technische universität dortmund

fakultät für chemie und chemische biologie

DISSERTATION

Characterization of protein structure and dynamics by solution and solid-state NMR

Zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

der

Fakultät Chemie und Chemische Biologie der

Technischen Universität Dortmund

Vorgelegt von

Sara Medina Gomez

aus Mexiko-Stadt

2023

Acknowledgments

When this journey started, I didn't know how it would be so far away from home in a place so different. After these years, I can only be thankful for all the people the Ph.D. journey brought to my life.

I want to thank Prof. Dr. Rasmus Linser for being my supervisor, for trusting me, and for all the lessons taught. I enjoyed learning and growing in his research group, as he has always been there to help with problem resolution and advising patiently when challenges became more complex. It is of great value to me all the time he invested in improving the science behind my projects with his advice and creating an enjoyable working atmosphere. I also thank the many opportunities I had to grow and learn in the scientific community thanks to all the workshops and conferences I was allowed and encouraged to attend.

My second thanks go to Prof. Dr. Daniel Rauh, not only for being the second reviewer but also for helping us to set up the conditions for starting to work on the MAP kinase p38α project. In particular, thanks for the many ideas, the meaningful discussion, and all the help from the group's beginning in TU Dortmund.

Like many other aspects of life, scientific work cannot be achieved alone. And for this, I am very grateful for having a great work atmosphere surrounded by supportive colleagues and friends. Thank Dr. Suresh K. Vasa and Dr. Alexander Klein for all the advice, help, and patience regarding all aspects of NMR spectroscopy: from experiments acquisition to data reconstruction and analysis. I thank Dr. Diana Carolina Pöllman for helping me gain confidence in the wet lab, for all the invaluable lessons and tips shared during the first months of my Ph.D. time, for always being available to help regardless of the distance, and for being one of my first friends in this new country. Many aspects of this thesis were completed thanks to all your very valuable help. For that, I will always be thankful.

To all current and past group members Dr. Petra Rovó, Dr. Himanshu Sing, Dr. Laura Kukuk, Ekatarina Burakova, Julia Kotschy, Benedikt Söldner, Anna Aucharova, and Suchandra R. Acharyya, I want to thank for the many times of scientific and personal exchange, as well for making the working atmosphere more enjoyable with our talks over coffee. Additionally, Christine Wehbrink and Vanessa Slota for administrative support. It has been a pleasure to share our international brunches and the opportunities to interact outside of the university setup.

The collaborative environment provided by the university creates the opportunity for fruitful collaborations. In this line, I would like to thank Dr. Mike Bührmann for all his help and advice for the p38α project and for providing plasmids and ligands, and Dr. Ilaria Visco for her help with the RhoGDI project and for providing the protein samples. Without our different scientific interactions, the research would not have gone this far.

I want to thank my parents for always encouraging me to do what makes me happy, believing in me, and supporting me in many ways. I would not be as interested in science as I am now if they had not encouraged my curiosity and fed my desire for continuous learning. Thank you for building a solid ground for your kids to grow passionate with all your love, understanding, and support. I want to thank Felipe and Chuy for always being there, listening to my strange ideas, and encouraging me to be myself. I can only be thankful for growing up with you.

My special thanks to Alex for all his support during the Ph.D. time and for being with me when I needed a hand or someone to talk to. For being not only my college but also my life partner. With your help, I made it to this point.

Last but not least, I want to acknowledge all the other persons I could interact with that are not mentioned here, as they have contributed to satisfactorily concluding this work.

Summary

The study of proteins for understanding biological processes or as an approach to treating diseases has become more relevant in recent years. Different structural biology approaches can help to answer some of these questions, including nuclear magnetic resonance (NMR) spectroscopy, which provides information at atomic resolution about the structure and dynamics on a broad range of timescales. Traditionally, protein studies are carried out by solution NMR, where many methodologies have been developed over the years that allow one to obtain a great variety of information. However, in solution NMR studies, the protein size is limited to proteins up to 50 kDa. Using different labeling schemes and different levels of protonation in the protein, with solid-state NMR and ^{1}H detection, we can access biologically relevant systems with a molecular size larger than 50 kDa and extract information of their structure and dynamics on a different timescale with atomic resolution. This thesis presents a combination of solution and solid-state NMR methodologies to understand the dynamical processes of the MAP kinase p38α and their changes upon binding of the ligand as a tool for characterizing the allosteric dynamic network present in the protein. All this was done by combining different labeling schemes and proving new methodologies for acquiring relaxation dispersion experiments that can be used for large protein systems. Furthermore, relaxation dispersion techniques in solid-state NMR are used to characterize the modulation of dynamics by inter- and intramolecular interaction for the SH3 domain of chicken α-spectrin by introducing a mutation in the RT loop.

Furthermore, this work presents the characterization of the secondary structural element for the N-terminal domain of RhoGDI1, a domain thought to be disordered in the apo form and ordered in complex with the GTPase client. However, with the experimental chemical-shift information, we proved it to have a secondary structure before complex formation.

Zussammenfassung

Die Untersuchung von Proteinen zum besseren Verständnis biologischer Prozesse sowie zur Behandlung von Krankheiten hat in der Vergangenheit zunehmend an Bedeutung gewonnen. Unterschiedliche Verfahren der Strukturbiologie können dabei helfen, einige der Fragen in diesem Zusammenhang zu beantworten. Die Kernmagnetresonanzspektroskopie (NMR) ist dabei eine Technik, die Informationen über Struktur und Dynamik über Zeitskalen einiger Größenordnungen mit atomarer Auflösung liefert. Üblicherweise kommt dabei die Lösung-NMR Spektroskopie zum Einsatz, bei dir mit Hilfe über Jahre entwickelter und etablierter Methoden eine Vielfalt von Informationen gewonnen werden können. Ein Problem der Lösungs-NMR Spektroskopie ist jedoch, dass Proteine von mehr als 50 kDa häufig nicht untersucht werden können. Allerdings können in der Festkörper-NMR Spektroskopie, durch den Einsatz unterschiedlicher Isotopenmarkierung und Protonierungsgraden, biologisch relevante Systeme mit Molekulargewichten von mehr als 50 kDa im Hinblick auf ihre Struktur und dynamischen Eigenschaften über verschiedene Zeitskalen hinweg untersucht werden. In dieser Arbeit werden durch eine Kombination verschiedener Methoden der Lösungs- und ^{1}H-detektierter Festkörper-NMR Spektroskopie die dynamischen Vorgänge der MAP-Kinase p38a sowie deren Veränderungen bei der Bindung von Liganden untersucht, um so die allosterische Regulierung durch ein dynamisches Netzwerk innerhalb des Proteins zu charakterisieren. Dazu werden verschiedene Methoden zur Isotopenmarkierung mit neuen Ansätzen zur Aufnahme von Relaxationsexperimenten für Proteine mit hohem Molekulargewicht eingesetzt. Darüber hinaus zeigt die Arbeit durch Relaxationsdispersionsexperimente der Festkörper-NMR Spektroskopie und Punktmutationen im RT-Loop, wie die inter- und intramolekularen Wechselwirkungen die dynamischen Eigenschaften der SH3-Domäne von Hühner-a-Spektrin modulieren.

Zusätzlichen wird in dieser Arbeit die Charakterisierung der Sekundärstruktur der N-terminalen Domäne von RhoGDI1 gezeigt, von der angenommen wurde, dass sie in ihrer apo-Form unstrukturiert vorliegt und nur im Komplex mit dem GTPase-Bindungspartner eine geordnete Struktur annimmt. Hier wird durch experimentell gemessene chemische Verschiebungen gezeigt, dass allerdings bereits vor der Formierung des Komplexes Sekundärstrukturelemente vorhanden sind.

Table of Contents

Chapter 1

Introduction

1. INTRODUCTION

1.1 NMR, a powerful tool for Structural Biology

Life on our planet is constructed by four indivisible units of macromolecules: nucleic acid, lipids, carbohydrates, and proteins. These four building blocks participate in the cell's construction, development, and function. A deep understanding of their function is important not only to comprehend better the biological processes in which they are involved but, more importantly, to prevent or treat a broad range of conditions[1].

Proteins are virtually involved in all biological processes in the cell. They can regulate the expression of genetic information, control the growth and differentiation of the cell, catalyze a broad range of reactions, transport ions, provide support and filamentous architecture, and others. Proteins are large-size polymers for which the structural units are the *amino acids*. [1, 2]. The study of the 3D structural characteristics and conformational dynamics of the protein is evaluated by structural biology techniques. These techniques provide information about their assembly, function, and interactions with other molecules. The characterization of protein structure is performed mainly by three techniques: X-ray crystallography, Cryo-electron microscopy (Cryo-EM), and Nuclear Magnetic Resonance spectroscopy (NMR). A complete characterization of proteins includes a detailed study of dynamic processes, as proteins are inherently flexible systems.

Nuclear magnetic resonance spectroscopy is an analytical technique used to determine the molecular structure at the atomic level of a molecule. This method was started in the field of physics, but in later developments and technologies, it became an essential tool for chemistry and in recent years evolved toward more diverse applications in biochemistry and biology. NMR spectroscopy is the method of choice to study both protein structure and dynamics. New advances in fast Magic Angle Spinning (MAS) and proton detection in the solid state have made it possible to get access to bigger and more biologically relevant systems by NMR [2, 3].

1.2 Basic physical concepts in NMR

Larmor frequency and chemical shift [4-6].

To understand the theory of the technique we can look at the quantum mechanics of the nuclear spin angular momentum, as this is the phenomenon that explains nuclear magnetism and NMR spectroscopy.

