



Universiteit
Leiden
The Netherlands

Applications of paramagnetic NMR spectroscopy for protein research

Lescanne, M.

Citation

Lescanne, M. (2018, May 29). *Applications of paramagnetic NMR spectroscopy for protein research*. Retrieved from <https://hdl.handle.net/1887/62455>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/62455>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/62455> holds various files of this Leiden University dissertation

Author: Lescanne, Mathilde

Title: Applications of paramagnetic NMR spectroscopy for protein research

Date: 2018-05-29

Applications of paramagnetic NMR spectroscopy for protein research

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C. J. J. M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 29 mei 2018
klokke 13:45 uur

door

Mathilde Lescanne

geboren te Bordeaux, Frankrijk
in 1989

Applications of paramagnetic NMR spectroscopy for protein research

Mathilde Lescanne

Doctoral thesis, Leiden University, 2018

Cover image: Cross section of PCS-isosurfaces of ntd-HSP90 V150 methyl groups

ISBN number: 978-94-91546-15-0

©2018, Mathilde Lescanne

Printed by Boekdruken.com, The Netherlands

Promotores: Prof. Dr. M. Ubbink
Prof.Dr. C. Luchinat (University of Florence, Italy)

Promotiecommissie: Prof. Dr. H.J.M De Groot
Prof. Dr. H.S. Overkleeft
Prof. Dr. D. Marion (Institute of Structural Biology of
Grenoble, France)
Dr. A. Pandit
Dr. W.M. Liu (Fu Jen Catholic University, Taiwan)

The research described in this thesis was performed in the Macromolecular Biochemistry Department of Leiden Institute of Chemistry at the Universiteit Leiden in Leiden, the Netherlands.

The research described in this thesis was financially supported by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n°317127.

*Quand le vent est au sud écoutez-le chanter
Le plat pays qui est le mien*

Jacques Brel Le plat pays

*Wanneer de zuidenwind er jubelt langs de baan
Dan juicht mijn land, mijn vlakke land*

Jacques Brel/Ernst van Altena Mijn vlakke land

To DALAN

TABLE OF CONTENTS

	List of abbreviations	6
	List of symbols	7
Chapter 1	Introduction	9
Chapter 2	Methyl group assignment using pseudocontact shifts with PARAssign	25
Chapter 3	Lanthanoid generated RDC for amide assignment with PARAssign	49
Chapter 4	Methyl group re-orientation under ligand binding probed by pseudo-contact shifts	55
Chapter 5	Flexibility in a ubiquitin dimer	73
Chapter 6	Multimeric paramagnetic tagging: Proteasome	83
Chapter 7	General discussion and outlook	93
	Reference list	99
	English summary	118
	Nederlandse samenvatting	120
	List of Publications	123
	Curriculum vitae	124

LIST OF ABBREVIATIONS

EDTA	Ethylenediaminetetraacetic acid
Ntd-HSP90	N-terminal domain of heat shock protein 90
HSQC	heteronuclear single quantum coherence
HMQC	heteronuclear multiple quantum coherence
CSP	chemical shift perturbation
Ln	lanthanoid
CLaNP-5	Caged lanthanoid NMR probe 5
CLaNP-7	Caged lanthanoid NMR probe 7
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
PCS	pseudo-contact shift
pNMR	paramagnetic NMR
PDB	protein data bank
PRE	paramagnetic relaxation enhancement
RACS	residual anisotropic chemical shift
RDC	residual dipolar coupling
TROSY	transverse relaxation-optimized spectroscopy
Ub	Ubiquitin
Ub2	Ubiquitin dimer
Ub2A	¹⁵ N-distal-labeled K48-Ub2 with distal Ub being K48R variant and proximal Ub being the A28C/D32C/D77 variant
Ub2B	¹⁵ N-proximal-labeled K48-Ub2 with distal Ub being the A28C/D32C/K48R variant, and proximal Ub being the D77 variant
WT	wild-type

LIST OF SYMBOLS

B_0	static magnetic field
g_e	electronic g-factor
\hbar	reduced Planck's constant
k_B	Boltzmann constant
r_{AB}	distance between nucleus A and nucleus B
S	spin quantum number
S_{LZ}	Lipari-Szabo parameter
T	temperature
γ_A	gyromagnetic ratio of nucleus A
$\Delta\chi_{ax}$	axial component of $\Delta\chi$ tensor
$\Delta\chi_{rh}$	rhombic component of $\Delta\chi$ tensor
μ_0	vacuum permeability
τ_c	total correlation time $\tau_c^{-1} = \tau_r^{-1} + \tau_e^{-1}$
τ_r	rotational correlation time
τ_e	longitudinal electronic relaxation time
β	Bohr magneton

Chapter 1

Introduction

Challenges in liquid-state NMR spectroscopy of proteins

Liquid-state NMR spectroscopy is widely used to study protein structure, dynamics and interactions with atomic detail. The technique inherently has a low sensitivity, although it has improved greatly over the years with the introduction of cooled electronics (cryo-probes) and very strong magnetic fields. Liquid-state NMR is one of the main techniques to probe the protein world, together with solid-state NMR spectroscopy, electron paramagnetic resonance (EPR) spectroscopy, X-ray diffraction of crystals (XRD), small-angle X-ray scattering (SAXS) and cryo-electron microscopy (cryoEM). NMR remains a high-expertise technique requiring expensive high-field spectrometers and expert researchers. Despite the great resolution of NMR spectroscopy, liquid-state NMR spectra quality remains highly dependent on the size of the protein or protein complex observed. Indeed, the bigger the system is, the higher the rotational correlation time τ_c is and the less the averaging of the dipole-dipole interactions are. Consequently, relaxation increases, leading to peak broadening. Moreover, bigger systems have more nuclei, so more signals are observed, rendering spectra very crowded. Oligomerization gives rise to symmetry, which means that a resonance cannot be attributed to a single nucleus in the system. Dynamic systems represent a challenge for NMR, as the acquired spectrum is the average resultant of all the sub-states visited by the protein in case of fast exchange or an overlap of spectra representative of each visited sub-state in case of slow exchange.

Assignment

NMR assignment is the exercise of matching resonance peaks to nuclei, **Figure 1.1**. In protein NMR the most frequently assigned nuclei are ^1H , ^{15}N and ^{13}C . In the nineteen seventies, multidimensional NMR started to be developed and NMR assignment benefited from 2D homonuclear NMR sequences such as COSY^{7,12}, TOCSY^{3,4} and NOESY¹ sequences. COSY and TOCSY use scalar $^3J_{\text{HH}}$ coupling to create magnetization modulated with two chemical shifts. While COSY experiments are ideal for small molecules presenting a simple coupling network, TOCSY extends magnetization transfer to spins without direct coupling, permitting the study of more complex coupling networks, making it useful for identifying protons in amino acids. At the same time, NOE-based experiments were developed, transferring magnetization through space, via dipolar coupling. Being very useful to identify spins undergoing cross-relaxation, i.e. spins close to each other ($< 5 \text{ \AA}$), NOESY experiments were used in combination with COSY and TOCSY for assignment purposes. Additionally, NOE were and are still used as structural restraints for protein structure calculation as a NOE decreases with r^6 , with r the distance between the two spins^{5,6}. In the nineteen eighties, heteronuclear coupling based experiments were designed and applied for the assignment of calmodulin (16.7 kDa) in 1990, using a combination of four through bond experiments HNC O ^{7,9}, HNCA^{7,8,10}, HCACO^{11,12} and HCA(CO)N^{6,12}. Through-bond

magnetization transfer to neighboring residues (instead of via NOE) and the use of J scalar couplings, which are much larger than J_{NH} , enabled efficient transfer magnetization. H_nNCO correlates the amide ^1H and ^{15}N shifts with ^{13}C shift of the carbonyl resonance of the preceding amino acid. H_nNCA correlates the amide ^1H and ^{15}N to the $\text{C}\alpha$ chemical shift from the same residue, and a weak correlation with the $\text{C}\alpha$ of the previous residue can also be observed. H_nCACO correlates the $\text{H}\alpha$ and $\text{C}\alpha$ with the carbonyl shift of the same residue. Finally, $\text{H}_n\text{CA}(\text{CO})\text{N}$ correlates $\text{H}\alpha$ and $\text{C}\alpha$ of one residue with ^{15}N of the following residue. In 1991, Ad Bax et al., published a new procedure to correlate the backbone atoms with the preceding residue using a combination of $\text{H}_n\text{N}(\text{CO})\text{CA}$ and H_nNCA .¹²

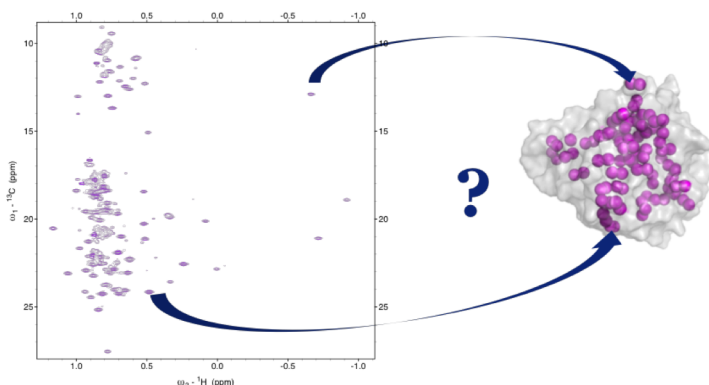


Figure 1.1. Assignment problem for methyl groups. On the right, the protein surface and the methyl groups are represented in grey and in purple spheres, respectively. On the left, the NMR spectrum correlating the ^1H to the ^{13}C of the methyl groups.

These NMR breakthroughs, were possible by the availability of uniformly ^{15}N and ^{13}C labelled proteins¹³. The protein is produced by recombinant gene expression in bacteria that are cultured on minimal medium, supplemented with $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -glucose as sole source of nitrogen and carbon, respectively. These advances enabled setting up routine experiments for protein assignment based on the amino acid sequence of the protein, called sequential assignment. A combination of NMR experiments such as $\text{CBCANH}/\text{HNCACB}$ ¹⁴ with $\text{CBCA}(\text{CO})\text{NH}/\text{HN}(\text{CO})\text{CACB}$ ¹⁴ and ^{15}N - ^1H HSQC enable the correlation of resonances of nuclei of residue i with those of residue $i-1$ and therefore, perform a sequential assignment of backbone amide, CA and CB chemical shifts. This method is very efficient and well-described and it benefits from semi-automated assignment procedures (such as AutoAssign^{15,16}, PINE^{16,17}), compatible with several NMR spectra analysis software packages, like CYANA^{18,19}, CCPNMR²⁰ and SPARKY²¹. Yet, this robust approach is limited to small proteins, for which tumbling time is fast enough to prevent extensive loss of signal during the long spin preparation periods in the multidimensional experiments. Liquid-state sequential NMR assignment becomes challenging for proteins larger than 25 kDa. To circumvent

this size limitation, different strategies have been developed over the years to make liquid-state NMR suitable to study proteins up to half a mega-dalton²². The first improvement is the use of mobile NMR observable probes, like methyl groups. From a spectroscopic point of view, the $^{13}\text{CH}_3$ methyl groups are valuable probes in very large systems because their inherent fast rotation (picosecond timescale)²³. Second, transverse relaxation optimized spectroscopy (TROSY), described by K. Pervushin and K. Wüthrich in 1997²⁴, and methyl-TROSY experiments²⁵, developed by Lewis Kay group in 2003, represented a breakthrough in liquid-state NMR and enabled studying proteins much larger than 25 kDa. Third, perdeuteration used to significantly decrease line broadening, is a major step forward. Additionally, the non-polar nature of methyl groups makes them suitable to probe the core of the protein²⁶ and study specific interactions they may have with other proteins, ligands or substrates²⁷. Thus, methyl groups have become important in bioNMR to probe biomolecular interactions²³ as well as the dynamics of proteins over a large range of timescales, from picoseconds to seconds²⁸. Furthermore, selective isotope labeling has become an established technique to diminish the crowding of the spectra. Various approaches have been developed to cover all possible labeling combinations and levels of protonation. The specific isotope labeling on the isoleucine, leucine and valine (ILV) methyl groups is an well-known approach²⁹ to reduce spectral crowding and to provide good spatial coverage, due to dispersion of these residues over the protein structure. Therefore, the use of methyl group resonances as NMR observables, possibly in combination with perdeuteration and methyl-TROSY experiments paved the way to the application of NMR to systems of up to hundreds of kDa. However, the assignment of ILV methyl groups is, non-trivial. To take advantage of standard J -coupling based assignments procedures discussed above for methyl group assignments, uniformly labeled samples are needed, requiring the production of additional samples and time-consuming 3D NMR experiments (if they work at all). Various alternative approaches have been published to assign methyl groups in large proteins, Table 1.1. Systematic site-directed mutagenesis, combined with stereospecific labeling on the methyl groups has been applied to TET2^{30,31} labeled at pro-S and pro-R valines methyl groups. Stereo-specific labelling has proved to be a powerful tool to limit the overcrowding of large protein spectra and was used in large proteins³¹⁻³³ and developed for cell-free protein expression³⁴. To circumvent the size limitation, the so-called “divide and conquer” approach also proved to be effective. Here, individual domains or subunits are first assigned, followed by a transfer of the assignments to the spectrum of the larger multidomain protein or multisubunit protein complex^{31,35,36}. Another approach is to use multidimensional NOESY spectra in combination with a crystal structure³⁷. This approach has been implemented in the MAP-XSII software³⁸ that automatically assigns methyl groups based on a known structure combined with NOEs measured on 4D ^1H - ^{13}C -HSQC-NOESY-HSQC experiments. Similarly, the program FLAMEnGO uses a

combination of 3D NOESY, methine-methyl TOCSY and paramagnetic relaxation enhancements (PRE) and a protein structure to assign methyl groups³⁹. Yet, in many cases, multi-dimensional experiments can lead to significant loss of signal. Therefore, assignment using only 2D experiments would be a powerful tool for such fast-relaxing proteins. Moreover, despite the access to non-linear sampling NMR, time-consuming multi-dimensional NMR experiments might not be suitable for non-stable proteins. The possibility to assign protein NMR spectra with sensitive and fast 2D HSQC experiments would be of interest for such challenging systems. Pseudocontact shifts, originating from a paramagnetic center, measured on 2D spectra have been used for assignment purposes for amide and methyl groups. Either used as an additional piece of information⁴⁰, or exclusively, with computer programs like Possum⁴¹ and PARAssign^{42,43}, PCS have proven to be useful and efficient to assign protein spectra. PARAssign software automatically assigns 2D spectra using PCS and a 3D structure, without any prior knowledge of the magnetic susceptibility tensor.

Table 1.1: NMR assignment software and methods.

Method	NMR experiments	Type of input data	assignable groups	Output	Pros	Cons
PARAssign	2D-HSQC	PCS/ RDC	Amides Methyl groups	Assignment Susceptibility tensors	No prior knowledge needed	Multiple paramagnetic centers
MAP-XSII	3D and 4D NOESY	NOE, PRE, RDC, PCS	Methyl groups	Assignment	No mutation needed	High-dimensional NMR
FlaMEngo	NOESY, TOCSY	NOE, PRE	Methyl groups	Assignment	User-friendly interface	Low sensitive experiment
Platypus	HSQC	PCS, PRE, RDC, CCR	Amides	Assignment	Multiple input kind	No tensor refinement
Possum	HCCH-TOCSY, HSQC	Chemical shifts	Methyl groups	Assignment	Methyl connectivity	No user-friendly interface
Directed-mutagenesis	HSQC	Chemical shifts	all		Highly reliable	Tedious and sample expensive
Divide and conquer	HSQC	Chemical shifts	all		Simplification to routine assignment	Only for multimer
Traditional 3D through backbone NMR	3D through backbone	Chemical shifts	all		Routine technique	Limited to relatively small proteins

Even though the size of the proteins reachable for assignment has increased tremendously with years, for large systems, methyl assignment is further complicated by

the necessary fractional or complete deuteration of the other hydrogen positions to reduce relaxation by dipolar couplings. The broad range of available techniques for assignment enables the user to choose the technique most suitable for the molecule of interest.

Spectral crowding

As seen in the previous paragraph, the development of multi-dimensional NMR was possible thanks to the development of isotopic labeling. From uniform ^{15}N labeling to selectively labelled stereospecific methyl groups in a deuterated background, labeling for protein NMR has become a research field in itself to bring further NMR studies on larger proteins. Uniform ^{15}N labeling is the easiest and cheapest kind of labeling and was used for the first times in 1990¹³ by McIntosh¹⁸. In the same publication, uniform ^{13}C labeling was described. A bit earlier in 1988, protein deuteration had been shown to enhance spectra resolution^{14,15}. Since then, deuteration has been used successfully for the determination of structures of proteins larger than 20 kDa¹⁶. On one hand, the signal-to-noise is improved by suppressing spin diffusion between neighboring ^1H replaced by ^2H . On the other hand, the relaxation rates of ^{13}C and ^{15}N are scaled proportional to $(\gamma_D/\gamma_H)^2 \approx 0.02$. This was used later to obtain narrower linewidths in methyl group experiments by observing CHD_2 instead of CH_3 ⁴⁷⁻⁴⁹ after the development of these isotopomers^{26,30}. Precursors for methyl group labeling for a wide range of amino-acid containing methyl groups³¹⁻³⁶ are now available and efforts have been carried out to reduce crowding spectra by labelling stereospecifically^{35,36}.

However, a disadvantage of methyl groups spectra is that they show less dispersion than those of amides. PCS can improve the methyl resonance dispersion, as will be discussed in chapter 4.

Multi-domain dynamics

Dynamics studies with NMR benefited from improvements relating to both assignment and labeling approaches. NMR can provide information about the dynamics at a wide range of timescales, from nanoseconds to hours. CPMG and CEST experiments have proven to bring highly valuable information to millisecond timescale for intrinsic motions. In the case of multi-domain proteins and encounter complexes, the relative orientation of domains or protein partners is being studied. In some cases, a single configuration is observed but in many cases one conformation does not agree with the observations. Rather, the system needs to be described by an ensemble of conformations, to match the experimental NMR observables. NOE, widely used for structure refinement and characterization of ligand binding^{57,58} are not suitable for study ensembles. The development of high fields, up to 1000 MHz nowadays, contributed to the development of RDC for structure elucidation and domain orientation

determination. And more generally new methods based on paramagnetic effects have shown to be very suitable for the study of such ensembles^{59,60}.

Paramagnetic effects

To make NMR experiments possible, nuclei of interest must have a magnetic moment. An unpaired electron also has a magnetic moment, about 658 times stronger than the one from hydrogen. The interactions of electron spins with nuclei generate the so-called paramagnetic effects. These effects are measurable with NMR spectroscopy. Well-described since the nineteen fifties and the nineteen sixties, they have been observed first in the protein NMR field in metalloproteins. Initially the effects were perceived as hindrance to NMR studies because of their contribution to line broadening, but later paramagnetic effects have been increasingly recognized as a useful NMR tool to probe protein structure, conformation and dynamics.

The magnetization of a molecule per unit volume M is proportional to the applied magnetic field H_0 : $M = \chi_v H_0$. χ_v is called magnetic susceptibility per unit volume. The magnetic susceptibility per mole is defined as $\chi_M = V_M \chi_v$. According to the Curie law, **equation 1.1**:

$$\chi_M = \mu_0 N_A \mu_B^2 g^2 \frac{S(S+1)}{3kT} \quad (1.1)$$

A magnetic susceptibility per molecule χ can be defined as the molar magnetic susceptibility divided by the Avogadro constant $\chi = \frac{\chi_M}{N_A}$. The anisotropy of the magnetic susceptibility originates from the electron orbital magnetic moment contribution. The magnetic susceptibility tensor parameters are difficult to predict from theory but the effects, easily measured by NMR, are used to determine these parameters.

Tensor representation and conventions

χ can be represented by a tensor, useful to describe properties that depend on directions. The susceptibility tensor, $\boldsymbol{\chi}$, is represented by a 3x3 matrix. In its principal frame the susceptibility tensor is diagonal. The tensor is divided in isotropic and anisotropic components and can be written as a sum of an isotropic and an anisotropic sub-tensor,

$$\boldsymbol{\chi} = \begin{bmatrix} \chi_{xx} & 0 & 0 \\ 0 & \chi_{yy} & 0 \\ 0 & 0 & \chi_{zz} \end{bmatrix} = \begin{bmatrix} \chi_{iso} & 0 & 0 \\ 0 & \chi_{iso} & 0 \\ 0 & 0 & \chi_{iso} \end{bmatrix} + \Delta\boldsymbol{\chi}$$

$$\chi = \begin{bmatrix} \chi_{iso} & 0 & 0 \\ 0 & \chi_{iso} & 0 \\ 0 & 0 & \chi_{iso} \end{bmatrix} + \begin{bmatrix} -\frac{1}{3\Delta\chi_{ax}} + \frac{1}{2\Delta\chi_{rh}} & 0 & 0 \\ 0 & -\frac{1}{3\Delta\chi_{ax}} - \frac{1}{2\Delta\chi_{rh}} & 0 \\ 0 & 0 & \frac{2}{3\Delta\chi_{ax}} \end{bmatrix}$$

with: $\Delta\chi_{ax} = \chi_{zz} - \frac{1}{2}(\chi_{xx} + \chi_{yy})$ and $\Delta\chi_{rh} = (\chi_{xx} - \chi_{yy})$

$\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ represent the axial and rhombic component of the tensor and thus define the shape of the tensor effects in space. While the isotropic part component χ_{iso} alters the Curie-spin relaxation rates R_1 and R_2 , the anisotropic component $\Delta\chi$ causes contact shifts, pseudo-contact shift (PCS), residual anisotropic chemical shift (RACS) and residual dipolar couplings (RDC). Eight parameters are necessary to fully describe a $\Delta\chi$ in space, the three position coordinates x , y and z of the paramagnetic center, the three Euler angles α , β , γ , used to rotate from a given frame (for instance the protein molecular frame) to the principal tensor frame, the axial component and the rhombic component of the $\Delta\chi$ tensor. In this thesis, the Euler angles are defined with the convention z-y-z with $\alpha \in [-\pi, \pi]$, $\beta \in [-\pi/2, \pi/2]$ and $\gamma \in [-\pi, \pi]$. $\Delta\chi$ can be represented by different manner. A representation by PCS iso-surfaces, as seen in **Figure 1.2**, illustrates the magnitude of $\Delta\chi$.

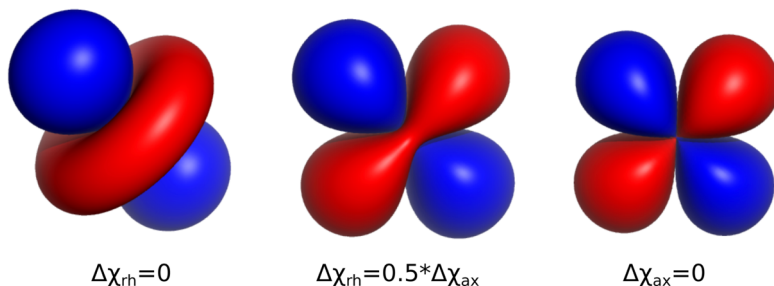


Figure 1.2. Tensor representation in space for a strictly axial, axial and rhombic and strictly rhombic tensor, from the left to the right respectively. The blue and red surfaces represent a positive and a negative PCS isosurface of the same absolute intensity, respectively.

$\Delta\chi$ can also be represented by its three main axes in a sinusoidal projection, with longitude λ and latitude φ . Such a representation illustrates the orientation of the tensor relative to the initial protein frame, x , y , z axes are shown in red, green, and blue throughout this thesis, **Figure 1.3**.

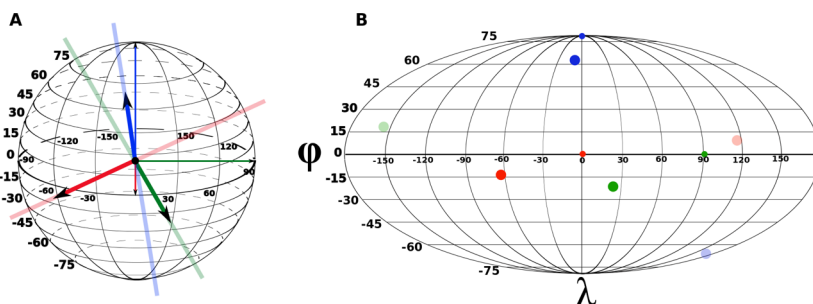


Figure 1.3. **a.** Origin axes and rotated frame axes in a 3D representation, in thin bright lines and in transparent lines with bright lines and arrows on top, respectively **b.** Origin axes are represented in small bright spheres and rotated frame axes are represented in larger spheres in flat colours, in a sinusoidal projection. For instance, the origin y axis is represented in the sinusoidal projection at the (λ, ϕ) coordinates $(90^\circ, 0^\circ)$.

Lanthanoid based tags

Among the large range of paramagnetic ions, lanthanoids, called in everyday language rare-earth elements, are the f-block elements. F-blocks elements have their 4f electron shell filled (except for the diamagnetic lanthanoids lutetium and lanthanum). Lanthanoids are very useful paramagnetic ions for biochemical purposes. From the diamagnetic lanthanum to the strong paramagnetic terbium, they offer a wide range of paramagnetic effects such as PCS, RDC, PRE and RACS. PCS RDC and RACS provide long range restraints, with PCS being on the order of ppm and RDC and RACS on the order of ppb. PRE are observable up to 45 Å and are particularly suitable for low-populated states study. Sometimes, natural metals can be replaced with lanthanoids⁶¹, but more widely applied are tags that are linked to the protein. Lanthanoid-binding peptides can also be engineered in the protein sequence and present the advantage of avoiding tagging issues^{62,63}. Numerous paramagnetic probes designed for protein chemistry have been published⁶⁴⁻⁶⁸, varying in diversity, by the nature of their attachment and applications that they are used for, from structure calculation to dynamics and, more recently, assignment^{82,43,40,70}. CLaNP-5^{64,65} is a two-armed probe that can be attached via disulfide bridges to cysteines spaced by 6 to 8 Å, **Figure 1.4.** Its two-armed geometry makes its rigidity optimal for PCS measurement, avoiding PCS averaging caused by probe mobility. In addition, CLaNP-5 chelates the whole range of Ln^{3+} ions and chelating it with Lu^{3+} provides a suitable diamagnetic reference.

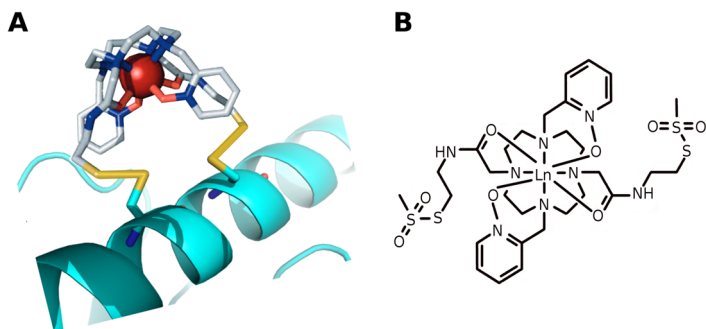


Figure 1.4. a. 3D model of CLaNP-5 attached via disulfide bridges to an alpha helix. b. Chemical structure of CLaNP-5.

Contact and pseudo-contact shifts

The average induced electron magnetic moment gives rise to a contribution to the chemical shift that is called the hyperfine shift. Experimentally, it is measured as the difference observed between the chemical shift of the paramagnetic molecule and the chemical shift of the diamagnetic molecule. It is composed of contact shift and pseudo-contact shift. Contact shifts result from the presence of unpaired electron spin density at the observed nucleus, while pseudo-contact shift result from the through-space dipole-dipole interaction between the electron and the nucleus.

Contact shifts

The contact shift is given by an additional magnetic field generated at the nucleus by the electron magnetic moment located at the nucleus itself, like an orbital contribution. The large coupling between the nucleus and the electron, the hyperfine coupling A , **equation 1.2**, in the order of MHz, is not visible because of the fast electron spin relaxation. Consequently, the nuclear resonance is the average of the two nucleus-electron coupled energy states, weighted according to their spin populations. Given the requirement for delocalisation, a contact shift δ^{con} , **equation 1.3**, is a short-range interaction, usually visible up to a few chemical bonds away from the metal, and is therefore not observable for nuclei in the protein when the paramagnetic ion is chelated in a molecule that is attached to the protein.

$$A = \frac{2\mu_0}{3} \hbar \gamma_I g_e \mu_B \rho \quad (1.2)$$

$$\delta^{\text{con}} = \frac{A g_e \mu_B S(S+1)}{\hbar 3\gamma_I kT} \quad (1.3)$$

Pseudo-contact shifts

A pseudo-contact shift (PCS) is a modification of the chemical shift of nuclei, δ^{PCS} , caused by unpaired electrons, present at the paramagnetic centers (p). The magnetic field arising from the unpaired electron at the position of the nucleus is a consequence of the dipolar through-space interaction between the dipoles of the nucleus and the

electron(s). The probability density of the electron at the nucleus is neglected. The r^{-3} dependence makes the PCS effective over long distances, compared to for example the NOE. A PCS can be described by its spherical coordinates r , θ and φ in the $\Delta\chi$ frame, where r is the distance between the nucleus and p, θ is the angle between the p-nucleus vector and the $\Delta\chi$ z-axis, and φ is the angle between the projection of the p-nucleus vector on the xy plane and the $\Delta\chi$ x-axis, **Figure 1.5**.

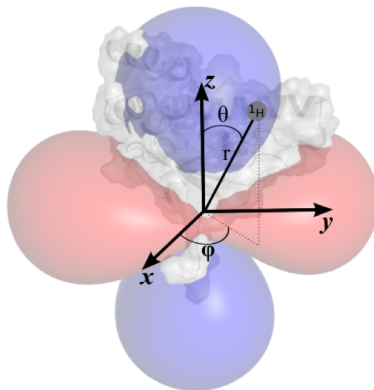


Figure 1.5. Representation of the spherical coordinates r , θ and φ , for a ^1H nucleus in a protein shown in grey. The displayed frame is the $\Delta\chi$ principal frame with the unpaired electron(s) localized at the origin. The red and blue surfaces represent a positive and negative PCS isosurfaces.

$\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic components, respectively, of the anisotropic susceptibility tensor, **equation 1.4**.

$$\delta^{\text{PCS}} = \frac{1}{12\pi r^3} [\Delta\chi_{ax}(3 \cos^2 \theta - 1) + \frac{3}{2}\Delta\chi_{rh} \sin^2 \theta \cos 2\varphi] \quad (1.4)$$

A PCS can also be described by its Cartesian coordinates, **equation 1.5**.

$$\delta^{\text{PCS}} = \frac{1}{12\pi r^3} \left(\Delta\chi_{ax} \frac{2z^2 - x^2 - y^2}{r^2} + \frac{3}{2}\Delta\chi_{rh} \frac{x^2 - y^2}{r^2} \right) \quad (1.5)$$

where x , y and z are the coordinates of the nucleus in the $\Delta\chi$ frame.

In protein NMR, PCS can be generated with paramagnetic ions such as lanthanoids. They are readily measurable as the chemical shift difference between the chemical shift in the diamagnetic control sample and the one in the paramagnetic sample. With an adequate lanthanoid, PCS can be observable up to 100 Å, considering 0.02 ppm as a resolution threshold detection, **Figure 1.6**, the tensor magnitudes were taken from the paper published by Pintacuda et al. in 2007⁶⁹. The PCS is easy to predict with a known protein structure and $\Delta\chi$ tensor, and they provide useful long-range restraints for protein NMR studies.

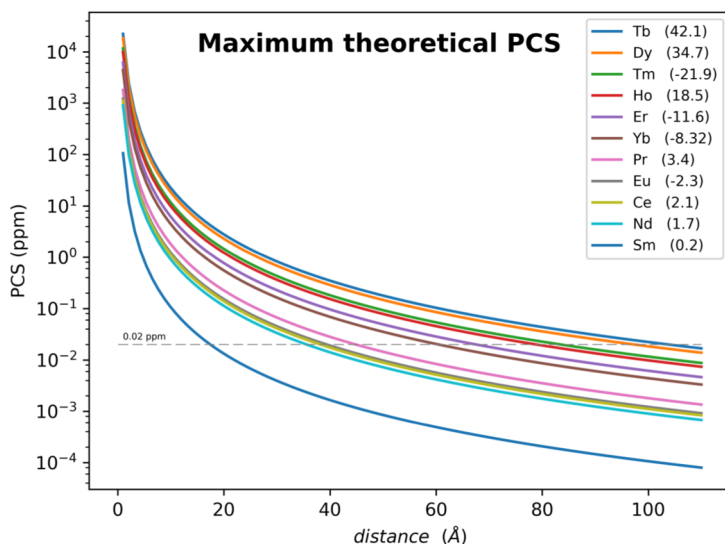


Figure 1.6. Maximum theoretical PCS, in absolute value, for 11 lanthanoids, according to the distance nucleus-paramagnetic center. In the legend, the axial component of $\Delta\chi$ are shown in parenthesis in 10^{-21} m^2 .

