen, D. E.; Cahill, S. M.; Girvin, M. E. *J. Am. Chem. Soc.* **2007**, *129*, 1846–1847.
75. Xu, X.; Keizers, P. H. J.; Reinle, W.; Hannemann, F.; Bernhardt, R.; Ubbink, M. *J. Biomol. NMR* **2009**, *43*, 247–254.
76. Chiliveri, S. C.; Louis, J. M.; Ghirlando, R.; Baber, J. L.; Bax, A. *J. Am. Chem. Soc.* **2018**, *140*, 34–37.
77. Barnes, C. A.; Shen, Y.; Ying, J.; Takagi, Y.; Torchia, D. A.; Sellers, J. R.; Bax, A. *J. Am. Chem. Soc.* **2019**, *141*, 9004–9017.
78. Gillespie, J. R.; Shortle, D. *J. Mol. Biol.* **1997**, *268*, 170–184.
79. Donaldson, L. W.; Skrynnikov, N. R.; Choy, W.-Y.; Muhandiram, D. R.; Sarkar, B.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 9843–9847.
80. Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 5879–5896.
81. Liang, B.; Bushweller, J. H.; Tamm, L. K. *J. Am. Chem. Soc.* **2006**, *128*, 4389–4397.
82. Iwahara, J.; Clore, G. M. *Nature* **2006**, *440*, 1227–1230.
83. Iwahara, J.; Tang, C.; Marius Clore, G. *J. Magn. Reson.* **2007**, *184*, 185–195.
84. Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108–4139.
85. Gansow, O. A.; Loeffler, P. A.; Davis, R. E.; Lenkinski, R. E.; Wilcott, M. R. *J. Am. Chem. Soc.* **1976**, *98*, 4250–4258.
86. Sherry, A. D.; Ren, J.; Huskens, J.; Brücher, E.; Tóth, É.; Gerales, C. F. C. G.; Castro, M. M. C. A.; Cacheris, W. P. *Inorg. Chem.* **1996**, *35*, 4604–4612.
87. Fonda, K. K.; Vallarino, L. M. *Inorg. Chim. Acta* **2002**, *334*, 403–410.

88. Di Pietro, S.; Piano, S. L.; Di Bari, L. *Coord. Chem. Rev.* **2011**, *255*, 2810–2820.
89. Cornilescu, G.; Bax, A. J. *Am. Chem. Soc.* **2000**, *122*, 10143–10154.
90. John, M.; Park, A. Y.; Pintacuda, G.; Dixon, N. E.; Otting, G. *J. Am. Chem. Soc.* **2005**, *127*, 17190–17191.
91. Orton, H. W.; Kuprov, I.; Loh, C.-T.; Otting, G. *J. Phys. Chem. Lett.* **2016**, *7*, 4815–4818.
92. Angyal, S. J.; Greeves, D.; Pickles, V. A. *J. Chem. Soc. Chem. Commun.* **1974**, 589–590.
93. Rodríguez-Rodríguez, A.; Esteban-Gómez, D.; de Blas, A.; Rodríguez-Blas, T.; Botta, M.; Tripiet, R.; Platas-Iglesias, C. *Inorg. Chem.* **2012**, *51*, 13419–13429.
94. Chen, K.; Tjandra, N. *Top. Curr. Chem.* **2012**, *326*, 47–67.
95. Hikone, Y.; Hirai, G.; Mishima, M.; Inomata, K.; Ikeya, T.; Arai, S.; Shirakawa, M.; Sodeoka, M.; Ito, Y. *J. Biomol. NMR* **2016**, *66*, 99–110.
96. Bertini, I.; Luchinat, C.; Parigi, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 249–273.
97. Schmitz, C.; Stanton-Cook, M. J.; Su, X. C.; Otting, G.; Huber, T. *J. Biomol. NMR* **2008**, *41*, 179–189.
98. Orton, H. W.; Huber, T.; Otting, G. *Magn. Reson. Discuss.* **2019**, 1–18.
99. Saio, T.; Yokochi, M.; Kumeta, H.; Inagaki, F. *J. Biomol. NMR* **2010**, *46*, 271–280.
100. Lescanne, M.; Skinner, S. P.; Blok, A.; Timmer, M.; Cerofolini, L.; Fragai, M.; Luchinat, C.; Ubbink, M. *J. Biomol. NMR* **2017**, *69*, 183–195.
