DOTTORATO DI RICERCA IN
SCIENZE CHIMICHE

CICLO XXXII

COORDINATORE Prof. ANDREA GOTI

Design, Synthesis and Optimization of Paramagnetic Tags for NMR Spectroscopy

Settore Scientifico Disciplinare CHIM/06

Dottorando
Dott. Maxime Denis

_____
(firma)

Tutore
Prof. Cristina Nativi

_____
(firma)

Coordinatore
Prof. Piero Baglioni

_____
(firma)

Anni 2016/2019

Summary

Abstract

Chapter 1: Paramagnetic restraints in NMR and how to generate them

Chapter 2: Synthesis of cross-bridge azamacrocycles

Chapter 3: Design of high-inertness paramagnetic tags

Chapter 4: Synthesis of enantiopures paramagnetic tags

Chapter 5: Application of paramagnetic tagging to the study in solution of *tc*Pex14

Abstract

Paramagnetic restraints have been used in biomolecular NMR for the last three decades in order to elucidate and refine structures but also to characterize protein/ligand interactions. Pseudo-Contact Shifts (PCS) are the most commonly used restraints, consisting in measuring a shift between a paramagnetic species and a diamagnetic reference, whose magnitude depends on the distance from the paramagnetic centre. A common technique to generate such restraints consists in the attachment of lanthanides ions to the protein via a Lanthanide-Binding-Tag (LBT). In order to design such LBTs, it is important to consider the efficiency and stability of the conjugation, the geometry of the complex (conformational exchanges and coordination) and the chemical inertness of the ligand.

In this thesis, we first focused on the synthesis of cross-bridge azamacrocyles. After having discussed their synthesis and chelation-properties, we proposed a paramagnetic tag based on the cross-bridge cyclam for which we proposed several synthetic pathways. This tag is based on the thiol-ene reaction as its conjugation method, the efficiency and catalysis of which will be discussed. A protocol using UV coupled to a radical iniator was developed, allowing the fast and cysteine-selective paramagnetic tagging of our model protein GB1 T53C. In a second time, efforts were concentrated on the synthesis of enantiopure paramagnetic tags. To this purpose we proposed two ligands, which were synthesized, tagged to the model protein GB1 T53C, and characterized via NMR and FFC Relaxometry. Finally, we used the thiol-ene based paramagnetic tagging of *tc*Pex14, a pharmaceutically relevant protein, as an application of paramagnetic NMR in drug discovery. Its crystal structure was refined, and the possibilities of characterizing the binding mode of small ligands using paramagnetic restraints were discussed.

Chapter 1: Paramagnetic restraints in NMR and how to generate them

Molecules containing unpaired electrons (i.e. paramagnetic) affect the magnetic environment of the nuclei in their proximity, and such effects can be followed and studied via Nuclear Magnetic Resonance (NMR) spectroscopy. Because these effects are strongly distance-dependent, they become a source of structural information, in that they give distance and angular restraints between the paramagnetic center and the nuclei of interest. In the last 3 decades, these effects have been widely used in NMR-based structural biology. In this chapter, we will first discuss some of the effects induced by paramagnets and their application in structural biology. We will then show how it is possible to generate such restraints with a focus on paramagnetic lanthanides and the design of synthetic Lanthanide-Binding-Tags (LBT).

1.1 Unpaired electrons and magnetic susceptibility:

A single paramagnetic center is characterized by its magnetic susceptibility tensor[1] χ, which describes the contribution of the unpaired electron's orientation-dependent magnetic moment, and which can be divided into two parts: an isotropic component χ_{iso}, and its anisotropic part, the $\Delta\chi$ tensor. It is important to remark that not all paramagnetic centers are anisotropic: The χ tensor is anisotropic if the magnetic moment of the paramagnetic center varies with different orientations in the magnetic field[2]. The anisotropicity of a tensor depends on its source: Mn^{2+} and Gd^{3+} are isotropic, for example, while paramagnetic lanthanides Ln^{3+} (except Gd^{3+}) are all anisotropic. We can then decompose the χ tensor as such:

$$\chi = \chi_{iso} + \Delta\chi$$

Where χ_{iso} is characterized by three orthogonal axes (χ_x, χ_y, χ_z) and the anisotropic component $\Delta\chi$ by axial and rhombic components:

$$\Delta\chi_{ax} = \chi_z - \frac{\chi_x + \chi_y}{2}$$

$$\Delta\chi_{rh} = \chi_x - \chi_y$$

It is clear from these equations that in the case of $\chi_x=\chi_y=\chi_z$, $\Delta\chi$ becomes null.

There are several mechanisms of interaction between a paramagnetic center (χ) and a nuclear spin, resulting in distance -and sometimes orientation-dependent effects that can be exploited as structural restraints. To be exploited, these restraints are calculated as the difference between an NMR spectrum in the paramagnetic state, and its diamagnetic counterpart (also called diamagnetic reference). From now on, we will distinguish two cases: whether the perturbation is induced by the isotropic component of χ or the anisotropic one. We will look separately at their mechanisms, effects and their applications in NMR.

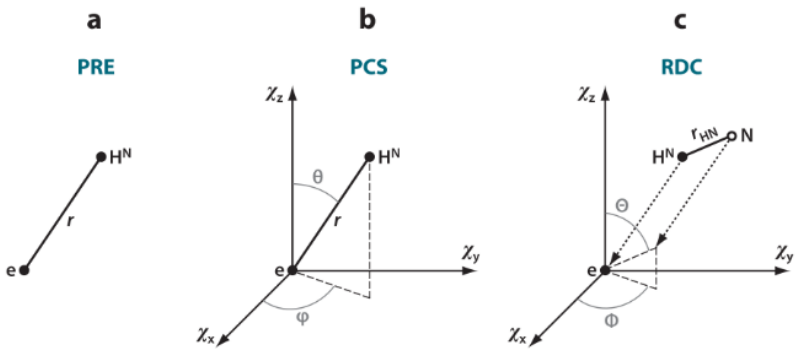


Figure 1.1: Geometric information contained in the paramagnetic effects discussed in this chapter. PRE (a) is described in eq. (2). PCS (b) is described in eq. (5). RDC (c) is described in eq. (6)[2]

1.2 Isotropic Paramagnetic restraints: Paramagnetic Relaxation Enhancement

In this section, we will consider an isotropic paramagnetic center with a tensor where $\Delta\chi = 0$, and therefore: $\chi = \chi_{iso}$

When a nucleus is close enough from such a paramagnetic center, its longitudinal and transversal relaxation rate (R_1 and R_2, respectively) will change[3]. In most cases, this will cause a broadening of the NMR signal, called Paramagnetic Relaxation Enhancement (PRE). These PRE are measurable, and it is possible to turn them into structural distances. Indeed, the intensity of the line broadening follows spheres centered on the paramagnetic center, the intensity and the decay of these spheres depend on several parameters such as the intensity of χ, the magnetic field strength, the electron spin correlational time τ_e, and the gyromagnetic ratio of the considered nucleus[4]. We will see in this section why and how by measuring the difference between the relaxation rates

of the same system in a paramagnetic and a diamagnetic state it is possible to derive distances between nuclei and the paramagnetic centers.

PRE stems from three main contributions: the Dipolar Relaxation (or Solomon Relaxation, or DD), the Curie relaxation (or Dipolar Shift Anisotropy, DSA) and the Curie-DD cross-correlated relaxation. There is also a mechanism of contact-relaxation, but in if the nucleus is far enough from the paramagnetic center (r > 4 Å), we can neglect the contact effects for simplification.

The dipolar relaxation usually dominates for long electron spin correlation times (for example Gadolinium). On the other hand, the Curie relaxation dominates for short electron spin correlation times (lanthanides other than Gadolinium). The relaxation rates are always positive and additive. When all paramagnetic effects are combined we can therefore sum up R_i^{para} as:

$$R_i^{para} = R_i^{DD} + R_i^{Curie} + R_i^{Curie,DD}$$

with i ∈{longitudinal ≡ 1; transverse ≡ 2}

Both longitudinal and transversal relaxation rates are affected by the presence of a paramagnet. The relation between longitudinal relaxation rate enhancement (R_1^{para}) and the distance from nucleus to paramagnet was described by Solomon[5]. R_1^{para} relaxation rates can be used to estimate distances from metals to protein nuclei, however this area has not been explored much due to experimental difficulties[6].

Measuring R_2^{para} and calculating distance using spin labels

The most widespread source of unpaired electrons (or spin labels) in this technique are organic radicals such as MTSL ((*S*-(1-oxyl-2,2,5,5-tetramethyl-

2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate))[7], and gadolinium-based probes. The approach mostly used to measure distances within proteins is called “single-point measurement” and has been described in 2000 by Wagner *et al*[8]. First, it is required to measure R_2^{para} and this can be done via NMR following the equation:

$$\frac{I_{para}}{I_{dia}} = \frac{R_2^{dia}\exp\left(-R_2^{para}t\right)}{R_2^{dia} + R_2^{para}} \qquad (\mathbf{1})$$

Where I_{para} and I_{dia} are the peak intensities measured via NMR and t the total time the magnetization evolves in the transverse plane during the INEPT transfer[4,8]. As we have seen earlier the contributions to R_2^{para} are additive and in the case of organic radicals and Gd-based spin labels the dipolar interaction is the prevalent mechanism. The mathematical relation between R_2^{para} and the electron-to-nucleus distance ***r*** derives from those rules and from the Solomon equations[5,9], and can be written as such[10]:

$$R_2^{para} = \frac{2}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^2 g_e^2 \mu_B^2 S(S+1)}{r^6}\left[\frac{7\tau_c}{1+\omega_S^2\tau_c^2} + \frac{3\tau_c}{1+\omega_I^2\tau_c^2}\right] \qquad (\mathbf{2})$$

To generate such restraints, a common techniques consists in the attachment of a spin label on the surface of the protein of interest. Because mobility of this tags is deleterious to PRE-measurement and analysis[11], continuous efforts are devoted to the development of more rigid and more stable spin labels[12].

1.3 The hyperfine shift and its effects: Pseudo-Contact Shifts (PCS) and Residual Dipolar Coupling (RDC)

In the last paragraph we have seen how unpaired electrons can affect the relaxation of surrounding nuclei and how it is relevant to us in terms of biomolecular NMR. In this part, we will focus on how the average electron-induced magnetic moment affects the energy of the nucleus. Indeed, it can be seen as a perturbation in the overall magnetic field sensed by the nucleus and therefore it will have an effect on chemical shift. This contribution is called the hyperfine shift[13]. We can expect such shift when there are energy levels for the considered ion close to the ground state. This causes orbital contributions to the ground state, and such contributions are orientation dependent, depending on the shapes of the contributing orbitals. In this case, the magnetic tensor is anisotropic.

1.3.1. Theoretical considerations on Contact Shifts and Pseudo-Contact Shifts

1.3.1.1 Spin Density and Contact shift

The electron has the particularity of not being present in a single point but in a Molecular Orbital (MO) which may be delocalized on a whole molecule. This is also true for unpaired electrons. We name spin density ρ the fraction of unpaired electron present in a given point of space. If ρ is not null at a nucleus, it will generate an additional magnetic field giving rise to what we name a Contact Shift (CS). It is clear that we will observe such shifts only on those nuclei that are very close to our paramagnetic metal or radical. Often, those shifts are difficult to be measured since the signals of the nuclei close to the paramagnetic centers tend to disappear because of PRE.

1.3.1.2 Dipolar interaction and Pseudo-Contact Shifts

Both nuclei and the unpaired electron have magnetic moments, and therefore there will be a dipolar interaction between them. This interaction, for two point-dipoles μ_1 and μ_2 separated by a vector **r** are characterized by the dipolar energy E^{dip} (equation (3)). The induced chemical shift given by this interaction is given by equation (4). We can see at that point that the value of this dipolar chemical shift perturbation depends on the distance but also the respective orientation of the two magnetic moments.

$$E^{dip} = -\frac{\mu_0}{4\pi}\left[\frac{3(\mu_1.r).(\mu_2.r)}{r^5} - \frac{\mu_1.\mu_2}{r^3}\right] \qquad (\mathbf{3})$$

$$\delta^{dip} = \frac{-\Delta E^{dip}}{\hbar\gamma_i B_0} \qquad (\mathbf{4})$$

By combining (3) and our previous definition of χ, it has been shown that in solution this shift, known as Pseudo-Contact Shift (PCS), can be written under the form:

$$\delta^{PCS} = \frac{1}{12\pi r^3}[\Delta\chi_{ax}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{rh}(\sin^2\theta\cos 2\varphi)] \qquad (\mathbf{5})$$

This equation will be very important in our work. In this case r, θ, and φ are the spherical coordinates of the nucleus in the frame of reference centered on the paramagnetic center and defined by χ. From this equation, we can predict the δ^{PCS} for a given nuclei. However, we may note that from one PCS value we cannot calculate one position as there is not only one solution to this equation. The ensemble of points which are solution of this equation is called an isosurface.

Practically, PCS are measured by subtracting the chemical shift of a nucleus in a diamagnetic environment to the one obtained in “paramagnetic conditions” (i.e. close enough to a paramagnetic center so that the dipolar interaction between the two is significant and may be observed). This analysis requires only simple spectra (1D, 2D HSQC…) and therefore eases access to a wealth of structural information easily.

PCS can be measured at distances up to 100 Å. As we can see from (5), in order to maximize the shift, we need to work in a system where $\Delta\chi$ is as large as possible. We can also note that PCS, unlike PRE, are not only distance-dependent but also orientation dependent.

1.3.2. Theoretical considerations on Partial Alignment and Residual Dipolar Couplings

The orientation of a molecule in a given magnetic field has an influence on the dipolar couplings between nuclei that compose the molecule. This is clear in solid-state NMR where signals are very broad. In solution, the fast isotropic reorientations of molecules average out the contributions corresponding to different nuclear positions with respect to the external magnetic field. This simplifies spectral analysis.

If the molecule is partially aligned in the magnetic field, however, it will reintroduce some effects related to the anisotropy of the nuclear interactions. One of these effects is called “Residual Dipolar Coupling” (RDC), and it is visible as a change in dipolar coupling values between a partially aligned and non-partially aligned sample. Partial alignment of a molecule causes the occurrence of residual dipolar couplings because when not all orientations have the same probability, the dipole-dipole interaction between nuclear magnetic moments does not average to zero.

There are several ways to induce partial alignment in a protein in solution. One of them consists in the attachment of a paramagnetic tag on the surface of the protein. The paramagnetic molecules will seek to minimize the dipole interaction with the magnetic field and the magnetic moment of the paramagnet, thus causing partial alignment of the protein in the magnetic field. When the difference is calculated between the splitting of coupled nuclei observed for the paramagnetic molecule and a diamagnetic analog, the contribution from the paramagnetic metal can be isolated and the equation to analyze the residual dipolar coupling data is:

$$\Delta\nu_{AB}^{RDC} = -\frac{1}{4\pi}\frac{B_0^2}{15kT}\frac{\gamma_A\gamma_B h S_{LS}}{2\pi r_{AB}^3}\left[\Delta\chi_{ax}(3cos^2 \propto -1) + \frac{3}{2}\Delta\chi_{rh}sin^2\alpha\, cos2\beta\right] \qquad (\mathbf{6})$$

Where A and B are two J-coupled nuclei, ($\mathbf{r}_{AB}$, α, β) the spherical coordinates of the **AB** vector, γ_A and γ_B the gyromagnetic ratios of A and B respectively, and $\Delta\chi_{ax}, \Delta\chi_{rh}$ the anisotropic parameters. Interestingly, we can see from this equation that unlike for PCS and PRE there is no dependence on the distance between the nuclei A, B and the paramagnetic center (fig. 1.1).

RDC are very sensitive to molecular motions, and this represented in eq. (6) by the Lipari-Szabo order parameter S_{LS}[14]. This parameter reflects the internal motion of the protein and scales down the values of RDC[15].

1.3.3. Practical considerations on PCS and RDC, and what can we get from them

To induce PCS and RDC in a given system we need an anisotropic magnetic tensor. To this purpose, the most commonly used unpaired electron sources are anisotropic lanthanide ions (corresponding to all paramagnetic lanthanides except Gadolinium, which is isotropic). However, diamagnetic

lanthanides are still of use: indeed, PCS and RDC are measured as the difference between peaks in a paramagnetic and in a diamagnetic environment. In order to differentiate the effects stemming from the paramagnetism of lanthanides from the chemical effects of the introduction of the ion, a diamagnetic control using a diamagnetic lanthanide is always performed and used to measure changes in chemical shifts and dipolar couplings.

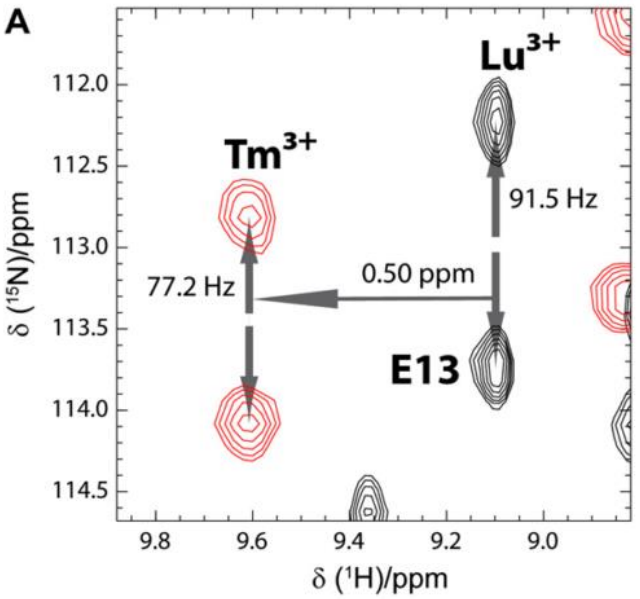


Figure 1.2: Example of paramagnetic effects observed via 2D ^{1}H-^{15}N HSQC for a protein tagged with diamagnetic Lutetium and paramagnetic Thulium. PCS is observed as an increase of chemical shift, and RDC as a decrease in the J-coupling. Reproduced from[6].

Based on equations (5) and (6), we can define tensors as 8-variable mathematical functions which correlate a given point in space with a value of a paramagnetic effect. Those functions are depending on the system studied (couple protein/tag) and the metal used as the paramagnetic center. Determining tensor parameters with precision is a fundamental step in analyzing paramagnetism-assisted NMR data. Indeed, once PCS or RDC have been

measured, calculating distances and orientations cannot be done directly. In order to associate a shift value to an isosurface using equations (5) and (6) , first need to be determined the parameters of the anisotropic tensor, $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$, but also the position of the paramagnetic center and the orientation of the tensor axes[16,17].

To determine these tensors, several pieces of software have been developed, two of which are Numbat[18] and FANTEN[17]. FANTEN is a server-based and "user-friendly" software which requires very few inputs. Those inputs are: a .pdb structure of the studied protein, a text file with the measured paramagnetic measurable, and some tensor-parameters which may be already known. The software will then perform a fitting of this input and the output will be the fitted tensor parameters, the calculated average position of the metal center, and a Q-factor correlated with the quality of the fitting. In order to perform a good fitting, it is important that 1) the protein structure as input is valid and 2) a sufficient number of paramagnetic measurable has been inserted.

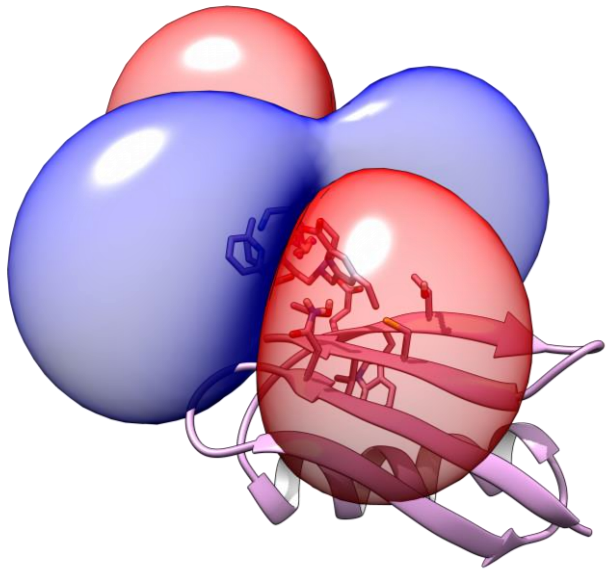


Figure 1.3: Example of an isosurface corresponding to a PCS value of +1 ppm (blue) and -1 ppm (red), for a protein with a Dysprosium ion attached on its surface

1.4 Applications of Paramagnetism-derived restraints

1.4.1. Applications in structure determination and refinement

Structure calculation by NMR mainly relies on the collection of short-range distance-restraints (up to ~5-6 Å) provided by the time-consuming and troublesome analysis of NOESY spectra[19]. The use of long-range paramagnetic distance restraints (up to ~100 Å), such as pseudo-contact shifts (PCS), induced by a paramagnetic ion, has been widely proposed to help in *de novo* structure determination by NMR[20–24]. One of the first example was Calbindin D_{9k}, its structure in solution was determined in 2001 by Bertini *et al.* showing that paramagnetic restraints were useful and complementary to traditionally used NOEs[20]. These restraints are implementable directly in some structure determination software such as Xplor[25].

Another application of paramagnetic restraints consists in the refinement of pre-existing X-ray structures[26,27]. In this way from the crystalline structure a more reliable model describing the protein in solution can be obtained. This combination allows to correct some possible bias obtained via X-Ray where the protein's structure is altered in the solid state, in multi domain proteins for example[28].

With the improvements of methods to generate large anisotropic tensors, it is possible to generate restraints on larger distances and especially between different domains of proteins[29], and even between different proteins[30–32].

1.4.2. Paramagnetism-assisted NMR of multi-domain protein and protein complexes

Paramagnetic restraints can also be used for the study of larger systems, like multi-domain proteins or even protein complexes[33]. Indeed, long-range through-space paramagnetic effects may spread from a paramagnetically tagged protein to another domain of the same protein or a different protein with which it is in complex. In such cases, paramagnetic restraints are not used to determine *de novo* structures, but to refine existing X-ray structures or to characterize the dynamic of such complexes, for example including paramagnetic restraints in docking procedures[6,34].

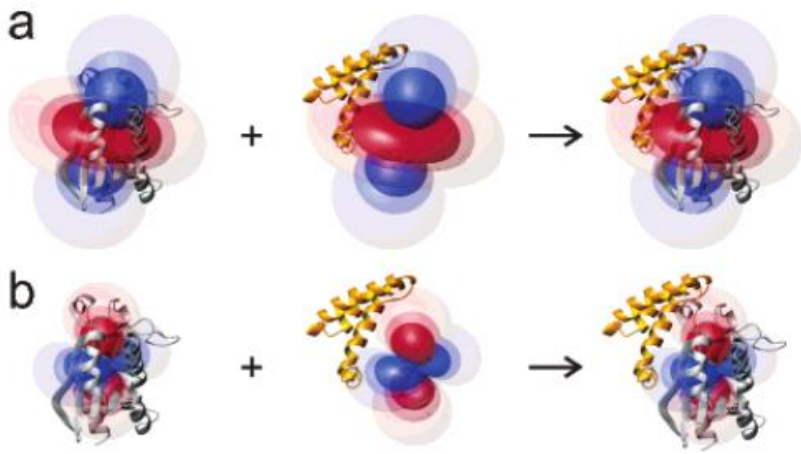


Figure 1.4: (a) Rigid body docking of the proteins ε186 (gray) and θ (orange) by superposition of the $\Delta\chi$ tensors of Dy^{3+} determined with respect to the individual proteins. The $\Delta\chi$ tensors are represented by PCS isosurfaces with positive and negative shifts indicated by blue and red colors, respectively. (b) Same procedure with Er^{3+} instead of Dy^{3+} [35].

An example of two-domain protein studied with paramagnetic restraints is Calmodulin[36]. Both domains were paramagnetically tagged and the maximum

occurrence (MO) of a set of calmodulin conformations that are representative of the overall conformational space possibly sampled by the protein, has been calculated from the paramagnetism-based restraints. The author then showed that the availability of paramagnetic restraints arising from metal ions placed on both domains, reduces the MO of the conformations, thereby helping to identify those conformations which are more representative of the ones adopted by the protein in solution.

1.4.3 Paramagnetism-assisted Drug Discovery

An important application of this source of structural restraints is found in the determination, refinement and dynamic studies of proteins in complex with small molecules. This is particularly useful in structure-based drug discovery projects, where being able to identify precisely a small binder's binding site on a protein surface, and possibly obtain the structure of the complex, is required to optimize hit-compounds[33]. This is often achieved through analysis of NOE contacts; however, the transfer of paramagnetic effects from a paramagnetically tagged protein to a small ligand may also yield additional structural restraints. Particularly, PRE and PCS have been used[33,37]. Here, we will focus on PCS as they are easier to interpret and more straight-forward to measure.

PCS in small ligands, just as for proteins, are measured as the difference in chemical shift between a diamagnetic and a paramagnetic sample. Often ^{1}H are enough to measure PCS and no 2D experiments are required. In the case of ligands in fast exchange, in order to calculate the PCS value for the ligand in the bound state, it is necessary to scale the measured value as such[38]:

$$\Delta\delta^{PCS}_{measured} = f_{bound}\,\Delta\delta^{PCS}_{bound} \qquad (7)$$

Where: $f_{bound} = \frac{c_p}{K_D + c_L}$

By using equation (7), we can then back-calculate the values of PCS corresponding to the bound state. It is this value which will be an input as a structural restraint. Recent examples have used this strategy to perform docking measuring PCS on ^{1}H but also on ^{19}F [39,40,41].

The measurement of a minimum of three independent sets of a PCS allows for a precise structural determination of the small-ligand binding-mode in a GPS-like fashion[33,42]. Indeed, using triangulation it is possible to determine the position of a nuclei with respect to the protein.

1.5 Lanthanide-Binding Tags as a source of paramagnetic restraints

1.5.1 Lanthanide-Binding Tags:

In order to induce meaningful and measurable effects on an NMR spectrum, it is important to consider the necessity of having the paramagnetic center fasten on the surface of the biomolecule of interest[2]. In the case of metal-binding protein, this can be easily done by substituting a naturally bound metal by a lanthanide[27,43]; In the case of non-naturally metal-binding-proteins, we can distinguish two kinds of techniques: natural metal-binding sites and synthetic lanthanide-chelating moieties. Metal-binding peptides, where lanthanides can be substituted, can be engineered at an extremity of a protein[44], so that paramagnetic shifts may be induced on the rest of the protein, however the introduction of such peptide may interfere with the global structure of the protein; moreover, the peptide can be engineered only at one end of the protein, leaving very little flexibility as to where to introduce the paramagnetic center.

Synthetic lanthanide-chelating tags, on the other hand, are used if a protein does not have the possibility to chelate a lanthanoid[45]. Such complexes were first used as non-covalent probes, also called shift reagents[46,47]. In the last 20 years however, we have been able to synthetize molecules capable of both chelating a lanthanoid and being conjugated to the target protein on a specific site[8,48]. This presents the advantage of having the paramagnetic center fixed on a specific point of the protein's surface, which allows for the determination of a single tensor and generally induces larger paramagnetic perturbations and on larger distances. Moreover, it gives the freedom to choose where to attach the tag on the protein's surface. These molecules will be referred hereafter as lanthanide-binding tags or simply paramagnetic tags.

An ideal lanthanide-binding-tag (LBT) requires three characteristics:

1) Be chemically inert
2) Be present as a single stereoisomer
3) Be as rigid as possible

By "chemically inert", we intend that the tag ought not to release the lanthanoid in solution, so as not to lose the desired paramagnetic effect and potentially catalyze protein hydrolysis. This is even more important in the case when longer experiment times (several days) are needed. As seen previously, the effects we want to measure are distance-dependent, and it implies that the tags used are enantiomerically pure. Indeed, the presence of two stereoisomers would create two distinct paramagnetic centers, inducing a doubling of the peaks observed on the spectrum.

As for rigidity, the more flexible a tag, the more the paramagnetic center will tumble in solution, causing an averaging of the paramagnetic effect and a diminution of the intensity of the transferred $\Delta\chi$ tensor. As seen in equation (5) and (6), the larger the $\Delta\chi$ tensor, the larger the measured PCS and RDC will be

and at greater distances. In order to gain structural restraints on bigger proteins or in order to monitor biological processes at the atomic level on a larger scale, it is therefore needed to optimize the location and the parameters of the anisotropic tensor. Are therefore needed paramagnetic tags with large intrinsic tensors and this can be achieved through the increase of the tags' rigidity. While 1) and 2) have been readily achieved in several tags, rigidity is the parameter towards which efforts are still required.

Another important parameter for the optimization of paramagnetic tags and their effects is the site of conjugation and the type of chemical linker used to attach the tag to the protein's surface. In order to be conjugated on the surface of proteins, paramagnetic tags need to present a functional group allowing for a conjugation reaction to take place. The thiol group of cysteine stands out as an ideal target due to its relative low abundance in protein primary structures and the possibility for a chemoselective modification with LBTs. In order to provide a rigid tagging site in an ordered region of a protein, cysteine residues in alpha-helixes are often chosen as suitable functional group.

For the tagging of cysteine residues in proteins, several suitable linkers have been developed. Two widely used systems are 2-mercaptopyridine as well as phenylsulphonyl activated linker (fig. 1.5). The use of 2-mercaptopyridine as an activating group for the formation of disulfide bonds between tags and cysteines has been widely employed[49,50]. It allows for a fast and selective tagging (generally around 4 hours of reaction) but yields a disulfide bond, which is reducible. This may constitute a problem for longer experiments and make this strategy unusable for in-cell applications. On the other hand, 4-phenylsulfonyl-pyridine has been introduced to yield a thioether bond between the free thiol or a protein and a tag[51]. The thioether bond is stable in reducing conditions, making this strategy attractive. However, reaction times necessary to give high conjugation yield are longer, typically one night is required.

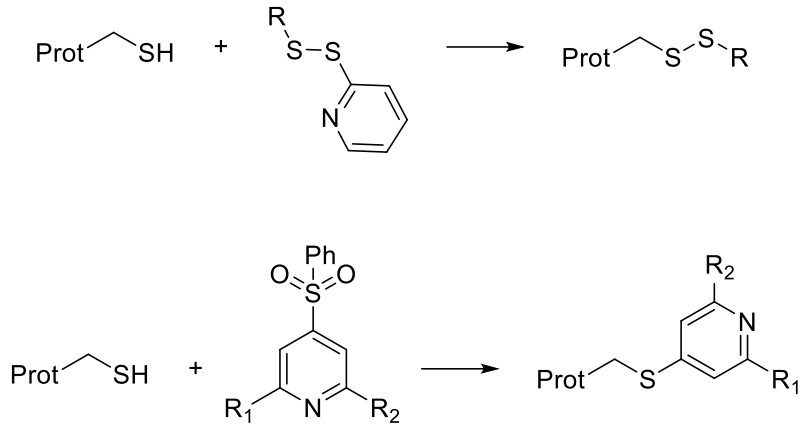


Figure 1.5: Conjugation to cysteines via disulfide formation (top) and through nucleophilic substitution of phenylsulfone (bottom). R, R_1 and R_2 are generic lanthanide-chelating moieties.

We could separate synthetic single-attachment-point LBTs in two families, whether the lanthanide-chelating moiety is cyclic (DOTA-like molecules) or not (e.g. EDTA-based tags).

1.5.1.1. Acyclic Paramagnetic tags

Some of the first examples of the design of single-attachment-point LBT include acyclic tags, e.g. EDTA-derived molecules[48,52]. However, they displayed different stereoisomers and therefore induced different set of peaks in the NMR spectra. Later, Cys-Ph-TAHA was introduced[49]. Enantiomerically pure, it displayed only one set of peaks upon conjugation; bearing a long linker it is however quite flexible and yields only small paramagnetic tensors.

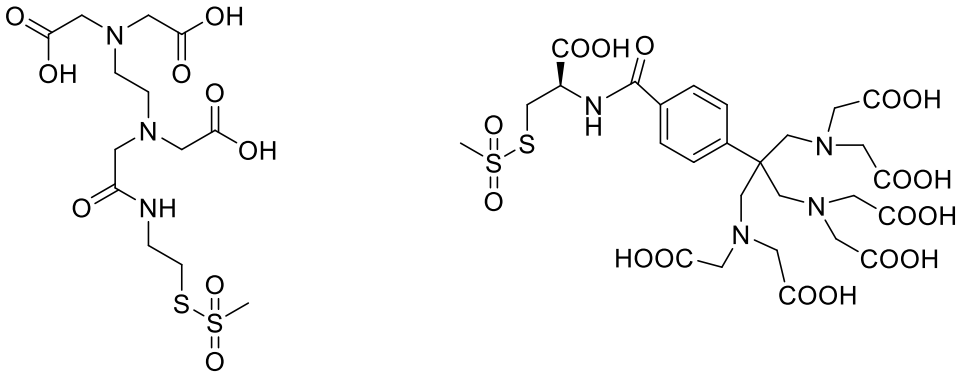


Figure 1.6: MTS-EDTA (left) and Cys-Ph-TAHA (right)

1.5.1.2. DOTA-like Paramagnetic tags

DOTA is a strong and almost universal chelator that has been used with many applications in imaging, diagnostic, theranostic, *etc*…[53,54] Because it has such a high affinity for lanthanoids, DOTA has also been used in paramagnetic NMR Spectroscopy[55]. DOTA's main drawback is that it can take two different conformations in solution: SAP (Square AntiPrismatic) and TSAP (Twisted Square AntiPrismatic). Those two conformations each behave as a different and independent paramagnetic center and therefore each lead to a different set of peaks in the NMR Spectrum, making much harder the analysis of the NMR spectra. Several attempts to overcome this drawback have been done. A notable strategy has been the introduction of several methyls groups around the cycle in order to block its conformation and slow down its motion. The main application of this strategy in the field of paramagnetic tagging was achieved by Daniel Haüsssinger, DOTA-M8[56]. However, it was later shown that this macrocyle was still in equilibrium between two diastereoisomers, the ratio of which is interestingly depending on the radius of the chelated Lanthanide[57]. For many

lanthanides, though, it adopts only one conformation and therefore yields only one set of peaks and large paramagnetic tensors.

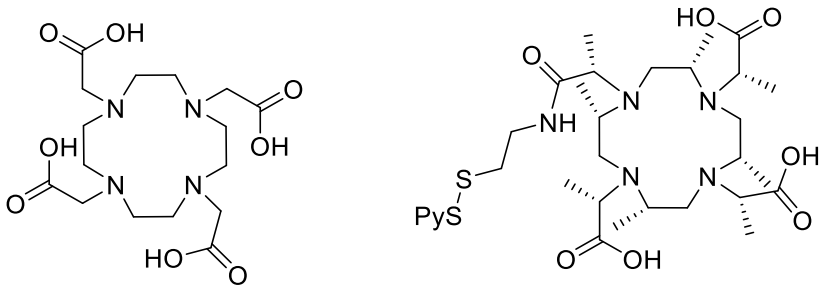


Figure 1.7: DOTA (left) and DOTA-M8 (right)

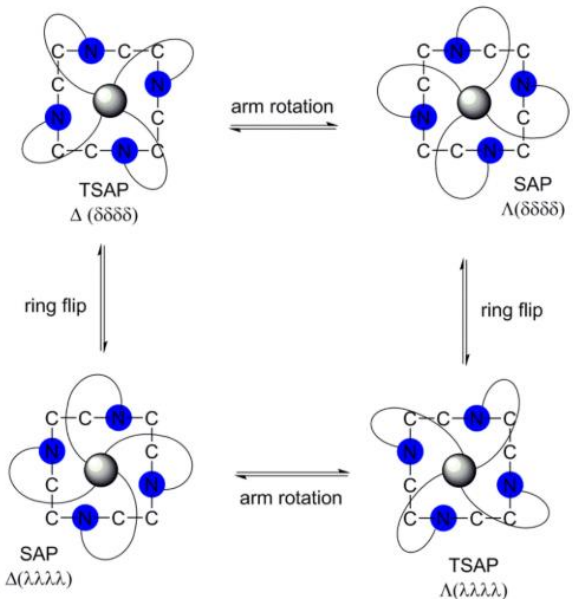


Figure 1.8: Conformations of DOTA-derived ligands upon metal-coordination[57]

1.5.2. Properties of Lanthanides:

Lanthanoids elements have been extensively used in NMR spectroscopy[2]. While bearing similarities in chemical behavior and reactivity, they offer a broad range of paramagnetic characteristics depending on the occupation states of their 4f orbital. Indeed, while Lu and La are diamagnetic (absence of unpaired electron), there is a broad range of magnitude of $\Delta\chi$ tensor from Cerium (Ce) to Dysprosium (Dy), offering the possibility to tune the desired effects by choosing a specific lanthanoid ion[58]. Gadolinium, extensively used in spectroscopy and in imaging, has the particularity of being isotropic, inducing strong relaxation enhancement but no PCS or RDC.

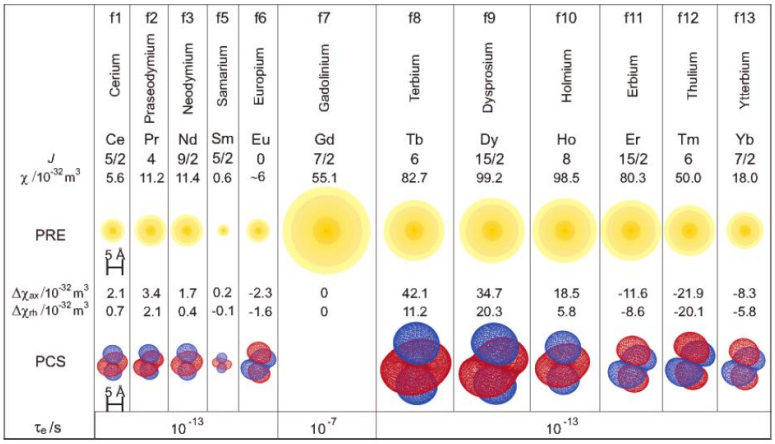


Figure 1.9: Paramagnetic properties of Ln3+ ions. Only paramagnetic and nonradioactive lanthanides are included. Representative isosurfaces are plotted for PCSs by (5 ppm using tensors reported by [16]) The radius of the yellow sphere indicates the distance from the metal ion at which ^{1}H NMR signals of macromolecules with a rotational correlation time of 15 ns would be broaden by 80 Hz due to paramagnetic relaxation enhancement (PRE) at a magnetic field strength of 18.8 T. Typical electronic relaxation times representative for this field strength are indicated at the bottom. For Eu^{3+}, the estimate of the relaxation enhancement includes a contribution from excited J manifolds[59]. Reproduced from[35].

Having similar chemical behavior but different paramagnetic properties, it is common to use the same tag loaded with different paramagnetic lanthanides. Doing so, it is possible to obtain different sets of structural restraints using only one tag/protein system.

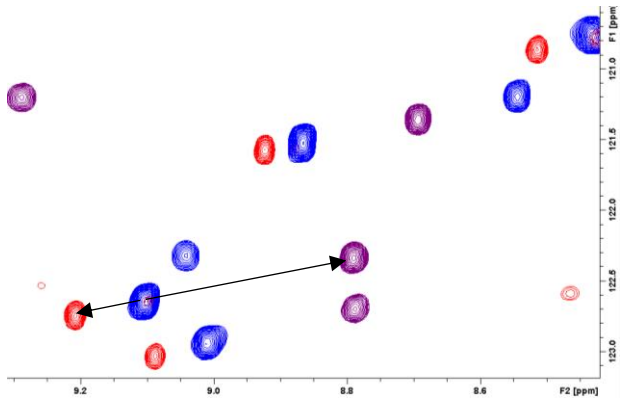


Figure 1.10: Example of 2D ^{1}H-^{15}N HSQCs of *a* protein tagged with Lutetium (blue), Ytterbium (red) and Dysprosium (purple). The PCS induced by Yb and Dy are of opposite signe, and the absolute value of the shift induced by Dy is larger than the one induced by Yb, what is in line with values reported in fig 1.9.

CHAPTER 2: Synthesis of cross-bridge azamacrocycles

Abstract

Cross-bridge azamacrocycles are metal chelators which have been shown to form very inert complexes with different transition metals. Particularly, cross-bridge cyclam derivative cb-tedpa has been shown to chelate lanthanides and those complexes display very high inertness, even in highly acidic media. With the scope of diversifying available highly rigid ligands for lanthanide complexation, we have developed a new strategy for the synthesis of diversely alkylated cross-bridge cyclam and cyclen derivatives. We have shown that this new strategy was more efficient than the previously described one. More, the stability of different complexes of Europium and azamacrocycles in acidic media have been studied and compared.

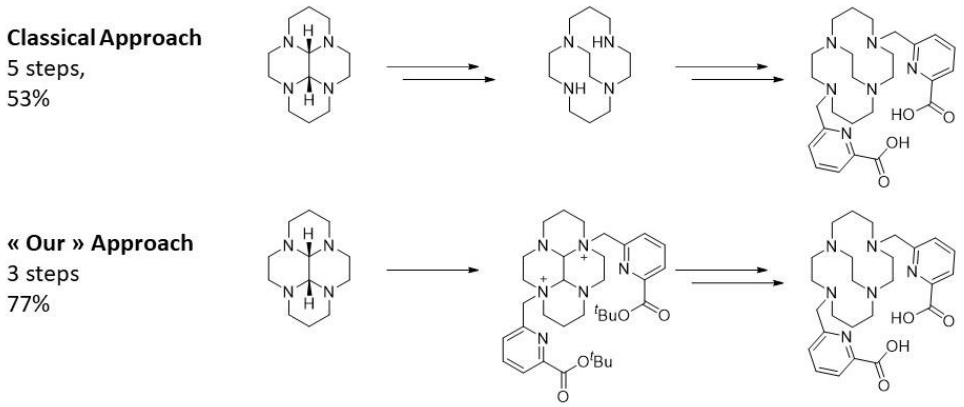


Figure 2.1: Synthetic strategies discussed in this chapter

2.1 Introduction

Cross-bridged cyclam (cb-cyclam) was first reported in 1990[60], and cross-bridge cyclen (cb-cyclen) in 1994[61]. Cb-cyclam and cb-cyclen rapidly gained interest as potential highly rigid transition metal chelators[62]. Indeed, it was shown that those molecules have very high affinities for Lewis Acids, and that they are able to adopt cleft-containing geometries around metal ions.

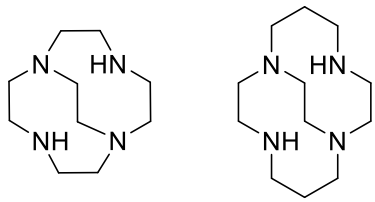


Figure 2.2: Structure of cross-bridge cyclen (left) and cross-bridge cyclam (right)

Structural studies showed that in both cases, the bridge adopts a cis conformation[62], in which the lone pairs of the nitrogen atoms all point towards the center of the molecules, creating a cavity. It is believed that the kinetic inertness of these molecules is endowed in part by the ligand's relatively strain-free coordinating conformation, which adopts a C2 symmetry.

It was shown that it was possible to further enhance the stability of cross-bridge azamacrocycles by adding ionizable pendant arms at the secondary amines[63]. These can serve to fully envelop six-coordinated transition metal cations, as well as to neutralize their charges. Two examples are cb-TE2A and cb-DO2A, which have found applications in imaging[64] (fig. 2.3).

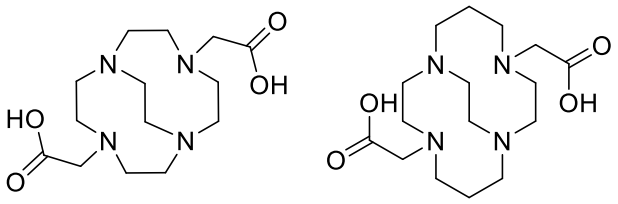


Figure 2.3: Structures of cb-DO2A (left) and cb-TE2A (right)

It has been reported by Niu *et al*. that cb-DO2A exhibits an enantiomeric interconversion in solution in complex with metal ions[65] (fig. 2.4). This dynamic process has been shown to be fast on the NMR timescale. It is hypothesized that this dynamic process can take place through two possible mechanisms: a dissociative one where pendant arms are “detached" and the ligand can then reorganize in its enantiomer while the ion is temporarily released, or a non-dissociative one with only a pseudo-rotation of the arms. We can expect that the thermodynamical and kinetical stabilities of the complexes will govern on which of these mechanisms takes place. As for cb-cyclam, it also exhibits dynamic enantiomerization. This process requires the ethylene bridge tucking through the 14-membered ring along with inversion of all four nitrogens (fig. 2.5)[66]. To the best of our knowledge, there is no published data on the enantiomerization of alkylated cb-cyclams.

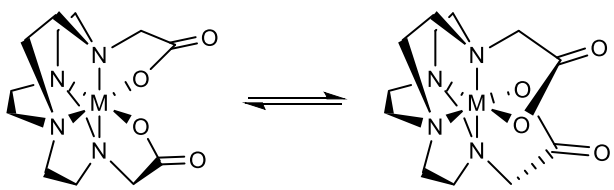


Figure 2.4: Enantiomerization of cb-DO2A in complex with metal ions

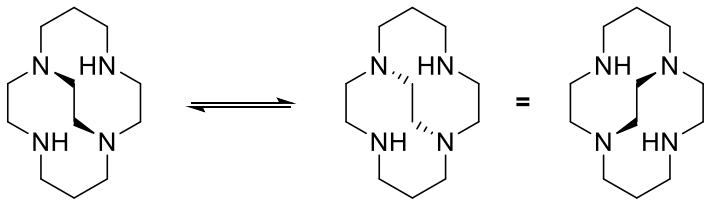


Figure 2.5: Enantiomerization of cross-bridge cyclam

It is only recently that cross-bridge azamacrocyles have started to be used for lanthanide complexation[67,68]. Rodiguez-Rodriguez *et al.* used picolinic acid pendant arms in order to increase the number of coordinating atoms in of the ligand while using cross-bridge cyclam as the central scaffold. The resulting cb-tedpa has been shown to fully envelop Europium in the solid state (fig. 2.6). From a kinetical point of view, [cb-tedpa(Ln)] complexes are extremely stable and no sign of ion releasing was observed even after months in HCl 2M. From a thermodynamical point of view, a comparative study in solution and via DFT showed that cb-tedpa was much more stable than its non-bridged counterpart Me_2tedpa[69]. They do not, however, discuss the stereochemistry of their new complexes.

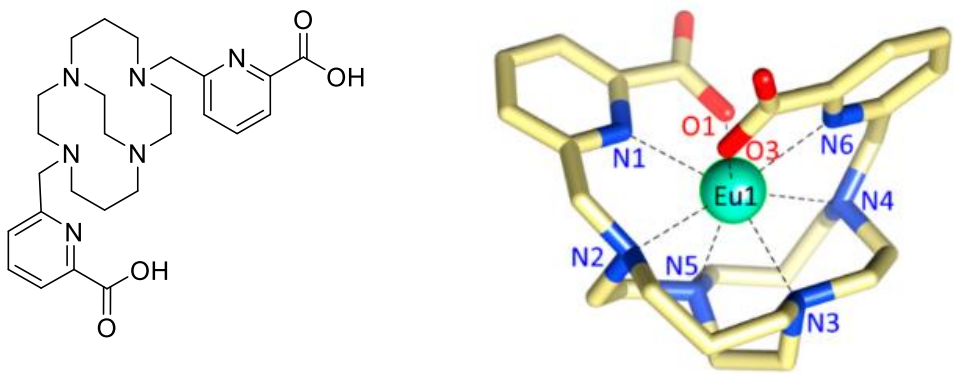


Figure 2.6: 2D structure of cb-tedpa (left) and 3D structure of [(Eu)cb-tedpa][67] (right)

Synthesis of cross-bridge tedpa (cb-tedpa)

In terms of synthesis, the classical approach to obtain alkylated cross-bridge cyclam consists in the direct alkylation of the secondary amines of cb-cyclam, itself can be obtained following the procedure reported by Wong et. al[63]. This procedure allows the obtention of cross-bridge cyclam in four steps starting from cyclam (58% total yield). Two interesting and important steps in this synthesis are the regioselective bisaminal alkylation and the reductive ring-opening. The regioselectivity in the bisaminal alkylation stems from the conformation of cis-glyoxal cyclam in which two nitrogen lone pairs are endo (non-reactive) and two are exo (reactive). Next, the reductive double ring expansion allows for the selective obtention of the macrocycle where all amines are tertiary.