According to the physical definition, the spin (I) describes the angular momentum (S) of an electron; this is quantized and occupies only discrete energy states with allowed values of S

(1.1)

$$S = \hbar\sqrt{I(I+1)} = \frac{h}{2\pi}\sqrt{\frac{n}{2}\frac{(n+2)}{2}}$$

Where I is the spin quantum number and $\hbar$ is the reduced plank constant. Because the angular momentum is a vector, the spin quantum number has both a magnitude (with values of $I = 0, 1/2, 1, 3/2, \ldots, 6$) and direction (+ or -). Nuclei with even mass numbers and even atomic numbers have spin numbers equal to zero and are inactive in NMR. Nuclei with odd mass numbers have a half-integer spin and nuclei with even mass numbers and odd atomic numbers have integral spin quantum numbers.

For all nuclei with I larger than ½, the charge distribution of protons will be asymmetric, giving them an electric quadrupole moment [4, 7]. All of these nuclei also have a nuclear magnetic moment whose z-component can be defined as

(1.2)

$$\mu_z = \gamma I_z = \gamma\hbar m,$$

in which the gyromagnetic ratio, γ, is a constant characteristic for a given nucleus and m is the magnetic quantum number. In the same way as the angular momentum, the nuclear magnetic moment is also quantized. In the absence of an external field, the spin angular momentum vector does not have a preferred orientation. However, when we apply an external magnetic field B_0 along the $z-axis$, the spins align along or against the magnetic field, and the spin states have energies given by

(1.3)

$$E = -\gamma I_z B_0 = -m\hbar\gamma B_0$$

For spin-half nuclei in an external magnetic field, there will be a splitting of the nuclear spin levels, known as nuclear Zeeman splitting, where we have two levels of energy

(1.4)

$$E_\alpha = -\frac{1}{2}\hbar\gamma B_0 \qquad E_\beta = +\frac{1}{2}\hbar\gamma B_0$$

Where the α state ($m = +1/2$) is described as 'spin up', while the β state (m=-1/2) is described as 'spin down'. Which of the states is considered the lower energy level depends on the gyromagnetic ratio of each nucleus, that is, for ^{1}H and ^{13}C γ is positive and for ^{15}N γ is negative [6]. Thus, the energy difference $\Delta E = \hbar\gamma B_0$ is dependent of the applied B_0.

For describing the behavior of a noninteracting spin-$1/2$ nuclei, the Bloch model can be used. This model describes a simple semiclassical vector model, where the evolution of the bulk magnetic moment, ***M(t)***, is represented as a vector quantity, and in the presence of a magnetic field, a torque applied to the magnetic moment causes a change in the same direction, which eventually makes it precess around the z-axis.

(1.5)

$$\frac{d\boldsymbol{M}(t)}{dt} = \gamma(\boldsymbol{M}(t) \times \boldsymbol{B}(t))$$

When we use the Bloch equation (1.5) in the fixed laboratory frame, ***M(t)*** precesses around ***B(t)*** with a frequency of $\omega = -\gamma B$. For a static field of strength B_0, the precessional frequency is called *Larmor frequency* and is given by

(1.6)

$$\omega_0 = -\gamma B_0$$
$$\nu_0 = -\frac{\gamma B_0}{2\pi}$$

The Larmor frequency is dependent not only on B_0 but also on the local electronic environment since the electrons are magnetic. The effect of the surrounding environment caused by the adjacent nuclei will perturbate the magnetic field, this is known as chemical shift. The effect of the nuclear shielding, σ, results in

(1.7)

$$\Delta E = -\gamma\hbar(1-\sigma)B_0$$

In which $(1-\sigma)B_0$ is the effective field experienced by the nucleus B_{eff}.

In a molecule, the different nuclei will feel a different effect corresponding to their local chemical environment; thus, nuclei with a high electron density will feel a higher shielding effect, which creates a weaker local field, resulting in a reduced Larmor frequency. On the contrary, nuclei with low electron density, i.e., in polar bonds, will have the opposite effect, resulting in larger Larmor frequencies. As the electronic environment generally has low symmetry and the chemical shift depends on the orientation of the molecule with respect to B_0, the chemical shift can be represented in a 3x3 matrix in the principal-axis system

(1.8)

$$\sigma = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix}.$$

After diagonalization, the diagonal represents the average nuclear shielding tensor, σ, which is the isotropic shielding constant for the nucleus that can be represented as

(1.9)

$$\sigma_{iso} = \sigma = \frac{1}{3}\left(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}\right)$$

The chemical shift anisotropy (CSA) is defined as

(1.10)

$$\sigma_{CSA} = \Delta\sigma = \frac{1}{2}\left(\sigma_{xx} - (\sigma_{yy} + \sigma_{zz})\right),$$

and the asymmetry of the tensor is defined as

(1.11)

$$\eta = \frac{3(\sigma_{yy} - \sigma_{zz})}{2\Delta\sigma}.$$

In liquids, where the collisions lead to a rapid reorientation of the molecule, the CSA is average out during the experimental timescale.

Traditionally the chemical shifts are measured in parts per million (ppm or δ) relative to a reference resonance signal from a reference compound, as be calculated from

(1.12)

$$\delta = \frac{\omega - \omega_{ref}}{\omega_0} \times 10^6 ppm.$$

In practice the reference compounds are tetramethyl silane (TMS) or 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as a regular convention.

Quantum mechanics and product operators [4-6]

In a real system, the nuclear spins are not alone and the interaction among each other should be considered to describe the dynamics of the spin systems. In quantum mechanics, the evolution of a system is governed by the Schrödinger equation

(1.13)

$$\frac{\partial \Psi(t)}{\partial t} = -\frac{i}{\hbar}\widehat{\mathcal{H}}\Psi(t)$$

The Schrödinger equations involves the motion of all the nuclei and all the electrons: the Hamiltonian operator, $\widehat{\mathcal{H}}$, involves the interactions between all these particles, and the wavefunction for the system, $\Psi(t)$, contains all the information like the position, velocities, and spin states of all the electrons and nuclei. In NMR only the wave function of a single spin is considered, however the complete wavefunction is defined as

(1.14)

$$\Psi(t) = \sum_{n=1}^{N} C_n \psi_n.$$

In which C_n are complex numbers and may depend upon time. For a single spin system, the superposition state can be then written as

(1.15)

$$\psi = c_\alpha \psi_\alpha + c_\beta \psi_\beta.$$

The solution for the time-dependent Schrödinger equation is an eigenvalue equation for the Hamiltonian operator:

(1.16)

$$E\psi = \widehat{\mathcal{H}}\psi$$

Here the eigenvalues of this equation are the energies of the system and the eigenfunctions are termed the stationary states of the system. The Hamiltonian operator for a single spin system can be defined by equation (1.3) and the expectation values for the angular momentum of the magnetization are defined as

(1.17)

$$\langle I_z \rangle = \frac{1}{2} c_\alpha^* c_\alpha - \frac{1}{2} c_\beta^* c_\beta$$

$$\langle I_x \rangle = \frac{1}{2} c_\alpha^* c_\beta + \frac{1}{2} c_\beta^* c_\alpha$$

$$\langle I_y \rangle = \frac{1}{2} i c_\beta^* C_\alpha - \frac{1}{2} i c_\alpha^* c_\beta^*.$$

The z-component of the angular momentum is represented by the operator I_z, the expectation value corresponds to the average of all the z-components measured from all the spins in a real sample, where the total z-magnetization is the sum of the z-components multiplied by γ.

(1.18)

$$M_z = \frac{1}{2}\gamma N\overline{\left(c_\alpha^* c_\alpha - c_\beta^* c_\beta\right)} = \gamma N\overline{\langle I_z\rangle}.$$

The same way, the x- and y- magnetization can be given by

(1.19)

$$M_x = \frac{1}{2}\gamma N\overline{\left(c_\alpha^* c_\beta + c_\beta^* c_\alpha\right)} = \gamma N\overline{\langle I_x\rangle}$$

$$M_y = \frac{1}{2}i\gamma N\overline{\left(c_\beta^* c_\alpha - c_\alpha^* c_\beta\right)} = \gamma N\overline{\langle I_y\rangle}.$$

Considering the ensemble average is found by adding up the contributions from all spins and dividing by N, we can interpret the population of a level as

(1.20)

$$n_\alpha = N c_\alpha c_\alpha^* \qquad n_\beta = N c_\beta c_\beta^*$$

At equilibrium, these populations are predicted by the *Boltzmann distribution*:

(1.21)

$$n_{\alpha,eq} = \frac{1}{2}Nexp\left(-\frac{E_\alpha}{k_B T}\right) \qquad n_{\beta,eq} = \frac{1}{2}Nexp\left(-\frac{E_\beta}{k_B T}\right)$$

Where E_α and E_β are the energies of the two levels in Joule, k_B is the Boltzmann constant, T is the temperature, and N is the total number of spins. For spins with positive γ, the α state has the lower energy and the Boltzmann distribution predicts $n_{\alpha,eq} > n_{\beta,eq}$.