Residual dipolar couplings RDC and residual anisotropic chemical shift RACS

In liquid-state NMR, the short tumbling time of molecules averages out the different dipole-dipole orientations. However, a fraction of dipolar coupling can be observed if the tumbling is not completely isotropic, so when certain molecule orientations are favored. Such alignment can be generated either by external alignment media or by self-alignment⁷⁰ if the interaction of a magnetic dipole with the external field is sufficiently strong, for example as observed for stacked nucleic acid bases^{71,72} or anisotropically distributed unpaired electrons. Under alignment, dipole-dipole couplings do not average to zero anymore and a residual dipolar coupling (RDC) can be measured. Even a 1:1000 alignment is enough to observed RDC because the dipolar couplings are strong. The higher the external magnetic field, the stronger the interaction and the greater the orientation, and thus the larger the RDC is. RDC for bonded atoms can be measured as the difference between the observed paramagnetic coupling and the observed diamagnetic coupling.

$$\text{RDC} = \text{observed coupling}_{\text{paramagnetic}} - \text{observed coupling}_{\text{diamagnetic}}$$

$$\text{RDC} = (J_{\text{scalar}} + \text{RDC}) - J_{\text{scalar}}$$

Coupling constants can be measured with IPAP-HSQC^{73,74}. This experiment consists of two datasets acquired in an interleaved manner. The in-phase (IP) experiment is a regular HSQC with ^1H coupled to nucleus ^{15}N during t_1 evolution. The ^{15}N chemical shifts are modulated by $\cos(Ct_1)$, where C is the J_{scalar} or the $J_{\text{scalar}} + \text{RDC}$, and the two components of the doublet have the same sign. The anti-phase (AP) experiment

contains an additional INEPT block to convert the ^{15}N antiphase term into an in-phase term before the ^{15}N chemical shift evolution. After ^{15}N evolution only the term modulated by $\sin(Ct)$ is detected and the two components of the doublet have opposite signs. The addition and subtraction of the IP and AP spectra yield spectra with single peaks, shifted by either $+C/2$ or $-C/2$ in the ^{15}N dimension, relative to the regular HSQC spectrum. The coupling constants C are readily measured as the difference of resonances. The RDC is then calculated as the difference between the coupling measured in the aligned and the non-aligned samples. RDC depends on the degree of the alignment of a nucleus, which will be the same for each nucleus in a rigid molecule. Thus, it is distance independent. In case of paramagnetic alignment, motions of the nucleus relative to the p. will lead to a reduction in the observed RDC. This effect can be used to establish mobility of domains or protein partners^{75,77}. The RDC ($\Delta\nu^{\text{RDC}}$ expressed in Hz) caused by an anisotropic paramagnetic center is given by **equation 1.6**.

$$\Delta\nu^{\text{rdc}} = -\frac{1}{4\pi} \frac{B_0^2}{15kT} \frac{\gamma_A\gamma_B\hbar S_{\text{LS}}}{2\pi r_{AB}^3} \left[\Delta\chi_{\text{ax}} \frac{2z_{AB}^2 - x_{AB}^2 - y_{AB}^2}{r_{AB}^2} + \frac{3}{2} \Delta\chi_{\text{rh}} \frac{x_{AB}^2 - y_{AB}^2}{r_{AB}^2} \right] \quad (1.6)$$

Where $x_{AB}^2 = (x_A - x_B)^2$, $y_{AB}^2 = (y_A - y_B)^2$ and $z_{AB}^2 = (z_A - z_B)^2$. A and B are the two nuclei for which the RDC is measured. S_{LS} is the order parameter defined using the Lipari-Szabo model-free⁷⁸, with $0 \leq S_{\text{LS}} \leq 1$ and reports on the motion of the bond between A and B^{69,70}. The higher is S_{LS} , the mores rigid is the protein, the larger is the range of RDC at a given degree of (maximal) alignment, **Figure 1.7**. Note also that the RDC scales with the square of the external field size.

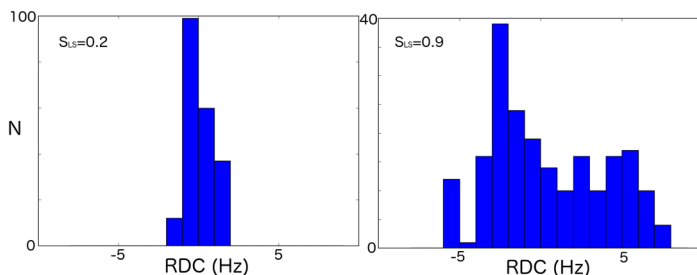


Figure 1.7. Histogram of RDC values for $S_{\text{LS}} = 0.2$ (flexible) and for $S_{\text{LS}} = 0.9$ (rigid), predicted for the backbone amide groups of the n-terminal domain of HSP90 using $\Delta\chi_{\text{ax}} = 8.5$ and $\Delta\chi_{\text{rh}} = 2.5$, as for ytterbium at $B_0 = 19.6\text{ T}$.

Another effect caused by partial alignment is the residual anisotropic chemical shift (RACS). Like the dipolar coupling the chemical shift is anisotropic but under fast isotropic tumbling the anisotropic component vanishes. Partial alignment re-introduces part of the anisotropy, resulting in a change of the chemical shift. Whereas ^{13}C and ^{15}N are affected by RACS, hydrogens are not strongly subject to RACS, due to the low anisotropy of the electron density for this nucleus. Therefore, PCS are measured

preferably in the proton dimension, additionally to the fact that the experimental resolution is better in this dimension.

Paramagnetic relaxation enhancement PRE

Paramagnetic relaxation enhancement, PRE, is a contribution to the longitudinal and transverse relaxation due to interactions between the nucleus and p. PRE can provide information about low-populated states as long as these low-populated states experience large PRE, and the exchange between the ground and low-populated state is faster than the PRE rate experienced by the low-populated state⁷⁹. Thus, it has been used to characterize encounter complexes^{80,81} and other molecular ensembles⁸². Aside from its great interest for dynamic systems, strong PRE is also problematic because excessive peak broadening can render peaks undetectable.

For nuclei outside the contact shift region, dipolar coupling between the unpaired electron and the nucleus are the primary source of relaxation. Two mechanisms contribute to this relaxation phenomenon, the Solomon mechanism and the Curie mechanism⁸³. For lanthanoids with anisotropic susceptibility tensors, which is the case for all paramagnetic lanthanoids except Gd³⁺, Curie relaxation is dominant for transverse relaxation. Curie relaxation scales with the square of the Larmor frequency, and thus with the field, **equation 1.7**. It is also stronger for large molecules, because it depends on the rotational correlation time τ . Lanthanoids with a large number of unpaired electrons (S , the spin quantum number), such as Tb³⁺ or Dy³⁺, show the largest Curie relaxation effects.

$$r^{\text{PRE}} = \frac{1}{5} \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\omega_I^2 g_e^4 \beta^4 S^2 (S+1)^2}{(3kT)^2 r^6} \left(4\tau_r + \frac{3\tau_r}{1 + \omega_I^2 \tau_r^2} \right) \quad (1.7)$$

From a more general perspective, paramagnetic NMR can help to make NMR on proteins more accessible for non-experts. The NMR theory is difficult and many experiments are highly complex, so set-up and data analysis can be complicated. NMR would benefit from a large bank of routine experiments and software yielding information to non-specialists⁸⁴. Paramagnetic NMR effects are applicable to structure determination but also for dynamics and protein-protein complexes. It is therefore a powerful and accessible tool for biochemists. The incorporation of the paramagnetic ions in the protein of interest can be challenging, but broadly available molecules and techniques make it easier to find an adequate method for the system of interest. In a way, using paramagnetic NMR corresponds to shifting from NMR complexity to biochemical engineering complexity. The usefulness of paramagnetic tools for NMR studies is reflected by the increasing number of programs incorporating paramagnetic effects: For structure calculation: CYANA¹⁹, REDCRAFT/ REDCAT^{85,86}, REFMACS^{5,87}, DipoCoup⁸⁸, ROSETTA^{89,90}, FLEXIBLE-MECANO⁹¹, Xplor-NIH^{92,93}; for dynamics and ligand docking: Xplor-NIH^{92,94}, HADDOCK^{95,96}, MaxOcc^{90,97}; for assignment:

POSSUM⁴¹, PARASSIGN^{42,43}, MAP-XSII³⁸; and for tensor refinement: FANTEN⁹⁸, NUMBAT⁹⁹, ECHIDNA¹⁰⁰).

Thesis outline

The aim of the research presented in this thesis was to develop new methods for challenging systems in liquid-state NMR using paramagnetic effects generated by the two-armed probe CLaNP-5. Chapter 1 is an introduction about NMR assignment development and challenges and a brief opening to the theory of paramagnetic effects and their applications. Chapter 2 describes the upgrade of PARAssign software and its ability to assign and assess methyl groups of a 25 kDa protein using experimental pseudocontact shifts from 1 to 3 paramagnetic centers. Chapter 3 describes the implementation of residual dipolar coupling into PARAssign software as extra dataset input in addition to the pseudocontact shifts. Applications are given for residue selective isotope labeled protein. Chapter 4 depicts the use of pseudocontact shifts used to probe small methyl group movements and re-orientation under ligand binding. Chapter 5 gives some clues about how paramagnetic effect such as RDC but also PCS can provide information for challenging highly dynamic system. Chapter 6 gives insights for PCS application to very large multimeric protein, the challenges and limitations are discussed. Chapter 7 provides a discussion and few perspectives about the work carried out and presented in this thesis.

Chapter 2

Methyl group assignment using pseudocontact shifts with PARAssign

Based on the research article: Lescanne M., Skinner SP., Blok A., Timmer M., Cerofolini L., Fragai M., Luchinat C., Ubbink M. (2017) *Methyl group assignment using pseudocontact shifts with PARAssign*. J. Biomol. NMR DOI: 10.1007/s10858-017-0136-3

Abstract

A new version of the program PARAssign has been evaluated for assignment of NMR resonances of the 76 methyl groups in leucines, isoleucines and valines in a 25 kDa protein, using only the structure of the protein and pseudocontact shifts (PCS) generated with a lanthanoid tag at up to three attachment sites. The number of reliable assignments depends strongly on two factors. The principle axes of the magnetic susceptibility tensors of the paramagnetic centers should not be parallel so as to avoid correlated PCS. Second, the fraction of resonances in the spectrum of a paramagnetic sample that can be paired with the diamagnetic counterparts is critical for the assignment. With the data from two tag positions a reliable assignment could be obtained for 60% of the methyl groups and for many of the remaining resonances the number of possible assignments is limited to two or three. With a single tag, reliable assignments can be obtained for methyl groups with large PCS near the tag. It is concluded that assignment of methyl group resonances by paramagnetic tagging can be particularly useful in combination with some additional data, such as from mutagenesis or NOE-based experiments. Approaches to yield the best assignment results with PCS generating tags are discussed.

Keywords

Assignment; Pseudocontact shift; Methyl groups; NMR spectroscopy; Paramagnetic tag; Heat shock protein

Abbreviations

PCS, pseudocontact shift; RDC, residual dipolar coupling; PRE, paramagnetic relaxation enhancement; HSQC, heteronuclear single quantum coherence; TROSY, transverse relaxation optimized spectroscopy; CLaNP-5, caged lanthanoid NMR probe 5; ntd-HSP90, N-terminal domain of heat shock protein 90

Introduction

Liquid state NMR is widely used to study protein structure, dynamics and interactions with atomic detail. Resonance assignment is a prerequisite for all these studies. Methyl groups have become important in biomolecular NMR to probe biomolecular interactions²³ as well as dynamics of proteins over a large range of timescales, from picoseconds to seconds²⁸. Furthermore, the hydrophobic nature of methyl groups makes them suitable to probe the core of a protein²⁶ and study specific interactions it may have with other proteins, ligands or substrates²⁷. From a spectroscopic point of view, ¹³CH₃ methyl groups are valuable probes in very large systems because their inherent fast rotation and geometry give them favorable relaxation properties that make methyl groups suitable for TROSY experiments^{24,25}. Specific isotope labeling on isoleucine, leucine and valine (ILV) methyl groups is an established approach²⁹ to reduce spectral crowding and to provide good spatial coverage due to dispersion of these residues through the protein. Assignment of ILV methyl groups is, however, non-trivial. To take advantage of standard *J*-coupling based assignments procedures, uniformly labeled samples are needed, requiring the production of additional samples and time-consuming 3D NMR experiments. For large systems, methyl assignment is further complicated by the need for fractional or complete deuteration of the other hydrogen positions to reduce relaxation by dipolar couplings. Various approaches have been published to assign methyl groups in large proteins. Systematic site-directed mutagenesis, combined with stereospecific labeling on methyl groups has been applied to TET2^{30,31}. To circumvent the size limitation, the so-called “divide and conquer” approach also proved to be effective. Here, individual domains or subunits are first assigned, followed by a transfer of assignments to the spectrum of the larger multidomain protein or multisubunit protein complex^{31,35,36}. Another approach is to use multidimensional NOESY spectra in combination with a crystal structure³⁷. This approach has been implemented in the MAP-XSII software³⁸ that automatically assigns methyl groups based on a known structure combined with NOEs measured on 4D NOESY experiments. Similarly, the program FLAMENGO uses a combination of 3D NOESY, methine-methyl TOCSY and paramagnetic relaxation enhancements (PREs) and a protein structure to assign methyl groups³⁹.

More generally, paramagnetic NMR has become attractive for assignment¹⁰¹, and software packages have been developed such as Platypus¹⁰², Echidna¹⁰⁰, Possum⁴¹ and PARAssign^{42,43}. Platypus and Echidna have been used to assign NH resonances on the basis of a crystal-structure and pseudo-contact shifts (PCS). Possum assigns methyl groups with PCS measured with paramagnetic centers for which the magnetic susceptibility tensor is known. This requires assignment of at least five amides or methyl groups to be available. PARAssign does not require any assignment or prior knowledge of the susceptibility tensors. It automatically assigns ¹H, ¹³C, or ¹⁵N resonances based on PCS from several paramagnetic centers obtained from 2D HSQC

or TROSY spectra and a 3D structure of the protein and it also provides an assessment of the reliability of the assignment. It has been shown that PARAssign software could assign the amides of pseudoazurin (PAZ) (125 residues) using PCS from 3 three different paramagnetic centers. PARAssign has been tested also on P450cam (414 residues) for methyl groups assignment using synthetic data⁴².

Here, we evaluate the ability of PARAssign to assign the 76 ILV methyl groups of a 25 kDa protein, the N-terminal domain of HSP90 (ntd-HSP90), using experimentally observed PCS. Ntd-HSP90¹⁰³⁻¹⁰⁵ is a well-studied chaperone protein, for which methyl group assignments are available³⁷. PCS were induced by the paramagnetic tag CLaNP-5, attached to the protein via disulfide linkages to engineered cysteine residues. We find that, using two judiciously placed paramagnetic tags, PARAssign can assign up to 60% of the 76 methyl groups of ntd-HSP90 with high reliability and give many correct suggestions for the remaining 40%. The quality of the input data is an important determinant for the number of reliable assignments. With a single tag, reliable assignments can be obtained for nearby nuclei with large PCS.

Material and methods

Protein preparation

Four double cysteine mutants of the ntd-HSP90 were designed on the surface of the protein, S50C/D54C, A101C/N105C, T149C/I187C, and M130C/Q133C. Ntd-HSP90 does not have any native cysteines. The protein was produced by overexpression of the gene for ntd-HSP90 (residues 8-233, UNIPROT entry P07900) with a cleavable His₆ tag in the PQTEV vector in *Escherichia coli* BL21 DE3. The plasmid was kindly provided by Dr Dipen Shah³⁷. Cultures of 500 mL M9 medium in 2 L Erlenmeyer flasks were incubated at 37 °C at 150 rpm. At an OD₆₀₀ of 0.6, IPTG (1 mM) was added to induce gene expression, the temperature was reduced to 18°C and the incubation was continued overnight. For ¹⁵N labeled protein, ¹⁵N ammonium chloride (0.3 g/L) was added as sole nitrogen source. For ¹³CH₃ methyl labeling, two samples were produced labeled either at Leu- δ 1- δ 2/Val- γ 1- γ 2-[¹³CH₃] or Ile- δ 1-[¹³CH₃]. Consequently, it was possible to identify unambiguously the residue type of the Ile methyl groups. Either 50 mg/L 2-keto-3,3-1,2,3,4-¹³C-butyrate (Ile ¹³CH₃ labeling) or 100 mg/L of 2-keto-3-methyl-3-1,2,3,4-¹³C butyrate (Leu and Val ¹³CH₃ labeling) were added to the M9 medium with 4 g/L ¹³C glucose, one hour before induction at OD₆₀₀ ~ 0.5. The cells were harvested by centrifugation at 6000 rpm for 20 min at 4°C. Cell lysis was performed using lysozyme. PMSF and DNase were added for protease inhibition and DNA breakdown. Protein purification was performed with a His-trap NTA-column, eluting with an imidazole gradient (5-500 mM imidazole in 100 mL). The His-tag was then cleaved with TEV protease during overnight incubation at 4°C. The His-tag and TEV protease were removed by a second elution from the column. CLaNP-5 was synthesized as described before^{106,107}. For CLaNP-5 attachment, the protein was reduced

for one hour on ice in the presence of 5 mM DTT. Then, DTT was removed with a Sephadex G-25 PD10 desalting column (GE Healthcare) and the protein solution was mixed immediately with four molar equivalents of either Yb³⁺-CLaNP-5 or Lu³⁺-CLaNP5 and incubated at room temperature for one hour. The sample was loaded on an analytical Superdex 200 gel filtration column (GE Healthcare) to remove excess CLaNP-5 and protein dimers.

NMR spectroscopy

NMR samples contained 50 - 100 μ M protein in 50 mM Tris-HCl, 100 mM NaCl, pH 7.5 and, 6% D₂O. All the spectra were recorded at a temperature of 298 K. ¹⁵N-HSQC spectra were recorded on a Bruker Ascend Avance III 850 HD MHz spectrometer, equipped with a TCI cryoprobe, with spectral widths of 20.00 ppm (17.00 kHz) and 35.00 ppm (3.015 kHz) in the ¹H and ¹⁵N dimensions, respectively. Leu- δ 1- δ 2/Val- γ 1- γ 2-[¹³CH₃] and Ile- δ 1-[¹³CH₃] protein spectra were recorded on Bruker 800 MHz and 850 MHz spectrometers equipped with TXI cryoprobes, with a spectral width of 20.00 ppm in the ¹³C dimension. Experiments were performed with 16 or 32 scans. Data were processed with Topspin 3.2 and NMRpipe 8.2¹⁰⁸ and spectra were analyzed with CCPNMR Analysis 2.4¹⁰⁹.

Determination of the PRE-derived cutoff distance

The cutoff distance, below which a signal would be too broad to be detected because of PRE, was calculated using **equation 2.1**. The line broadening due to PRE (f_2) was estimated to render nuclei within 9 Å from the Yb³⁺ ion invisible, based on the expected Curie relaxation, **equation 2.1**¹¹⁰, which is the dominant transverse relaxation mechanism for ¹H by lanthanoids with an anisotropic magnetic susceptibility.

$$r_{\text{cutoff}} = \sqrt[6]{\frac{1}{5T_2} \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\omega_H^2 (g_e \mu_B)^4 J^2 (J+1)^2}{(3k_B T)^2} \left(4\tau_r + \frac{3\tau_r}{1 + \omega_H^2 \tau_r^2}\right)} \quad (2.1)$$

Here, μ_0 is the permeability of a vacuum, ω_H is the Larmor frequency of the hydrogen nucleus, g_e is the Landé g -factor, μ_B is the Bohr magneton, J is the total spin quantum number of the paramagnetic lanthanoid, k_B is the Boltzmann constant, T is the absolute temperature, and τ_r is the rotational correlation time of the molecule.

Methyl group assignments

Methyl group assignments from PARAssign were compared to those obtained with traditional 3D NMR procedures. The published methyl group assignments³⁷ were checked and completed using CccoNH/HccoNH/HSQC spectra, kindly provided by Dr. Dipen Shah and Dr. Eiso Ab.

Assignment procedure

The PARAssign code has been refactored for the sake of readability and easier maintenance. A new reliability criterion has been implemented (see Results section). The core procedure remains unchanged¹². The PARAssign user must supply the PCS file containing all the measured proton PCS, a PDB file and the JSON configuration

file with information about the assignment to be performed (atom names, stereospecificity of the input, double Cys mutation sites, starting point for the tensor parameters, input file directory, output file directory). ^{13}C PCS were not considered as input, because the resolution in this dimension is lower and the residual chemical shift anisotropy is significant for carbon but can be neglected for protons. The PARAssign output includes the assignment plot, tensor parameters, tensor PDB file(s) and predicted PCS for all the assignable groups in the structure. The PARAssign input consisted of the ntd-HSP90 structure, PDB entry 3T0Z ¹¹¹, and the PCS datasets acquired for the S50C/D54C, A101C/N105C and T149C/I187C mutants tagged with Lu^{3+} -CLaNP-5 as diamagnetic control and Yb^{3+} -CLaNP-5 as the paramagnetic center. Although we used two different samples for isoleucines and leucines + valines, PARAssign can select the predicted type(s) of each assignable peak by entering the residue type code in the input file ('ILE' or 'VAL LEU' in our case but all possible combinations are acceptable). A run of PARAssign comprises an iteration of two consecutive procedures, the assignment procedure and the tensor optimization procedure using PCS calculated during the previous assignment procedure. In the assignment procedure of the PARAssign software, the Hungarian method for minimal cost assignment ¹¹² is used to perform the assignment of the input PCS, using a scoring factor cost function to populate the assignment matrix:

$$\text{Scoring factor} = \frac{1}{P_c} * \sum_{i=1}^{P_c} \sqrt{\frac{(\delta_{\text{PCS},i}^{\text{pred}} - \delta_{\text{PCS},i}^{\text{exp}})^2}{(|\delta_{\text{PCS},i}^{\text{pred}}| + |\delta_{\text{PCS},i}^{\text{exp}}|)^2}} \quad (2.2)$$

where P_c is the number of paramagnetic centers. PARAssign attributes a scoring function of 1 to assignments for which experimental and predicted PCS have opposite signs, thus disfavoring the assignment. A scoring factor threshold is user-defined (default 0.2) and is used to filter out assignments for the optimization procedure for which the scoring factor is too high. A high threshold yields many assignments with poor accuracy and a low threshold yields few assignments. The optimal threshold is the lowest possible value that yields the maximum number of assignments. For instance, when starting with a scoring factor of 0.7 and decreasing it progressively, the optimum threshold is passed when the number of assigned groups starts to drop. The tensor parameters optimization procedure consists of the minimization of the square difference between calculated and experimental PCS for all paramagnetic centers. This sequential least squares programming procedure ¹¹³ is performed by the python function `scipy.optimize.fmin_slsqp` (<http://www.scipy.org>). The output of PARAssign gives the susceptibility tensors of the paramagnetic centers and the assignment of the diamagnetic protein spectrum with a reliability scoring for each assignment (discussed below). The output figures included in this article may not reflect the true output quality. In the original .png image, the user can zoom in to read all the details. For this chapter, it is relevant to be able to judge the quality of the assignment at a glance.

Results and discussion

Selection of tagging sites

Assignment based on PCS requires accurate PCS that are not affected by averaging effects due to tag mobility. For this reason, the two-armed caged lanthanoid NMR probe (CLaNP) #5 was used^{107,114}, which requires two Cys residues located close together on the surface of the protein for attachment. Ntd-HSP90 has no Cys residues, so pairs of cysteines were engineered. The criteria for selection were (I) good distribution of the sites over the protein surface, (II) limited mobility of the attachment site by using secondary structure elements, (III) sufficiently exposed sites without steric hindrance of large side-chains. Four double cysteine mutations were designed, three of which are on α -helices (S50C/D54C, C α -C α = 6.0 Å; A101C/N105C, C α -C α = 5.9 Å; M130C/Q133C, C α -C α = 5.7 Å) and one was across neighboring strands of a beta-sheet (T149C/I187C, with C α -C α = 5.0 Å). Out of the four double cysteine variants, the three with the highest yields were selected, which were S50C/D54C, A101C/N105C and T149C/I187C, with yields of 30 mg/L, 30 mg/L and 20 mg/L in minimal medium, respectively. Variant M130C/Q133C yielded only 1 mg/L and was not considered further.

Reference $\Delta\chi$ -tensor determination using backbone amides

For reference, the lanthanoid location and the orientation of the $\Delta\chi$ -tensor were first determined on the basis of amide proton PCS and using Numbat software⁹⁹ with a crystal-structure of ntd-HSP90 (PDB entry 3T0Z¹¹¹), **Table 2.1**.

Table 2.1 Susceptibility $\Delta\chi$ -tensors obtained by fitting amide PCS of ntd-HSP90 (PDB entry 3T0Z)

Mutant tagged with Yb ³⁺ - CLaNP-5	50C/54C	101C/105C	149C/187C
$\Delta\chi_{\text{ax}}$ (10^{-32} m ³)	8.4 (0.1)	8.0 (0.2)	8.3(0.3)
$\Delta\chi_{\text{rh}}$ (10^{-32} m ³)	2.0 (0.3)	1.6(0.5)	2.2(0.3)
α (°)	173(5)	180(2)	176(3)
β (°)	155(1)	88(2)	148(1)
γ (°)	150(3)	170(3)	15(4)
x (Å)	12.4(0.5)	25.3(0.2)	-1.8(0.2)
y (Å)	1.3(0.5)	-9.5(0.2)	-5.8(0.5)
z (Å)	25.2(0.3)	6.8(0.2)	-1.2(0.2)
Restraints	80	63	69
Q value*	0.056	0.068	0.26
Q ^a value*	0.028	0.034	0.12
Average pcs deviation (ppm)	0.012	0.009	0.035

* The Q-values were calculated using:

$$Q = \sqrt{\frac{\sum(\delta_{\text{PCS},i}^{\text{pred}} - \delta_{\text{PCS},i}^{\text{exp}})^2}{\sum(\delta_{\text{PCS},i}^{\text{exp}})^2}}$$

$$Q^a = \sqrt{\frac{\sum(\delta_{\text{PCS},i}^{\text{pred}} - \delta_{\text{PCS},i}^{\text{exp}})^2}{\sum(|\delta_{\text{PCS},i}^{\text{pred}}| + |\delta_{\text{PCS},i}^{\text{exp}}|)^2}}$$

Diamagnetic assignments for the amide groups were available ¹¹⁵, BMRB entry 7003. The three mutants were ¹⁵N-labeled and tagged with the diamagnetic control tag Lu³⁺-CLaNP-5 or the paramagnetic tag Yb³⁺-CLaNP-5 and PCS were obtained from ¹⁵N-¹H HSQC spectra. To ameliorate the pairing problem, spectra of the ¹⁵N labelled ntd-HSP90 tagged with CLaNP-5 loaded with Eu³⁺ were acquired for mutants 50C/54C and 101C/105C. The two lanthanoid $\Delta\chi$ tensors are oriented in a nearly identical orientation and only the sizes of $\Delta\chi_{ax}$ and $\Delta\chi_{th}$ differ. Thus, the paramagnetic peaks are on the one correlation line with the diamagnetic peak and the size of the CS scales with the strength of the $\Delta\chi$ tensor. The alignment of the three peaks presents an easy way to pair paramagnetic peaks to diamagnetic ones, **Figure 2.1** For mutants S50C/D54C and A101C/N105C, the agreement between observed and calculated PCS was excellent, **Figure 2.2**. For mutant T149C/I187C the fit was of somewhat lower quality, for reasons discussed later. To check whether significant PCS could be expected for most methyl groups, PCS were predicted for the three $\Delta\chi$ -tensors derived from the amide PCS. **Figure 2.3** shows that with Yb³⁺-CLaNP-5, PCS ≥ 0.04 ppm are expected for most methyl groups for each of the three tag locations. These amide-based $\Delta\chi$ -tensors were used solely to generate the simulated data and for comparison with the ones refined by PARAssign for the experimental methyl group data, see the sections below.

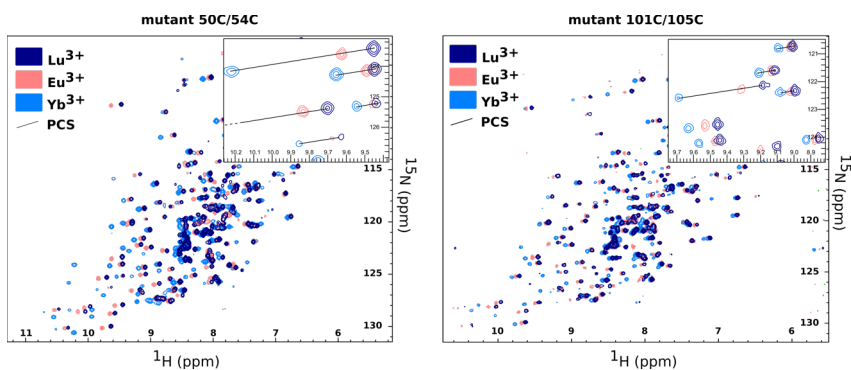


Figure 2.1. Overlay of HSQC spectra of the two ntd-HSP90 mutants 50C/54C and 101C/105C tagged with Lu³⁺-CLaNP-5, Eu³⁺-CLaNP-5, Yb³⁺-CLaNP-5, in dark blue, pink and light blue, respectively. Zoom of the alignment of the three peaks are shown enclosed.

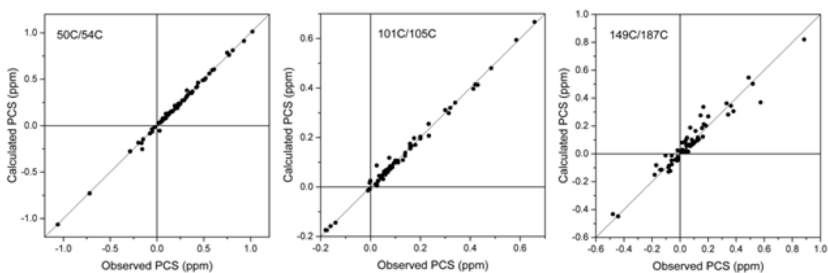


Figure. 2.2 Quality of $\Delta\chi$ fits. The calculated PCS are plotted against the observed PCS of amide groups of ntd-HSP90 tagged with Yb³⁺-CLaNP-5 at the indicated positions. The solid lines represent a perfect correlation ($x=y$).

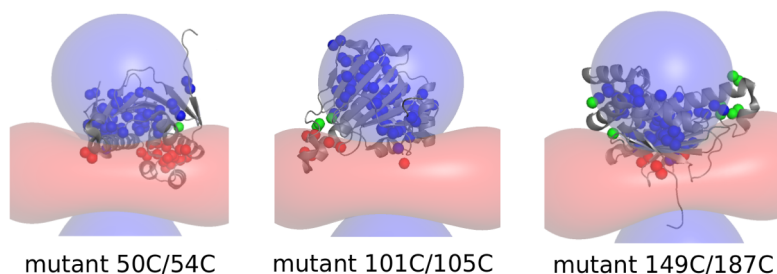


Figure. 2.3 Predicted methyl group PCS. Methyl groups that are predicted to experience positive PCS, negative PCS and PCS smaller than 0.04 ppm are represented as blue, red and green spheres, respectively. The -0.04 ppm and 0.04 ppm isosurfaces of the $\Delta\chi$ -tensors based on the amide PCS are shown as red and blue surfaces.

Tagging effects on methyl group spectra

To establish the effects of a double cysteine mutation and the attachment of CLaNP-5 on methyl resonances, ¹³C-H-HSQC spectra were acquired for the WT and Lu³⁺-CLaNP-5 tagged mutant ntd-HSP90 with ILV methyl-labeling. The peaks from the mutant protein were matched to those of the wild-type. Most peaks could be readily matched and showed only very small chemical shift changes, **Figure 2.4a**. Two peaks, one for mutant 50C/54C and one for mutant 101C/105C, could not be matched with certainty to equivalent peaks in the spectrum of wild-type ntd-HSP90. These methyl groups were located close to the tag, **Figure 2.4b**. In mutant 149C/187C one Ile was replaced by a Cys for attachment of the tag. The limited chemical shift perturbations between WT and mutant ntd-HSP90 indicate that no major structural changes occur because of the mutations and the tagging.

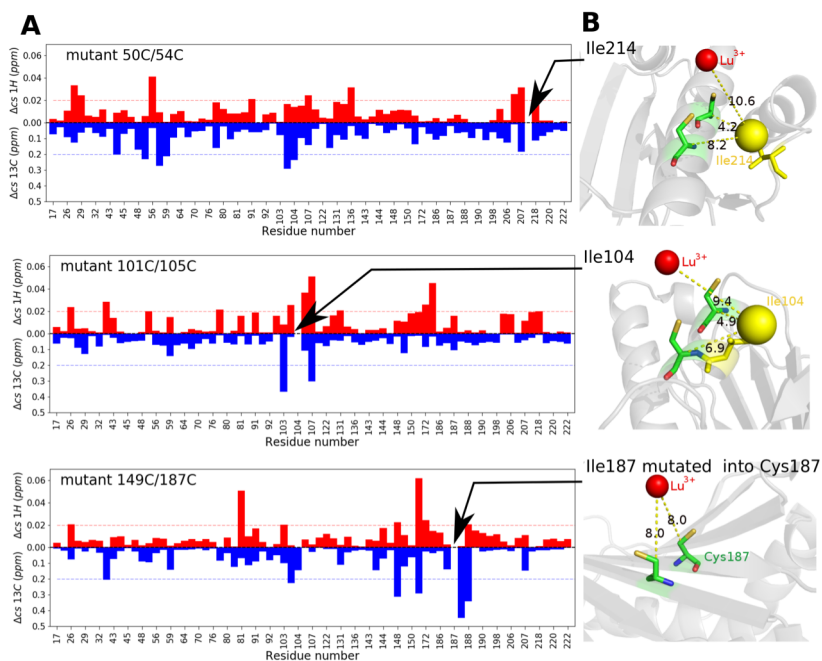


Figure. 2.4 Mutation and tagging effects. **a** Absolute chemical shift differences between WT and Lu^{3+} -tagged mutant ntd-HSP90 resonances for ^1H (red) and ^{13}C (blue) are plotted for each methyl group. The dotted lines mark significant changes, with 0.2 ppm for ^{13}C and 0.02 for ^1H . **b** Location of methyl groups with unmatched resonances relative to the tag site. Distances are expressed in Ångström.

The pairing problem

PCS are measured as differences in chemical shifts of resonances observed in a paramagnetic and the diamagnetic spectrum. The problem to be solved is thus to pair the diamagnetic and paramagnetic resonances belonging to the same methyl group. PCS expressed in ppm units do not depend on the gyromagnetic ratio of the observed nucleus. Consequently, when the distance between the nucleus and the paramagnetic center, as well as the orientation of the vector connecting them in the $\Delta\chi$ -tensor frame are similar for ^1H and ^{13}C , which is almost always true for a $^{13}\text{C}^1\text{H}_3$ methyl group, the ^{13}C and ^1H nuclei of a methyl group experience nearly the same PCS. In **Figure 2.5**, the predicted PCS for all methyl ^{13}C and ^1H nuclei are plotted for ntd-HSP90 tagged at the three positions (224 points in total). The ^1H PCS are averaged over the three PCS for the individual protons in the methyl group to account for the fast methyl rotation. The ^1H PCS differ by less than 5% from the ^{13}C PCS for PCS < 0.5 ppm and up to 10% for PCS larger than 0.5 ppm. Note that we neglected the effects of residual anisotropic chemical shifts and residual dipolar couplings due to partial alignment. In the plot, two exceptions can be seen. The angles of the metal-methyl vector of methyl groups L89

C δ 1 and V92 C γ 1 are very close to the angle where the PCS changes sign in mutant 149C/187C, resulting in a very high PCS gradient, **Figure 2.5**.