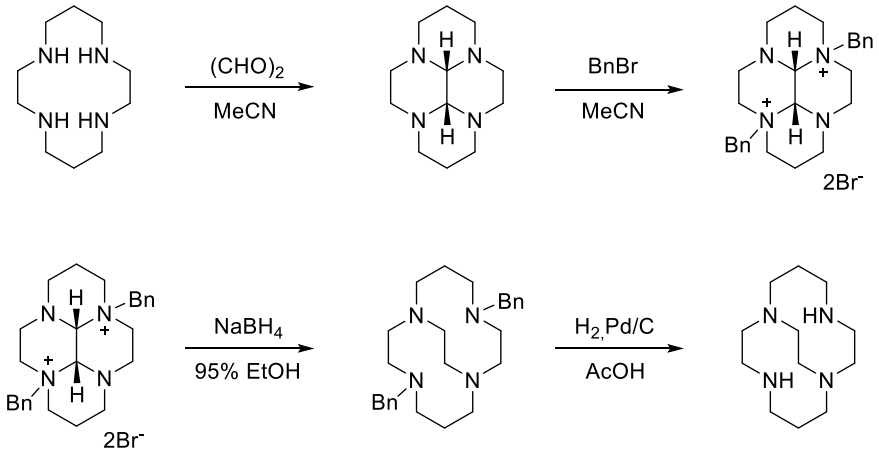


Figure 2.7: Synthesis of cb-cylam

The last step consists in the deprotection of the secondary amines. At this point, cb-cyclam can be symmetrically alkylated to yield cb-TE2A or cb-tedpa. Some examples also report the mono-alkylation of cb-cyclam to create

asymmetrical ligands[70,71]. This is extremely useful for the design of scaffolds for bioconjugation; indeed, it allows us to design an asymmetric LBT with a cb-cyclam moiety, two arms, one of which bearing a functional group for protein conjugation[71,72].

In this chapter, we will introduce and discuss a new pathway for the synthesis of cb-tedpa. We will compare it to the classical synthesis and discuss its applicability for the design and the faster synthesis of novel asymmetric cb-cyclam and cyclen based scaffolds. As we have seen lanthanides have many interesting properties useful not only in paramagnetic NMR but in many other techniques. In order to harness the power of cross-bridge azamacrocyles, it is very important that we can control the chemical environment (coordination, rigidity) of the lanthanides. Therefore, it becomes useful to easily introduce different arms on the macrocycle. In the specific case of paramagnetic tagging, it would allow us to design an asymmetric LBT with a cb-cyclam moiety, two arms, one of which bearing a functional group for protein conjugation.

2.2 Synthesis of cb-tedpa through direct bisaminal alkylation

2.2.1 New synthetic strategy

The synthesis of cb-tedpa as reported by Rodriguez-Rodriguez goes through the alkylation of the secondary amines of cross-bridge cyclam, an intermediate obtained from cis-glyoxal cyclam through reductive double ring expansion. In order for the ring expansion to proceed correctly and give the desired cross-bridge, it is important that two amines in trans are quaternary. To achieve this, the bisaminal alkylation with benzyl bromide (BnBr) is performed which requires debenzylation once the glyoxal moiety has been "reductively opened" to the cross bridge. In this section we will discuss the synthesis of cb-tedpa through direct bisaminal alkylation. Indeed, rather than using the bisaminal alkylation with BnBr as a protecting step, we will try to use this step to directly introduce the desired pendant arm on cis-glyoxal cyclam, in order to avoid the removing of the benzyl groups.

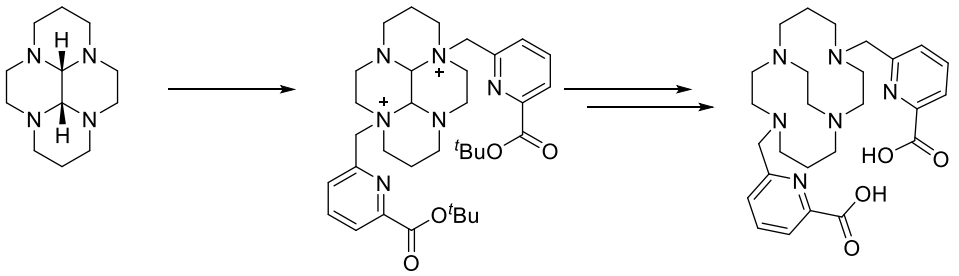


Figure 2.8: New strategy for the synthesis of cb-tedpa

To do so, we need to have a picolinic acid derivative with a leaving group in position 6 and a carboxylic acid protected as an ester that can resist the harsh conditions of the reductive double ring expansion. For this purpose, we decided

to use **1**, with a bromomethyl in position 6 an ^{t}Bu ester to protect the carboxylic acid from reduction.

2.2.2 Synthesis of cb-tedpa

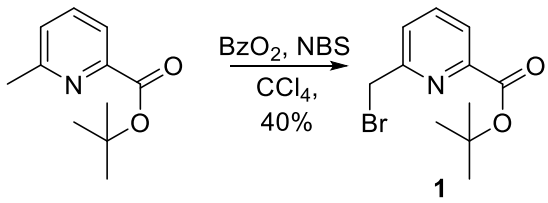


Figure 2.9: Synthesis of **1**

1 was synthesized in one step from *tert*-butyl 6-methyl picolinate (fig. 2.9). With **1** in hand, we first tried to mono-alkylate cis-glyoxal cyclam. Interestingly, we obtained mono-alkylated **2** in 95% yield and no trace of dialkylated cyclam was observed. Mono-alkylated **2** was then put in presence of 4 equivalents of **1**; formation of dialkylated **3** was observed although slower- what was expected. Reaction was complete in ten days.

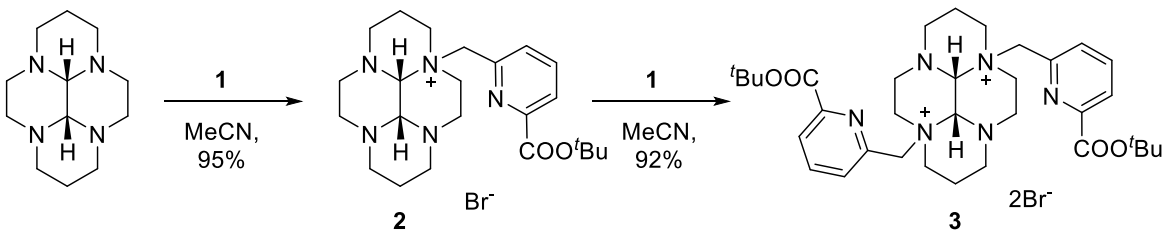


Figure 2.10: Synthesis of dialkylated glyoxal cyclam **3**

Next, we tried the reductive double ring expansion, hoping that the reaction would go as well with picolinc ester on the quartenary amines as with benzyl groups and that the esters wouldn't be reduced. We noticed that not only the reaction was faster (3 days instead of 14), but only the desired product was formed, in good yield (88%). Finally, the ^{t}Bu-esters were easily deprotected in Trifluoroacetic acid (TFA), yielding the desired product as the TFA salt (fig. 2.11).

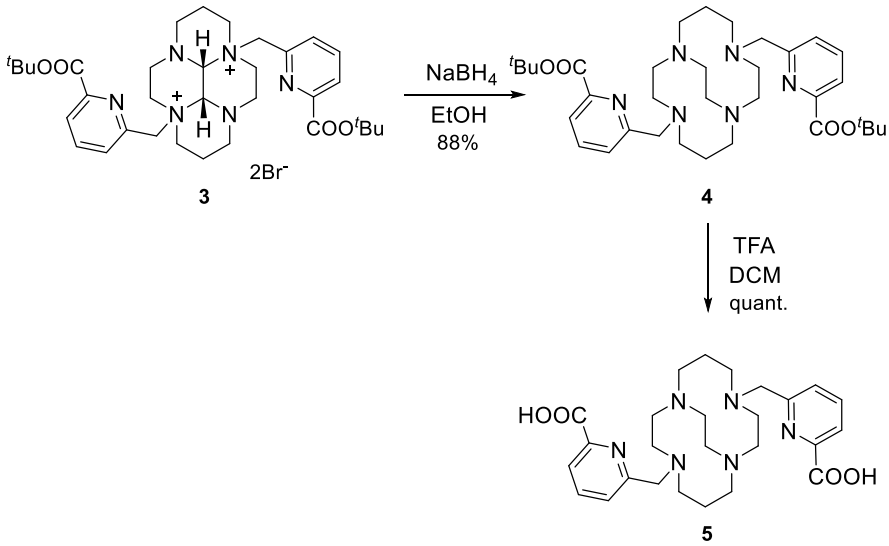


Figure 2.11: Synthesis of cb-tedpa **5**

We divided the first and the second alkylation in two steps for the sake of understanding its reactivity; however, those steps can clearly be combined in order to obtain the desired ligand **5** in only 3 steps starting from B and commercial cis-glyoxal cyclam, instead of 5 with the classical route[67]. Alkylation of cis-glyoxal cyclam is faster with **1** than with BnBr. This route also offers the advantage of avoiding the hydrogenation to remove the benzyl groups and presents higher yields.

2.2.3. Structure of cis-glyoxal cyclam and kinetics of alkylation

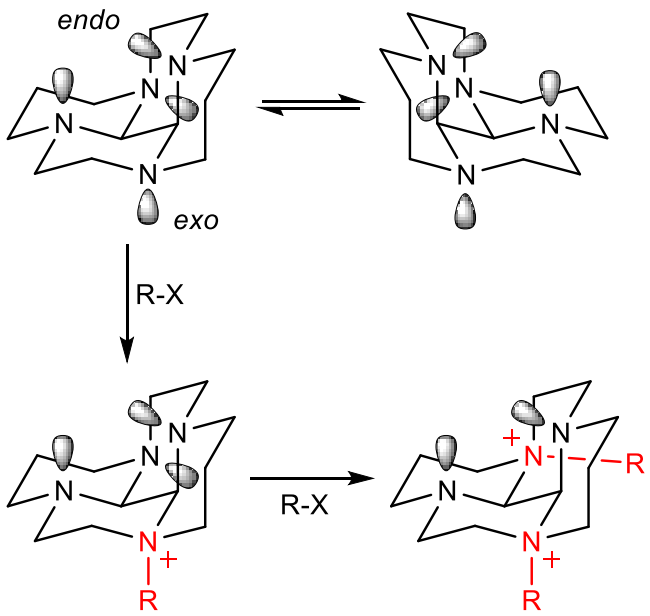


Figure 2.12: Regioselective bisaminal alkylation of cis-glyoxal cyclam

As we have seen earlier, bisaminal alkylation of cis-glyoxal cyclam is both regioselective and very slow. An explanation to this peculiar reactivity may reside in the 3D structure of cis-glyoxal cyclam. The molecule possesses 4 nitrogen atoms: two homotopic exo nitrogens which have lone pairs which protrude from the convex face of the molecule, and two homotopic endo nitrogens which have lone pairs which are more sterically concealed, pointing towards the concave face (fig. 2.12). Enantiomerization of cis-glyoxal cyclam, which also exchanges the exo and endo nitrogens, is fast on the laboratory time scale[63].

The first alkylation will take place on one of the exo nitrogens, more accessible, which will block the enantiomerization process[63]. The highly

regioselective second alkylation, which is much slower than the first, occurs on the remaining exo nitrogen, which is less sterically hindered than the two remaining endo nitrogens.

2.2.4 Complex stability in very acidic conditions

Lanthanide-complexes' behavior is often studied in acidic conditions, and that is because lanthanide decomplexation is often acid-catalyzed[73]. [Eu(cb-tedpa)] was shown to be stable in acidic conditions[67]. In order to better characterize the behavior of [Ln(cb-tedpa)] complexes in solution and to compare it to other complexes we prepared [Eu(cb-tedpa)].

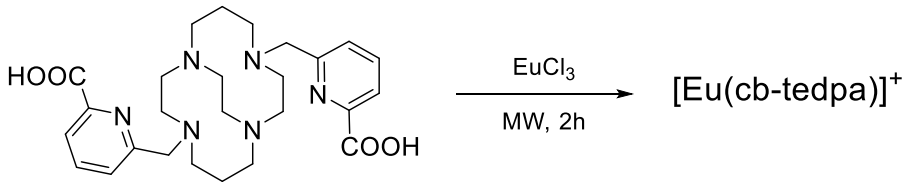


Figure 2.13: Complexation of Eu by cb-tedpa

We then recorded the NMR spectra of [Eu(cb-tedpa)] in D_2O at 25°C, at pH 7 and in acidic conditions (h=0.2 M). The spectra are reported in figure 2.14. We can see that the peaks are very sharp, which indicates very little dynamic mobility. This is true both in neutral and acidic conditions. Upon chelation, NMR signals are largely shifted, mostly downfield, because of the paramagnetism of Europium. There are some in differences in chemical shift between the two spectra, but peaks remain sharp, which suggests that there is no acid-catalyzed chemical exchange between, for example, a protonated and a deprotonated form.

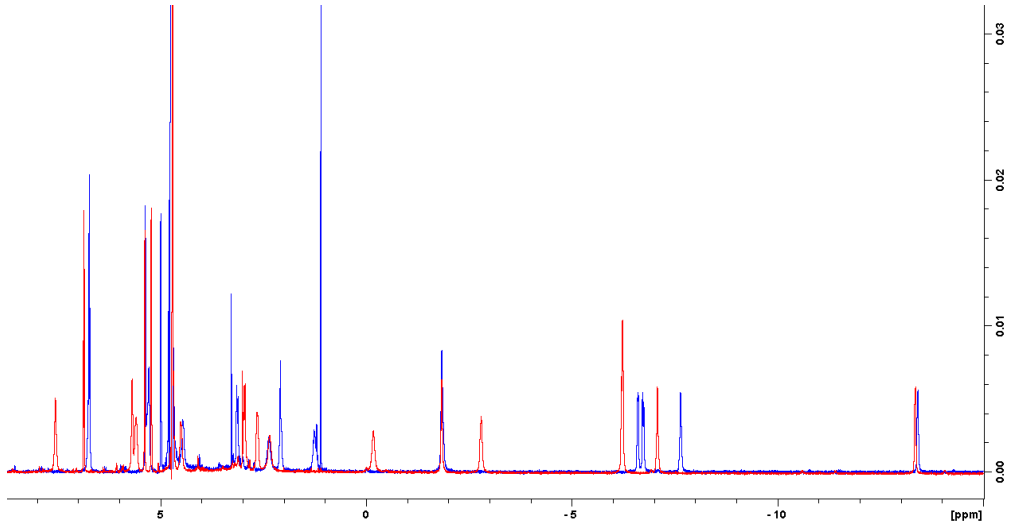


Figure 2.14: ^{1}H NMR Spectra of [Eu(cb-tedpa)] at pH 7 (red) and with [D^{+}]=0.2 M (blue)

In order to compare the stability of [Eu(cb-tedpa)] with [Eu(DOTA)], a standard in the field, the DOTA complex was also prepared. Its ^{1}H 1D NMR spectra was recorded (fig. 2.24). A first observation is that the peaks of [Eu(DOTA)] are much wider than the ones of [Eu(cb-tedpa)], and that is probably linked to the conformational mobility of DOTA complexes in solution[74]. Furthermore, we can clearly see two sets of peaks corresponding to the two isomers TSAP/SAP (c.f. Fig. 1.8). Both complexes were then dissolved in the same solution of D_2O with h= 0.2 M, and incubated at 37°C. Their ^{1}H 1D NMR spectra were then recorded and followed over 4 days. Interestingly, the spectra of [Eu(cb-tedpa)] were all identical, indicating no evident sign of release over this period (fig. 2.15). On the other hand, in the spectra of [Eu(DOTA)], we could after 4 hours already see peaks of free DOTA, and the release was almost complete after 4 days (fig. 2.25). Those results are in line with previous work from Wang *et.al*[75] who studied the acid-catalyzed decomposition of [Ln(DOTA)] complexes.

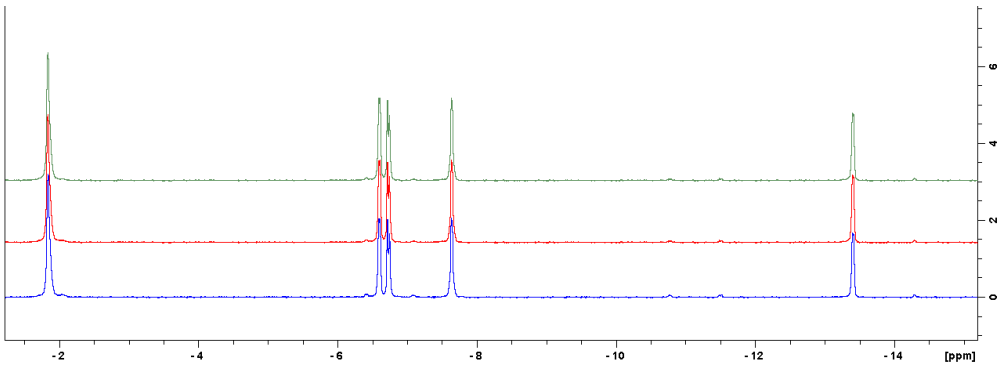


Figure 2.15: ^{1}H NMR Spectra of [Eu(cb-tedpa)] at [H^{+}]=0.2 M and incubated at 37°C for 4h (blue), 20h (red) and 4 days (green)

The fact that no apparent release of Europium by cb-tedpa can be observed while [Eu(DOTA)] decomplexation is almost complete in the same conditions further illustrate the potential and the interest of cross-bridge cyclam-based lanthanide chelators. NMR data indicates that [Eu(cb-tedpa)] does not undergo chemical exchange even under very acidic conditions and that this rigidity probably impedes protonation of the macrocycle which would favor lanthanide-decomplexation.

2.3 Synthesis of asymmetric cross-bridge cyclam derivatives

In an effort to expand the utility of cross-bridged cyclams as lanthanide chelators, we sought to synthesize new asymmetric cb-cyclam derivatives. We were interested in studying the impact of the coordination on the stability and the physical properties of our ligands, and therefore we selected **9** as a target molecule. **9** has two carboxylates, as cb-tedpa, which is useful to compensate the +III charge of the lanthanide, although it lacks one pyridine and therefore one coordinating atom.

As we have seen in **2.2**, the mono-alkylation of cis-glyoxal cyclam is quantitative and no sign of dialkylated compound was detected via NMR. We therefore tried to apply the synthetic methodology developed in 2.2 to synthesize **9** by selectively alkylating one position and then the other with two different alkylating agents.

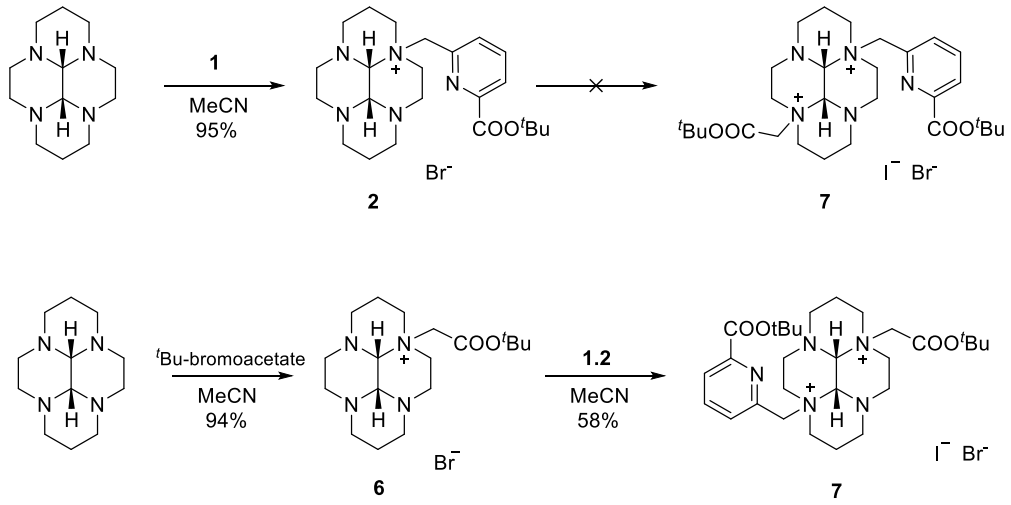


Figure 2.16: Reactivity of the bisaminal alkylation of cis-glyoxal cyclam

In a first time, we decided to start from **2** and use ^{t}Bu-iodoacetate to get to **7**. Even though -I was used as the leaving group instead of -Br, the reaction was found to be very slow. Since ^{t}Bu-bromoacetate seemed to be less reactive than **1**, we then chose to alkylate first with ^{t}Bu-bromoacetate and then with tert-butyl-6-(iodomethyl)picolinate **1.2**. Indeed, we found the second alkylation being faster with **1.2** than with ^{t}Bu-iodoacetate. The two deprotection steps were conducted as in section 2.2., reductive double ring opening was performed in good yield and the ^{t}Bu-esters were deprotected in TFA. Final **9** was obtained as a TFA salt after purification via semi-preparative HPLC. Thus showing that the new synthetic strategy described in this section also allows for the fast synthesis of asymmetrically alkylated cross-bridge cyclams.

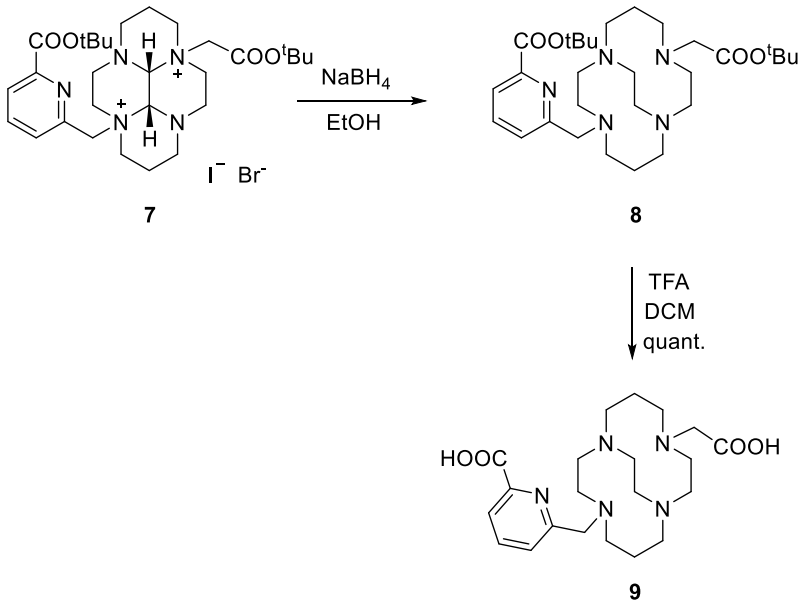


Figure 2.17: Synthesis of **9**

2.4 Synthesis of cross-bridge-cyclen-based ligands through direct bisaminal alkylation

As we have seen in the introduction, cross-bridge cyclens have also been used as chelators for transition metals. They have however, to the best of our knowledge, never been used as chelating platforms for lanthanides. In this section, we will try and apply the methodology described above to synthetize and characterize cross-bridge cyclen derivatives and their complex with lanthanides.

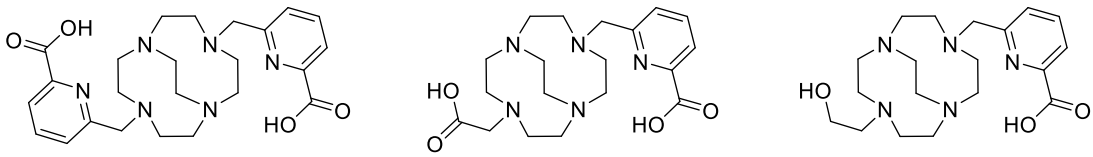


Figure 2.18: Novel cross-bridge cyclen-based ligands

2.4.1. Synthesis of cb-dodpa 12

In a first time we were interested in synthesizing symmetrical cb-dodpa **12**. Following the procedure developed in section 2.2, we could isolate the desired ligand in high yields (fig. 2.19). Interestingly, the reactions of alkyations and reductive double ring expansion were much faster than the ones with cb-cyclam (three days instead of ten). Those results are in line with findings from Niu[65], in which they obtained cross-bridge cyclen with shorter reaction times than for cross-bridge cyclam.

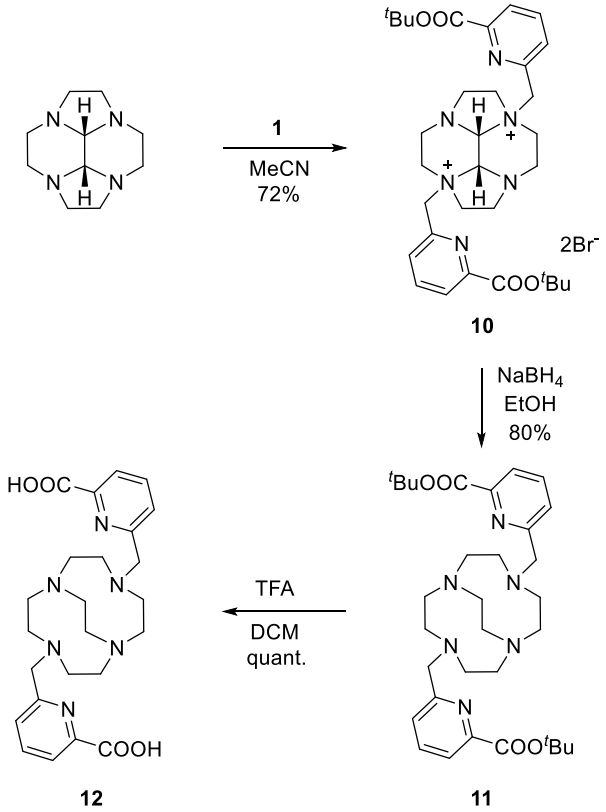


Figure 2.19: Synthesis of **12**

2.4.2. Synthesis of asymmetrical ligands

Having symmetrical **12** in hand, we then sought to synthesize asymmetric ligands, while still following the same strategy. Our first plan was to use **13** as a common intermediate for the synthesis of diverse ligands with different pendant arms. However, the second alkylation being much slower than the first one, we found it is very important to consider the reactivity of the alkylating agent when planning the synthesis. We found that regarding the synthesis of **14**, it was possible to alkylate first with the picolinic ester and then with ^{t}Bu-bromo acetate (fig. 20).

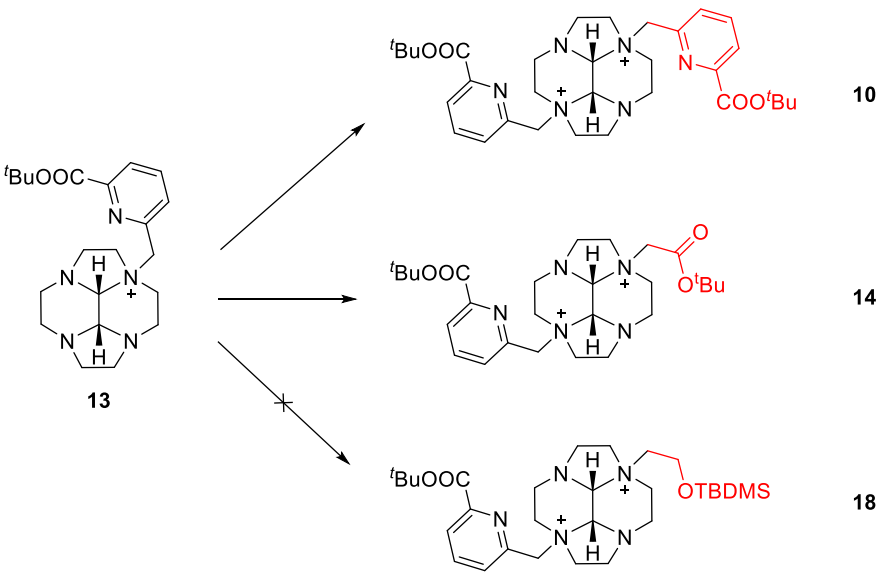


Figure 2.20: Reactivity of **13** towards alkylation

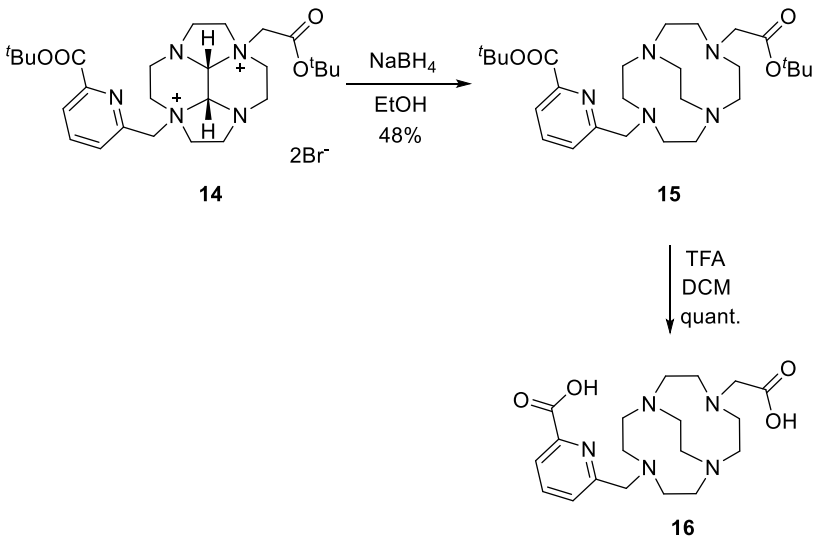


Figure 2.21: Synthesis of asymmetric **16**

Following our strategy, we were then able to isolate asymmetric **16**, bearing a picolinic acid and an acetate as pendant arms. On the other hand, we found that TBDMS-bromoethanol was much less reactive and therefore not usable as a second alkylating agent. However, by using first TBDMS-bromoethanol and then **1** it was possible to isolate asymmetric **18**, which was then deprotected following our strategy to yield ligand **20**.

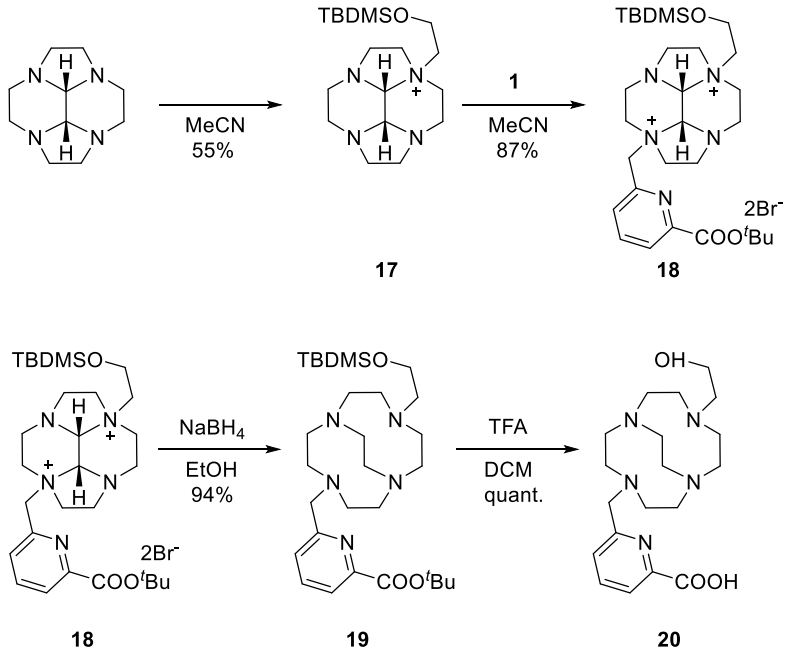


Figure 2.22: Synthesis of **20**

In this section, we have shown that our new synthetic strategy towards alkylated cross-bridge cyclams was also applicable to the synthesis of alkylated cross-bridge cyclen ligands. We have shown that it was possible to synthesize different ligands and that even in the case of asymmetric ligands, little effort was required to isolate the desired product. Last, because of the intrinsic reactivity of cis-glyoxal cyclen towards alkylation, reaction times appeared to be much shorter for cyclen-derivatives than for cyclam ones. Future experiments would

consist in the study of complexation by those ligands of lanthanides and other transition metals. It would be interesting to compare the properties of those complexes with the ones reported for [Ln(cb-tedpa)]. Those experiments are yet to be performed.

2.5 Conclusion

In this chapter, we have presented and discussed a novel synthetic strategy for the synthesis of cross-bridge azamacrocycles. Taking the example of cb-tedpa, we showed that this new strategy was more efficient than the already-described one. More, we were able to use our new pathway to synthesize a novel asymmetric cross-bridge cyclam-based ligand in only six steps from commercial reagents. However, we have also identified the rate of the second alkylation of cis-glyoxal cyclam as being a potentially limiting step in this synthesis and the reactivity of the alkylating agent is an important parameter to consider when planning such synthesis.

Novel cross-bridge cyclen-based ligands were also isolated following the same pathway. Once again, we have shown that the rate of the second alkylation of cis-glyoxal cyclen is a potential limitation in this synthesis; however, we also illustrated that the reaction times are much shorter in the case of cyclen-based ligands than in the case of cyclam-based ligands.

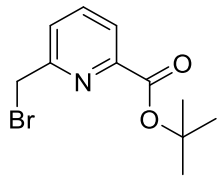
Last, having synthesized the complex [Eu(cb-tedpa)], we showed that its inertness in acidic media is higher than the one of [Eu(DOTA)], a reference in this domain. This further illustrates the potential of cross-bridge azamacrocyles in the field of lanthanide-complexation and therefore enhances the interest and potential applications of the synthetic strategy presented in this chapter. This strategy is attractive because it allows for the tuning of the coordination environment of lanthanides by easily synthesizing different ligands from similar

building blocks. It could help improving the properties and applicability of lanthanide complexes in structural biology and spectroscopy.

2.6 Experimental procedures

2.6.1 Synthetic procedures

Tert-butyl 6-(bromomethyl)picolinate **1**

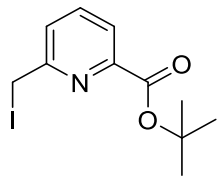


tert-butyl 6-methylpicolinate (1.46 g, 7.56 mmol), NBS (943 mg, 5.30 mmol), and Benzoyl Peroxide (LUPEROX 75%, 21 mg) are dissolved in CCl_4 (10 mL). The reaction mixture is kept at reflux for 4 hours after which it is filtrated. The filtrate is then purified via Flash Chromatography (7/3 Petroleum ether/Ethyl Acetate). Obtention of 810 mg of the pure product (40%).

^{1}H NMR (500 MHz, CDCl3): δ 7.95 (1H, d, J=7.81Hz), 7.84 (1H, t, J=7.80 Hz), 7.67 (1H, d, J=7.78 Hz), 4.67 (2H, s), 1.65 (9H, s)

ESI-MS: m/z=271.83 Th

Tert-butyl 6-(iodomethyl)picolinate **1.2**



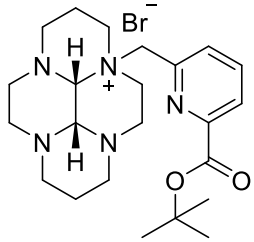
1 (130 mg, 0.48 mmol) was dissolved in dry acetone (640 µL). Sodium Iodide (145 mg, 2 eq) was added and the reaction was let stirred for 22 hours covered

in aluminum foil. The reaction mixture was then diluted with 5 mL DCM and washed with 3 mL H_20. The aqueous phase is then re extracted with 2x3 mL DCM. Organic phases are combined and dried over $MgSO_4$ to yield 159 mg of pure product (quant.).

^{1}H NMR (500 MHz, CDCl3): δ 7.86 (1H, d, J=7.82Hz), 7.76 (1H, t, J=7.79 Hz), 7.60 (1H, d, J=7.78 Hz), 4.60 (2H, s), 1.62 (9H, s)

^{13}C NMR (125 MHz, CDCl3): δ 163 .63, 158.91, 148.93, 137.90, 126.34, 123.44, 82.41, 28.09, 5.55

(3a1S,5a1R)-5a-((6-(tert-butoxycarbonyl)pyridin-2-yl)methyl) dodecahydro-1H-3a,5a,8a,10a-tetraazapyren-5a-ium bromide **2**



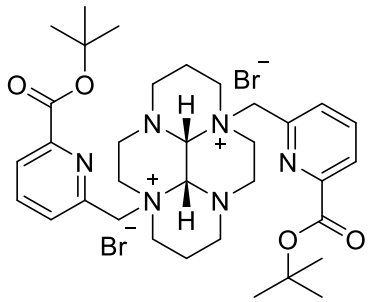
Cis-glyoxal cyclam (700 mg) was dissolved in acetonitrile (24 mL), and **1** (1 eq, 857 mg), was added. After 2 days at room temperature, solvent was removed and the obtained solid was triturated with diethyl ether to yield 1.47 g of a white solid (95%)

^{1}H NMR (500 MHz, DMSO-d6): δ 8.15 (1H, t, 7.76 Hz), 8.11 (1H, d, J=8.04 Hz), 7.85 (1H, d, J=7.74 Hz), 5.33 (1H, d, J=13.46 Hz), 4.89 (1H, d, J=13.46 Hz), 4.24 (1H, dt, J=12.71 Hz, 3.36 Hz), 4.20 (1H, d, J=2.04 Hz), 3.92 (1H, s), 3.61 (1H, dt, J=13.01 Hz, 3.09 Hz), 3.49 (1H, dt, J=11.49 Hz, 2.37 Hz), 3.20 (1H, dt, J=13.4 Hz, 2.36 Hz), 2.97-2.89 (6H, m), 2.54 (1H, dt, J=12.05 Hz, 3.01 Hz), 2.38 (1H, dt, J=11.62 Hz, 3.22 Hz), 2.29-2.20 (2H, m), 2.08 (2H, s), 1.66 (1H, d, J=14.21 Hz), 1.57 (9H, s), 1.25 (1H, d, J=14.21 Hz)

^{13}C NMR (125 MHz, DMSO-d6): δ 163.45, 149.69, 149.20, 139.74, 131.11, 125.51, 118.54, 83.03, 82.19, 69.69, 60.94, 60.15, 54.27, 52.52, 52.40, 49.55, 46.60, 42.70, 19.33, 1.63

ESI-MS: m/z= 414.33 Th

mono((3a1R,5a1R)-5a,10a-bis((6-(tert-butoxycarbonyl)pyridin-2-yl) methyl)tetradecahydro-3a,5a,8a,10a-tetraazapyrene-5a,10a-diium) monobromide **3**

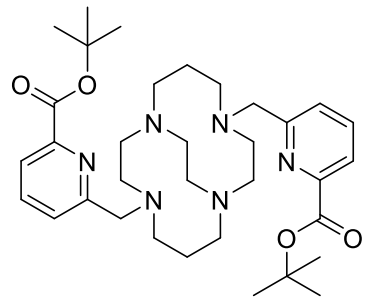


3 (232 mg) was dissolved in acetonitrile (3 mL), and **1** (4 eq, 600 mg), was added. After 10 days at room temperature, solvent was removed and the oil was triturated with diethyl ether to yield 353 mg of a white solid (92%)

^{1}H NMR (500 MHz, DMSO-d6): δ 8.22 (t, J= 7.75 Hz, 2H), 8.15 (d, J=7.75 Hz, 2H), 7.95 (d, J=7.75 Hz, 2H), 5.31 (d, J=13.45 Hz, 2H), 5.11 (d, J=13.45 Hz, 2H), 4.45 (t, J=13.21, 2H), 3.66 (d, J=11.36 Hz, 2H), 3.47 (t, J=12.61 Hz, 4H), 2.98 (d, J=11.34 Hz, 2H), 2.79 (t, J=12.69 Hz, 2H), 2.23 (q, J=13.28 Hz, 2H), 1.78 (d, J=14.63 Hz, 2H), 1.47 (s, 18H).

ESI-MS: m/z= 303.25 Th

di-tert-butyl 6,6'-((1,4,8,11-tetraazabicyclo[6.6.2]hexadecane-4,11-diyl) bis(methylene))dipicolinate **4**

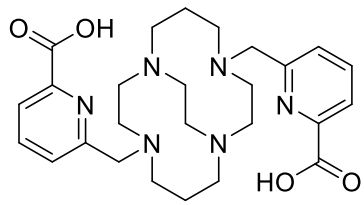


3 (350 mg, 0.43 mmol) was dissolved in ethanol 95% (12 mL), and Sodium Borohydride (6 eq, 98 mg), were added in portion. After 3 days at room temperature, 12 mL H_2O are added to the solution and the pH is lowered to 2 with HCl 3%. The solution is saturated with KOH and extracted with 3x50 mL DCM. Organic phases are combined and dried over Na_2SO_4. Solvent is removed to yield 230 mg of a white solid (88%)

^{1}H NMR (500 MHz, MeOD d4): δ 7.90 (2H, d, J=7.75 Hz), 7.85 (2H, t, J=7.67 Hz), 7.50 (2H, d, J=7.71 Hz), 3.99 (2H, d, J=15.09 Hz), 3.84 (2H, d, J=15.09 Hz), 3.15-3.04 (4H, m), 3.00-2.84 (12H, m), 2.65-2.55 (4H, m), 1.82 (broad s, 4H), 1.64 (18H, s)

^{13}C NMR (125 MHz, MeOD-d4): δ 158.41, 148.69, 138.04, 133.78, 126.73,123.18, 82.12, 58.43, 54.57, 38.33, 38.16, 37.99, 37.83, 26.98, 26.92

6,6'-((1,4,8,11-tetraazabicyclo[6.6.2]hexadecane-4,11-diyl)bis(methylene) dipicolinic acid **cb-tedpa 5**



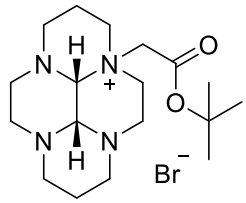
4 (100 mg, 0.16 mmol) was dissolved in DCM (2 mL), and Trifluoro acetic acid (1mL) was added. After 5 hours at room temperature, DCM and TFA are evaporated in vacuo, yielding the product as the TFA salt (oil, quant.)

^{1}H NMR (500 MHz, D2O): δ 7.79 (2H, d, 7.68 Hz), 7.56 (2H, t, J=7.93 Hz), 7.03 (2H, d, J=7.82 Hz), 4.97 (2H, d, J=16.34 Hz), 4.08 (2H, dt, J=14.04 Hz, 3.87 Hz), 4.04 (2H, d, J=16.34 Hz), 3.92 (2H, dt, J=14.38 Hz, 2.93 Hz), 3.62-3.41 (6H, m), 3.17 (2H, broad d, J=15.37 Hz), 2.99 (2H, broad d, 14.14 Hz), 2.84 (2H, bd, J=12.99 Hz), 2.63 (2H, broad d, J=13.47 Hz), 2.57 (2H, q, 5.52 Hz), 2.52 (2H, broad q, 14.52 Hz), 1.62 (2H, bd, 16.78 Hz),

^{13}C NMR (125 MHz, D2O): δ 167.70, 151.38, 147.00, 138.96, 126.19, 125.33, 56.17, 55.06, 54.95, 53.33, 50.16, 48.39, 18.38

ESI-MS: m/z= 497.58 Th

(3a1S,5a1R)-5a-(2-(tert-butoxy)-2-oxoethyl)dodecahydro-1H-3a,5a,8a,10a-tetraazapyren-5a-ium bromide **6**

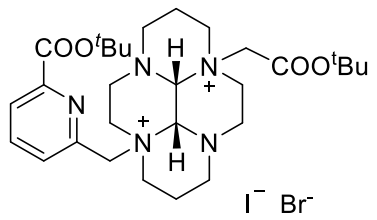


Cis-glyoxal cyclam was dissolved in acetonitrile (3mL), and *tert*-butyl bromoacetate (1 eq, 88 mg, 67 µL), was added. After 3 days at room temperature, solvent was removed and the transparent oil was triturated with diethyl ether to yield 176 mg of a white solid (94%)

^{1}H NMR (500 MHz, DMSO-d6): δ 4.83 (1H, d, J=16.2 Hz), 4.49 (1H, d, J=16.1 Hz), 4.28 (1H, dt, J=12.75 Hz, 3.83 Hz), 3.92 (2H, d, J=12.75 Hz), 3.77 (1H, d, J=12.42), 3.60 (1H, d, J=12.42 Hz), 3.46 (2H, t, J=12.47 Hz), 3.01-2.71 (m, 10H), 2.34 (1H, dt, J=11.54 Hz, 3.83 Hz), 2.23 (2H, d, J=11.71 Hz), 1.75 (1H, d, J=12.10 Hz), 1.50 (9H, s), 1.21 (1H, d, 13.10 Hz)

^{13}C NMR (125 MHz, DMSO-d6): δ 164.02, 84.98, 82.72, 69.25, 60.75, 57.70, 54.19, 53.86, 52.49, 52.30, 50.26, 46.01, 42.57, 28.03, 19.23, 19.12

(3a1R,5a1R)-10a-(2-(tert-butoxy)-2-oxoethyl)-5a-((6-(tert-butoxycarbonyl) pyridin-2-yl)methyl)tetradecahydro-3a,5a,8a,10a-tetraazapyrene-5a,10a-diium bromide iodide **7**

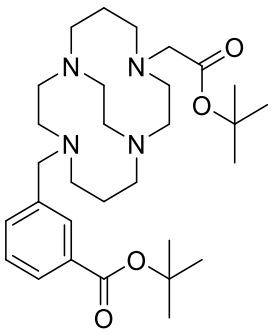


6 (170 mg, 0.41 mmol) was dissolved in acetonitrile (5mL), and **1.2** (4 eq, 523 mg), was added. After 10 days at room temperature, solvent was removed and

the hygroscopic brown oil was triturated with diethyl ether to yield 176 mg of a brown solid (58%)

ESI-MS: m/z= 264.17 Th

tert-butyl 3-((11-(2-(tert-butoxy)-2-oxoethyl)-1,4,8,11-tetraazabicyclo [6.6.2]hexadecan-4-yl)methyl)benzoate **8**

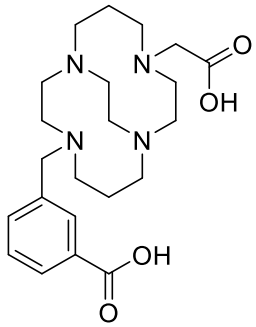


7 (176 mg) was dissolved in EtOH 95% (10 mL), and $NaBH_4$ (120 mg), was added. After 3 days at room temperature, reaction mixture was diluted with 20 mL H_2O and pH was lowered to 3 with HCl 3%. The solution was washed with 2x50 mL DCM, the aqueous phase was saturated with KOH, and then extracted with 3x50 mL DCM. Organic layers were combined and dried. Mixture was then purified by semi-preparative HPLC to yield 48 mg of product.

^{1}H NMR (500 MHz, CDCl3): δ 8.19 (d, J=7.87 Hz, 1H), 8.13 (t, J=7.87 Hz, 1H), 7.70 (d, J=7.67 Hz, 1H), 5.06 (d, J=14.67 Hz, 1H), 5.00 (d, 16.81 Hz, 1H), 4.32 (d, J=14.42 Hz, 1H), 4.26 (t, J=14.63, 1H), 4.14 (t, J=13.50 Hz, 1H), 3.91 (t, J=13.71 Hz, 1H), 3.75 (t, J=14.19 Hz, 1H), 3.60 (m, 1H), 3.49 (d, J=16.87 Hz, 1H), 3.30-3.13 (m, 6H), 3.11-2.86 (m, 5H), 2.83-2.73 (m, 2H), 2.64-2.55 (bd, 3H), 1.80 (bd, J=15.87 Hz, 1H), 1.69 (s, 9H), 1.03 (s, 9H)

ESI-MS: m/z=532.67 Th

3-((11-(carboxymethyl)-1,4,8,11-tetraazabicyclo[6.6.2]hexadecan-4-yl)methyl)benzoic acid **9**

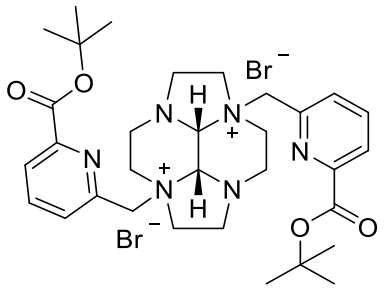


8 (48 mg) was dissolved in DCM (1mL), and 1 mL TFA was added. After 8 hours, reaction was complete and quantitative.