As time evolution is of great importance in NMR, knowing the variations of the Hamiltonian for the time-dependent Schrödinger equation is necessary for determining the time-evolution of the wavefunction. For a single spin, the Hamiltonian for free evolution in the rotating frame is

(1.22)

$$\widehat{\mathcal{H}} = \Omega\hat{\mathrm{I}}_\mathrm{z}$$

We can solve how the superposition state evolves over time under the free precession Hamiltonian by solving the time-dependent Schrödinger equation. The results show that the coefficients vary according to

(1.23)

$$c_\alpha(t) = c_\alpha(0)\exp\left(-\frac{1}{2}i\Omega t\right) \qquad c_\beta(t) = c_\beta(0)\exp\left(\frac{1}{2}i\Omega t\right)$$

In which $c_\alpha(0)$ and $c_\beta(0)$ are the coefficients at time zero. Each coefficient oscillates in phase at a frequency which depends on the correspondent eigenvalue of the Hamiltonian

(1.24)

$$\Omega\hat{I}_z|\alpha\rangle = \frac{1}{2}\Omega\ |\alpha\rangle \qquad \Omega\hat{I}_z|\beta\ \rangle = -\frac{1}{2}\Omega\ \ |\beta\rangle$$

When we apply a radio frequency pulse (RF) along the xy-plane we rotate the magnetization away from the equilibrium in the z-component. Thus, the Hamiltonian in the rotating frame for an on-resonance or strong pulse about the x-axis is

(1.25)

$$\widehat{\mathcal{H}}_{pulse} = \omega_1\hat{I}_x$$

After solving the time-dependent Schrödinger equation we obtain the differential equations for the coefficients

(1.26)

$$\frac{dc_\alpha(t)}{dt} = -\frac{1}{2}i\omega_1 c_\beta(t) \qquad \frac{dc_\beta(t)}{dt} = -\frac{1}{2}i\omega_1 c_\alpha(t)$$

And the solutions for the equations are

(1.27)

$$c_\alpha(t) = \cos\left(\tfrac{1}{2}\omega_1 t\right)c_\alpha(0) - i\ sin\left(\tfrac{1}{2}\omega_1 t\right)c_\beta(0)$$

$$c_\beta(t) = \cos\left(\frac{1}{2}\omega_1 t\right)c_\beta(0) - i\sin\left(\frac{1}{2}\omega_1 t\right)c_\alpha(0)$$

As the value of c_α at time t depends on both $c_\alpha$and $c_\beta$at time zero, thus the Hamiltonian is causing a mixing between the states $|\alpha\rangle$ and $|\beta\rangle$, which causes an oscillatory effect of the coefficients.

The expectation values of the x-, y-, and z-components of the angular momentum at time zero can be represented as

(1.28)

$$\langle I_x\rangle(t) = \langle I_x\rangle(0)$$

$$\langle I_y\rangle(t) = \langle I_y\rangle(0)\cos(\omega_1 t) - \langle I_z\rangle(0)\sin(\omega_1 t)$$

$$\langle I_z\rangle(t) = \langle I_z\rangle(0)\cos(\omega_1 t) + \langle I_y\rangle(0)\sin\ (\omega_1 t)$$

As we have applied a pulse in about the x-axis, we can see that the x-component is unaffected and does not change. In the case of the y-component, we expect an oscillatory interchange of the y- and z- components. Thus, when we apply a 90° pulse ($\pi/2$ radians) about the x-axis we will see a flip of the magnetization from the z-component to the y-component. As for the z-component, we will see the vector rotating in the yz-plane at frequency ω_1.

A simplified way to analyze the ensemble average is the used of the density operator, or density matrix, which is defined as

(1.29)

$$\hat{\sigma} = \overline{|\psi\rangle\langle\psi|}$$

And in which the superposition states, $|\psi\rangle$ and $\langle\psi|$, are given by $|\psi\rangle = c_\alpha|\alpha\rangle + c_\beta|\beta\rangle$ and $\langle\psi| = c_\alpha^*\langle\alpha| + c_\beta^*\langle\beta|$, respectively. The matrix representation of the density operator is

(1.30)

$$\sigma = \begin{pmatrix} \overline{c_\alpha c_\alpha^*} & \overline{c_\alpha c_\beta^*} \\ \overline{c_\beta c_\alpha^*} & \overline{c_\beta c_\beta^*} \end{pmatrix}$$

The *Liouville-von Neumann* equation describes the time-evolution of the density operator in the laboratory frame:

(1.31)

$$\frac{d\hat{\sigma}(t)}{dt} = -i\left(\widehat{\mathcal{H}}\hat{\sigma}(t) - \hat{\sigma}(t)\widehat{\mathcal{H}}\right),$$

for which, the solution is

(1.32)

$$\hat{\sigma}(t) = \exp\left(-i\widehat{\mathcal{H}}t\right)\hat{\sigma}(0)\exp\left(i\widehat{\mathcal{H}}t\right)$$

Where $\hat{\sigma}(t)$ is the density operator at time t and $\hat{\sigma}(0)$ is the density operator at time zero.

Because NMR is a technique sensitive to the chemical environment of the studied nuclei, we can observe different type of interaction, which we can divide into spin-field and spin-spin interactions. The spin-field interactions include the Zeeman interaction, magnetic susceptibility, chemical shift, RF fields. The spin-spin interactions refer to the isotropic scalar coupling (*J*-coupling), dipolar coupling and quadrupolar coupling [8]. All these interactions can be described by their respective Hamiltonian and will describe the full NMR Hamiltonian

(1.33)

$$\widehat{\mathcal{H}} = \widehat{\mathcal{H}}_{Zeeman} + \widehat{\mathcal{H}}_J + \widehat{\mathcal{H}}_{cs} + \widehat{\mathcal{H}}_{DD} + \widehat{\mathcal{H}}_Q$$

Dipolar coupling: The direct dipole-dipole coupling between two spins, is an interaction originating from the fields between the nuclear spins propagated through space, without involving the electron clouds. These interactions can be inter- or intramolecular, and the Hamiltonian can be represented as

(1.34)

$$\hat{\mathcal{H}}_{jk}^{DD,full} = b_{jk}\left(3\left(\hat{I}_j \cdot e_{jk}\right)\left(\hat{I}_k \cdot e_{jk}\right) - \hat{I}_j \cdot \hat{I}_k\right),$$

where $e_{jk}$is a unit vector parallel to the line joining the centers of the two nuclei involved, which is defined to be $e_{jk} \cdot e_{jk} = 1$, and $b_{jk}$is the dipole-dipole coupling constant, which describes the magnitude of the through-space interaction and is given by

(1.35)

$$b_{jk} = -\frac{\mu_0}{4\pi}\frac{\gamma_j \gamma_j \hbar}{r_{jk}^3},$$

where γ_j and γ_k are the gyromagnetic ratios of the two spins and r_{jk} is the spin-spin distance between the two spins.

The dipole-dipole spin Hamiltonian is orientation dependent since the value of the vector e_{jk}changes direction as the molecule rotates. As the dipole-dipole coupling depends only on known physical constants and the inverse cube of the distance r_{jk}^{-3}, it can be used for structural studies. Experiments that make use of this property are the NOESY and ROESY experiments.

Scalar coupling: The J-coupling is mediated by chemical bonds, two spins will have a measurable J-coupling when they are linked together through a small number of chemical bonds, including hydrogen bonds. The full Hamiltonian of the intramolecular J-coupling is defined by

(1.36)

$$\hat{\mathcal{H}}_{jk}^{J,full} = 2\pi \hat{I}_j \cdot J_{jk} \cdot \hat{I}_k$$

Where J_{jk} is the J-coupling tensor, which is a 3 x 3 real matrix, and depends on the molecular orientation. In isotropic liquids the J-coupling tensor is averaged out by the molecular tumbling and is equal to the average of the diagonal elements. Here we can distinguish two cases, the homonuclear and heteronuclear case, dependent on the type of nuclei involved in this interaction.

It is known that the intensity of the signal depends on the gyromagnetic ratio, γ, and the field B_0 with the dependence

(1.37)

$$\frac{signal}{noise} \propto |\gamma|^{\frac{5}{2}}(B_0)^{\frac{3}{2}}.$$

We can think about two of the important nuclei in protein studies: ^{1}H and ^{15}N. The ^{1}H nuclei will provide a signal-to-noise ratio ~300 time larger than an equal number of ^{15}N nuclei at the same magnetic field. Thus, for the ^{15}N case it will take longer to acquire a spectrum with the same signal-to-noise ratio compared to the ^{1}H spectrum. There are strategies that can be used for enhancing the signal when working with nuclei with small γ. We can transfer the polarization from the higher magnetic ratio proton spins to the weakly magnetic ^{15}N nuclei. One of the most popular pulse sequences for this purpose is the Insensitive Nuclei Enhanced by Polarization Transfer (INEPT) [5, 6].

The INEPT can be used as a building block for other pulse sequences. For studies with proteins in solution NMR it is common to use pulse programs that include the INEPT blocks and thus make use of the high γ of ^{1}H. One of the basic experiments used in small molecules and proteins is the HSQC (heteronuclear single-quantum correlation), which allows us to visualize the correlation between ^{1}H and 15N, i.e., amide group of amino acids, or ^{1}H-^{13}C, i.e., methyl groups of certain amino acids [6, 9].