101. Lescanne, M.; Ahuja, P.; Blok, A.; Timmer, M.; Akerud, T.; Ubbink, M. *J. Biomol. NMR* **2018**, *71*, 275–285.
102. Saio, T.; Ogura, K.; Shimizu, K.; Yokochi, M.; Burke, T. R.; Inagaki, F. *J. Biomol. NMR* **2011**, *51*, 395–408.
103. Shannon, R. D. *Acta Crystallogr. A* **1976**, *32*, 751–767.
104. Bertini, I.; Donaire, A.; Jimenez, B.; Luchinat, C.; Parigi, G.; Piccioli, M.; Poggi, L. *J. Biomol. NMR* **2001**, *21*, 85–98.
105. Bertini, I.; Janik, M. B. L.; Lee, Y.-M.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **2001**, *123*, 4181–4188.
106. Mironov, V. S.; Galyametdinov, Y. G.; Ceulemans, A.; Görlner-Walrand, C.; Binnemans, K. *Chem. Phys. Lett.* **2001**, *345*, 132–140.
107. Funk, A. M.; Finney, K.-L. N. A.; Harvey, P.; Kenwright, A. M.; Neil, E. R.; Rogers, N. J.; Kanthi Senanayake, P.; Parker, D. *Chem. Sci.* **2015**, *6*, 1655–1662.
108. Mason, K.; Rogers, N. J.; Sutorina, E. A.; Kuprov, I.; Aguilar, J. A.; Batsanov, A. S.; Yufit, D. S.; Parker, D. *Inorg. Chem.* **2017**, *56*, 4028–4038.
109. Sutorina, E. A.; Mason, K.; Geraldes, C. F. G. C.; Kuprov, I.; Parker, D. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 12215–12218.
110. Harnden, A. C.; Sutorina, E. A.; Batsanov, A. S.; Senanayake, P. K.; Fox, M. A.; Mason, K.; Vonci, M.; McInnes, E. J. L.; Chilton, N. F.; Parker, D. *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 10290–10294.
111. Ma, C.; Opella, S. J. *J. Magn. Reson.* **2000**, *146*, 381–384.
112. Nitz, M.; Sherawat, M.; Franz, K. J.; Peisach, E.; Allen, K. N.; Imperiali, B. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 3682–3685.
113. Su, X.-C.; Huber, T.; Dixon, N. E.; Otting, G. *ChemBioChem* **2006**, *7*, 1599–1604.
114. Su, X.-C.; McAndrew, K.; Huber, T.; Otting, G. *J. Am. Chem. Soc.* **2008**, *130*, 1681–1687.
115. Ikegami, T.; Verdier, L.; Sakhaii, P.; Grimme, S.; Pescatore, B.; Saxena, K.; Fiebig, K. M.; Griesinger, C. *J. Biomol. NMR* **2004**, *29*, 339–349.
116. Pintacuda, G.; Moshref, A.; Leonchiks, A.; Sharipo, A.; Otting, G. *J. Biomol. NMR* **2004**, *29*, 351–361.
117. Prudêncio, M.; Rohovec, J.; Peters, J. A.; Tocheva, E.; Boulanger, M. J.; Murphy, M. E. P.; Hupkes, H.-J.; Kusters, W.; Impagliazzo, A.; Ubbink, M. *Chem. – Eur. J.* **2004**, *10*, 3252–3260.
118. Consortium, T. U. *Nucleic Acids Res.* **2018**, *47*, D506–D515.
119. Yang, Y.; Wang, J.-T.; Pei, Y.-Y.; Su, X.-C. *Chem. Commun.* **2015**, *51*, 2824–2827.
120. Jiang, W.-X.; Gu, X.-H.; Dong, X.; Tang, C. *J. Biomol. NMR* **2017**, *67*, 273–282.
121. Liu, W.-M.; Keizers, P. H. J.; Hass, M. A. S.; Blok, A.; Sarris, A. J. C.; Overhand, M.; Ubbink, M. *J. Am. Chem. Soc.* **2012**, *134*, 17306–17313.
122. Liu, W.-M.; Skinner, S. P.; Timmer, M.; Blok, A.; Hass, M. A. S.; Filippov, D. V.; Overhand, M.; Ubbink, M. *Chem. – Eur. J.* **2014**, *20*, 6256–6258.
123. Lee, M. D.; Dennis, M. L.; Swarbrick, J. D.; Graham, B. *Chem. Commun.* **2016**, *52*, 7954–7957.
124. Lee, M. D.; Dennis, M. L.; Graham, B.; Swarbrick, J. D. *Chem. Commun.* **2017**, *53*, 13205–13208.
125. Dvoretzky, A.; Gaponenko, V.; Rosevear, P. R. *FEBS Lett.* **2002**, *528*, 189–192.
126. Leonov, A.; Voigt, B.; Rodríguez-Castañeda, F.; Sakhaii, P.; Griesinger, C. *Chem. – Eur. J.* **2005**, *11*, 3342–3348.