^{1}H NMR (500 MHz, D2O): δ 8.10 (d, J=7.77 Hz, 1H), 7.98 (t, J=7.88 Hz, 1H), 7.53 (d, J=7.77 Hz, 1H), 4.97 (d, J=17.11 Hz, 1H), 4.23 (dd, J=14.08, 5.10 Hz, 2H), 3.96 (t, J=14.33 Hz, 1H), 3.87 (t, J=14.09 Hz, 1H), 3.74 (t, J=14.43 Hz, 1H), 3.59 (td, J=13.05, 4.27 Hz, 1H), 3.53-3.36 (m, 3H), 3.33-3.25 (m, 3H), 3.19-2.98 (m, 5H), 2.92 (s, 1H), 2.89-2.77 (m, 4H), 2.75 (d, J=2.82 Hz, 2H), 2.62-2.55 (m, 1H), 2.50-2.30 (m, 3H), 1.64 (bd, J=16.23 Hz, 2H)

ESI-MS: m/z= 420.42 Th

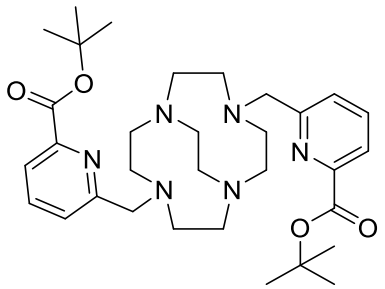
((2a1R,4a1R)-4a,8a-bis((6-(tert-butoxycarbonyl)pyridin-2-yl)methyl)dodecahydro-2a,4a,6a,8a-tetraazacyclopenta[fg] acenaphthylene-4a,8a-diium) dibromide **10**



Cis-glyoxal cyclen (93 mg, 0.48 mmol) was dissolved in dry acetonitrile (5 mL) and was added 1 (300 mg, 2.3 eq). The mixture was stirred overnight at room temperature, during which a white precipitate was form. Solid was isolated by

filtration, washed with acetonitrile and diethyl ether to yield 524 mg of the pure product (72%)

di-tert-butyl 6,6'-((1,4,7,10-tetraazabicyclo[5.5.2]tetradecane-4,10-diyl)bis(methylene))dipicolinate **11**

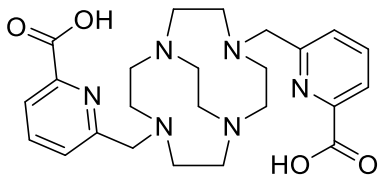


10 (250 mg, 0.34 mmol) was dissolved in 95% EtOH (10 mL), and Sodium Borohydride (4 eq, 52 mg), was added. The mixture was stirred overnight at room temperature. Reaction is then poured on 20 mL H_2O and extracted with 3x20 mL DCM. Comnined organic phases are dried over Na_2SO_4 and solvent is removed in vacuo. 157 mg (80%)

^{1}H NMR (500 MHz, MeOD): δ 7.95 (2H, d, J=7.75 Hz), 7.89 (2H, t, J=7.76 Hz), 7.71 (2H, d, J=7.76 Hz), 4.17 (4H, s), 3.23-3.15 (12H, m), 3.09-2.97 (8H, m), 1.61 (18H, s)

^{13}C NMR (125 MHz, MeOD): δ 164.12, 159.32, 148.57, 137.71, 126.39, 123.25, 82.14, 59.21, 55.87, 51.68, 27.11

6,6'-((1,4,7,10-tetraazabicyclo[5.5.2]tetradecane-4,10-diyl)bis(methylene))dipicolinic acid **12**



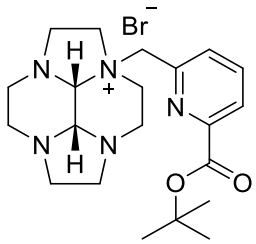
11 (175 mg, 0.27 mmol) was dissolved in DCM (2 mL), and Trifluoroacetic acid (30 eq, 625 uL), was added. The mixture was stirred at room temperature for 3 hours when DCM and TFA are removed in vacuo. Obtention of 157 mg of the product as the TFA salt.

^{1}H NMR (500 MHz, D2O): δ 8.27 (2H, bt, J=8Hz), 8.17 (2H, bd, J=8Hz), 7.82 (2H, bd, J=8Hz), 4.67 (4H, s), 3.58-3.35 (12H, m), 3.24 (4H, s), 3.07 (4H, bd, J=12.73 Hz),

^{13}C NMR (125 MHz, D2O): δ 176.55, 164.81, 152.88, 143.69, 127.77, 125.75, 58.32, 55.05, 43.86, 20.48

ESI-MS: m/z= 373.25 Th

(2a1S,4a1R)-4a-((6-(tert-butoxycarbonyl)pyridin-2-yl)methyl)decahydro-3H-2a,4a,6a,8a-tetraazacyclopenta[fg]acenaphthylen-4a-ium bromide **13**



Cis-glyoxal cyclen (500 mg) was dissolved in dry acetonitrile (20 mL), and **1** (1 eq, 700 mg), was added. The mixture was stirred overnight at room temperature, solvent was removed to yield 1.2 g of a white solid (quant)

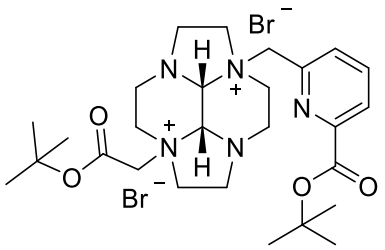
^{1}H NMR (500 MHz, CDCl3): δ 8.67 (2H, d, J=7.73 Hz), 8.09 (2H, d, J=7.73 Hz), 7.98 (2H, t, J=7.73 Hz), 5.87 (2H, d, J=13.55 Hz), 5.28 (2H, d, J=13.68 Hz), 4.86 (2H, dt, J=11.17, 3.86 Hz), 4.78 (2H, d, J = 2.78 Hz), 4.50 (2H, bd, J=12.17 Hz), 3.80 (2H, dt, J= 12.45, 7.38 Hz), 3.67 (2H, d, J=2.85 Hz), 3.50 (2H, dt, J=11.84, 7.82 Hz), 3.41-3.24 (12H, m), 3.07 (2H, bd, J=3.24 Hz), 2.93-2.83 (4H, m), 2.80-2.73 (4H, m), 2.48 (2H, m), 1.64 (18, s)

^{1}H NMR (500 MHz, D2O): δ 8.20 (2H, d, J=7.87 Hz), 8.09 (2H, t, 7.85 Hz), 7.86 (2H, d, J=7.84 Hz), 4.99-4.91 (6H, m), 4.31 (2H, bt, J=11.59 Hz), 4.00 (2H, bt,

11.62 Hz), 3.90-3.78 (4H, m), 3.60-3.54 (2H, bm), 3.42-3.27 (4H, bm), 3.06 (2H, bq, 8.02 Hz), 1.52 (18H, s)

^{13}C NMR (125 MHz, D2O): δ 164.82, 149.41, 147.51, 140.16, 130.60, 126.52, 84.71, 77.59, 61.95, 60.95, 56.77, 46.34, 43.03, 27.24

mono((2a1R,4a1R)-8a-(2-(tert-butoxy)-2-oxoethyl)-4a-((6-(tert-butoxycarbonyl)pyridin-2-yl)methyl)dodecahydro-2a,4a,6a,8a-tetraazacyclopenta[fg]acenaphthylene-4a,8a-diium) monobromide **14**

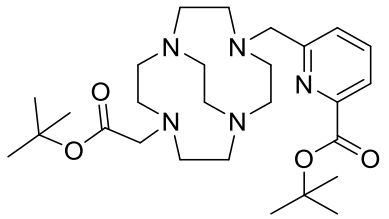


13 (800 mg, 1.71 mmol) was dissolved in dry acetonitrile (18 mL), and bromo-*tert*-butylacetate (2.4 eq, 804 mg), were added. The mixture was stirred at room temperature for 2 days, during which a white precipitate was form. Solid was isolated by filtration, washed with diethyl ether to yield 995 mg of the pure product (89%)

^{1}H NMR (500 MHz, DMSO): δ 8.18 (t, J=7.78 Hz, 1H), 8.15 (d, J=7.82 Hz), 7.98 (d, J=7.71 Hz, 1H), 5.17 (d, J=14.07 Hz, 1H), 4.98 (d, J=14.05 Hz, 1H), 4.76 (bs, 1H), 4.72 (d, J=17.17 Hz, 1H), 4.62 (bs, 1H), 4.55 (d, J=17.33 Hz, 1H), 4.36 (bt, J=11,02 Hz, 1H), 4.22-4.11 (m, 2H), 3.95-3.73 (m, 5H), 3.50 (dt, J=8.28, 3.58 Hz, 1H), 3.43 (dt, J=8.28, 3.58 Hz, 1H), 3.22 (bd, J=13.54 Hz, 1H), 3.10 (bt, J=12.75 Hz, 1H), 2.96 (q, J=8.71 Hz, 1H), 2.74 (q, J=8.58 Hz, 1H), 1.59 (s, 9H), 1.51 (ss, 9H)

^{13}C NMR (125 MHz, DMSO): δ 163.89, 163.08, 149.44, 149.18, 140.15, 130.51, 125.94, 85.10, 82.49, 79.19, 76.22, 65.38, 63.93, 61.95, 59.9

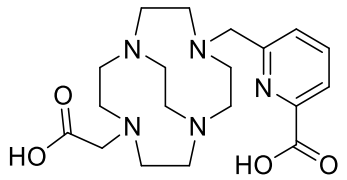
tert-butyl 6-((10-(2-(tert-butoxy)-2-oxoethyl)-1,4,7,10tetraazabicyclo[5.5.2] tetradecan-4-yl)methyl)picolinate **15**



14 (900 mg, 1.36 mmol) was dissolved in EtOH (40 mL), and $NaBH_4$ (8 eq, 412 mg), were added. The mixture was stirred at room temperature overnight, after which the pH was lowered to 3 using HCl 3%. 50 mL H_2O were then added, and the solution was saturated with KOH. Aqueous solution was extracted with 3x100 mL DCM. Organic layers were combined and dried to yield 330 mg of the pure product (48%)

^{1}H NMR (500 MHz, MeOD): δ 8.06-8.01 (m, 2H), 7.69 (dd, J=6.55, 2.38 Hz, 1H), 4.43 (bs, 2H), 3.63 (bs, 2H), 3.44-3.38 (bm, 2H), 3.26-3.17 (m, 8H), 3.17-3.10 (m, 5H), 3.04 (bt, J=3.41 Hz, 1H), 3.02 (bt, J=3.41 Hz, 1H), 1.66 (s, 9H), 1.42 (s, 9H)

6-((10-(carboxymethyl)-1,4,7,10-tetraazabicyclo[5.5.2]tetradecan-4-yl)methyl)picolinic acid **16**



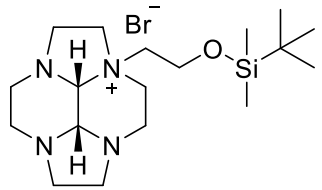
15 (300 mg, 0.6 mmol) was dissolved in DCM (4 mL), and TFA (1.4 mL), was added. The mixture was stirred at room temperature for 4 hours and DCM and TFA were removed in vacuo, yielding 470 mg of the pure product as a TFA salt (quant.)

^{1}H NMR (500 MHz, D2O): δ 8.01-7.94 (m, 2H), 7.56 (bd, J=6.98 Hz, 1H), 4.59 (bs, 2H), 3.72 (bs, 2H), 3.52-3.39 (bm, 4H), 3.39-3.27 (bm, 2H), 3.19-2.87 (bm, 14H)

^{13}C NMR (125 MHz, D2O): δ 174.58, 165.58, 151.39, 145.18, 141.77, 127.22, 125.62, 58.98, 58.90, 57.35, 55.21, 54.67, 53.48, 45.08

ESI-MS: m/z=392.42 Th

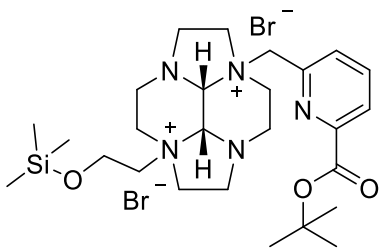
(2a1S,4a1R)-4a-(2-((tert-butyldimethylsilyl)oxy)ethyl)decahydro-3H-2a,4a,6a,8a-tetraazacyclopenta[fg]acenaphthylen-4a-ium bromide **17**



Cis-glyoxal cyclen (400 mg, 2.06 mmol) was dissolved in acetonitrile (5 mL), and 2-bromoethoxy-*tertbutyl* dimethyl silane (1 eq, 493 mg), was added. The mixture was stirred at room temperature, and after three days conversion was of around 58%. Solvent was removed and the obtained oi was triturated with Et_2O to yield 490 mg of an off-white solid (55%)

^{1}H NMR (500 MHz, DMSO): δ 4.13-4.00 (m, 2H), 3.93-3.82 (m, 4H), 3.76-3.58 (m, 4H), 3.19-3.08 (m, 2H), 3.07-3.00 (m, 2H), 2.78-2.58 (m, 6H), 2.41-2.34 (m, 2H), 2.26-2.20 (m, 2H), 0.89 (s, 9H), 0.12 (s, 6H)

mono((2a1R,4a1R)-4a-((6-(tert-butoxycarbonyl)pyridin-2-yl)methyl)-8a-(2-((trimethylsilyl)oxy)ethyl)dodecahydro-2a,4a,6a,8a-tetra azacyclopenta[fg]acenaphthylene-4a,8a-diium) monobromide **18**

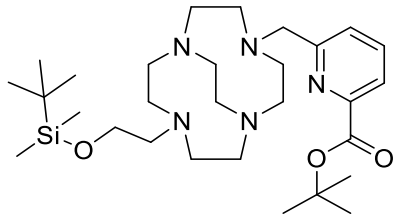


17 (470 mg, 1.08 mmol) was dissolved in dry acetonitrile (10 mL), and **1** (1.5 eq, 442 mg), was added. The mixture was stirred at room temperature for 2 days, during which a white precipitate was form. Solid was isolated by filtration, washed with diethyl ether to yield 660 mg of the pure product (87%)

^{1}H NMR (500 MHz, DMSO): δ 8.18 (t, J=7.78 Hz, 1H), 8.15 (d, J=7.82 Hz), 7.98 (d, J=7.71 Hz, 1H), 5.17 (d, J=14.07 Hz, 1H), 4.98 (d, J=14.05 Hz, 1H), 4.76 (bs, 1H), 4.72 (d, J=17.17 Hz, 1H), 4.62 (bs, 1H), 4.55 (d, J=17.33 Hz, 1H), 4.36 (bt, J=11,02 Hz, 1H), 4.22-4.11 (m, 2H), 3.95-3.73 (m, 5H), 3.50 (dt, J=8.28, 3.58 Hz, 1H), 3.43 (dt, J=8.28, 3.58 Hz, 1H), 3.22 (bd, J=13.54 Hz, 1H), 3.10 (bt, J=12.75 Hz, 1H), 2.96 (q, J=8.71 Hz, 1H), 2.74 (q, J=8.58 Hz, 1H), 1.59 (s, 9H), 1.51 (ss, 9H)

^{13}C NMR (125 MHz, DMSO): δ 163.89, 163.08, 149.44, 149.18, 140.15, 130.51, 125.94, 85.10, 82.49, 79.19, 76.22, 65.38, 63.93, 61.95, 59.9

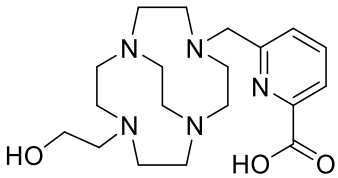
tert-butyl 6-((10-(2-((tert-butyldimethylsilyl)oxy)ethyl)-1,4,7,10-tetraazabicyclo[5.5.2]tetradecan-4-yl)methyl)picolinate **19**



19 (630 mg, 0.89 mmol) was dissolved in EtOH 95% (25 mL), and $NaBH_4$ (6 eq, 203 mg), were added. The mixture was stirred at room temperature overnight, after which the pH was lowered to 5 using HCl 3%. 25 mL H_2O were then added, and the solution was saturated with KOH. Aqueous solution was extracted with 3x60 mL DCM. Organic layers were combined and dried to yield 470 mg of the pure product (94%)

^{1}H NMR (500 MHz, MeOD): δ 7.93 (t, J=7.89 Hz, 1H), 7.83 (d, J=7.88 Hz, 1H), 7.59 (d, J=7.91 Hz, 1H), 3.85 (s, 2H), 3.60 (t, J=6.01 Hz, 2H), 3.09-2.83 (m, 12H), 2.82-2.60 (m, 20H), 1.55 (s, 9H), 0.87 (s, 9H), 0.36 (s, 6H)

6-((10-(2-hydroxyethyl)-1,4,7,10-tetraazabicyclo[5.5.2]tetradecan-4-yl)methyl)picolinic acid **20**



19 (470 mg, 0.86 mmol) was dissolved in DCM (6 mL), and TFA (2 mL), was added. The mixture was stirred at room temperature for 4 hours and DCM and TFA were removed in vacuo, yielding 470 mg of the pure product as a TFA salt (quant.)

^{1}H NMR (500 MHz, D2O): δ 8.26-8.20 (m, 2H), 7.81 (dd, J=7.35, 1.74 Hz, 1H), 4.58 (s, 2H), 3.92 (bt, J=4.81 Hz, 2H), 3.71 (bt, J=3.54 Hz, 1H), 3.68 (bq, J=2.42 Hz, 1H), 3.61-3.53 (m, 2H), 3.51-3.43 (m, 6H), 3.23-3.15 (m, 4H), 3.07-2.92 (m, 8H),

2.6.2 Lanthanide Chelation and NMR

Synthesis of ***[Eu(cb-tedpa)]***

Cb-tedpa (50 mg) was dissolved in 2 mL H2O and 1 mL MeCN. pH was adjusted to 7 and $EuCl_3$.6H2O (50 mg) was added to the solution. The mixture was placed in a micro-wave vial irradiated at 150W, 150°C for 2 hours. The crude mixture was then purified via semi-preparative HPLC.

^{1}H NMR (500 MHz, D2O): δ 7.54 (bs), 6.85 (t, J=7.21 Hz), 5.68 (bs, 5.58 (bs), 5.38 (d, J=7.20 Hz), 5.21 (d, J=7.89 Hz), 4.47 (bs), 3.00 (s), 2.95 (bd, J=15.49 Hz), 2.64 (bs), 2.35 (bs), -0.15 (bs), -1.84 (s), -2.80 (bs), -6.24 (m, 2H), -7.09 (bs), -13.35 (bs)

ESI-MS: m/z= 647.50 (100%), 645.58 (85%)

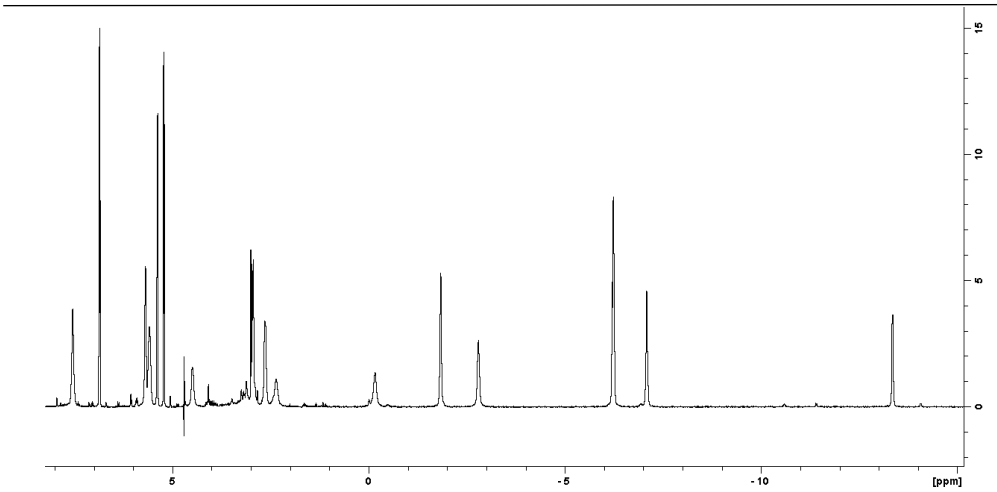


Figure 2.23: ^{1}H 1D NMR Spectra of [Eu(cb-tedpa)]

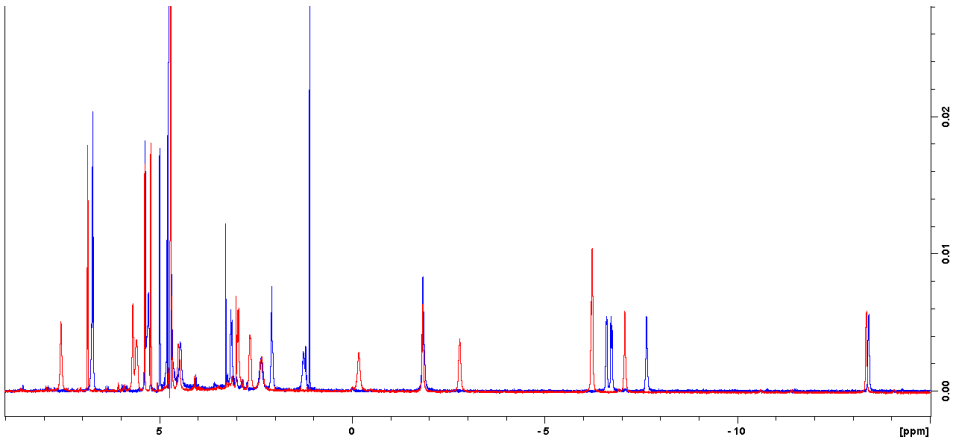


Figure 2.14: 1H NMR Spectra of [Eu(cb-tedpa)] at pH 7 (red) and with [H^{+}]=0.2 M (blue)

*Synthesis of **[Eu(DOTA)]***

DOTA (50 mg) was dissolved in 2 mL H_2O. pH was adjusted to 7 and $EuCl_3$.6H2O (50 mg) was added to the solution. The mixture was placed refluxed overnight. The crude mixture was then purified via semi-preparative HPLC.

^{1}H NMR (500 MHz, D2O): δ 33.64, 13.42, 1.15, -1.25, -4.39, -6.89, -8.20, -9.77, -14.68, -15.98

ESI-MS: m/z= 555.42 (100%), 553.33 (80%), 556.42 (78%), 554.42 (60%)

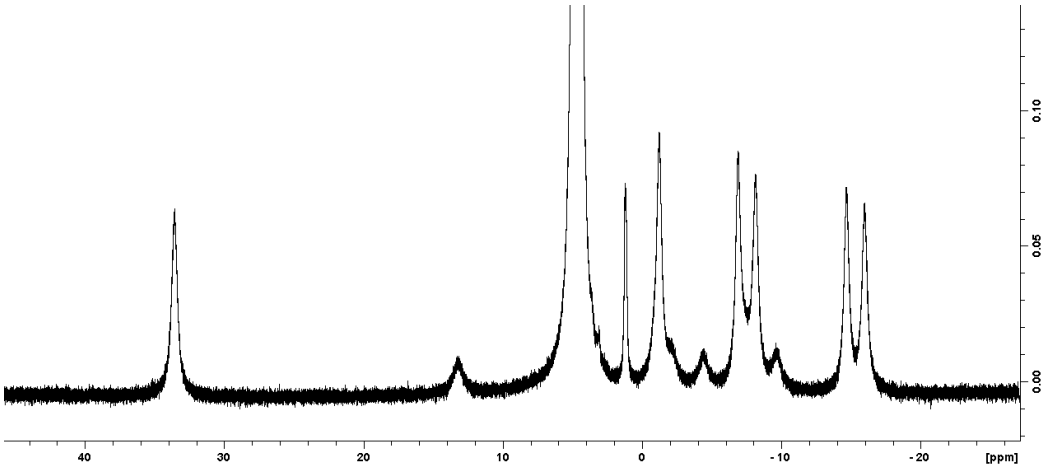


Figure 2.24: ^{1}H 1D NMR Spectra of [Eu(DOTA)]

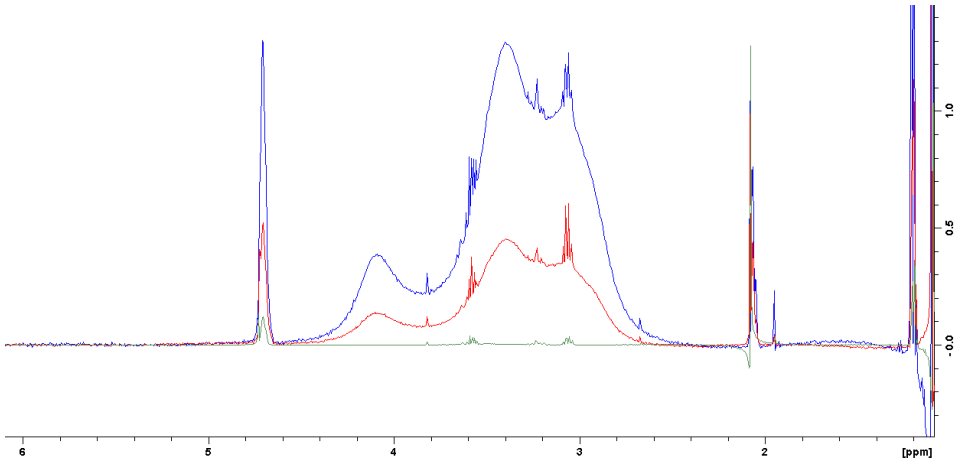
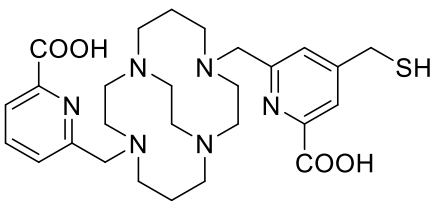


Figure 2.25: ^{1}H 1D NMR Spectra of [Eu(DOTA)] at [H^{+}]=0.2 M after incubation at 37°C for 5h (green), 20h (red), 4 days (blue).
The broad signals at 3.4 ppm and 4.1 ppm are the signals of diamagnetic free DOTA, which appears upon decomplexation of Europium.

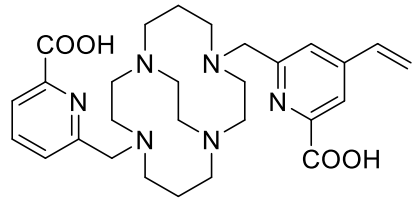
CHAPTER 3: Design of high-inertness paramagnetic tags

Abstract

In the first chapter of this thesis we discussed the requirements for making a “good” paramagnetic tag for NMR Spectroscopy. A paramount quality needed for the development of such ligands is the chemical inertness of the complex, and as we have seen in chapter 2, cross-bridged cyclam complexes are extremely inert. In this chapter, we will discuss the synthesis of two ligands, T1 and T2, both based on cb-tedpa, and with two different functionalization on the position 4 of one of the two pyridine pendant arms. We designed T1 with a free thiol and T2 with a double bond as the functional groups for bioconjugation. In chapter 2 we discussed two synthetic pathways for the obtention of akylated cross-bridged cyclams. Here, we will discuss how both pathways can be used depending on which conjugation method is chosen. Last, we will use our tag on model protein GB1 T53C in order to test it and evaluate its paramagnetic properties.



T1



T2

Figure 3.1: T1 and T2, ligands discussed in this chapter

3.1 Introduction

A desired property which is required in all the applications of organo-lanthanide complexes is inertness[76]. It is of particular interest in our case because in order to obtain an homogeneous sample and to prevent any undesired paramagnetic effect due to the presence of free Ln^{3+} in solution, we need our tags to be inert for the duration of the experiments (typically between 24 and 48 hours). A lot of efforts have been displayed in this direction, and the golden standard remains DOTA. Indeed, many DOTA-based Lanthanide-Binding-Tags (LBT) have been used in the last 30 years[56,77], and one of the reasons which pushed researchers in this direction in the first place was the inertness of DOTA-Lanthanide complexes[78]. As we have seen in chapter 2, though, cross-bridge cyclams seem to be even more inert than DOTA and therefore we thought it could be used as a new platform for the development of new lanthanide-binding tags.

Another important feature for the design of tags is the bioconjugation method that we will choose to attach our complexes to the surface of proteins[45]. On the protein side, cysteine residues' side chains are often preferred as the attaching point. Indeed, cysteines are not as present as other amino acids within proteins, and it is easier to create cysteine-selective chemistry as the properties of sulfhydryls are somehow peculiar, especially in terms of oxydo-reduction. As a result, the formation of a disulphide bond between a cysteine and a thiol on a tag has been the method of choice for many years. This method presents the advantage of being fast and absolutely cysteine selective. However, disulphides are not always stable and can be easily reduced, which can diminish the stability of the sample; this is becoming a growing concern especially with the rise of in-cell NMR[79]- an environment where disulphides get easily reduced[80]. Moreover, handling thiols is not always easy from a synthetic point of view.

An alternative to the formation of disulphide bonds for the conjugation of tags to biomacromolecules is the thiol-ene reaction[81]. The thiol-ene coupling or TEC is a reaction occurring through a radical mechanism in an Anti-Markovnikov fashion between a double bond and a free sulfhydryl[82]. However, it has been shown that in some cases the reaction could also happen following a Michael-addition-like mechanism[83]. In the case of bioconjugation, the free thiol is the one of the cysteine and the double bond lies on the molecule to be attached. This reaction offers the vantages of being extremely residue selective towards cysteines, and to yield a thioether bond instead of a disulphide; thioether cannot be reduced and are therefore more stable even in reducing environments. TEC, has been long used in polymer chemistry and has also been introduced in the field of biomolecular conjugation in the last years[84].

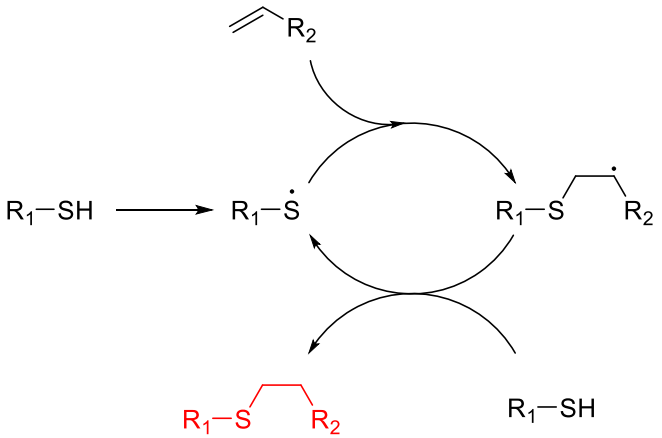


Figure 3.2: Radicalar mechanism of the Thiol-ene coupling (TEC)

The kinetically limiting step of this reaction is the radical generation[81]; the S· radical may form spontaneously but slowly. There are, however, two ways of catalysing this process: or by heating the reaction mixture, or by generating radicals in situ with the help of an initiator. TEC has been applied only once for the paramagnetic tagging of proteins by Su *et al*. This method is attractive in our field because 1) it creates a non-reducible thioether group and 2) the tether formed between the tag and the protein is of only 2 methylenes, thus giving little

mobility to the system. The aim of this chapter is to discuss the possibility of designing LBT combining the inertness and the rigidity of cb-cyclams and or the speed of reaction of the disulphide formation, or the stability and rigidity of the thioether formation through TEC.

3.2 Synthesis of T1

3.2.1 Retrosynthetic considerations

In a first time we decided to work on the synthesis of T1, which bears a free sulfhydryl group, designed to react with free cysteines in order to attach T1 through the formation of a disulphide bond. We decided to try and use the new synthetic strategy described in chapter 2. In order to do so, we needed to carefully protect the carboxylic acids and the thiol group so that they can resist through the harsh conditions of the reductive double ring expansion.

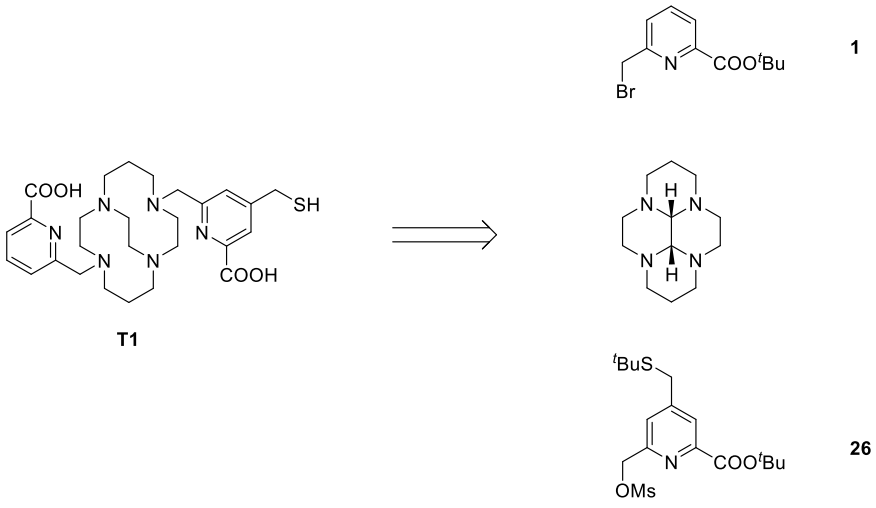


Figure 3.3: Retrosynthesis of T1, building blocks

As we have seen previously, ^{t}Bu esters offer good protection for the carboxylic acids. As for the free thiol, we decided to protect it as a ^{t}Bu thioether, a functional group stable in very basic conditions but deprotectable in strong acidic conditions, just like ^{t}Bu esters. We therefore designated **26** as a key intermediate in our synthesis.

3.2.2 Synthesis of 29

The synthesis of **23** was already described[85], but some adjustments were made. The introduction of the hydromethyl group in position 4 of the pyridine proceeds with a low yield (20%). The activation of the hydroxyl group with MsCl is nearly quantitative, and the introduction of the ^{t}BuSH group was performed using DBU instead of NaH, thus increasing the yield to 97%.

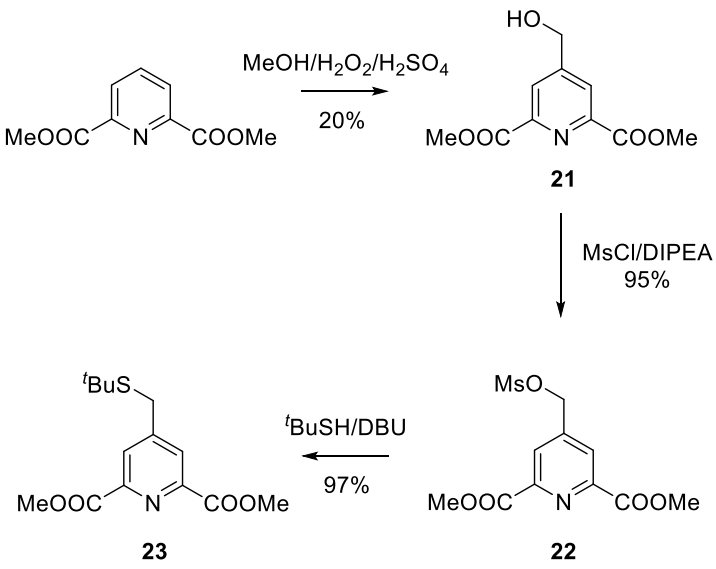


Figure 3.4: Synthesis of **23**

Selective mono-hydrolysis of **23** was achieved by using one equivalent of KOH. Next, we introduced one ^{t}Bu ester using Boc_2O and DMAP and obtained asymmetrically esterificated **24**. ^{t}Bu esters being unaffected by $NaBH_4$, we used it to reduce selectively the methylester and free hydroxyl was subsequently activated with a Mesyl group.

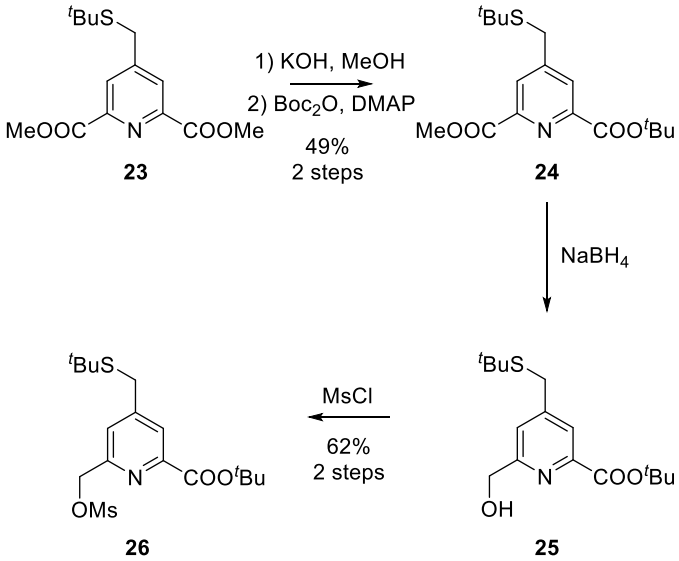


Figure 3.5: Synthesis of **26**

The total yield over those seven steps is only of 5.4%; however, it rises to 28% if we do not take into account the first, low-yielding step. The biggest challenge in this synthesis is the selective transesterification of one only of the two methyl esters. This problem was overcome thanks to the reactivity of dimethoxy pyridine-2,6-dicarboxylate derivatives, where the first methyl ester is much more reactive than the second one; that way it is possible to have a good control over the regioselectivity of the reaction. With **26** in hand it was possible to start alkylating cis-glyoxal cyclam.

3.2.3 Alkylation and deprotection

While working on asymmetrically substituted cb-cyclam derivatives, it is important having a very good control over the alkylation steps in order to sequentially alkylate with two different pendant arms, and possibly not having to perform purification step. As it has been shown in Chapter 2, the direct alkylation of cis-glyoxal cyclam offers the advantage of being extremely selective towards mono-alkylation.

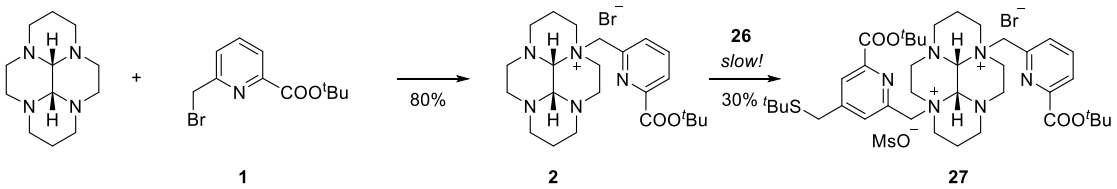


Figure 3.6: Synthesis of **27**

A disadvantage of this synthetic route, however, is that the second alkylation becomes very slow. Indeed, after a month of reaction we reached only 30% of conversion (reaction followed via HPLC/MS and NMR). Next, reductive double ring opening was performed on **27**. As expected, the ^{t}Bu groups were stable in those conditions and the reaction proceeded in satisfying yield. It is worth noting that at this point, it is possible to perform a purification via flash chromatography.

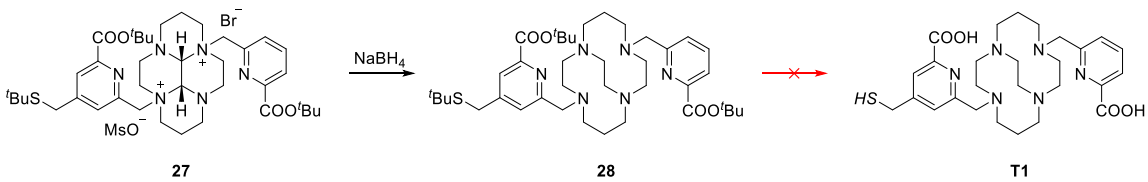


Figure 3.7: Synthesis of **T1**

In order to deprotect both ^{t}Bu esters and the ^{t}Bu thioether of **28**, very acidic conditions were required. In a first time, TFA was used; this led to the deprotection of ^{t}Bu esters, but not the ^{t}Bu thioether. Next, HCl 37% was used. After 4 hours, NMR showed a mixture of species. Further investigation revealed that while ^{t}Bu esters had been deprotected, ^{t}Bu thioether was not fully deprotected. More, the fraction that was deprotected seemed to be degraded, and signs of an aldehyde were identified via NMR. Similar experiment was conducted on intermediate **23** and similar results were obtained.

It was therefore concluded that this deprotection strategy was not applicable for the synthesis of T1 and that another protection strategy needed to be chosen. Other protecting groups could be envisaged for the protection of the free thiol, which would be easier to deprotect and in milder conditions. However, it was decided to focus on the synthesis of T2. Indeed, in T2, it was expected that the double bond, functional group designed to attach the tags to cysteines, would not need a protection/deprotection cycle and that it would be more stable in water.

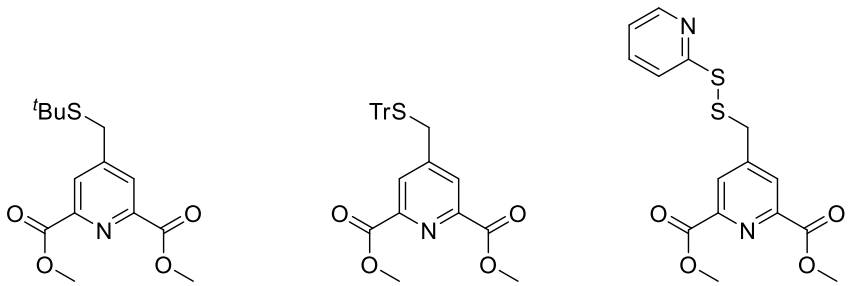
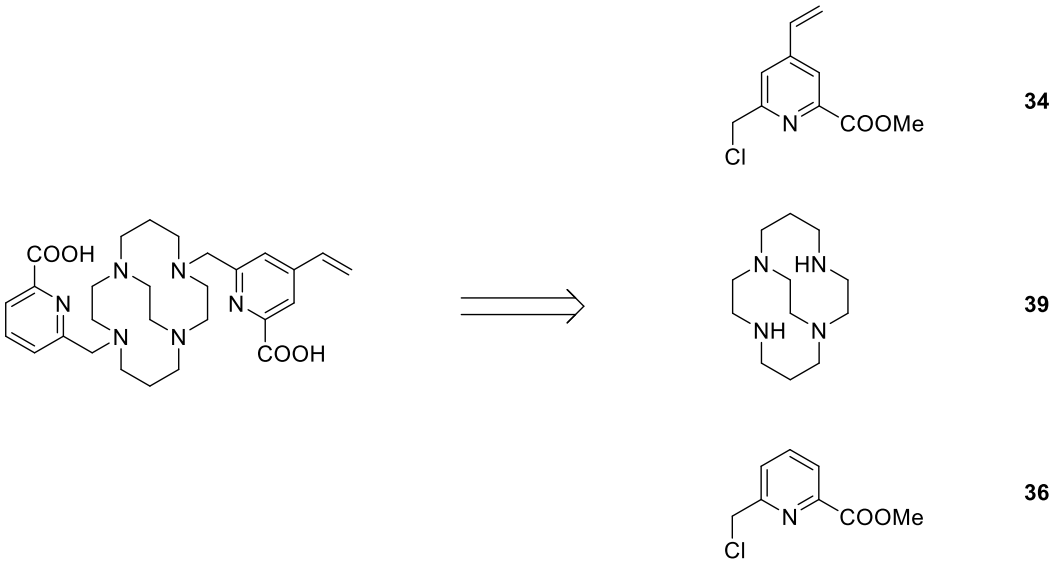


Figure 3.8: Possible protecting groups for the methanethiol group in position 4 of dimethoxy-2,6-pyridinedicarboxylate

3.3 Synthesis of T2

3.3.1 Retrosynthetic considerations

In chapter 2, were presented and discussed two synthetic pathways towards alkylated cross-bridge cyclams. Here, to decide which of those two pathways to follow, were taken into account the reactivity and protection of the reactive group for the bioconjugation. In the case of T2, there is a double bond which is not stable in the conditions of the reductive double ring opening (large excess of $NaBH_4$). An option could be to try and protect this double bond to follow the “novel” pathway presented in chapter 2. However, in order not to have to protect the double bond and the methyl ester, it was decided to use the “classical” pathway instead. Were then identified three key intermediates: **34**, **36** and **39**, with **34** and **36** to be sequentially attached on cb-cyclam **39** through secondary-amine alkylation.



34

39

36

Figure 3.9: Retrosynthesis of T2

3.3.2 Synthesis of 34

3.3.2.1 First strategy

Our first strategy was to first synthesize **30** starting from chelidamic acid, which we obtained in three steps and 60% yield. Those reactions were performed up to a scale of 10g. It has been shown that the mono-reduction of one only of the two esters of dipicolinic ester-like molecules was easily achievable with a slight excess of $NaBH_4$. However, since the vinylic double bond of **30** is reduced by $NaBH_4$, we explored other reducing conditions that would be chemoselective for esters, hoping we would conserve the same selectivity towards the mono-reduction. The results are summed up in Table 3.1.

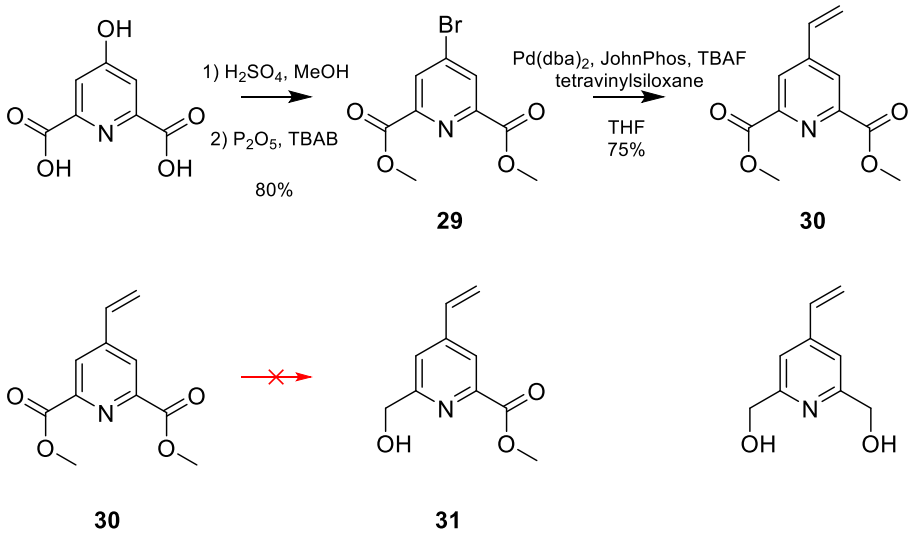


Figure 3.10: First strategy for the synthesis of **34**

It was interesting to observe that the adding of Cerium Chloride before $NaBH_4$ induced a total change in the reactivity and totally prohibited the reduction of the alkene. It was hypothesized that the lanthanide is chelated by the ester-nitrogen-ester motif, thus increasing the electrophilicity of the Cα, making them more reactive towards reduction. However, adding $CeCl_3$ to the reaction mixture lead to a loss of selectivity towards the mono-reduction of the diester. NMR analysis of the mixtures generated by the reaction of **30** with $NaBH_4$ in presence of $CeCl_3$ revealed a mixture of **30**, **31** and the di-reduced product. Reactions conducted with DIBAL-H also led to the reduction of esters instead of the alkene, however the reaction stopped at the aldehyde formation and in the ended a mixture of reagent, aldehyde and di-aldehyde products were recovered. It was concluded that this first synthetic strategy could not lead to the isolation of **34**.

Reducing Reagent	Chemoselectivity
NaBH4	No
DIBAL	Yes
CeCl3/NaBH4	Yes

Table 3.1: Reduction of **30**

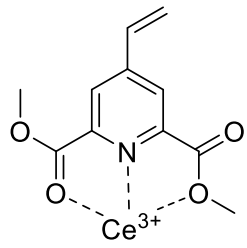


Figure 3.11: Complex of **30** with Cerium III

3.3.2.2 Second strategy

Since it was not possible to perform the ester reduction after the introduction of the vinylic double bond, we decide to explore another route where we would first try and perform the mono-reduction of one of the two esters only and then introduce the double bond in position 4.

The strategy we followed is shown in fig. 3.12. Because of the absence of competitive group, the reduction of **29** with $NaBH_4$ yielded the desired **32** nearly quantitatively. The conditions of the Suzuki-like coupling which yielded **30** unfortunately did not allow us to isolate **33** and therefore had to be adapted, with moderate success (68%). Last, the free hydroxyl group was activated with Mesyl Chloride. In conclusion, we isolated **34** in 5 steps and 50% yield.

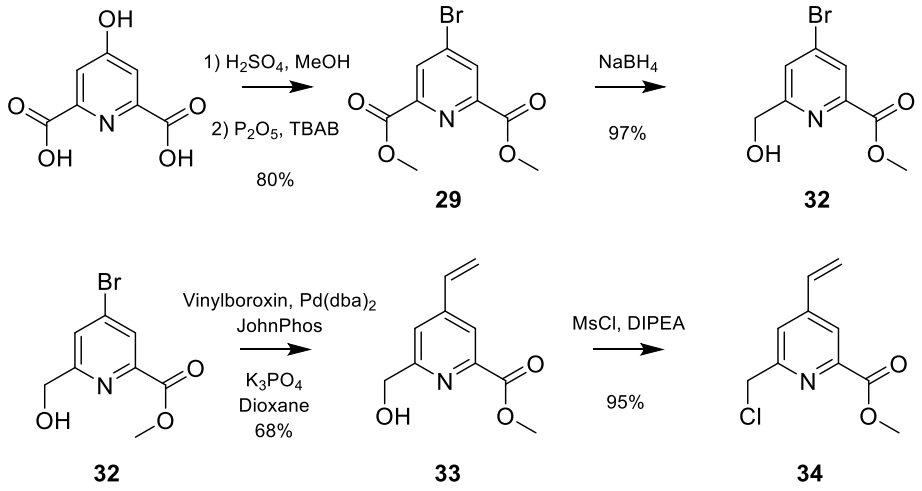
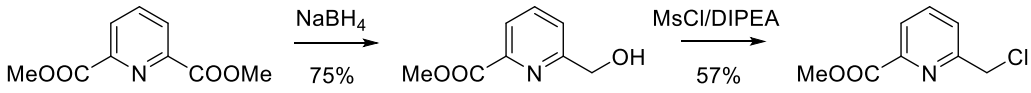


Figure 3.12: Second strategy for the synthesis of **34**

3.3.3 Synthesis of 36 and 39

3.3.3.1 Synthesis of **36**

36 was easily isolated in 2 steps starting from commercial dimethoxy-2,6-pyridinedicarboxylate. Reduction of one of the two methylesters was easily



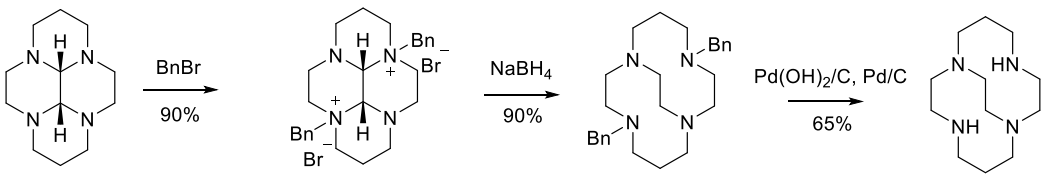
35 **36**

achieved and the hydroxyl group was activated as a Chloride leaving group.