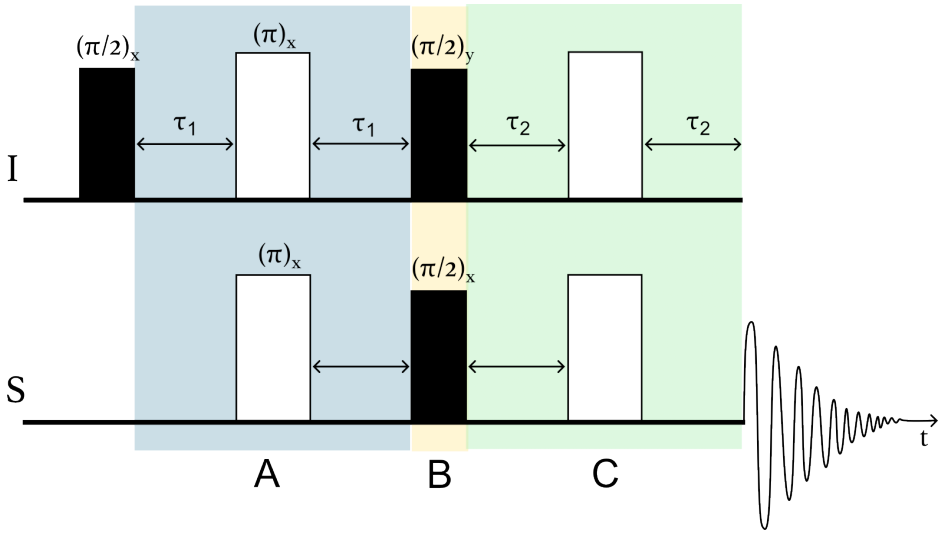


Figure 1.1. Pulse sequence for refocused INEPT. The INEPT starts with a 90°$_x$ pulse on the spin I, which generates an anti-phase state on the I spin during A, and then transferring the magnetization to the S spin with the second 90°$_y$ pulse in B. During the second evolution period (C), the anti-phase state evolves into an in-phase state, with an optimum value for both delays τ_1 and τ_2 being $1/(4J_{IS})$.

For large proteins, we encounter additional problems for obtaining good-quality spectra when working with proteins in the solution state. For the amide ^{15}N-^{1}H pairs, the effect of cross-correlation between the CSA and dipolar relaxation gives origin to two lines of a doublet with different linewidths. In the conventional HSQC experiment these two lines are collapsed by a 180° pulse on spin *I* and by decoupling the spin *S*.This results in a line for which the width is the average of the two lines of the doublet, if one of the lines is much broader, we will observe a decrease of the peak height. For those cases where the cross-correlation effects are significant, it is best not to remove the splitting and rather selectively 'discard' the unwanted terms to obtain a higher signal-to-noise ratio because of the height of the sharp line. This method is called Transverse Relaxation Optimized Spectroscopy (TROSY), which, like INEPT, can be introduced for many pulse programs [10-13].

1.3 Solid-state NMR and ^{1}H-detection

Traditional solid-state NMR [14-16]

The NMR interactions discussed above include the isotropic and anisotropic interactions involved in the NMR signals. As described the anisotropic interactions are orientation dependent, including the scalar coupling, dipolar coupling, quadrupolar coupling, and chemical shift anisotropy (CSA). In isotropic liquids these interactions are averaged out by fast exchange on the molecule orientation as a result of the molecular tumbling intrinsically present in solution. Due to the nature of the solid samples, the anisotropic interactions are no longer averaged out, as they take part in a geometric lattice that prevents the fast change of orientation. The largest anisotropic contribution in the solid state is the dipolar coupling, which the Hamiltonian is given by

(1.38)

$$\widehat{\mathcal{H}}_{dd} = -\left(\frac{\mu_0}{4\pi}\right)\gamma_I\gamma_S\hbar\left(\frac{\hat{I}\cdot\hat{S}}{r^3} - 3\frac{(\hat{I}\cdot r)(\hat{S}\cdot r)}{r^5}\right).$$

There are two possible cases for the dipolar coupling: homonuclear and heteronuclear, depending on the species involved in the interaction. The homonuclear dipolar coupling is given by

(1.39)

$$\widehat{\mathcal{H}}_{dd}^{homo} = -d\cdot\frac{1}{2}(3\cos^2\theta - 1)\left[3\hat{I}_z\hat{S}_z - \hat{I}\cdot\hat{S}\right],$$

where d is the dipolar coupling constant

(1.40)

$$d = \hbar \left(\frac{\mu_0}{4\pi}\right) \frac{1}{r^3} \gamma_I \gamma_S.$$

And the heteronuclear dipolar coupling is defined as

(1.41)

$$\widehat{\mathcal{H}}_{dd}^{hetero} = -d(3\cos^2\theta - 1)\,\hat{I}_z \hat{S}_z.$$

The presence of homonuclear dipolar coupling in a multistate system will change the linewidth in the NMR spectra, giving origin to a broader line than the one obtained in solution. While the heteronuclear dipolar coupling will cause the characteristic powder pattern on which the split corresponds to the value of the coupling constant.

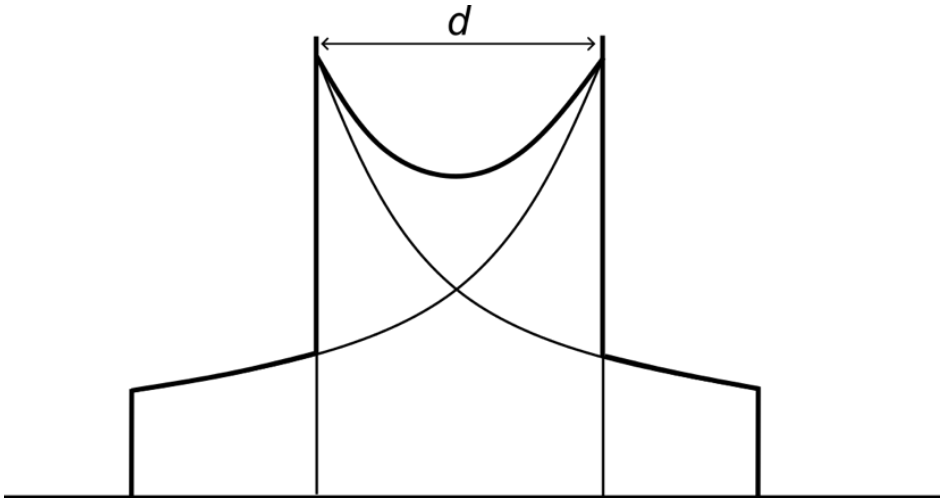


Figure 1.2. Representation of the powder pattern, the shape of the powder line for the *I* spin in a two-spin heteronuclear system. [14]

Magic-angle spinning [14, 15]

It is possible to artificially average out the anisotropic interactions in the solid state by introducing Magic-angle spinning. Due to the molecular orientation dependence of the anisotropic interactions to the form $3\cos^2\theta - 1$, where θ is the angle which describes the orientation of the spin interaction tensor. If we spin the samples at an angle $\theta_R$with respect to the B_0 field, then θ varies with time as the molecule rotates and the average orientation of the nuclear spin interaction is

(1.42)

$$\langle 3\cos^2\theta - 1\rangle = \frac{1}{2}(3\cos^2\theta_R - 1)(3\cos^2\beta - 1),$$

in which β is the angle between the principal z-axis of the shielding tensor and the spinning axis, θ_R is the angle between the applied field and the spinning axis, and θ is the angle between the principal z-axis of the interaction tensor and the applied field B_0. If θ_R is set to 54.7°, then $(3\cos^2\theta_R - 1) = 0$, and so is the average. Provided the spinning rate is fast so that θ is averaged rapidly, the anisotropic interaction is averaged to zero as well.

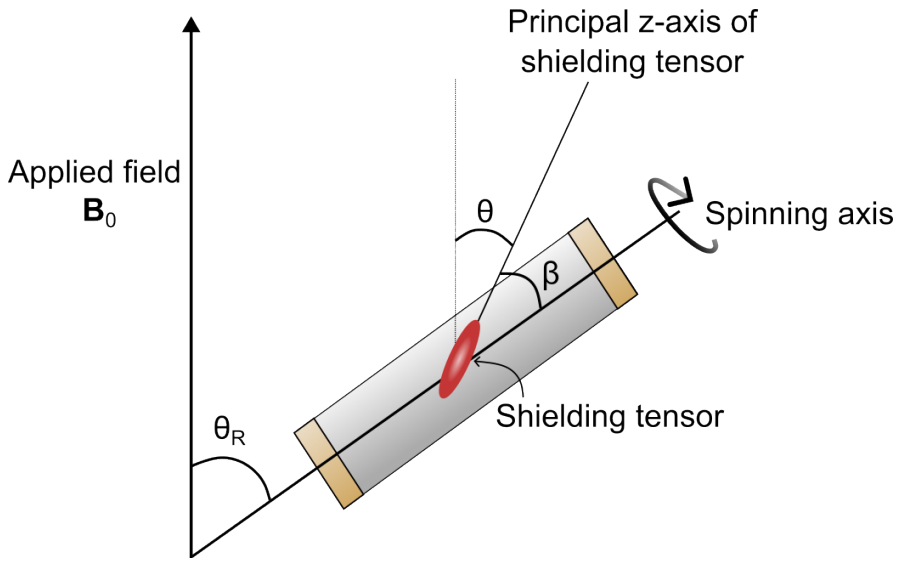


Figure 1.3. Magic-angle spinning, the rotor is oriented along an axis tilted by 54.7° away from the z-axis.

In the same way that we can enhance the sensitivity of our experiment using INEPT, in the solid state we can also make use of the high γ of the ^{1}H and transfer the initial polarization to other nuclei using the cross-polarization technique. In solids, the cross-polarization (CP transfer) is achieved by using rf irradiation to induce heteronuclear mutual spin flips. The abundant spins with a high gyromagnetic ratio have a larger population number than the rare spins with a low gyromagnetic ratio. This process will cause the polarization of the rare spins to be equal to that of the abundant spins. There are two common experimental schemes for cross-polarization, involving either adiabatic demagnetization in the rotating frame (ADRF) of the abundant spin reservoir or by a simultaneous spin locking of the two species, as proposed by Hartmann and Hahn [17].