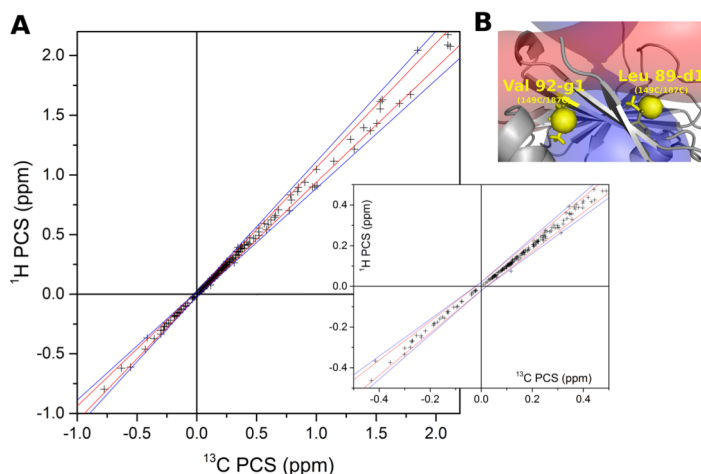


Figure. 2.5 Correlation between the ^1H PCS and ^{13}C PCS of methyl groups. **a** The ^1H and ^{13}C methyl PCS predicted for the three tag positions on ntd-HSP90 are plotted. The PCS differing less than 5% and 10% are contained between the red and blue lines, respectively. The inset shows a detail of the plot for small PCS. The lines intercept the axes at $|0.02|$ ppm, taken as error margin. **b** Location of the two methyl groups, for the mutant 149C/187C, for which the deviation between ^1H and ^{13}C PCS is bigger than 10%. Isosurfaces are shown for PCS = 0.25 and -0.25 ppm. Both methyl groups are very close to the region where the PCS changes sign.

Thus, in general, the chemical shifts for the ^1H and ^{13}C are expected to be on a line with slope 0.95 - 1.05 for small PCS and 0.9 - 1.1 for large PCS, the correlation line. This property is used to identify which diamagnetic and paramagnetic resonances derive from the same nucleus and, thus, yield an experimental PCS (the pairing procedure). For well-dispersed parts of the spectrum pairing is trivial but in more crowded areas it can be ambiguous when more than one paramagnetic peak is on the correlation line. Pairing can either be done fully automatically with a script implemented in PARAssign based on these criteria, or in combination with manually paired peaks. When pairing manually, differences in peak intensity and shape can sometimes help, **Figure 2.6**.

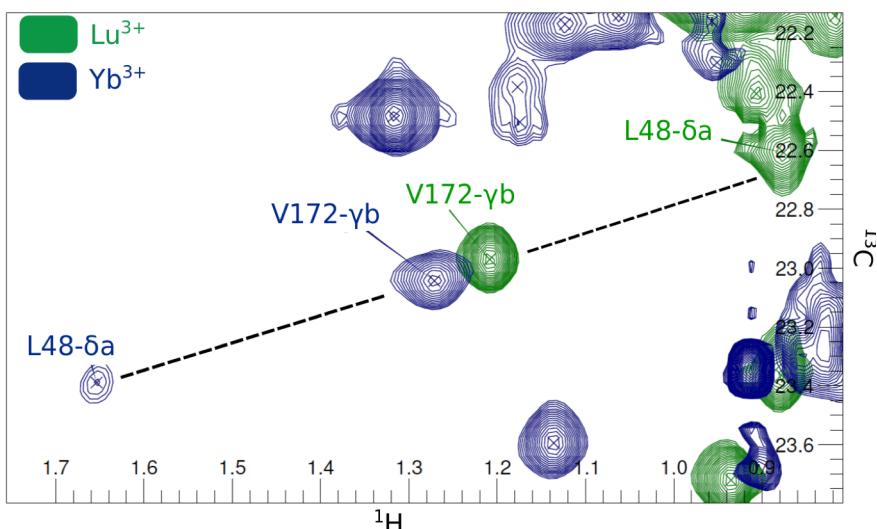


Figure. 2.6 Illustration of the peak pairing problem. The peaks for V172 and L48 are on the same correlation line in the spectrum of Yb^{3+} tagged ntd-HSP90, making the pairing to the peaks in the spectrum of Lu^{3+} tagged protein problematic. In this case, the observation that the V172 in the

Pairing is done iteratively, by pairing additional peaks using the predicted PCS after a run of PARAssign. In this way, 97% of ILV peaks were matched for both the mutants 50C/54C and 101C/105C and 58% of the mutant 149C/187C. The lower pairing level for the latter mutant as well as the relatively poor correlation between measured and calculated amide PCS suggest that the data for this mutant are of lower quality. It is unlikely that the tag causes a major structural modification because only small chemical shift differences between the spectra of the WT and Lu^{3+} -CLaNP tagged mutant are observed, **Figure 2.4a**. The tag is also connected via both arms, because the size of the $\Delta\chi$ -tensor is similar to that of the other mutants, **Table 2.1**. Single armed CLaNP yields strongly reduced $\Delta\chi$ -tensor values¹¹. However, some unexpected broadening and the presence of double peaks in the paramagnetic spectrum suggest that the tag may assume more than one orientation relative to the protein. The tag in mutant 149C/187C CLaNP-5 crosslinks two β -strands, a type of attachment that has not been reported before, to the best of our knowledge. Perhaps this type of attachment results in some degree of conformational freedom of the tag. The $C\alpha$ - $C\alpha$ distance is about 1 Å shorter for this mutant than for the two other considered mutants. In the two other mutants, the CLaNP tag connects two consecutive turns on alpha-helices, **Figure 2.4b**.

A new criterion for reliability

A consequence of the pairing problem is that the input data will, in general, not be complete, contrary to the synthetic data used in the simulation described in a previous paper¹². To reduce possible artefacts in the assignment due to local minima in the search space, a jackknife procedure was implemented. Assignment is performed 100 times while leaving out 5% of the input PCS, randomly selected for each of the runs.

During each run, PARAssign performs the assignment automatically using an iterative procedure. The back-calculated PCS are matched with the input PCS, followed by the optimization of the $\Delta\chi$ -tensor parameters and lanthanoid location. PARAssign assesses the reliability of the assignments with a starring system. If all (and at least two) PCS for a given methyl group are >0.02 ppm and the combined fit of the PCS has a scoring factor, **equation 2.2**, lower than the average, the assignment is 2*. If only one of the criteria is met, the assignment gets 1*, otherwise 0*. After 100 runs, the fraction of occurrence (out of the 100 runs) of each assignment is combined with the average starring of the assignment to get an overall reliability, indicated with a color code, green, yellow and white for highly reliable, suggestion and not reliable, respectively, illustrated in **Figure 2.7**.

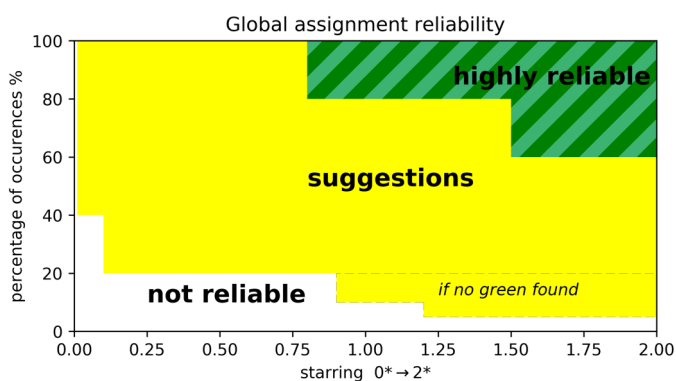


Figure. 2.7 Assignment reliability. The average quality score from the 100 assignment runs is indicated on the horizontal axis and the occurrence percentage of the runs in which the assignment is found is plotted vertically. Dark and light green refer to stereospecific and non-stereospecific assignment of Leu and Val methyl groups.

To be classified as highly reliable, colored in green, a residue assignment must occur in more than 60% of the runs with an average starring higher than 1.5 or occur more than 80% with an average starring higher than 0.8. Methyl groups that are assigned to only one of the two methyl groups in Val or Leu, i.e. methyls that are assigned stereospecifically, are shown in dark green. If the residue is reliably assigned but both methyl groups are frequently identified as the correct assignment in the 100 consecutive assignment runs, the bar is shown in light green, indicating that the resonance has been assigned reliably but non-stereospecifically to a residue. It should be noted that the correctness of stereospecificity could not be assessed because no stereospecific assignments were available for ntd-HSP90. Assignments occurring in less than 5% of the runs are deemed unreliable (white). For all the other assignments, colored in yellow, PARAssign finds one or more assignments with reasonable occurrence frequencies (higher than 40% but not already green, higher than 20% with a starring >0.1 , and, if no highly reliable assignment has been found, extra yellow assignments are proposed, see **Figure 2.7**). Such a set of suggestions can still be very useful when

combined with other strategies. Data from other sources (mutagenesis, NOE) could be used to decide which of the assignments is the correct one.

Methyl group assignment with simulated data

Ntd-HSP90 as represented in PDB entry 3T0Z contains 20 isoleucines labeled on the carbon $\delta 1$, 18 leucines labeled on both carbon $\delta 1$ and carbon $\delta 2$ and 10 valines labeled on both carbon $\gamma 1$ and carbon $\gamma 2$. Thus, in total 76 methyl groups need to be assigned. To establish whether PARAssign can solve the assignment problem, first the PCS for the 76 methyl groups were calculated using the $\Delta\chi$ -tensors for the three tag positions obtained from the amide PCS. PARAssign was run using the WT protein structure 3T0Z and the 3 simulated PCS datasets for input. The starting parameters of the three $\Delta\chi$ -tensors were $\Delta\chi_{ax} = 8.5$ and $\Delta\chi_{az} = 2.0$. The starting angles and metal position were predicted by PARAssign purely based on reported positions relative to the Cys C α and C γ atoms and $\Delta\chi$ -tensor orientations for CLaNP-5, which has been shown to be a good approximation^{42,107}. Therefore, no prior experimental knowledge of the tensors is necessary. Using the simulated data for three paramagnetic centers as input, the vast majority of assignments was classified as highly reliable and all these ‘green’ assignments were correct. In addition, for the three assignments classified as ‘yellow’ the correct assignment was proposed, **Table 2.2**.

Table 2.2 Assignment of the methyl groups. The terms m1, m2 and m3 refer to the mutants 50C/54C, 101C/105C and 149C/187C, respectively. For the green category, the numbers represent the correct/total number of assignments. For the yellow category, the first number indicates for how many residues the correct residue was among the assignment suggestions, the second represents the total number of assignments in this category.

Number of Paramagnetic centers		3			2		1	
Mutant(s):		m1, m2, m3	m1, m2	m1, m3	m2, m3	m1	m2	m3
input		76	76	76	76	76	76	76
Complete simulated dataset	Highly reliable	73/73	61/61	6/6	73/73	0/0	5/5	7/10
	Suggestion	3/3	15/15	69/69	3/3	22/52	69/70	16/53
	Not reliable	0	0	1	0	24	1	13
input		76	76	76	76	72	74	44
Experimental dataset	Highly reliable	27/27	46/46	0/0	9/9	1/1	6/6	1/3
	Suggestion	45/47	27/28	22/55	58/66	25/57	40/56	13/33
	Not reliable	2	2	21	1	14	12	8

The PARAssign output is displayed in **Figure 2.8**. When using the data from only two tags, two of the three pairs of tags (50C/54C with 101C/105C and 101C/105C with 149C/187C) yield similar results as with three tags, **Table 2.2**, **Figure 2.9**. Combining the data of mutants 50C/54C and 149C/187C yields mostly assignments with more than one possibility (‘yellow’), on the whole. The reason for this striking difference will be discussed below. We can conclude from the assignment runs with synthetic datasets that the assignment problem can be solved by PARAssign with complete and error-free data input from two or three tag positions.

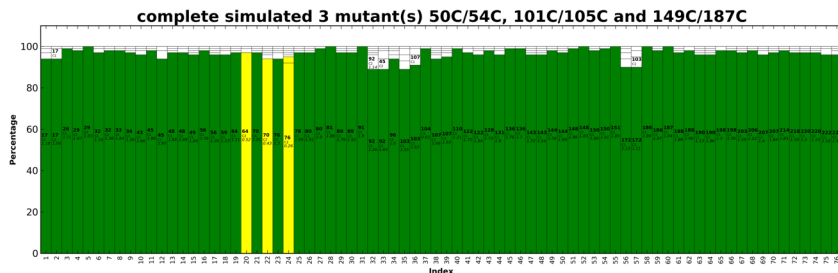


Figure. 2.8 PARAssign output for a complete, simulated dataset. In green the highly reliable assignments and in yellow the suggested assignments are indicated. In each bar, the proposed residue and methyl group as well as the average starting value are shown. In this example, all the assignments are correct. See also the enlargement in **Figure 2.10**.

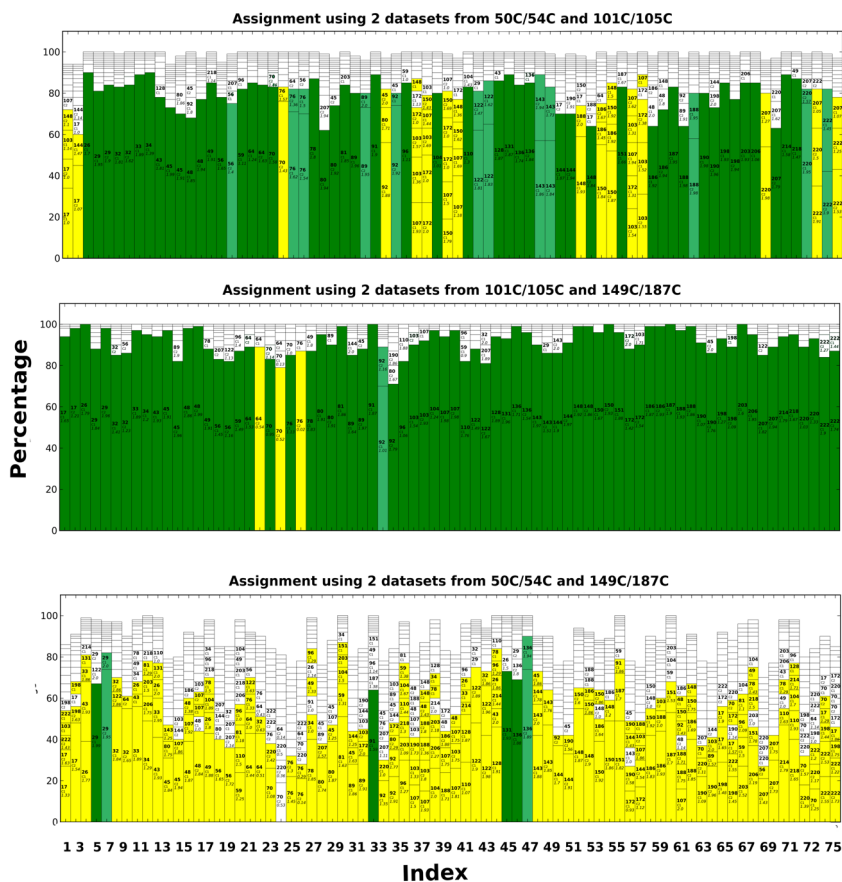


Figure. 2.9 PARAssign output for the three combinations of two of the complete, simulated datasets.

Assignment of methyl groups with experimental datasets

First, the methyl resonances of ntd-HSP90 were assigned using experimental datasets from the three paramagnetic centers. The PCS pairing procedure yielded 72, 74 and 44 of the expected 76 PCS for the mutants 50C/54C, 101C/105 and 149C/187C, respectively. As we had collected the spectra of Ile labeled protein samples separately from those labeled with Leu and Val (see Materials and Methods), it was possible to identify Ile methyl groups but we could not distinguish Leu and Val methyl groups. Therefore, in the input file the two possible residue types were entered as “LEU VAL”. In case of other methyl group labeling schemes, the residue type can be entered with its amino acid code, for example, “MET” for the methionine methyl group labeling. PARAssign runs were performed as described above, by using either the data from three, two, or one paramagnetic center(s). The results are listed in **Table 2.2** and PARAssign output using experimental datasets of three mutants is displayed **Figure 2.10**.

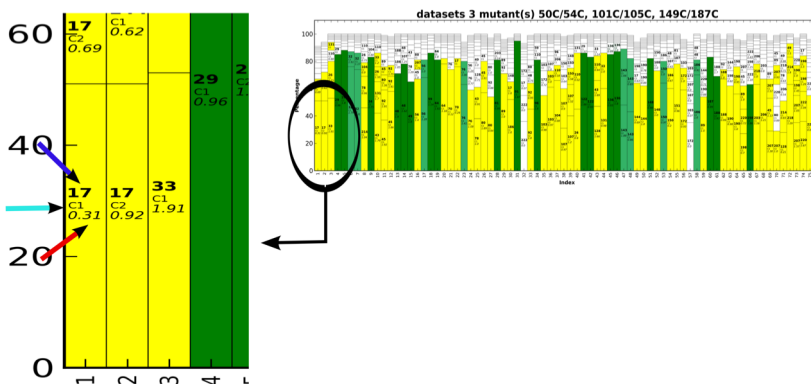


Figure 2.10 PARAssign output for the complete experimental dataset. The highly reliable assignments are shown in green and the proposed reliable assignments are in yellow. In this example, all the green assignments are correct and 95% of the yellow assignments contain the correct assignment (among 1 to 4 possibilities). In the close-up the blue, cyan and red arrows indicate the residue number, the stereospecificity and the scoring factor, respectively.

The tensor parameters are given in **Table 2.3** and **Figures 2.11** and **2.12**.

Table 2.3 Tensor parameters refined for the different combinations of datasets. In parenthesis are the standard deviations. The Euler angles α , β , γ are defined according to the z - y - z convention. A * indicates that there are other conformations. See also **Figures 2.11** and **2.12**.

	3 paramagnetic centers			2 paramagnetic centers			1 paramagnetic center					
	50C/54C	101C/105C	149C/187C	50C/54C – 101C/105C	50C/54C – 149C/187C	101C/105C – 149C/187C	50C/54C	101C/105C	149C/187C	50C/54C	101C/105C	149C/187C
ΔX_{h} (10^{-32} m ³)	8.1 (0.7)	7.5 (0.6)	8.2 (0.9)	7.8 (0.5)	7.6 (0.8)	7.2 (0.5)	7.6 (0.8)	7.6 (0.6)	8.5 (0.8)	7.3 (0.7)	7.3 (0.5)	8.3 (0.6)
ΔX_{h} (10^{-32} m ³)	2.0 (0.4)	2.4 (0.3)	2.4 (0.3)	2.1 (0.4)	2.2 (0.4)	2.4 (0.3)	2.2 (0.4)	2.0 (0.3)	2.2 (0.3)	2.4 (0.2)	2.0 (0.4)	2.2 (0.3)
α (°)	-144 (9)*	-177 (12)	15 (81)	-38 (5)	-113 (28)	-2 (38)	-174 (16)*	-47 (15)	118 (24)	30 (4)	160 (2)*	44 (12)
β (°)	146 (9)	103 (8)	142 (11)	31 (5)	64 (18)	97 (38)	104 (6)	151 (7)	153 (6)	98 (4)	98 (4)	153 (6)
γ (°)	21 (6)	176 (6)	19 (16)	-160 (5)	-136 (23)	159 (10)*	176 (10)	6 (7)	-1 (14)	-142 (14)	160 (2)*	-1 (14)
x (Å)	14.0 (2.3)	22.1 (5.4)	-3.8 (3.6)	13.0 (1.3)	14.9 (2.7)	-2.3 (6.2)	24.9 (0.4)	-1.6 (2.2)	-4.9 (1.3)	11.2 (1.7)	25 (0.4)	-4.9 (1.3)
y (Å)	-1.6 (2.6)	-9.0 (4.3)	-4.4 (2.7)	-1.0 (1.4)	-2.4 (2.6)	-4.5 (6.5)	-7.4 (1.3)	-5.8 (1.4)	-8.4 (1.7)	-1.6 (1.1)	-6.6 (1.7)	-8.4 (1.7)
z (Å)	24.4 (1.9)	4.2 (2.9)	0.1 (1.6)	24.5 (1.9)	22.3 (3.3)	-0.2 (5.1)	4.8 (1.3)	-1.4 (1.0)	-0.2 (1.0)	25.6 (0.7)	5.6 (1.1)	-0.2 (1.0)

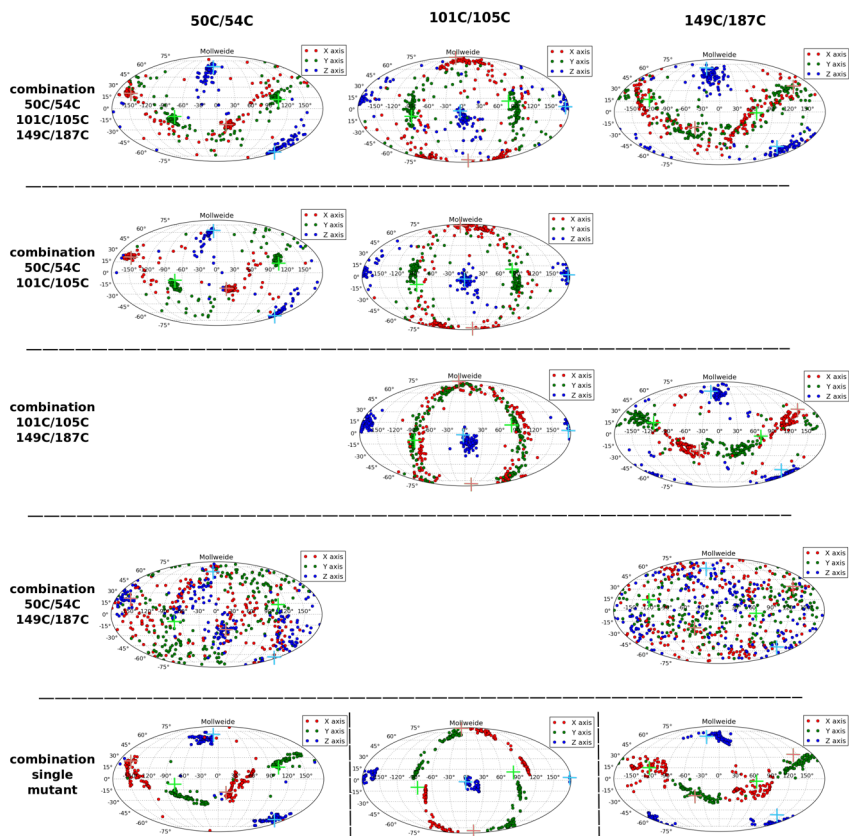


Figure 2.11. Mollweide sinusoidal representation of the tensor axes refined with PARAssign for all 7 combinations. The light crosses are the axis coordinates of the tensors refined with amide PCS, Table S1. A clear correlation can be observed between the number of highly reliable assignments for a given correlation and the quality of the tensor refinement.

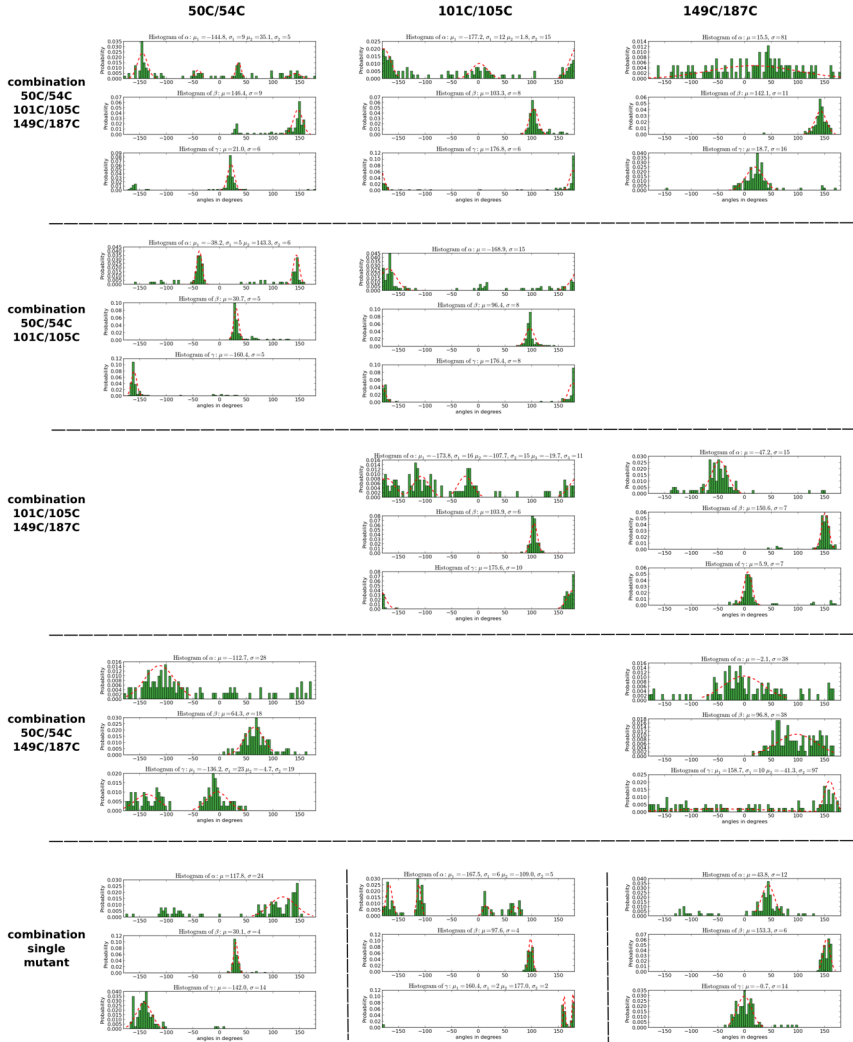


Figure 2.12. Distribution of $\Delta\chi$ Euler angles (convention z-y-z) for the 7 combinations. For all angle sets, the top, middle and bottom subplots are the histograms of α , β and γ , respectively. It should be noted that the yz Euler representation $(\alpha, \beta, \gamma) \setminus \beta \in [0, 180[$ is equivalent to $(\alpha+180, -\beta, -\gamma) \setminus \beta \in [-180, 0[$ yz Euler representation. Additionally, the xz and yz planar symmetries of $\Delta\chi$ lead to equivalence by 180° rotation.

In the results based on three datasets, all of the 27 highly reliable assignments were correct and 95% of the yellow assignments contained the correct assignment (1 to 4 possibilities). The assignment procedure was also performed with the input data of only two of the three paramagnetic centers, **Figure 2.13**. In each case, all the highly reliable (green) assignments were correct. Interestingly, however, the number of highly reliable assignments was very different for the three different combinations of paramagnetic

centers. With the combination (50C/54C-101C/105C), 46 methyl groups were assigned with high reliability, whereas only 9 assignments were highly reliable for the combination (101C/105C-149C/187C) and none for the combination (50C/54C-149C/187C). Furthermore, this combination yielded very few reliable (yellow) assignments. The much smaller input data set for mutant 149C/187C partly explains why fewer reliable assignments were obtained for combinations with this mutant. However, the large difference in assignment quality between the two combinations that contain the 149C/187C data must have another reason. We ascribe this to the relative orientations of the $\Delta\chi$ -tensors, **Figure 2.13**.

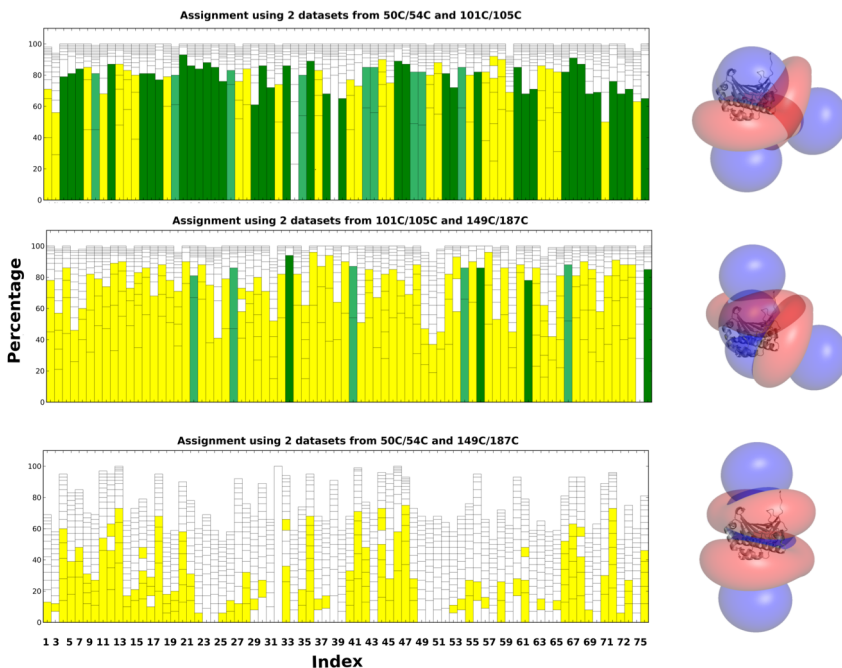


Figure 2.13 PARAssign output for the three combinations of two of the experimental data sets. Labels were removed for clarity. On the right the corresponding relative orientations of the two CLaNP-5 tensors are shown.

For parallel $\Delta\chi$ -tensors the assignment quality is poorer than for orthogonal $\Delta\chi$ -tensors, which was confirmed by a run of PARAssign with complete, simulated datasets for which the combination 50C/54C-149C/187C yielded many fewer highly reliable assignments than the two other combinations, **Table 2.2**. The assignment runs were also performed using data from only one paramagnetic center. As expected, the number of correct assignments drops dramatically. Nevertheless, PARAssign identified 1 to 6 highly reliable assignments for the two mutants 50C/54C and 101C/105C, which were all correct. These assignment calculations were performed using a PRE cutoff of 9 Å, **equation 2.1**, to prevent methyl groups for which the PRE leads to extensive broadening, being considered for assignment. The 149C/187C mutant did not yield

correct assignments because of the limited input set due to low quality of the spectra. Therefore, when using PARAssign with only one dataset, the quality of the input becomes essential to avoid false positive assignments. From this study, we conclude that with a dataset containing less than 60% of the expected PCS, PARAssign fails to give highly reliable assignments that are all correct. The highly reliably assigned nuclei from the two mutants with good input data correspond to methyl groups that are localized close to the lanthanoid, displaying sizeable PCS, as illustrated in **Figure 2.14**. Thus, the use of a single tag could be considered for localized assignment, for instance of methyl groups in an active site.

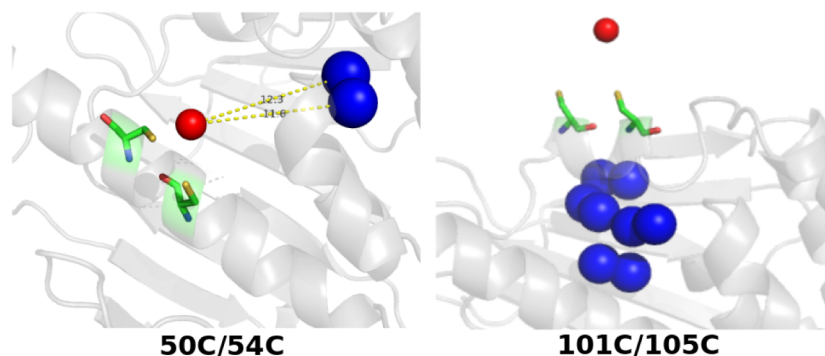


Figure. 2.14 Highly reliably assigned methyl groups, represented in blue spheres, using a single paramagnetic center for the assignment. The numbers in the 50C/54C panel indicate the distances from the lanthanoid to the methyl groups in Å.

The quality of the assignment is reflected in the accuracy and precision of the $\Delta\chi$ tensors parameters refined by PARAssign. The sinusoidal projection and histogram representation of the Euler angles of refined tensors are shown **Figures 2.11** and **2.12**, respectively. The combination (mutant 50C/54C - 101C/105C) leading to the best assignments is also the one for which tensor orientations are the best refined. This is to be expected, because high reliability is obtained only if the same assignments are found consistently over the 100 runs. The combinations of data sets that lead to poor assignment reliability also yield a large spread in the Euler angles, indicating a poor reproducibility of tensor determination over the 100 runs. A comparison of the precision of the 50C/54C tensor calculated when using the single data set as compared to the combination with 149C/C187C shows that the former is higher. This analysis strongly suggests that the addition of a relatively poor dataset of a parallel $\Delta\chi$ -tensor can worsen the assignment process.

Conclusions

The results presented in this study demonstrate possibilities and limitations for use of PCS in the assignment of protein methyl group resonances. The simulations show that, with complete input, the search problem can be solved with data from three and even from only two paramagnetic tags. The quality of the input data is vitally important. One

of the three tags yielded broader peaks and poorer spectra, resulting in a significant decrease in the number of paramagnetic peaks that could be paired with their diamagnetic counterparts. Such an incomplete PCS data set strongly affects the quality of the assignment. In our hands, CLaNP attached to two consecutive turns of an α -helix gives the best results with minimal disturbance of the protein structure. It was also found that the $\Delta\chi$ tensors of different tags should not be oriented along the same main axis, because this results in correlated PCS and reduced assignment quality. These factors can easily be taken into account when designing the tag positions. With two-armed tags such as CLaNP-5, the lanthanoid location and approximate directions of the axes of the $\Delta\chi$ tensor can be estimated well. It is then possible to predict the PCS for different tags and use these in PARAssign to predict whether complete assignment will be achieved if good spectra can be acquired.

PARAssign version 2 runs multiple times with a small part of the input data removed to yield more robust assignments. It was demonstrated that, with data from two or three independent tags, assignments classified as highly reliable are correct and that in those classified as 'yellow', the correct assignment is nearly always among the few suggested possibilities. Such information could be combined with data from NOE, PRE or mutagenesis experiments to find a unique assignment.

The critical point in obtaining complete PCS input sets is the pairing of peaks in the paramagnetic and diamagnetic spectra. Clearly, the larger the number of peaks in a spectrum, the more complicated this process is, in particular for methyl resonances, which show little dispersion. For larger systems, preparing several samples with only one type of amino acid labeled will reduce the overlap^{116,118}. Labeling a single methyl group per amino acid also helps²⁹. All the PCS can then be combined in PARAssign, with the amino acid types indicated. Another approach to assist the pairing of peaks is to prepare a sample with the same tag containing a weakly paramagnetic lanthanoid, such as Eu³⁺. The $\Delta\chi$ tensors of different lanthanoids in the same tag are oriented to good approximation in the same direction, so that a weaker lanthanoid results in a resonance on the correlation line between the diamagnetic and paramagnetic peak, at a position determined by the ratio of the sizes of the $\Delta\chi$ tensors of the stronger and weaker lanthanoids. With Eu³⁺ the PCS are small and more readily attributable to the corresponding diamagnetic peak.