Figure 3.13: Synthesis of **36**

3.3.3.2 Synthesis of **39**

39 was synthesised following the "classical" route described by Wong *et al*[63]. Starting from commercial cis-glyoxal cyclam, transaminal alkylation and reductive double ring opening were performed following the described procedures. The described protocol for the deprotection of the benzylated amines did not work in our hands and the procedure was adapted, increasing the amount of catalyst used. The overall yield over three steps is of 53%.



37 **38** **39**

Figure 3.14: Synthesis of **39**

3.3.4 Alkylation of 39 and deprotection

3.3.4.1 First Alkylation

In a first time we explored the selectivity of the mono-alkylation of cb-cyclam above di-alkylation. We observed that by adding a stoichiometric amount of **36** to **39**, the formation of an important amount of dialkylated compound was observed by NMR (around 20%). By decreasing the amount of **36** to 0.9 equivalent, we managed to decrease the amount of dialkylated **41** formed and increased the yield to 90%.

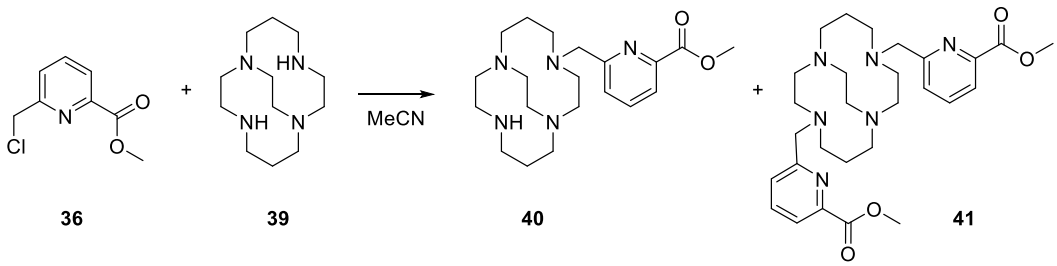


Figure 3.15: cross-bridge cyclam monoalkylation

3.3.4.2 Second Alkylation and Deprotection

With **40** and **34** in hand, the second alkylation was performed in good yields and did not require particular optimisation. At this point it is possible to purify the compound via Silica Flash Chromatography. Deprotection of the methyl esters was achieved with a slight excess of LiOH in a water/THF mixture (fig. 3.16). As expected, the vinylic double bond was stable in those conditions. The deprotected compound was purified via semi-preparative HPLC.

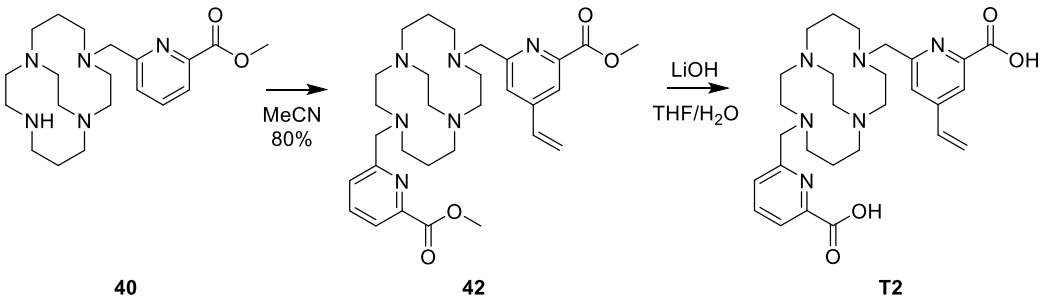


Figure 3.16: Synthesis of **T2**

3.3.5 Lanthanide Chelation

3.3.5.1 Chelation Reaction

With the ligand T2 in hand, next was to be performed the chelation of various lanthanides. Chelation reactions were performed in a micro-wave (MW) apparatus at 150 W and 150°C. Reactions were generally completed after 3 cycles of 2 hours. Reactions were followed and purified via HPLC. It was found that the pH of the reaction was a very important factor in the kinetic of the reaction. Indeed, if the pH gets acidic, we found that the reaction would slow down drastically. On the other if the pH gets too high, Ln^{3+} ions may precipitate as $Ln(OH)_3$, insoluble in water. Interestingly, it was noted that the pH increases during the reaction, probably because the amines deprotonate upon chelation thus increasing the concentration of $[H^+]$ in solution.

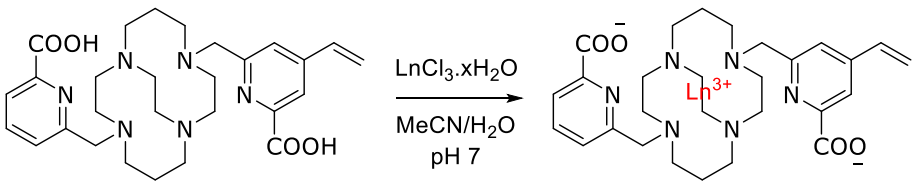


Figure 3.17: Chelation of lanthanides

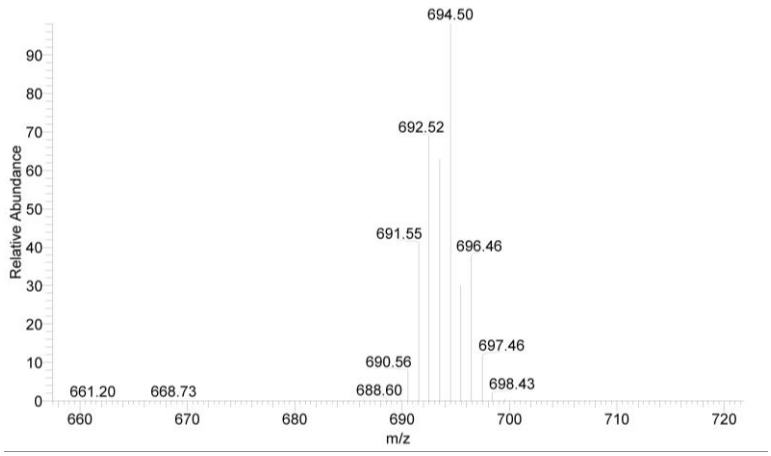


Figure 3.18: Isotope distribution of [T2(Yb)] revealed by Mass Spectrometry

3.3.5.2 Stereochemistry

An aspect which has to be taken into account at this point is the stereochemistry of the complex. Indeed, it was hypothesized that T2 in complex with a lanthanide could present four different configurations. Cross-bridge cylam is in equilibrium between two conformations, which are enantiomers and therefore no distinction is observed by NMR, only one set of peaks is seen in the ^{1}H 1D spectra of **39**. When T2 complexes a metal, its arms become stereogenic because they must take an orientation around the metal and there are two possible orientations for this, one where the arms turn “leftwise” and one where they turn “rightwise” (c.f. fig 3.19). Thus, [Ln(T2)] could adopt a total of four configurations. Those four possible stereoisomers, which would form two diastereoisomeric pairs of enantiomers, were modelled, and they are displayed in fig. 3.20.

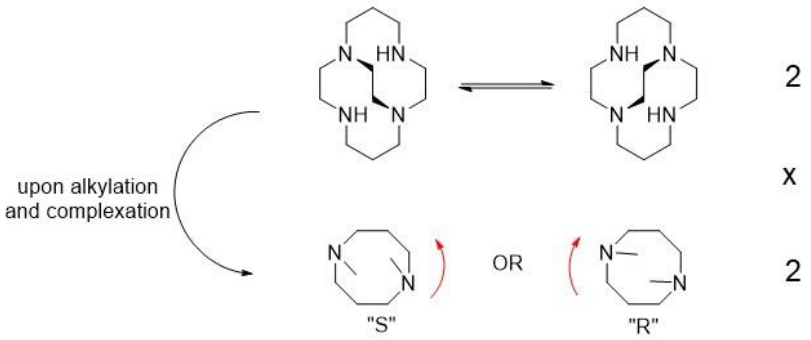


Figure 3.19: Stereogenicity in alkylated cb-cyclams

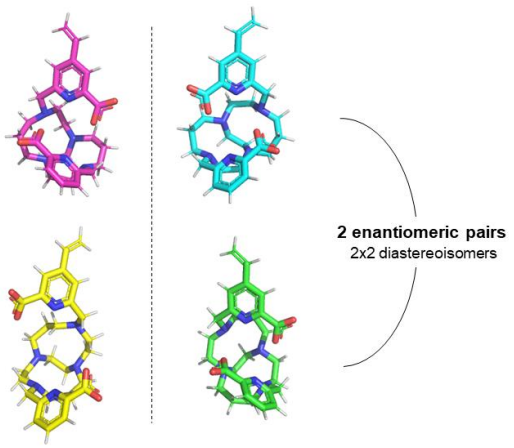


Figure 3.20: The 4 possible configurations of T2 upon complexation.
Two couples of enantiomers, respectively diastereoisomers

This hypothesis was confirmed by HPLC-MS experiments. Indeed, following the chelation reaction, two peaks relatively close by HPLC with the exact same MS spectra were observed (c.f. fig. 3.20). This indicates that at least

two species, respectively diastereoisomers, exists. Those peaks may actually contain mixtures of enantiomers, if indeed the complex could take the four stereoisomeric forms described in figure 3.20. From now on we will call “A” and “B” those two diastereoisomeric species. Interestingly we found the A/B ratio to be lanthanide-dependent with a correlation with the radius of the ion. The results obtained are summed up in table 3.2. Species A and B were readily separated via semi-preparative HPLC.

Lanthanide	Crystal Radius (Å)	%A	%B
La	**1.185**	**0**	**100**
Dy	**1.062**	**13.5**	**86.5**
Yb	**1.008**	**21**	**79**

Table 3.2: Ratios of diastereoisomers

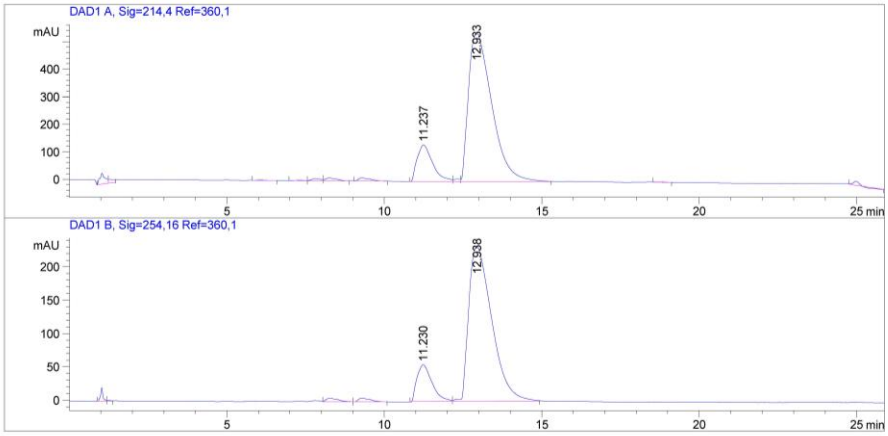


Figure 3.21: HPLC separation of diastereoisomers of [Dy(T2)]

3.3.5.3 Structural investigations

On figure 3.22, we can see two 3D-structures of [Ln(T2)] that are diastereoisomers. Lanthanides have been removed from the structures for the sake of clarity. Our goal is to assign those structures to species A and B. To do so, we performed ^{1}H-COSY and NOESY experiments on [T2(La)]. First, spin-systems were isolated and assigned to the different structural components of [T2(La)]. Using ^{1}H-COSY, we could not assign each and every proton, however, we managed to separate and assign spin systems corresponding to: the pendant arms, the benzylic protons (corresponding to the methylene located between the pyridines and the cyclam's nitrogens), cyclam's ethylene and propylene subunits, and the cross-bridge. Next, NOESY experiments were performed and analysed.

From the structures in figure 3.22, we can see that the benzylic protons should have different magnetic environments whether in one configuration or the other. Looking at NOE contacts within [La(T2)], which corresponds to configuration B (Table 2), it was found that while one benzylic proton had contacts with protons from both the spin systems corresponding to the ethylene and the propylene subunits, the other one had contact only with the propylenic spin system. If we look at the first 3D-structure in figure 3.22 (orange), we can expect that for this structure one proton would have NOE contacts with protons from both the spin systems corresponding to the ethylene and the propylene subunits, but that the other benzylic proton would have contact only with protons from the ethylenic subunit. On the other hand, in the structure at the bottom, we can expect that one proton has contacts with both spin-systems and that the other one would have contacts only with protons from the propylene, which is the case in the spectra of [La(T2)]. We could therefore assign the first structure and its enantiomer (not represented in fig. 3.22) to species A, and the structure at the bottom and its enantiomer (not represented in fig. 3.22) to species B.

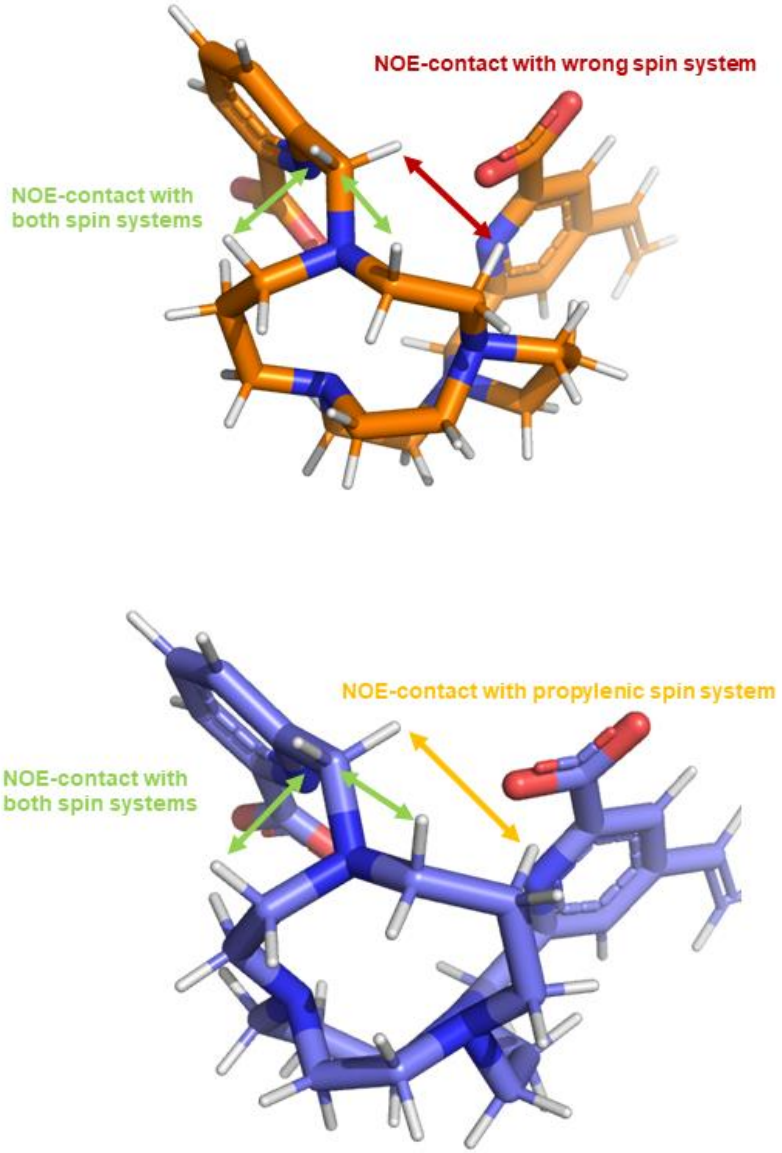


Figure 3.22: Conformational analysis of [Ln(T2)] via NOESY

3.4 Conjugation of T2 to GB1 T53C and NMR

3.4.1 Prerequisites

In **3.3**, the synthesis of ligand T2 was discussed and were characterised its complexes with some Lanthanides. Furthermore, it was shown that upon complexation two diastereoisomers were found and that they were separable. Those two populations, though, may not be enantiomerically pure.

In order to test T2 and use it as a paramagnetic tag, it was decided to use mutant T53C of GB1 as a model. Indeed, GB1 is a small globular protein, stable under various conditions, and the mutant T53C presents only one cysteine, making it easier to monitor the reaction and study the effects of Paramagnetism.

Complex T2 is meant to be attached to free cysteines via the thiol-ene reaction. This reaction proceeds through a radical mechanism which may be catalysed via heating[84], but also following a Michael-addition-like mechanism. In this section we will discuss the conjugation reaction, its kinetic, as well as the behaviour of the tag in terms of paramagnetic properties but also in terms of stereochemistry.

3.4.2 Conjugation

In a first attempt to conjugate [Ln(T2)] (species B) to GB1, we followed the protocol of Su[86]. However, NMR showed that after one night nothing had happened. Indeed, it can be hypothesized that without in situ radical-generation, the reaction proceeds in a Michael-addition-like fashion. In the reaction described by Su[86], they used vinyl-dipicolinic acid (VDPA) as the substrate,

which bears two carboxylic acids on its vinyl pyridine instead of one in the case of T2. Those carboxylic acids probably increase the electrophilicity of the double bond and therefore increase the kinetic rate of the Michael addition, therefore making T2 less reactive than VDPA. To confirm this hypothesis, we performed the conjugation of GB1 T53C with VDPA following the same protocol and we did obtain a conjugation of around 60%, thus showing that VDPA is more reactive than T2.

In order to try and speed up the reaction, and because both our tag and GB1 are extremely thermostable, we decided to heat the reaction and leave it at 70°C for one night. This led to the complete tagging of the protein, with all three lanthanides used (La, Dy, Yb), as shown by NMR (fig. 3.24).

Upon conjugation, pronounced changes on the NMR signals of our protein were observed via 2D ^{1}H-^{15}N HSQC experiments; those changes were attributed to the paramagnetism of T2, and consisted mostly in PCS. However, it was observed that those peaks were belonging to two different sets, and that for each diamagnetic peak, two peaks were observed in the paramagnetic samples. Also, the intensities of the peaks from each set were identical. This indicated that the sample was composed of a 50/50 mixture of two species, both paramagnetically tagged, with the same number of residues. This confirmed us that the main peak isolated via HPLC after complexation is a racemic mixture, which resolve upon conjugation to the protein, thus validating our hypothesis that T2 can take 4 different conformations upon complexation.

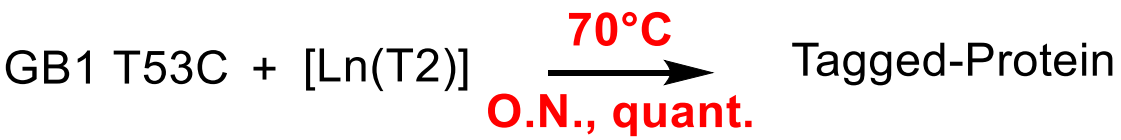


Figure 3.23: Tagging of GB1 T53C with T2

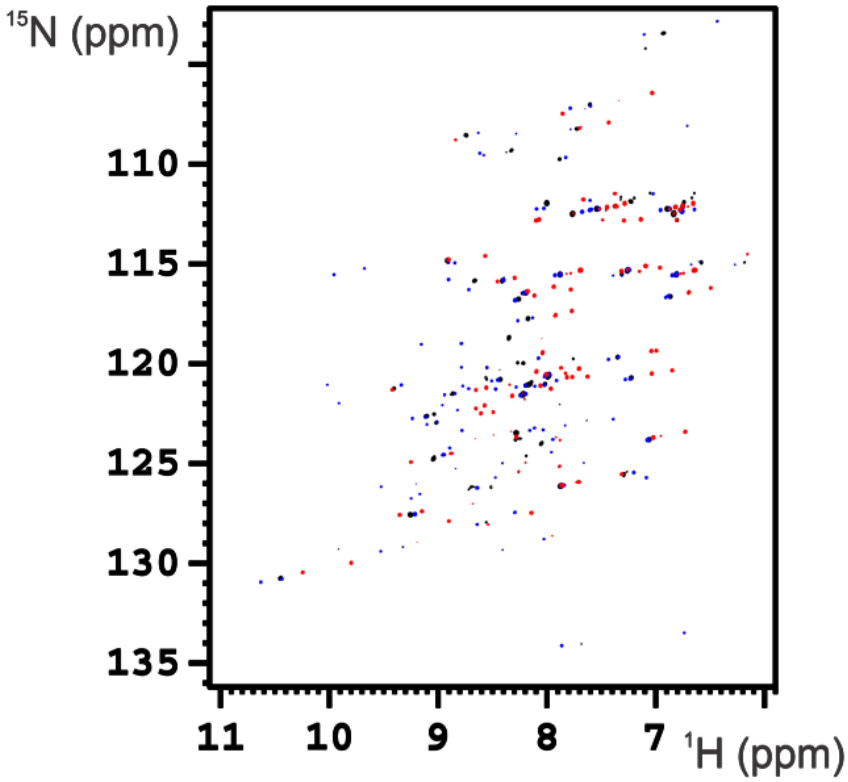


Figure 3.24: 2D ^{1}H-^{15}N HSQC of GB1 T53C tagged with [T2(La)], [T2(Yb)], [T2(Dy)]

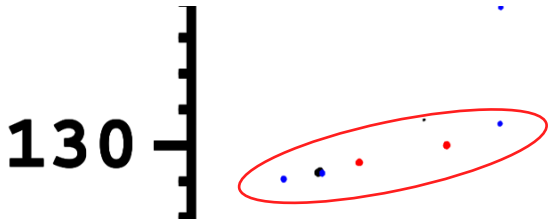


Figure 3.25: Each residue splits into 2 peaks upon conjugation

3.4.3 PCS measurements and Tensor evaluation

The main paramagnetic effects considered in paramagnetic NMR, and also the most straightforward to measure, are PCS. Looking as those shifts it is immediate observing that the protein is fully paramagnetically tagged because not a single H-N cross-peak was not shifted in the protein, and therefore no residual signal corresponding to the apo protein was found. As seen in chapter 1, PCS follow the equation:

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

where *r*, θ and φ are the spherical coordinates of the nucleus in the frame where the anisotropy tensor is diagonal and with origin onto the metal position. $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic anisotropies of the tensor, defined as:

$$\Delta\chi_{ax} = \chi_{zz} - \frac{\chi_{xx} + \chi_{yy}}{2}$$

$$\Delta\chi_{rh} = \chi_{xx} - \chi_{yy}$$

To measure precisely those PCS, differences in chemical shift were measured in both H and N dimensions between the spectra of the paramagnetic and the diamagnetic species. In this case, this task was a bit complicated by the presence of two different sets of peaks. Indeed, it made it harder to assign resonances in the paramagnetic samples to the ones in the diamagnetic sample. However, two sets of peaks were identified, and PCS were calculated and treated separately.

The results were then analysed with the software FANTEN. The program FANTEN can be used to obtain the best fit $\boldsymbol{\Delta\chi}$ tensor (consisting in 5 parameters: $\Delta\chi_{ax}$, $\Delta\chi_{rh}$ and the three Euler angles defining the frame where the tensor is diagonal) and the coordinates of the metal ion from the PCS values and the coordinates of the protein nuclei. Both sets of peaks, stemming from both enantiomers, were treated separately. First, we noticed that for both sets of data the fitting was good (low Q-value). This confirmed that both sets had been separated properly, referred to hereafter as T2.1 and T2.2, respectively. The value of anisotropy of these tensors ($\Delta\chi_{ax}$ and $\Delta\chi_{rh}$) are very high. Interestingly, the two enantiomers gave tensors with very different intensities (Table 3.3).

	Metal	Q	$\Delta\chi_{ax}$ (10^{-32} m^3)	$\Delta\chi_{rh}$ (10^{-32} m^3)
T2.1	Yb	0.088	3.31	-1.82
	Dy	0.138	-17.33	9.17
T2.2	Yb	0.063	-16.07	7.77
	Dy	0.041	57.94	-28.22

Table 3.3: Tensor values of T2 tagged to GB1 T53C

Since T2.1 and T2.2 are two enantiomers, the lanthanides chelated within will have the exact same coordination environment. However, it is clear that they yield very different tensors once tagged to the protein, and it can be assumed that interactions between the tag and the protein may give rise to such differences. It was hypothesized that in the case of T2.2, an interaction between the protein and T2 stabilises the tag, decreasing its motion in solution and thus

yielding a very large tensor on the protein. To get a better understanding of the results given by FANTEN, we superposed the metal positions fitted to the PCS values with a structure of GB1. The results are shown in fig 3.26.

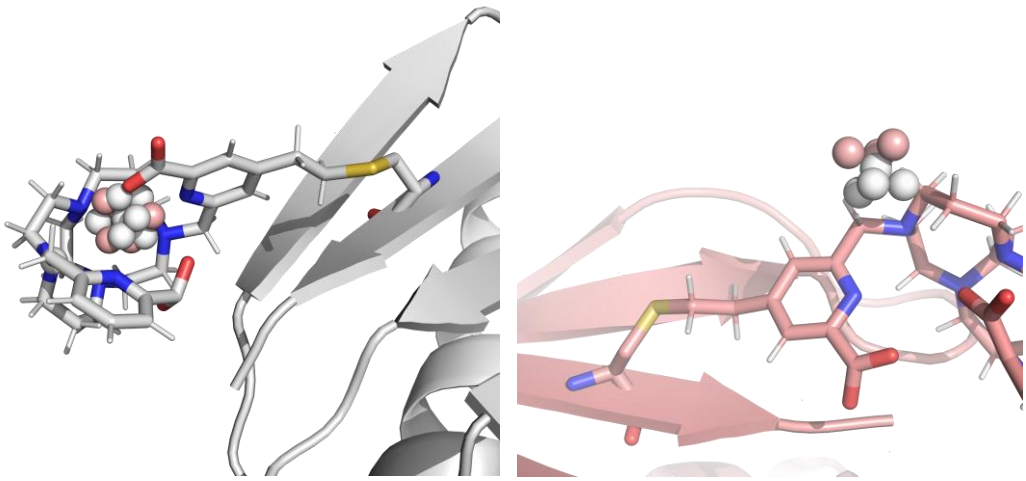


Figure 3.26: Metal position for T2.1 (left) and T2.2 (right) as given by FANTEN. The metal positions are superimposed with a structure of GB1 tagged with the two enantiomers of T2, generated with MAESTRO.

It can be noted that the metal seems to be very close from the C-terminus of the protein. In this area, there are 3 carboxylates which may form an electrostatic interaction between the tag and the protein. It has been shown that such interactions could enhance the rigidity of the system by stabilising the tag and therefore increase the tensor transferred from the tag to the protein[87]. Lanthanides are 9-coordinated in water[88]; since lanthanides in T2 are only 8-coordinated, it could therefore allow for one extra non-covalent interaction with the protein's surface what would explain the high tensor-values measured and explain the intensity of the paramagnetism transferred from the tag to the protein.

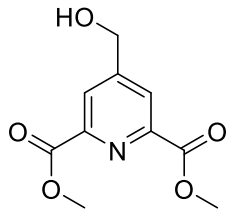
3.5 Conclusion

In this chapter we have designed and planned the synthesis of two potential LBT, T1 and T2, based on cb-tedpa. Unable to isolate T1 because of our unability to deprotect the ^{t}Bu-sulphides (designed to be the point of conjugation of our tag to cysteines), we were on the other hand able to isolate T2 in good yields. Lanthanide-complexes of T2 were synthesized and characterised, and particular attention was devoted to the study of the complexes' stereochemistry, showing that upon complexation [Ln(T2)] complexes adopt 4 different configurations. T2 was subsequently tagged to GB1 T53C via the thiol-ene reaction, which allowed us to measure large PCS via 2D ^{1}H-^{15}N HSQC. Last, those PCS were analysed, and high-quality tensors were determined, displaying very large anisotropy values. Those results show that cb-tedpa is a promising platform for the development of paramagnetic tags. However, future efforts should be concentrated on the obtention of enantiopures cb-tedpa-based ligands.

3.6 Experimental Procedures

3.6.1 Synthetic procedures

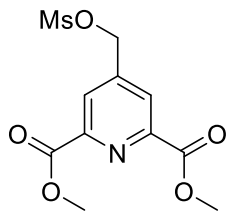
dimethyl 4-(hydroxymethyl)pyridine-2,6-dicarboxylate **21**



Dimethoxy-2,6-pyridinedicarboxylate (5g) was suspended in MeOH (26.8 mL), and 26.8 mL of H_2SO_4 (30% v/v) were added. At 0°C, 21.4 mL of H_2O_2 (30% v/v) and 26.8 mL of a saturated solution of Fe_2SO_4 were added simultaneously and slowly, so that the temperature of the reaction never rised above 15°C. Once the adding was complete, the reaction was let stirred at room temperature for 20 mns, after which is was neutralized with K_2CO_3 and filtrated. The aqueous mixture mas then extracted with 4x150 mL Ethyl acetate. The organic layers were combined, dried, and purified with Flash Chromatography with $CHCl_3$:2/1:EtOAc to yield 1.15 g of pure product.

^{1}H NMR (500 MHz, CDCl3): δ 8.32 (s, 2H), 4.92 (s, 2H), 4.04 (s, 6H)

dimethyl 4-(((methylsulfonyl)oxy)methyl)pyridine-2,6-dicarboxylate **22**



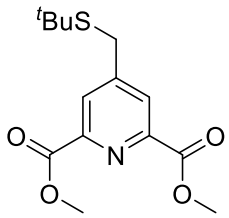
21 (1.1g, 4.88 mmol) was dissolved in dry DCM (95mL), and DIPEA (2.55 mL, 13.9 mmol) was added. MsCl (567µL, 7.33 mmol) was then added dropwise over an ice-bath. Reaction mixture was let stirred for 35 mns at room temperature and

was then washed with 2x80 mL H_2O. Organic layer was dried over Mg_2SO_4 and solvents were removed in vacuo, yielding an orange oil. The crude mixture was filtered over a pad of silicagel with 3/1:$CHCl_3$/EtOAc as the mobile phase to yield 1.3 g. of pure product (89%).

^{1}H NMR (500 MHz, CDCl3): δ 8.34 (s, 2H), 5.39 (s, 2H), 4.06 (s, 6H), 3.16 (s, 3H)

^{13}C NMR (125 MHz, CDCl3): δ 164.66, 149.03, 145.98, 126.03, 67.11, 53.44, 38.37

dimethyl 4-((tert-butylthio)methyl)pyridine-2,6-dicarboxylate **23**

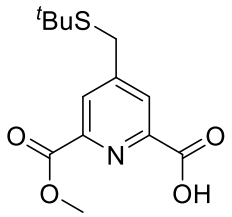


22 (3.2 g., 10.5 mmol) was dissolved in dry DCM (55 mL), under N_2. *Tert*-butylthiol (3.37 mL, 28.4 mmol) and then DBU (1.86 mL, 21.1 mmol) were added slowly. After 1h under N_2 and at room temperature, the reaction mixture was diluted with 100 mL and washed with 4x100 mL H_2O. Organic layers were combined, dried over Mg_2SO_4, and solvents were evaporated to yield 3.12 g. of pure product (97%)

^{1}H NMR (500 MHz, CDCl3): δ 8.33 (s, 2H), 4.04 (s, 6H), 3.86 (s, 2H), 1.36 (s, 9H)

^{13}C NMR (125 MHz, CDCl3): δ 165.08, 151.91, 148.41, 128.36, 53.21, 43.87, 32.31, 30.90

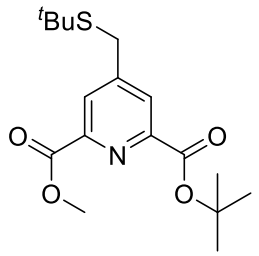
4-((tert-butylthio)methyl)-6-(methoxycarbonyl)picolinic acid



23 (286 mg, 0.96 mmol) was dissolved in MeOH (6 mL) over an ice-bath. KOH (54 mg, 0.96 mmol) were added. The mixture was stirred for 2h at 0°C, after which MeOH was evaporated. The crude mixture was then suspended in water and filtrated over a fritt #4. Filtrate was evaporated to yield 230 mg of product (75%)

^{1}H NMR (500 MHz, D2O): δ 8.15 (s, 1H), 8.03 (s, 1H), 3.94 (s, 3H), 3.91 (s, 2H), 1.27 (s, 9H)

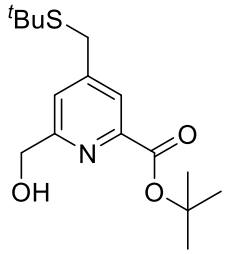
2-(tert-buty) 6-methyl 4-((tert-butylthio)methyl)pyridine-2,6-carboxylate **24**



4-((tert-butylthio)methyl)-6-(methoxycarbonyl)picolinic acid (500 mg, 1.76 mmol), was dissolved in dry DCM and under N_2. DMAP (42.5 mg, 0.35 mmol), and then Boc-anhydride (942 µL, 4.1 mmol) were added to the reaction mixture, which was let stirred overnight at room temperature. DCM was evaporated and crude yellow oil was dissolved in EtOAc (40 mL), washed with 3x10 mL sat. NH_4Cl. Aqeous layers were combined and re-extracted with 40 mL DCM. Organic layers were combined, dried over Mg_2SO_4, and solvent was removed. The obtained mixture was further purified by Flash Chromatography ($CHCl_3$/MeOH : 9/1) to yield 400 mg of product (67%).

^{1}H NMR (500 MHz, CDCl3): δ 8.29 (s, 1H), 8.21 (s, 1H), 4.03 (s, 3H), 3.68 (s, 2H), 1.67 (s, 9H), 1.38 (s, 9H)

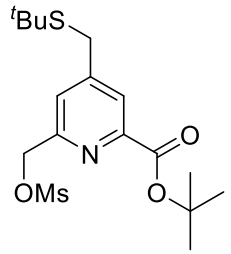
tert-butyl 4-((tert-butylthio)methyl)-6-(hydroxymethyl)picolinate **25**



24 (470 mg, 0.98 mmol), was dissolved in MeOH (27 mL) and DCM (9 mL) under N_2 and over an ice-bath. $NaBH_4$ (148 mg, 3.92 mmol) was then added slowly. After 3 hours at room temperature, reaction mixture was diluted with 40 mL EtOAc, and washed with 30 mL brine. Aqueous layer was then re-extracted with 3x40 mL EtOAc. Organic layers were combined, dried over Mg_2SO_4, and solvent was removed to yield 430 mg of product.

^{1}H NMR (500 MHz, CDCl3): δ 7.89 (s, 1H), 7.47 (s, 1H), 4.82 (s, 2H), 3.76 (s, 2H), 1.62 (s, 9H), 1.34 (s, 9H)

tert-butyl 4-((tert-butylthio)methyl)-6-(((methylsulfonyl)oxy)methyl)picolinate **26**

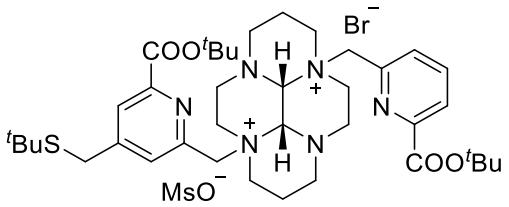


25 (430 mg) was dissolved in dry DCM (13 mL) under N_2 and over an ice-bath. DIPEA (513 µL, 2.94 mmol) and Methansulfonyl Chloride (152 µL, 1.96 mmol) were added. After 1h30, reaction was over and the crude mixture was purified via Flash Chromatography (8/2 : CHCl3/AcOEt) to yield 223 mg of product (62% over 2 steps)

^{1}H NMR (500 MHz, CDCl3): δ 7.95 (s, 1H), 7.57 (s, 1H), 5.35 (s, 2H), 3.76 (s, 2H), 3.16 (s, 3H), 1.59 (s, 9H), 1.31 (s, 9H)

^{13}C NMR (125 MHz, CDCl3): δ 163.49, 154.30, 151.02, 149.41, 124.98, 124.89, 82.51, 71.23, 43.67, 38.23, 32.28, 30.90, 28.05

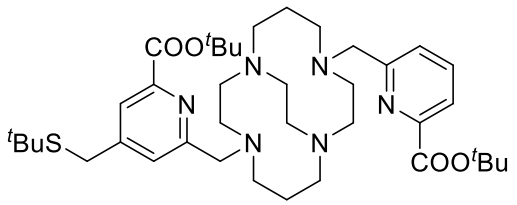
(3a1R,5a1R)-5a-((6-(tert-butoxycarbonyl)-4-((tert-butylthio)methyl)pyridin-2-yl)methyl)-10a-((6-(tert-butoxycarbonyl)pyridin-2-yl)methyl)tetradecahydro-3a,5a,8a,10a-tetraazapyrene-5a,10a-diium bromide methanesulfonate **27**



2 (160 mg, 0.32 mmol) and **26** (120 mg, 0.32 mmol) were dissolved in 1.5 mL MeCN. Reaction was monitored via NMR and MS. After one month the conversion was of 30%. Solvent was removed and the crude mixture of **2** and **27** was used without further purification.

ESI-MS: m/z= 354.35 Th

tert-butyl6-((11-((6-(tert-butoxycarbonyl)pyridin-2-yl)methyl)-1,4,8,11-tetraaza bicyclo[6.6.2]hexadecan-4-yl)methyl)-4-((tert-butylthio)methyl)picolinate **28**



Mixture of **2** and **27** was dissolved in 9 mL EtOH 95%. $NaBH_4$ (48.5 mg, 1.28 mmol) was then added in portions and the mixture was let stirred overnight at room temperature. The reaction mixture was then poured onto 20 mL of water, and saturated with KOH. The aqueous solution was then extracted with 3x20 mL

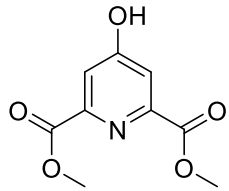
DCM. Organic layers were combined and dried. The crude mixture was then further purified via semi-preparative HPLC (TFA 0.1% in H2O/MeCN).

^{1}H NMR (500 MHz, CDCl3): δ 11.11 (bs, 2H), 7.74 (d, J=7.48 Hz, 1H), 7.61 (t, J=7.51 Hz, 1H), 7.58 (s, 1H), 7.11 (bd, J=7.48 Hz, 1H), 7.08 (s, 1H), 5.25 (bt, J=13.24 Hz, 2H), 4.52 (bs, 1H), 4.41 (bs, 3H), 3.92 (bs, 2H), 3.73 (bs, 2H), 3.64-3.54 (m, 4H), 3.52-3.39 (m, 6H), 2.92 (bd, J=11.64 Hz, 2H), 2.70 (bs, 4H), 2.55 (bs, 2H), 1.67 (s, 9H), 1.65 (s, 9H), 1.38 (s, 9H).

^{13}C NMR (125 MHz, CDCl3): δ 163.18, 160.97, 160.77, 160.57, 152.12, 151.87, 148.36, 148.04, 138.73, 126.18, 125.26, 124.95, 117.12, 115.47, 113.80, 84.11, 83.91, 56.51, 54.93, 54.71, 54.09, 50.00, 49.09, 43.95, 31.74, 30.77, 28.01, 18.79

ESI-MS: m/z=711.50 Th

dimethyl 4-hydroxypyridine-2,6-dicarboxylate

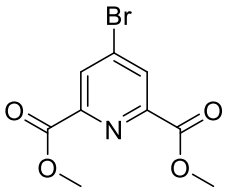


Chelidamic acid (3.3g., 16.41 mmol) was dissolved in MeOH (500mL). H_2SO_4 (1.5mL) was added. The mixture was heated to reflux overnight. It was then quenched with sat. $NaHCO_3$ over an ice-bath. MeOH was removed *in vacuo.* HCl 3% was then added to lower the pH to 2. The aqueous solution was then extracted with 3x200 mL $CHCl_3$. Organic layers were combined and dried over Na_2SO_4. $CHCl_3$ was removed to afford the pure compound as a white solid (3.4 g., 98%)

^{1}H NMR (500 MHz, DMSO-d6): δ 3.37 (s, $COOCH_3$)

^{13}C NMR (125 MHz, DMSO-d6): δ 165.32, 149.76, 115.91, 79.69, 53.05

dimethyl 4-bromo-2,6-dicarboxylate **29**

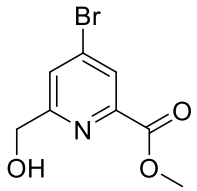


dimethyl 4-hydroxypyridine-2,6-dicarboxylate (3.3g., 15.6 mmol) and TBAB (6.04g., 18.72 mmol) were suspended in Toluene. Reaction mixture was heated to 40°C and P2O5 (7.75g., 54.6 mmol) was. The mixture was then stirred at 100°C for 3 hours. It was then diluted with 100 mL Toluene and quenched with sat. NaHCO3 up to pH=3. The two phases were separated, and the organic layer was further washed with 50 mL water. Organic Layer was dried over Na_2SO_4. Toluene was removed to afford the pure compound as a white solid (4.15 g., 97%)

^{1}H NMR (500MHz, $CDCl_3$): δ 8.48 (s, 2H), 4.06 (s, 6H)

^{13}C NMR (125MHz, CDCl3): δ 164.05, 149.09, 135.14, 131.32, 53.51

methyl 4-bromo-6-(hydroxymethyl)picolinate **32**

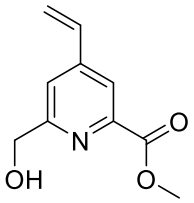


29 (3g., 10.95 mmol) was dissolved in MeOH (270 mL) and DCM (82 mL) over an ice-bath. $NaBH_4$ (994 mg, 26.27 mmol) was added in portions. After 1 hour at room temperature, reaction was quenched with HCl 3%. Organic solvents were removed in vacuo. The mixture was then dilluted in 100 mL water and extracted with 3x300 mL EtOAc. Organic layers were combined and dried over Na_2SO_4. Solvent was removed to afford the pure compound (2.01g., 75%)

^{1}H NMR (500MHz, $CDCl_3$): δ 8.21 (d, 1.8 Hz, 1H), 7.78 (d, 1.8 Hz, 1H), 4.87 (s, 2H), 4.03 (s, 3H)

^{13}C NMR (125MHz, $CDCl_3$): δ 164. 47, 161.71, 148.12, 134.51, 127.35, 12.48, 64.38, 53.19

methyl 6-(hydroxymethyl)-4-vinylpicolinate **33**

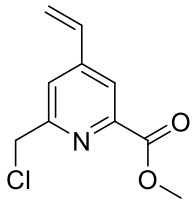


32 (1.8g., 7.3 mmol), K_3PO_4 (3.1g., 14.6 mmol) and trivinylboroxin-pyridin complex (1.3g., 5.5 mmol) were suspended in dioxane (51 mL). N_2 was bubbled through the solution for 5 mns and JohnPhos (6%, 131 mg) and $Pd(dba)_2$ (0) were added, still under N_2. The reaction mixture was heated at 85°C for 2 hours. Dioxane was removed in vacuo and the crude mixture was purified via column chromatography with a mixture of CH_2Cl_2 and EtOAc. Obtention of 960 mg (68%) of the pure product.

^{1}H NMR (500MHz, $CDCl_3$): δ 8.07 (s, 1H), 7.50 (s, 1H), 6.76 (dd, 18 Hz,11 Hz, 1H), 6.12 (d, 18 Hz, 1H), 5.62 (d, 11 Hz, 1H), 4.88 (d, 5.9 Hz, 2H), 4.03 (s, 3H)

^{13}C NMR (125MHz, $CDCl_3$): δ 165.69, 160.57, 147.55, 146.85, 134.00, 121.10, 120.88, 120.28, 64.67, 53.93

methyl 6-(chloromethyl)-4-vinylpicolinate **34**



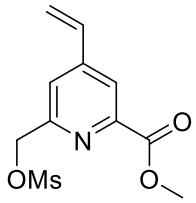
33 (140 mg, 0.72 mmol) was dissolved in DCM (mL) and DIPEA (380uL, 2.17 mmol) then MsCl (112 uL, 1.45 mmol) were added at 0°C. Reaction mixture was stirred for 1h30 at room temperature. DCM was removed and the crude oil was

purified via column chromatography with a mixture of Toluene and EtOAc to yield the product as a yellow solid (145 mg, 95%)

^{1}H NMR (500MHz, $CDCl_3$): δ 7.91 (s, 1H), 7.52 (s, 1H), 6.58 (dd, 18 Hz,11 Hz, 1H), 5.95 (d, 18 Hz, 1H), 5.45 (d, 11 Hz, 1H), 4.62 (s, 2H), 3.85 (s, 3H)

^{13}C NMR (125MHz, $CDCl_3$): δ 165.21, 157.52, 147.84, 147.16, 133.60, 122.99, 121.37, 120.40, 52.86, 46.16

In some cases, methyl 6-(((methansulfonyl)oxy)methyl)-4-vinylpicolinate was also isolated from the crude mixture. It was then further transformed in **34** as such:

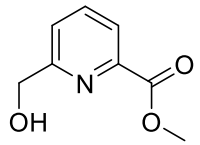


methyl 6-(((methansulfonyl)oxy)methyl)-4-vinylpicolinate (23 mg, 0.08 mmol) and LiCl (10.8 mg, 0.25 mmol) were dissolved in dry DMF (1.6 mL). The reaction mixture was let stirred overnight at room temperature and was then diluted with 20 mL $CHCl_3$ and washed with 2x10 mL H_2O. Organic layer was dried over Na_2SO_4, $CHCl_3$ was removed in vacuo and residual DMF was removed with a high vacuum, pump. Yield: 15 mg (89%)

^{1}H NMR (500MHz, $CDCl_3$): δ 8.10 (s, 1H), 7.62 (s, 1H), 6.58 (dd, 18 Hz,11 Hz, 1H), 6.11 (d, 18 Hz, 1H), 5.63 (d, 11 Hz, 1H), 5.42 (s, 2H), 4.00 (s, 3H), 3.15 (s, 3H)

^{13}C NMR (125MHz, $CDCl_3$): δ 165.27, 154.84, 148.21, 147.50, 133.60, 122.20, 121.99, 120.96, 71.02, 53.07, 38.03

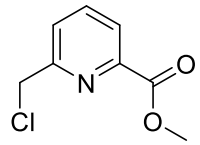
methyl 6-(hydroxymethyl)picolinate **35**



2,6-dimethoxycarbonylpyridine (2.1g., 10.8 mmol) was dissolved in MeOH (283 mL) and DCM (95 mL) over an ice-bath. $NaBH_4$ (817 mg, 21.6 mmol) was added in portions. After 3 hours at room temperature, organic solvents were removed *in vacuo*. The mixture was then suspended in 100 mL water and extracted with 3x150 mL EtOAc. Organic layers were combined and dried over Na_2SO_4. Solvent was removed to afford the pure compound (1.9g., quant.)