Hartmann-Hahn cross-polarization is achieved when for a system with two nuclear spins, I and S, it is simultaneously irradiated with two rf fields B_{1I} and B_{1S} of frequencies ω_I and ω_S close to the Larmor frequencies ω_{0I} and ω_{0S}. The Hartmann-Hahn condition is met when the

rf power of the two nuclear species are adjusted, and the nutation frequencies ω_{1I} and ω_{1S} around the two fields are such that [15, 18]

(1.43)

$$\omega_{1I} = \gamma_I B_{1I} = \gamma_S B_{1S} = \omega_{1S}$$

This is that the length of a $\pi/2$ pulse is the same for both spin systems.

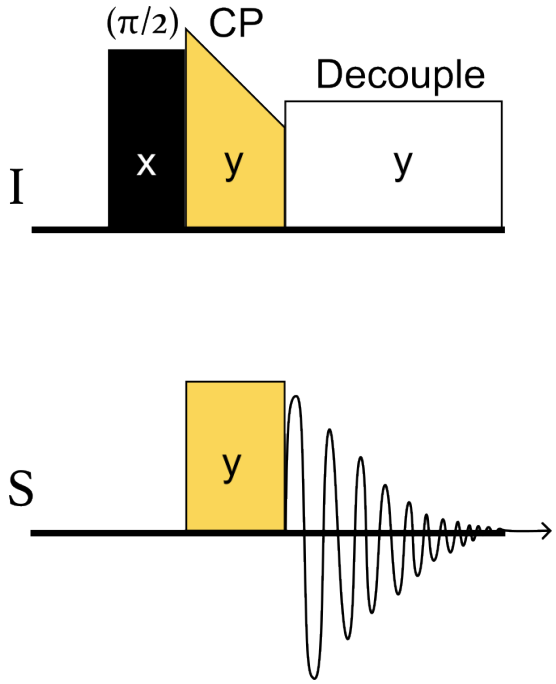


Figure 1.4. Basic Hartmann-Hahn cross polarization pulse sequence.

For the Hartmann-Hahn cross-polarization, it is necessary to consider the effect of dipolar coupling between the two spins. For this, the dipolar coupling operator can be described for a heteronuclear dipolar interaction as

(1.44)

$$\hat{H}_{HX} = -\sum_i d_i \cdot (3\cos^2\theta_i - 1) \cdot \hat{I}_{iz}^H \hat{S}_z^X,$$

were d_i is the dipolar-coupling constant for the coupling between the ith ^{1}H and the X spin. This operator remains unaltered by rotations about z, then both spins are quantized in their respective rotating frame, and thus the dipole-dipole coupling operator cannot affect the net energy of the spin system in the rotating frame, as the spin system is determined by fields parallel to the spin. Similarly, the dipole-dipole coupling operator cannot alter the net spin polarization parallel to the quantization axis. So, the dipole-dipole operator couples the ^{1}H and X spin and acts in a way to conserve both the energy and the angular momentum, which in the Hartmann-Hahn condition the energy gap between the spin states in the rotating frame are equal

and the transition absorbing energy can be compensated by a transition releasing energy on X spin.

When MAS is used, there is a time dependency is introduced into the dipolar coupling. The strength of an interaction under MAS is described by terms which oscillate at $\pm\omega_R$ and $\pm 2\omega_R$, where ω_R is the spinning rate. The Hamiltonian in the toggling frame is described by

(1.45)

$$\widehat{H}^*(t) = \widehat{H}^*_{HH} + \widehat{H}^*_{HX}$$

In which

$$\widehat{H}^*_{HH} = -\frac{1}{2}\sum_{i>j} C_{ij}^{HH}\left(\hat{I}_i^H \hat{I}_j^H - 3\hat{I}_{iX}^H \hat{I}_{jX}^H\right)$$

$$\widehat{H}^*_{HX} = -\sum_i C_i^{HX}\left[\left(\sum_i \hat{I}_{iz}^H S_z^X + \sum_i \hat{I}_{iy}^H S_y^X\right)\cos(\omega_1^H - \omega_1^X)\,t + \left(\sum_i \hat{I}_{iZ}^H \hat{S}_y^H + \sum_i \hat{I}_{iy}^H \hat{S}_z^X\right)\sin(\omega_1^H - \omega_1^X)\,t\right]$$

$$C_{ij}^{HH} = \left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_H^2}{r_{ij}^3}\,\frac{1}{2}\,(3\cos^2\theta_{ij} - 1)$$

$$C_i^{HX} = \left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_{1H}\gamma_X}{r_i^3}\frac{1}{2}(3\cos^2\theta_i - 1)$$

The isotropic term of the dipolar coupling does not remain under MAS, and the term that induces cross-polarization between the ^{1}H and X spins is $\widehat{H}^*_{HX}$. It is important to notice the C_i^{HX} terms are no longer constant, they oscillate at $\pm\omega_R$ and $\pm 2\omega_R$, therefore the normal Hartmann-Hahn match $\omega_1^H \cong \omega_1^X$. If the rotor period is small, with relation to the contact time, the net cross-polarization term in the Hamiltonian averages to zero. As the $\widehat{H}^*_{HX}$ contains other oscillatory terms, $\cos(\omega_1^H - \omega_1^X)\,t$ and $\sin(\omega_1^H - \omega_1^X)\,t$, the oscillatory behavior under MAS can be cancelled by matching $(\omega_1^H - \omega_1^X)$ to $\pm\omega_R$ or $\pm 2\omega_R$, this is known as the *sideband match conditions*. This effect decreases as the spinning rate increases and the dipolar coupling is more effectively averaged.

The effect of modulating the cross-polarization rate by the spinning frequency ($\omega_1^H = \omega_1^X \pm n\omega_R$) is defined as a rotary resonance phenomenon and similar phenomenon can also be observed under other conditions. Taking into account Equation (43) the rotary resonance condition can be written as [14, 19, 20]:

(1.46)

$$n\omega_R = -\gamma B_{eff}.$$

As the anisotropic interactions are reduced by MAS, some of the useful parts of this interaction are lost, such as dipolar coupling, which provides information about the dynamics and structure of the analyzed molecule. Recoupling experiments can be used to access the information about dipolar coupling, thus measuring distances, torsion angles and as transfer magnetization blocks [21].

Protons in NMR: ^{1}H-detection

Traditionally, solid-state NMR studies are performed using ^{13}C detected techniques, although this is an insensitive nucleus. Recent technological advances in smaller rotors (≤1.3mm outer diameter) that allow for faster spinning (>50 kHz), in combination with the implementation of techniques to dilute the proton concentration in the protein by extensive deuteration, have allowed the development of methodologies for proton detection [3, 22-24]. Some of advantages of the ^{1}H detection are the narrow proton linewidths and the reduction of the effective proton-proton dipolar coupling, which translate into the reduction of the rf energy irradiated to the sample because there is no need for a high-power proton decoupling. Additionally, the number of scans per time can be increased to the point that it is limited by the proton T_1 time, which can also decrease after paramagnetic doping [3, 25].

The experiments recorded using ^{1}H detection have a strong resemblance to the experiments recorded in solution NMR, where it is possible to use techniques such as the INEPT or CP transfer. There are established methods to record assignment experiments in multiple dimensions. Experiments in 3D such as hCANH, hCONH, hcaCBcaNH, hCAcoNH, hCOcaNH, and others, can be recorded by transferring the magnetization from ^{1}H to carbons and back via ^{15}N in an out-and-back fashion, and the obtained spectra can be used similarly to the solution assignment experiments [3]. 4D experiments have been developed in recent years to overcome problems when working with proteins of high molecular size, where the level of signal overlap represents an important problem for protein assignment [26, 27]. By including additional dimensions and combining the chemical shift information of the backbone and sidechain nuclei it is possible to reduce overlap, and in combination with ultrafast MAS,

paramagnetic doping, and perdeuteration it is possible to accelerate the process of protein assignment, which is regularly the bottleneck for NMR studies [27, 28].

1.4 Relaxation

The thermodynamical laws establish that when a system is perturbed from the equilibrium state by an external, eventually will come back to the equilibrium state. The same principle applies to the phenomenon observed by NMR. For the bulk z-magnetization there is an energetic preference for the z-component of the magnetic moment to be aligned along the B_0 field direction. When an RF pulse is applied to the equilibrium magnetization, the z-magnetization is rotated towards the transverse plane and the z-component is reduced in size. After the pulse the resulting transverse magnetization precesses about the field direction. The process of the z-magnetization returning to equilibrium is called *longitudinal relaxation* (T_1) and the process by which transverse magnetization decays away to its equilibrium value of zero is referred as *transverse relaxation* (T_2) [6].

Relaxation theory [8, 29-31]

The longitudinal and transverse relaxation processes can be described using the Bloch equations (1.5). Here, it is assumed that the longitudinal relaxation is characterized by the first-order rate expression.

$$\frac{dM_z}{dt} = R_1[M_0 - M_z(t)] \Rightarrow M_z(t) = M_0 - [M_0 - M_z(0)]\exp(-R_1 t), \quad (1.47)$$

where R_1 is the longitudinal relaxation rate constant ($R_1 = \frac{1}{T_1}$).

The transverse relaxation is also characterized by a first-order rate expression

$$\frac{dM_x(t)}{dt} = -R_2 M_x(t) \Rightarrow M_x(t) = M_x(0)\exp(-R_2 t) \quad (1.48)$$

$$\frac{dM_y(t)}{dt} = -R_2 M_y(t) \Rightarrow M_y(t) = M_y(0)\exp(-R_2 t),$$

in which R_2 is the transverse relaxation rate constant ($R_2 = \frac{1}{T_2}$) and $M_x(0)$ and $M_y(0)$ are the values of the transverse magnetization at $t = 0$.