If some assignments are already known, these can be introduced in PARAssign to improve the assignment results. To confirm the correctness of the available assignments of the methyl resonances in ntd-HSP90, we mutated L45, L80 and L89 one by one to Ile. When these assignments were fixed in the PARAssign calculations, the number of highly reliable assignments improved by up to 15% for the best combination.

A single tag cannot be used to obtain an overall assignment, because of the symmetric nature of the paramagnetic effect, in particular for axial tensors. However, in

combination with the protein structure, resonances of nearby methyl groups can still be assigned. This approach can be helpful to highlight a certain region of a protein, for example the active site in an enzyme. The assigned nuclei can then be used to report on dynamics or structural changes that occur in its proximity.

Acknowledgements

We thank Dr Marta Carneiro for her help to assign the methyl groups of ntd-HSP90 with sparky software.

Chapter 3

Lanthanoid generated RDC for amide
assignment with PARAssign

Abstract

For some biological systems NMR assignment remains a challenging step. However, it is essential for further NMR studies, such as protein structure determination, dynamics, interaction with substrates. PARAssign software has been shown to assign efficiently ntd-HSP90 methyl groups using two PCS datasets and a crystal structure. Up to 60% of methyl groups were assigned with high reliability. Assignment of amide resonances using experimental datasets measured on pseudoazurin (125 residues) showed the ability of PARAssign to avoid false positives, even though the performance was low, with 38 amide groups were assigned with high reliability. Additionally, three paramagnetic centers were necessary to assign such a small protein. In this study, it was explored whether the number of required paramagnetic centers for amide groups assignment could be reduced. Amide assignment is of particular interest for proteins with selective amide labelling, because standard sequential methods cannot be applied in that case. It is investigated whether selective labelled amide nuclei in ntd-HSP90 can be assigned using sets of PCS and RDC generated by a single paramagnetic center. On the basis of simulated data with added errors, it is shown that with a combination of PCS and RDC the number of highly reliable assignments can significantly be enhanced, as compared to PCS alone. The possible hurdles for application of this approach to experimental data sets are discussed.

Introduction

Paramagnetic effects, and especially PCS, have been proven to be powerful for protein NMR assignment^{41-43,101}. PARAssign software uses PCS datasets and a crystal or NMR structure to assign NMR spectra of proteins. One of the limitations of such a tool is the protein sample preparation. To measure a set of PCS, two samples need to be prepared, tagged with a diamagnetic or paramagnetic tag. The engineering of double cysteine mutation on the surface of the protein for CLaNP-5^{64,65} attachment can be challenging, so minimizing the number of sites to be tagged is desirable. When tagged with a molecule that has an anisotropic component in the magnetic susceptibility, proteins experience a slight alignment that prevents the averaging of dipolar couplings to zero. Residual dipolar couplings (RDC) can then be measured as the difference between the couplings of two nuclei measured on the paramagnetic and diamagnetic protein samples^{119,120}. RDC report on orientation of the vectors connecting the nuclei. In the case of an amide group, the vector is aligned with the ¹H-¹⁵N bond, and, thus, the RDC can be used as structural restraints¹²¹. RDC can be measured using the same samples that are used to obtain PCS, providing a good way to get additional, independent data that can be used for assignment. Assignment of amides is of particular use for spectra of selectively labelled samples. Labelling of specific residues has been used to simplify backbone spectra and enable the study of bigger proteins¹²²⁻¹²⁵. The ability of PARAssign to assign selectively labeled residues without any knowledge of backbone assignment would allow skipping the tedious task of multidimensional through backbone assignment.

Incorporation of RDC in PARAssign

PARAssign uses the Hungarian method¹²⁶ for minimal cost assignment. The scoring function to populate the matrix is described by **equation 3.1**. This function, formerly used only with PCS, was modified to include RDC, **equation 3.2**. The tensor optimization function of PARAssign was not modified, to ensure that PCS remain the only data used for the tensor optimization. The accuracy of RDC is usually poorer than for PCS measurement, so a second scoring function S_i^{RDC} is defined for the RDC with a well function that depends on the user-defined well threshold error Γ_{RDC} (in Hz), **equation 3.3**. Additionally, a user-defined α (0 - 1) gives the weight of RDC compared to PCS. These modifications result in a final score that depend on the experimental RDC error, the user-defined weight given to RDC and the two scoring functions S_i^{PCS} and S_i^{RDC} , defined in **equation 3.3a** and **3.3b**. S_i^{RDC} allocates 0 if the predicted and experimental RDC differ less than the error, otherwise it yields the scoring factor as calculated in **equation 3.1**.

$$\text{Scoring}_{\text{factor}} = \frac{1}{P_c} * \sum_{i=1}^{P_c} \sqrt{\frac{(\delta_{dt,i}^{\text{pred}} - \delta_{dt,i}^{\text{exp}})^2}{(|\delta_{dt,i}^{\text{pred}}| + |\delta_{dt,i}^{\text{exp}}|)^2}} \quad dt = \text{PCS or RDC} \quad (3.1)$$

$$\text{score} = (1 - \alpha) * S_f^{\text{PCS}} + \alpha * S_f^{\text{RDC}} \quad (3.2)$$

$$S_f^{\text{RDC}} = \begin{cases} 0, & |\text{RDC}_{\text{pred}} - \text{RDC}_{\text{exp}}| < \text{error}_{\text{RDC}} \\ \text{Scoring}_{\text{factor}}, & |\text{RDC}_{\text{pred}} - \text{RDC}_{\text{exp}}| \geq \text{error}_{\text{RDC}} \end{cases} \quad (3.3a)$$

$$S_f^{\text{PCS}} = \text{Scoring}_{\text{factor}} \quad (3.3b)$$

Assignment with simulated datasets

The new scoring function was tested on simulated datasets. The idea here was to use PARAssign to assign a protein ^{15}N labelled only on some types of residue, such as isoleucine and alanine, for instance. Ntd-HSP90 (see **chapter 2**) contains 20 isoleucines and 15 alanines, yielding a total of 35 assignable amide groups. PCS and RDC were simulated using a tensor refined previously with Numbat software¹²⁷ based on experimental PCS measured on fully ^{15}N -labelled 101C/105C mutant tagged with Lu^{3+} -CLaNP-5 or Yb^{3+} -CLaNP-5, **Table 3.1**. Protons were added to the 3T0Z crystal structure¹¹¹ using h_added function in Pymol¹²⁸. Uniform random errors were added of 0.02 ppm, to PCS, and 1.0, 2.0 or 4.0 Hz, to RDC. The well function error threshold was set to the RDC error value. The weight of the RDC in the scoring function (α) was varied from 0% to 50%. At 0% RDC were effectively not used for assignment. The results of the simulations are reported in **Table 3.2**.

Table 3.1 $\Delta\chi$ tensor of mutant 101C/105C with Yb^{3+} -CLaNP-5 determined with experimental amide PCS and PDB entry 3T0Z.

Parameter	Value
$\Delta\chi_{\text{ax}}$ (10^{-32} m^3)	8.0 (0.2)
$\Delta\chi_{\text{th}}$ (10^{-32} m^3)	1.7 (0.3)
x (Å)	25.3 (0.2)
y (Å)	-9.6 (0.3)
z (Å)	6.8 (0.3)
α (°)	179 (1.5)
β (°)	88.0 (1.5)
γ (°)	171.0 (5)

Table 3.2. Assignment results. The highly reliable (class green) assignments are in bold (correct/total). The suggested assignments (class yellow) are in italic. Total number of nuclei was 35.

RDC error weight α	No RDC	1 Hz	2 Hz	4 Hz
0	9/9			
0.1	<i>14/24</i>	9/9	11/11	10/10
0.3		<i>23/24</i>	<i>19/22</i>	<i>19/23</i>
0.5		<i>17/17</i>	<i>15/15</i>	<i>12/12</i>
		<i>14/14</i>	<i>15/16</i>	<i>17/21</i>
		1/1	13/13	11/11
		<i>33/33</i>	<i>18/19</i>	<i>18/20</i>

In the absence of RDC information, the PCS data from a single tag allow nine out 35 amides to be assigned with high reliability. The assignment improved already by the inclusion of the RDC data with $\alpha = 0.1$, by obtaining more correct assignment suggestions in the yellow class (see **chapter 2**), from 14 out of 24 to 19 or more. With $\alpha = 0.3$, the use of RDC increased significantly the number of highly reliable assignments

(green class), from 9 without RDC to 17, 15 and 12 when using errors of 1, 2 and 4 Hz, respectively. An error of 2 Hz may be considered realistic. For that simulation, the 15 highly reliable assignments were all correct and in 15 cases out of 16, the correct assignment was among the suggestions in the yellow class. At first sight, the performance for the run with $\alpha = 0.5$ and only 1 Hz error looks much worse than expected, as only a single highly reliable assignment was found. However, a closer look shows that most of the yellow suggestions contain only one suggestion with a high scoring, although not high enough to fulfill the green criterion (**chapter 2**), **Figure 3.1a**. Besides, the ensemble of 100 tensor values from this simulation showed many clear outliers (**Figure 3.1b**), which points to the poor convergence for some PARAssign runs, preventing the classification of assignments into the green class. With the same input datasets, PARAssign was run once more but with a 2 Hz error threshold in the well function for the RDC. Interestingly, the quality of the assignment increased tremendously from only one highly reliable assignment to 23, which were all correct, **Figure 3.1c**. The improved fit correlates with a decrease in the number of tensor outliers, **Figure 3.1d**.

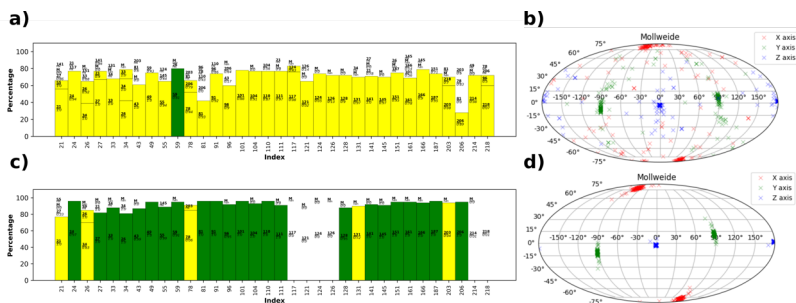


Figure 3.1 Assignments (a, c) and tensor angle distributions (b, d) for simulated data using PCS (with 0.02 ppm error) and RDC ($\alpha = 0.5$) from a single paramagnetic center, 101C/105C on nt-HSP90 with an RDC error levels of 1 Hz (a, b) or 2 Hz (c, d).

The well function must have a threshold of a minimum of 2 Hz no matter what is the accuracy of the input data. Below this threshold, the assignment worsens and hampers the refinement of the tensor based on PCS in the procedure. The use of input RDC with 4 Hz error led to assignment not as good as when the error is of 2 Hz but interestingly showed an improvement of 2-3 extra highly reliable green assignments compared to the assignment ran without any RDC, **table 3.1**.

Conclusion

The study using simulated data showed that in principle RDC can be used to improve the quality of the assignment and avoid false positives, when using a single paramagnetic center. The improvement is optimal when the error in the RDC is not higher than 1 Hz. Nevertheless, even with the addition of RDC using an error 2 Hz or even 4 Hz PARAssign provides more highly reliable assignments than using the same

PCS dataset without RDC. RDC can be measured with high precision (> 1 Hz), at least for small proteins. However, when crystal structures are used for the assignment, they could represent the limiting factor. Ntd-HSP90 seems to present an unfavorable example. An overlay of PDB entries 3T0Z and 3T0h shows large conformational changes in some of the secondary elements. RDC predicted for amide groups, obtained after building protons into these structures, differ up to 10 Hz for some regions. A priori, it is not clear which structure is the correct one. In such cases, it is best to do the assignment twice using the different structures as input. Furthermore, the accuracy of building hydrogens into a crystal structure is limited by the accuracy of crystallographic model and, thus, the X-ray data quality. A realistic error for the RDC is estimated to be 2 Hz, which is a combination of structure limitations and experimental error. In the presence of flexible parts in the protein, limiting the use of RDC for assignments to residues in α -helices and β -sheets could be advisable. This would have to be defined within PARAssign, by setting the well function error values for some residues much higher than for others, or by defining α values on a per-residue basis.

Chapter 4

Methyl group reorientation under
ligand binding probed by
pseudocontact shifts

Based on the research article: Lescanne M., Ahuja P., Blok A., Timmer M., Akerud T., Ubbink M. *Methyl group reorientation under ligand binding probed by pseudocontact shifts. (2018) J. Biomol. NMR (accepted)*

Abstract

Liquid-state NMR spectroscopy is a powerful technique to elucidate binding properties of ligands on proteins. Ligands binding in hydrophobic pockets are often in close proximity to methyl groups and binding can lead to subtle displacements of methyl containing side chains to accommodate the ligand. To establish whether pseudocontact shifts can be used to characterize ligand binding and the effects on methyl groups, the N-terminal domain of HSP90 was tagged with caged lanthanoid NMR probe 5 at three positions and titrated with a ligand. Binding was monitored using the resonances of leucine and valine methyl groups. The pseudocontact shifts (PCS) caused by ytterbium result in enhanced dispersion of the methyl spectrum, allowing more resonances to be observed. The effects of tag attachment on the spectrum and ligand binding are small. Significant changes in PCS were observed upon ligand binding, indicating displacements of several methyl groups. By determining the cross-section of PCS isosurfaces generated by two or three paramagnetic centers, the new position of a methyl group can be estimated, showing displacements in the range of 1 to 3 Å for methyl groups in the binding site. The information about such subtle but significant changes may be used to improve docking studies and can find application in fragment-based drug discovery.

Keywords

Isotope labeling; Pseudocontact shift; Methyl groups; NMR spectroscopy; Paramagnetic tag; Heat shock protein

Abbreviations

PCS, pseudocontact shift; HSQC, heteronuclear single quantum coherence; TROSY, transverse relaxation optimized spectroscopy; CLaNP-5, caged lanthanoid NMR probe 5; ntd-HSP90, N-terminal domain of heat shock protein 90

Introduction:

Fragment-based drug discovery (FBDD) has proven to be an effective method to develop medicinal drugs¹. FBDD is based on finding very small molecules that bind to the target weakly but with high efficiency (large binding energy per heavy atom). Such fragments are then elaborated to molecules with higher affinity, either by linking fragment hits or by growing them^{129,130}. For the elaboration of hits into lead compounds FBDD depends heavily on structural analysis of fragment-target complexes by X-ray diffraction of crystals or NMR spectroscopy. The former technique is most commonly used but NMR is an alternative for structure determination and offers complementary information. Structure determination of the complex by NMR requires a complete NOE analysis of protein and ligand, which is tedious but can be used in cases where X-ray crystallography fails⁴. Other, less demanding methods are based on transferred NOEs, paramagnetic relaxation enhancements (PRE) or pseudocontact shifts (PCS) to obtain information about the ligand bound state while benefitting from the narrow linewidths of the ligand in the free state¹³¹⁻¹³⁴.

We aim to investigate the possibilities of PCS to study ligand-protein interactions. Before, we demonstrated that ligands that are in fast exchange between bound and free state can exhibit transferred PCS caused by lanthanoid tags on the protein, which can be used to determine a low-resolution model of the ligand in the binding site, provided that a structure of the protein is available and under the assumption that ligand binding does not result in backbone conformational changes⁶. Fragments, as well as larger compounds often bind in hydrophobic pockets on the protein where they do not alter the positions of backbone atoms significantly. Methyl groups are often found in such pockets and, thus, are in direct contact with the ligand. They are prone to experience chemical shift changes due to changes in the chemical environment upon ligand binding, and may also more readily than backbone atoms show conformational changes due to the rotational freedom of side-chains²³. In this way, they can help to accommodate ligand binding, enabling it to form optimal interactions. Therefore, we wondered whether such changes could be detected by using PCS. These shifts are caused by the interaction of the nuclear spin and the spin of unpaired electron(s) in a paramagnetic center. They depend on the distance between the spin and the center to the third power as well as on the orientation of the spin in the frame of the anisotropic component of the magnetic susceptibility of the unpaired electron(s)^{135,136}. With a probe that is rigid relative to the protein and a proper diamagnetic control, the PCS can be predicted and measured with very high accuracy and small changes in the location of the spin relative to the paramagnetic center can result in measurable PCS changes. Methyl PCS can be observed in sensitive 2D NMR spectra, potentially also on large systems by applying selective labelling in a deuterated background and by using TROSY-based experiments^{25,26,137}.

Here, we describe the use of PCS as structural restraints to probe at the same time binding kinetics and structural changes of the protein ntd-HSP90 upon fragment binding. HSP90 is a target protein against cancer¹³⁸ and its ATP binding site located in the N-terminal domain (ntd) is targeted for inhibition¹⁶. HSP90 is a molecular chaperone essential to prevent client proteins from ubiquitin-proteasome system degradation. More than 200 client proteins of HSP90 have been identified, including oncoproteins¹³⁹. Therefore, HSP90 is a cancer-target protein and inhibitors have been found to bind the N-terminal domain and/or the C-terminal domain¹⁴⁰. Several potent molecules are clinical candidates for cancer treatment through inhibition of the ATPase activity and FBDD has been successfully applied to HSP90, which led to a clinical trial¹³⁹. We find that ligand binding is only marginally affected by attaching the two-armed lanthanoid tag CLaNP-5^{106,107} to ntd-HSP90 at three locations. Methyl group resonances show extensive chemical shift perturbations in the binding site, as well as further in the hydrophobic core of the protein. Several significant PCS changes are observed upon ligand binding, which can be interpreted as movements of the methyl groups of a few Ångström. These changes can be translated into structural restraints that may be used in ligand docking studies.

Materials and methods

Sample preparation

Three double cysteine mutants of the ntd-HSP90 were designed on the surface of the protein, S50C/D54C, A101C/N105C and T149C/I187C²². Ntd-HSP90 does not have any native cysteines. The protein was produced labelled with Leu- δ 1- δ 2/Val- γ 1- γ 2- [¹³CH₃] and purified and tagged with CLaNP-5 according to a published protocol²². CLaNP-5 was synthesized as described before^{106,107}.

NMR titration

Ntd-HSP90 in 50 mM Tris-HCl and 50 mM NaCl buffer, pH 7.7, was titrated with 4-(2-Fluorophenyl)-2-pyrimidinamine, **1**, **Figure 4.1**, a 189 Da known ligand of ntd-HSP90 that was kindly provided by AstraZeneca (Göteborg, Sweden). Titrations were performed with three ntd-HSP90 mutants, S50C/D54C, A101C/N105C and T149C/I187C tagged with Lu³⁺-CLaNP-5 or Yb³⁺-CLaNP-5. The concentrations of S50C/D54C, A101C/N105C T149C/I187C were 20 μ M, 103 μ M and 65 μ M, respectively, for both diamagnetic and paramagnetic forms of the protein. Concentrations (μ M) of **1** for titrations with S50C/D54C, A101C/N105C and T149C/I187C were [0, 39, 59, 88, 132, 198, 296, 444, 667, 1000], [0, 20, 40, 121, 364, 1111] and [0, 40, 89, 200, 442, 665, 1008, 1897], respectively. The NMR sample volume was 595 μ L for all samples, and dilution was neglected, because the biggest volume of ligand solution added was less than 5 μ L. ¹³C-¹H HSQC¹¹¹⁻¹⁴³ spectra were acquired at each titration point, on a Bruker Avance III 800 MHz spectrometer, equipped with a cryogenically cooled TXI-probe head, operating at 298 K. Spectra

were processed with nmrpipe¹⁰⁸ using the exponential EM apodization function for analysis with TITAN¹⁴⁴.

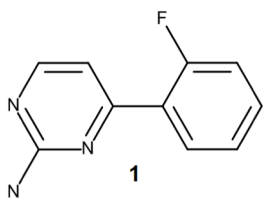


Figure 4.1. Structure of **1** used to titrate ntd-HSP90

Assignments

Methyl groups assignments have been performed before with traditional through-backbone NMR techniques and confirmed by PARAssign⁴³. PARAssign provided the stereo-specific assignment with high reliability for 14 Leu/Val methyl groups²², based on PCS generated by CLaNP-5 attached at two distinct positions, S50C/D54C and A101C/N105C.

Cross-section refinement

A home-written python script was used to define possible locations of the bound methyl groups locations. PCS iso-surfaces were calculated for a grid of 15x15x15 Å, with 100 points per dimension and centered on the methyl group position in the crystal structure of free ntd-HSP90 (PDB entry 3t0h¹⁰). Cross-sections of iso-surfaces from different tags were defined by finding the positions within the grid that matched the required PCS of all tags within an error of 0.02 ppm (0.03 ppm in one case, see below).

Q and Qa factors

Q and Qa factors were used to quantify deviation between experimental and predicted data. Q and Qa were calculated according to **equation 4.1** and **4.2**, respectively. Q is the usual measure for goodness of fit, Qa is, however, less sensitive to bias toward cases in which the predicted value is much larger than the observed one, as compared to the opposite case, in which the predicted value is much smaller than the observed one²⁹. In cases of a good fit, $Qa \approx 0.5Q$.

$$Q = \sqrt{\frac{\sum(\delta_{PCS,i}^{pred} - \delta_{PCS,i}^{exp})^2}{\sum(\delta_{PCS,i}^{exp})^2}} \quad (4.1)$$

$$Qa = \sqrt{\frac{\sum(\delta_{PCS,i}^{pred} - \delta_{PCS,i}^{exp})^2}{\sum(|\delta_{PCS,i}^{exp}| + |\delta_{PCS,i}^{pred}|)^2}} \quad (4.2)$$

Results

Tagging effects

Ntd-HSP90 was tagged at three sites using the Caged Lanthanoid NMR probe #5 (CLaNP-5), containing either Lu³⁺ as a diamagnetic control or Yb³⁺ as a paramagnetic

center. The tagging sites, double mutants 50C/54C, 101C/105C and 149C/187C, have been described before²². Methyl ^{13}C - ^1H HSQC spectra of WT and CLaNP-5(Li^{3+}) tagged mutants are very similar except for the resonance of a few methyl groups very close to the tags²², indicating that the tags do not have significant effects on the structure of the protein. A first comparison of the methyl group spectra of the paramagnetic and diamagnetic samples shows an increased dispersion of the resonances for the paramagnetic samples, **Figure 4.2**.

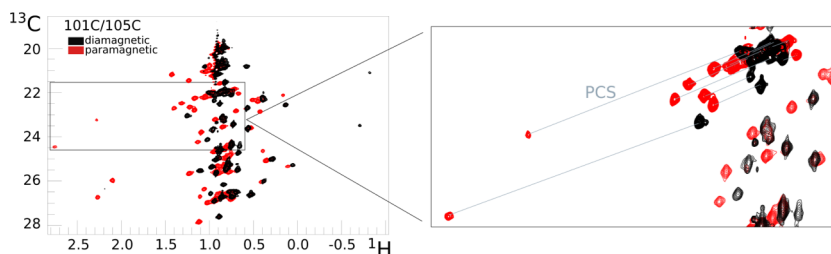


Figure 4.2. Enhancement of spectral dispersion by PCS. An overlay is shown of Leu/Val methyl HSQC spectra of mtHSP90 101C/105C tagged with CLaNP-5 loaded with Li^{3+} (black contours) or Yb^{3+} (red contours). The inset shows a detail and the lines connect equivalent resonances.

For example, a crowded spectral region with a width of 0.5 ppm in the ^1H dimension in the spectrum of the diamagnetic sample, disperses over 1.0 ppm for mutant 50C/54C and up to 3.0 ppm for mutant 101C/105C in the spectrum of the paramagnetic sample. Increased dispersion is of interest for methyl group HSQC spectra, because the resonances are often more crowded than in amide HSQC spectra, in particular for Leu and Val methyl hydrogens and to a lesser extent methyl carbons. In principle, lanthanoids with larger paramagnetic effects (Tm^{3+} , Dy^{3+} , Tb^{3+}) provide even more dispersion but also cause considerable paramagnetic relaxation farther from the metal than in the case of Yb^{3+} . Thus, such lanthanoids are more appropriate to generate dispersion in spectra of bigger proteins. Magnetic susceptibility tensors $\Delta\chi$ were refined previously using amide proton PCS²². Methyl group PCS were predicted based on these tensors and the ligand-free structure (PDB entry 3t0h¹⁰) and compared to the experimental ones. For mutants 50C/54C and 101C/105C most predicted PCS fit the experimental values well, **Figure 4.3**.

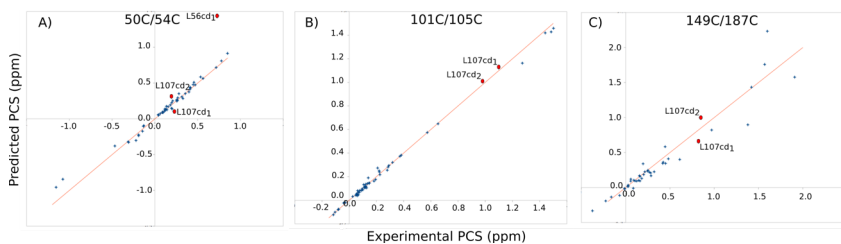


Figure 4.3. Prediction of methyl group PCS. The PCS for the Leu/Val methyl groups of ntd-HSP90 were predicted using the published amide based $\Delta\chi$ tensor sizes and orientations and lanthanoid locations²² and the structure with PDB entry 3t0h¹⁶ and plotted against the experimental PCS. No fitting was performed. A) For mutant 50C/54C the Q factor (eq. 1) is 0.28 ($Q_s = 0.14$, eq. 2). After exclusion of the PCS for L56 and L107 $Q = 0.14$ and $Q_s = 0.07$. B) For mutant 101C/105C $Q = 0.05$ ($Q_s = 0.025$). C) For mutant 149C/187C $Q = 0.24$ ($Q_s = 0.12$). The red line represents a perfect correlation.

Mutant 149C/187C shows a poorer fit but that is in line with the results for the amides, which was attributed to the fact that the tag crosslinks two β -strands and appears to assume two conformations²². To illustrate how differences between experimental and predicted PCS translate into differences in the expected locations of the nuclei in the protein, cross-sections of experimental PCS iso-surfaces for the different mutants were calculated, **Figure 4.4**. The iso-surface identifies all locations around a paramagnetic center with a given PCS. The cross-sections are the overlap areas of two or three iso-surfaces. A cross-section was calculated such that its thickness reflects 0.02 ppm uncertainty. Thus, large cross-sections, **Figure 4.4a**, indicate a weak PCS gradient, as is observed far from the paramagnetic centers. Thin cross-sections, **Figure 4.4b**, report on a steep PCS gradient, closer to the paramagnetic center or close to where the PCS changes sign. Cross-sections for the iso-surfaces of pairs of mutants are shown in **Figure 4.4**, panels A and B, for the methyl groups from residues Leu70 and Val150. The predicted PCS of these methyl groups match the experimental PCS within 0.02 ppm. The methyl groups observed in the crystal structure are located at the intersection of the two cross-sections, i.e. at the position that matches the experimental PCS for all three mutants within 0.02 ppm. For some methyl groups the discrepancy between experimental PCS and PCS predicted on the basis of the crystal structure is larger than 0.02 ppm. For instance, the experimental PCS of Val136 γ_1 and γ_2 are -1.03 and -1.12 ppm, respectively, and the predicted values show deviations of 0.19 (-0.84 ppm) and 0.16 ppm (-0.96 ppm) for mutant 50C/54C for γ_1 and γ_2 , respectively. Because of the strong PCS gradient, these large PCS deviations translate in only a small displacement of about 0.6 Å as compared to the crystal structure, **Figure 4.4c**. In the case of the mutant 101C/105C, the experimental PCS of Val136 γ_1 and γ_2 equal 0.27 and 0.27 and the deviations from the predicted values are only 0.03 (0.23 ppm) and 0.05 ppm (0.22 ppm), respectively, yet these translate into sizeable displacements of 1.5 Å, due to the weak PCS gradient at this position. This remarkable difference is visualized in **Figure 4.5**. Thus, these findings demonstrate that deviations larger than the

measurement error of PCS (usually 0.02 ppm or less) can be caused by very subtle differences in structure, whereas small deviations may still reflect a more sizeable mismatch between observed PCS and the structure used for PCS prediction. The PCS gradient at the methyl group position can be used to evaluate the structural relevance for observed differences between experimental and predicted PCS.

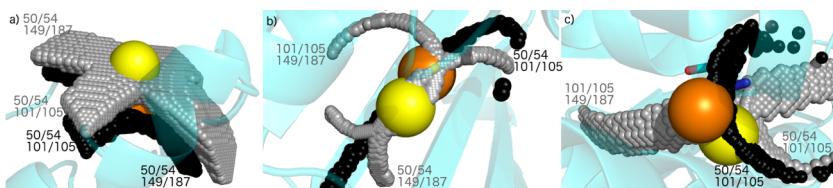


Figure 4.4. Cross-sections of iso-surfaces of experimental PCS. Yellow and orange spheres represent the crystal structure locations of methyl groups, centred on the carbon methyl group with radius of 1 Å, for Val γ 1/Leu δ 1 and Val γ 2/Leu δ 2, respectively. Grey and black spheres represent the experimental PCS cross-sections for Val γ 1/Leu δ 1 and Val γ 2/Leu δ 2, respectively. Each cross-section was calculated using 0.02 ppm error on the PCS, for a cubic grid with sides of 50 points over 5 Å, centred on the methyl group of interest. (a) Leu70 methyl groups. The two grey and black areas are cross-sections from mutant 50C/54C with mutant 101C/105C and mutant 50C/54C with 149C/187C for Leu70 δ 1 and δ 2, respectively. (b) Val150 methyl groups. The grey areas are the γ 1 cross-sections from mutant 101C/105C with mutant 149C/187C and from mutant 50C/54C with 149C/187C, the black area is the V150 γ 2 cross-section for mutant 50C/54C with mutant 101C/105C. (c) Val136 methyl groups. In black spheres the PCS iso-surfaces cross-section of mutant 50C/54C with mutant 101C/105C is shown for methyl γ 2. In grey spheres are the two PCS iso-surface cross-sections of mutant 50C/54C with 101C/105C and mutant 101C/105C with mutant 149C/187C. Both crystal structure methyl groups are on the edge of the cross-section. Methyl γ 1 is 3.1 Å from the triple cross-section area centre. For the γ 2 methyl, the minimal distance from the crystal structure location to the cross-section is 1.3 Å.

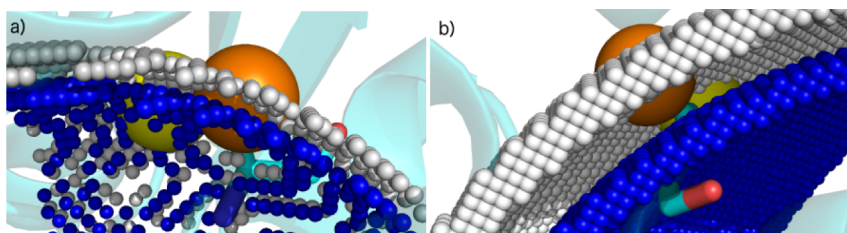


Figure 4.5. PCS gradient effects. In light grey are the iso-surfaces of the predicted PCS, in blue are the iso-surfaces of the experimental PCS, for Val136 γ 2, represented as an orange sphere. All iso-surfaces were calculated with an uncertainty of 0.01 ppm, for clarity. a) Iso-surfaces for mutant 50C/54C, the difference between the two surfaces (0.6 Å) represents a PCS difference of 0.16 ppm. b) Iso-surfaces for mutant 149C/187C, the difference between the two surfaces (1.5 Å) represents a PCS difference of 0.05 ppm.

For residue Leu107, the predicted PCS correlated poorly with the experimental ones for mutants 50C/54C and 149C/187C. The predicted PCS are clearly distinct for the δ 1 and δ 2 methyl groups (0.072 ppm and 0.252 ppm for mutant 50C/54C and 0.68 ppm and 0.94 ppm for mutant 149C/187C), whereas the experimental values are very similar (0.221 ppm and 0.194 ppm for mutant 50C/54C and 0.822 ppm and 0.842 ppm for mutant 149C/187C). The similarity of the two values suggests a kind of

averaging. Although the linewidth of the resonances in the diamagnetic sample also suggests a form of exchange, the δ_1 and δ_2 methyl groups have distinct chemical shifts, indicating that they do not populate the three rotamers of the side chain equally, as this would result in the same, averaged chemical shift for both δ methyl groups. To establish whether population of more than one rotamer explains the deviating PCS, the populations of each of the three rotamers were determined. Each rotamer was modelled in the structure (ignoring steric clashes) and the PCS were predicted. The best fit was found for an exchange between two rotamers populated at 53% and 47%, with the third rotamer not being populated, **Figure 4.6**. The Qa fit quality factor (see eq. 2) for Leu107 drops from 0.14 to 0.03 when predicted PCS are calculated as a combination of rotamer 2 ($+0^\circ$) and rotamer 3 ($+120^\circ$), for mutants 50C/54C and 149C/187C. For mutant 101C/105C, rotamer 2 fits the data well and admixture of rotamer 3 reduces the fit quality. However, Leu107 is located very close to the tag in this mutant, so either the methyl group location or the exchange populations of rotamers could be influenced by the CLaNP tag. It is interesting that the PCS seem to provide evidence for rotamer exchange and can be used to estimate populations. Such dynamics is not obvious from the chemical shift, unless the side chain sample all rotamers equally.