^{1}H NMR (500MHz, $CDCl_3$): δ 8.07 (d, 7.8 Hz, 1H), 7.88 (t, 7.8 Hz, 1H), 7.54 (d, 7.8 Hz, 1H), 4.89 (s, 2H), 4.03 (s, 3H)

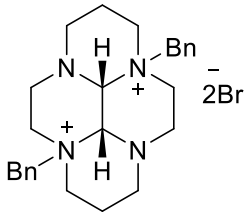
methyl 6-(chloromethyl)-picolinate **36**



38 (1.9 g., 11.3 mmol) was dissolved in dry DCM (150mL) and DIPEA (6 mL, 34.4 mmol) then MsCl (2.7 mL, 34.8 mmol) were added at 0°C. Reaction mixture was stirred for 1h30 at room temperature. DCM was removed and the crude oil was purified twice via column chromatography with a mixture of DCM and EtOAc to yield the pure product (1.4 g. 67%)

^{1}H NMR (500MHz, $CDCl_3$): δ 8.10 (d, 7.8 Hz, 1H), 7.92 (t, 7.8 Hz, 1H), 7.76 (d, 7.8 Hz, 1H), 4.80 (s, 2H), 4.04 (s, 3H)

(3a1R,5a1R)-5a,10a-dibenzyltetradecahydro-3a,5a,8a,10a-tetraazapyrene-5a,10a-diium bromide **37**



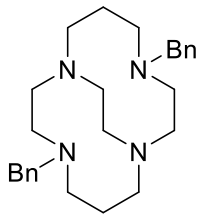
Cis-glyoxal-cyclam (640 mg, 2.88 mmol), Benzyl bromide (5.5 mL), were dissolved in dry MeCN (14 mL). The mixture was stirred at room temperature for 20 days, after which the suspension was filtered on #4 fritt. The precipitate was washed with dry MeCN. Obtention of 1.48 g. of product as a white solid (94%)

^{1}H NMR (500 MHz, D2O): δ 7.59-7.46 (m, 10H), 5.19 (d, J=13.07 Hz, 2H), 5.00 (s, 2H), 4.35 (dt, J=13.45, 3.94 Hz, 2H), 3.66 (dt, J=13.45, 3.94 Hz, 2H), 3.53-3.33 (m, 6H), 3.16 (m, 4H), 2.74 (dt, J=12.52, 3.02 Hz, 2H), 2.20 (m, 2H), 1.83 (bd, J=14.82 Hz, 2H)

^{13}C NMR (125 MHz, D2O): δ 133.19, 131.45, 129.49, 124.69, 76.87, 62.39, 60.54, 51.27, 46.76, 46.09, 17.96, 0.85

ESI-MS: m/z= 202.02 Th

4,11-dibenzyl-1,4,8,11-tetraazabicyclo[6.6.2]hexadecane **38**



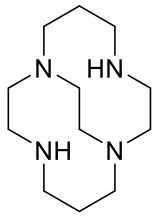
37 (1.5 g., 2.73 mmol) is dissolved in 70 mL 95% EtOH. $NaBH_4$ (5.2 g, 139.2 mmol), is then added in portions at 0°C. After 2 weeks, the reaction was quenched with HCl 3% over an ice-bath. EtOH was evaporated, the aqueous solution was saturated with KOH and then extracted with 5x50 mL DCM. Organic layers w recombined, dried over Na_2SO_4 and solvent was evaporated to yield 1g of product (90%)

^{1}H NMR (500 MHz, CDCl3): δ 7.38-7.22 (m, 10H), 3.98 (td, J=11.59, 4.59 Hz, 2H), 3.81 (d, J=13.54 Hz, 2H), 3.25-3.18 (m, 4H),

^{13}C NMR (125 MHz, CDCl3): δ 141.02, 128.98, 128.02, 126.52, 77.23, 60.03, 57.66, 56.57, 54.24, 52.09, 28.08

ESI-MS: m/z=407.42 Th

1,*4,8,11-tetraazabicyclo[6.6.2]hexadecane* **39**

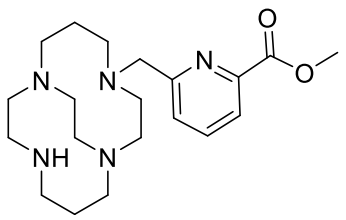


38 (1.5 g, 3.7 mmol), is suspended in water (45 mL), and 3 mL HCl 37% are added. Pd(OH)2/C (1g) and Pd/C (0.5 g) are then added and the mixture is put under H2 atmosphere. After 2 days, the reaction mixture was filtered on a pad of celite and the filatrate was saturated with KOH. The aqueous solution was then extracted with 3x120 mL DCM. Organic layers were combined, dried over Na2SO4, and DCM was evaporated to yield 550 mg of the product as an colorless oil (65%).

^{1}H NMR (500 MHz, CDCl3): δ 3.15 (bt, J=12.62 Hz, 2H), 2.94-2.86 (m, 2H), 2.65-2.73 (m, 6H), 2.70-2.62 (m, 6H), 2.46-2.39 (m, 4H), 1.98-1.87 (m, 2H), 1.4-1.3 (m, 2H)

ESI-MS: m/z= 227.25 Th

methyl 6-((1,4,8,11-tetraazabicyclo[6.6.2]hexadecan-4-yl)methyl)picolinate **40**



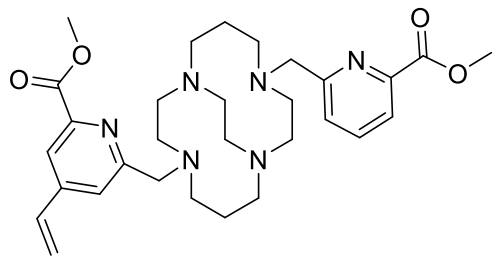
36 (119 mg, 0.64 mmol) was added as solution (2mL MeCN) in a solution of 15 mL (MeCN) of **39** (160 mg, 0.71 mmol). Mixture was stirred at room temperature overnight. MeCN was then removed and the mixture was purified with FlashChromatography using a mixture of DCM and MeOH and 1% NH_3/H_2O. Obtention of 210 mg, 90 %.

^{1}H NMR (500 MHz, CDCl3): δ 8.07 (d, J=8.72 Hz, 1H), 7.95 (t, 7.62 Hz, 1H), 7.61 (d, J=7.91 Hz, 1H), 4.26 (d, J=13.28 Hz, 1H), 3.61 (d, J=13.28 Hz, 1H), 3.27 (m, 1H), 3.21-3.02 (m, 3H), 3.01-2.89 (m, 4H), 2.87-2.76 (m, 4H), 2.74-2.56 (m, 4H), 2.53-2.47 (m, 1H), 1.99 (m, 1H), 1.74 (m, 1H), 1.56 (m, 1H), 1.25 (m, 1H).

^{13}C NMR (125 MHz, CDCl3): δ 165.50, 159.33, 147.26, 138.64, 128.70, 124.17, 63.38, 57.42, 56.10, 55.43, 53.92, 53.71, 53.28, 52.30, 50.87, 50.15, 47.32, 44.73, 26.82, 21.73

ESI-MS: m/z=386.42 Th

methyl 6-((11-((6-(methoxycarbonyl)pyridin-2-yl)methyl)-1,4,8,11-tetraaza bicyclo [6.6.2]hexadecan-4-yl)methyl)-4-vinylpicolinate **42**



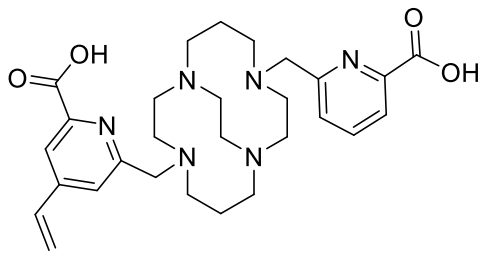
40 (430 mg, 1.14 mmol), **34** (242 mg, 1.14 mmol) and K_2CO_3 (633 mg, 4.58 mmol) were dissolved in MeCN (5 mL). Reaction mixture was stirred at room

temperature overnight. MeCN was then removed and the mixture was purified via Flash Chromatography using a mixture of DCM and MeOH and 2% NH_3/H_2O. Obtention of 510 mg, 81 %.

^{1}H NMR (500 MHz, CDCl3): δ 10.63 (bs, 1H), 8.00 (m, 2H), 7.81 (t, J=7.80 Hz, 1H), 7.53 (d, J=7.59 Hz, 1H), 7.47 (s, 1H), 6.72 (dd, J=17.83, 11.16 Hz, 1H), 6.1 (d, J=17.62 Hz, 1H), 5.62 (d, J=10.96 Hz, 1H), 4.03 (d, J=15.06 Hz, 1H), 4.00 (s, 3H), 3.99 (s, 3H), 3.88 (d, J=14.58 Hz, 2H), 3.49-3.38 (m, 2H), 3.36-3.25 (m, 4H), 3.24-3.11 (m, 4H), 3.05-2.93 (m, 6H), 2.70-2.62 (m, 2H), 1.77 (m, 8H)

^{13}C NMR (125 MHz, CDCl3): 165.67, 158.68, 158.20, 148.21, 147.83, 146.61, 144.79, 137.45, 133.81, 17.17, 124.13, 123.90, 120.96, 120.53, 58.81, 54.65, 54.56, 52.82, 51.97, 51.83, 30.95, 24.37, 24.26

methyl 6-((11-((6-(methoxycarbonyl)pyridin-2-yl)methyl)-1,4,8,11-tetraaza bicyclo[6.6.2]hexadecan-4-yl)methyl)-4-vinylpicolinate **T2**



42 (190 mg, 0.34 mmol) was dissolved in 2.3 mL THF. LiOH (58 mg, 1.38 mmol) was dissolved in 2.3 mL H_2O. Solutions were mixed and stirred at room temperature overnight. HPLC-MS reveals complete conversion. Solvents are removed in vacuo and the mixture is purified via semi-preparative HPLC with a H_2O+0.1%TFA/MeCN 100/0 to 50/50 gradient. Obtention of 100 mg of pure product.

^{1}H NMR (500 MHz, D2O): δ 7.44 (m, 3H), 6.79 (dd, J=6.62, 2.31 Hz, 1H), 6.76 (s, 1H), 6.49(dd, J=17.76, 10.66 Hz, 1H), 5.90 (d, J=17.76 Hz, 1H), 5.49 (d, J=10.91 Hz, 1H), 4.92 (d, J=16.50 Hz, 1H), 4.81 (d, J=16.12 Hz, 1H), 4.10 (m, 2H), 3.94-3.81 (m, 4H), 3.68-3.52 (m, 4H), 3.45 (m, 2H), 3.13 (bd, J=14.4 Hz,

1H), 2.99 (bt, 2H), 2.78 (bd, 2H), 2.62-2.51 (m, 4H), 2.49-2.36 (m, 2H), 1.61 (m, 2H)

ESI-MS: m/z= 386.42 Th

Formation of **[Ln(T2)]**

T2 (20 mg) was dissolved in 1 mL H_2O and 1 mL MeCN. pH was adjusted to 7 and $EuCl_3$.6H2O (30 mg) was added to the solution. The mixture was placed in a micro-wave vial irradiated at 150W, 150°C for 6 hours. The crude mixture was then purified via semi-preparative HPLC.

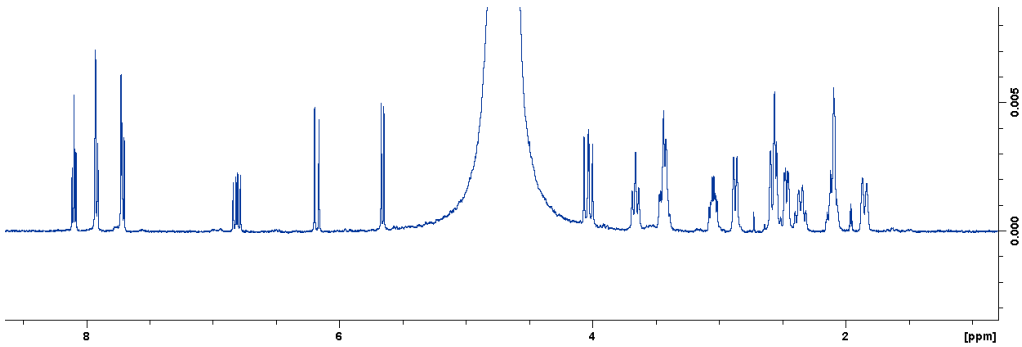
[La(T2)] ^{1}H NMR (500 MHz, D2O): δ 8.10 (t, J=7.66 Hz, 1H), 7.92 (m, 2H), 7.72 (m, 2H), 6.81 (dd, J=17.93, 11.08 Hz, 1H), 6.18 (d, J=17.53 Hz, 1H), 5.66 (d, J=10.85 Hz, 1H), 4.04 (dd, J=15.18, 11.59 Hz, 2H), 3.66 (bt, J=13.69 Hz, 2H), 3.44 (m, 4H), 3.09-3.00 (m, 2H), 2.87 (bd, J=13.25 Hz, 2H), 2.62-2.53 (m, 4H), 2.50-2.44 (bdd, 2H), 2.40-2.30 (bq, 2H), 2.16-2.05 (m, 4H), 1.85 (bd, 2H),

ESI-MS: [Yb(T2)]: 694.50 (100%), 692.69 (80%), 693.68 (76%), 691.71 (60%)

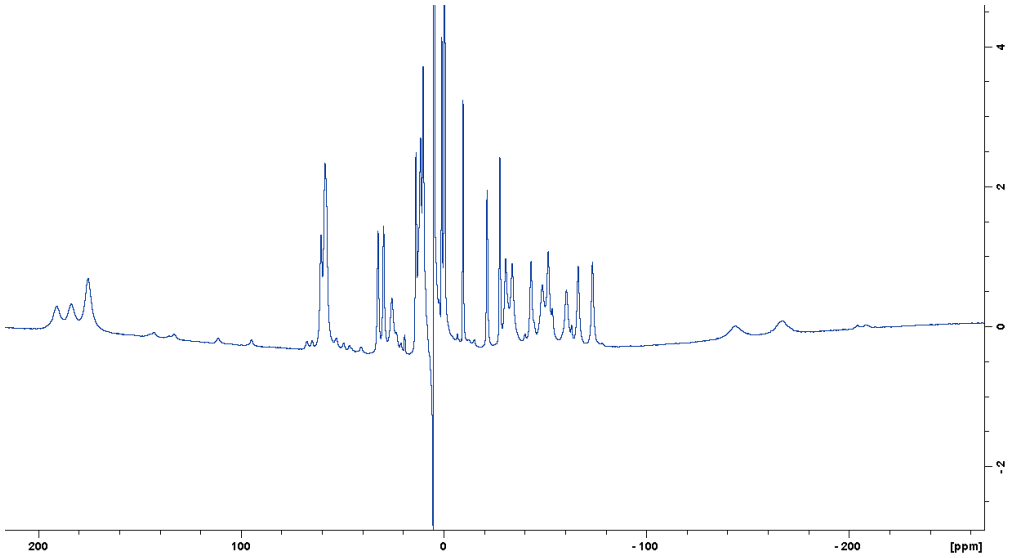
[La(T2)]: 659.47 (100%)

[Dy(T2)]: 684.43 (100%), 683.53 (95%), 682.56 (89%)

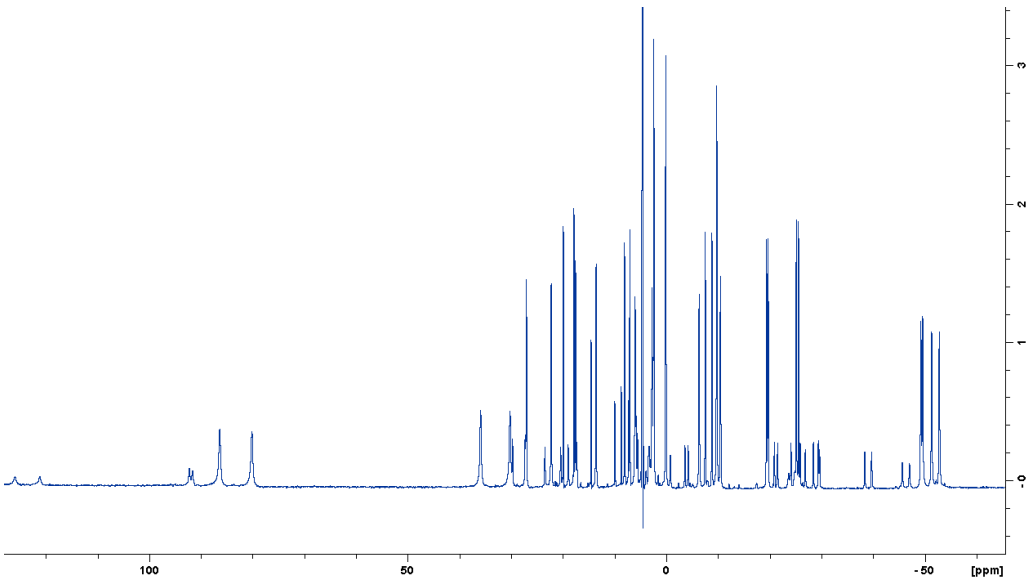
3.6.2. NMR Spectra



NMR spectrum of [La(T2)]



NMR spectrum of [Dy(T2)]



NMR spectrum of [Yb(T2)]

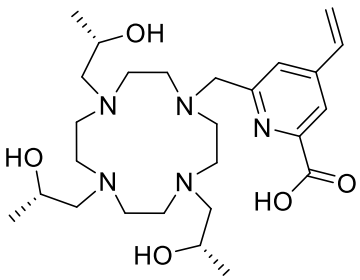
3.6.3. Conjugation

Reduced GB1 T53C was buffer exchanged to NaPi 20mM pH 7.5, at a concentration of 200 µM. 10 equivalents of [Ln(T2)] were added and the mixture was placed at 70°C and let overnight, after which reaction was complete. Excess tag was removed via buffer exchange and NMR spectra were recorded.

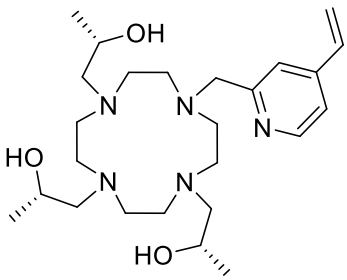
CHAPTER 4: Synthesis of enantiopures paramagnetic tags

Abstract:

As we have seen in the previous chapter, the thiol-ene reaction is a powerful and promising methodology for the conjugation of paramagnetic tags to proteins. Driven by the necessity of designing an enantiomerically pure lanthanide-chelating tag, we have decided to explore other scaffolds that could be conjugated via the thiol-ene reaction. THP-like scaffolds have been shown to be usable as LBTs by Otting *et. al.*[80], being thermodynamically stable and displaying only one set of peaks in the NMR spectra. In this chapter, we will discuss the design of THP-like paramagnetic tags bearing a vinyl-pyridine moiety for the attachment to cysteines. We will focus on two scaffolds, T3 and T4, differing by the number of atoms coordinating the lanthanide ion. Their synthesis and performance on GB1 T53C are presented and the influence of the coordination environment will be discussed.



T3



T4

Figure 4.1: T3 and T4, ligands discussed in this chapter

4.1 Introduction

4.1.1 THP

((2S,2'S,2''S,2'''S)-1,1',1'',1'''-(1,4,7,10- tetraazacyclododecane1,4,7,10 -tetrayl)tetrakis(propan-2-ol)) (also called (S)-THP) is a cyclen derivative featuring four chiral (S)-2-hydroxypropyl pendants. It was introduced by Chin *et al*. as a chelating agent for lanthanides[89]. It has been shown that it forms inert complexes at neutral pH with early, middle, and late lanthanides. (S)-THP is interesting in that it bears chirality within its pendant arms, which is expected to have an influence on its dynamic, conformational space and chelation properties[90], and that its synthesis is quite simple.

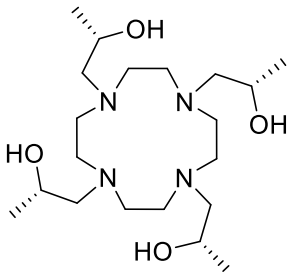


Figure 4.2: (S)-THP, an enantiopure ligand

It has been shown via NMR that (S)-THP-Ln complexes present only one conformation, as only one set of peaks is observed via 1D ^{1}H and ^{13}C NMR spectroscopy. Indeed, the (S)-THP-Yb^{3+} complex specifically has been shown to adopt a Λ(λλλλ) TSAP geometry in solution[90]. This is of course of interest for the design of enantiopures paramagnetic tags for NMR spectroscopy. Another interesting feature of (S)-THP is the apparent absence of dynamic process on the NMR timescale. This for example, is not the case of DOTA its NMR spectra

shows important temperature-dependent differences in terms of shifts and of peak-broadening[74].

In 2015, Lee *et al*. proposed (S)-THP based paramagnetic tags for NMR spectroscopy (fig. 4.3)[85]. They designed different linkers, one of which particularly short. They showed that the tags kept the interesting properties of (S)-THP in terms of chelation and of stereochemistry. They were able to transfer medium to large anisotropic tensors on different proteins. Those tags, however, have the limitation that they use disulfide formation as the conjugation reaction. Seeking to combine the properties of (S)-THP in terms of stereochemistry and the stability of the thioether bond product of the thiol-ene reaction, we designed T3. In order to study the effects of coordination on the complexation of lanthanides but also on the tag-s behavior once attached on a protein, we designed T4, displaying only 8 coordinating atoms.

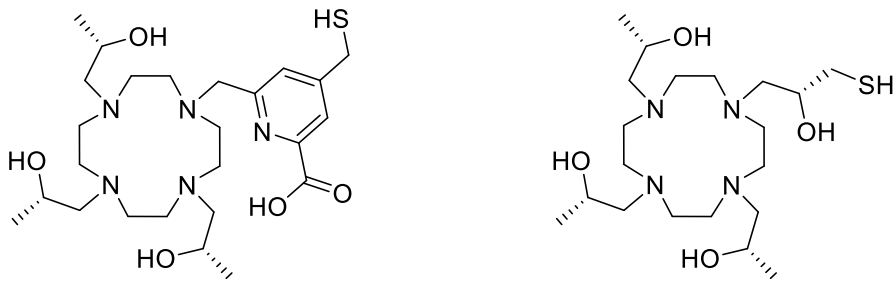


Figure 4.3: C5 and C7**,** described LBT[85]

4.1.2 Relaxometry and lanthanide-coordination:

Fast-Field Cycling NMR (FFC), also called Nuclear Magnetic Resonance Dispersions (NMRD), consists in the measurement of the spin-lattice relaxation rate constant as a function of the frequency of the applied magnetic field[91]. It is used to measure the relaxation rate enhancement induced by the presence in solution of paramagnetic metal ions. An example of NMRD profile is given in Figure 4.4, where the relaxivity of a paramagnetic Iron III (measured in s^{-1} mM^{-1}) is displayed as the function of the Larmor Frequency in two different solvents.

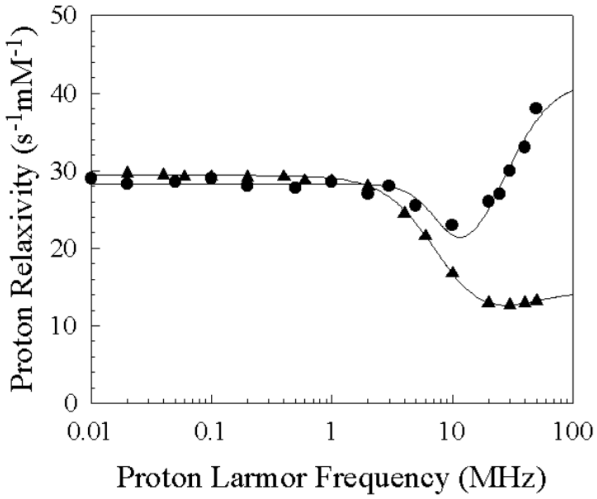


Figure 4.4: Paramagnetic enhancements to solvent 1H NMRD profiles for $Fe(H_2O)_6^{3+}$ solutions at 298 K with (▲) pure water and (●) 60% glycerol[91]

The enhancement in water proton relaxation rate due to the presence of a paramagnetic complex in solution at a given magnetic field is a linear function of the paramagnetic metal ions' concentration[92]. At a fixed concentration, the enhancement depends on several parameters, among which, the magnetic field, the nature of the metal ion, the availability of coordination sites for solvent molecules, the temperature, the size of the paramagnetic molecule and the viscosity of the solution.

Furthermore, the analysis of NMRD profiles may also provide information at a molecular level on the structure and the dynamic of the studied paramagnetic system[92]. Parameters accessible are: geometrical parameters which determine the position of protons relative to the paramagnetic site, the number of protons in the first coordination sphere, and possibly in the second coordination sphere[93], as well as the distance of closest approach of diffusing water protons. Those structural parameters are linked to dynamic parameters such as the reorientational correlation time τ_R and the life-time τ_M. We will discuss which of those parameters are of interest for us in the specific case of lanthanide-

small molecule complexes and how can that information can be useful for the design of paramagnetic tags.

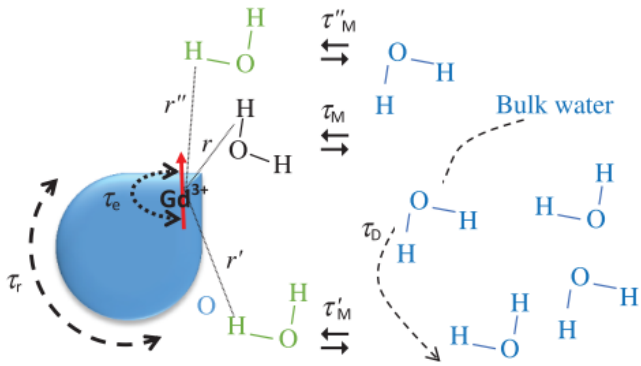


Figure 4.5: The relaxivity of bulk water protons can have contributions from first sphere water molecules (in black), second sphere water molecules (in green) and from the free diffusion of outer-sphere molecules (in blue). Reproduced from Bertini *et al.*[92]

Controlling the coordination of lanthanides in paramagnetic tags is paramount. Indeed, the coordinating environment not only will have a direct influence on the intrinsic anisotropic tensor of the paramagnetic lanthanide, but it will also play a role in the dynamic of the complex[73,94,95]. As discussed in the literature as well as in chapter 3, electrostatic interactions between protein side-chains and lanthanide complexes have the possibility to stabilize greatly those complexes on the surface of the protein. This generally occurs between tags (positively charged) and negatively charged side-chains (carboxylates). In order for this interaction to happen, the tag needs to be placed at a suitable distance from a negatively charged side-chain[87]; however, this interaction can take place only if the coordination of the complexed lanthanide allows it. This means that some “coordinating space” must be available around the lanthanide. This can be measured via FFC NMR. Indeed, parameters accessible via NMRD profile analysis are the presence of water molecules in the first coordination spheres

and its time of residency τ_M. The presence of one or more first-coordination sphere water around the paramagnetic ion indicates that coordination by the ligand is not complete and further coordination interaction are possible[92]. The time of residency τ_M indicates how easily displaceable this water is.

Because NMRD profile-analysis allows us to study the interaction between a metal-complex and its chemical environment, and because it allows the evaluation of the presence or coordinated water molecules around the metal, FFC could be of a great help when designing paramagnetic tags. This technique is already well-established for the development of Gadolinium-based probes in MRI imaging[92]. Gadolinium is used because it bears 7 unpaired electrons and has a very long electronic relaxation time ($T_{1e} \geq 2 \times 10^{-10}$ s, at room temperature), which allows for the design of contrast agents with very high relaxivity. Since lanthanides have similar chemical properties, it can be assumed that [Gd-ligand] complexes are representatives in terms of coordination for the complexes of the same ligand with the whole series of lanthanides.

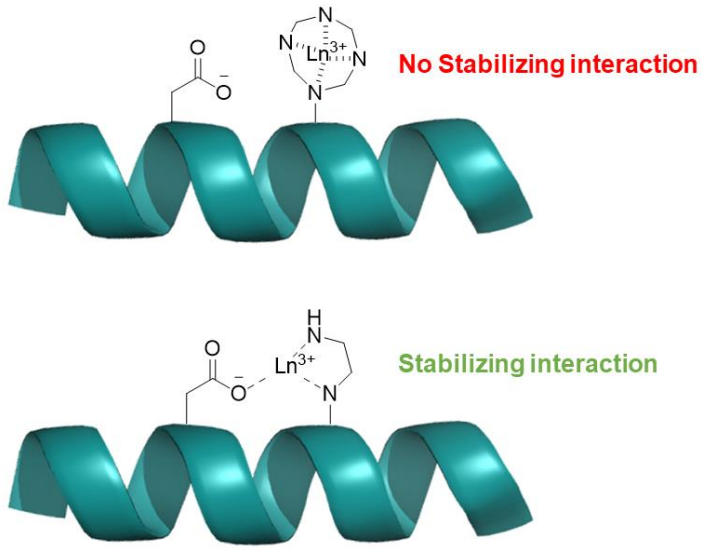


Figure 4.6: Interaction between a side-chain and a lanthanide may stabilize a paramagnetic tag on a protein surface

4.2 Synthesis of T3, T4 and their Lanthanide complexes

4.2.1 Synthesis of T3

4.2.1.1 Retrosynthesis

Mono-alkylation of cyclen has been reported in the literature. The approach chosen to synthetize **T3** was to first alkylate cyclen with a picolinic-like arm and then the remaining 3 nitrogen atoms of the macrocycle. To this purpose, it is possible to use intermediate **34** described in chapter 3. The chiral arms can then be introduced using (S)-propylen Oxide as shown by Lee *et al*[85].

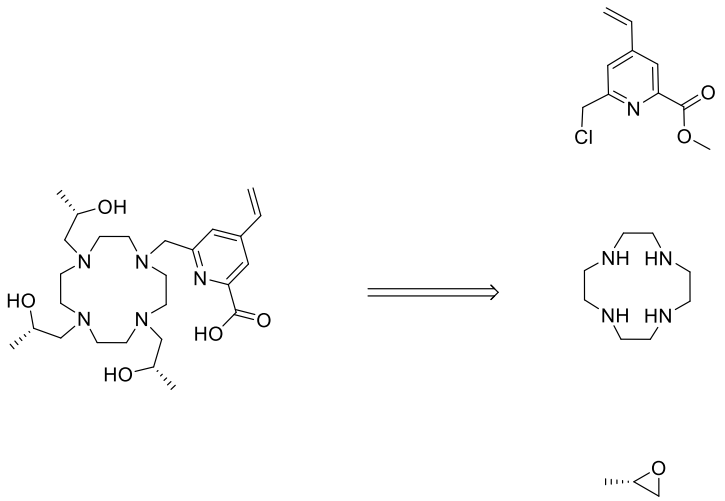


Figure 4.7: Retrosynthesis of **T3**

4.2.1.2 Synthesis of T3

The synthesis of **34** had been already optimized and its synthesis has been described in chapter 3. The first step consists in the selective mono-alkylation of cyclen and is achieved by using a large excess of it (10 eq). Using a smaller amount of cyclen leads to formation of dialkylated cyclen. Next, alkylation of the remaining secondary amines using enantiomerically pure (S)-propylene oxide was achieved. The formation of a single product was confirmed by NMR and HPLC. Last, methyl ester was deprotected to yield **T3** (fig. 4.8), which was purified via semi-preparative HPLC. Overall, T3 was obtained in high yields and only 8 steps from commercial reagents.

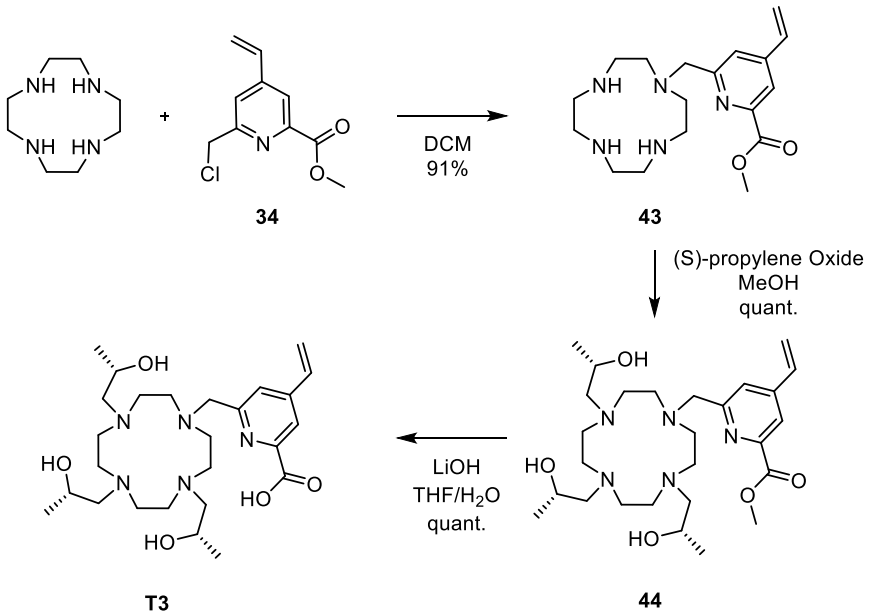


Figure 4.8: Synthesis of **T3**

4.2.1.3 Formation of Lanthanide complexes

Lanthanide-complexation was achieved by refluxing T3 with an excess of $LnCl_3$ overnight. No lanthanide-dependence in the reactivity for the complexation was observed. Once again, only one peak was observed via HPLC. This indicates that upon complexation no diastereoisomeric species are generated. Since 3 chiral centers have their configuration blocked, if the chelation induced stereogenicity it would lead to the creation of diasteroisomers. At this point it is then possible to hypothesize that the [Ln(T3)] complexes only have one configuration possible and are therefore enantiopures.

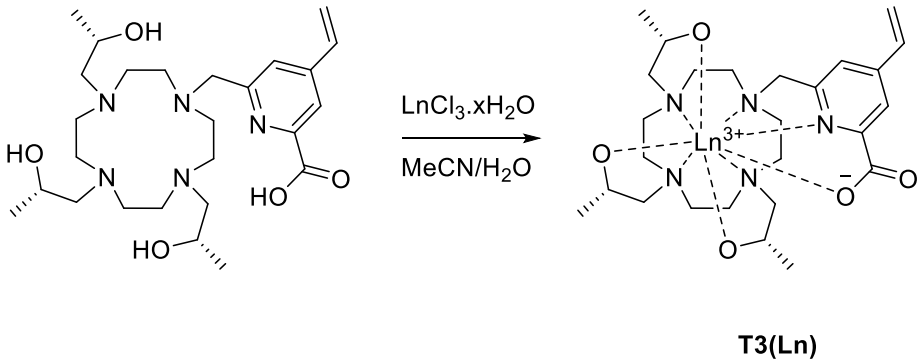


Figure 4.9: Lanthanide Complexation

On figure 4.10 hereafter, we can see the 1D-^{1}H NMR spectrum of [T3(Yb)]. Main features of this spectrum can be attributed to paramagnetic effects: the important line-broadening is caused by Paramagnetic Relaxation Enhancement (PRE), while the large shifts (up to 36 ppm) are mainly due to Pseudo-Contact Shifts (PCS). Only one set of peaks is individuated. The samples studied via NMR are stable over days at neutral pH, what is consistent with published data[60,80].

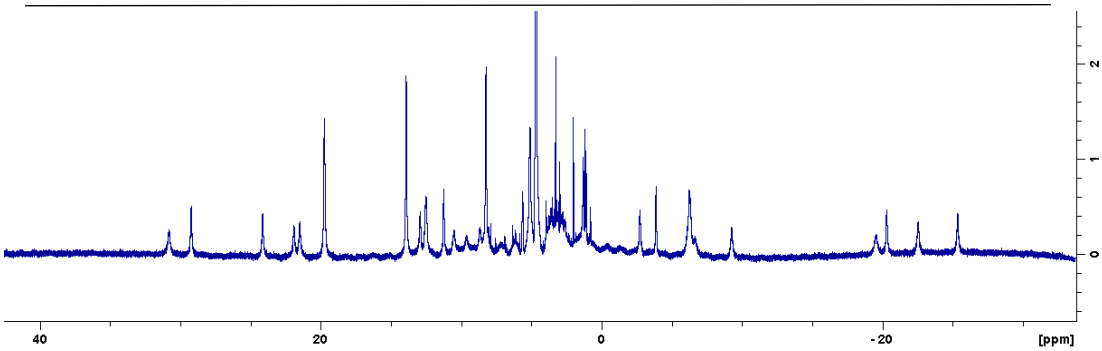


Figure 4.10: ^{1}H 1D NMR spectra of [Yb(T3)]

4.2.2 Synthesis of T4

4.2.2.1 Retrosynthesis

It was chosen to follow the same synthetic pathway as for T3. The first step would be the alkylation of cyclen with the vinyl-pyridinic arm and then of the remaining 3 nitrogen atoms of the macrocycle. To this purpose, it was required to isolate intermediate **46**. The chiral arms can then be introduced using (S)-propylen Oxide as shown before.

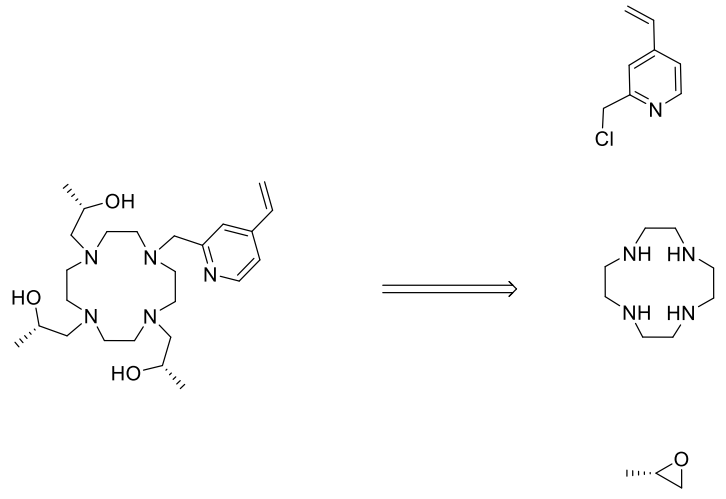


Figure 4.11: Retrosynthesis of **T4**

4.2.2.2 Synthesis of T4

A similar strategy to the one used to obtain **34** was designed to isolate **46.** Since there is in this case no carboxylic acid to protect, **46** could be obtained in only two steps starting from commercial 4-bromo-2-pyridinemethanol. First, the double bond was introduced in the position 4 of the pyridine via a Suzuki-like coupling. Then, the benzylic position in **45** of the pyridine was activated using Mesyl Chloride. The product spontaneously transformed in the chloride derivative and isolated as such.

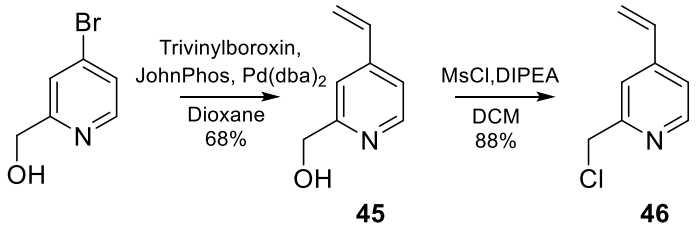


Figure 4.12: Synthesis of **46**

With intermediate **46** in hand, next was performed cyclen alkylation. Once again, using an excess of cyclen allowed for the isolation of mono-alkylated compound. Alkylation with (S)-propylene Oxide yielded the desired product which displayed only one peak via HPLC. Not having to deprotect a carboxylic acid, T4 was then obtained in only 4 steps from commercial reagents.

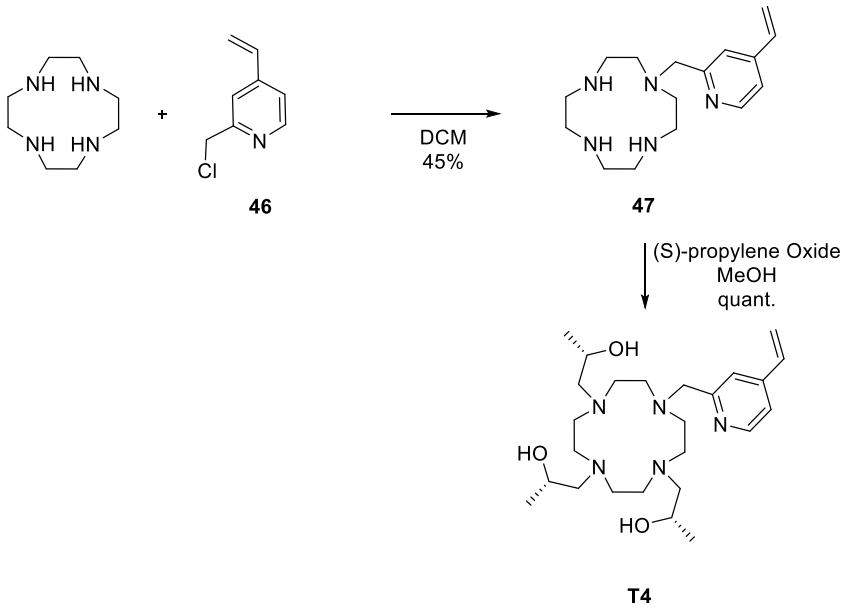
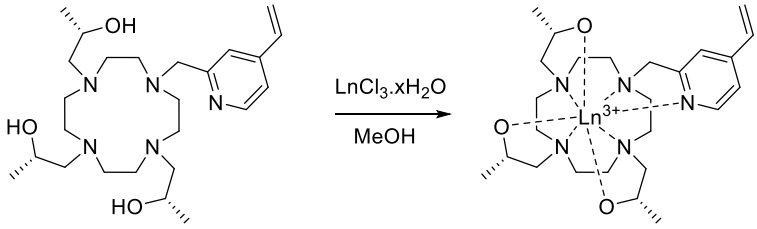


Figure 4.13: Synthesis of **T4**

4.2.2.3. Formation of Lanthanide complexes

Lanthanide-complexation was achieved by refluxing T3 with an excess of $LnCl_3$ overnight. However, it was noted that the reaction did not proceed equally with all lanthanides. In fact, after several trials with different metals it seemed than the reaction proceeded better with early lanthanides (c.f. Table 1). We can see that T4 seems to have a certain “preference” in terms of chelation for early and middle lanthanides. In terms of coordination properties, lanthanides are very similar and the main difference between them is their ionic radius, governed by the lanthanide contraction. It can be hypothesized that for geometrical reasons, bigger lanthanides fit better in the cavity of T4 and the complexes are therefore more prone to formation. Such behavior is rare but has been reported in the literature[96]

.



T4

Figure 4.14: Lanthanide Complexation

Lanthanide	Ionic Radius (pm)	Chelation (%)
La	**103**	**100**
Ce	**102**	**100**
Gd	**93.8**	**100**
Dy	**91.2**	**100**
Yb	**86.8**	**86**
Lu	**86.1**	**40**

Table 1: Chelation of diverse lanthanides by **T4**

4.3 Photocatalyzed Thiol-ene-based conjugation to GB1

4.3.1 Photocatalysis for the Thiol-Ene Coupling

As shown in chapter 3, TEC is not always spontaneous. If it is considered that the reaction goes through a Michael-addition-like mechanism, both the nucleophilicity of the cysteine and the electrophilicity of the double-bond need to be considered when evaluating the feasibility of the reaction. In chapter 3 we showed that heating the reaction led to a significant increase in the reaction rate, but whether it favorishes the Michael-addition-like pathway or the radical one is unknown.

It has been shown that the radicalar pathway for TEC can be catalyzed by the use of photoactivated radical initiators[81]. Indeed, such molecules can upon irradiation generate radicals *in situ* and therefore initiate radical chain reactions such as TEC. An indicated photoinitiator for biomolecular TEC is Dimethoxyacetophenone (DPAP), which is activated at a wavelength of 365 nm. Upon irradiation, the C-C bond in α of the ketone breaks in an homolytic fashion, thus generating radicalar species, which will then rearrange and transfer their radical to the sulfhydryl of a cysteine (fig. 4.15). DPAP has the advantage of being activated at 365 nm, wavelength at which proteins generally do not absorb and therefore are not damaged by this irradiation.

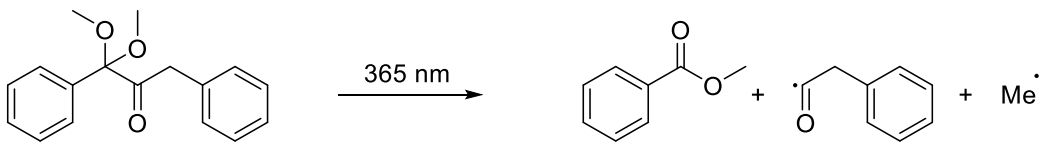


Figure 4.15: Homolytic rupture of DPAP under UV

This methodology has been successfully applied to the generation of glycoconjugates[84]. To the best of our knowledge, it has however never been applied to the paramagnetic tagging of proteins.

4.3.2. Reaction conditions optimization

As a model to assess the possibility of catalyzing the TEC-based paramagnetic tagging of proteins, it was decided to study the reaction between GB1 T53C and vinyl dipicolinic acid (VDPA) in presence of DPAP under irradiation at 365 nm. When optimizing the reaction, 3 parameters were particularly studied: the amount of VDPA, the amount of DPAP, and the presence of O_2 in the reaction mixture. In a first attempt, 5 equivalents DPAP and 5 equivalents VDPA were used. After 30 mns, the tagging was of around 60%. In a second time, 10 equivalents of DPAP and of DPA were used. The conjugation was higher after 30 mns (around 75%) but some minor species seemed to have appeared in the spectra, maybe due to non-covalent interactions between small molecules and the protein. It was found that by using 5 equivalents of DPAP and VDPA, irradiating the solution 1h at 365 nm and then repeating the procedure, spectra were cleaner and the conjugation higher (over 90%) than by using directly 10 equivalents of DPAP/VDPA. The presence of O_2 in the mixture seems to slow the reaction down and therefore further experiments were all conducted after having bubbled Argon through the mixture.

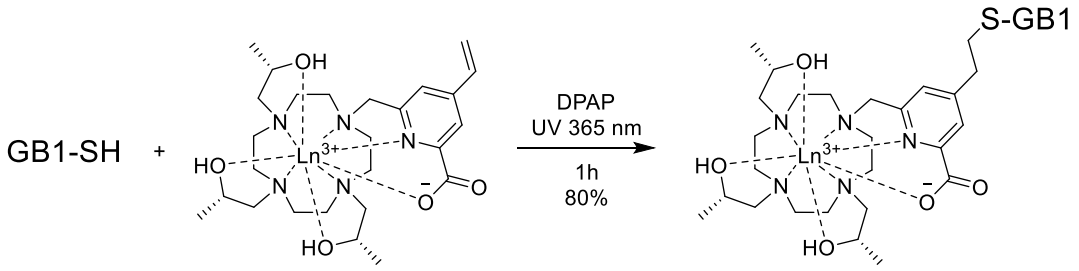


Figure 4.16: Conjugation of [Ln(T3)] to GB1 T53C

4.3.3 Conjugation to GB1 T53C

4.3.3.1 Conjugation of T3 to GB1 T53C

In order to investigate the reactivity of T3, a first experiment consisted in an attempt to tag GB1 T53C at room temperature. After incubating the protein with 10 equivalents of [T3(Yb)], NMR analysis of the sample (via 1D ^{1}H and 2D ^{15}N-^{1}H HSQC) revealed that the protein had not reacted and was left intact. Signals corresponding to [T3(Yb)] were also found and no sign of releasing of the lanthanide were observed, what indicates that the tag was stable under these conditions. Since T3 bears the same 4-vinyl-picolinic acid derivative pendant arm that T2, this result was expected.

Next, possibilities to catalyze the reaction with DPAP were explored. Conditions used were the same as optimized in 4.3.2. To a solution of GB1 were added 5 equivalents of tag and 5 of DPAP. The mixture was irradiated with a UV lamp at 365 nm for 1 h after which the sample was analyzed by NMR. After 1h, tagging was of 80% (as revealed by 2D 1H–15N HSQC spectrum). Another 5 equivalents of DPAP and of tag may then be added to the mixture and re-irradiated in order to push the reaction further.

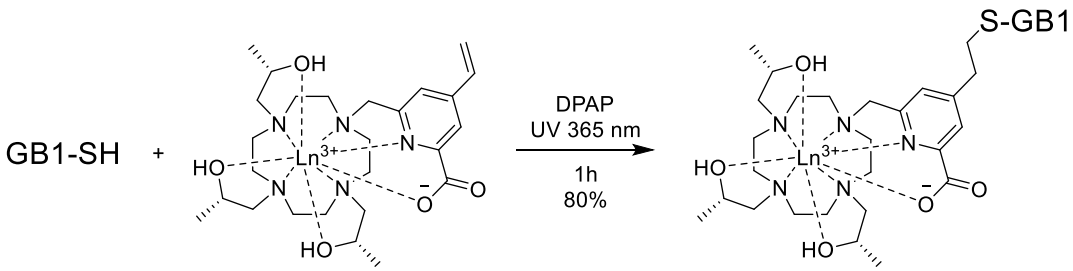


Figure 4.17: Conjugation of [Ln(T3)] to GB1 T53C

Tagging of GB1 T53C with paramagnetic [T1(Yb)] and [T1(Dy)] resulted in PCS in the 2D 1H–15N HSQC spectrum (fig. 4). As expected, the two paramagnetic metal ions provide shifts in opposite directions. The intensities of the shift were medium; it was relatively easy to assign the peaks of the paramagnetically-tagged species to the ones of apo protein. In order to differentiate between the chemical shift perturbation stemming from the tagging itself and the paramagnetic contribution to the shifts (PCS), GB1 was also tagged with the diamagnetic [T1(Lu)]. The resonances which were shifted after the tagging of the diamagnetic compound corresponded to residues located close to the tagging-site. This confirms that T3 was attached only on one site and that this site was the one which we expected, the Cysteine in position 53.