The semiclassical derivation of the master equation of relaxation theory formulated by Wangsness, Bloch, and Redfield [32-34] gives the full Hamiltonian system in the laboratory frame:

(1.49)

$$\mathcal{H}(t) = \mathcal{H}_0(t) + \mathcal{H}_1(t),$$

where $\mathcal{H}_0(t)$ is the deterministic Hamiltonian and the $\mathcal{H}_1(t)$ is the stochastic Hamiltonian. The deterministic Hamiltonian consists of a static part that describes the Zeeman and isotropic interactions, and a time-dependent part that represents the radio frequency irradiation of the spins. The stochastic Hamiltonian contains the time-dependent interactions originated by stochastic processes and are the source of relaxation phenomena. The description of the time evolution of the density operator in the laboratory frame is given by the Liouville-von Neumann equation (1.30) and after the mathematical treatment we obtain

(1.50)

$$\frac{d}{dt}\overline{\tilde{\sigma}(t)} = -\int_0^t \overline{\left[\widetilde{\mathcal{H}}_1(t), \left[\widetilde{\mathcal{H}}_1(t-\tau), \tilde{\sigma}(t-\tau) - \sigma_{eq}\right]\right]} d\tau,$$

where the bar represents the ensemble average. When we assume that the time scale of the random fluctuations described by τ is much shorter than the variations in $\bar{\bar{\sigma}}$ which is on the time scale of t. This is known as the 'weak collision' limit or Readfield limit. In solution NMR, the 'weak collision' limit is typically fulfilled because the molecular tumbling is shorter than the relaxation processes. However, in solids, as the molecular tumbling is absent, the validity of this assumption must be evaluated individually.

When $\tilde{\sigma}(t-\tau)$ is replaced by $\tilde{\sigma}(t)$ and the upper integration boundary is set to infinity, since $\tau \ll t$, we obtained the master equation for relaxation in the interaction-frame

(1.51)

$$\frac{d}{dt}\tilde{\sigma}(t) = -\int_0^\infty \overline{\left[\widetilde{\mathcal{H}}_1(t), \left[\widetilde{\mathcal{H}}_1(t-\tau), \tilde{\sigma}(t) - \sigma_{eq}\right]\right]} d\tau$$

We can extract the information of the correlation time, τ_c, which is the approximate average time for the molecule to rotate by 1 radian. The correlation time depends primarily on the molecular size and shape, as well as the solvent viscosity and temperature. We can define the correlation function as

(1.52)

$$G(t) = \frac{1}{5}\exp\left(-\frac{t}{\tau_c}\right).$$

The corresponding spectral density function is

(1.53)

$$J(\omega) = \frac{2}{5}\frac{\tau_c}{(1+\omega^2\tau_c^2)}.$$

Due to the dependence of the spectral density function on the Larmor frequency, when comparing the Larmor frequency and the correlation time we find a maximum in $J(\omega_0)$ occurs when $\tau_c = 1/\omega_0$. From this comparison we can define two motion regimes: the fast motion and ($\omega_0\tau_c \ll 1$) and slow motions ($\omega_0\tau_c \gg 1$). The fast-motion regime becomes

(1.54)

$$J(\omega_0) = \frac{2}{5}\tau_c = J(0),$$

and the slow regime is given by

(1.55)

$$J(\omega_0) = \frac{2}{5}\frac{\tau_c}{\omega^2\tau_c^2}.$$

Fast dynamics and Spectral Density Mapping

For studying the fast timescale of motion in solution we can use the longitudinal (R_1), transverse (R_2), and the dipole-dipole cross-relaxation (σ_{IS}) rate constants. The rate constants are given by

(1.56)

$$R_1 = \left(\frac{d^2}{4}\right)[J(\omega_I - \omega_S) + 3J(\omega_s) + 6J(\omega_I + \omega_S] + c^2J(\omega_S)$$

(1.57)

$$R_2 = \left(\frac{d^2}{8}\right)[4J(0) + J(\omega_I - \omega_S) + 3J(\omega_S) + 6J(\omega_I) + 6J(\omega_I + \omega_S)]\left(\frac{c^2}{6}\right)[4J(0) + 3J(\omega_S)]$$

(1.58)

$$\sigma_{IS} = \left(\frac{d^2}{4}\right)[6J(\omega_I + \omega_S) - J(\omega_I - \omega_S)],$$

in which $d = \left(\frac{\mu_0 h\gamma_I\gamma_S}{8\pi^2}\right)\langle r_{IS}^{-3}\rangle$ and $c = \frac{\Delta\sigma\omega_S}{\sqrt{3}}$; μ_0 is the permeability of free space.

The relaxation constants are linear combinations of $J(\omega)$ sampled at the eigenfrequencies of the spin systems. Thus, the values of the spectral density function can be obtained from the experimental relaxation data by different mathematical methods. For the backbone amide ^{15}N nuclei we can use the reduced spectral density mapping, for which information can be extracted from the experimental values of T_1 and T_2 and heteronuclear steady-state experiments. These values are commonly analyzed by fitting a simple 'model-free' functional form for $J(\omega)$ that contains a limited number of free parameters to the relaxation or spectral density data. By using the Lipari and Szabo formulation

(1.59)

$$J(\omega) = \frac{S^2\tau_m}{1+\omega^2\tau_m^2} + \frac{(1-S^2)\tau}{1+\omega^2\tau^2},$$

in which S^2 is the order parameter, τ_m is the isotropic overall rotational correlation time of the protein, $\tau_e$is the effective internal correlation time, and $\tau = (1/\tau_m + 1/\tau_e)^{-1}$.

The generalized order parameter, S^2, is a measure of the equilibrium distribution of the orientations of the vector, $\mu(t)$, in a molecular reference frame. Completely restricted motion will have $S^2 = 1$, and completely unrestricted motion have a $S^2 = 0$. Changes on the S^2 after conformational transitions or upon ligan binding in proteins, reflect altered flexibility and consequent changes in conformational entropy.

Slow timescale dynamics [29, 35]

Dynamical processes that include conformational changes, ligand binding, protein folding, and other biological processes occur on the microsecond to millisecond time scale. The chemical or conformational exchange described by the reaction scheme

(1.60)

$$A_1 \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} A_2, \qquad A_1 \underset{k_{31}}{\overset{k_{13}}{\rightleftharpoons}} A_3, \qquad \dots \; A_{n-1} \underset{k_{n(n-1)}}{\overset{k_{(n-1)n}}{\rightleftharpoons}} A_n$$

In which k_{jk} is the reaction rate constant. The chemical kinetic rate laws are written in matrix form, and for a couple set of first-order chemical reactions between N species, the equation is generalized to

(1.61)

$$\frac{d\mathrm{A}(t)}{dt} = \mathrm{KA}(t)$$

For which the matrix elements of the rate matrix, K, are given by

(1.62)

$$K_{jk} = k_{jk} \qquad (k \neq j),$$

$$K_{kk} = -\sum_{\substack{j=1 \\ j\neq k}}^{N} k_{kj}$$

The chemical exchange can be studied by solution and solid-state NMR. In solution NMR, it is common to quantify the chemical reaction kinetics using longitudinal magnetization exchange, line shape analysis, R_2CPMG relaxation dispersion and $R_{1\rho}$ relaxation dispersion [29, 35]

^{15}N CPMG relaxation dispersion[29]

The ^{15}N CPMG TROSY experiment analysis allows us to know which residues are involved in chemical exchange, in particular, to study the conformational exchange and the structure of 'invisible' minor states on the microsecond to millisecond time scale. The two-site exchange model of dynamics process can be used to describe the populations (p_a and p_b), the exchange rate (K_{ex}), and the difference in chemical shift difference (Δω) between the two sites. In a CPMG experiment, relaxation of transfer magnetization is observed during a $(\tau_{cp}/2 - 180° - \tau_{cp} - 180° - \tau_{cp}/2)_n$ spin-echo sequence (**Figure 1.5**)

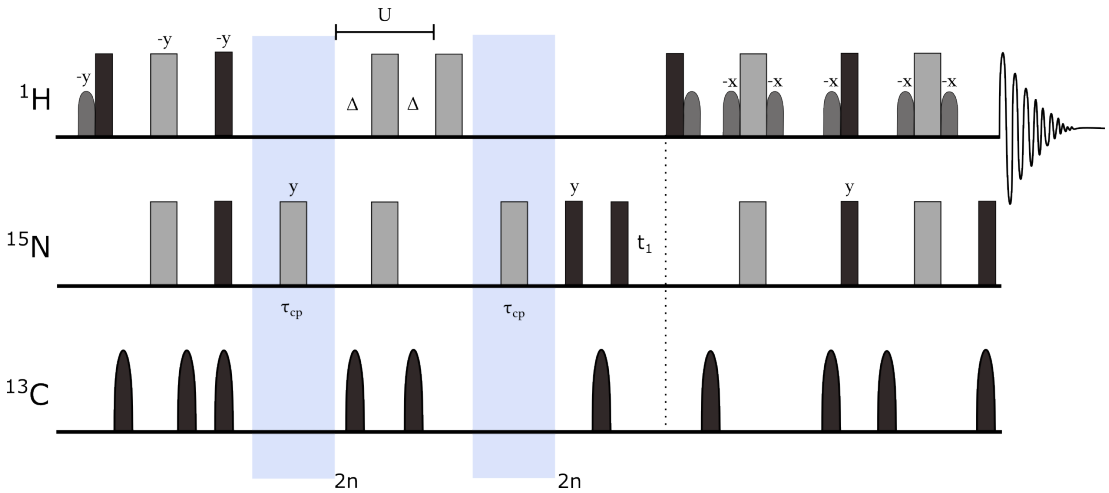


Figure 1.5. Pulse sequence for the ^{15}N TROSY-CPMG experiment. Narrow black bars represent 90° pulse, gray wide bars represent 180° pulses, short dark gray curved bars are water-selective 90° pulses. The delays are $\tau = \tau_{cp}/2$ and $\Delta = 1/(4J_{IS})$. The relaxation period is $T = 4n\tau_{cp}$.