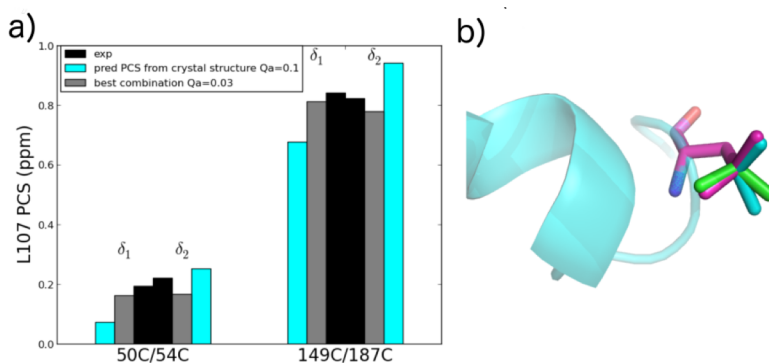


Figure 4.6. a) Experimental (black) and predicted PCS for the two methyl groups of L107 for rotamer 2, observed in the crystal structure (PDB entry 3t0h) in cyan and for the combination of rotamer 2 and 3 in grey, for the two mutants 50C/54C and 149C/187C. b) Residue L107. The rotamer found in the crystal structure is shown in cyan sticks. Rotamers 1 and 3 are in green and magenta, respectively. The rotamers were generated with Pymol. Note that the idealized, generated rotamers shows slightly different angles compared to the crystal structure rotamer.

Ligand titration

First, ntd-HSP90 mutants tagged with diamagnetic CLaNP-5 were titrated with the weakly binding ligand **1**. Three regions of the protein exhibit chemical shift perturbations (CSP) upon titration, **Figure 4.7**, revealing a localized binding site but

with small CSP being observed relatively far away in the core of the protein, **Figure 4.8**. The affected regions are similar for the three mutants, **Figure 4.7**, suggesting that the tag at different locations does not alter the binding location.

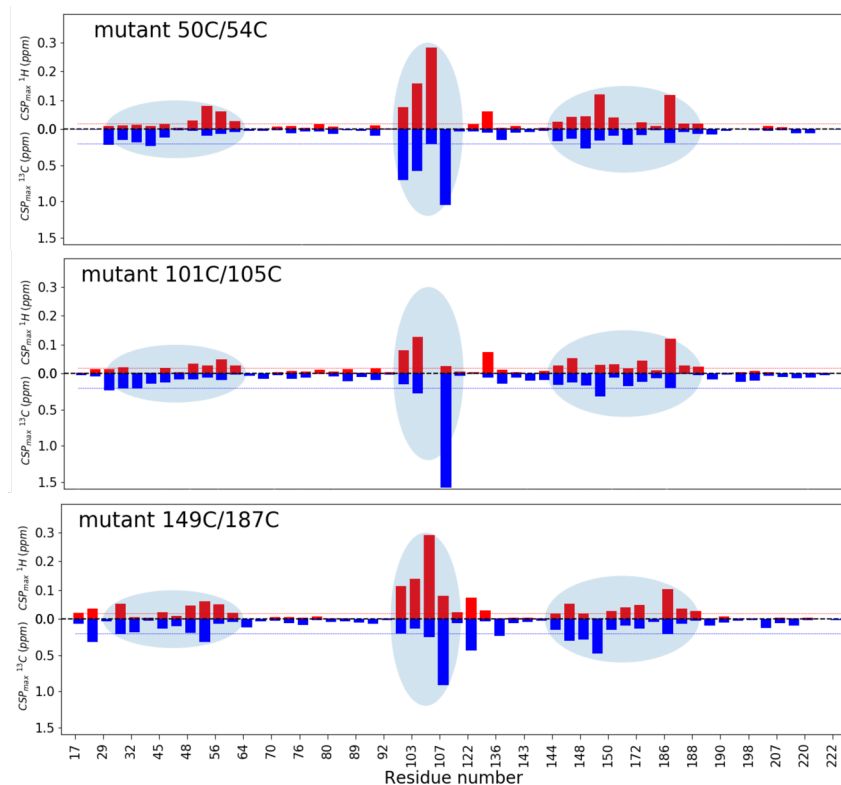


Figure 4.7. Ligand binding. |CSP| are plotted for methyl group resonances upon binding of **1**. The ^1H and ^{13}C CSP are shown in red and blue bars, respectively. The CSP have been extrapolated to the 100% bound state. The red and blue lines mark 0.02 ppm and 0.2 ppm, respectively. The light blue ovals highlight the most perturbed methyl groups.

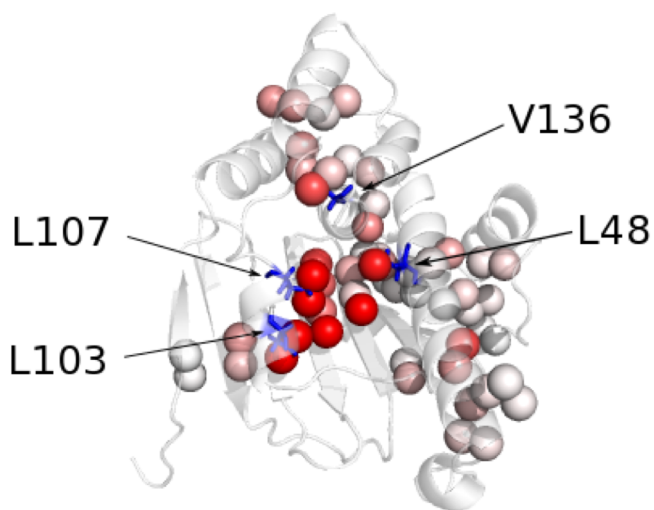


Figure 4.8. CSP map for binding of **1** to ntd-HSP90. The spheres represent the Leu and Val methyl groups, coloured in a gradient of white to red for increasing CSP as observed in the titration of mutant 50C/54C tagged with CLaNP-5 (Lu^3). Plotted CSP were calculated as the sum of 1/10 of the carbon CSP and the proton CSP. The residues L48, L103, L107 and V136, harboring a methyl group with $|\Delta\text{PCS}| > 0.04$ ppm for at least two mutants, are shown in blue sticks. The structure is taken from PDB entry 3t0h.

Moreover, most peaks show similar CSP directions in the two-dimensional HSQC spectra, **Figure 4.9**, indicating similar changes in the chemical environment of the methyl groups upon ligand binding. An exception is the resonance for Leu103 $\delta 1$, for which the CSP is not the same for all mutants. The discrepancy is largest for mutant 101C/105C, in which the Leu is very close to the tag, being located in between the two engineered Cys residues. Thus, the tag at this position appears to have some effect on ligand binding.

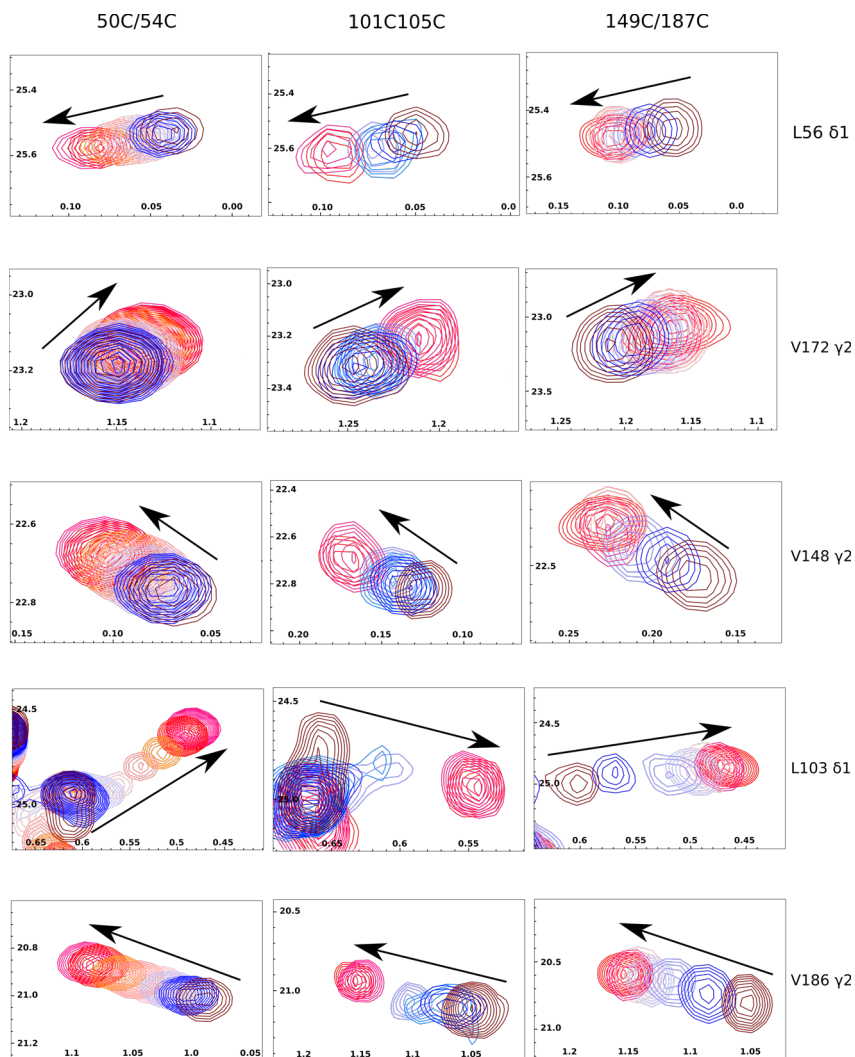


Figure 4.9. CSP patterns. Details of the overlay of the HSQC spectra of ntd-HSP90 titrations with **1** are shown for 5 methyl group resonances. The arrows indicate the direction of shifts. Except for methyl group L103 δ 1, the CSP patterns are similar for the three mutants.

Binding parameters

The dissociation constant K_D , as well as the dissociation rate constant (k_{OFF}) were calculated with TITAN software¹⁴⁴ using a two-site binding model. A global fit of the resonances of five residues was performed and then one at the time was taken out and the fit repeated. The ranges of values obtained in this way are reported in **Table 4.1** and provide a more realistic error range than the fit error. Differences are observed between the different mutants, suggesting that the affinity is influenced by the tags to some degree. K_D values of 200 μ M and 150 μ M were used for the calculation of the

100% bound state CSP for mutant 50C/54C and 101C/105C respectively. For these two mutants, a 50 μM variation of the K_D results in a change of the extrapolated CSP of 0.01 ppm at most, half the error used for the further calculations. For mutant 149C/187C a K_D of 50 μM was used to extrapolate the CSP.

Table 4.1. Parameters for binding of **1** to ntd-HSP90 derived from the titrations with TITAN software, for the three mutants.

	K_D (μM)	k_{off} (s^{-1})
50C/54C	195-238	1500-2300
101C/105C	140-171	1075-1300
149C/187C	41-45	1300-1900

Methyl group re-orientation

The titrations were repeated with ntd-HSP90 tagged with CLaNP-5 (Yb^{3+}). The binding characteristics were the same as for the diamagnetic sample. To obtain the PCS for the ligand bound state, the maximal CSP, representing ntd-HSP90 fully saturated with **1**, were calculated for both diamagnetic and paramagnetic proteins by extrapolation from the CSP data in the last point of the titration, by using the K_D values. The extrapolated CSP were subtracted to obtain the PCS of the bound state. Ligand binding causes changes in some PCS up to 0.1 ppm, **Figure 4.10**.

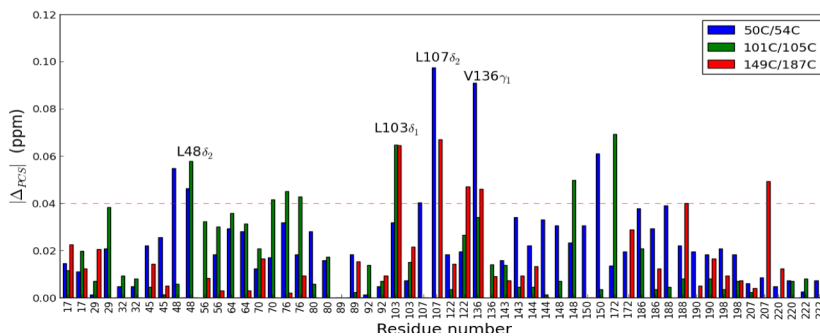


Figure 4.10. PCS changes upon ligand binding. $|\Delta\text{PCS}|$ for Leu/Val methyl groups for the bound state are shown as blue, green and red bars for the mutants 50C/54C, 101C/105C and 149C/187C, respectively. The red dashed line represents the threshold of 0.04 ppm.

Methyl groups exhibiting significant ΔPCS , larger than 0.04 ppm for at least two of the paramagnetic centers, were selected for further analysis, comprising L48 $_{\delta_2}$, L103 $_{\delta_1}$, L107 $_{\delta_2}$ and V136 $_{\gamma_1}$, **Table 4.2**.

Table 4.2. Significant ΔPCS (ppm). Methyl groups for which two $|\Delta\text{PCS}| > 0.04$ ppm are listed.

	50C/54C	101C/105C	149C/187C
L48 $_{\delta_2}$	-0.046	-0.058	No data
L103 $_{\delta_1}$	-0.032	-0.065	0.064
L107 $_{\delta_2}$	0.097	No data	0.067
V136 $_{\gamma_1}$	0.091	0.039	-0.046

These residues surround the binding site of ntd-HSP90, **Figure 4.8**. The presence of significant Δ PCS for methyl groups in the binding site suggest that ligand binding causes an adjustment in the position of the methyl groups to accommodate the ligand. We wondered whether the PCS can provide information about the distance range of the rearrangement as well as the new location of the methyl groups. Using the PCS of the bound form the possible new positions can be calculated as an iso-surface of the new PCS around the lanthanoid. By determining the cross-sections of such iso-surfaces from two or even three mutants, the location space can be reduced, as was shown above for the free protein.

All cross-sections for the ligand bound state could be determined using an uncertainty of 0.02 ppm, except for Leu107, for which the uncertainty had to be raised to 0.03 ppm to find a cross-section. The cross-sections are shown in **Figure 4.11** in grey (γ 1/ δ 1) and black (γ 2/ δ 2) for the PCS of free protein and in yellow (γ 1/ δ 1) and orange (γ 2/ δ 2) for the PCS observed for the bound state.

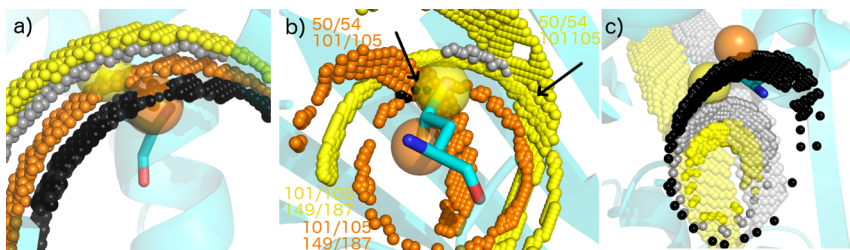


Figure 4.11. Positions for methyl groups in the ntd-HSP90 bound to ligand 1. Small spheres represent PCS iso-surfaces cross-section in the free state, grey for Val γ 1/Leu δ 1 and black for Val γ 2/Leu δ 2, and bound state, yellow for Val γ 1/Leu δ 1 and orange for Val γ 2/Leu δ 2. The residues as found in the crystal structure of the free protein (PDB entry 3t0h¹⁶) are shown in cyan sticks, with the methyl groups in the analogous colors shown as 0.7 Å radius spheres centred on the carbon methyl. (a) Leu48, with cross-sections for PCS from mutants 50C/54C and 101C/105C. The minimal distance from the crystal structure conformation to the free cross-sections is 0.8 Å (δ 1) and 0.6 Å (δ 2), and to the bound cross-sections is 1.3 Å (δ 1) and 1.2 Å (δ 2). (b) Leu103, the small yellow and orange spheres represent the cross-sections of mutant 50C/54C with 101C/105C and 101C/105C with 149C/187C. The triple-mutant cross-sections for the bound state are indicated with arrows. The grey and black spheres represent the triple-mutant cross-sections for the free state, at a distance of 1.6 Å (δ 1) and 2.3 Å (δ 2) from the crystal structure positions. (c) Val136, the small spheres represent the cross-section for methyl group γ 1 between mutant 50C/54C and 101C/105C in bright colors and between mutant 101C/105C and 149C/187C in soft colors. The orange cross-section is not shown because the γ 2 methyl group showed insignificant PCS changes.

For Leu48, iso-surfaces could be calculated for mutants 50C/54C and 101C/105C and the cross-sections are shown for the δ 1 and δ 2 methyl groups, **Figure 4.11a**. The conformation of the Leu in the free protein is shown in sticks. The Leu methyl groups need to move at least 1.3 Å to move into the cross-section area of the bound state. Similarly, **Figure 4.11b** shows the cross-sections for Leu103. In this case, data from all three mutants result in only two possible positions for the Leu that can match

simultaneously the PCS of both methyl groups to the experimental ones. The center of the triple mutant cross-section area for the $\delta 1$ methyl is shifted by 3 Å upon binding of the ligand, whereas for the $\delta 2$ methyl it does not change, which could indicate that the sidechain rotates. Similarly, the $\gamma 1$ methyl group of Val136 experiences a significant change in PCS, whereas the $\gamma 2$ methyl group does not. The cross-section center of the $\gamma 1$ methyl group for mutant 50C/54C with mutant 101C/105C shifted 1.2 Å from the position in the free protein. In **Figure 4.11c**, it can be that the cross-section area for the bound state (yellow) has moved relative to the one for the free state (light gray). Inclusion of the cross-section of mutant 101C/105C and 149C/187C did not reduce the cross-section area much (soft colors in **Figure 4.11c**). For Leu107, the two iso-surfaces, from mutants 50C/54C and 149C/187C are similar because the two tensors happen to be nearly parallel²². Consequently, the cross-section is a curve area rather than a line like in the cases discussed above **Figure 4.12**.

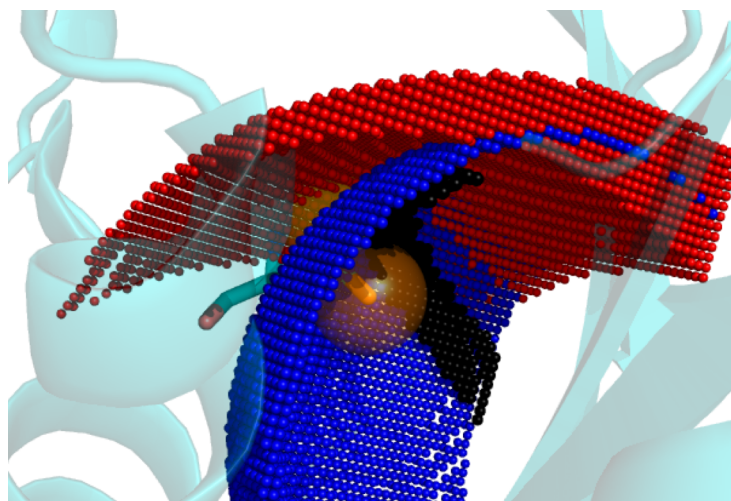


Figure 4.12. Leu107 $\delta 2$ methyl group. The blue and red spheres represent the $\delta 2$ iso-surfaces for the experimental PCS of the bound state for mutants 50C/54C and 149C/187C, respectively. The cross-section between these two iso-surfaces is obtained using an uncertainty of 0.03 ppm. In the free state, the cross-section between 50C/54C and 149C/187C, in small black spheres, is at a minimum distance of 0.6 Å from the crystal state methyl carbon.

This cross-section area in the bound state is at a large distance from the conformation of Leu107 in the crystal structure of the free protein, the minimal distance between the $\delta 2$ methyl group and the cross-section being more than 5 Å, while it is only 0.6 Å for the free state. Thus, a large change in position is suggested by the observed change in the PCS. In different crystal structures of ntd-HSP90, Leu 107 is found in very different orientations, with 3.0 Å between C α atoms in the structure of the free protein (PDB entry 3t0h¹⁶) and the one with ntd-HSP90 bound to a close analogue of **1** (PDB entry: 2xdk²⁰). This observation suggests that Leu107 is very sensitive to ligand binding. Therefore, the observed change in methyl group position is plausible, although it

should be treated with caution because of the high degree of correlation between the two $\Delta\chi$ tensors involved. The data for mutant 101C/105C cannot be used in this case because the tag is too close to this Leu. In conclusion, the small but significant changes in the PCS yield consistent information about the change in methyl group position that can be in the range of 1-5 Å. The iso-surface approach yields a limited set of possible conformations, in particular in the case when data from three tag locations are available. These data could be translated into distance restraints, for example for ligand docking studies.

Conclusion and discussion

Methyl groups are widely used NMR probes but the applications of PCS to study them remain limited¹⁴⁵. A straightforward application of paramagnetic tagging is the additional dispersion induced by PCS. Paramagnetic tags have already been used to disperse resonances of an intrinsically disordered protein, presenting a crowded amide proton spectrum¹⁴⁶. We showed here that additional dispersion is also relevant for folded, larger proteins that are being studied with methyl group NMR probes, because two-dimensional spectra usually show considerable crowding for methyl resonances in the ¹H region of 0.8 - 1.0 ppm. Ytterbium was the lanthanoid of choice for a 25 kDa protein. For larger proteins, lanthanoid of different paramagnetic ‘strength’ could be used to benefit from a high dispersion, while limiting loss of information because of PRE.

Furthermore, PCS were used to provide evidence for small-scale movements (1 to 3 Å) of methyl groups upon ligand binding, by using a triangulation approach. This makes the PCS a powerful tool to observe re-arrangement of the side chains solely on the basis of 2D NMR spectra. Of course, assignments are required for the interpretation of such data. Recently, we have demonstrated that the same paramagnetic constructs used in the ligand titrations could be used to obtain partial assignments of the methyl groups²². Thus, this application of PCS could complement the use NOESY experiments, or indeed crystal structures, for structure determination of ligand-protein complexes. An advantage is that for measurement of PCS only low sample concentrations are required. It is clear that a single paramagnetic center is not sufficient to define a relevant area for the methyl group position in the bound state. With two paramagnetic centers and the use of both methyl groups of valine and/or leucine the positions can be approximated, because of the additional steric constraint that the two methyl groups must be at a distance of 2.5 Å. With three tags, the location becomes quite restrained, although two or three quite different positions may be found due to the shape of PCS iso-surfaces. Attention should be paid to the relative orientations of the susceptibility tensors used to generate the PCS. In case of tensors with parallel main axes, cross-intersections of iso-surfaces are less resolved or non-existing. When using PCS for methyl localization, it is important to realize that the PCS is highly anisotropic

and falls off with the third power of the distance between the nucleus and the paramagnetic center. Localized, high PCS gradients make the PCS very sensitive to the methyl group position. That is an advantage but also a danger. The higher the PCS gradient, the larger the effect of any PCS error will be. The use of several tag positions can counteract this problem. However, this approach also demonstrated that positions predicted from PCS from two or three tags do not always match perfectly with those observed in the crystal state of the protein, even in the absence of ligand. The reasons are not clear. It could be that the average position of these methyl groups is slightly different in the solution state but it is also possible that (one of) the tags have subtle structural effects causing a small mismatch between PCS of the native and tagged protein variants. Consequently, to study the effects of ligand binding it is recommended to use the difference in PCS between free and bound states to derive distance restraints for methyl movement, as was discussed above. To obtain reliable PCS the use of a probe that is rigid relative to the protein is important. The high rigidity of two-armed CLaNP-5 enables the measurement accurate PCS and the prediction of the tensors parameters accurately. The choice of the lanthanoid determines the optimal region. Nuclei too close to the tag will experience PRE, nuclei too far experience small changes in PCS upon displacement. With Yb^{3+} used here, distances between nucleus and tag in the range of 50 Å yielded PCS changes of 0.04 - 0.1 ppm that appear to match with displacements of 1 - 3 Å. The ability to observe and characterize small but significant changes in the methyl group positions can potentially also be applied to large proteins, because it is well-established that with deuteration methyl groups can be detected in very large systems^{147,148}. We also note that this approach is not limited to ligands that are in fast exchange. As long as assignments are available for free and bound states of the protein, PCS can also be obtained in slow-exchange systems.

Chapter 5

Flexibility in a ubiquitin dimer

Abstract

Highly dynamic multi-domain proteins remain challenging systems to study because any observables are averaged, hiding the ensemble of conformations sampled by the system. The difficulty is to extract quantitative information about the weight and the role of single conformations sampled by the system from averaged observables, and to determine a realistic ensemble. K48 linked ubiquitin dimer is an essential protein for eukaryotes. It is most active in open state(s), observed at acidic pH. At neutral pH, the closed state is more populated but remaining activity suggests an exchange between these two states. It is shown how RDC and PCS, generated by the rigid paramagnetic probe CLaNP-5 chelated with holmium, can report on dynamics and aid in the determination of an ensemble of conformations.

Introduction

Ubiquitination is the attachment of a ubiquitin (Ub) chain to a target protein. Ub is a small (8.5 kDa), highly conserved¹⁴⁹ protein and essential in many eukaryotic biochemical pathways¹⁵⁰⁻¹⁵². Ub chains can be linked via one of the 7 Ub Lys residues to the carboxyl group of the Gly 76. The different links and lengths of the Ub chains lead to different function^{153,154} possibly by different specific recognition of the ubiquitination^{155,156}. The canonical K48 and K63 linked di-Ub are known to be involved in proteasomal protein degradation^{157,158}, for example in case of protein misfolding¹⁵². The non-canonical linkages^{159,160}, such as linkage via K6, K11¹⁶¹, K29, K27¹⁶² and K33¹⁶³ are known to be involved in DNA repair processes¹⁶⁴, cell division¹⁵², regulation of mRNA stability¹⁶⁵ and immune system regulation¹⁶⁶, respectively. Linkage can be homogeneous or heterogeneous, linear and/or branched^{154,155,167}. The multiplicity of possible linkages makes the understanding of the full processes challenging¹⁶⁸. K48-linkage is involved in proteasomal degradation. It has been shown that in the dimer, multiple conformations co-exist that rapidly interconvert¹⁶¹. The flexibility of such linkages has been proven to be essential for activity and recognition of the linkage by receptors^{156,169}. K48-linked di-Ub has been studied under different external conditions, such as temperature and pH. The equilibrium between a closed, binding incompetent and open, binding competent state(s) was shown to be pH dependent, with a shift from an open state at acidic pH to a closed state for pH-values higher than 6.8¹⁷⁰. NMR studies have been performed to elucidate the role and the weight of the different states for the recognition by receptors¹⁷¹. An open structure (PDB entry 3NS8²²) of Ub2 was crystallised at pH 4.5 (PDB entry 1AAR¹⁷²) and at pH 7.5. As opposite of what was observed liquid-state with NMR, the dimer was observed in a closed conformation at the acidic pH and in an open state at neutral pH. The authors suggest that the crystal packing might favour the open state at such a low pH. On the other hand, the presence of minor liquid state is evidence that the dimer exists in the two forms at acidic and neutral pH.

Solution-state NMR is a suitable method to study highly dynamic systems. In particular, RDC¹⁷³ have been used to refine protein structures^{174,177} and study multi-domain proteins⁷⁶ to characterize the orientation of different domains relative to each other¹²¹. RDC can be measured in solution when the protein is slightly aligned, preventing the complete average of dipolar couplings and resulting in residual dipolar couplings. Multiple ways of alignment are now available^{119,178} and here, paramagnetic alignment is used¹⁷⁹⁻¹⁸¹. Paramagnetic effects have already been used to probe two-domain proteins, mostly PRE³⁹ and RDC^{97,182,183}. RDC, generated by alignment media, were used to study the closed-state of the dimer in solution and used as restraints¹⁸⁴ in the HADDOCK software^{95,96}. We propose here to study the dynamics and the balance between closed and open states, in solution, of the K48 di-Ub at two pH values. To do so, RDC and PCS were acquired for both distal and proximal di-Ub, tagged with the paramagnetic molecule Ho³⁺-CLaNP-5⁶⁵.

Material and methods

Protein preparation

Mono-Ub with the double mutation A28C/D32C, as well as K48-di-Ub, D77 variant, with the same A28C/D32C mutation either in the proximal or the distal subunit were produced and purified by Carlos A. Castañeda, from David Fushman's lab in University of Maryland, USA, according to an already published protocol¹⁸⁵. The use of D77-variant and K48D mutant, containing an additional Asp after Gly 76, allows controlled synthesis of K48 linked Ub chains to obtain the desired length of 2, dimeric Ub chain. The three samples were tagged with Lu³⁺-CLaNP-5 and Ho³⁺-CLaNP-5 by Monika Timmer, according to the published protocol⁶⁵. For the first dimer, referred as Ub2 A, the proximal subunit is the A28C/D32C/D77 variant, non-labelled. The distal subunit is the K48R mutant, ¹⁵N labelled. For the second dimer, referred as Ub2 B, the distal subunit is the A28C/D32C/K48R, non-labelled. The proximal subunit is the D77 variant, ¹⁵N-labelled, **Figure 5.1**.

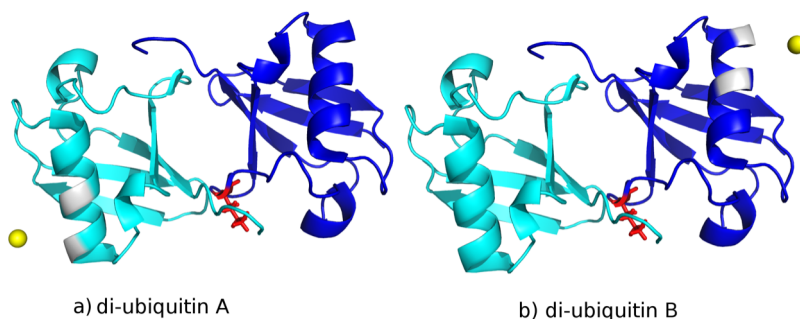


Figure 5.1. Representation of dimer of K48 di-Ub, structure 3gbf. The locations of A28C/D32C used for CLaNP-5 tagging are shown in grey. The paramagnetic center Ho³⁺-CLaNP-5 is shown as a yellow sphere. K48 is shown in red stick. In all di-Ub NMR experiments the subunit opposite of the tagged one ¹⁵N-labelled.

NMR spectroscopy

NMR samples contained 160-180 μM protein in 20 mM NaPi, pH 6.8 or pH 7.6 and, 6% D_2O . All the spectra were recorded at a temperature of 296 K. The ^{15}N -HSQC and ^{15}N -IPAP²¹ spectra were recorded on a Bruker Ascend Avance III 850 MHz spectrometer, equipped with a TCI cryoprobe, with 1024 points in the ^{15}N dimension, and zero filling up to 4096 in both dimensions. The data were processed with Topspin 3.2 and NMRpipe 8.2¹⁰⁸. The spectra were analyzed with CCPNMR Analysis 2.4¹⁰⁹. Assignment was provided by Carlos A. Castañeda and David Fushman.

Results and Discussion

Ubiquitin monomer

First, the $\Delta\chi$ tensor was determined for the monomer Ub tagged on the A28C/D32C with Lu^{3+} -CLaNP-5 and Ho^{3+} -CLaNP-5, at pH 6.8. Due to PRE, caused by the strong lanthanoid Ho^{3+} , only 19 backbone amide PCS out of 76 were observed on the monomer, **Figure 5.2**.

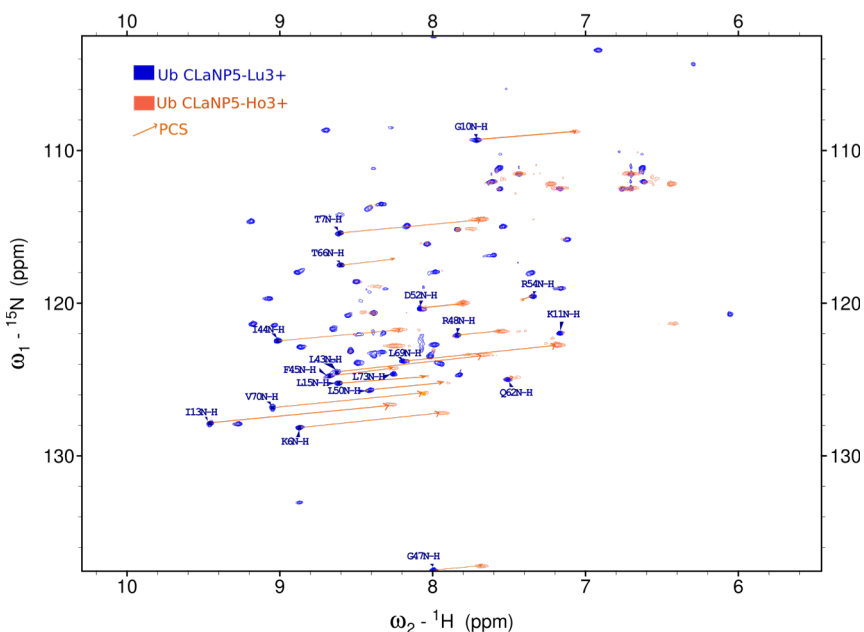


Figure 5.2. Spectra of the mono-Ub, tagged with Lu^{3+} -CLaNP-5 and Ho^{3+} -CLaNP-5 in blue and orange, respectively.

The $\Delta\chi$ tensor was determined with Numbat¹²⁷ using the 1UBQ¹⁸⁶ and 1D3Z structures and an eight parameter fit. The paramagnetic center was located too close to the surface of the protein to be realistic. Therefore, Ho^{3+} positions were estimated based on four sets of likely C β -C γ bond orientations of the engineered cysteines, and tensors were calculated with a five-parameter fit. The paramagnetic center positions relative to

the mutated α -helix are displayed in **Figure 5.3** and the relative orientation of the refined tensors for the two structures are shown in **Figure 5.4**.

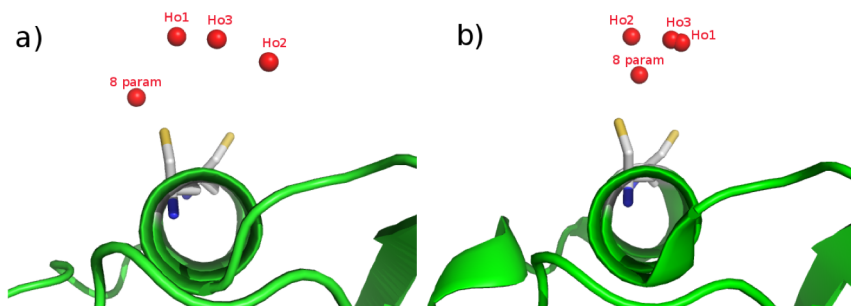


Figure 5.3. Possible paramagnetic center positions, obtained for an eight-parameter (8 param) fit to the 19 PCS and calculated from the structure for different cysteine side chain orientations (Ho1, Ho2, Ho3) for structure 1UBQ (a) and structure 1D3Z (b).