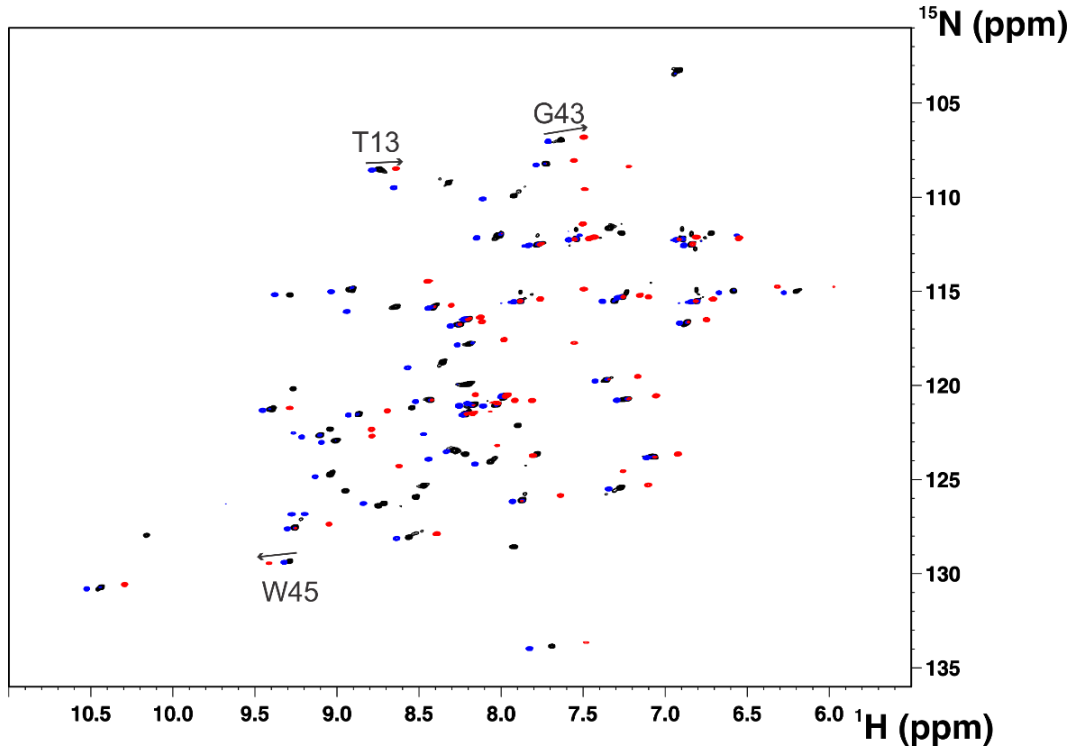


Figure 4.18: 2D ^{1}H–^{15}N HSQC acquire on GB1 T53C tagged with T3(Lu) (black), T3(Yb) (blue), T3(Dy) (red) at 700 MHz and 298 K

We noted that for each cross-peak in the 2D 1H–15N HSQC of diamagnetic samples there was only one corresponding cross-peak in the spectra of paramagnetic species. We therefore concluded that T1 was indeed present as a single stereoisomer or a single conformation which gave rise to a single set of paramagnetically shifted peaks.

4.3.3.2 Conjugation of T4 to GB1 T53C

The same procedure yielding tagging of T3 to GB1 T53C was repeated with T4, loaded with Lanthanum (La), Cerium (Ce), and Dysprosium (Dy), respectively. Once again, the reaction was followed by NMR spectroscopy. The reaction mixture was first analyzed after 1 hour of reaction under UV at 365 nm. We observed many different species in the spectra, and the shifts observed did not seem to be PCS, because they were not diagonal. Procedure was repeated in different conditions and with different amount of DPAP and of tag, but the results were similar: we observed the presence of many species, and spectra proved difficult to interpret. In order to characterize this mixture, MALDI experiments were conducted. The results are shown in figure 4.20. First, we can notice the presence of many peaks corresponding to many different species, what was expected from the NMR experiments. We can see the peak corresponding to the protein tagged with [Dy(T4)] (6913 Th). However, the main peak corresponds to the mass of the protein tagged with two complexes (7537 Th).

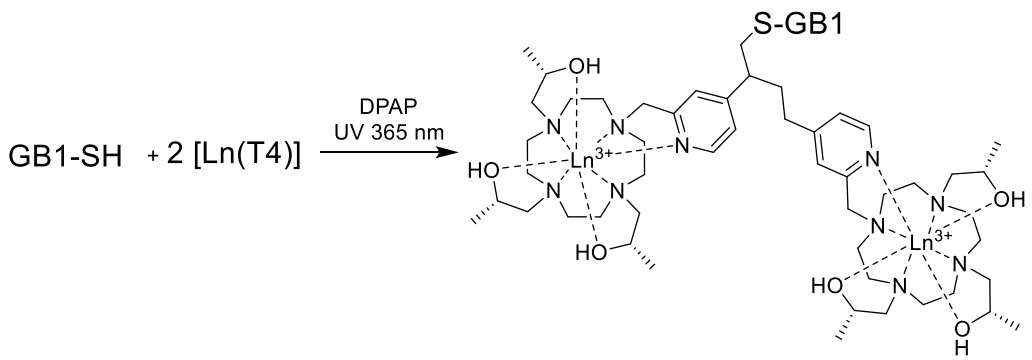


Figure 4.19: Proposed conjugation of T4 to GB1 T53C

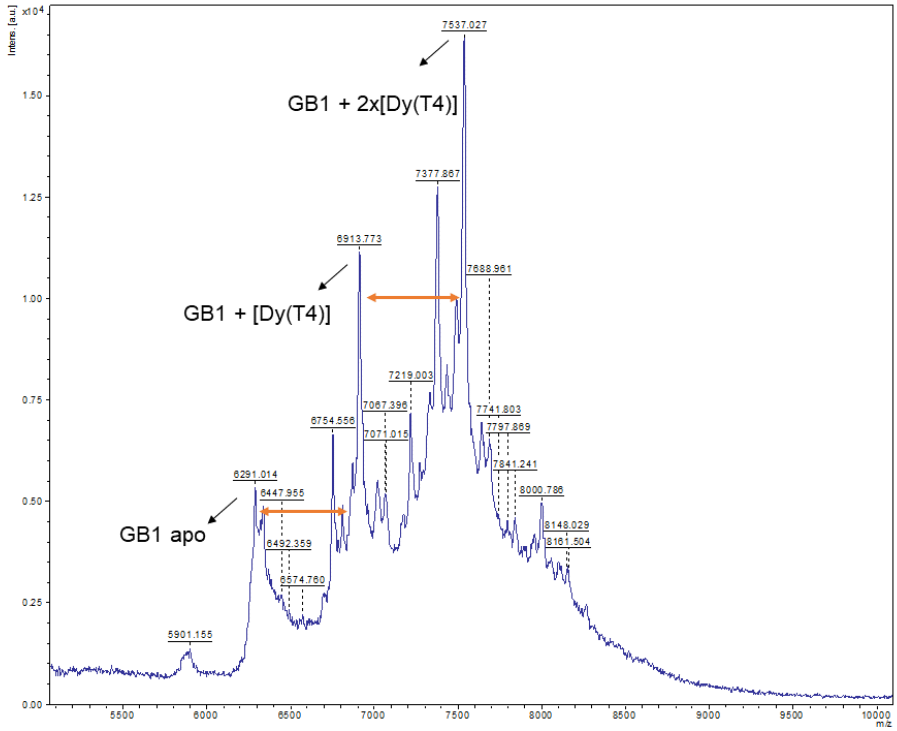


Figure 4.20: MALDI analysis of GB1 T53C tagged with **T4**

A first hypothesis would be that another site of the protein was tagged; however, we didn't detect any other reactive site during reaction with T3, and this would also be in contradiction with data from the literature. Another hypothesis would be that the radical generated upon the addition of the radical thiol on the double bond was more stable than the one generated by the addition of the radical thiol to T3 because of the absence of carboxylate on the pyridine. The radical formed may then be able to perform a second addition on another free molecule of tag, thus yielding a double-tagged molecule (fig. 4.19). Looking at the MALDI spectrum, we may discern peaks corresponding to a triple-tagged protein (8000 and 8148 Th) what could indicate that this process could go on for a third addition. This is in line with literature, where the polymerization of vinyl-pyridine is described[97]. The vinyl-pyrindine in T4 therefore seems incompatible with the radical reaction used in this tagging protocol.

4.3.4. Evaluation of paramagnetic effects

As discussed in the first chapter, the PCS are commonly described as a function of the nuclear coordinates and of an anisotropy tensor ($\boldsymbol{\Delta\chi}$), according to the equation:

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

where *r*, θ and φ are the spherical coordinates of the nucleus in the frame where the anisotropy tensor is diagonal and with origin onto the metal position. $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic anisotropies of the tensor, defined as:

$$\Delta\chi_{ax} = \chi_{zz} - \frac{\chi_{xx} + \chi_{yy}}{2}$$

$$\Delta\chi_{rh} = \chi_{xx} - \chi_{yy}$$

The program FANTEN can be used to obtain the best fit $\boldsymbol{\Delta\chi}$ tensor (consisting in 8 parameters: $\Delta\chi_{ax}$, $\Delta\chi_{rh}$, the three Euler angles defining the frame where the tensor is diagonal, and the coordinates of the metal ion from the PCS values and the coordinates of the protein nuclei).

First, the resonances of 19 peaks in the spectra of the protein tagged with [T3(Yb)] were assigned unambiguously, and the PCS were evaluated. The anisotropy tensor for the Yb^{3+} tagged protein was determined from the best fit of these 19 HN PCS to the X-ray structure (PDB 1IGD), using the program FANTEN. During these tensor calculations, the position of the metal was also obtained. New NMR peaks could then be assigned taking advantage of the PCS predicted by the program for the other nuclei, so that a total of 36 PCSs could

be obtained for the [**T3**(Yb)] tagged protein. The assignment of the spectra of the protein tagged with [**T3**(Dy)] was obtained by comparison with the assigned spectra of the [**T3**(Yb)] and [**T3**(Lu)] tagged proteins (fig. 4.18); in this way 32 PCS could be obtained for [**T3**(Dy)].

The PCS originated by [**T3**(Yb)] and [**T3**(Dy)] were first analysed separately. The metal positions obtained from the two sets of data were similar, as well as the direction of the main axes of the best fit tensors (fig. 4.21). Then, the two sets of PCSs were evaluated jointly by constraining both metals to reside in the same position: the orientations and magnitudes of the two tensors were almost unaffected with respect to the values obtained from the separate fits, thus showing the consistency of the two sets of paramagnetic data (Table 2 and fig. 4.20). The two tensors are almost coaxial, with angles of 11.5° between the two Z-axes, 5° between the two X-axes and 12° between the two Y-axes. The agreement between experimental and back-calculated PCS obtained from the fit of both sets of data is very good, with Q factors of 0.064 and 0.072, for [**T3**(Yb)] and [**T3(**Dy)], respectively (fig. S3).

Metal	**Q factor**	$\Delta\chi_{ax}$ **(10^{-32} m^3)**	$\Delta\chi_{rh}$ **(10^{-32} m^3)**
Yb (alone)	0.061	1.91 ± 0.02	0.68 ± 0.25
Dy (alone)	0.072	-6.51 ± 0.02	2.54 ± 0.89
Yb (+ Dy)	0.064	1.83 ± 0.03	-0.59 ± 0.28
Dy (+Yb)	0.072	-6.81 ± 0.21	3.06 ± 0.89

Table 2: Tensor parameters calculated with the program FANTEN using the PCSs measured with [T3(Yb)] and [T3(Dy)] implemented separately or jointly in the evaluation of the metal position.

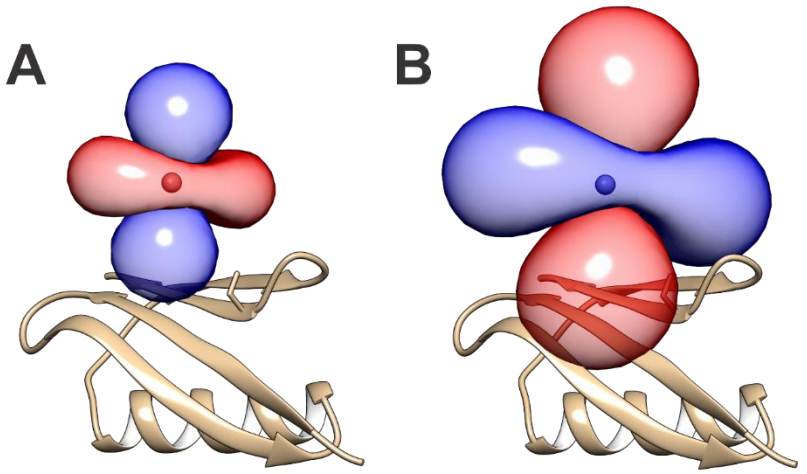


Figure 4.21: Isosurfaces (1 ppm) of [Yb(T3)] (A) and [Dy(T3)] (B)

It has been shown that some LBTs can be immobilized on the protein surface by an electrostatic interaction between the lanthanide cage and a carboxylate belonging to an ASP or GLU residue[87,98]. This interaction may decrease the tag's mobility and therefore increase the effective tensor anisotropies. In the case of GB1 T53C, there are two carboxylates within 9 and 11 Å from the cysteine (E42, E56); we therefore expected to see such stabilizing interaction take place between our tag and our protein. However, we noted that the axial anisotropies of the two tensors are a factor 4-5 smaller than the values expected for the anisotropies of the magnetic susceptibility tensors of the Yb^{3+} and Dy^{3+} ions usually calculated from rigid systems containing these paramagnetic ions. This likely indicates that some motional averaging of the magnetic susceptibility tensors occurred, so that the best fit tensors result from the motional averaging of the magnetic susceptibility anisotropy tensors.

Interestingly, the position of the metal calculated with FANTEN is relatively far from the protein surface. Those observations could actually indicate that the tag is not interacting with the protein, thus being very mobile, in

agreement with the observed small tensor anisotropies. This indicates that in **T3**, 9-coordinated lanthanide ions have no possibility to interact with the carboxylate placed on the protein's surface.

4.4 Characterization via FFC Relaxometry

4.4.1 NMRD Profiles

In order to characterize the coordination of the lanthanides in T3 and T4, FFC relaxometry experiments were performed. The ^{1}H nuclear magnetic relaxation dispersion (NMRD) profiles of the [T3(Gd)] and [T4(Gd)] complexes in water solution at 10, 25 and 37 °C are shown in figure 4.22. The dispersions are somewhat smoother than predicted by a Lorentzian spectral density function; no sizable contributions from outer-sphere relaxation is anyway expected, because when the profiles were fitted using one regularly coordinated first-sphere water molecule and diffusion water molecules, a good fit could only be achieved for distances of closest approach larger than 8 Å. The profiles were thus fit by considering one regularly coordinated water molecule (with metal-proton distances of 3.05 Å) and a second water molecule at a distance r_2 from the gadolinium ion free to be adjusted in the best fit analysis.

The best fit parameters are reported in Table 3 and the corresponding profiles in Figure 4.22. Due to the low sensitivity to the correlation time for electron relaxation, the values of τ_v were kept fixed to values typically observed in gadolinium complexes. Although there exists some covariance among the different parameters, the analysis indicates that the exchange rates of the regularly coordinated water molecule decrease on passing from [Gd(T4)] to [Gd(T3)], up to a lifetime $\tau_{M(1)}$ of several microseconds (ca. 4 μs at 25 °C). The analysis also shows contributions from fast exchanging second-sphere water molecules, with lifetime in the picosecond timescale, thus representing the

correlation time which modulates the dipole-dipole interaction energy between the magnetic moment of the unpaired electrons of Gd^{3+} and the magnetic moments of these water protons. The number of these second sphere water molecules is totally covariant with the metal-proton distance r_2; if one water molecule is considered, r_2 is about 3.4-3.5 Å.

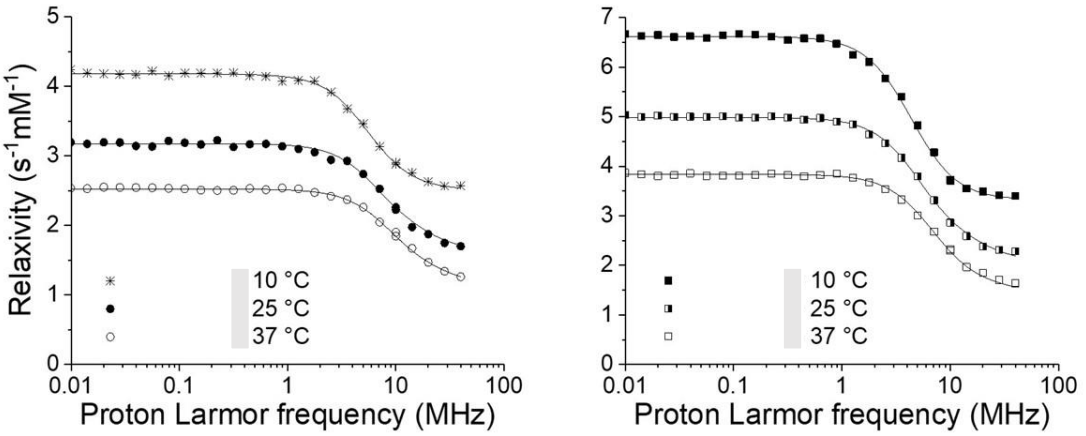


Figure 4.22: NMRD Profiles of **T3(Gd)** and **T4(Gd)** measured at different temperatures

	[Gd(T4)]			[Gd(T3)]		
	37 °C	25 °C	10 °C	37 °C	25 °C	10 °C
τ_R (ps)	40	61	110	35	57	112
ω_t (cm^{-1})	0.016			0.023		
τ_v (ps)	15	20	25	15	20	25
$\tau_{M(1)}$ (µs)	0.88	1.4	2.6	3.4	3.8	4.6
r_2 (Å)	3.37			3.48		
$\tau_{M(2)}$ (ps)	5.2	16	79	8.8	18	51

Table 3. Parameters for the fitting of NMRD profiles for [Gd(T3)] and [Gd(T4)]

4.4.2. Analysis of data for T3 and correlation with NMR

The fast exchanging second-sphere water protons may be related to water molecule(s) hydrogen bonded in position(s) allowing for a large mobility. The presence of two protons, exchanging in the microsecond time scale, at 3.05 Å, i.e. at the distance expected for the protons of a water molecule regularly coordinated to the Gd^{3+} ion, or at a somewhat larger distance (as 3.3 Å) could be related to one water molecule hydrogen-bonded to the oxygen atoms of hydroxyl groups or of the carboxylate group. Its lifetime is much longer (several microseconds) than the lifetime of the water molecule coordinated to the Gd^{3+} ion in DO3A-like or DOTA-like complexes (in the submicrosecond timescale[92]) likely because of the bipositive charge of this complex which disfavours water exchange with respect to neutral or negative complexes. A long lifetime of this water molecule indicates a relatively “compact” and stable environment surrounding the Gd^{3+} ion, which may prevent the replacement of the coordinated water by negatively charged groups present on the protein surface.

Alternatively, these slow exchanging protons may be the hydrogen atoms of the three hydroxyl groups coordinated to the Gd^{3+} ion. In this case, no visible contribution from a first-sphere water would be observed. This would indicate that no water is directly coordinated to the Gd^{3+} ion and therefore that no potential interaction with a residue of a protein is to be expected.

Both those hypotheses are in agreement with the observed tag mobility that affects the magnitude of the PCS-determined tensor. Indeed, FFC NMR indicates two scenarios: or lanthanides chelated by T3 do not present first-sphere water coordinated and therefore it is quite likely that the structure of T3 will impede an interaction with a carboxylate on the protein’s surface. Or, there is a water present in the first sphere of coordination of Gd^{3+} but it is so tightly bound that it is unlikely that a carboxylate on the protein’s surface would be able to displace it.

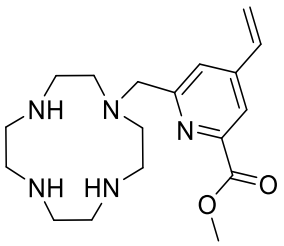
4.5 Conclusions

In this chapter, we described the synthesis of 2 enantiopures ligands as potential tags for NMR spectroscopy, **T3** and **T4**. Their complexes with several lanthanides, stable at neutral pH, were described and characterized. A protocol for the paramagnetic tagging of proteins based on the photo-catalyzed thiol-ene reaction was developed for the first time. Although the reactivity of the vinyl-pyridine substituent in **T4** proved incompatible with this reaction, the conjugation of **T3** to GB1 T53C proceeded in high-yield. The paramagnetic conjugates were analyzed via NMR Spectroscopy and a single set of peaks was observed. PCS were recorded and fitted using the software FANTEN, giving tensors of moderate anisotropy but of high consistency. The coordination environment of lanthanides chelated in **T3** and **T4** were analyzed by means of FFC Relaxometry, and for **T3** the results obtained were correlated to the values and the positions of tensors obtained by NMR. This shows that NRMD profiles analysis is useful for the interpretation of NMR data and the study of paramagnetic tags. As for the thiol-ene reaction, our experiments show that this reaction is a powerful tool for the paramagnetic tagging of proteins; even though vinyl-pyridine derivatives' reactivity can be troublesome, other linkers for paramagnetic tags could be imagined in order to circumvent those limitations.

4.6 Experimental Procedures

4.6.1 Synthesis

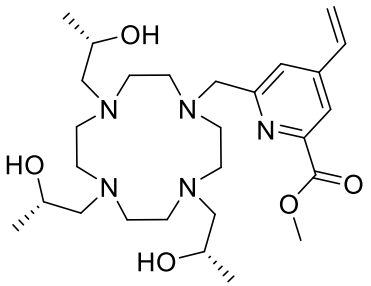
methyl 6-((1,4,7,10-tetraazacyclododecan-1-yl)methyl)-4-vinylpicolinate **43**



34 (100 mg, 0.47 mmol) was dissolved in DCM (10 mL) and added to a solution of cyclen (815 mg, 10 eq) in 30 mL DCM. The mixture was stirred at room temperature for 3 days. The reaction mixture is then washed with 30mL 1M NaOH and 2x30 mL H_2O. The organic layer is dried over MgSO4 and solvent is removed. Obtention of a yellow oil (151 mg, 91%)

MS: Predicted: 348.24 Th, Found: 348.50

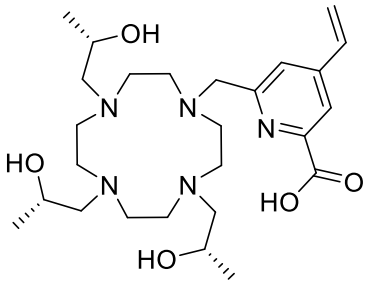
methyl 6-((1,4,7,10-tetraazacyclododecan-1-yl)methyl)-4-vinylpicolinate **44**



43 (170 mg, 0.49 mmol) was dissolved in MeOH (4 mL). (S)-Propylene Oxide (420 µL) was added and the mixture was stirred at room temperature for 4 days. Obtention of a yellow oil (220 mg, quant.)

MS: Predicted: 522.36 Th, Found: 522.62

methyl 6-((1,4,7,10-tetraazacyclododecan-1-yl)methyl)-4-vinylpicolinate **T3**



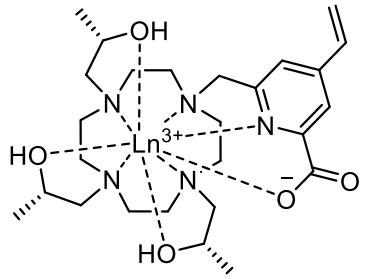
44 (220 mg, 0.4 mmol) was dissolved in 2.8 mL THF. To this mixture was added a solution of 67.2 mg $LiOH.H_2O$ in 2.8 mL H_2O. After 6 hours at room temperature the reaction was over. The reaction mixture was purified via semi-preparative HPLC.

^{1}H NMR (500MHz, D_2O): δ 8.36 (s, 1H), 8.03 (s, 1H), 6.85 (dd, 18 Hz, 11 Hz, 1H), 6.38 (d, 18 Hz, 1H), 5.87 (d, 11 Hz, 1H), 4.23 (d, 14 Hz, 1H), 4.067 (m, 2H), 3.88 (m, 2H), 3.73 (d, 14 Hz, 1H), 3.57 (bm, 3H), 2.59-3.47 (m, 18H), 1.26 (bs, 1.5 H), 1.06 (d, 6.3 Hz, 3H), 1.03 (d, 6.3 Hz, 3H), 0.73 (bs, 1.5 H)

^{13}C NMR (125MHz, D_2O): δ 155.15, 150.67, 132.26, 128.54, 127.51, 123.10, 63.04, 60.72, 60.40, 59.80, 59.73, 59.36, 59.16, 58.88, 54.57, 50.69, 50.24, 49.43, 48.55, 47.18, 46.75, 45.32, 20.53, 19.98, 19.43, 19.22

MS: Predicted: 508.35 Th, Found: 508.69

Synthesis of [T3(Ln)]



T3 (70 mg) and $LnCl_3$ nH_2O (100 mg) were dissolved in 6 mL H_2O/ MeCN (50/50). The pH was adjusted to 7 and the mixture was refluxed. Chelation was typically quantitative overnight, as shown by LC/MS analysis. Reaction mixtures were purified via semi-preparative HPLC.

[Lu(T3)]:

^{1}H NMR (500MHz, D_2O): δ 7.99 (s, 1H), 7.69 (s, 1H), 6.82 (dd, 18 Hz, 11 Hz, 1H), 6.22 (d, 18 Hz, 1H), 5.71 (d, 11 Hz, 1H), 4.30 (d, 14 Hz, 2H), 3.78 (bs, 1H), 3.30-3.67 (m, 8H), 2.96-3.16 (m, 4H), 2.07-2.71 (m, 12H), 1.37 (d, 6.23 Hz, 3H), 1.17 (d, 5.89 Hz, 3H), 0.80 (d, 5.91 Hz, 3H).

MS: Predicted (100%): 340.64 Th, Found 340.97

[Yb(T3)]:

^{1}H NMR (500MHz, D_2O): δ 30.87, 29.27, 24.14, 21.91, 21.50, 19.7, 13.91, 12.91, 12.49, 11.24, 10.46, 9.69, 8.23, 5.61, 5.10, 3.43, 1.29, -2.64, -3.88, -6.27, -6.73, -9.38, -19.74, -20.36, -22.58, -25.38

MS: Predicted (100%): 340.14 Th, Found 340.32

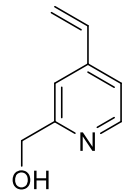
[Dy(T3)]:

MS: Predicted (100%): 335.48 Th, Found 335.13

[Gd(T3)]:

MS: Predicted (100%): 331.63 Th, Found 331.56

methyl 6-(hydroxymethyl)-4-vinylpicolinate **45**

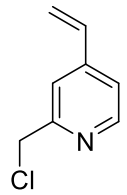


4-bromopyridine methanol (0.50 g., 2.66 mmol), K_3PO_4 (1.13g., 5.32 mmol) and trivinylboroxin-pyridin complex (0.70g., 2.91 mmol) were suspended in dioxane (18 mL). N_2 was bubbled through the solution for 5 mns and JohnPhos (6%, 47 mg) and $Pd(dba)_2$ (0) (3%, 46 mg) were added, still under N_2. The reaction mixture was heated at 90°C for 3 hours. Dioxane was removed *in vacuo* and the crude mixture was purified via flash chromatography with a mixture of DCM and MeOH (95/5) Obtention of 245 mg (68%) of the pure product.

^{1}H NMR (500MHz, $CDCl_3$): δ 8.40 (d, J= 4.7 Hz, 1H), 7.30 (s, 1H), 7.13 (d, 4.7 Hz, 1H), 6.62 (dd, 18Hz, 11Hz, 1H), 5.94 (d, 18 Hz, 1H), 5.45 (d, 11Hz, 1H), 4.73 (s, 1H)

^{13}C NMR (125 MHz, $CDCl_3$): δ 160.44, 148.77, 145.70, 134.68, 119. 42, 119.01, 117. 81, 64.42

methyl 6-(chloromethyl)-4-vinylpicolinate **46**

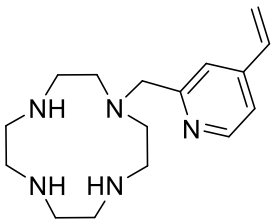


45 (240 mg, 1.78 mmol) was dissolved in DCM (24 mL) and DIPEA (1.2 mL, 7.1 mmol) then MsCl (415 uL, 5.3 mmol) were added at 0°C. Reaction mixture was stirred for 3h at room temperature. DCM was removed and the crude oil was purified via flash chromatography with a gradient of Dichloromethane/MeOH 100/0 to 95/5. Obtention of the pure product as a darkoil (240 mg, 88%)

^{1}H NMR (500MHz, $CDCl_3$): δ 8.54 (d, J= 5.29 Hz, 1H), 7.48 (s, 1H), 7.25 (d, 5.29 Hz, 1H), 6.71 (dd, 17Hz, 11Hz, 1H), 6.03 (d, 18 Hz, 1H), 5.55 (d, 11Hz, 1H), 4.71 (s, 2H)

^{13}C 156.93, 149.74, 145.93, 134.40, 120.08 119.86 119.24, 46.68

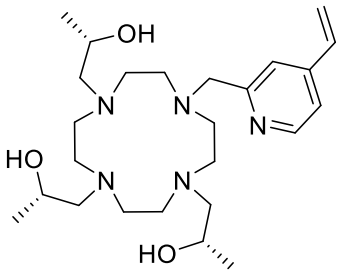
1-((4-vinylpyridin-2-yl)methyl)-1,4,7,10-tetraazacyclododecane **47**



46 (240 mg, 1.56 mmol) was dissolved in DCM (50mL) and the mixture was added to a solution of cyclen (2.7 g., 50 mL). Reaction mixture was stirred for 3 days at room temperature. The mixture was then washed with 100 mL 1M NaOH and 2x100 mL H_2O. The organic layered was dried on $MgSO_4$ and the solvent was removed to yield 200 mg of pure product (45%).

MS: m/z=290.21 (100%)

(2S,2'S,2''S)-1,1',1''-(10-((4-vinylpyridin-2-yl)methyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triyl)tris(propan-2-ol) **T4**

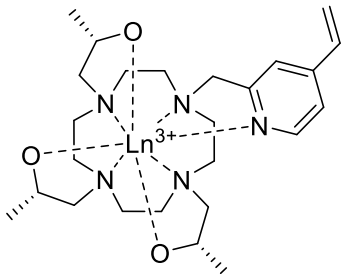


47 (200 mg, 0.69 mmol) was dissolved in MeOH (6 mL) and 800 µL of (S)-Propylene Oxide were added. After 3 days, LC-MS revealed the presence of only the desired product. Methanol is removed in vacuo. Obtention of 319 mg of product that was used without purification (quant.).

^{1}H NMR (400MHz, D_2O): δ 8.26 (d, J=5.38 Hz, 1H), 7.37 (s, 1H), 7.27 (d, J=5.38 Hz, 1H), 6.60 (dd, J=17Hz, J=11Hz, 1H), 5.95 (d, J=17Hz, 1H), 5.42 (d, J=11Hz, 1H), 3.96-3.80 (m, 4H) 3.74 (d, 13Hz, 1H), 3.41 (d, 13Hz, 1H),

MS: m/z= 464.55 Th (100%)

Synthesis of [T4(Ln)]



T4 (70 mg) and $LnCl_3$ *n*H_2O (100 mg) were dissolved in 6 mL MeOH, and the mixture was refluxed. Chelation was typically quantitative overnight, as shown by LC/MS analysis. Reaction mixtures were purified via semi-preparative HPLC (H_2O+0.1% TFA/MeCN gradient).

[La(T4)]:

^{1}H NMR (500MHz, D_2O): δ 8.58 (d, J=5.33 Hz, 1H), 7.55 (s, 1H), 7.51 (bd, J=6.21 Hz, 1H), 6.74 (dd, J=17.21, 10.20 Hz, 1H), 6.15 (d, J=17.21 Hz, 1H), 5.64 (d, J=10.21 Hz, 1H), 3.76-3.49 (m, 9H), 3.14-2.96 (m, 4H), 2.62-2.22 (m, 13H), 1.28 (d, J=5.67 Hz, 3H), 1.20 (d, J=5.67 Hz, 3H), 1.17 (d, J=5.67 Hz, 3H)

MS: 714.29 (100%), 357.63 (49%)

[Ce(T4)]:

MS: 358.33 (100%), 715.55 (34%)

[Dy(T4)]:

MS: 370.03 (100%), 369.04 (89%), 738.12 (76%), 737.13 (74%)

Conjugation of T3 to GB1 T53C

Prior to conjugation, GB1 T53C was buffer-exchanged into NaPi 20 mM, pH 7.5, then concentrated to 180 µM. 5 equivalents each of DPAP and [**Ln(T3)**] were added to the protein, as well as 10% D_2O. The mixture was transferred into a 5-mm NMR tube and argon was gently bubbled through the solution for 5 minutes. The tube was placed under a UV Lamp (UVGL-55 Mineralight 26W) at 365 nm for 1h, after which the sample was measured using NMR spectroscopy and the spectra analyzed. Excess of small molecules were then washed away by buffer exchange and NMR spectra were recorded again.

4.6.2 NMR, FFC, MALDI

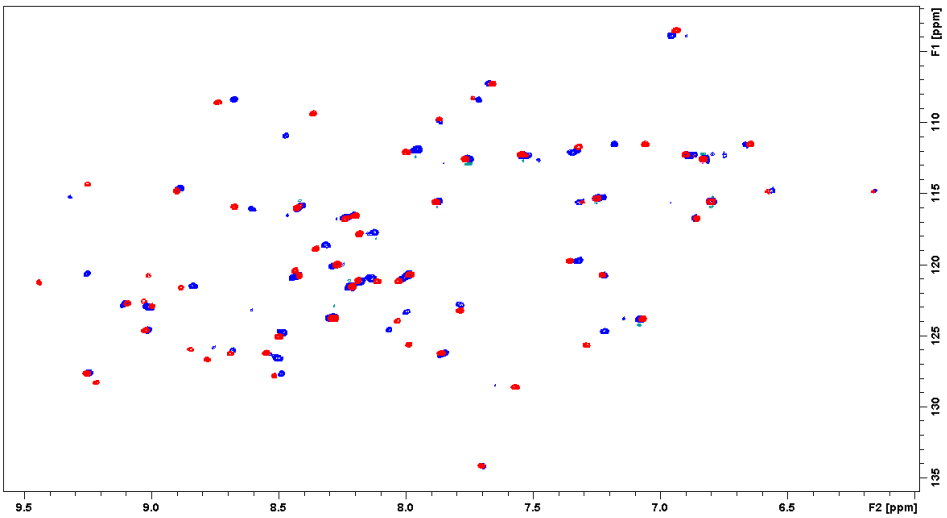


Figure 4.23: 2D 1H-15N HSQC of GB1 T53C (red) and GB1 T53C tagged with VDPA (blue)

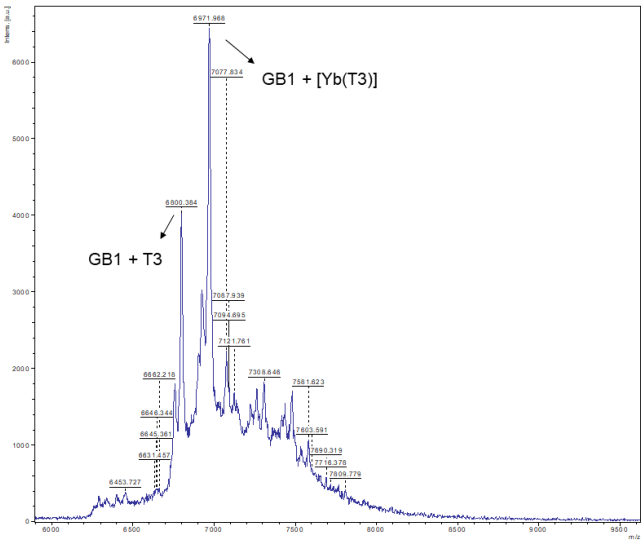


Figure 4.24: MALDI analysis of GB1 T53C tagged with **T3**

	[Gd(T4)]			[Gd(T3)]		
	37 °C	25 °C	10 °C	37 °C	25 °C	10 °C
τ_R (ps)	40	61	110	35	57	112
τ_t (cm^{-1})	0.016			0.023		
τ_v (ps)	15	20	25	15	20	25
$\tau_{M(1)}$ (μs)	0.88	1.4	2.6	3.4	3.8	4.6
r_2 (Å)	3.37			3.48		
$\tau_{M(2)}$ (ps)	5.2	16	79	8.8	18	51

Table 3. Parameters for the fitting of NMRD profiles for [Gd(T3)] and [Gd(T4)]

Chapter 5: Application of paramagnetic tagging to the study in solution of *tc*Pex14

Abstract

*tc*Pex14 is a small (8 kDa), medicinally relevant protein involved the peroxisomal-import pathway in *Trypanosoma* cruzi, a parasite responsible for Chagas disease. In order to better understand its behavior in solution, we have tagged *tc*Pex14 with DOTA-M8 loaded with Thulium (Tm, paramagnetic) and Lutetium (Lu, diamagnetic). Important differences in the 2D ^{1}H-^{15}N HSQC between the paramagnetic species and the diamagnetic reference made it difficult to transfer the assignment from the diamagnetic spectra to the paramagnetic ones. In order to diminish the extent of the PCS and therefore ease assignment-transfer we have attached vinyl dipicolinic acid (VDPA) to *tc*Pex14's surface, and used it to chelate Lutetium, Ytterbium and Thulium. PCS and RDC were recorded, thus allowing for the structure refinement of the protein. In this chapter, we will discuss the influence of the tag and of the paramagnetic lanthanide ion on the intensity of PCS and RDC, and we show that tagging a protein with smaller and more mobile tags allows to tune the paramagnetism of the system and to adjust it according to the size of the studied protein, in order to optimize the measurement of its effects. Last, implications of those findings in paramagnetism-assisted structure-based drug discovery are discussed.

5.1 Introduction

5.1.1 Trypanosimiasis

Human African trypanosomiasis (HAT, sleeping sickness) and Chagas disease are caused by the protist parasites Trypanosoma brucei, and Trypanosoma cruzi, respectively. Chagas disease is most prevalent in South America but is spreading internationally due to increased migration, and potential climate change–induced vector spreading[99,100]. The number of drugs already available for trypanosomiases is low. Existing medications (Suramin, Pentamidine, Melarsoprol, Benzidazole, and Nifurtimox) have serious side effects, require long treatment schedules, and often fail to eliminate parasitemia. Therefore, new therapeutic opportunities are needed.

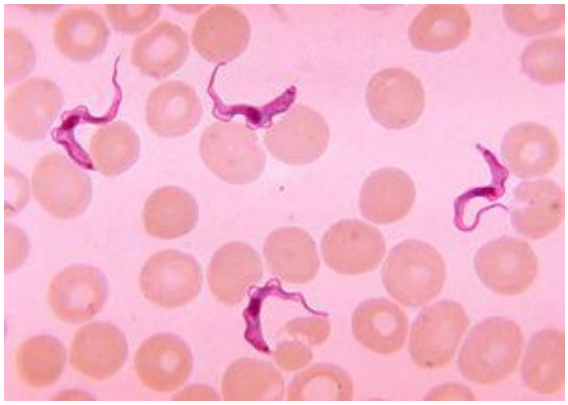


Figure 5.1: Blood smear from patient with African Trypanosomiasis

5.1.2 Glycosomes and Protein import

Peroxisomes are organelles present in all eukaryotic cells and facilitate enzymatic reactions that require oxidizing conditions, such as involving catalases or D-amino acid oxidases[101]. Although not all peroxisomal proteins (Pex, peroxins) that are involved in peroxisomal matrix import are conserved across species, the main pathways of peroxisomal import share some common features (Fig. 5.2). The import of peroxisomal matrix proteins is based on some unique aspects, making it different from other import-pathways[102]. It has been shown that the nature of the pore is highly dynamic, thereby rationalizing how a wide size range of cargo can be translocated, including folded proteins and oligomers but even particles of 9 nm diameter[103].

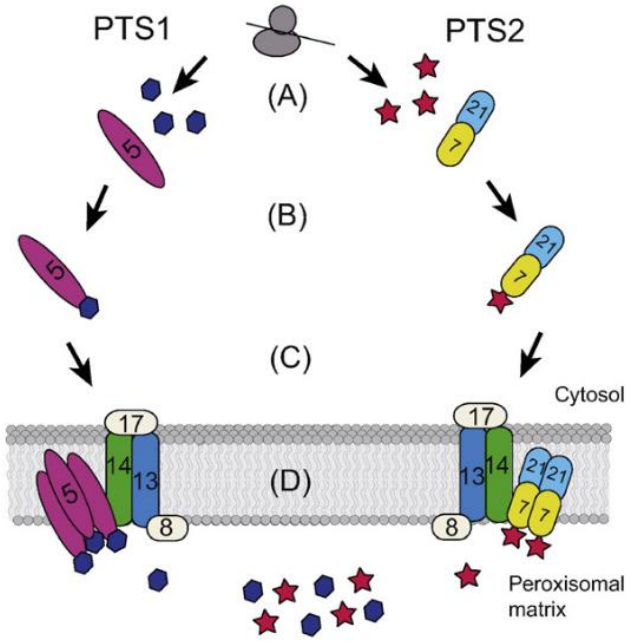


Figure 5.2: Mechanism of import in Peroxisomes. PTS stands for Peroxisomal Targeting Signal. Reproduced from Emanouilidis *et al.*[102]

Glycosomes are peroxisome-related organelles containing enzymes required for glucose metabolism and parts of other intermediary metabolic pathways. They are essential and characteristic to a few protists, including Trypanosoma. Their mechanism of cargo-mediated import is the same as for the other peroxisomes. It has been shown that inhibiting cargo-mediated import in glycosomes in Trypanosoma was an efficient way to kill the parasite[104].

The vast majority of peroxisomal matrix proteins are targeted based on peroxisomal targeting signals (PTS1, PTS2) located at either the C-terminus (PTS1) or the N-terminus (PTS2) of the proteins. PTS1 and PTS2 cargo proteins reach their partners via different pathways, although basic steps and some of the proteins involved are shared between these pathways (c.f. fig. 5.2). Peroxisomal cargo proteins carrying the C-terminal PTS1 signal peptide are most abundant. PTS1 cargo is transported by the Pex5 protein[105], which is partially disordered and acts both as a chaperone and a receptor for PTS1 cargo. At the membrane, the cargo-loaded Pex5 complex interacts with the membrane-associated Pex14 and Pex13 proteins which form the docking complex. In this complex the interaction of Pex5 involves short WxxxF sequence motifs in its N-terminal region (Fig. 5.3), which bind to a small, globular domain in the N-terminal region of membrane-bound Pex14[106]. Pex14 and Pex13 play distinct roles in the import process. It has been suggested that the presence of cargo proteins influences the membrane interaction of Pex5. Pex14 preferentially binds to cargo-loaded Pex5 receptor, while Pex13 has higher affinity to the free receptor. Functional assays indicate that only Pex14 and Pex5 are implicated in the gating of the peroxisomal pore. This suggests that Pex14 and Pex5 together are responsible for pore formation, while Pex13 may be involved in later steps, such as in coupling of translocation to receptor recycling (Fig. 5.2). The Pex17 protein is an additional factor associated with the Pex13/Pex14 docking complex in yeast, although its specific role is not well understood.

5.1.3 The Pex14/Pex5 Interface

The Pex14/Pe5 interaction is crucial for glycosome biogenesis and glycosomal protein targeting in Trypanosoma. From a structural point of view, this interaction takes place between the N-terminal domain of membrane-bound Pex14, soluble and globular, and the N-terminal domain of Pex5, disordered and rich of a WxxxF motif; it is this di-aromatic WxxxF motif which mediates the interaction with Pex14[107].

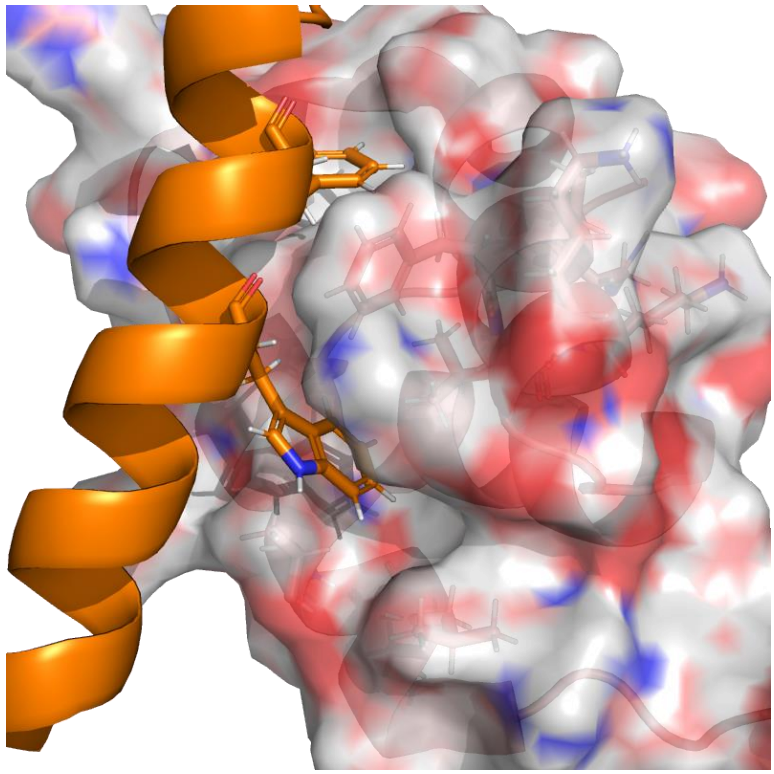


Figure 5.3: The Pex14/Pex5 interface
Pex5 (orange) anchors onto the surface of Pex14 with aromatic residues (PDB 2W84[107])

It has been shown via Pex14-RNA interference (RNAi) studies that glucose becomes toxic to *T. brucei* when glycosomal import is disrupted, thus indicating that this complex is a potential drug target. More, it has been recently shown that the small-molecule-mediated disruption of the Pex14/Pex5 leads to accumulation of glycosomal enzymes in the cytosol, adenosine triphosphate diminution, glucose toxicity, and metabolic disruption resulting in *T. brucei* parasite death.

Small molecules inhibitors have recently attracted much attention for the inhibition of Pex14 in Trypanosoma. Indeed, Dawidoski *et al*.[108] showed that small molecules were able to mimic the WxxxF motif of Pex5 and to inhibit the formation of the Pex14/Pex5 complex.

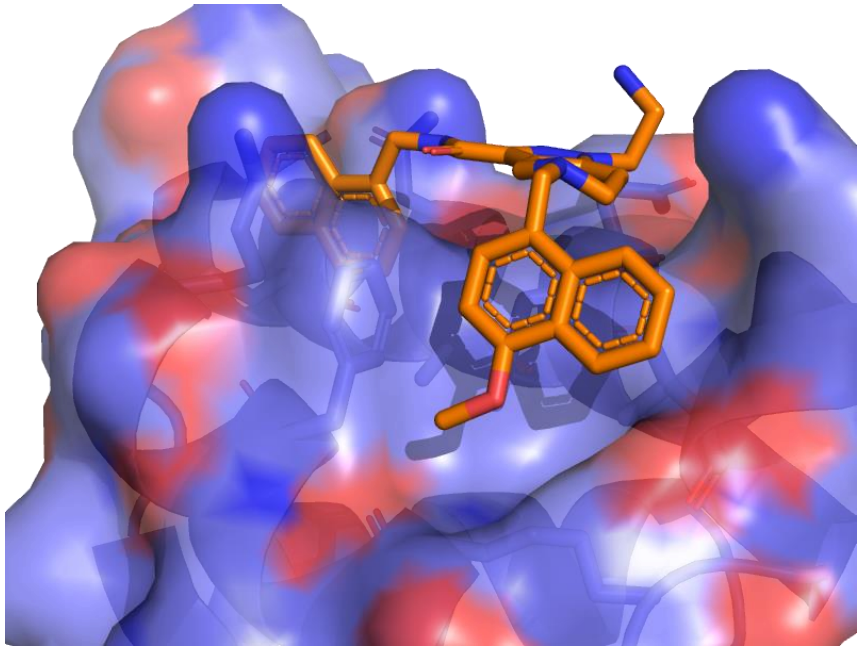


Figure 5.4: Pex14 in complex with an inhibitor, PDB 5N8V[108]

Several of those molecules have been shown to efficiently display trypanosomicidal activity. Furthermore, those molecules were shown to be “on-target”, meaning that the death of Trypanosoma induced by those compounds are related to inhibition of the Pex14/Pex5 interaction[108].

Peroxisomes proteins and particularly Pex14 and Pex5 are also present in humans and have similar role and structures. Therefore, the issue of selectivity must be raised. Structures of human and Trypanosoma Pex14 were compared: the overall fold are very similar, but Pex14 in Trypanosoma exhibits an additional C-terminal helix. The two hydrophobic pockets and the two phenylalanine residues in the binding surface are conserved in *T. brucei*. However, characteristic amino acid differences are observed in the PEX5 binding pockets of Trypanosoma PEX14 (Arg28, Asn31, Glu34, and Asp38) compared with its human parent (Leu28, Thr31, Lys34, and Asn38). This indicated that inhibitors can be designed to selectively target Pex14 in Trypanosoma using specific polar interactions. Those structural considerations were confirmed by trypanosomicidal activity tests, which showed no correlation between mammalian cell toxicity and HsPex14 affinity.

The molecules developed by Dawidowski et al. are potent and selective inhibitors[108]. However, being quite lipophilic, they suffer from high plasma-binding values, making them poor clinical candidates. Their mode-of-action is however potent and selective towards Trypanosoma, making this case an important proof-of-concept for the future development of clinical candidates inhibitors of Pex14. In this chapter, we will discuss how the paramagnetic tagging of *tc*Pex14 can help characterize the binding-mode of novel small ligands in order to optimize them.