CPMG experiments are sensitive to the chemical exchange process if values of $1/\tau_{cp}$ near k_{ex} are achieve experimentally. For the case of two-site exchange, a general expression for the transverse relaxation rate constant for site 1 ($p_1 > p_2$), $R_2(1\tau_{cp})$,

(1.63)

$$R_2\left(\frac{1}{\tau_{cp}}\right) = \bar{R}_2^0 + \frac{1}{2}\left(k_{ex} - \frac{1}{\tau_{cp}}\cosh^{-1}[D_+\cosh(\eta_+) - D_-\cos(\eta_-)]\right),$$

in which

$$D_\pm = \frac{1}{2}\left[\pm 1 + \frac{\psi + 2\Delta\omega^2}{(\psi^2 + \zeta^2)^{\frac{1}{2}}}\right]^{\frac{1}{2}}$$

$$\eta_\pm = \frac{\tau_{cp}}{2}\left[\pm\psi + (\psi^2 + \zeta^2)^{\frac{1}{2}}\right]^{1/2}$$

$$\psi = k_{ex}^2 + \Delta\omega^2 \text{ and } \zeta = -2\Delta k_{ex}(p_1 - p_2).$$

For the fast exchange limit is defined as

(1.64)

$$R_2\left(\frac{1}{\tau_{cp}}\right) = \bar{R}_2^0 + \frac{p_1 p_2 \Delta\omega^2}{k_{ex}}\left(1 - \frac{2tanh\left[\frac{k_{ex}\tau_{cp}}{2}\right]}{k_{ex}\tau_{cp}}\right).$$

There are two types of acquisition for CPMG experiments, self-compensating and constant time. For accurate results from multiexponential decay the use of constant-time experiments is preferred. In this approach the relaxation rate constants are determined by measuring experiments with 0 and T relaxation delay. If the relaxation is multiexponential due to multiple dipole-dipole interaction it can be described by

(1.65)

$$R_{eff}\left(\frac{1}{\tau_{cp}}\right) = R_{ex}\left(\frac{1}{\tau_{cp}}\right) - \left(\frac{1}{T}\right) In\left[\sum_j a_j exp[-R_j T]\right],$$

in which a_i and R_i are the amplitude and rate constant for the ith term in the multiexponential expansion.

CPMG experiments are an important tool for assessing the intermediate time scale motion in solution NMR samples.

Relaxation in the solid state

Recall that the anisotropic interactions are averaged out in the solution state because of the overall rotational diffusion of the molecules, which causes all the possible orientations to be sampled. On the other hand, in the solid state the molecules are static, and the overall tumbling is absent, the anisotropic interactions are not intrinsically averaged. This leads to broad spectral features as a result of the dipolar coupling and chemical-shift interaction. MAS is used to average anisotropic interactions only in a first-order average-Hamiltonian treatment by the periodic modulation or sample rotation. Thus, the averaging method is different: in solution, this is a stochastic and fast (nanoseconds) process, while for solids it is due to a determinist process of sample rotation. With the use of appropriate recoupling sequences, anisotropic interactions can be measured to obtain information about the dynamical processes that occur on the intermediate time scale of motion [8, 36, 37].

Relaxation in the presence of a spin-lock field: $T_{1\rho}$ [8, 23, 36, 38-42]

The dynamical processes in the microsecond to millisecond time scale can be probed by measuring the rotating frame relaxation in the presence of radiofrequency fields ($R_{1\rho}$) by solution and the solid state. In solid-state the coherent contributions, originated from incomplete averaging of strongly coupled networks of anisotropic interactions, can be suppressed using deuteration, the MAS frequency and the RF field strength. For a quantitative description of the $R_{1\rho}$ relaxation rate, there are three time-dependent processes that needs to be considered: MAS frequency, RF field amplitude of the spinlock, and the stochastic process, which modulates the anisotropic interactions and leads to relaxation. Therefore, we can describe the $R_{1\rho}$ spin relaxation induced by fluctuations of the heteronuclear ($I-S$) dipolar interaction and the spin CSA as

(1.66)

$$R_{1\rho}^{II} = \frac{3}{16} d_{II}^2 \{J(2\omega_e - 2\omega_r) + 2J(2\omega_e - \omega_r) + 2J(2\omega_e + \omega_r) + J(2\omega_e + 2\omega_r) + 10J(\omega_I) + 4J(2\omega_I)\}$$

(1.67)

$$R_{1\rho}^{IS} = \frac{1}{8} d_{IS}^2 \left\{\frac{2}{3}J(\omega_e - 2\omega_r) + \frac{4}{3}J(\omega_e - \omega_r) + \frac{4}{3}J(\omega_e + \omega_r) + \frac{2}{3}J(\omega_e + 2\omega_r) + J(\omega_I - \omega_S) + 3J(\omega_S) + 6J(\omega_I) + 6J(\omega_I + \omega_S)\right\}$$

(1.68)

$$R_{1\rho}^{CSA} = \frac{\omega_S^2 \Delta\sigma^2}{18}\left\{\frac{2}{3}J(\omega_e - 2\omega_r) + \frac{4}{3}J(\omega_e - \omega_r) + \frac{4}{3}J(\omega_e + \omega_r) + \frac{2}{3}J(\omega_e + 2\omega_r) + 3J(\omega_S)\right\},$$

in which $d_{II} = -\frac{\mu_0 \gamma_I \gamma_I \hbar}{4\pi\langle r_{II}^3\rangle}$ and $d_{IS} = -\frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi\langle r_{IS}^3\rangle}$ are the homo- and heteronuclear dipolar coupling constant, $\Delta\sigma$ is the reduced chemical shift anisotropy and ω_I and ω_S are the Larmor frequency of the involved nuclei, for studies of ^{15}N relaxation dispersion correspond to the Larmor frequency of ^{1}H and ^{15}N, respectively [8, 37, 38, 40].

In addition to the part induced by CSA and dipolar-coupling fluctuations, the contribution due to the isotopic CS fluctuations are termed exchange contribution and are defined by

(1.69)

$$R_{1\rho} = R_{1\rho}^{DD,CSA} + R_{1\rho}^{ex}.$$

Similarly, to solution NMR analysis, in solids the spectral density function can be defined as

(1.70)

$$J(\omega) = \frac{2}{5}\left(\frac{(1 - S^2)\tau_c}{1 + (\omega\tau_c)^2}\right),$$

where τ_c is the correlation time of the motion and S^2 is the generalized order parameter, which characterizes the spatial restriction of the motion, thus $S^2 = 0$ represents a system fully disorder and $S^2 = 1$ a system completely rigid.

In the solid state, the $R_{1\rho}$ relaxation rate constant depends on the MAS frequency and the rf-field amplitude. Relaxation dispersion (RD) techniques make use of these dependences; thus, the $R_{1\rho}$ relaxation rate constants are obtained at different effective irradiation field strengths

(1.71)

$$\omega_e = \sqrt{{\omega_1}^2 + \Delta\Omega^2},$$

on an offset frequency

(1.72)

$$\Delta\Omega = \Omega - \omega_{rf},$$

at an effective tilt angle

(1.73)

$$\beta_e = \arctan\left(\omega_1 / \Delta\Omega\right),$$

where ω_1 and ω_{rf} are the on-resonance field strength and the carrier frequencies, respectively [40, 41].

Solid-state applications based on the effective field strength (ω_e) can divide in two different regimes where $R_{1\rho}$is sensitive to µs-ms motion: The Block-McConnell Relaxation Dispersion (BMRD) and the Near-Rotary Relaxation Dispersion (NERRD).

The Bloch-McConnell RD is performed in the low ω_e regime, for which the dispersion is a consequence of the slow timescale modulation of the isotropic chemical shift tensor that is present in coherent dephasing. For the NERRD analysis, the experiments are recorded at higher radio-frequency field strengths, this is near the rotary resonance condition.

(1.74)

$$\omega_{RF}^{eff} = \omega_{MAS}$$

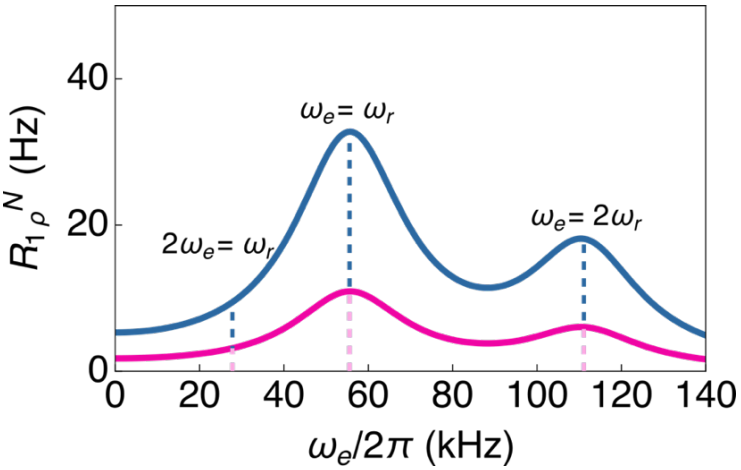


Figure 1.6. Recoupling conditions for hypothetical amide experiencing slower motion (blue) and faster motion (pink).