127. Haberz, P.; Rodríguez-Castañeda, F.; Junker, J.; Becker, S.; Leonov, A.; Griesinger, C. *Org. Lett.* **2006**, *8*, 1275–1278.
128. Parker, D.; Dickins, R. S.; Puschmann, H.; Crossland, C.; Howard, J. A. K. *Chem. Rev.* **2002**, *102*, 1977–2010.
129. Chen, J.-L.; Zhao, Y.; Gong, Y.-J.; Pan, B.-B.; Wang, X. Su, X.-C. *J. Biomol. NMR* **2018**, *70*, 77–92.
130. Man, B.; Su, X.-C.; Liang, H.; Simonsen, S.; Huber, T.; Messerle, B. A.; Otting, G. *Chem. – Eur. J.* **2010**, *16*, 3827–3832.
131. Jia, X.; Maleckis, A.; Huber, T.; Otting, G. *Chem. – Eur. J.* **2011**, *17*, 6830–6836.
132. Li, Q.-F.; Yang, Y.; Maleckis, A.; Otting, G.; Su, X.-C. *Chem. Commun.* **2012**, *48*, 2704–2706.
133. Swarbrick, J. D.; Ung, P.; Chhabra, S.; Graham, B. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 4403–4406.
134. Swarbrick, J. D.; Ung, P.; Su, X.-C.; Maleckis, A.; Chhabra, S.; Huber, T.; Otting, G.; Graham, B. *Chem. Commun.* **2011**, *47*, 7368–7370.
135. Yagi, H.; Maleckis, A.; Otting, G. *J. Biomol. NMR* **2013**, *55*, 157–166.
136. Loh, C.-T.; Graham, B.; Abdelkader, E. H.; Tuck, K. L.; Otting, G. *Chem. – Eur. J.* **2015**, *21*, 5084–5092.
137. Yang, Y.; Li, Q.-F.; Cao, C.; Huang, F.; Su, X.-C. *Chem. – Eur. J.* **2013**, *19*, 1097–1103.
138. Huang, F.; Pei, Y.-Y.; Zuo, H.-H.; Chen, J.-L.; Yang, Y.; Su, X.-C. *Chem. – Eur. J.* **2013**, *19*, 17141–17149.
139. Tweedle, M. F.; Hagan, J. J.; Kumar, K.; Mantha, S.; Chang, C. A. *Magn. Reson. Imaging* **1991**, *9*, 409–415.
140. Maacke, H. R.; Reubi, J. C. *J. Nucl. Med.* **2011**, *52*, 841–844.
141. Stasiuk, G. J.; Long, N. J. *Chem. Commun.* **2013**, *49*, 2732–2746.
142. Opina, A. C. L.; Strickland, M.; Lee, Y.-S.; Tjandra, N.; Byrd, R. A.; Swenson, R. E.; Vasalaty, O. *Dalton Trans.* **2016**, *45*, 4673–4687.
143. Opina, A. C. L.; Strickland, M.; Lee, Y. S.; Tjandra, N.; Swenson, R. E.; Vasalaty, O. *Inorg. Chem.* **2019**, *58*, 15788–15800.
144. Müntener, T.; Thommen, F.; Joss, D.; Kottelat, J.; Prescimone, A.; Häussinger, D. *Chem. Commun.* **2019**, *55*, 4715–4718.
145. Sutorina, E. A.; Häussinger, D.; Zimmermann, K.; Garbuio, L.; Yulikov, M.; Jeschke, G.; Kuprov, I. *Chem. Sci.* **2017**, *8*, 2751–2757.
146. Neumann, C. N.; Hooker, J. M.; Ritter, T. *Nature* **2016**, *534*, 369–373.
147. Vlasie, M. D.; Comuzzi, C.; van den Nieuwendijk, A. M. C. H.; Prudêncio, M.; Overhand, M.; Ubbink, M. *Chem. – Eur. J.* **2007**, *13*, 1715–1723.
148. Keizers, P. H. J.; Desreux, J. F.; Overhand, M.; Ubbink, M. *J. Am. Chem. Soc.* **2007**, *129*, 9292–9293.
149. Brath, U.; Swamy, S. I.; Veiga, A. X.; Tung, C.-C.; Van Petegem, F.; Erdélyi, M. *J. Am. Chem. Soc.* **2015**, *137*, 11391–11398.
150. Gao, J.; Liang, E.; Ma, R.; Li, F.; Liu, Y.; Liu, J.; Jiang, L.; Li, C.; Dai, H.; Wu, J.; Su, X.; He, W.; Ruan, K. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 12982–12986.