5.2 Paramagnetic tagging of *tc*Pex14 with DOTA-M8

5.2.1. Structural considerations on *tc*Pex14

In this chapter we will focus on the ortholog of Pex14 of *Trypanosoma cruzi*, *tc*Pex14. In order to gain deeper structural knowledge of the protein and with the idea of conducting Paramagnetic restraints-driven structure-based hit optimization, the paramagnetic tagging of *tc*Pex14 has been conducted.

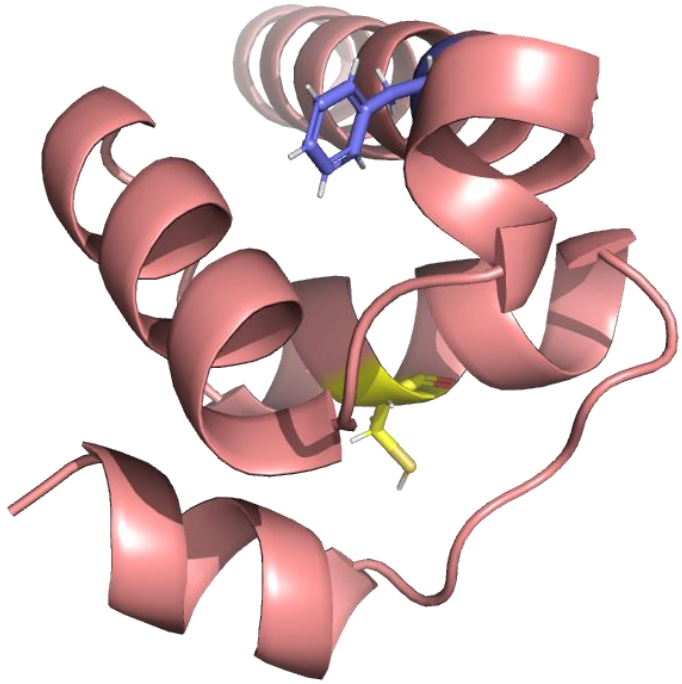


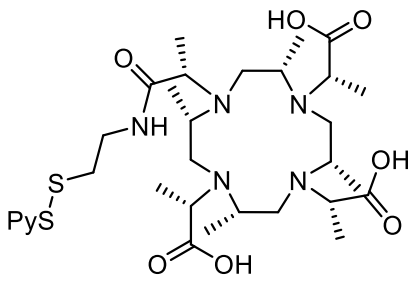
Figure 5.5: Crystal Structure of *tc*Pex14
Cysteine (yellow) and binding site (blue)

*tc*Pex14 naturally bears one cysteine, which would allow for the tagging of this residue only. This tagging-site is on the other side of the

protein compared to the binding site, and therefore no major change in the protein's structure is expected (c.f. Fig. 5.5). However, due to the protein's small dimensions, it is still quite near (ca. 15 Å), and therefore it can be expected that paramagnetic effects could, upon tagging, be easily transferred on the ligand's NMR signal. The combination of those characteristics makes *tc*Pex14 an ideal scenario for the use of paramagnetic NMR.

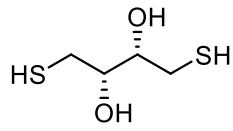
5.2.2. Tagging of *tc*Pex14 with DOTA-M8

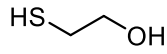
In a first time, we decided to tag *tc*Pex14 with DOTA-M8. DOTA-M8 is meant to be attached through disulfide formation, thus reacting selectively with cysteines. Before starting the reaction, the protein must first be reduced in order to have its cysteine free for tagging; commonly BME (Beta-Mercapto-Ethanol) or DTT (DiThioThreitol) are used. Then, the reducing agent must be removed in order to protect the tag, and the reaction must be conducted soon enough to that the protein does not have the time to reform disulfide bonds with itself.



DOTA-M8

DTT





BME

Figure 5.6: Molecules discussed in this section

First trials of this reaction proved unsuccessful, as no change were observed via NMR analysis of the protein. Next, we dialyzed the protein in a buffer containing TCEP instead of BME as the reducing agent. Indeed, TCEP presents the advantage of being a stronger and irreversible reducing agent. The reaction then proceeded as expected and tagging was complete in 4 hours.

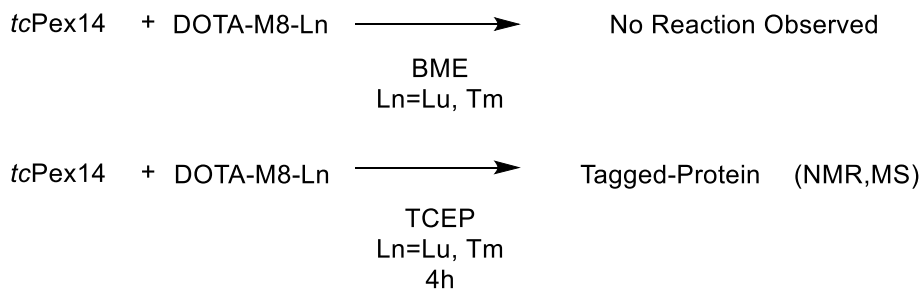


Figure 5.7: Conditions of tagging of *tc*Pex14 with DOTA-M8

*tc*Pex14 was tagged with DOTA-M8 containing Thulium (Tm, paramagnetic) and Lutetium (Lu, diamagnetic). Changes in the 2D ^{1}H-^{15}N HSQC spectra of the protein after reaction with the tag were clear and after 4 hours no residual trace of the untagged protein was observed. Tagging with [Lu(DOTA-M8)] induced changes in the chemical shifts only for the residues neighboring the tagging site. On the other hand, tagging with [Tm(DOTA-M8)] yielded important changes on almost of all of the residues of the proteins, attributed to the paramagnetism of Tm. Most of the peaks underwent PCS, and the signals of the residues close to the tagging site decreased in intensity or disappeared because of PRE. The tagging of the protein was confirmed by ESI-quadrupole MS.

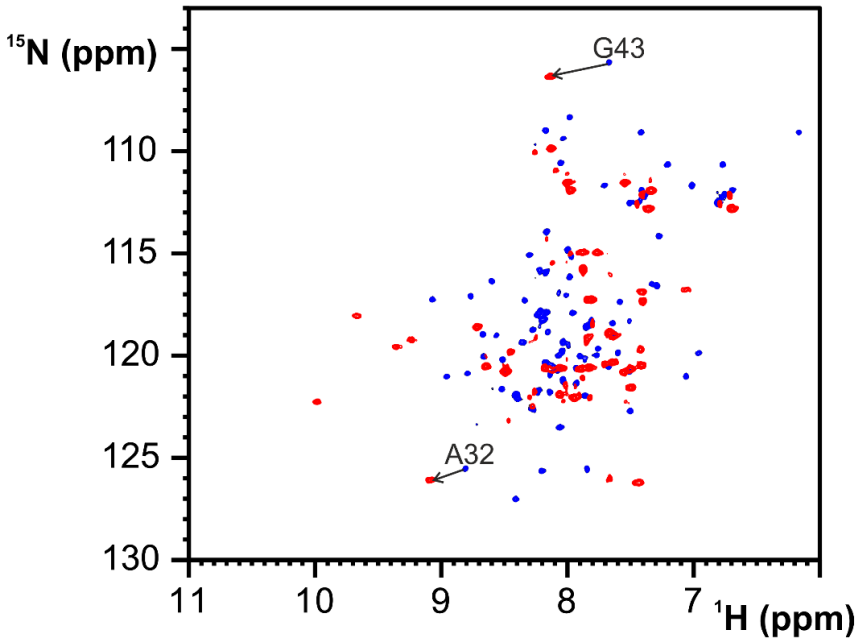


Figure 5.8: 2D ^{1}H-^{15}N HSQC of *tc*Pex14 tagged with [Lu(DOTA-M8)] (blue) and with [Tm(DOTA-M8)] (red)

5.2.3. NMR Analysis and temperature titration

Upon tagging with [Tm(DOTA-M8)], very large PCS were recorded. Because the protein is small (8 kDa) all the residues are considered near the tag and therefore they all undergo large PCS. Since ^{15}N and ^{1}H, directly bond to each other, are at almost the same distance from the metal center, they are equally affected by PCS and paramagnetism-induced shifts in the 2D ^{1}H-^{15}N HSQC spectra take place along parallel lines that pass through the peak of the diamagnetic species. This is useful to transfer the assignment from the diamagnetic species to the paramagnetic spectrum. However, in this case, this assignment transfer was not so trivial for two reasons: the shifts recorded were very large and were found not to always follow those diagonals, making ambiguous the assignment transfer.

In order to make easier the assignment transfer, we decided to perform a temperature titration. Indeed, the extent of the anisotropic magnetic susceptibility is inversely proportional to the temperature, as shown in eq. (7) (cit). Therefore, by gradually increasing the temperature we will decrease the values of the measured PCS. The cross-peaks of each residue will then follow diagonals “pointing” towards the position of the corresponding diamagnetic residue, thus helping with the assignment transfer. We did observe such tendency, nevertheless we did not manage to significantly improve the quality of the assignment and did not manage to fit the PCS values together.

$$\chi_{a,a} = \frac{\mu_0 \mu_B^2 J(J+1)}{3\kappa T} g_{aa}^2 \qquad (7)$$

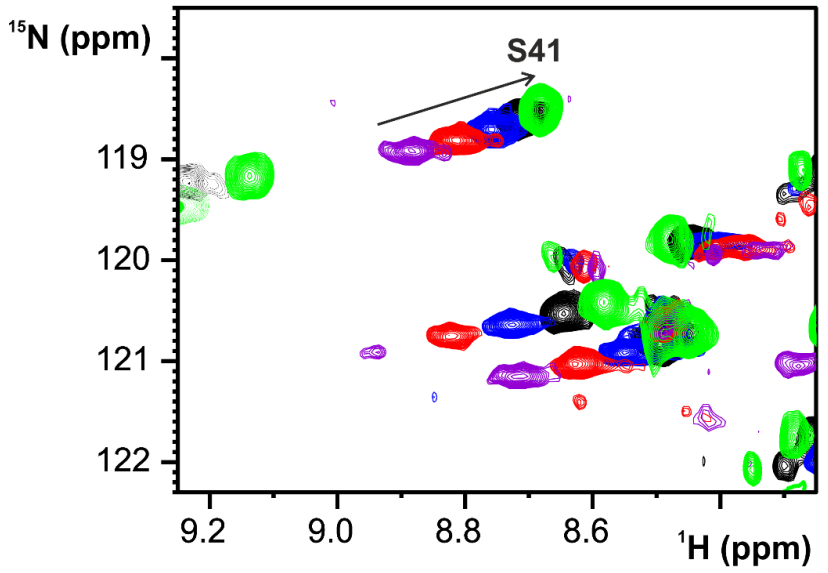


Figure 5.9: Temperature-titration of *tc*Pex14 tagged with [Tm(DOTA-M8)] 293°K (purple), 298°K (red), 303°K (blue), 308°K (black), 313°K (green)

5.3 Paramagnetic tagging of *tc*Pex14 with VDPA

5.3.1. Tagging reaction

To simplify the assignment of the paramagnetic spectra of *tc*PEX14 tagged with [Tm(DOTA-M8)] without the use of expensive ^{13}C isotopically enriched samples and acquisition of time-consuming classical assignment experiments, we decided to tag *tc*Pex14 with VDPA. Indeed, being VDPA tag smaller and more mobile, we expect an averaging of the magnetic susceptibility anisotropy which will result in smaller tensors and smaller diagonal paramagnetic shift. Moreover, not having metals already chelated in the tag, by choosing different lanthanides according to their paramagnetic properties, we could tune the paramagnetic effects on our protein.

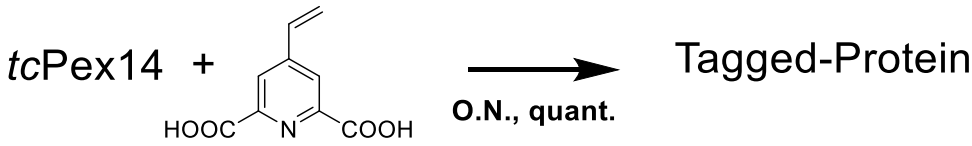


Figure 5.10: Tagging of *tc*Pex14 with VDPA

We therefore performed the tagging following the protocol of Su *et al.* and tagging was quantitative overnight. In the case of GB1, instead, tagging was only of 50% in the same conditions. Based on the hypothesis that without in situ radical generation the reaction goes through a Michael-Addition-like mechanism, this suggests that the cysteine of *tc*Pex14 is more nucleophilic than the one of GB1 T53C.

Then, titration of the tagged protein with solutions of different lanthanides ($LuCl_3$, $YbCl_3$ and $TmCl_3$) was performed. No peak broadening due to the presence in solution of free paramagnetic ions was observed. Final spectra are showed in fig. 5.11.

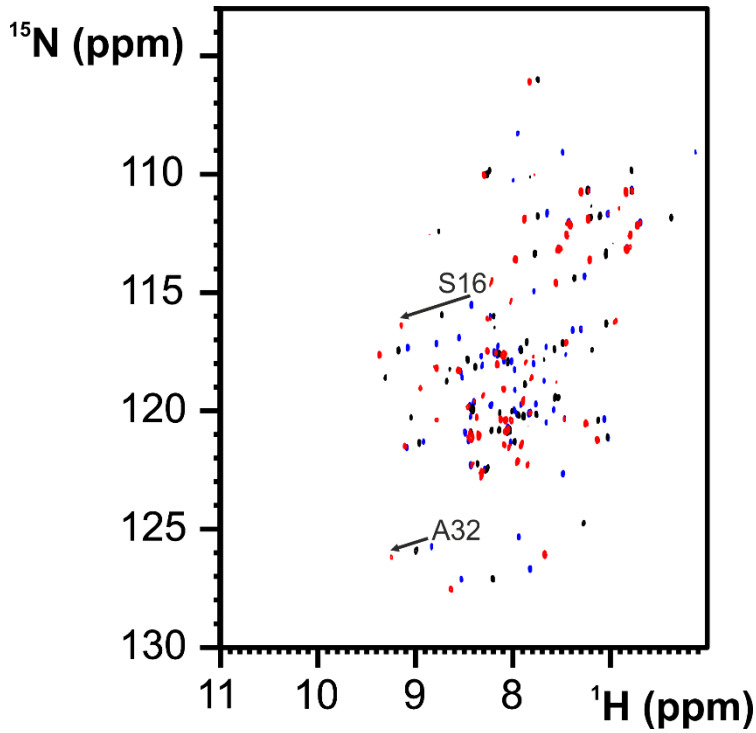


Figure 5.11: 2D ^{1}H-^{15}N HSQC of *tc*Pex14 tagged with VDPA Lu (blue), Yb (black) and Tm (red)

5.3.2. Tensor analysis

PCS on the protein's NMR peaks were recorded via 2D ^{1}H-^{15}N HSQC spectra. The shifts measured were smaller than with [Tm(DOTA-M8)], what made easier the assignment transfer. As expected, shifts were

larger when the protein was tagged with Tm than with Yb. Moreover, RDC were also recorded for all three metals. In total, we assigned a total of 21 values of PCS and 16 RDC when the protein was tagged with Yb, and a total of 16 PCS and 10 RDC when the protein was tagged with Tm. The values were analyzed and fitted to the structure of *tc*Pex14 using FANTEN, first considering the 4 sets of data separately. Results are displayed in Table 1 and the fitting in fig. 5.12.

	Q factor	**$\Delta\chi_{ax}$ (10^{-32} m^3)**	**$\Delta\chi_{rh}$ (10^{-32} m^3)**
Yb (PCS)	0.13	-5.25	3.32
Yb (RDC)	0.495	3.78	-1.30
Tm (PCS)	0.09	-9.98	2.25
Tm (RDC)	0.405	-5.21	2.73

Table 1: Tensors values for *tc*Pex14 tagged with VDPA

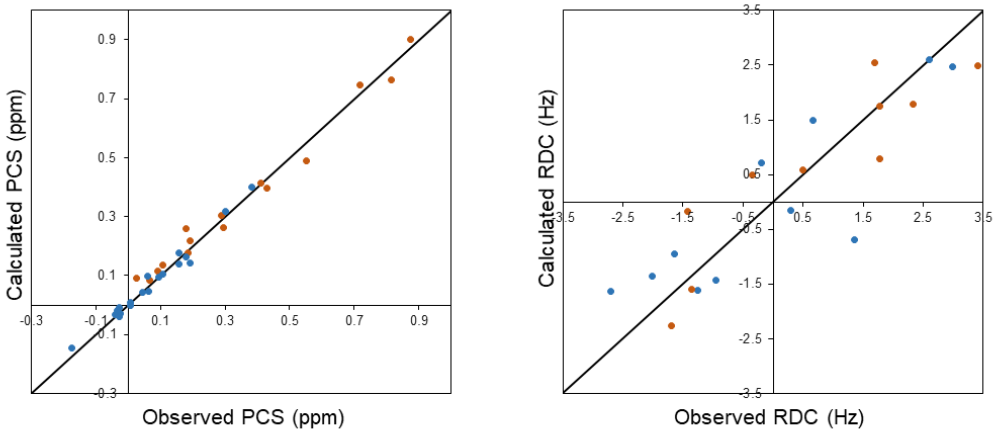


Figure 5.12: Correlation plots between experimental and calculated PCS (left) and RDC (right) for *tc*Pex14 tagged with VDPA and Yb (blue) or Tm (Orange). Calculated values were calculated based on the crystal structure. Data sets were treated independently.

As we can see in Table 1, the fitting of PCS values is good (low Q-factors) for both Yb and Tm, but Q-factors values are much higher in the case of RDC. The values of anisotropies for the 4 tensors are lower than average considering the lanthanides used, which was expected since VDPA is small and mobile. Particularly, the values of anisotropy of the RDC tensors are decreased with respect to the ones of PCS-tensors even more than by the expected factor of 0.9, which usually is used to account for the ^{N}H-N intrinsic dynamics. Next, we fitted for each metal the RDC and the PCS data sets together into one single tensor, to see if the quality of the fitting improved. RDC were scaled up by 10% to compensate for the local mobility of ^{N}H–N vectors. The results are shown in table 2 and the fitting in fig. 5.13.

	Q factor	**$\Delta\chi_{ax}$ (10^{-32} m^3)**	**$\Delta\chi_{rh}$ (10^{-32} m^3)**
PCS (Yb)	0.33	4.19	-1.66
RDC (Yb)	0.50		
PCS (Tm)	0.225	-8.92	4.20
RDC (Tm)	0.51		

Table 2: Tensors values from FANTEN with PCS and RDC fitted together

We can see from Table 2 that the quality of the fitting, when a unique tensor for PCS and RDC is considered, worsens. This is true for the fitting of RDCs and in a greater extent for PCS. This indicates a deviation between the paramagnetic data and the input crystal structure used for the fitting.

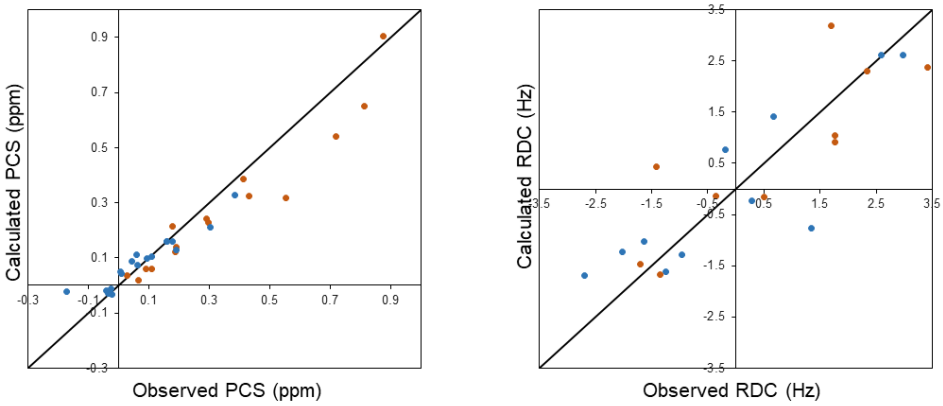


Figure 5.13: Correlation plots between experimental and calculated PCS (left) and RDC (right) for *tc*Pex14 tagged with VDPA and Yb (blue) or Tm (Orange). Calculated values were calculated based on the crystal structure. In both fitting, for each metal the PCS and RDC data sets were fitted together into one single tensor and the RDC tensor was scaled down by 0.9 to compensate for the local mobility of ^{N}H–N vectors

5.3.3. Structure refinement

PCS and RDC provide structural information because they depend on the position of the observed nuclei and on the orientation of the vectors connecting coupled nuclei, respectively. They can thus be used for assessing whether an available structure is in agreement with these data, and possibly for calculating a solution structure in better agreement. As we have seen in 5.3.2., it seems that in our case, the crystal structure of *tc*Pex14 and our paramagnetic data is not in good agreement with RDC data.

In order to estimate the extent of the structural changes needed to reproduce the paramagnetism-based restraints, the initial X-ray structural

model was first adapted to fulfill all chemical bond constraints (in terms of bond angles and lengths) of the library of the program CYANA (c.f. Fig. 5.15, RMSD between X-Ray structure and initial CYANA structure of 0.188 Å). This was done by including in the protein sequence a ‘‘pseudoprotein residue’’, composed by as many pseudoatoms as the number of atoms of the protein, each of them labeled according to its residue number and atom name. These pseudoatoms have coordinates equal to the coordinates of the corresponding atoms in the model structure and no van der Waals radius. The pseudoprotein residue was linked to the protein sequence through dummy residues, which have the function of allowing the pseudoprotein residues to freely move with respect to the protein residues. A simulated annealing calculation was performed with CYANA with upper distance limits of 0.1 Å (with weight 0.1) between all the heteroatoms of the protein and the corresponding atoms of the pseudoprotein residue. The dihedral φ and ψ angles were also restrained to vary within ±10° around the value in the model structure. A further conjugate gradient minimization was then performed with the same restraints and with the weight of the upper distance limits reduced to 0.01. In this way, the protein atoms are positioned as close as possible to the starting structure, being at the same time constrained to the bond lengths and angles defined in the internal library.

All PCS and RDC data were then introduced as restraints in the assumption that a unique tensor for each metal is responsible for all the observed PCS and RDC values, as occurring in the absence of motion. The usual local mobility of ^{N}H–N vectors was considered by using an order parameter S_{LS} of 0.9 for RDC. The weights of the restraints were of 0.5 for RDC and 100 for PCS. The higher weight used for PCSs is to compensate for their much smaller numerical absolute value. Moreover, being more robust in providing magnetic susceptibility anisotropy tensor values, it was decided to increase their weight in order to increase their influence in the tensor-optimization process. Upper distance limits between the protein heteronuclei and the pseudoatoms of the pseudoprotein residue were also

included in the same way as described for the calculations in the absence of the paramagnetic restraints, for anchoring the position of the protein atoms to the coordinates of the selected model. A simulated annealing followed by a conjugated gradient minimization were performed with PARAMAGNETIC CYANA. The output for the tensors' parameters and Q-factors is displayed in Table 3 and the fitting of the values in fig. 5.14.

	Q factor	$\Delta\chi_{ax}$ (10^{-32} m^3)	$\Delta\chi_{rh}$ (10^{-32} m^3)
PCS (Yb)	0.15	4.67	-3.08
RDC (Yb)	0.20		
PCS (Tm)	0.13	-11.7	1.92
RDC (Tm)	0.15		

Table 3: Tensors values after structure refinement in CYANA

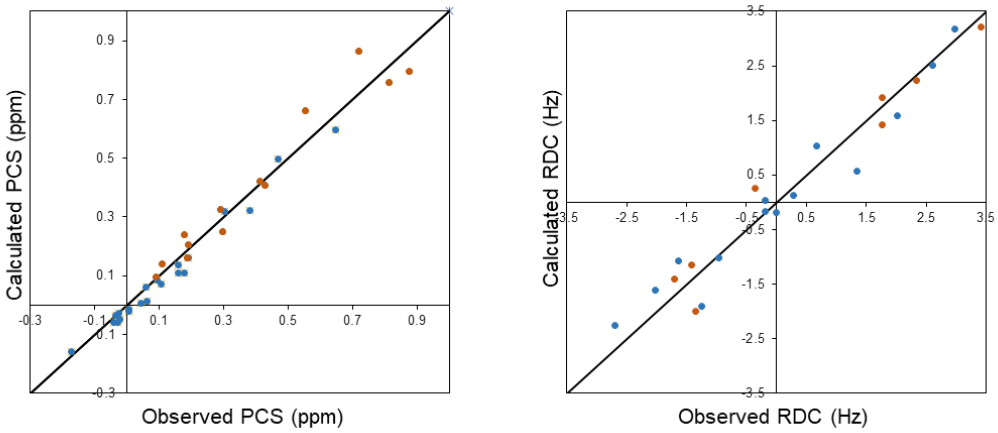


Figure 5.14: Correlation plots between experimental and calculated PCS (left) and RDC (right) for *tc*Pex14 tagged with VDPA and Yb (blue) or Tm (Orange). Fitting was performed with refined structure from CYANA. In both fitting, for each metal the PCS and RDC data sets were fitted together into one single tensor and the RDC tensor was scaled down by 0.9 to

compensate for the local mobility of ^{N}H–N vectors.

We can see that after structure refinement in CYANA the fitting of the PCS and RDC data obtained from both paramagnetic lanthanides improves, especially for RDCs (lower Q-values, c.f. Table 3 and fig. 5.14). The RMSD of the refined structure compared to the starting CYANA structure is of 0.260, and the RMSD of the refined structure compared to the original crystal structure is of 0.364. Such small RMSD values indicate that only little structural optimization was necessary to improve the fitting of the paramagnetic NMR data in solution. Most of those structural changes consist in variation of the orientation ^{N}H–N vectors.

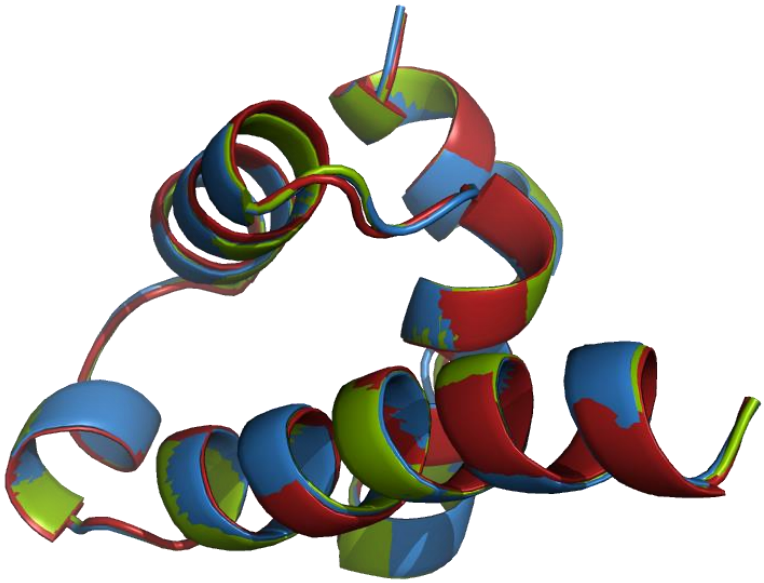


Figure 5.15: Superimposed structures of: *tc*Pex14 crystal structure (blue), CYANA structure before refinement (green) and after refinement (red)

As we can see in figure 5.15, structures before and after refinement are superimposable. This indicates that the X-ray structure and the refined solution structure are in agreement. If it were not the case, the refinement would not have been possible and Q-values would have remained high[109].

In conclusion, we have tagged our target protein with VDPA which we have used to chelate diamagnetic Lu and paramagnetic Yb and Tm. By using a smaller and more mobile tag we managed in decreasing the anisotropy values of our paramagnetic system and this allowed us to more easily transfer assignment from the diamagnetic to the paramagnetic spectra and help in the assignment of the spectra of the protein tagged with [Tm(DOTA-M8)]. PCS and RDC were first analyzed and fitted separately. In order to fit them together, refinement of the protein's crystal structure was required, and performed using CYANA. In the end, we obtained tensors with improved fitting for both Yb and Tm datasets. Small RMSD (0.364 Å) between the refined solution structure and the X-ray structure indicates that both structures are in good agreement.

5.4 Outlook and applications to drug discovery

In this chapter we have shown that a tuning of the intensity of the paramagnetic effects could be needed when tagging a small protein such as *tc*Pex14.

As we have introduced in chapter 1, paramagnetic restraints are useful in structure-based drug design as the transfer of paramagnetic effects from a paramagnetically tagged protein to the NMR signals of a ligand in slow or fast exchange yields structural information on its binding mode and its site of interaction. To do so, we need to have a tensor its position and its intensity will induce shifts in the area of the protein's

surface where interactions and/or binding are expected. In the case of fast-exchanging ligands, even larger tensors are required since the paramagnetic effects are scaled according to the bound/unbound fraction. In order not to interfere with the binding process of the studied ligands, it may be worth considering a tagging site far enough from the binding site. It becomes clear that a balance needs to be found between the need to have a paramagnetic system inducing large-enough shifts in the binding area and a tagging site far enough not to interfere with the process; thus justifying the development of paramagnetic tags with higher anisotropy values in order to induce larger paramagnetic effects and at greater distances.

In the case of *tc*Pex14, the protein is small which limits the possibilities of tagging sites. However, the naturally occurring cysteine is in a good position, since it is close (15 Å) from but on the other side of the protein compared to the binding site. In order to characterize the binding modes of small ligands in fast exchange with the protein, a large tensor is needed; although as we have seen with DOTA-M8, a too large tensor on a small protein can be deleterious as PCS on the protein are hard to follow and therefore it is hard to fit the paramagnetic data to a tensor. The small VDPA tag, however, offers an interesting platform. As we have seen previously, it offers the possibility to tune the paramagnetic effects to the desired intensity. Tagging with Yb and Tm, we obtained medium tensors which induced PCS in the binding-site. Studying the interaction between *tc*Pex14 and weak binders, though, those tensors might not be large enough. In any case, having a refined structure is already an advantage for a structure-based drug design approach, as for the identification and the optimization of potential strong binders. Indeed, having optimized the orientation of some vectors and especially ^{N}H-N, we have a more accurate description of the protein's surface, thus offering the possibility to better compute interaction with small molecules.

5.5 Experimental Procedures

Protein was expressed and purified as reported in [108].

The tagging with DOTA-M8 was performed as such: Protein (previously reduced with 1mM TCEP) was buffer-exchanged to 6.5 mM NaPi, 20 mM NaCl, pH 6.5. 5 equivalents of tag were added. After 4 hours, reaction was complete as revealed by NMR. Excess tag was removed with a PD-10 column and longer NMR spectra were recorded.

The tagging with VDPA was performed in 6.5 mM HEPES, 20 mM NaCl, pH 6.5 following the protocol described in [86].

Spectra were acquired on a Bruker AVANCE NEO NMR spectrometer operating at 700 MHz (^{1}H Larmor frequency) at 298 K. RDC were measured on an AVANCE 900 MHz Bruker spectrometer using the IPAP method[110]. Spectra were analysed with CARA[111].

Refinement was performed using CYANA. The weights of the restraints were of 0.5 for RDC and 100 for PCS, and S_{LS} for RDC were set as 0.9.

ACKNOWLEDGMENTS

First, I would like to thank the AEGIS ITN program, funded by European Union's Framework Program Horizon 2020. Especially, I would like to thank Prof. Sattler and Dr. Schlosser for making this program such a rich and stimulating 3-year formation, on both the scientific and the human level. A warm thought goes to all my AEGIS friends, their presence and support have been a cheerful light from our very first seminar in Uppsala up to this very day.

I would have never made it through those 3 years without the support of all my colleagues, in the group of Prof. Nativi, in GiottoBiotech and at CERM. I would like to particularly thank Francesco, Marco and Sabrina. Thanks for being there for me, more often than I wish to admit, both inside and outside the lab. Oscar, I know proof-reading my thesis was so much fun, you are welcome. The sofa should be named after me as I already own more than half of it. Take care of Milanesi and don't make him frequentare too much. My infinite gratitude goes to Linda, I am forever grateful for what you have done for me, from the beginning and until the very end of those three years. A special thank you also goes to Stefano Cardelli, his e-mails del 10 del mese were always more than welcome. This acknowledgment expands to the whole team of Giotto for their support. Last, I would like to thank all the persons whom I have met at the polo scientifico during those 3 years, at the department of Chemistry and at CERM, and who have been accompanying me along this journey.

This thesis is dedicated to everybody who has been, willingly or not, contributing to it. This is to Sant'Ambrogio, Roberto e i Papi, for being home. This is to Gabriele, my favorite franco-calabrese and maybe my favorite at all, after all. This is to Mestola 62. This is to Ale, l'agila e il Mantra. This is to Gianlu. This is to le famiglie Pratesi, Billeri, e Del Vita. This is to ATAF linea 59. This is to LabUnico and the rotavapor del notturno. This is also to both sides of the Alps. This is to les Skuuu. This is to les 4F. This is to the AegisBoyz. This to Nantes and le Croisic, bien sûr. This is to les Denis bruns, les Denis blonds, les Coutant, les autres aussi. This to Griffon. This is to Ginevra.

I would like to carefully thank Matteo Gentili and Cristina Nativi, for having been not only scientific supervisors but also mentors for the past three years. Thank you for giving me this opportunity.

Last, I would like to thank my parents, not for being my scientific supervisors, but for having mentored me for the last twenty-six years. Merci, beaucoup.

In fondo spero che mia madre abbia avuto ragione,
e che davvero un laureato,
conti più di un cantante.

ANNEX 1: Generalities on experimental procedures

All the reagents were purchased from Sigma-Aldrich, except cyclen and cyclam derivatives which was purchased from Chematech. Small-molecules NMR experiments were recorded on a Bruker AVANCE II 500 MHz (^{1}H Larmor frequency) at 298 K. HPLC was performed with an Agilent 1200 Series with ZORBAX 300SB-18 analytical and semi-preparative. Mass spectrometry was performed on a Thermo-Fisher LTQ-XL ESI.

All protein-NMR -spectra were acquired on a Bruker AVANCE NEO NMR spectrometer operating at 700 MHz (^{1}H Larmor frequency) at 298 K equipped with a 5 mm TCI 3 channels HCN cryo-probehead. All the spectra were processed with the Bruker TopSpin 4.0.7 software package and analysed with the program CARA[111] (ETH Zürich).

MALDI was performed on a Bruker Daltonics Ultraflex III TOF/TOF.

ANNEX 2

The photo-catalysed thiol-ene reaction: a strategy tag to yield fast, selective and irreversible paramagnetic tagging of proteins

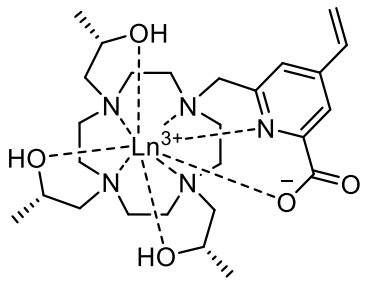
Maxime Denis[a,b], Charlotte Softley[c,d], Stefano Giuntini[b,e], Matteo Gentili[a], Enrico Ravera[e], Giacomo Parigi[b,e], Marco Fragai[b,e], Grzegorz Popowicz[d], Michael Sattler[c,d], Claudio Luchinat[b,e], Linda Cerofolini[e], Cristina Nativi[b*]*

[a]GiottoBiotech,S.R.L, Via Madonna del piano 6, 50019, Sesto Fiorentino (FI) Italy ; [b]Department of Chemistry "Ugo Schiff", Via della Lastruccia 3, 50019, Sesto Fiorentino (FI), Italy; [c]Biomolecular NMR, Deparment Chemie, Technische Universitaet Muenchen, Lichtenbergstrasse 4, 85747, Garching, Germany, [d]Institute of Structural Biology, Helmholtz Zentrum Muenchen, Neuherberg, Germany, [e]Magnetic Resonance Center (CERM), University of Florence, and Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine (C.I.R.M.M.P), Via L. Sacconi 6, 50019 Sesto FIorentino (FI), Italy

*to whom correspondence should be addressed: cerofolini@cerm.unifi.it, cristina.nativi@unifi.it

Paramagnetic restraints have been used in biomolecular NMR for the last three decades to elucidate and refine biomolecular structures, but also to characterize protein-ligand interactions. A common technique to generate such restraints in proteins, which do not naturally contain a (paramagnetic) metal, consists of attachment to the protein of a lanthanide-binding-tag (LBT).

In order to design such LBTs, it is important to consider the efficiency and stability of the conjugation, the geometry of the complex (conformational exchanges and coordination) and the chemical inertness of the ligand. Here we describe a photo-catalyzed thiol-ene reaction for the cysteine-selective paramagnetic tagging of proteins. As a model, we designed an LBT with a vinyl-pyridine moiety which was used to attach our tag to the protein GB1 in a fast and irreversible fashion. Our tag **T1** yielded medium tensors with different lanthanides and was carefully characterized via NMR and relaxometry measurements.



T1

INTRODUCTION

Nuclear magnetic resonance (NMR) is a standard technique used for the study of biomacromolecules in solution. Indeed, it offers the opportunity to investigate structure, behavior, internal motions and mechanism of action of proteins, as well as their interactions with small molecules and other biomolecules at the atomic level[112,19].

Structure calculation by NMR mainly relies on the collection of short-range distance restraints (up to ~5-6 Å) provided by the time-consuming and troublesome analysis of NOESY spectra. The use of long-range paramagnetic distance restraints (up to ~40 Å), such as pseudo-contact shifts (PCS), induced by a paramagnetic ion, has been widely proposed to help in *de novo* structure determination by NMR [2,10,20,87] or in the refinement of pre-existing X-ray structures[35,113,114]. In this way, a more reliable model describing the protein in solution can be obtained from the crystal structure [26,33].

Furthermore, PCS restraints, together with residual dipolar coupling (RDC) restraints originating from the same tensor, have been exploited in the analysis of the internal dynamics of multi-domain proteins[43,115–117] and in the investigation of the interaction of proteins with their partners or ligands[118].

One approach to incorporate paramagnetic metal ions into metallo-proteins is via the exchange of the naturally occurring diamagnetic metal ion[16,119]. As a more widely applicable alternative, any protein, not necessarily binding metal ions, can be made paramagnetic by attaching metal binding peptides or organic small synthetic ligands chelating paramagnetic metal ions[45,50,56,109,120,121]. The idea is to specifically attach the metal at a chosen position to monitor protein dynamics or protein-protein and protein-ligand interactions[6].

To generate such restraints, we often rely on chemically synthesized lanthanide-binding paramagnetic tags (LBT)[45]. An important feature of such molecules lies in their ability to react quickly and selectively with the desired amino acid. Apart from non-natural amino acids, paramagnetic tags generally target cysteines. The standard strategy is conjugation

by formation of a disulfide bond: this technique has been widely used but it is hampered by the stability of the bond formed[79,122]. Thus, in the last years, much effort has been concentrated on the design of paramagnetic tags that are stable under reducing conditions once tagged[123,124], and which could also be used in-cell[80]. For this purpose, thiol-ene coupling (TEC) has emerged as a useful tool in protein chemical conjugation[84]. It creates, through a radical mechanism, a stable thioether C-S bond between the free thiol of a cysteine and a double bonded carbon in an irreversible fashion. This reaction can be spontaneous but may also be photo-catalyzed[82]. There are a few examples of the use of this reaction for the paramagnetic tagging of proteins[86] but the reactions are slow (overnight). Here, we describe a new vinyl-pyridine-based paramagnetic tag and its conjugation to proteins via the presented photo-catalyzed thiol-ene reaction to demonstrate that this reaction increases the applicability of paramagnetic tagging by making it suitable for proteins that are not stable for extended periods of time in vitro.

So as to design LBTs, it is important to consider the geometry of the complex and the chemical inertness of the ligand. ((2S,2'S,2''S,2'''S)-1,1',1'',1'''-(1,4,7,10-tetraazacyclo dodecane-1,4,7,10-tetrayl)tetrakis (propan-2-ol)), (S)-THP has been proposed as a simple yet enantiomerically pure[90] platform for the design of LBTs[85]. Hence, we have designed a THP-like LBT (**T1**) with a vinyl-pyridine group as the single point of attachment, designed to be conjugated to proteins via TEC.

GB1 has been chosen as model protein and engineered with a cysteine residue. The effect of the designed LBT, chelating different paramagnetic metal ions (Lu^{3+}, Yb^{3+} and Dy^{3+}), on the resonances of the protein has been investigated by solution NMR. Relaxometry measurements were also performed to obtain information on the lanthanide coordination site in **T1**(Ln).

RESULTS AND DISCUSSION

Synthesis of T1

In order to synthesize **T1,** was needed a picolinic-like intermediate bearing the double-bond for conjugation and a leaving group for cyclen alkylation. **A** was designated as a key intermediate in the synthesis.

Derivative **A** was synthesized from commercial chelidamic acid (fig. 1). The carboxylic acids of chelidamic acid were protected as methyl esters and position 4 was brominated with tetrabutylammonium bromide (TBAB) and P_2O_5. The reduction of a single methyl ester was easily achieved thanks to the peculiar reactivity of pyridine 2,6 diester. The double bond in position 4 was then introduced via a Suzuki-like coupling[125] and the benzylic hydroxyl was subsequently replaced by a chloride as a good leaving group, giving **A** with a yield of 50% calculated over 5 steps.

The intermediate **A** was then added to an excess of cyclen to isolate the mono functionalized cyclen derivative **B** (fig. 2). **B** was subsequently alkylated using (S)-propylene oxide and the methyl ester was deprotected to yield **C**. Lanthanide-chelation was achieved quantitatively, by refluxing overnight **C** in the presence of $LnCl_3$ salts in a H_2O/MeCN mixture.

Conjugation to GB1 T53C

In order to test our LBT we chose the GB1 T53C mutant as a model system. GB1 is a small globular protein which is stable under many conditions and engineered to bear one cysteine. Vinyl picolinic acids have been shown to react through TEC with cysteines overnight by Su *et al*[86,83], but this reaction is quite slow (overnight). To reduce the reaction time, we used a photo-activated radical initiator, 2,2-dimethoxy-2-phenylacetophenone (DPAP), which has been shown to catalyze TEC when activated with UV at 365 nm.

To a solution of GB1, 5 equiv. of tag and 5 equiv. of DPAP were added. The mixture was irradiated by a UV lamp at 365 nm for 1 hour after which the sample was analyzed by NMR. We used **T1** loaded with Lu^{3+}, Yb^{3+} and Dy^{3+}. The protein tagging was evaluated recording 2D ^{1}H-^{15}N HSQC spectra which showed after 1 hour a conjugation of 80% (fig. 3). In each experiment, only one set of peaks was observed, indicating that no side-reaction or protein degradation took place.

Evaluation of paramagnetic effects

In our case, the tagging of GB1 T53C with paramagnetic [**T1**(Yb)] and [**T1**(Dy)] resulted in PCS in

the 2D $^{1}H–^{15}N$ HSQC spectrum (fig. 4). As expected, the two paramagnetic metal ions provide shifts in opposite directions[16]. In order to differentiate between the chemical shift perturbation stemming from the tagging itself and the paramagnetic contribution to the shifts (PCS), we also tagged GB1 with the diamagnetic [T1(Lu)]. The resonances which shifted after the addition of the diamagnetic compound correspond to residues located close to the tagging-site (C.f. S.I. fig. S1). We noted that for each cross-peak in the 2D $^{1}H–^{15}N$ HSQC of diamagnetic samples there was only one corresponding cross-peak in the spectra of paramagnetic species. We therefore concluded that **T1** was indeed present as a single stereoisomer or a single conformation which gave rise to a single set of paramagnetically shifted peaks.

The PCS are commonly described as a function of both the nuclear coordinates and an anisotropy tensor ($\boldsymbol{\Delta\chi}$)[126,127], according to the equation

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

where r, θ and φ are the spherical coordinates of the nucleus in the frame in which the anisotropy tensor is diagonal and has its origin at the metal position. $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic anisotropies of the tensor, defined as

$$\Delta\chi_{ax} = \chi_{zz} - \frac{\chi_{xx} + \chi_{yy}}{2}$$

$$\Delta\chi_{rh} = \chi_{xx} - \chi_{yy}$$

The program FANTEN[17] can be used to obtain the best fit $\boldsymbol{\Delta\chi}$ tensor (consisting of 5 parameters: $\Delta\chi_{ax}$, $\Delta\chi_{rh}$ and the three Euler angles defining the frame in which the tensor is diagonal) and the coordinates of the metal ion from the PCS values and the coordinates of the protein nuclei.

First, the resonances of 19 peaks in the spectra of the protein tagged with [T1(Yb)] were unambiguously assigned, and the PCS were evaluated. The anisotropy tensor for the Yb^{3+} tagged protein was determined from the best fit of these 19 H^N PCS to the X-ray structure (PDB 1IGD)[128], using the program FANTEN. During these tensor calculations, the position of the metal was also obtained. New NMR peaks could then be assigned, taking advantage of the PCS values predicted by the program for the other nuclei, so

that a total of 36 H^N PCSs could be obtained for the [**T1**(Yb)] tagged protein.

The assignment of the spectrum of the protein tagged with [**T1**(Dy)] was obtained by comparison with the assigned spectra of the [**T1**(Yb)] and [**T1**(Lu)] tagged proteins (fig. 4); in this way 32 PCS could be obtained for [**T1**(Dy)].

The PCS originating from [**T1**(Yb)] and [**T1**(Dy)] were first analyzed separately. The metal positions obtained from the two sets of data were similar, as well as the direction of the main axes of the best fit tensors (fig. 5A and 5B). Then, the two sets of PCSs were evaluated jointly by constraining both metals to reside in the same position: the orientations and magnitudes of the two tensors were almost unaffected with respect to the values obtained from the separate fits, thus showing high consistency between the two sets of paramagnetic data (Table 1 and fig. 5). The two tensors are almost coaxial, with angles of 11.5° between the two Z-axes, 5° between the two X-axes and 12° between the two Y-axes. The agreement between experimental and back-calculated PCS obtained from the fit of both sets of data is very good, with Q factors of 0.064 and 0.072, for [**T1**(Yb)] and [T1(Dy)], respectively (fig. S3).

It has been shown that some LBTs can be immobilized on the protein surface by an electrostatic interaction between the lanthanide cage and a carboxylate belonging to an Asp or Glu residue[86,129]. This interaction may decrease the tag mobility and therefore increase the effective tensor anisotropies. In the case of GB1 T53C, there are two carboxylates at distances of 9 and 11 Å from the cysteine (E42, E56); we therefore expected to see such a stabilizing interaction taking place between our tag and our protein. However, we noted that the axial anisotropies of the two tensors are a factor 4-5 smaller than the values expected for the anisotropies of the magnetic susceptibility tensors of the Yb^{3+} and Dy^{3+} ions usually calculated from rigid systems containing these paramagnetic ions[36,130]. This likely indicates that some motional averaging of the magnetic susceptibility tensors occurred, so that the best fit tensors result from the averaging of the magnetic susceptibility anisotropy tensors due to tag mobility.

Interestingly, the position of the metal calculated with FANTEN is

relatively far from the protein surface. This observation corroborates the idea that the tag is not interacting with the protein surface, thus being very mobile, in agreement with the small tensor anisotropies observed. We suppose that in **T1** the lanthanide ions have a coordination number of 9 and have no further coordination site with which to interact with the carboxylate positioned on the protein surface.

Relaxometry

In order to characterize the coordination of the lanthanides in **T1**, FFC relaxometry experiments were performed[131-98]. The ^{1}H nuclear magnetic relaxation dispersion (NMRD) profiles of the [**T1**(Gd)] complex in water solution at 10, 25 and 37 °C are shown in figure 7. The profiles are characterized by dispersions somewhat smoother than predicted by the Lorentzian spectral density function; however, any attempt to reproduce them by considering inner sphere and outer-sphere contributions failed (with the diffusion coefficients constrained to values in the range expected for water solutions) unless the distance of closest approach between paramagnetic ion and diffusive water molecules was larger than 8 Å. Therefore, no sizable contribution from outer-sphere relaxation is apparent, and the profiles were fit by including two protons at a distance r_1 and other two protons at a distance r_2 from the Gd^{3+} ion (r_1 and r_2 were left free to be adjusted in the best fit analysis).

The best fit parameters are reported in Table 2 and the corresponding profiles in Figure 7 as solid lines. Due to the low sensitivity to the correlation time τ_v for electron relaxation, the values of τ_v were kept fixed to values typically observed in gadolinium complexes. Although there exists some covariance among the different parameters, the analysis indicates that the lifetime of the two protons at r_1 = 3.05 Å, $\tau_{M(1)}$, is as long as several microseconds (ca. 4 μs at 25 °C). The analysis also shows contributions from fast exchanging second-sphere water molecules, with a lifetime on the picosecond timescale. The number of these second-sphere water molecules is totally covariant with the metal-proton distance r_2; if one water molecule is considered, r_2 is about 3.5 Å.