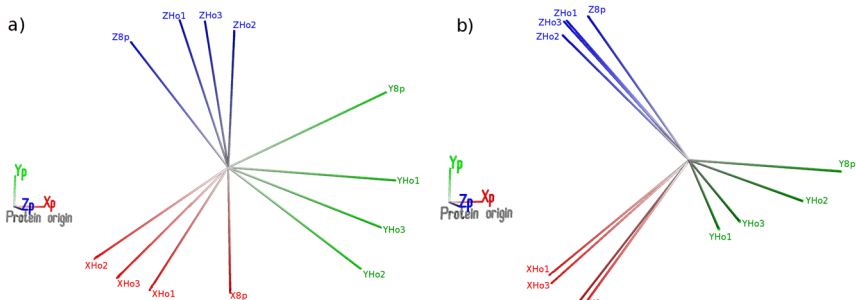


Figure 5.4. Relative orientations of the tensor frames based on five-parameter fits using the positions shown in Figure. 5.3 for 1UBQ (left) and on 1D3Z (right).

The values are listed in **Table 5.1**. The axes of the tensors when calculated with 1D3Z are less dispersed than when refined with 1UBQ, indicating that in the former the exact location has less effect on the tensor orientation. The PCS back-predicted with these $\Delta\chi$ tensors are compared to the 19 experimental PCS in **Figure 5.5**. The maximum PCS deviation for the 1D3Z-based tensors, 0.11 ppm is smaller than for the 1UBQ-based tensors, 0.18 ppm. This suggests that the structure 1D3Z is more representative of the solution state that we are observing with NMR. As expected, the eight-parameter fit gives better agreement with the experimental data than the five-parameter fit, but brings the CLaNP tag too close to the protein surface. For the rest of this study, the 1D3Z-based tensor with position Ho1 is used.

Table 5.1. $\Delta\chi$ tensor parameters, calculated with structures 1UBQ and 1D3Z with eight free parameters, or five parameters with fixed position in Ho1, Ho2 or Ho3.

Model	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	x	y	z	α	β	γ
1UBQ Ho1	-22.9	-7.8	44.871	22.734	9.631	123	69	117
1UBQ Ho1	-22.9	-7.8	44.871	22.734	9.631	124	69	117
1UBQ Ho2	-19.3	-7.4	41.130	23.092	6.700	115	52	100
1UBQ Ho3	-22.2	-7.4	43.445	22.612	8.171	119	62	107
1UBQ 8 param	-15.7	-7.6	44.715	25.369	12.350	138	79	138
1D3Z Ho1	-19.9	-3.7	58.222	-95.416	-6.574	135	97	65
1D3Z Ho2	-20.9	-7.1	59.144	-94.799	-8.948	142	87	85
1D3Z Ho3	-20.2	-4.1	58.462	-95.348	-7.000	136	95	70
1D3Z 8 param	-15.3	-6.0	57.090	-93.912	-8.896	136	82	101

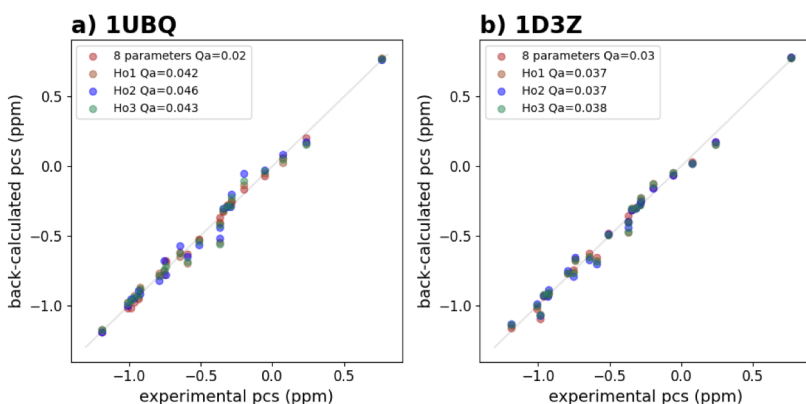


Figure 5.5. Back-calculated versus experimental PCS measured for the mono-Ub at pH 6.8 for 1UBQ (left) and 1D3Z (right). The best five-parameter fit is the one for 1D3Z and position Ho1. Q_a was calculated according to [equation 4.2](#)

Di-ubiquitin

IPAP-HSQC spectra of both Ub2 A and Ub2 B at pH 6.8 and 7.6. were acquired. Secondary peaks overlapping in both diamagnetic and paramagnetic spectra were observed for several residues, [Figure 5.6](#). These minor peaks were present in both dia- and para-magnetic sample spectra but did not experience pcs, suggesting that they stem from untagged protein. These peaks can affect the measurement of RDC by shifting the peak position center, in case the resolution does not permit to distinguish them. Therefore, the spectra were acquired with high resolution and processed with ample zero-filling to obtain optimal resolution and peak shape in the ^{15}N dimension. RDC were measured as the difference of the ^{15}N couplings measured in the spectra of paramagnetic and diamagnetic samples. The RDC profiles were different for the two di-Ub systems, [Figure 5.7](#) panel A and B, and the two pH values, panels C and D, suggesting different dynamics, difficult to quantify at this stage. PCS and RDC were predicted for the opposite subunit for both Ub2 A and Ub2 B, using structure 2BGF-model 1 and the tensors calculated with 1D3Z-Ho1. The alpha helix of structure

2BGF on which CLaNP-5 is located was aligned with the same alpha helix of 1D3Z. The RMS was 0.35 Å, good enough to assume that the tensor will be the same for the monomer and the dimer.

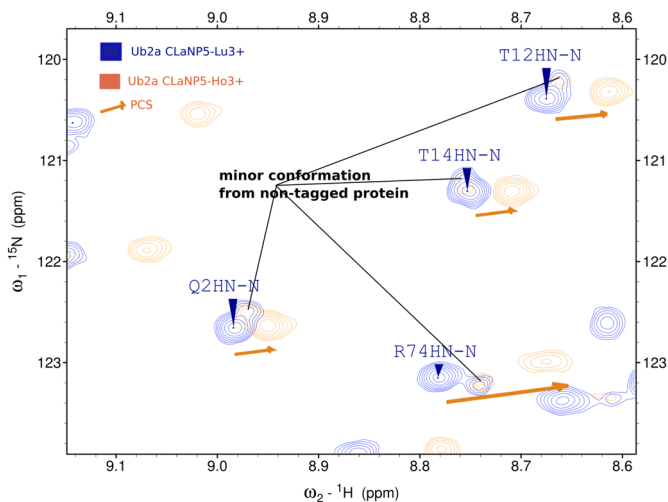


Figure 5.6 Overlay of the up-field area of ^{15}N HSQC spectra of $\text{Lu}^{3+}\text{CLaNP-5}$ (blue) and $\text{Ho}^{3+}\text{CLaNP-5}$ (orange) tagged Ub2a.

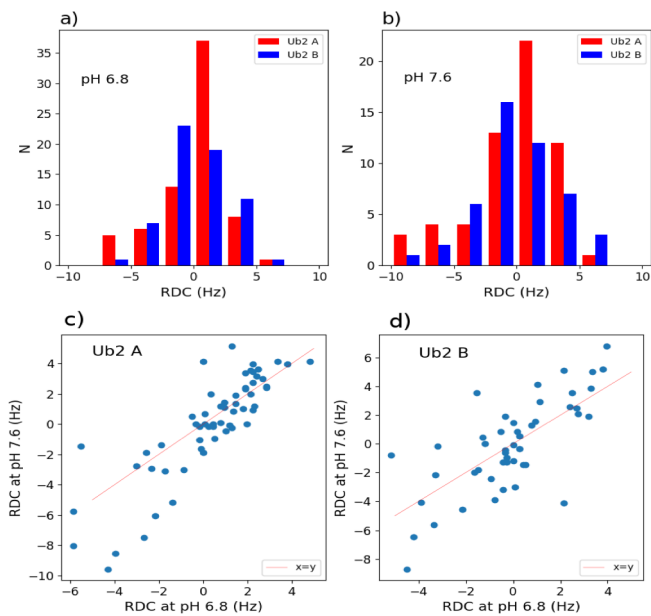


Figure 5.7. Histogram of RDC measured on CLaNP5- Ho^{3+} tagged di-Ub A and B in red and blue, respectively at pH=6.8 (a) and pH=7.6 (b). RDC measured at pH 6.8 versus RDC measured at pH 7.6 for Ub2 A (c) and Ub2 B (d). The red line represents a perfect correlation.

The results for the data are shown **Figure 5.8** and a detailed plot is shown for Ub2 A in **Figure 5.9**.

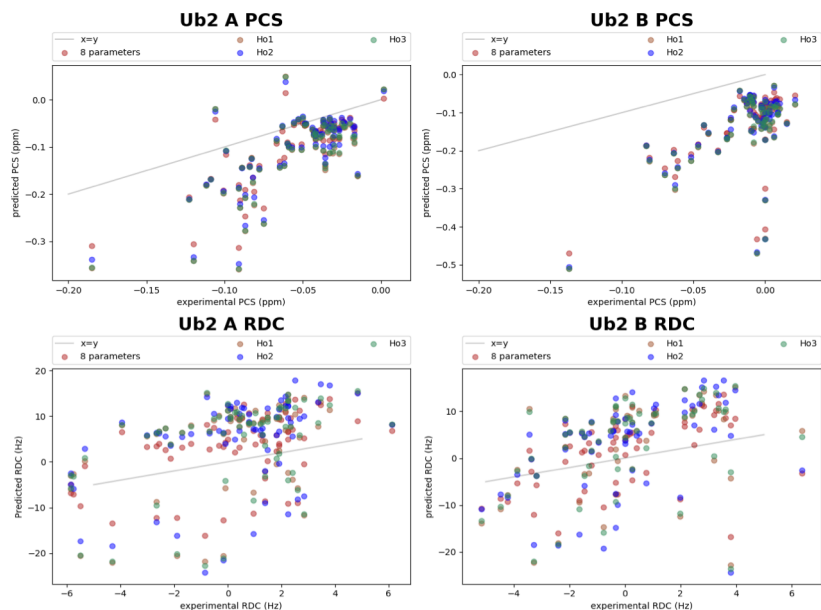


Figure 5.8. Experimental versus PCS (top) and RDC (bottom) for the different tensors calculated with the Ub monomer 1D3Z for Ub2 A (left) and Ub2 B (right). The grey lines represent a perfect correlation.

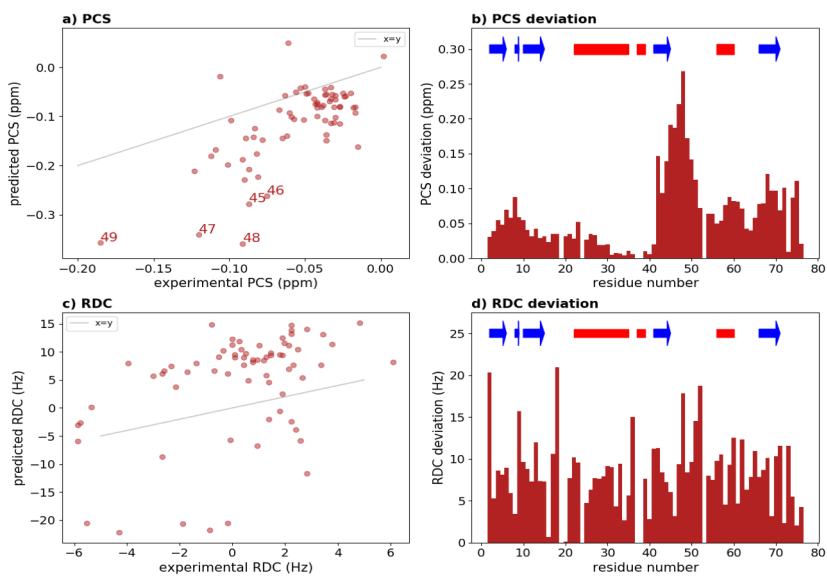


Figure 5.9. Detail of the correlation between the predicted and experimental PCS (a, b) and RDC (c, d) for Ub2 A. Predicted values are for the 1D3Z with Ho1 position refined tensor. Absolute deviations are plotted on the right-hand side panels for PCS and RDC. The blue arrows and red bars represent β -strands and α -helices, respectively.

For the PCS plot, the values of five residues stood out by large deviations. These five residues, 45-49, are located around the inter-subunit peptide bond, a highly flexible part of the dimer. The average absolute difference between the experimental PCS and the predicted ones for the closed state is 0.08 ppm for both di-Ub, indicating a poor match. However, the data show a correlation ($R = 0.5$), reflecting a reduction of PCS of about two-fold for Ub2 A and three-fold for Ub2 B. PCS depends on both the distance between the nucleus and the paramagnetic center and the orientation of the nucleus in the tensor frame. In a dynamic system of two linked subunits, the distance will be less variable than the orientation, leading to an overall reduction of the PCS with roughly similar size for each nucleus, explaining the observed correlation. Thus, the overall PCS reduction is evidence for dynamics occurring in the dimer. This is corroborated by the observed RDC, which are about a factor of three smaller than predicted for a rigid dimer. It can be concluded that the closed structure 2BGF model 1 does not represent the solution state. Using the same $\Delta\chi$ tensor, PCS and RDC were predicted for the open structure 3NS8 and compared to the experimental PCS and RDC obtained at pH 6.8, **Figure 5.10**.

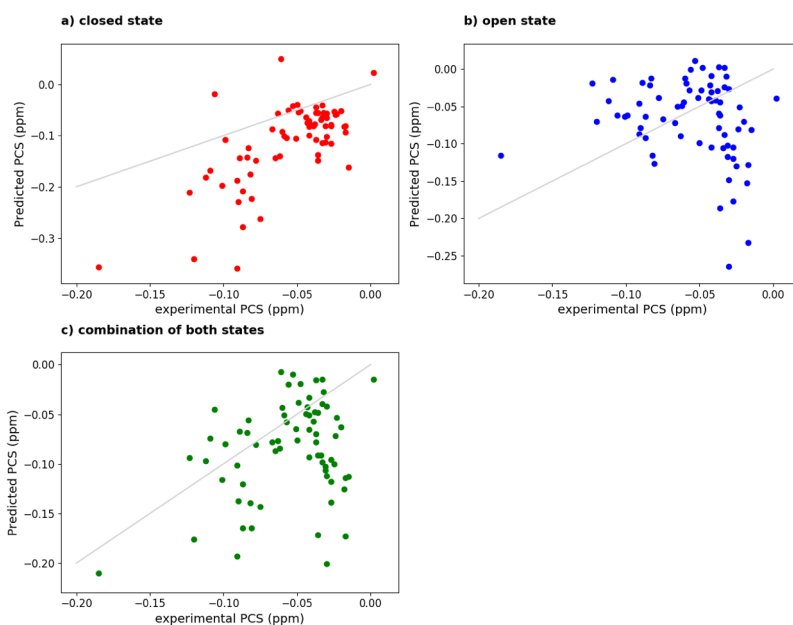


Figure 5.10. PCS fits. Predicted PCS are plotted against the experimental PCS for Ub2 A, using for prediction the closed state (2BGF) (a), the open state (3NS8) (b) and a linear combination of closed (61%) and open (39%) states.

The PCS correlation dropped to $R = 0.04$ for both di-Ub. The correlation for the RDC is even lower. The average absolute deviation is still high, 0.07 ppm, simply indicating that the PCS ranges for the experimental and predicted values are of the same magnitude. If conformations of the dimer would be in fast exchange, the experimental

PCS would be a linear combination of PCS experienced by the nuclei in the different conformations visited by the di-Ub. To see if a combination of the closed, 2BGF, and open state, 3NS8 represented better the solution state at pH 6.8, a least-square function was used in a home-written python script to determine the fraction a of the open state ($1-a$ being the fraction of the closed state). The best fit was found for $a = 0.61$. As expected because of the additional fit parameter a , the fit was somewhat better, from a standard deviation of 0.08 ppm and 0.07 ppm for the closed and open states, respectively, to a standard deviation of 0.06 ppm. However, the fit is still rather poor. Most likely, the subunits are dynamic relative to one other, assuming an ensemble of orientations of which the closed and open states may or may not be elements.

Conclusion

K48-linked Ub2 were successfully tagged. The tagging efficiency was deemed high, and was qualitatively assessed by the weak intensities of what was assumed to be untagged protein resonances. Ho^{3+} was an interesting lanthanoid as it was strong enough to measure RDC up in a range of ~ 15 Hz for the dimer but also weak enough to observe PCS on the monomer allowing the refinement of the experimental susceptibility tensor. However, the susceptibility tensor accuracy would benefit from an extra PCS dataset, generated with the same paramagnetic probe CLaNP-5 but chelated with a weaker lanthanoid such as Yb^{3+} . The orientation would remain the same and the magnitude would simply be scaled roughly by a factor 2.5. Berlin et al. developed the Sparse Ensemble Selection (SES) for determining multiple conformational states from limited number of observation. It is a model-free method that was applied to probe K48-linked Ub2 dynamics. They showed that an ensemble of already 3 structures was determined with a correlation of 1 between the experimental RDC and back-calculated ones. The published results reported on a strong correlation between the ensemble properties and the k48 Ub2 recognition by receptors. A next step in this project would be to test this SES method for our set of RDC, generated by a paramagnetic alignment. To go further the current data could be the basis for ensemble calculations, for example using the MaxOcc method⁶⁰, which previously has been shown to be powerful in describing domain dynamics¹⁸⁷ and even encounter state of two proteins using RDC and PCS as input¹⁸⁸. The acquisition of an additional independent set, using the same double mutated cysteine site but with a different tag such as CLaNP-7¹⁸⁹, would be very beneficial to polish the ensemble conformations.

Chapter 6

Multimeric paramagnetic tagging:
Proteasome

Abstract

The combination of methyl labelling in a deuterated background with the use of methyl-HMQC-TROSY experiments has paved the way to study very large protein complexes with liquid-NMR state. Paramagnetic effects could be very useful to probe structure and interaction of such complexes. We investigated the potential of a rigid ytterbium paramagnetic tag to generate PCS for a highly symmetric $\alpha 7\alpha 7$ proteasome-derived complex of fourteen identical subunits. It is shown that PCS can be detected. However, it is difficult to reach 100% tagging of the molecules, which results in a population of molecules with multiple tags at various positions. Consequently, the spectra become very crowded and impossible to interpret. As an alternative, minimizing the number of tags per complex was investigated. This approach yields PCS that are in line with the prediction but suffers from low sensitivity and high background of peaks from diamagnetic species.

Introduction

Proteasomes are enzymatic complexes found in eukaryotes, archaea and in some bacteria¹⁹⁰. They are composed of a 20S catalytic core complex that can bind regulator complexes such as, for example, 19S to form the 26S proteasome. The 20S core is composed of four stacked rings of seven subunits each, two inner rings (β) and two outer rings (α). Proteasomes targets ubiquitin labelled proteins, and, when associated with the 11S regulatory subunit, also non-ubiquitinated proteins^{191,192}, for degradation carried out by the two β -rings, containing the protease enzymes subunits. In eukaryotic proteasomes, the α -rings are each composed of seven different alpha subunits, and the inner rings are each composed of seven different beta subunits, whereas the proteasome in archaea like *Thermoplasma acidophilum* has only one type of α and one type of β subunit^{190,193}. Nowadays, it is possible to study dynamics and interactions of such a large complex with liquid-state NMR. TROSY (transverse relaxation optimized spectroscopy)^{24,194} experiments applied to methyl groups²⁵ has paved the way to binding to study protein complexes up to at least 670 kDa¹³⁷ because of the favourable linewidths of methyl resonances in a perdeuterated environment. For such large systems, not only the linewidth is a problem but also the sheer number of resonances. Even when using selective labelling of methyl groups, the spectra would be very crowded for a monomeric protein complex of that size. The two α rings of the proteasome of *Th. acidophilum* can form an artificial complex, the $\alpha 7\alpha 7$ complex of 350 kDa, which is a 14-mer of identical subunits. This $\alpha 7\alpha 7$ assembly was crystalized in 2015, PDB entry 5dsv¹⁹⁵. This complex thus has a strongly reduced spectra complexity, similar to that of a single α subunit (21 kDa), making it attractive for solution state NMR²⁸.

PCS would be beneficial for the study of large complexes, because they can be measured readily with sensitive NMR spectra that do not require high concentrations of sample and PCS can provide useful information for assignments and about structure and interactions. The goal of the work described in this chapter was to investigate whether PCS generated with the paramagnetic Yb³⁺-CLaNP-5^{63,196} can be observed for the 14-meric $\alpha 7\alpha 7$ complex of 350 kDa.

Materials and methods

Five double cysteine mutants of the $\alpha 7\alpha 7$ complex were designed in such a way that the engineered Cys residues are accessible on the surface, **Figure 6.1**. The mutants were K53C/R55C, T216C/N219C, T191C/K195C, A171C/S174C, and E29C/K32C. The double cysteine mutations were designed with Pymol and PCS were predicted for Yb³⁺-CLaNP-5 using a Pymol wizard written by Dr Stephan Theisgen. The mutant proteins were produced and purified by Rui Huang, in Professor Lewis Kay's laboratory in Toronto, Canada. According to a published protocol²⁸.The protein

samples were ILVM-methyl-labelled, the labelling was U-[^2H] Ile-Leu, [δ , ^{13}C], Leu,Val-[^{13}C], [^{13}C], Met-[^{13}C]. The NMR samples contained 1.2, 0.5, 0.9, 0.2, 0.75 mM of the 14-mer of K53C/R55C, T216C/N219C, T191C/K195C, A171C/S174C, and E29C/K32C mutants, respectively, in 25 mM KPi, pH 7.4, 50 mM NaCl in D₂O. No experimental paramagnetic data were recorded for mutant A171C/S174C, only the spectrum with no tag was acquired. The samples were tagged in Leiden with Lu³⁺-CLaNP-5 and Yb³⁺-CLaNP-5 according to protocol already published⁶⁵. Different excess of CLaNP-5 were used and are discussed in the Results part.

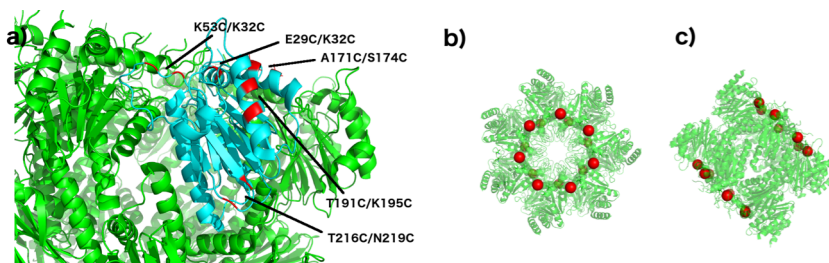


Figure 6.1 a) designed mutated cysteine in red cartoon b-c) $\alpha 7\alpha 7$ 14-mer top and side view, respectively (PDB entry 5dsv). The 14 Yb³⁺-CLaNP-5 paramagnetic centers of mutant E29C/K32C are represented in red.

Spectra were acquired on a Bruker Ascend Avance III 850 MHz spectrometer, equipped with a TCI cryoprobe, using 2D CH₃ TROSY SOFAST-HMQC¹⁹⁷ at 296 K. All spectra were acquired with 230 complex in the 13C dimension for a spectral width of 19.1 ppm, with 64 scans. The data were processed with Topspin 3.2 and NMRpipe 8.2⁹⁸. The spectra were analyzed with CCPNMR Analysis 2.4¹⁹⁹.

Results and discussion

Half of the ~1.2 mM K53C/R55C mutant sample was tagged with five molar equivalents of Yb³⁺-CLaNP-5. About 35% of the protein precipitated during the tagging reaction. To avoid more protein loss, a diamagnetic control spectrum was first acquired on the untagged protein. Final concentrations of the two samples were 0.8 mM and 0.3 mM, for the diamagnetic control and the paramagnetic sample, respectively. The spectrum of the paramagnetic sample showed PCS, as well as apparent line broadening, **Figure 6.2**, panel A.

To establish whether this paramagnetic broadening was caused by the high field, responsible for more important Curie relaxation²⁰⁰, the same paramagnetic sample was measured at a lower field, 600 MHz instead of 850 MHz, using the same SOFAST-HMQC methyl TROSY experiment¹⁹⁷. The weaker alignment due to the lower magnetic field did not reduce the broadening effect. The broadening could also be explained by a partial tagging of the monomers. If the tagging occurs at all sites, a nucleus experiences fourteen different PCS contributions, originating from all subunits

of the $\alpha 7\alpha 7$. However, partial tagging generates a mixture of non-identically tagged $\alpha 7\alpha 7$ molecules, resulting in different PCS contributions, and therefore many or at least several observable PCS. The PCS falls off with the third power of the distance between the tag and the nucleus, so the main PCS contribution arises from the nearest paramagnetic center, and others paramagnetic centers are responsible for a smaller positive or negative PCS contribution, causing a range of slightly different PCS. Yb is a modestly strong lanthanoid, so that each nucleus is expected to feel no more than two or at most three probes and not all fourteen. For mutant T191C/K195C, multiple PCS could be identified for some peaks, **Figure 6.3**.

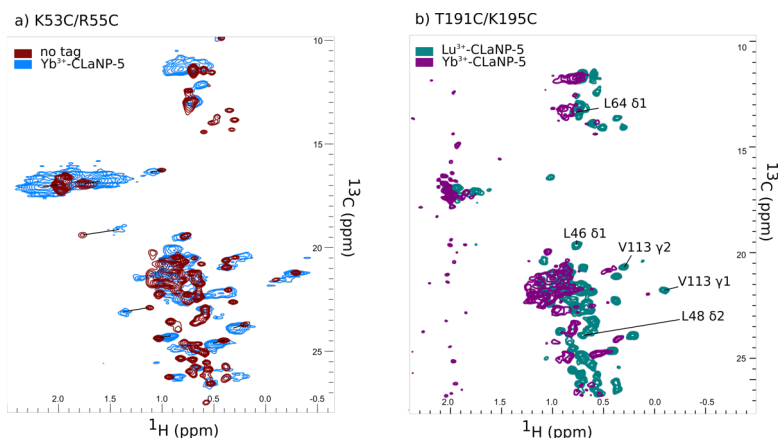


Figure 6.2. NMR spectra of a) the two K53C/R55C mutant spectra untagged, in blue, and tagged with Yb³⁺-CLaNP-5 b) the two T191C/K195C tagged with Lu³⁺-CLaNP-5, in teal and with Yb³⁺-CLaNP-5 in purple.

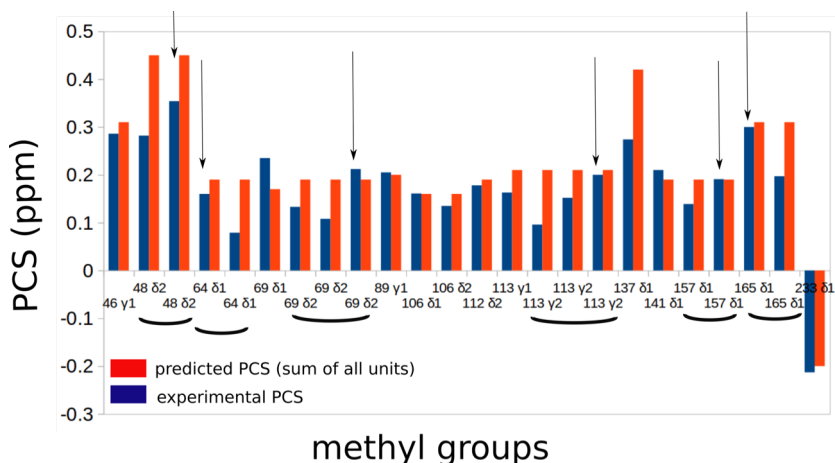


Figure 6.3 Experimental PCS for mutant T191CK195C. The experimental PCS are in blue and the PCS predicted on subunit A from Yb³⁺-CLaNP-5 tagged on the 14 subunits are in red. The brackets highlight methyl groups for which more than one PCS are observed experimentally, the arrows indicate on the best experimental PCS that are the closest to the predicted PCS.

For L48_{δ2}, I64_{δ1}, I157_{δ1}, I165_{γ1} two PCS were identified and three PCS were measured for L69_{δ2}, V113_{γ2}. To circumvent the effects of partial tagging, which yields complex spectra, a low-ratio tagging strategy was performed for mutants T216C/N219C and E29C/K32C. The idea is to decrease the targeted tagging ratio to 1:14, to have only one subunit tagged. It is expected to observe then the diamagnetic spectrum resulting from the 13 non-paramagnetic subunits and the PCS resulting from paramagnetic tagged subunit. Only the contribution from one subunit would be observed, with an obvious drawback in reduced sensitivity of the paramagnetic peaks, which would be reduced 14-fold. On the other hand, less precipitation is expected from the tagging process, so less material is expected to be lost. The tagging process follows a Bernoulli trial f , **equation 6.1**, with two outputs of the event “tagging” noted k . p is the probability of the successful event “the subunit is tagged”, and $q = 1-p$ is the probability of the unsuccessfully event, “the subunit is not tagged”.

$$f(k, p) = p \quad \text{if the subunit is tagged (event } k=1)$$

$$f(k, p) = 1-p \quad \text{if the subunit is not tagged (event } k=0)$$

The Bernoulli trial is repeated fourteen times for all the fourteen subunits subject to tagging. The probability of tagged subunits can be calculated according to **equation 6.1**.

$$p(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (6.1)$$

where k is the number of tagged subunits, p is the tagging probability (for instance =1/14 if the tagging ratio is 1:14), and $n=14$ is the number of repetitions. The probability of tagged subunits was calculated for four tagging ratios, 2/14, 1/14, 0.7/14, 0.5/14, **Figure 6.4**. The highest yield of $\alpha7\alpha7$ molecules with a single tag is found for a CLaNP-5: subunit ratio equal to 1. However, this ratio leads to 20% molecules with two tags. To minimize these doubly tagged molecules without reducing the singly tag fraction, a ratio of 0.7:14 was chosen, for the next tagging strategy.

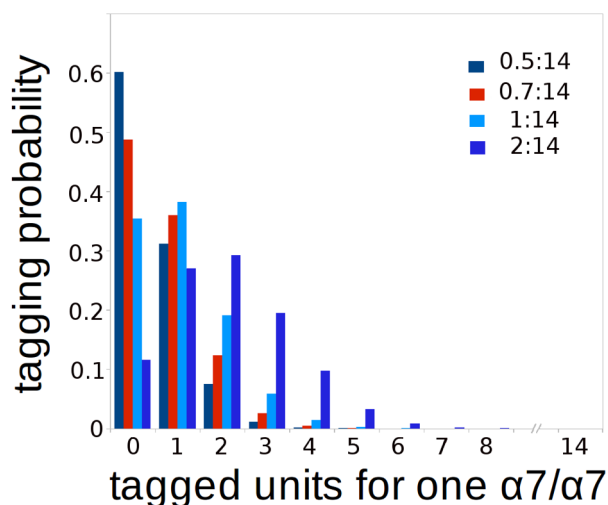


Figure 6.4 Tagging probability calculated for Bernoulli trials on the fourteen subunits of $\alpha 7/\alpha 7$, for different CLaNP-5:subunit ratio, 0.5:14, 0.7:14, 1:14 and 2:14.

CLaNP-5 tagging efficiency is not 100% and has been used before with 5-fold excess to achieve near to 100% tagging⁶⁵. Therefore, the protein sample was tagged with $0.7/14 * 5 = 0.25$ molar equivalents CLaNP-5 per subunit. E29C/K32C and T216C/N219C were tagged, using this 0.25:1 CLaNP: subunit ratio, both with Lu^{3+} -CLaNP-5 and Yb^{3+} -CLaNP-5. As could be expected, PCS on the edge of the crowded spectra could be easily be identified. The remaining strong diamagnetic peaks mask the other paramagnetic peaks. For the mutant E29C/K32C, at first glance the agreement between the experimental PCS and the predicted ones seems better when the predicted PCS are the sum of the contribution coming from all the proteasome subunits, **Figure 6.5**.

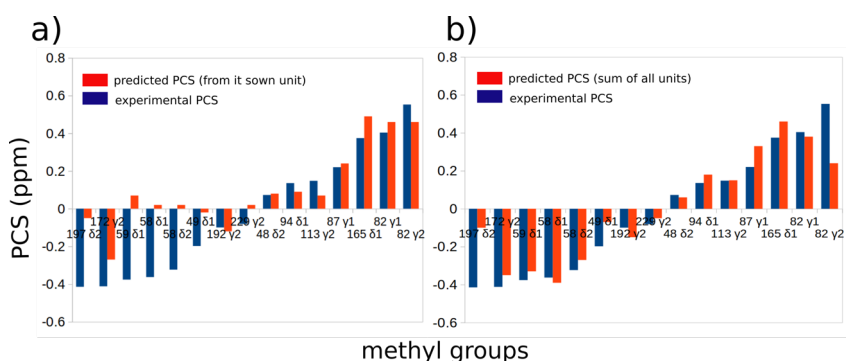


Figure 6.5 Experimental PCS for mutant E29C/K32C tagged with Yb^{3+} CLaNP-5. a) The experimental PCS are in blue and the PCS predicted on subunit A from Yb^{3+} CLaNP-5 tagged on subunit A are in red. b) The experimental PCS are in blue and the PCS predicted on subunit A from Yb^{3+} CLaNP-5 tagged on the 14 subunits are in red, this configuration is expected when the tagging ratio is high and tend to 100%.

This is unexpected as with the low-tagging ratio only intrasubunit PCS should be observed. However, the methyl groups for which the discrepancy is higher are the methyl groups L58_{δ1} and I59_{δ1}. For those ones, the predicted PCS coming from their own subunit is smaller than the PCS generated by the neighbouring subunit, **Table 6.1**.

Table 6.1. Predicted PCS contribution of the seven same-alpha-ring subunits on of L58 and I59 for mutant E29C/K32C.

Methyl group	PCS engineered by subunit							Sum of all PCS
	A	B	C	D	E	F	G	
L58 _{δ1}	0.02	-0.36	-0.04	-0.01	0	0	0	-0.39
I59 _{δ1}	0.08	-0.25	-0.03	-0.01	0	0	0.01	-0.2

The sum of all the pcs is close to the major contribution not stemmed from subunit A but from subunit B. This explains the bad agreement observed previously, **Figure 6.5**, between the experimental PCS and the predicted PCS from subunit A. On the spectrum, this is clearly illustrated for methyl group I59_{δ1} for which the two PCS from subunit A (predicted to be 0.08 ppm) and subunit B (predicted to be -0.25 ppm) are observed, **Figure 6.6**. Mutant T216C/N219C shows the same discrepancy between experimental PCS and predicted PCS, **Figure 6.7**. Similarly, the methyl groups for which the agreement is not good with the PCS from their own subunits are mainly stemmed from the neighbouring subunits, **Table 6.2**.