151. Xu, D.; Li, B.; Gao, J.; Liu, Z.; Niu, X.; Nshogozo, G.; Zhang, J.; Wu, J.; Su, X.-C.; He, W.; Ma, R.; Yang, D.; Ruan, K. *J. Phys. Chem. Lett.* **2018**, *9*, 3361–3367.
152. Chen, W.-N.; Nitsche, C.; Pilla, K. B.; Graham, B.; Huber, T.; Klein, C. D.; Otting, G. *J. Am. Chem. Soc.* **2016**, *138*, 4539–4546.
153. Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30–36.
154. Zheng, Y. J.; Merz, K. M. *J. Am. Chem. Soc.* **1992**, *114*, 10498–10507.
155. Supuran, C. T. *Nat. Rev. Drug Discov.* **2008**, *7*, 168–181.
156. Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108*, 946–1051.
157. Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins Struct. Funct. Bioinf.* **1988**, *4*, 274–282.
158. Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3467–3474.

159. Schmid, M.; Nogueira, E. S.; Monnard, F. W.; Ward, T. R.; Meuwly, M. *Chem. Sci.* **2012**, *3*, 690–700.
160. Kanamori, K.; Roberts, J. D. *Biochemistry* **1983**, *22*, 2658–2664.
161. Alward, W. L. M. *N. Engl. J. Med.* **1998**, *339*, 1298–1307.
162. Bunzel, H. A.; Garrabou, X.; Pott, M.; Hilvert, D. *Curr. Opin. Struct. Biol.* **2018**, *48*, 149–156.
163. Zeymer, C.; Hilvert, D. *Annu. Rev. Biochem.* **2018**, *87*, 131–157.
164. Arnold, F. H. *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 14420–14426.
165. Monnard, F. W.; Heinisch, T.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *Chem. Commun.* **2011**, *47*, 8238–8240.
166. Monnard, F. W.; Nogueira, E. S.; Heinisch, T.; Schirmer, T.; Ward, T. R. *Chem. Sci.* **2013**, *4*, 3269–3274.
167. Heinisch, T.; Pellizzoni, M.; Dürrenberger, M.; Tinberg, C. E.; Köhler, V.; Klehr, J.; Häussinger, D.; Baker, D.; Ward, T. R. *J. Am. Chem. Soc.* **2015**, *137*, 10414–10419.
168. Rebelein, J. G.; Cotellet, Y.; Garabedian, B.; Ward, T. R. *ACS Catal.* **2019**, *9*, 4173–4178.
169. Bereau, T.; Kramer, C.; Monnard, F. W.; Nogueira, E. S.; Ward, T. R.; Meuwly, M. *J. Phys. Chem. B.* **2013**, *117*, 5460–5471.
170. Benej, M.; Pastorekova, S.; Pastorek, J. Carbonic Anhydrase IX: Regulation and Role in Cancer. In *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*; Frost, S. C., McKenna, R., Eds.; Springer Netherlands: Dordrecht, 2014; pp 199–219.
171. Zhao, J.; Kajetanowicz, A.; Ward, T. R. *Org. Biomol. Chem.* **2015**, *13*, 5652–5655.
172. Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. *Nature* **1958**, *181*, 662–666.
173. Merk, A.; Bartsaghi, A.; Banerjee, S.; Falconieri, V.; Rao, P.; Davis, M. I.; Pragani, R.; Boxer, M. B.; Earl, L. A.; Milne, J. L. S.; Subramaniam, S. *Cell* **2016**, *165*, 1698–1707.
174. Wang, J.; Sanchez-Rosello, M.; Acena, J. L.; del Pozo, C.; Sorochinsky, A. E.; Fustero, S.; Soloshonok, V. A.; Liu, H. *Chem. Rev.* **2014**, *114*, 2432–2506.
175. Avvaru, B. S.; Kim, C. U.; Sippel, K. H.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. *Biochemistry* **2010**, *49*, 249–251.
176. Eddy, M. T.; Didenko, T.; Stevens, R. C.; Wuthrich, K. *Structure* **2016**, *24*, 2190–2197.
177. Kim, C. Y.; Chang, J. S.; Doyon, J. B.; Baird, T. T.; Fierke, C. A.; Jain, A.; Christianson, D. W. *J. Am. Chem. Soc.* **2000**, *122*, 12125–12134.
178. Kurland, R. J.; McGarvey, B. R. *J. Magn. Reson.* **1970**, *2*, 286–301.
179. Oliphant, T. E. *Comput. Sci. Eng.* **2007**, *9*, 10–20.
180. Chandra, P. P.; Jain, A.; Sapse, A. M. *J. Mol. Model.* **2004**, *10*, 1–5.
181. Gómez-Reyes, J. F.; Ariza-Castolo, A. *J. Math. Chem.* **2013**, *51*, 1961–1980.