Fits of equivalent quality, however, can be obtained also for increasing values of r_1 (and decreasing values of r_2). When the condition $r_1 = r_2 = 3.3$ Å is met, almost indistinguishable best fit

profiles are calculated, the complex reorientation time τ_R, the transient ZFS Δ_t and the lifetimes $\tau_{M(1)}$ and $\tau_{M(2)}$ being somewhat increased (τ_R = 84 ps, Δ_t = 0.027 cm^{-1}, $\tau_{M(1)}$ = 5.3 μs and $\tau_{M(2)}$ = 28 ps, at 25 °C).

The fast exchanging second-sphere water protons may be related to one or more water molecules hydrogen bonded in positions allowing for a large mobility.

The presence of two protons exchanging in the microsecond time scale at 3.05 Å, i.e. at the distance expected for the protons of a water molecule regularly coordinated to the Gd^{3+} ion, or at a somewhat larger distance (such as 3.3 Å), could be related to one water molecule hydrogen-bonded to the oxygen atoms of hydroxyl groups or of the carboxylate group. Its lifetime is much longer (several microseconds) than the lifetime of the water molecule coordinated to the Gd^{3+} ion in DO3A-like or DOTA-like complexes (on the sub-microsecond timescale), likely because of the bipositive charge of this complex which disfavours water exchange with respect to neutral or negative complexes[93]. A long lifetime of this water molecule indicates a relatively "compact" and stable environment surrounding the Gd^{3+} ion, which may prevent the replacement of the coordinated water by negatively charged groups present on the protein surface. This is in agreement with the observed tag mobility that affects the magnitude of the PCS-determined tensor. Alternatively, these slow-exchanging protons may be the hydrogen atoms of the three hydroxyl groups coordinated to the Gd^{3+} ion.

CONCLUSIONS

In this work, we presented the synthesis and the performance of an enantiopure paramagnetic tag for NMR spectroscopy designed to efficiently conjugate proteins under mild conditions using UV irradiation. In particular, we have demonstrated that UV-catalyzed TEC was applicable to the paramagnetic tagging of GB1 T53C with a high yield and short reaction times. The creation of the stable thioether bond required much shorter reactions times than previously described methods[51,86].

The conjugation involves the protein's cysteine side-chain and forms thioether bonds without affecting the paramagnetic properties of the tag. Being easily obtained, and in high yields, the formation of non-reducible thioether bonds is

convenient and makes this strategy a method of choice in the field of paramagnetic tagging, opening the way to development of more novel paramagnetic tags.

EXPERIMENTAL SECTIONOrganic synthesis

All the reagents were purchased from Sigma-Aldrich, except cyclen which was purchased from Chematech. Small-molecules NMR experiments were recorded on a Bruker AVANCE II 500 MHz (^{1}H Larmor frequency) at 298 K. HPLC was performed with an Agilent 1200 Series with ZORBAX 300SB-18 analytical and semi-preparative. Mass spectrometry was performed on a Thermo-Fisher LTQ-XL ESI.

Formation of Lanthanide Complexes

C (70 mg) and $LnCl_3$ $n H_2O$ (100 mg) were dissolved in 6 mL H_2O/ MeCN (50/50). The pH was adjusted to 7 and the mixture was refluxed. Chelation was typically quantitative overnight, as shown by LC/MS analysis. Reaction mixtures were purified via semi-preparative HPLC.

Protein Conjugation

GB1 T53C was expressed and purified as reported in the supplementary information. Prior to conjugation, the protein was buffer-exchanged into NaPi 20 mM, pH 7.5, then concentrated to 180 µM. 5 equivalents each of DPAP and [**T1**(Ln)] were added to the protein, as well as 10% D2O. The mixture was transferred into a 5-mm NMR tube and argon was gently bubbled through the solution for 5 minutes. The tube was placed under a UV Lamp (UVGL-55 Mineralight 26W) at 365 nm for 1h, after which the sample was measured using NMR spectroscopy and the spectra analyzed. Excess of small molecules were then washed away by buffer exchange and NMR spectra were recorded again.

NMR measurements and PCS analysis

All the experiments were acquired on a Bruker AVANCE NEO NMR spectrometer operating at 700 MHz (^{1}H Larmor frequency) at 298 K equipped with a 5 mm TCI 3 channels HCN cryo-probehead. All the spectra were processed with the Bruker TopSpin 4.0.7 software package and analysed with the program CARA[111] (ETH Zürich).
The spectra were collected using a protein concentration of ~ 180 µM in buffered solution (20 mM sodium phosphate pH 7.5).
The assignment of GB1 was taken from the literature[132].

The PCS values were calculated from the difference in the value of chemical shift of each amino acid peak between the paramagnetic [**T1**(Yb) or **T1**(Dy)] and diamagnetic [**T1**(Lu)] 2D ^{1}H-^{15}N HSQC spectra acquired.
The fitting of the PCS tensor was carried out using the program FANTEN[17].

ACKNOWLEDGMENTS

This worked received funding from the European Union's framework program for Research and Innovation Horizon 2020 (2014-2020) under the Marie–Slodowska Curie Grant agreement number 675555, Accelerated Early-staGe drug dIScovery (AEGIS). The authors also acknowledge MIUR Italy PRIN 2012SK7ASN, "Progetto Dipartimenti di Eccellenza 2018–2022" to the Department of Chemistry "Ugo Schiff" and iNEXT, Grant number 653706, funded by the Horizon 2020 program of the European Union.

Keywords Paramagnetic tags, photo-catalyzed thiol-ene reaction, lanthanide coordination, Biomolecular NMR, protein structure refinement.

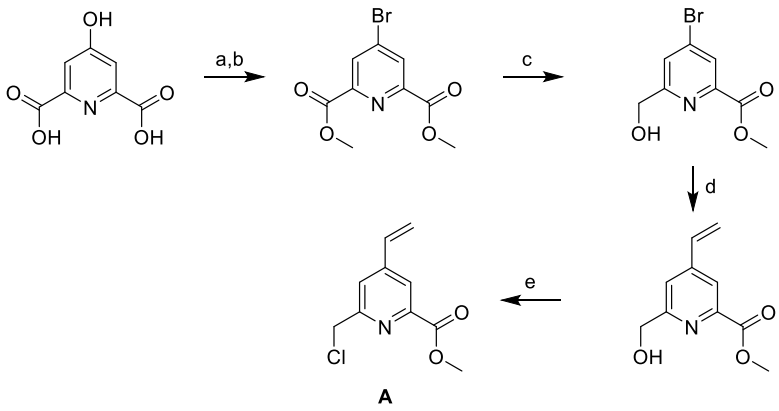


Figure 1: Synthesis of intermediate **A**

a. H_2SO_4, MeOH, 18h, 98%. b. P_2O_5, TBAB, Tol, 3h, 97%. c. $NaBH_4$, MeOH, 75% d. trivinyl-boroxin, K_3PO_4, JohnPhos, $Pd(dba)_2$, dioxane, 2h, 68% e. MsCl, DIPEA, DCM, 95%

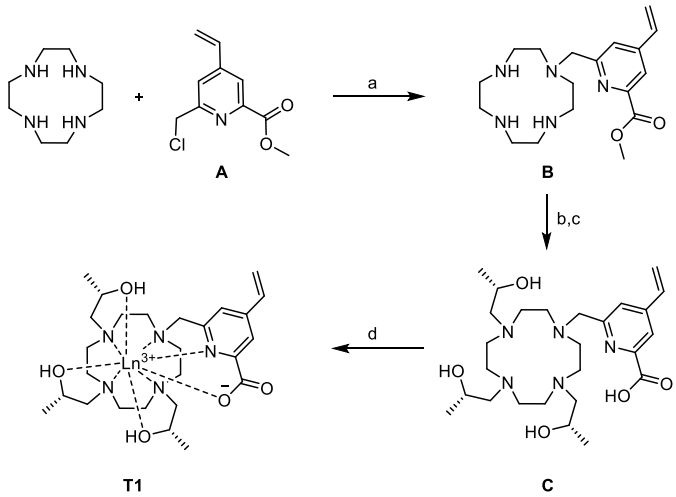


Figure 2: Synthesis of **T1**

a. DCM, 3 days b. (S)-propylen Oxide, MeOH, 4 days, quant. c. LiOH, THF/H_2O, 18h, quant. d. $LnCl_3$ $n H_2O$, H_2O/MeCN

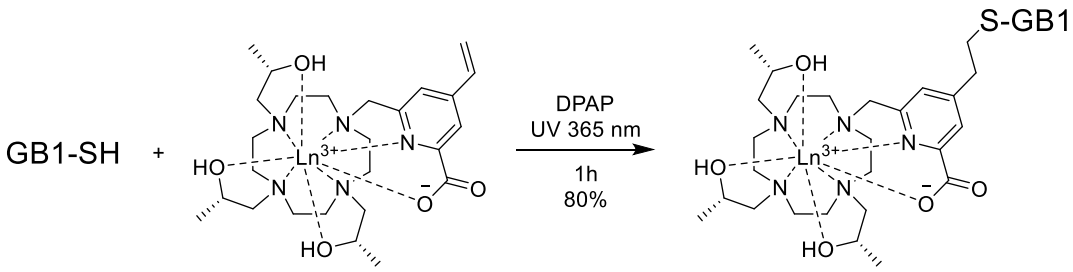


Figure 3: Tagging of GB1 with **T1** using the thiol-ene reaction.

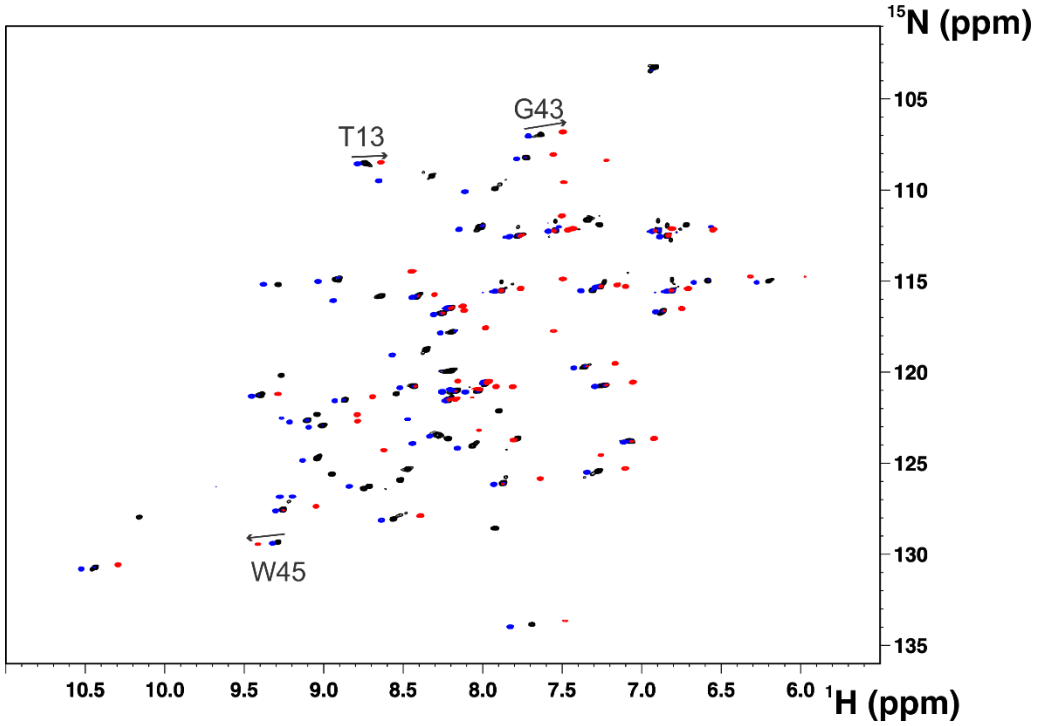


Figure 4: 2D ^{1}H–^{15}N HSQC of GB1 T53C tagged with T1(Lu) (black), T1(Yb) (blue) and T1(Dy) (red), acquired at 700 MHz and 298 K.

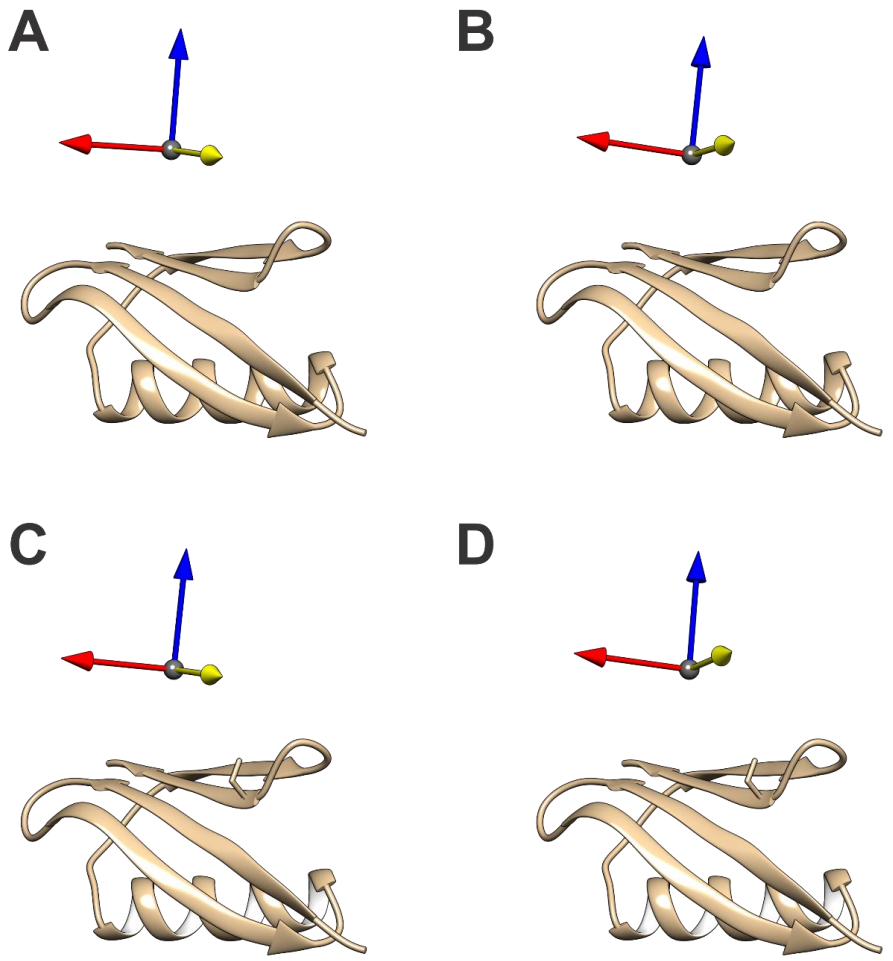


Figure 5. Graphical representation of the magnetic susceptibility anisotropy tensor orientations. The x, y and z axes (corresponding in turn to the directions with the smallest, intermediate and largest magnetic susceptibility) are represented as red, yellow and blue arrows, respectively. The panels A and C show the orientation of Yb tensors when only its own PCS dataset was considered, and when both datasets of Yb and Dy were both taken into account in FANTEN, respectively. The panels B and D show the orientation of the Dy tensors when only its PCS-data-set was considered, and when data sets of Yb and Dy where both taken into account in FANTEN, respectively.

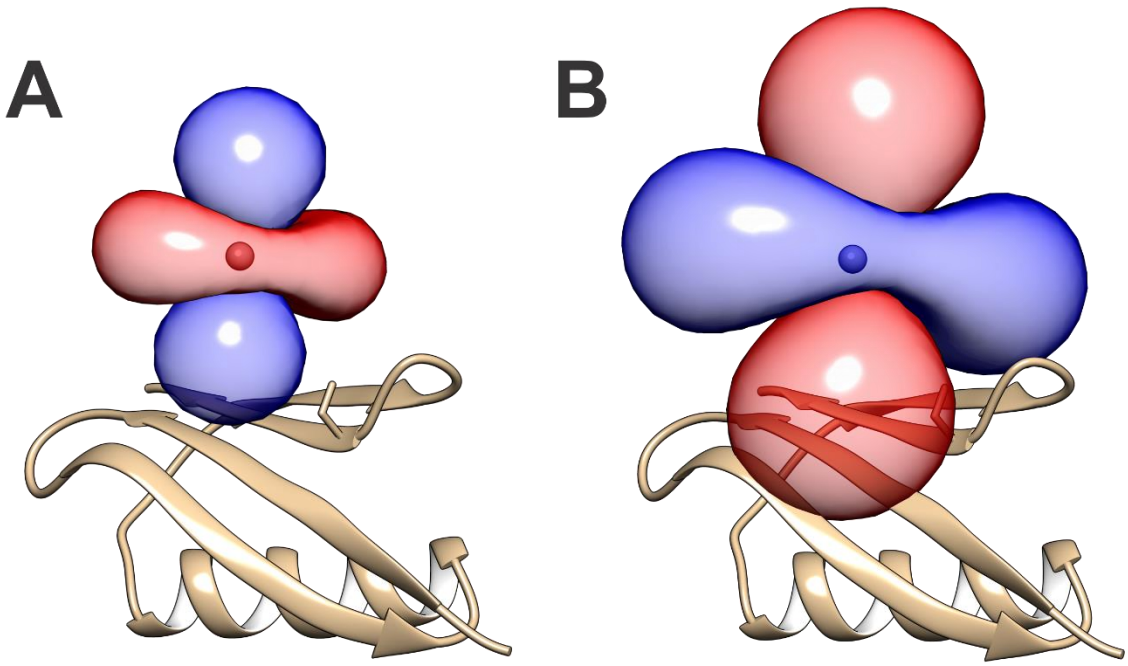


Figure 6. Graphical representation of PCS iso-surfaces of 1 (blue) and -1 (red) ppm obtained for Yb^{3+}-T1 (A) and Dy^{3+}-T1 (B) using the program FANTEN.

Table 1. Tensor parameters calculated with the program FANTEN using the PCS values measured with T1(Yb) and T1(Dy) implemented separately or jointly in the evaluation of the metal position.

Metal	PCS restraints source	Q factor	$\Delta\chi_{ax}$ (10^{-32} m^3)	$\Delta\chi_{rh}$ (10^{-32} m^3)
Yb	Yb	0.061	1.91 ± 0.02	0.68 ± 0.25
Dy	Dy	0.072	-6.51 ± 0.02	2.54 ± 0.89
Yb	Yb & Dy	0.064	1.83 ± 0.03	-0.59 ± 0.28
Dy	Yb & Dy	0.072	-6.81 ± 0.21	3.06 ± 0.89

Table 2. Best fit parameters obtained from the NMRD profiles of T1(Gd), shown in Figure 7. Contributions from two protons in the first-coordination and second-coordination spheres were considered.

	37 °C	25 °C	10 °C
τ_R (ps)	35	57	112
Δ_t (cm^{-1})	0.023		
τ_V (ps)	15	20	25
r_1 (Å)	3.05		
$\tau_{M(1)}$ (μs)	3.4	3.8	4.6
r_2 (Å)	3.48		
$\tau_{M(2)}$ (ps)	8.8	18	51

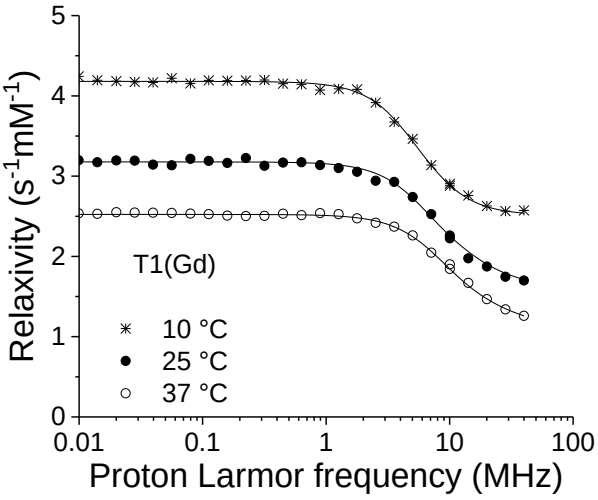


Figure 7: NMRD Profiles of **T1(Gd)** measured at different temperatures.

REFERENCES

1. Bertini, I., Luchinat, C. & Parigi, G. Magnetic susceptibility in paramagnetic NMR. **40**, 249–273 (2002).

2. Otting, G. Protein NMR Using Paramagnetic Ions. *Annu. Rev. Biophys.* **39**, 387–405 (2010).

3. Bertini, I., Luchinat, C., Parigi, G. & Pierattelli, R. NMR spectroscopy of paramagnetic metalloproteins. *ChemBioChem* **6**, 1536–1549 (2005).

4. Koehler, J. & Meiler, J. Expanding the utility of NMR restraints with paramagnetic compounds: Background and practical aspects. *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 360–389 (2011).

5. Solomon, I. Relaxation Processes in a System of Two Spins. *Phys. Rev.* **953**, 559–565 (1955).

6. Keizers, P. H. J. & Ubbink, M. Paramagnetic tagging for protein structure and dynamics analysis. *Prog. Nucl. Magn. Reson. Spectrosc.* **58**, 88–96 (2011).

7. Berliner, L. J., Grunwald, J., Hankovszky, H. O. & Hideg, K. A novel reversible thiol-specific spin label: Papain active site labeling and inhibition. *Anal. Biochem.* **119**, 450–455 (1982).

8. Battiste, J. L. & Wagner, G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355–5365 (2000).

9. Solomon, I. & Bloembergen, N. Nuclear magnetic interactions in the HF molecule. *J. Chem. Phys.* **25**, 261–266 (1956).

10. Barbieri, R., Luchinat, C. & Parigi, G. Backbone-only protein solution structures with a combination of classical and paramagnetism-based constraints: A method that can be scaled to large molecules. *ChemPhysChem* **5**, 797–806 (2004).

11. Iwahara, J., Schwieters, C. D. & Clore, G. M. Ensemble Approach for NMR Structure Refinement against 1H Paramagnetic Relaxation Enhancement Data Arising from a Flexible Paramagnetic Group Attached to a Macromolecule. *J. Am. Chem. Soc.* **126**, 5879–5896 (2004).

12. Karthikeyan, G. *et al.* A Bioresistant Nitroxide Spin Label for In-Cell EPR Spectroscopy: In Vitro and In Oocytes Protein Structural Dynamics Studies. *Angew. Chemie - Int. Ed.* **57**, 1366–1370 (2018).

13. Bertini, I., Luchinat, C., Parigi, G. & Ravera, E. *The hyperfine shift*. (2017). doi:10.1016/B978-0-444-63436-8/00002-8

14. Woźniak-Braszak, A., Jurga, K. & Baranowski, M. The Lipari-Szabo Model-Free Analysis as a Method for Study of Molecular Motion in Solid State Heteronuclear Systems Using NMR Off-Resonance. *Appl. Magn. Reson.* **47**, 567–574 (2016).

15. Bertini, I., Luchinat, C., Parigi, G. & Ravera, E. The effect of partial orientation: residual dipolar couplings. *Solut. NMR Paramagn. Mol.* 61–76 (2017). doi:10.1016/b978-0-444-63436-8.00003-x

16. Bertini, I., Janik, M. B. L., Lee, Y., Luchinat, C. & Rosato, A. Magnetic Susceptibility Tensor Anisotropies for a Lanthanide Ion Series in a Fixed Protein Matrix. *J. Am. Chem. Soc.* 4181–4188 (2001). doi:10.1021/ja0028626

17. Rinaldelli, M., Carlon, A., Ravera, E., Parigi, G. & Luchinat, C. FANTEN : a new web-based interface for the analysis of magnetic anisotropy-induced NMR data. 21–34 (2015). doi:10.1007/s10858-014-9877-4

18. Schmitz, C., Stanton-Cook, M. J., Su, X. C., Otting, G. & Huber, T. Numbat: An interactive software tool for fitting δχ-tensors to molecular coordinates using pseudocontact shifts. *J. Biomol. NMR* **41**, 179–189 (2008).

19. I. Bertini, K.S. McGreevy, G. P. *NMR of Biomolecules: Towards Mechanistic Systems Biology*. (2012).

20. Bertini, I. *et al.* Paramagnetism-based versus classical constraints: An analysis of the solution structure of Ca Ln calbindin D9k. *J. Biomol. NMR* **21**, 85–98 (2001).

21. Bertini, I., Luchinat, C. & Parigi, G. Paramagnetic constraints: An aid for quick solution structure determination of paramagnetic metalloproteins. *Concepts Magn. Reson. Part A Bridg. Educ. Res.* **14**, 259–286 (2002).

22. Kemple, M. D., Ray, B. D., Lipkowitz, K. B., Prendergast, F. G. & Rao, B. D. N. The use of lanthanides for solution structure determination of biomolecules by NMR. Evaluation of the methodology with EDTA derivatives as model systems. *J. Am. Chem. Soc.* **110**, 8275–8287 (1988).

23. Bahramzadeh, A., Huber, T. & Otting, G. Three-Dimensional Protein Structure Determination Using Pseudocontact Shifts of Backbone Amide Protons Generated by Double-Histidine Co2+-Binding Motifs at Multiple Sites. *Biochemistry* **58**, 3243–3250 (2019).

24. Pilla, K. B., Otting, G. & Huber, T. Protein Structure Determination by Assembling Super-Secondary Structure Motifs Using Pseudocontact Shifts. *Structure* **25**, 559–568 (2017).

25. Banci, L., Bertini, I., Cavallaro, G., Giachetti, A. & Luchinat, C. Paramagnetism-based restraints for Xplor-NIH. 249–261 (2004).

26. Carlon, A., Ravera, E., Parigi, G., Murshudov, G. N. & Luchinat, C. Joint X-ray / NMR structure refinement of multidomain / multisubunit systems. *J. Biomol. NMR* **73**, 265–278 (2019).

27. Allegrozzi, M. *et al.* Lanthanide-induced pseudocontact shifts for solution structure refinements of macromolecules in shells up to 40 Å from the metal ion. *J. Am. Chem. Soc.* **122**, 4154–4161 (2000).

28. Shaanan B. Gronenborn A. Combining experimental information from crystal and solution studies: joint X-ray and NMR refinement. *Science (80-.).* **257**, 961–964 (1992).

29. Bertini I, Del Bianco C, Gelis I, Katsaros N, Luchinat C, Parigi G, Peana M, Provenzani A, Z. M. Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. *Proc Natl Acad Sci U S A.* 6841,6 (2004).

30. Xingfu Xu, Peter H. J. Keizers, Wolfgang Reinle, Frank Hannemann, Rita Bernhardt, M. U. Intermolecular dynamics studied by paramagnetic tagging. *J. Biol. NMR* **43**, 257–264 (2009).

31. Pintacuda, G., Park, A. Y., Keniry, M. A., Dixon, N. E. & Otting, G. Lanthanide labeling offers fast NMR approach to 3D structure determinations of protein-protein complexes. *J. Am. Chem. Soc.* **128**, 3696–3702 (2006).

32. Andrałojć, W. *et al.* Identification of productive and futile encounters in an electron transfer protein complex. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1840–E1847 (2017).

33. Nitsche, C. & Otting, G. NMR studies of ligand binding. *Curr. Opin. Struct. Biol.* **48**, 16–22 (2018).

34. Xu, X. *et al.* Intermolecular dynamics studied by paramagnetic tagging. *J. Biomol. NMR* **43**, 247–254 (2009).

35. Pintacuda, G., John, M., Su, X. C. & Otting, G. NMR structure determination of protein - Ligand complexes by lanthanide labeling. *Acc. Chem. Res.* **40**, 206–212 (2007).

36. Dasgupta, S. *et al.* Narrowing the conformational space sampled by two-domain proteins with paramagnetic probes in both domains. *J. Biomol. NMR* **51**, 253–263 (2011).

37. Saio, T. *et al.* Ligand-driven conformational changes of MurD visualized by paramagnetic NMR. *Sci. Rep.* **5**, 1–11 (2015).

38. John, M., Pintacuda, G., Park, A. Y., Dixon, N. E. & Otting, G. Structure determination of protein-ligand complexes by transferred paramagnetic shifts. *J. Am. Chem. Soc.* **128**, 12910–12916 (2006).

39. Zimmermann, K., Joss, D. & Nogueira, E. S. Localization of ligands within human carbonic anhydrase II using 19 F pseudocontact shift analysis †. *Chem. Sci.* 5064–5072 (2019). doi:10.1039/c8sc05683h

40. Xu, D. *et al.* Ligand Proton Pseudocontact Shifts Determined from Paramagnetic Relaxation Dispersion in the Limit of NMR Intermediate Exchange. *J. Phys. Chem. Lett.* **9**, 3361–3367 (2018).

41. Guan, J. Y. *et al.* Small-molecule binding sites on proteins established by paramagnetic NMR spectroscopy. *J. Am. Chem. Soc.* **135**, 5859–5868 (2013).

42. Chen, W. N. *et al.* Sensitive NMR Approach for Determining the Binding Mode of Tightly Binding Ligand Molecules to Protein Targets. *J. Am. Chem. Soc.* **138**, 4539–4546 (2016).

43. Cerofolini, L. *et al.* Examination of Matrix Metalloproteinase-1 in Solution. *J. Biol. Chem.* **288**, 30659–30671 (2013).

44. Su, X. C., Huber, T., Dixon, N. E. & Otting, G. Site-specific labelling of proteins with a rigid lanthanide-binding tag. *ChemBioChem* **7**, 1599–1604 (2006).

45. Joss, D., Häussinger, D., Joss, D. & Häussinger, D. Design and applications of lanthanide chelating tags for pseudocontact shift NMR spectroscopy with biomacromolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* (2019). doi:10.1016/j.pnmrs.2019.08.002

46. Ziegler, W., Lindstrom, M. & McTavish, R. Quantitative Determination of Mononucleotide Conformations in Solution using Lanthanide Ion Shift and Broadening NMR Probes. *Nature* **26**, 584–585 (1971).

47. Dean Sherry, A., Geraldes, C. F. G. C. & Cacheris, W. P. 31P and 23Na NMR lanthanide induced shifts in axially symmetric macrocyclic phosphonate complexes. *Inorganica Chim. Acta* **139**, 137–139 (1987).

48. Ikegami, T. *et al.* Novel techniques for weak alignment of proteins in solution using chemical tags coordinating lanthanide ions. *J. Biomol. NMR* **29**, 339–349 (2004).

49. Peters, F. *et al.* Cys-Ph-TAHA: A lanthanide binding tag for RDC and PCS enhanced protein NMR. *J. Biomol. NMR* **51**, 329–337 (2011).

50. Keizers, P. H. J., Saragliadis, A., Hiruma, Y., Overhand, M. & Ubbink, M. Design, synthesis, and evaluation of a lanthanide chelating protein probe: CLaNP-5 yields predictable paramagnetic effects independent of environment. *J. Am. Chem. Soc.* **130**, 14802–14812 (2008).

51. Yang, Y., Wang, J.-T., Pei, Y.-Y. & Su, X.-C. Site-specific tagging proteins via a rigid, stable and short thiolether tether for paramagnetic spectroscopic analysis. *Chem. Commun.* **51**, 2824–2827 (2015).

52. Dvoretsky, A., Gaponenko, V. & Rosevear, P. R. Derivation of structural restraints using a thiol-reactive chelator. *FEBS Lett.* **528**, 189–192 (2002).

53. Alexander, V. Design and Synthesis of Macrocyclic Ligands and Their Complexes of Lanthanides and Actinides. *Chem. Rev.* **95**, 273–342 (1995).

54. Wainwright, K. P. *Synthetic and structural aspects of the chemistry of saturated polyaza macrocyclic ligands bearing pendant coordinating groups attached to nitrogen*. *Coordination Chemistry Reviews* **166**, (Elsevier Science S.A., 1997).

55. Keizers, P. H. J., Desreux, J. F., Overhand, M. & Ubbink, M. Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. *J.*

Am. Chem. Soc. **129**, 9292–9293 (2007).

56. Häussinger, D., Huang, J. R. & Grzesiek, S. DOTA-M8: An extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 14761–14767 (2009).

57. Benton, C. B. *et al.* Analysis of the isomer ratios of polymethylated-DOTA complexes and the implications on protein structural studies. *Dalt. Trans.* **95**, 222–242 (2016).

58. Peters, J. A., Huskens, J. & Raber, D. J. Lanthanide induced shifts and relaxation rate enhancements. *Prog. Nucl. Magn. Reson. Spectrosc.* **28**, 283–350 (1996).

59. Bleaney, B. Nuclear magnetic resonance shifts in solution due to lanthanide ions. *J. Magn. Reson.* **8**, 91–100 (1972).

60. Weisman, G. R., Rogers, M. E., Wong, E. H., Jasinski, J. P. & Paight, E. S. Cross-Bridged Cyclam. Protonation and Li+ Complexation in a Diamond-Lattice Cleft. *J. Am. Chem. Soc.* **112**, 8604–8605 (1990).

61. Bencini, A. *et al.* Proton inclusion properties of a new azarnacrocycle. Synthesis, characterization and crystal structure of [H3L][C1]3.2H20 (L = 4, 10-dimethyl-l, 4, 7, 10-tetraazabicyclo [5.5.2] tetradecane). *Supramol. Chem.* **3**, 141–146 (1994).

62. Weisman, G. R. *et al.* Synthesis and transition-metal complexes of new cross-bridged tetraamine ligands. *Chem. Commun.* 947–948 (1996).

63. Wong, E. H. *et al.* Synthesis and characterization of cross-bridged cyclams and pendant-armed derivatives and structural studies of their copper(II) complexes. *J. Am. Chem. Soc.* **122**, 10561–10572 (2000).

64. Sun, X. *et al.* Radiolabeling and in vivo behavior of copper-64-labeled cross-bridged cyclam ligands. *J. Med. Chem.* **45**, 469–477 (2002).

65. Niu, W. *et al.* Structural and Dynamic Studies of Zinc , Gallium , and Cadmium Complexes of a Dicarboxylate Pendant-Armed Cross-Bridged Cyclen. *Eur. J. Inorg. Chem.* 3310–3315 (2004). doi:10.1002/ejic.200400075

66. Odendaal, A. Y. A modified synthesis , C-functionalization , resolution and racemization kinetics of cross - bridged tetraazamacrocycles. *Dr. Diss.* (2009).

67. Rodríguez-Rodríguez, A. *et al.* Lanthanide(III) complexes with a reinforced cyclam ligand show unprecedented kinetic inertness. *J. Am. Chem. Soc.* **136**, 17954–17957 (2014).

68. Grenier, L. *et al.* Highly Stable and Inert Complexation of Indium(III) by Reinforced Cyclam Dipicolinate and a Bifunctional Derivative for Bead Encoding in Mass Cytometry. *Chem. - A Eur. J.* **25**, 15387–15400 (2019).

69. Rodríguez-Rodríguez, A. *et al.* Complexation of Ln3+ Ions with Cyclam Dipicolinates: A Small Bridge that Makes Huge Differences in Structure, Equilibrium, and Kinetic Properties. *Inorg. Chem.* **55**, 2227–2239 (2016).

70. Lima, L. M. P. *et al.* Monopicolinate cross-bridged cyclam combining very fast complexation with very high stability and inertness of its copper(II) complex. *Inorg. Chem.* **53**, 5269–5279 (2014).

71. Zeng, D., Ouyang, Q., Cai, Z., Xie, X. Q. & Anderson, C. J. New cross-bridged cyclam derivative CB-TE1K1P, an improved bifunctional chelator for copper radionuclides. *Chem. Commun.* **50**, 43–45 (2014).

72. Camus, N. *et al.* Full control of the regiospecific N-functionalization of C-functionalized cyclam bisaminal derivatives and application to the synthesis of their TETA, TE2A, and CB-TE2A analogues. *J. Org. Chem.* **79**, 1885–1899 (2014).

73. Dai, L. *et al.* Chiral DOTA chelators as an improved platform for biomedical imaging and therapy applications. *Nat. Commun.* **9**, 1–10 (2018).

74. Aime, S. *et al.* Conformational and Coordination Equilibria on DOTA Complexes of Lanthanide Metal Ions in Aqueous Solution Studied by ^{1}H-NMR Spectroscopy. *Inorg. Chem.* **36**, 2059–2068 (1997).

75. Wang, X. *et al.* A Kinetic Investigation of the Lanthanide DOTA Chelates. Stability and Rates of Formation and of Dissociation of a Macrocyclic Gadolinium(III) Polyaza Polycarboxylic MRI Contrast Agent. *Inorg. Chem.* **31**, 1095–1099 (1992).

76. Nielsen, L. G., Junker, A. K. R. & Sørensen, T. J. Composed in the f-block: solution structure and function of kinetically inert lanthanide(III) complexes. *Dalt. Trans.* **47**, 10360–10376 (2018).

77. Joss, D., Bertrams, M.-S. & Häussinger, D. A sterically overcrowded, isopropyl-substituted lanthanide chelating tag for protein PCS NMR spectroscopy: Synthesis of its macrocyclic scaffold and benchmarking on ubiquitin S57C and hCA II S166C. *Chem. – A Eur. J.* chem.201901692 (2019). doi:10.1002/chem.201901692

78. Parker, D., Dickins, R. S., Puschmann, H., Crossland, C. & Howard, J. A. K. Being excited by lanthanide coordination complexes: Aqua species, chirality, excited-state chemistry, and exchange dynamics. *Chem. Rev.* **102**, 1977–2010 (2002).

79. Su, X. C. & Chen, J. L. Site-Specific Tagging of Proteins with Paramagnetic Ions for Determination of Protein Structures in Solution and in Cells. *Acc. Chem. Res.* **52**, 1675–1686 (2019).

80. Hikone, Y. *et al.* A new carbamidemethyl-linked lanthanoid chelating tag for PCS NMR spectroscopy of proteins in living HeLa cells. *J. Biomol. NMR* **66**, 99–110 (2016).

81. Dondoni, A. The Emergence of Thiol – Ene Coupling as a Click Process for Materials and Bioorganic Chemistry. *Angew. Chemie - Int. Ed.* 8995–8997 (2008). doi:10.1002/anie.200802516

82. Staderini, S., Chambery, A., Marra, A. & Dondoni, A. Free-radical

hydrothiolation of glycals : a thiol-ene-based synthesis of S -disaccharides. *Tetrahedron Lett.* **53**, 702–704 (2012).

83. Ma, F. H. *et al.* Kinetic assay of the michael addition-like thiol-ene reaction and insight into protein bioconjugation. *Chem. - An Asian J.* **9**, 1808–1816 (2014).

84. Dondoni, A., Massi, A., Nanni, P. & Roda, A. A new ligation strategy for peptide and protein glycosylation: Photoinduced thiol-ene coupling. *Chem. - A Eur. J.* **15**, 11444–11449 (2009).

85. Lee, M. D. *et al.* Compact, hydrophilic, lanthanide-binding tags for paramagnetic NMR spectroscopy. *Chem. Sci.* **6**, 2614–2624 (2015).

86. Li, Q.-F., Yang, Y., Maleckis, A., Otting, G. & Su, X.-C. Thiol–ene reaction: a versatile tool in site-specific labelling of proteins with chemically inert tags for paramagnetic NMR. *Chem. Commun.* **48**, 2704 (2012).

87. Su, X. C. *et al.* A dipicolinic acid tag for rigid lanthanide tagging of proteins and paramagnetic NMR spectroscopy. *J. Am. Chem. Soc.* **130**, 10486–10487 (2008).

88. Parker, D., Puschmann, H., Batsanov, A. S. & Senanayake, K. Structural Analysis of Nine-Coordinate Lanthanide Complexes: Steric Control of the Metal-Water Distance Across the Series. *Inorg. Chem.* **42**, 8646–8651 (2003).

89. Chin, K. O. A., Morrow, J. R., Lake, C. H. & Churchill, M. R. Synthesis and Solution Properties of Lanthanum(III), Europium(III), and Lutetium(III) THP Complexes and an X-ray Diffraction Study of a Crystal Containing Four Stereoisomers of a Europium (III) THP Complex (THP = 1,4,7,10-Tetrakis(2-hydroxypropyl)-1,4,7,1. *Inorg. Chem.* **33**, 656–664 (1994).

90. M. Lelli, G. Pintacuda, A. Cuzzola, L. B. Monitoring Proton Dissociation and Solution Conformation of Chiral Ytterbium Complexes with Near-IR CD. *Chirality* **211**, 201–211 (2005).

91. I. Bertini, C. Luchinat, G. P. 1H NMRD profiles of paramagnetic complexes and metalloproteins. *Adv. Inorg. Chem.* **57**, 105–172 (2005).

92. Bertini, I., Luchinat, C., Parigi, G. & Ravera, E. *Relaxometry and contrast agents for MRI*. *Solution NMR of Paramagnetic Molecules* (2017). doi:10.1016/b978-0-444-63436-8.00010-7

93. Toth, E. et al. Tuning water-exchange rates on (carboxymethyl)iminobis-(ethylenenitrilo)tetraacetate (dtpa)-type gadolinium(. *J. Chem. soc. Dalt. Trans.* 1587–1594 (1997).

94. Clough, T. J., Jiang, L., Wong, K. L. & Long, N. J. Ligand design strategies to increase stability of gadolinium-based magnetic resonance imaging contrast agents. *Nat. Commun.* **10**, 1–14 (2019).

95. Yagi, H. *et al.* Tunable paramagnetic relaxation enhancements by [Gd(DPA)3] 3- for protein structure analysis. *J. Biomol. NMR* **47**, 143–153 (2010).

96. Roca-Sabio, A. *et al.* Macrocyclic receptor exhibiting unprecedented selectivity

for light lanthanides. *J. Am. Chem. Soc.* **131**, 3331–3341 (2009).

97. Xia, J., Zhang, X. & Matyjaszewski, K. Atom transfer radical polymerization of 4-vinylpyridine. *Macromolecules* **32**, 3531–3533 (1999).

98. Pierre, V. C., Botta, M., Aime, S. & Raymond, K. N. Tuning the coordination number of hydroxypyridonate-based gadolinium complexes: Implications for MRI contrast agents. *J. Am. Chem. Soc.* **128**, 5344–5345 (2006).

99. Buhaya, M. H., Galvan, S. & Maldonado, R. A. Incidence of Trypanosoma cruzi infection in triatomines collected at Indio Mountains Research Station. *Acta Trop.* **150**, 97–99 (2015).

100. Chitanga, S. *et al.* High prevalence of drug resistance in animal trypanosomes without a history of drug exposure. *PLoS Negl. Trop. Dis.* **5**, (2011).

101. Haanstra, J. R., González-Marcano, E. B., Gualdrón-López, M. & Michels, P. A. M. Biogenesis, maintenance and dynamics of glycosomes in trypanosomatid parasites. *Biochim. Biophys. Acta - Mol. Cell Res.* **1863**, 1038–1048 (2016).

102. Emmanouilidis, L., Gopalswamy, M., Passon, D. M., Wilmanns, M. & Sattler, M. Structural biology of the import pathways of peroxisomal matrix proteins. *Biochim. Biophys. Acta - Mol. Cell Res.* **1863**, 804–813 (2016).

103. Brul, S. *et al.* Genetic heterogeneity in the cerebrohepatorenal (Zellweger) syndrome and other inherited disorders with a generalized impairment of peroxisomal functions. A study using complementation analysis. *J. Clin. Invest.* **81**, 1710–1715 (1988).

104. Moyersoen, J., Choe, J., Fan, E., Hol, W. G. J. & Michels, P. A. M. Biogenesis of peroxisomes and glycosomes: Trypanosomatid glycosome assembly is a promising new drug target. *FEMS Microbiol. Rev.* **28**, 603–643 (2004).

105. Bottger, G. *et al.* Saccharomyces cerevisiae PTS1 receptor Pex5p interacts with the SH3 domain of the peroxisomal membrane protein Pex13p in an unconventional, Non-PXXP-related manner. *Mol. Biol. Cell* **11**, 3963–3976 (2000).

106. Freitas, M. O. *et al.* PEX5 protein binds monomeric catalase blocking its tetramerization and releases it upon binding the N-terminal domain of PEX14. *J. Biol. Chem.* **286**, 40509–40519 (2011).

107. Neufeld, C. *et al.* Structural basis for competitive interactions of Pex14 with the import receptors Pex5 and Pex19. *EMBO J.* **28**, 745–754 (2009).

108. Dawidowski, M. *et al.* Inhibitors of PEX14 disrupt protein import into glycosomes and kill Trypanosoma parasites. *Science (80-.).* **355**, 1416–1420 (2017).

109. Bertini, I. *et al.* The catalytic domain of MMP-1 studied through tagged lanthanides. *FEBS Lett.* **586**, 557–567 (2012).

110. Ottiger, M., Delaglio, F. & Bax, A. Measurement of J and Dipolar Couplings from Simplified Two-Dimensional NMR Spectra. *J. Magn. Reson.* **131**, 373–378 (1998).

111. Rochus L.J. Keller. *The Computer Aided Resonance Assignment Tutorial*.

112. Nilges, M. & Markwick, P. R. L. Structural Biology by NMR : Structure , Dynamics , and Interactions. *PLoS Comput. Biol.* **4**, (2008).

113. Carlon, A. *et al.* Improved Accuracy from Joint X-ray and NMR Refinement of a Protein-RNA Complex Structure. *J. Am. Chem. Soc.* **138**, 1601–1610 (2016).

114. Rinaldelli, M. *et al.* Simultaneous use of solution NMR and X-ray data in REFMAC5 for joint refinement/detection of structural differences. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70**, 958–967 (2014).

115. Bertini, I. *et al.* MaxOcc: A web portal for maximum occurrence analysis. *J. Biomol. NMR* **53**, 271–280 (2012).

116. Ravera, E., Sgheri, L., Parigi, G. & Luchinat, C. A critical assessment of methods to recover information from averaged data. *Phys. Chem. Chem. Phys.* **18**, 5686–5701 (2016).

117. Russo, L., Maestre-Martinez, M., Wolff, S., Becker, S. & Griesinger, C. Interdomain dynamics explored by paramagnetic NMR. *J. Am. Chem. Soc.* **135**, 17111–17120 (2013).

118. Nitsche, C. & Otting, G. Pseudocontact shifts in biomolecular NMR using paramagnetic metal tags. *Prog. Nucl. Magn. Reson. Spectrosc.* **98**–**99**, 20–49 (2017).

119. Cerofolini, L. *et al.* Long - range paramagnetic NMR data can provide a closer look on metal coordination in metalloproteins. *JBIC J. Biol. Inorg. Chem.* **23**, 71–80 (2018).

120. Baldoneschi, V. *et al.* Active-Site Targeting Paramagnetic Probe for Matrix Metalloproteinases. *Chempluschem* **81**, 1333–1338 (2016).

121. Luca Frullano, P. C. Strategies for the Preparation of Bifunctional Gadolinium(III) Chelators. *Curr. Org. Synth.* **8**, 535–565 (2011).

122. Miao, Q. *et al.* A Double-Armed, Hydrophilic Transition Metal Complex as a Paramagnetic NMR Probe. *Angew. Chemie Int. Ed.* **58**, 13093–13100 (2019).

123. Müntener, T., Kottelat, J., Huber, A. & Häussinger, D. New Lanthanide Chelating Tags for PCS NMR Spectroscopy with Reduction Stable, Rigid Linkers for Fast and Irreversible Conjugation to Proteins. *Bioconjug. Chem.* 1–46 (2018). doi:10.1021/acs.bioconjchem.8b00512

124. Yang, F., Wang, X., Pan, B.-B. & Su, X.-C. Single-armed phenylsulfonated pyridine derivative of DOTA is rigid and stable paramagnetic tag in protein analysis. *Chem. Commun.* **52**, 11535–11538 (2016).

125. Norio Miyaura, A. S. Stereoselective synthesis of arylated (E)-alkenes by the reaction of alk-1-enylboranes with aryl halides in the presence of palladium catalyst. *J. Chem. Soc., Chem. Commun* 866–867 (1979).

126. Parigi, G., Ravera, E. & Luchinat, C. Magnetic susceptibility and

paramagnetism-based NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **114**–**115**, 211–236 (2019).

127. Shishmarev, D. & Otting, G. How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study. *J. Biomol. NMR* **56**, 203–216 (2013).

128. Jeremy P.Derrick, D. B. W. The Third IgG-Binding Domain from Streptococcal Protein G: An Analysis by X-ray Crystallography of the Structure Alone and in a Complex with Fab. *J. Mol. Biol.* **243**, 906–918 (1994).

129. Gempf, K. L., Butler, S. J., Funk, A. M. & Parker, D. Direct and selective tagging of cysteine residues inpeptides and proteins with 4-nitropyridyl lanthanide complexes. *Chem. Commun.* 9104–9106 (2013). doi:10.1039/c3cc45875j

130. Ravera, E., Parigi, G. & Luchinat, C. What are the methodological and theoretical prospects for paramagnetic NMR in structural biology? A glimpse into the crystal ball. *J. Magn. Reson.* **306**, 173–179 (2019).

131. Fragai, M., Ravera, E., Tedoldi, F., Luchinat, C. & Parigi, G. Relaxivity of Gd-Based MRI Contrast Agents in Crosslinked Hyaluronic Acid as a Model for Tissues. *ChemPhysChem* **20**, 2204–2209 (2019).

132. Wilton, D. J., Tunnicliffe, R. B., Kamatari, Y. O., Akasaka, K. & Williamson, M. P. Pressure-induced changes in the solution structure of the GB1 domain of protein G. 1432–1440 (2007). doi:10.1002/prot.21832