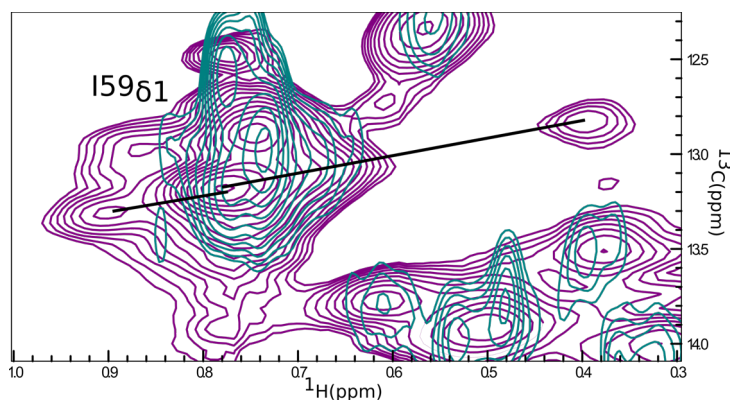


Figure 6.6. Detail of the E29C/K32C mutant spectra showing the methyl group I59_{δ1}. The spectra of the diamagnetic and paramagnetic samples are in teal and purple, respectively. Two PCS, indicated by the lines, are observed for the methyl group I59_{δ1}. The CLaNP5:subunit ratio during the tagging reaction was 0.25.

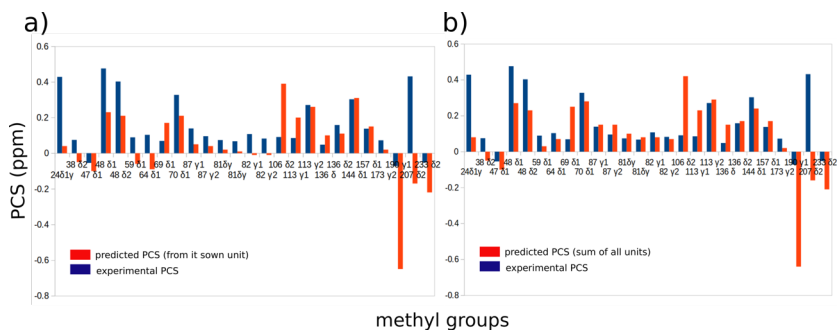


Figure 6.7 Experimental PCS for mutant T216C/N219C tagged with Yb³⁺CLaNP-5. a) The experimental PCS are in blue and the PCS predicted on subunit A from Yb³⁺CLaNP-5 tagged on subunit A are in red. b) The experimental PCS are in blue and the PCS predicted on subunit A from Yb³⁺CLaNP-5 tagged on the 14 subunits are in red, this configuration is expected when the tagging ratio is high and tend to 100%.

Table 6.2. Predicted PCS contribution of the seven same-alpha-ring subunits on of L58 and L59 for mutant T216C/N219C.

Methyl group	PCS engineered by subunit:							Sum of all PCS
	A	B	C	D	E	F	G	
L47 _{δ1}	-0.1	0.02	0.01	0	0	0	0.03	-0.1
I64 _{δ1}	-0.09	0.14	0.02	0.01	0	0	-0.01	-0.07
I59 _{δ1}	-0.06	0.07	0.02	0.01	0	0	-0.01	0.03
L38 _{δ2}	-0.08	0.02	0.01	0	0	0	-0.03	-0.08
V82 _{δ1}	-0.01	0.07	0.02	0.01	0	0	-0.03	0.08
V82 _{δ2}	-0.01	0.06	0.02	0.01	0	0	-0.01	0.07
L81 _{δ1}	0.02	0.04	0.02	0.01	0.01	0	0	0.1
L81 _{δ2}	0.01	0.04	0.02	0.01	0	0	0	0.08

Conclusion

PCS were observed for 4 double-cystein mutants of the $\alpha 7\alpha 7$ annulus. Despite the bad tagging efficiency, it was possible to identify PCS, which matched well with the predicted PCS. However, the multiplicity of peaks, plus the remaining diamagnetic untagged subunits spectra hampered the possibility to use of those PCS for PARAssign assignment purposes. Different strategies could be followed to improve the paramagnetic complex spectra. The tagging efficiency could be improved by a 10-fold CLaNP excess instead of 5-fold, but the risk is to engender more precipitation during the tagging process. On other option would be to take benefit from the 3+ charge of CLaNP-5 with an affinity column to remove the uncharged, and therefore untagged, proteasomes. At the opposite, without charge repulsion, the tagging yield might be improved, thus tagging with the neutral CLaNP-7 could be tested.

To reach 100% CLaNP-5 attachment, the tagging could be done on the monomers. After incubation of the monomeric protein with the tag, a size-exclusion column and an affinity column could be used to remove possible dimers and untagged monomers, respectively. The collected tagged fractions would be assembled to form $\alpha 7$ rings. Last but not least, different paramagnetic tagging strategies, such as lanthanoid-binding peptide tags (LBT) would eliminate tagging issues. If the affinity of the peptide for the lanthanoid is tight, all predicted sites would host a lanthanoid.

Chapter 7

General Discussion and Outlook

Since NMR was discovered in 1946^{201,202}, NMR has continuously developed to become an interdisciplinary technique used in chemistry²⁰³⁻²⁰⁵, bio-chemistry²⁰⁶⁻²⁰⁸, pharmaceuticals^{209,210}, biology and medicine with the MRI. The use of NMR in biochemistry and particularly for protein studies became possible thanks to the constant technical progress, namely higher magnetic fields to gain in sensitivity and resolution and cryo-probes that yield 4-6 fold better signal-to-noise ratios^{211,212}. From a theoretical point of view, optimized NMR²¹ pulse sequence were developed to increase sensitivity²¹³ as well as resolution²¹⁴. From the wet lab point of view, new sample preparation protocols, such as perdeuteration^{215,216} and other isotopic labeling¹¹⁷, ensured the possibility to study bigger and bigger proteins, by preventing undesired relaxation phenomena and incrementing the number of dimensions for less crowding. Development of non-linear sampling coupled with appropriate processing^{217,218} now allows the acquisition of spectra up to 5D²¹⁹ in a limited amount of time²²⁰.

Paramagnetic NMR was part of the progress to develop liquid-state NMR further to tackle even more challenging proteins, such as large protein complexes and dynamic proteins. Originally observed in metalloproteins^{221,222}, paramagnetic effects were initially treated as drawbacks, mainly because paramagnetic relaxation enhancement (PRE) hampered the quality of NMR spectra. In 1994, the first liquid-state structure of a paramagnetic protein was refined by Banci et al. using NOE²²³. Since then, paramagnetic effects were used as structural restraints for protein structure determination^{89,224,225}, as well as dynamics^{226,227} and assignment¹¹⁻¹³.

PARAssign applications

Several methods have been developed to assign protein spectra using different strategies and research into procedures to expand the assignment toolbox still continues to offer the NMR-user a wide choice of techniques to apply to the system of interest^{228,229}. In this thesis, the possibility to assign methyl spectra using a crystal structure and PCS measured from sensitive 2D HSQC spectra was demonstrated. Assignment of spectra from simple and sensitive HSQC-like spectra is of major interest for proteins suffering from low solubility or stability. It was found that the use of paramagnetic centers generating parallel susceptibility tensors must be avoided for PARAssign assignment. This alignment configuration led to redundant PCS datasets, which do not provide extra information, **chapter 2**. Consequently, the use of the same tag chelated with a different lanthanoid is predicted not to be useful for PARAssign. A simple change of the lanthanoid in the same tag would not be helpful because the $\Delta\chi$ tensor orientation is expected to be the same and the magnitude only scaled up or down, depending on the strength of the lanthanoid. However, the same mutated site could be used to work with a different tag, generating a $\Delta\chi$ tensor with different orientations. For

example, it has been shown that CLaNP-5 and CLaNP-7 give sufficiently different $\Delta\chi$ tensor orientations to yield independent data sets¹⁸⁹.

For full assignment, the completeness of datasets remains an issue, as discussed in **chapter 2**, affecting the number of highly reliable assignments. However, despite a reduction of the fraction of assigned nuclei, those classified as highly reliable are all correct, which accounts for the robustness of the approach. In case of incomplete datasets, combining PCS with other data from other sources to identify missing assignments would be advantageous. PARAssign provides small sets of reliable suggestions for assignments, which could readily be checked by mutagenesis.

PARAssign could be used with PCS measured from TROSY spectra²³⁰. However, it is important to realize that an amide peak in the TROSY spectrum is shifted by $(J_{\text{HN}}+\text{RDC})/2$ compared to the HSQC resonance in the spectrum of paramagnetic proteins because of the partial alignment caused by the lanthanoid. In the spectrum of the diamagnetic protein, the peak is shifted by only $(J_{\text{HN}})/2$. As a consequence, the difference of the proton chemical shift between paramagnetic and diamagnetic protein spectra is $(\text{PCS}+\text{RDC})/2$. In order to measure PCS accurately, a measure of the RDC with an IPAP experiment would be necessary. However, a RDC of 10 Hz would lead to an effect on the PCS of 0.012 ppm on the proton dimension at a field of 20T (850 MHz). These effects could be neglected without affecting the quality of the assignment. In methyl-TROSY, the selected long-lived signal derives from a component that is not split by ^1H - ^{13}C coupling. Therefore, it is possible to measure “RDC-free” PCS using methyl-TROSY²³¹. Note that for heteronuclei the relative contribution of the RDC to the observed shift is much larger, because the PCS expressed in Hz scales with the gyromagnetic ratio, whereas the RDC does not. Thus, PCS derived from hydrogens are more accurate than from heteronuclei.

An interesting extension of PARAssign would be the ability to assign nuclei in multimeric proteins, as discussed in **chapter 6** for the proteasome-derived $\alpha 7\alpha 7$ system. In many biological pathways²³², proteins associate to form protein complexes^{233,234}. Homo-oligomerization²³⁵ in particular concerns a large fraction of all proteins. Given that 100% tagging is difficult to achieve, especially in multimeric proteins, due to charge effects of the probe, analytical solutions to deal with multiple tagged protein species are desirable. Such analysis would need to take into account the possible species and the resulting PCS values. A nucleus in a dimer can exhibit four peaks, of untagged protein, or protein tagged at the near site, tagged at the far site or tagged at both sites, yielding a PCS that is the sum of the PCS caused by the tags at the near and the far sites. With higher order multimers, the number of possibilities grows rapidly, although many peaks will have very weak intensities.

PARAssign could also be enhanced to fit better experimental labeling strategies nowadays available. Methyl groups can be labelled stereo-specifically^{34,236}. In the program, currently, all observable stereoisomers must be specified for each type

residue, in the PARAssign configuration file. In the future, it would be better to indicate this property for each methyl group. In that way, two sets of PCS measured on two stereo-specifically labelled samples could be used together as an input with already the knowledge of the stereospecificities.

A derived application of PARAssign could ensue from its ability to determine $\Delta\chi$ tensors, using one or more paramagnetic centers, without known assignments. PARAssign's ability to determine the tensors using a PCS dataset generated from only one paramagnetic center was shown in **chapter 2**. Using methyl group PCS, $\Delta\chi$ tensors were refined to within 5 degrees error for beta angle, reporting on the main axis orientation, and between 15-25 degrees for alpha and gamma, reporting on the secondary axes orientations. Even with an incomplete set of PCS, 44 out of 76 PCS, the tensor could be refined. The possibility to determine the $\Delta\chi$ tensor of a single paramagnetic center, with a known crystal structure, without having to carry out the protein assignment first, would be a non-negligible gain of time. It could be of special interest in pharmaceutical field to quickly determine $\Delta\chi$ tensors with a reduced set of PCS, without prior knowledge of assignment. For determining ligand binding sites and orientations with transferred PCS, knowledge of the tensor parameters yields more accurate results²³⁷. This would enable the study of ligand binding with proton ligand NMR on large systems for which no backbone NMR spectra can be acquired, using methyl group labeling.

Structural restraints and dynamics observed with paramagnetic effects

PCS have been used already for structure calculation using proton amide PCS⁸⁹ but also proton methyl groups PCS²³⁸. They can also report on structural changes caused by ligand binding to a protein tagged with a paramagnetic center. Incremental changes in PCS observed during a titration with ligand report on the re-orientation of methyl group only or a more significant rearrangement of the backbone. Attention should be paid to possible tensor orientation changes upon ligand binding. Such an effect can be detected by tensor determination for the bound state and a comparison with the tensor found for the ligand-free state. In general, a global change of PCS all over the protein reports on a change in the $\Delta\chi$ tensor orientation. A local perturbation of the PCS reflects local structural changes, **chapter 4**. The mobility amplitude of the observed nucleus depends obviously on the observed PCS but also largely on the nucleus PCS gradient. The PCS gradient can be calculated to appreciate its influence on displacements for a given Δ PCS. The data in **chapter 4** illustrate the point that small changes in PCS can reflect large structural changes, in the case of a weak gradient and, analogously, large PCS changes may be caused by only tiny structural changes when the gradient is strong. The calculation of PCS gradients provides additional insight in the meaning of a PCS deviation in terms of structure. In the future, tensor calculation

software could take such gradients into account as weights for the different input PCS when calculating the tensor based on a crystal or NMR structure.

Liquid-state NMR and X-ray crystallography are techniques of choice to study proteins and protein complexes at atomic resolution. However, both suffer from different limitations. X-ray crystallography, acknowledged to be very powerful to grasp the protein ground state, fails to capture potential secondary, lowly-populated states. Crystal packing can cause misrepresentations of the structure in solution. As for liquid-state NMR, it is limited by the protein/protein complex size, and the quality of the spectrum in case of non-stable or lowly soluble protein. Paramagnetic NMR is positioned to help filling the gap between these two techniques, being capable to probe low-populated states with PRE and determine structures from PCS and RDC. Paramagnetic NMR is nowadays an essential part of NMR, providing different kinds of effect that can be tuned easily to the protein of interest with different lanthanoids. The wide range of observable effects, reporting on distances as well as on dynamics is advantageously incorporated into more user-friendly software making paramagnetic NMR useful and attractive for non-NMR spectroscopists.

References

1. Bax, A. et al. Practical Aspects of Proton Carbon Carbon Proton 3-Dimensional Correlation Spectroscopy of C-13-Labeled Proteins. *J Magn Reson* **87**, 620-627 (1990).
2. Kay, L.E., Ikura, M. & Bax, A. Proton Proton Correlation Via Carbon Carbon Couplings - a 3-Dimensional Nmr Approach for the Assignment of Aliphatic Resonances in Proteins Labeled with C-13. *J Am Chem Soc* **112**, 888-889 (1990).
3. Bax, A., Clore, G.M. & Gronenborn, A.M. H-1-H-1 Correlation Via Isotropic Mixing of C-13 Magnetization, a New 3-Dimensional Approach for Assigning H-1 and C-13 Spectra of C-13-Enriched Proteins. *J Magn Reson* **88**, 425-431 (1990).
4. Anet, F.A.L. & Bourn, A.J.R. Nuclear Magnetic Resonance Spectral Assignments from Nuclear Overhauser Effects. *J Am Chem Soc* **87**, 5250-& (1965).
5. Ni, F. & Scheraga, H.A. Use of the Transferred Nuclear Overhauser Effect to Determine the Conformations of Ligands Bound to Proteins. *Accounts Chem Res* **27**, 257-264 (1994).
6. Ikura, M., Kay, L.E. & Bax, A. A Novel-Approach for Sequential Assignment of H-1, C-13, and N-15 Spectra of Larger Proteins - Heteronuclear Triple-Resonance 3-Dimensional Nmr-Spectroscopy - Application to Calmodulin. *Biochemistry-Us* **29**, 4659-4667 (1990).
7. Kay, L.E., Ikura, M., Tschudin, R. & Bax, A. 3-Dimensional Triple-Resonance Nmr-Spectroscopy of Isotopically Enriched Proteins. *J Magn Reson* **89**, 496-514 (1990).
8. Grzesiek, S. & Bax, A. An Efficient Experiment for Sequential Backbone Assignment of Medium-Sized Isotopically Enriched Proteins. *J Magn Reson* **99**, 201-207 (1992).
9. Muhandiram, D.R. & Kay, L.E. Gradient-Enhanced Triple-Resonance 3-Dimensional Nmr Experiments with Improved Sensitivity. *J Magn Reson Ser B* **103**, 203-216 (1994).
10. Farmer, B.T., Venters, R.A., Spicer, L.D., Wittekind, M.G. & Muller, L. A Refocused and Optimized Hnca - Increased Sensitivity and Resolution in Large Macromolecules. *J Biomol NMR* **2**, 195-202 (1992).
11. Clubb, R.T., Thanabal, V. & Wagner, G. A Constant-Time 3-Dimensional Triple-Resonance Pulse Scheme to Correlate Intraresidue H-1(N), N-15, and C-13() Chemical-Shifts in N-15-C-13-Labeled Proteins. *J Magn Reson* **97**, 213-217 (1992).
12. Bax, A. & Ikura, M. An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the

- preceding residue in uniformly $^{15}\text{N}/^{13}\text{C}$ enriched proteins. *J Biomol NMR* **1**, 99-104 (1991).
13. McIntosh, L.P. & Dahlquist, F.W. Biosynthetic incorporation of ^{15}N and ^{13}C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q Rev Biophys* **23**, 1-38 (1990).
 14. Grzesiek, S. et al. ^1H , ^{13}C , and ^{15}N NMR backbone assignments and secondary structure of human interferon-gamma. **31**, 8180-90 (1992).
 15. Zimmerman, D.E. et al. Automated analysis of protein NMR assignments using methods from artificial intelligence. *J Mol Biol* **269**, 592-610 (1997).
 16. Bahrami, A., Assadi, A.H., Markley, J.L. & Eghbalnia, H.R. Probabilistic interaction network of evidence algorithm and its application to complete labeling of peak lists from protein NMR spectroscopy. *PLoS Comput Biol* **5**, e1000307 (2009).
 17. Lee, W., Westler, W.M., Bahrami, A., Eghbalnia, H.R. & Markley, J.L. PINE-SPARKY: graphical interface for evaluating automated probabilistic peak assignments in protein NMR spectroscopy. *Bioinformatics* **25**, 2085-7 (2009).
 18. Guntert, P., Mumenthaler, C. & Wuthrich, K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* **273**, 283-98 (1997).
 19. Guntert, P. Automated NMR structure calculation with CYANA. *Methods Mol Biol* **278**, 353-78 (2004).
 20. Skinner, S.P. et al. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. **66**, 111-124 (2016).
 21. Lee, W., Tonelli, M. & Markley, J.L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-7 (2015).
 22. Mainz, A. et al. NMR Spectroscopy of Soluble Protein Complexes at One Mega-Dalton and Beyond. *Angew Chem Int Edit* **52**, 8746-8751 (2013).
 23. Wiesnerl, S. & Sprangers, R. Methyl groups as NMR probes for biomolecular interactions. *Curr Opin Struct Biol* **35**, 60-67 (2015).
 24. Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* **94**, 12366-71 (1997).
 25. Tugarinov, V., Hwang, P.M., Ollershaw, J.E. & Kay, L.E. Cross-correlated relaxation enhanced H-1-C-13 NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* **125**, 10420-10428 (2003).

26. Tugarinov, V. & Kay, L.E. Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins. *ChemBiochem* **6**, 1567-77 (2005).
27. Gross, J.D., Gelev, V.M. & Wagner, G. A sensitive and robust method for obtaining intermolecular NOEs between side chains in large protein complexes. *J Biomol NMR* **25**, 235-42 (2003).
28. Sprangers, R. & Kay, L.E. Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **445**, 618-22 (2007).
29. Tugarinov, V., Kanelis, V. & Kay, L.E. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc* **1**, 749-54 (2006).
30. Amero, C. et al. A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J Biomol Nmr* **50**, 229-236 (2011).
31. Mas, G., Crublet, E., Hamelin, O., Gans, P. & Boisbouvier, J. Specific labeling and assignment strategies of valine methyl groups for NMR studies of high molecular weight proteins. *J Biomol NMR* **57**, 251-262 (2013).
32. Tugarinov, V. & Kay, L.E. Quantitative NMR studies of high molecular weight proteins: application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G. *J Mol Biol* **327**, 1121-33 (2003).
33. Kerfah, R., Hamelin, O., Boisbouvier, J. & Marion, D. CH₃-specific NMR assignment of alanine, isoleucine, leucine and valine methyl groups in high molecular weight proteins using a single sample. *J Biomol NMR* **63**, 389-402 (2015).
34. Kainosho, M. et al. Optimal isotope labelling for NMR protein structure determinations. **440**, 52-7 (2006).
35. Gelis, I. et al. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* **131**, 756-769 (2007).
36. Velyvis, A., Schachman, H.K. & Kay, L.E. Assignment of Ile, Leu, and Val methyl correlations in supra-molecular systems: an application to aspartate transcarbamoylase. *J Am Chem Soc* **131**, 16534-43 (2009).
37. Venditti, V., Fawzi, N.L. & Clore, G.M. Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear overhauser enhancement spectroscopy. *J Biomol NMR* **51**, 319-328 (2011).
38. Xu, Y.Q. & Matthews, S. MAP-XSII: an improved program for the automatic assignment of methyl resonances in large proteins. *J Biomol NMR* **55**, 179-187 (2013).

39. Chao, F.A. et al. FLAMEnGO 2.0: An enhanced fuzzy logic algorithm for structure-based assignment of methyl group resonances. *J Magn Reson* **245**, 17-23 (2014).
40. Velyvis, A., Schachman, H.K. & Kay, L.E. Assignment of Ile, Leu, and Val methyl correlations in supra-molecular systems: an application to aspartate transcarbamoylase. **131**, 16534-43 (2009).
41. John, M. et al. Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. *J Am Chem Soc* **129**, 13749-13757 (2007).
42. Skinner, S.P., Moshev, M., Hass, M.A.S. & Ubink, M. PARAssign-paramagnetic NMR assignments of protein nuclei on the basis of pseudocontact shifts. *J Biomol NMR* **55**, 379-389 (2013).
43. Lescanne, M. et al. Methyl group assignment using pseudocontact shifts with PARAssign. *J Biomol NMR* **69**, 183-195 (2017).
44. LeMaster, D.M. & Richards, F.M. NMR sequential assignment of Escherichia coli thioredoxin utilizing random fractional deuteration. **27**, 142-50 (1988).
45. LeMaster, D.M. Deuterium labelling in NMR structural analysis of larger proteins. *Q Rev Biophys* **23**, 133-74 (1990).
46. Sattler, M. & Fesik, S.W. Use of deuterium labeling in NMR: Overcoming a sizeable problem. *Structure* **4**, 1245-1249 (1996).
47. Otten, R., Chu, B., Krewulak, K.D., Vogel, H.J. & Mulder, F.A.A. Comprehensive and Cost-Effective NMR Spectroscopy of Methyl Groups in Large Proteins. *J Am Chem Soc* **132**, 2952-2960 (2010).
48. Bouvignies, G. & Kay, L.E. A 2D C-13-CEST experiment for studying slowly exchanging protein systems using methyl probes: an application to protein folding. *J Biomol NMR* **53**, 303-310 (2012).
49. Rennella, E., Huang, R., Velyvis, A. & Kay, L.E. (CHD2)-C-13-CEST NMR spectroscopy provides an avenue for studies of conformational exchange in high molecular weight proteins. *J Biomol NMR* **63**, 187-199 (2015).
50. Ishima, R. & Torchia, D.A. Estimating the time scale of chemical exchange of proteins from measurements of transverse relaxation rates in solution. *J Biomol NMR* **14**, 369-372 (1999).
51. Kurauskas, V., Schanda, P. & Sounier, R. Methyl-Specific Isotope Labeling Strategies for NMR Studies of Membrane Proteins. *Methods Mol Biol* **1635**, 109-123 (2017).
52. Fischer, M. et al. Synthesis of a C-13-methyl-group-labeled methionine precursor as a useful tool for simplifying protein structural analysis by NMR spectroscopy. *Chembiochem* **8**, 610-612 (2007).

53. Ayala, I. et al. An optimized isotopic labelling strategy of isoleucine-gamma(2) methyl groups for solution NMR studies of high molecular weight proteins. *Chem Commun* **48**, 1434-1436 (2012).
54. Isaacson, R.L. et al. A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. *J Am Chem Soc* **129**, 15428+ (2007).
55. Tugarinov, V. & Kay, L.E. An isotope labeling strategy for methyl TROSY spectroscopy. *J Biomol NMR* **28**, 165-172 (2004).
56. Tugarinov, V. & Kay, L.E. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J Am Chem Soc* **125**, 13868-13878 (2003).
57. Shah, D.M. et al. Rapid Protein-Ligand Costructures from Sparse NOE Data. *J Med Chem* **55**, 10786-10790 (2012).
58. Pellecchia, M. et al. NMR-based structural characterization of large protein-ligand interactions. *J Biomol NMR* **22**, 165-173 (2002).
59. Dasgupta, S. et al. Narrowing the conformational space sampled by two-domain proteins with paramagnetic probes in both domains. *J Biomol NMR* **51**, 253-63 (2011).
60. Bertini, I. et al. MaxOcc: a web portal for maximum occurrence analysis. *J Biomol NMR* **53**, 271-80 (2012).
61. Lee, L. & Sykes, B.D. Strategies for the uses of lanthanide NMR shift probes in the determination of protein structure in solution. Application to the EF calcium binding site of carp parvalbumin. *Biophys J* **32**, 193-210 (1980).
62. Saio, T., Ogura, K., Yokochi, M., Kobashigawa, Y. & Inagaki, F. Two-point anchoring of a lanthanide-binding peptide to a target protein enhances the paramagnetic anisotropic effect. *J Biomol NMR* **44**, 157-66 (2009).
63. Su, X.C., McAndrew, K., Huber, T. & Otting, G. Lanthanide-binding peptides for NMR measurements of residual dipolar couplings and paramagnetic effects from multiple angles. *J Am Chem Soc* **130**, 1681-7 (2008).
64. Keizers, P.H., Desreux, J.F., Overhand, M. & Ubbink, M. Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. **129**, 9292-3 (2007).
65. Keizers, P.H., 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. **130**, 14802-12 (2008).
66. Rodriguez-Castaneda, F., Haberz, P., Leonov, A. & Griesinger, C. Paramagnetic tagging of diamagnetic proteins for solution NMR. *Magn Reson Chem* **44 Spec No**, S10-6 (2006).

67. Su, X.C. et al. A dipicolinic acid tag for rigid lanthanide tagging of proteins and paramagnetic NMR spectroscopy. *J Am Chem Soc* **130**, 10486-7 (2008).
68. Wohnert, J., Franz, K.J., Nitz, M., Imperiali, B. & Schwalbe, H. Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. *J Am Chem Soc* **125**, 13338-13339 (2003).
69. Pintacuda, G., John, M., Su, X.C. & Otting, G. NMR structure determination of protein-ligand complexes by lanthanide labeling. *Acc Chem Res* **40**, 206-12 (2007).
70. Robinson, C. Liquid-Crystalline Structures in Polypeptide Solutions. *Tetrahedron* **13**, 219+ (1961).
71. Vavrinska, A. et al. Impact of nucleic acid self-alignment in a strong magnetic field on the interpretation of indirect spin-spin interactions. *J Biomol NMR* **64**, 53-62 (2016).
72. Vavrinska, A. et al. Erratum to: Impact of nucleic acid self-alignment in a strong magnetic field on the interpretation of indirect spin-spin interactions. *J Biomol NMR* **65**, 49 (2016).
73. Ottiger, M., Delaglio, F. & Bax, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson* **131**, 373-8 (1998).
74. Yao, L., Ying, J. & Bax, A. Improved accuracy of ¹⁵N-¹H scalar and residual dipolar couplings from gradient-enhanced IPAP-HSQC experiments on protonated proteins. **43**, 161-70 (2009).
75. Tolman, J.R. & Ruan, K. NMR residual dipolar couplings as probes of biomolecular dynamics. *Chem Rev* **106**, 1720-1736 (2006).
76. Xu, X. et al. Intermolecular dynamics studied by paramagnetic tagging. *J Biomol NMR* **43**, 247-54 (2009).
77. Ravera, E. et al. Insights into Domain-Domain Motions in Proteins and RNA from Solution NMR. *Accounts Chem Res* **47**, 3118-3126 (2014).
78. Lipari, G. & Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity. *J Am Chem Soc* **104**, 4546-4559 (1982).
79. Clore, G.M. Exploring sparsely populated states of macromolecules by diamagnetic and paramagnetic NMR relaxation. *Protein Sci* **20**, 229-46 (2011).
80. Volkov, A.N., Worrall, J.A., Holtzmann, E. & Ubbink, M. Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. *Proc Natl Acad Sci U S A* **103**, 18945-50 (2006).
81. Iwahara, J. & Clore, G.M. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. **440**, 1227-30 (2006).

82. Clore, G.M. Visualizing lowly-populated regions of the free energy landscape of macromolecular complexes by paramagnetic relaxation enhancement. *Mol Biosyst* **4**, 1058-69 (2008).
83. Bertini, I., Luchinat, C., Parigi, G. & Ravera, E. NMR of paramagnetic molecules. Vol. 2 (Elsevier, 2015).
84. Billeter, M., Wagner, G. & Wuthrich, K. Solution NMR structure determination of proteins revisited. *J Biomol NMR* **42**, 155-158 (2008).
85. Bryson, M., Tian, F., Prestegard, J.H. & Valafar, H. REDCRAFT: a tool for simultaneous characterization of protein backbone structure and motion from RDC data. *J Magn Reson* **191**, 322-34 (2008).
86. Valafar, H. & Prestegard, J.H. REDCAT: a residual dipolar coupling analysis tool. *J Magn Reson* **167**, 228-41 (2004).
87. Rinaldelli, M. et al. Simultaneous use of solution NMR and X-ray data in REFMAC5 for joint refinement/detection of structural differences. *Acta Crystallogr D Biol Crystallogr* **70**, 958-67 (2014).
88. Meiler, J., Peti, W. & Griesinger, C. DipoCoup: A versatile program for 3D-structure homology comparison based on residual dipolar couplings and pseudocontact shifts. *J Biomol NMR* **17**, 283-94 (2000).
89. Schmitz, C., Vernon, R., Otting, G., Baker, D. & Huber, T. Protein structure determination from pseudocontact shifts using ROSETTA. *J Mol Biol* **416**, 668-77 (2012).
90. Loffler, P., Schmitz, S., Hupfeld, E., Sterner, R. & Merkl, R. Rosetta:MSF: a modular framework for multi-state computational protein design. *PLoS Comput Biol* **13**, e1005600 (2017).
91. Ozenne, V. et al. Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. **28**, 1463-70 (2012).
92. Schwieters, C.D., Kuszewski, J.J., Tjandra, N. & Clore, G.M. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65-73 (2003).
93. Schwieters, C.D., Kuszewski, J.J. & Clore, G.M. Using Xplor-NIH for NMR molecular structure determination. *Prog Nucl Mag Res Sp* **48**, 47-62 (2006).
94. Schwieters, C.D., Bermejo, G.A. & Clore, G.M. Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci* **27**, 26-40 (2018).
95. Dominguez, C., Boelens, R. & Bonvin, A.M.J.J. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125**, 1731-1737 (2003).

96. van Zundert, G.C.P. et al. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **428**, 720-725 (2016).
97. Bertini, I. et al. Conformational space of flexible biological macromolecules from average data. *J Am Chem Soc* **132**, 13553-8 (2010).
98. Rinaldelli, M., Carlon, A., Ravera, E., Parigi, G. & Luchinat, C. FANTEN: a new web-based interface for the analysis of magnetic anisotropy-induced NMR data. *J Biomol NMR* **61**, 21-34 (2015).
99. Schmitz, C., Stanton-Cook, M.J., Su, X.C., Otting, G. & Huber, T. Numbat: an interactive software tool for fitting Delta chi-tensors to molecular coordinates using pseudocontact shifts. *J Biomol NMR* **41**, 179-189 (2008).
100. Schmitz, C. et al. Efficient chi-tensor determination and NH assignment of paramagnetic proteins. *J Biomol NMR* **35**, 79-87 (2006).
101. Pan, Y.Z. et al. Sequence-specific assignment of methyl groups from the neuronal SNARE complex using lanthanide-induced pseudocontact shifts. *J Biomol NMR* **66**, 281-293 (2016).
102. Pintacuda, G. et al. Fast structure-based assignment of N-15 HSQC spectra of selectively N-15-labeled paramagnetic proteins. *J Am Chem Soc* **126**, 2963-2970 (2004).
103. Pearl, L.H. The HSP90 Molecular Chaperone-An Enigmatic ATPase. *Biopolymers* **105**, 594-607 (2016).
104. Karagoz, G.E. et al. N-terminal domain of human Hsp90 triggers binding to the cochaperone p23. *P Natl Acad Sci USA* **108**, 580-585 (2011).
105. Li, J., Soroka, J. & Buchner, J. The Hsp90 chaperone machinery: Conformational dynamics and regulation by co-chaperones. *Bba-Mol Cell Res* **1823**, 624-635 (2012).
106. Keizers, P.H.J., Desreux, J.F., Overhand, M. & Ubink, M. Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. *J Am Chem Soc* **129**, 9292-+ (2007).
107. Keizers, P.H.J., Saragliadis, A., Hiruma, Y., Overhand, M. & Ubink, 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).
108. Delaglio, F. et al. Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. *J Biomol Nmr* **6**, 277-293 (1995).
109. Vranken, W.F. et al. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59**, 687-696 (2005).
110. Bertini, I., Luchinat, C., Parigi, G. & Ravera, E. NMR of paramagnetic molecules. Vol. 2 261 (Cathleen Sether, Elsevier Science, 2015).

111. Li, J. et al. Structure insights into mechanisms of ATP hydrolysis and the activation of human heat-shock protein 90. *Acta Bioch Bioph Sin* **44**, 300-306 (2012).
112. Kuhn, H.W. The Hungarian method for the assignment problem. *NAVRES LOG* **2**, 83-97 (1955).
113. Kraft, D. A software package for sequential quadratic programming (trans: Center DGA). *Tech Rep. DFVLR-FB. Institute for Flight Mechanics: Kohn*, 88-28 (1988).
114. Keizers, P.H., Desreux, J.F., Overhand, M. & Ubbink, M. Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. *J Am Chem Soc* **129**, 9292-3 (2007).
115. Jacobs, D.M. et al. NMR backbone assignment of the N-terminal domain of human HSP90. *J Biomol NMR* **36**, 52-52 (2006).
116. Pederson, K. et al. NMR characterization of HtpG, the E. coli Hsp90, using sparse labeling with ¹³C-methyl alanine. *J Biomol NMR* (2017).
117. Zhang, H.Y. & van Ingen, H. Isotope-labeling strategies for solution NMR studies of macromolecular assemblies. *Curr Opin Struc Biol* **38**, 75-82 (2016).
118. Miyanoiri, Y. et al. Highly efficient residue-selective labeling with isotope-labeled Ile, Leu, and Val using a new auxotrophic E-coli strain. *J Biomol NMR* **65**, 109-119 (2016).
119. Tolman, J.R., Flanagan, J.M., Kennedy, M.A. & Prestegard, J.H. Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. **92**, 9279-83 (1995).
120. Tjandra, N., Grzesiek, S. & Bax, A. Magnetic field dependence of nitrogen-proton J splittings in N-15-enriched human ubiquitin resulting from relaxation interference and residual dipolar coupling. *J Am Chem Soc* **118**, 6264-6272 (1996).
121. Chen, K. & Tjandra, N. The Use of Residual Dipolar Coupling in Studying Proteins by NMR. *Top Curr Chem* **326**, 47-67 (2012).
122. Senn, H., Eugster, A., Otting, G., Suter, F. & Wuthrich, K. ¹⁵N-labeled P22 c2 repressor for nuclear magnetic resonance studies of protein-DNA interactions. **14**, 301-6 (1987).
123. Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E. & Dahlquist, F.W. Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. **177**, 44-73 (1989).
124. Lee, K.M., Androphy, E.J. & Baleja, J.D. A Novel Method for Selective Isotope Labeling of Bacterially Expressed Proteins. *J Biomol Nmr* **5**, 93-96 (1995).

125. Waugh, D.S. Genetic tools for selective labeling of proteins with alpha-N-15-amino acids. *J Biomol Nmr* **8**, 184-192 (1996).
126. HW, K. The Hungarian method for the assignment problem. (1955).
127. Schmitz, C., Stanton-Cook, M.J., Su, X.C., Otting, G. & Huber, T. Numbat: an interactive software tool for fitting Deltachi-tensors to molecular coordinates using pseudocontact shifts. **41**, 179-89 (2008).
128. DeLano, W.L. PyMOL molecular viewer: Updates and refinements. *Abstr Pap Am Chem S* **238**(2009).
129. Bohacek, R.S., McMartin, C. & Guida, W.C. The art and practice of structure-based drug design: a molecular modeling perspective. *Med Res Rev* **16**, 3-50 (1996).
130. Hajduk, P.J. & Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* **6**, 211-219 (2007).
131. Viegas, A., Manso, J., Nobrega, F.L. & Cabrita, E.J. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *J Chem Educ* **88**, 990-994 (2011).
132. Guan, J.Y. et al. Small-molecule binding sites on proteins established by paramagnetic NMR spectroscopy. *J Am Chem Soc* **135**, 5859-68 (2013).
133. Jahnke, W. et al. Second-site NMR screening with a spin-labeled first ligand. *J Am Chem Soc* **122**, 7394-7395 (2000).
134. Jahnke, W., Rudisser, S. & Zurini, M. Spin label enhanced NMR screening. *J Am Chem Soc* **123**, 3149-3150 (2001).
135. Otting, G. Protein NMR Using Paramagnetic Ions. *Annu Rev Biophys* **39**, 387-405 (2010).
136. Liu, W.M., Overhand, M. & Ubbink, M. The application of paramagnetic lanthanoid ions in NMR spectroscopy on proteins. *Coordin Chem Rev* **273**, 2-12 (2014).
137. Sprangers, R. et al. TROSY-based NMR evidence for a novel class of 20S proteasome inhibitors. *Biochemistry-Us* **47**, 6727-6734 (2008).
138. Nagaraju, G.P. et al. Epigenetic effects of inhibition of heat shock protein 90 (HSP90) in human pancreatic and colon cancer. *Cancer Lett* **402**, 110-116 (2017).
139. Murray, C.W. et al. Fragment-based drug discovery applied to Hsp90. Discovery of two lead series with high ligand efficiency. *J Med Chem* **53**, 5942-55 (2010).
140. Den, R.B. & Lu, B. Heat shock protein 90 inhibition: rationale and clinical potential. *Ther Adv Med Oncol* **4**, 211-8 (2012).
141. Palmer, A.G., Cavanagh, J., Wright, P.E. & Rance, M. Sensitivity Improvement in Proton-Detected 2-Dimensional Heteronuclear Correlation Nmr-Spectroscopy. *J Magn Reson* **93**, 151-170 (1991).

142. Kay, L.E., Keifer, P. & Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J Am Chem Soc* **114**, 10663-10665 (1992).
143. Schleucher, J. et al. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J Biomol NMR* **4**, 301-6 (1994).
144. Waudby, C.A., Ramos, A., Cabrita, L.D. & Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Sci Rep-Uk* **6**(2016).
145. Brewer, K.D. et al. Dynamic binding mode of a Synaptotagmin-1-SNARE complex in solution. *Nat Struct Mol Biol* **22**, 555-+ (2015).
146. Gobl, C. et al. Increasing the Chemical-Shift Dispersion of Unstructured Proteins with a Covalent Lanthanide Shift Reagent. *Angew Chem Int Edit* **55**, 14847-14851 (2016).
147. Saio, T., Guan, X., Rossi, P., Economou, A. & Kalodimos, C.G. Structural basis for protein antiaggregation activity of the trigger factor chaperone. **344**, 1250494 (2014).
148. Kerfah, R., Plevin, M.J., Sounier, R., Gans, P. & Boisbouvier, J. Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr Opin Struct Biol* **32**, 113-22 (2015).
149. Zuin, A., Isasa, M. & Crosas, B. Ubiquitin signaling: extreme conservation as a source of diversity. **3**, 690-701 (2014).
150. Pickart, C.M. & Eddins, M.J. Ubiquitin: structures, functions, mechanisms. *Bba-Mol Cell Res* **1695**, 55-72 (2004).
151. Harper, W. et al. Quantifying Ubiquitin Signaling for Mitophagy. *Mol Cell Proteomics* **16**, S39-S39 (2017).
152. Wickliffe, K., Williamson, A., Jin, L.Y. & Rape, M. The Multiple Layers of Ubiquitin-Dependent Cell Cycle Control. *Chem Rev* **109**, 1537-1548 (2009).
153. Chen, Z.J. & Pickart, C.M. A 25-Kilodalton Ubiquitin Carrier Protein (E2) Catalyzes Multi-Ubiquitin Chain Synthesis Via Lysine-48 of Ubiquitin. *J Biol Chem* **265**, 21835-21842 (1990).
154. Pickart, C.M. & Fushman, D. Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* **8**, 610-616 (2004).
155. Fushman, D. & Wilkinson, K.D. Structure and recognition of polyubiquitin chains of different lengths and linkage. *F1000 Biol Rep* **3**, 26 (2011).
156. Lee, A.E., Castaneda, C.A., Wang, Y., Fushman, D. & Fenselau, C. Preparing to read the ubiquitin code: a middle-out strategy for characterization of all lysine-linked diubiquitins. *J Mass Spectrom* **49**, 1272-8 (2014).
157. Husnjak, K. & Dikic, I. Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu Rev Biochem* **81**, 291-322 (2012).

158. Varshavsky, A. The Ubiquitin System, Autophagy, and Regulated Protein Degradation. *Annu Rev Biochem* **86**, 123-128 (2017).
159. Ikeda, F. & Dikic, I. Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. *EMBO Rep* **9**, 536-42 (2008).
160. Castaneda, C.A. et al. Linkage-specific conformational ensembles of non-canonical polyubiquitin chains. *Phys Chem Chem Phys* **18**, 5771-5788 (2016).
161. Castaneda, C.A., Kashyap, T.R., Nakasone, M.A., Krueger, S. & Fushman, D. Unique structural, dynamical, and functional properties of k11-linked polyubiquitin chains. **21**, 1168-81 (2013).
162. Gatti, M. et al. RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. *Cell Rep* **10**, 226-38 (2015).
163. Peng, J.M. et al. A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **21**, 921-926 (2003).
164. Pinato, S., Gatti, M., Scandiuizzi, C., Confalonieri, S. & Penengo, L. UMI, a novel RNF168 ubiquitin binding domain involved in the DNA damage signaling pathway. *Mol Cell Biol* **31**, 118-26 (2011).
165. Geng, F., Wenzel, S. & Tansey, W.P. Ubiquitin and proteasomes in transcription. *Annu Rev Biochem* **81**, 177-201 (2012).
166. Park, Y., Jin, H.S., Aki, D., Lee, J. & Liu, Y.C. The ubiquitin system in immune regulation. *Adv Immunol* **124**, 17-66 (2014).
167. Varadan, R., Walker, O., Pickart, C. & Fushman, D. Structural properties of polyubiquitin chains in solution. *J Mol Biol* **324**, 637-47 (2002).
168. Komander, D. & Rape, M. The ubiquitin code. *Annu Rev Biochem* **81**, 203-29 (2012).
169. Kristariyanto, Y.A., Abdul Rehman, S.A., Weidlich, S., Knebel, A. & Kulathu, Y. A single MIU motif of MINDY-1 recognizes K48-linked polyubiquitin chains. *EMBO Rep* **18**, 392-402 (2017).
170. Lai, M.Y., Zhang, D., Laronde-Leblanc, N. & Fushman, D. Structural and biochemical studies of the open state of Lys48-linked diubiquitin. *Biochim Biophys Acta* **1823**, 2046-56 (2012).
171. Berlin, K. et al. Recovering a Representative Conformational Ensemble from Underdetermined Macromolecular Structural Data. *J Am Chem Soc* **135**, 16595-16609 (2013).
172. Cook, W.J., Jeffrey, L.C., Carson, M., Chen, Z. & Pickart, C.M. Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J Biol Chem* **267**, 16467-71 (1992).
173. Prestegard, J.H., Al-Hashimi, H.M. & Tolman, J.R. NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Q Rev Biophys* **33**, 371-424 (2000).

174. Lipsitz, R.S. & Tjandra, N. Residual dipolar couplings in NMR structure analysis. *Annu Rev Bioph Biom* **33**, 387-413 (2004).
175. Prestegard, J.H., Mayer, K.L., Valafar, H. & Benison, G.C. Determination of protein backbone structures from residual dipolar couplings. *Methods Enzymol* **394**, 175-209 (2005).
176. Bertini, I., Luchinat, C. & Parigi, G. Paramagnetic constraints: An aid for quick solution structure determination of paramagnetic metalloproteins. *Concept Magnetic Res* **14**, 259-286 (2002).
177. Tolman, J.R., Flanagan, J.M., Kennedy, M.A. & Prestegard, J.H. Nuclear magnetic dipole interactions in field-oriented proteins - information for structure determination in solution. **92**, 9279-9283 (1995).
178. Tjandra, N. & Bax, A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. **278**, 1111-4 (1997).
179. Longinetti, M., Luchinat, C., Parigi, G. & Sgheri, L. Efficient determination of the most favoured orientations of protein domains from paramagnetic NMR data. *Inverse Probl* **22**, 1485-1502 (2006).
180. Gardner, R.J., Longinetti, M. & Sgheri, L. Reconstruction of orientations of a moving protein domain from paramagnetic data. *Inverse Probl* **21**, 879-898 (2005).
181. Clore, G.M., Tang, C. & Iwahara, J. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr Opin Struct Biol* **17**, 603-616 (2007).
182. Bertini, I. et al. Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. *J Am Chem Soc* **129**, 12786-94 (2007).
183. Bertini, I. et al. Accurate solution structures of proteins from X-ray data and a minimal set of NMR data: calmodulin-peptide complexes as examples. *J Am Chem Soc* **131**, 5134-44 (2009).
184. van Dijk, A.D.J., Fushman, D. & Bonvin, A.M.J.J. Various strategies of using residual dipolar couplings in NMR-driven protein docking: Application to Lys48-linked Di-ubiquitin and validation against N-15-relaxation data. *Proteins* **60**, 367-381 (2005).
185. Castaneda, C.A., Spasser, L., Bavikar, S.N., Brik, A. & Fushman, D. Segmental Isotopic Labeling of Ubiquitin Chains To Unravel Monomer-Specific Molecular Behavior. *Angew Chem Int Edit* **50**, 11210-11214 (2011).
186. Vijay-Kumar, S., Bugg, C.E. & Cook, W.J. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* **194**, 531-44 (1987).
187. Ravera, E. et al. Insights into domain-domain motions in proteins and RNA from solution NMR. **47**, 3118-26 (2014).

188. Andralojc, 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).
189. Liu, W.M. et al. A pH-Sensitive, Colorful, Lanthanide-Chelating Paramagnetic NMR Probe. *J Am Chem Soc* **134**, 17306-17313 (2012).
190. DeMartino, G.N. & Gillette, T.G. Proteasomes: Machines for all reasons. *Cell* **129**, 659-662 (2007).
191. Rechsteiner, M., Realini, C. & Ustrell, V. The proteasome activator 11 S REG (PA28) and class I antigen presentation. *Biochem J* **345 Pt 1**, 1-15 (2000).
192. Hill, C.P., Masters, E.I. & Whitby, F.G. The 11S regulators of 20S proteasome activity. *Curr Top Microbiol Immunol* **268**, 73-89 (2002).
193. Lowe, J. et al. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. **268**, 533-9 (1995).
194. Ollerenshaw, J.E., Tugarinov, V. & Kay, L.E. Methyl TROSY: explanation and experimental verification. *Magn Reson Chem* **41**, 843-852 (2003).
195. Ishii, K. et al. Disassembly of the self-assembled, double-ring structure of proteasome alpha7 homo-tetradecamer by alpha6. *Sci Rep* **5**, 18167 (2015).
196. Prudencio, M. et al. A caged lanthanide complex as a paramagnetic shift agent for protein NMR. **10**, 3252-60 (2004).
197. Schanda, P., Kupce, E. & Brutscher, B. SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* **33**, 199-211 (2005).
198. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. **6**, 277-93 (1995).
199. Vranken, W.F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. **59**, 687-96 (2005).
200. Bertini, I., Luchinat, C. & Parigi, G. Magnetic susceptibility in paramagnetic NMR. *Prog Nucl Mag Res Sp* **40**, 249-273 (2002).
201. Bloch, F., Hansen, W.W. & Packard, M. The Nuclear Induction Experiment. *Phys Rev* **70**, 474-485 (1946).
202. Purcell, E.M., Torrey, H.C. & Pound, R.V. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys Rev* **69**, 37-38 (1946).
203. Holzgrabe, U., Deubner, R., Schollmayer, C. & Waibel, B. Quantitative NMR spectroscopy--applications in drug analysis. *J Pharm Biomed Anal* **38**, 806-12 (2005).
204. Jungnickel, J.L. & Forbes, J.W. Quantitative Measurement of Hydrogen Types by Integrated Nuclear Magnetic Resonance Intensities. *Anal Chem* **35**, 938-& (1963).

205. Pauli, G.F., Jaki, B.U. & Lankin, D.C. Quantitative ^1H NMR: development and potential of a method for natural products analysis. *J Nat Prod* **68**, 133-49 (2005).
206. Wuthrich, K. Protein structure determination in solution by NMR spectroscopy. *J Biol Chem* **265**, 22059-62 (1990).
207. Simon, B., Madl, T., Mackereth, C.D., Nilges, M. & Sattler, M. An efficient protocol for NMR-spectroscopy-based structure determination of protein complexes in solution. *Angew Chem Int Ed Engl* **49**, 1967-70 (2010).
208. Campbell, E. et al. The role of protein dynamics in the evolution of new enzyme function. *Nat Chem Biol* **12**, 944-950 (2016).
209. Gossert, A.D. & Jahnke, W. NMR in drug discovery: A practical guide to identification and validation of ligands interacting with biological macromolecules. *Prog Nucl Magn Reson Spectrosc* **97**, 82-125 (2016).
210. Pellicchia, M. et al. Perspectives on NMR in drug discovery: a technique comes of age. *Nat Rev Drug Discov* **7**, 738-45 (2008).
211. Styles, P. et al. A High-Resolution Nmr Probe in Which the Coil and Preamplifier Are Cooled with Liquid-Helium. *J Magn Reson* **60**, 397-404 (1984).
212. Kovacs, H., Moskau, D. & Spraul, M. Cryogenically cooled probes - a leap in NMR technology. *Prog Nucl Mag Res Sp* **46**, 131-155 (2005).
213. Morris, G.A. & Freeman, R. Enhancement of Nuclear Magnetic-Resonance Signals by Polarization Transfer. *J Am Chem Soc* **101**, 760-762 (1979).
214. Mulleti, S. et al. Super-Resolved Nuclear Magnetic Resonance Spectroscopy. *Sci Rep* **7**, 9651 (2017).
215. Rosen, M.K. et al. Selective methyl group protonation of perdeuterated proteins. *J Mol Biol* **263**, 627-36 (1996).
216. Gardner, K.H. & Kay, L.E. The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* **27**, 357-406 (1998).
217. Kazimierczuk, K. & Orekhov, V.Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew Chem Int Edit* **50**, 5556-5559 (2011).
218. Kazimierczuk, K. & Orekhov, V. Non-uniform sampling: post-Fourier era of NMR data collection and processing. *Magn Reson Chem* **53**, 921-6 (2015).
219. Kosinski, K., Stanek, J., Gorka, M.J., Zerko, S. & Kozminski, W. Reconstruction of non-uniformly sampled five-dimensional NMR spectra by signal separation algorithm. *J Biomol NMR* **68**, 129-138 (2017).
220. Freeman, R. & Kupce, E. New methods for fast multidimensional NMR. *J Biomol NMR* **27**, 101-13 (2003).
221. Arnesano, F., Banci, L. & Piccioli, M. NMR structures of paramagnetic metalloproteins. *Q Rev Biophys* **38**, 167-219 (2005).

222. Bertini, I., Luchinat, C., Parigi, G. & Pierattelli, R. Perspectives in paramagnetic NMR of metalloproteins. *3782-90* (2008).
223. Banci, L. et al. The three-dimensional structure in solution of the paramagnetic high-potential iron-sulfur protein I from *Ectothiorhodospira halophila* through nuclear magnetic resonance. *Eur J Biochem* **225**, 715-25 (1994).
224. Saio, T., Yokochi, M., Kumeta, H. & Inagaki, F. PCS-based structure determination of protein-protein complexes. *J Biomol NMR* **46**, 271-80 (2010).
225. Boisbouvier, J., Gans, P., Blackledge, M., Brutscher, B. & Marion, D. Long-range structural information in NMR studies of paramagnetic molecules from electron spin-nuclear spin cross-correlated relaxation. *J Am Chem Soc* **121**, 7700-7701 (1999).
226. Luchinat, C., Nagulapalli, M., Parigi, G. & Sgheri, L. Maximum occurrence analysis of protein conformations for different distributions of paramagnetic metal ions within flexible two-domain proteins. *J Magn Reson* **215**, 85-93 (2012).
227. Hass, M.A.S., Keizers, P.H.J., Blok, A., Hiruma, Y. & Ubbink, M. Validation of a Lanthanide Tag for the Analysis of Protein Dynamics by Paramagnetic NMR Spectroscopy. *J Am Chem Soc* **132**, 9952-9953 (2010).
228. Evangelidis, T. et al. Automated NMR resonance assignments and structure determination using a minimal set of 4D spectra. *Nat Commun* **9**, 384 (2018).
229. Robson, S.A. et al. Mixed pyruvate labeling enables backbone resonance assignment of large proteins using a single experiment. *Nat Commun* **9**, 356 (2018).
230. Crick, D.J. et al. Integral membrane protein structure determination using pseudocontact shifts. *J Biomol NMR* **61**, 197-207 (2015).
231. Sprangers, R. & Kay, L.E. Probing supramolecular structure from measurement of methyl (1)H-(13)C residual dipolar couplings. *J Am Chem Soc* **129**, 12668-9 (2007).
232. Alberts, B. The cell as a collection of protein machines: Preparing the next generation of molecular biologists. *Cell* **92**, 291-294 (1998).
233. Gavin, A.C. et al. Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631-636 (2006).
234. Levy, E.D., Pereira-Leal, J.B., Chothia, C. & Teichmann, S.A. 3D complex: A structural classification of protein complexes. *Plos Comput Biol* **2**, 1395-1406 (2006).
235. Andre, I., Strauss, C.E., Kaplan, D.B., Bradley, P. & Baker, D. Emergence of symmetry in homooligomeric biological assemblies. *Proc Natl Acad Sci U S A* **105**, 16148-52 (2008).

236. Gans, P. et al. Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chem Int Ed Engl* **49**, 1958-62 (2010).
237. Saio, T. et al. An NMR strategy for fragment-based ligand screening utilizing a paramagnetic lanthanide probe. *J Biomol NMR* **51**, 395-408 (2011).
238. Pearce, B.J.G. et al. Structure restraints from heteronuclear pseudocontact shifts generated by lanthanide tags at two different sites. *J Biomol NMR* **68**, 19-32 (2017).

ENGLISH SUMMARY

Nuclear magnetic resonance spectroscopy (NMR) has proven to be of great use in biochemistry. In the last decades, applications of paramagnetic NMR (pNMR) have been developed for protein research. pNMR has shown to be an especially powerful tool for protein structural refinement and for the assignment of protein NMR resonances to the corresponding nuclei. For instance, PARAssign software automatically assigns protein resonances using paramagnetic effects and a known structure of the protein. pNMR effects are caused by the presence of an unpaired electron described by a susceptibility tensor composed of an isotropic and an anisotropic component $\Delta\chi$, responsible for the effects observed in this thesis. pNMR effects can be generated on protein nuclei by paramagnetic probes such as Caged Lanthanide NMR Probe 5 (CLaNP-5) attached on the protein surface to two cysteines. Observable pNMR effects are, among other effects, peak shifts in NMR spectra called pseudo-contact shifts (PCS), and residual dipolar couplings (RDC). Thanks to its two-arm attachment, CLaNP-5 is highly rigid and its $\Delta\chi$ tensor can be predicted but also refined using experimental PCS.

Chapter 1 contains a short history of NMR assignment and an introduction to the techniques used in the presented work.

Chapter 2 describes an update of PARAssign software using PCS measured on methyl groups. A new criterion was developed to assess the reliability of every found assignment. The final assignments are plotted in a histogram, each coloured either in green when highly reliable, yellow when multiple assignments are suggested and white when no reliable assignment is found. The global quality of all the assignments can be then judged at a glance. Using 1 to 3 PCS datasets, generated with CLaNP-5 attached to up to 3 different double cysteine-mutants, PARAssign assigned up to 60% of the methyl groups with high reliability. It was shown that the two main limiting factors for the assignment quality were the completeness of the input datasets and the alignment of the different susceptibility tensors.

Chapter 3 introduces the incorporation of RDC as an independent dataset in addition to PCS in PARAssign. The goal was to get an extra independent dataset per paramagnetic centre to improve the assignment, without having to introduce another paramagnetic centre. After updating PARAssign program with extra required functions, preliminary tests were ran on simulated isoleucine and alanine ^{15}N labelled protein. PCS and RDC from a single paramagnetic centre were used. First results, using simulated data with up to 4 Hz error, showed improvement in the assignment when

using RDC in combination with PCS. However, the low accuracy of amide hydrogen atom positions in crystal structure models leads to large errors in the RDC prediction, limiting its usefulness for resonance assignment.

Chapter 4 illustrates the usefulness of protein methyl group PCS evolution for probing structural changes when the protein undergoes a ligand titration. In this chapter, a ligand from a fragment-based drug discovery library was used. Four methyl groups experienced a PCS difference between the free and bound state larger than 0.04 ppm. These PCS differences report on a modification of the distance between the methyl group and the paramagnetic centre. Titrations with diamagnetic and paramagnetic protein samples were performed with CLaNP-5 attached at 3 different locations. PCS isosurface cross-sections, for the free and the bound states were determined and visualised on the structure. Structural displacement in the range of 1-3 Å were observed between free and bound states. This application could be used to better understand re-orientation of methyl groups upon binding of small ligands for further drug hit development.

Chapter 5 describes the use of RDC to characterize highly dynamic ubiquitin dimer in solution. Diamagnetic RDC have already been used for such purposes but in this chapter paramagnetic RDC generated with CLaNP-5 were used, for which accurate $\Delta\chi$ tensor can be measured based on PCS. RDC were measured for both proximally and distally labelled di-ubiquitin at two pH values 6.8 and 7.2, to observe a possible switch from open to closed state. The preliminary results did not match either a single or a combination of two crystal structures. Ensembles of at least three conformations need to be tested to see whether they describe the system better.

Chapter 6 presents preliminary results of the use of PCS measured on a highly symmetric 14-mer protein for assignment with PARAssign. PCS were measured for 4 double cysteine mutants of the $\alpha7\alpha7$ ring proteasome. Measured PCS were compared to those predicted based on the properties of CLaNP-5 and its predictable orientation relative to the cysteines, and showed good correlation. CLaNP-5 tagging was not 100% for the 14 subunits, leading to multiple observed PCS, causing an extensive overlap of the resonances. A low-percentage tagging strategy was attempted to observe 13/14 diamagnetic peaks overlapping with 1/14 paramagnetic peaks but it was not successful due to low sensitivity.

Chapter 7 is a general discussion about the work presented in this thesis and an outlook for further development of PARAssign and the use of paramagnetic NMR for protein research in general.

NEDERLANDSE SAMENVATTING

Kernspinresonantiespectroscopie (*nuclear magnetic resonance spectroscopy*, NMR) heeft bewezen van grote waarde te zijn voor de biochemie. In de laatste decennia zijn ook toepassingen van paramagnetische NMR (pNMR) ontwikkeld voor eiwitonderzoek. pNMR is een bijzonder krachtig hulpmiddel gebleken bij de verfijning van eiwitstructuren en voor de toewijzing van de verschillende NMR resonanties van eiwitten aan de bijbehorende atoomkernen. De software PARAssign, bijvoorbeeld, voert deze toewijzing automatisch uit op basis van paramagnetische effecten en een bekende structuur van het eiwit. pNMR effecten worden veroorzaakt door de aanwezigheid van een ongepaard elektron. Ze worden beschreven door een gevoeligheidstensor, die bestaat uit een isotrope component en een anisotrope component, $\Delta\chi$, die verantwoordelijk is voor de effecten die beschreven worden in deze thesis. pNMR effecten op eiwitkernen kunnen gegenereerd door een paramagnetische sonde, zoals de “gekooide lanthanoïde NMR sonde 5” (*Caged Lanthanoid NMR Probe 5*, CLaNP-5), aan twee cysteïnes op het eiwitoppervlak te bevestigen. Onder de observeerbare pNMR effecten vallen onder andere verschuivingen van resonantiepieken in NMR spectra genaamd pseudo-contactverschuivingen (*pseudo-contact shifts*, PCS) en resterende dipolaire koppelingen (*residual dipolar couplings*, RDC). Dankzij zijn tweearmige bevestiging aan het eiwit is CLaNP-5 niet erg buigbaar en kan zijn $\Delta\chi$ tensor zowel voorspeld als experimenteel verfijnd worden.

Hoofdstuk 1 bevat een korte geschiedenis over het toewijzen van NMR spectra, evenals een inleiding tot de technieken die in dit werk gebruikt zijn.

Hoofdstuk 2 beschrijft een update van de software PARAssign, waardoor het PCS van methylgroepen kan gebruiken. Er is een nieuw criterium ontwikkeld om de betrouwbaarheid van iedere gevonden toewijzing te beoordelen. Deze beoordeling wordt weergegeven in het toewijzingshistogram door middel van een kleurcodering: resonanties met zeer betrouwbare toewijzing zijn groen gekleurd, resonanties waarvoor verschillende toewijzingen worden gesuggereerd zijn geel en resonanties zonder betrouwbare toewijzing zijn wit. Hierdoor kan de algehele kwaliteit van alle toewijzingen in één oogopslag beoordeeld worden. Door gebruik te maken van 1 tot 3 verschillende PCS datasets, gemeten op mutanten met CLaNP-5 ieder op een verschillende plek eraan bevestigd, heeft PARAssign tot 60% van de methylgroepen met hoge betrouwbaarheid kunnen toewijzen. Het bleek hierbij dat de compleetheid van de invoerdata en het tegen elkaar uitlijnen van de verschillende gevoeligheidssensoren de voornaamste limiterende factoren voor dit proces zijn.

Hoofdstuk 3 introduceert het gebruik van RDC in PARAssign, als onafhankelijke dataset naast PCS. Het idee is dat hierdoor voor ieder paramagnetisch centrum een extra, onafhankelijke dataset wordt verkregen, waardoor de toewijzing kan worden verbeterd zonder dat er extra paramagnetische centra geïntroduceerd hoeven te worden. Nadat de benodigde functionaliteit was ingebouwd in PARAssign, werd de software getest op gesimuleerd eiwit met ¹⁵N-gelabelde isoleucine en alanine. PCS en RDC van slechts één paramagnetisch centrum werden gesimuleerd met een afwijking tot 4 Hz. De eerste resultaten hiermee lieten zien dat, bij gebruik van RDC in combinatie met PCS, de toewijzing verbetert. De lage precisie waarmee de posities van amide-waterstofatomen in kristalstructuren bekend zijn leidt echter tot grote afwijkingen in de RDC-voorspelling, wat de bruikbaarheid ervan voor resonantietoewijzing blijkt te limiteren.

Hoofdstuk 4 illustreert het gebruik van PCS-ontwikkeling bij titratie met liganden om structurele veranderingen in eiwitten te volgen. In dit hoofdstuk is een ligand gebruikt uit een verzameling die gebruikt wordt om nieuwe medicijnen te ontdekken. Bij titratie van die ligand bij het eiwit voelden vier methylgroepen een PCS-verschil van meer dan 0,04 ppm tussen de vrije en de gebonden staat. Deze PCS-verschillen betekenen dat de afstand tussen de methylgroep en het paramagnetische centrum verandert. Er zijn titraties uitgevoerd met diamagnetische en paramagnetische CLANP-5, bevestigd op 3 verschillende locaties in het eiwit. Hierdoor konden PCS-isovlakken en hun doorsneden worden bepaald en gevisualiseerd op de eiwitstructuur, voor zowel de gebonden als de ongebonden toestand. Verplaatsingen van de methylgroepen, van 1 tot 3 Å tussen de twee toestanden, werden op die manier geobserveerd. Deze toepassing kan gebruikt worden bij de bepaling van het effect van eventuele nieuwe medicijnen op een eiwitstructuur.

Hoofdstuk 5 beschrijft het gebruik van RDC om de dynamische ubiquitine-dimeer in oplossing te beschrijven. Zoiets is met diamagnetische RDC eerder gedaan, maar in dit hoofdstuk wordt paramagnetische RDC met CLaNP-5 gebruikt. De $\Delta\chi$ tensor kan daarbij accuraat worden gemeten door de PCS te gebruiken. RDC van proximaal en distaal gelabelde di-ubiquitine werden ieder gemeten bij pH's 6,8 en 7,2. Bij het gebruik van deze data om een mogelijke conformatieverandering van een open naar een gesloten staat te bestuderen, bleken de resultaten noch met één van de twee kristalstructuren, noch met een combinatie ervan overeen te komen. Ensembles van ten minste drie conformaties zullen getest moeten worden om te zien of zij het systeem beter kunnen beschrijven.

Hoofdstuk 6 presenteert voorlopige resultaten van het gebruik van PCS voor de toewijzing met PARAssign van een 14-delig eiwit met een hoge graad van symmetrie. PCS werden gemeten met CLaNP-5 op 4 verschillende posities op het $\alpha7\alpha7$ ring proteasoom. De resultaten werden vergeleken met de PCS die op basis van de eigenschappen van CLaNP-5 voorspeld waren en vertoonden goede correlatie. Het bevestigen van CLaNP-5 was echter niet 100% effectief voor alle 14 eiwitdelen, wat leidde tot de observatie van meerdere PCS, die veel overlap tussen de verschillende paramagnetische resonanties veroorzaakten. Om dit effect te omzeilen werd een strategie met minder CLaNP-5 labeling geprobeerd, maar dit was niet succesvol vanwege de lage gevoeligheid.

Hoofdstuk 7 bevat een brede discussie over het werk dat gepresenteerd is in deze thesis, evenals een vooruitblik op de toekomstige ontwikkeling van PARAssign en het gebruik van paramagnetische NMR voor eiwitonderzoek in het algemeen.

LIST OF PUBLICATIONS

Lescanne M., Skinner SP., Blok A., Timmer M., Cerofolini L., Fragai M., Luchinat C., Ubbink M. (2017) *Methyl group assignment using pseudocontact shifts with PARAssign*. J. Biomol. NMR DOI: 10.1007/s10858-017-0136-3

Lescanne M., Ahuja P., Blok A., Timmer M., Akerud T., Ubbink M. (2018) *Methyl group reorientation under ligand binding probed by pseudocontact shifts*. (accepted for publication in J. Biomol. NMR)

Castañeda C., **Lescanne M.**, Timmer M., Ubbink M., Fushman D., *Flexibility in ubiquitin dimer probed by paramagnetic RDC*. (in preparation)

CURRIVULUM VITAE

Mathilde Lescanne was born the 2nd of April 1989 in Bordeaux, France. After her baccalauréat received with honours and two years of highly selective classes, Mathilde moved to Grenoble to start an engineering school. She finally completed a Master in physics with mention in life sciences. She discovered NMR during her first master internship with Dominique Marion at the Institute of Structural Biology of Grenoble where she worked on non-linear sampled NMR data processing. For her second master internship, still under the supervision of Dominique Marion, she worked on exchange phenomenon in a half-mega Dalton protein studied by liquid-state NMR. After graduating in 2013, she moved to Leiden to start her PhD with Prof. Dr. Marcellus Ubbink in the Macromolecular Biochemistry group of Leiden Institute of Chemistry. She worked on applications of paramagnetic NMR for protein research. She discovered there how to work in the wetlab and how to monitor a spectrometer. Within the European pNMR program, she had the opportunity to go abroad for two secondments. The first one took place in Florence, Italy, where she stayed two months to produce ¹³C-methyl group labelled proteins. The second one took place in AstraZeneca in Goteborg, Sweden, and lasted one month, to work on ligand/protein interaction probed by Δ PCS. The results of these projects are described in this thesis.