This experiment is sensitive to the angular fluctuations of the H-N bond vector, and if the bond angle fluctuations are involved in conformational exchange, the reorientation of the dipolar coupling and CSA interaction are present, and these fluctuations induce relaxation. As the experimental conditions approach the rotary resonance condition, the $R_{1\rho}$ rates will increase because of the sampling of spectral density function at a few kHz frequencies. Non-flat NERRD profiles are present specifically when motion occurs in the microsecond range, thus non-flat NERRD profiles are a straightforward sign of microsecond reorientation dynamics [40, 42-44].

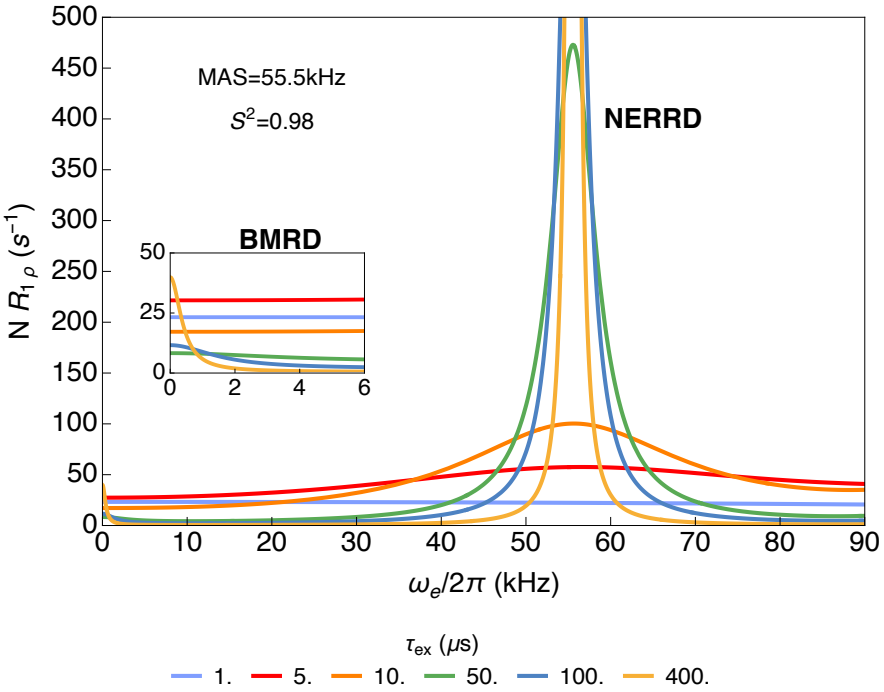


Figure 1.7. Theoretical full on-resonance $R_{1\rho}$relaxation dispersion profile of a hypothetical amide ^{15}N involved in a motion with slow timescale order parameter of 0.98 at spinning frequency of 55.5 kHz, which shows the Bloch-McConnell RD regime and the NERRD regime. The different timescale of motion (τ_{ex}) are shown in different color, the colors are indicated in the figure.

1.5 Protein Structure, function, and dynamics

The biological function of proteins is closely related to their structure and the conformational changes they experience. The complete structure of a protein can be described at different levels of complexity. To begin with the description, the primary structure is defined as the linear amino acid sequence. The primary amino sequence of a protein determines the higher order of structure due to the different chemical interactions that occur among the amino acids. In the next level, the secondary structure is defined as the local spatial conformation of the polypeptide backbone, without considering the side chains. The most common secondary structure elements include α-helices, β-sheets, and turns, which are generally defined both by a characteristic main chain ϕ and ψ dihedral angles, or torsion angles between the backbone atoms (**Figure 1.8**) [45, 46].

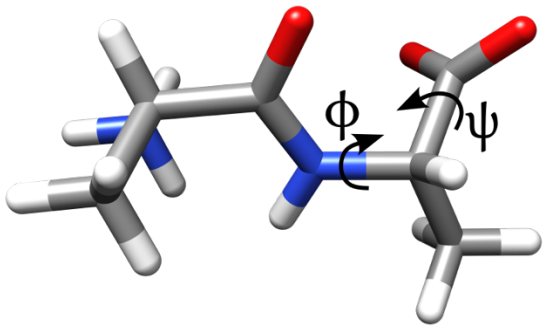


Figure 1.8. Representation of the ϕ and ψ dihedral angles for the Ala-Ala dipeptide. The ϕ angles describes the angle between the main chain amide and the carbon α atom, and the ψ angles describes the angle between the main chain carbonyl and the carbon α atom.

All structural elements have specific characteristics. For an α-helix to be formed, the backbone adopts a right-handed helical conformation with 3.6 residues per turn with a set of hydrogen bonds between the main chain carbonyl of the i^{th} residue and the main chain HN of the $(i + 4)^{\text{th}}$ residue. In the cases where the hydrogen bond is formed between the residues i and $i + 3$, a 3_{10}-helix will be observed. The optimal backbone ϕ and ψ dihedral angles for a right-handed α-helix are -57° and -47°, and for a 3_{10} helix are -49° and -26°, respectively [45, 46]. The β-sheets consists of at least two β-strands, each approaching an extended backbone conformation with dihedral angled confined to the region where the ϕ torsion angle is between -60° and -180°, and the ψ torsion angle is between 30° and 180°. In this structural element, all the main chain CO-to-HN hydrogen bonds are between adjacent strands [45, 46]. The tertiary structure refers to the three-dimensional arrangement of all the atoms that constitute a protein molecule; these three-dimensional folded shapes depend on their primary and secondary structures that make long-range interactions between amino acids possible [46-48]. The last level of organization of a protein is the quaternary structure, which is the association of several protein chains or subunits into a closely packed arrangement, each of the units has its own primary, secondary, and tertiary structure, and they are held together by hydrogen bonds and van der Waals forces [47].

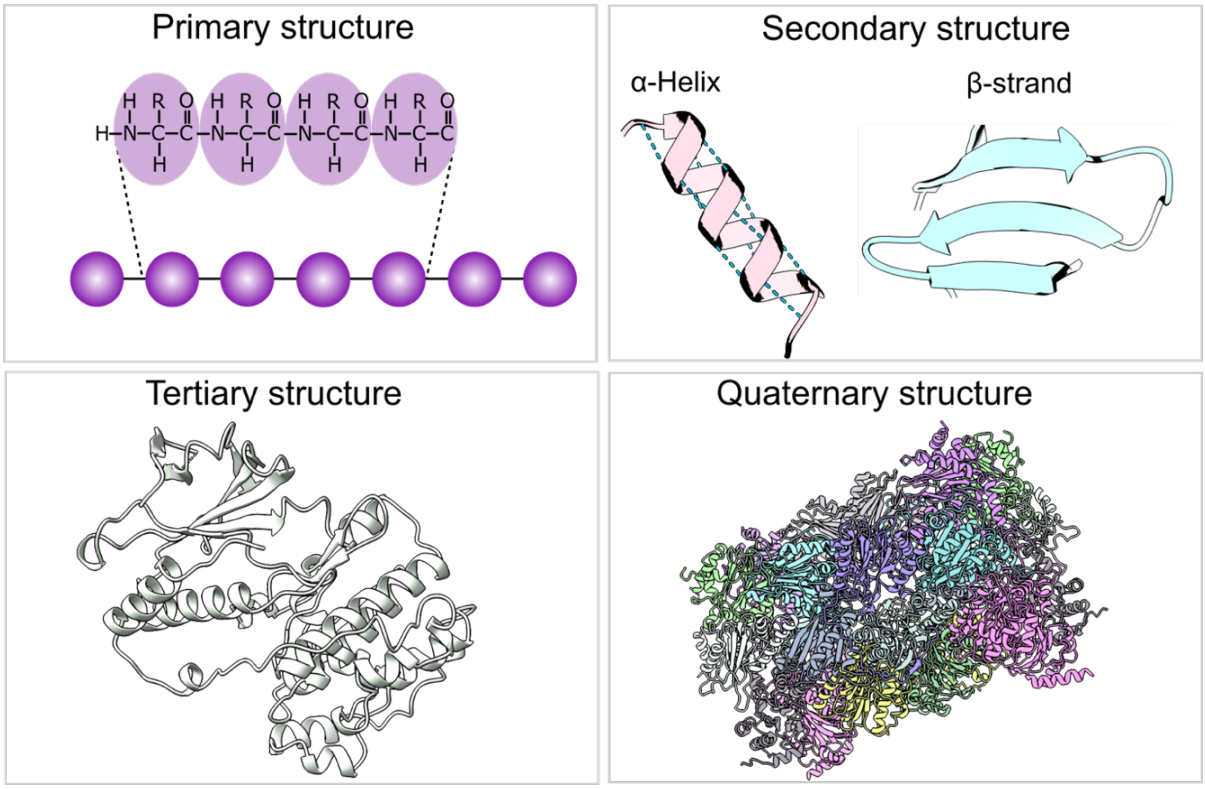


Figure 1.9. Representation of the levels of organization for proteins, with pictorial representations of the primary, secondary, tertiary, and quaternary structure.

Because the secondary structural elements of a protein originate by interaction among the atoms in the amino acid sequence, the individual effects of the chemical environment of the nuclei will influence the chemical shifts that are observed by NMR. Thus, chemical shifts are important tools for probing the secondary structural elements in a protein. The relationship between chemical shifts and the local protein structure of folded proteins has been established [49-51]. Different software to predict the secondary structural elements from the experimental chemical shift are available in different formats, from which it can be highlighted Talos+. Talos+ establishes the relationship between the chemical shifts of ^{13}C, ^{15}N and ^{1}H and the backbone torsion angles ϕ and ψ by using a neural network trained on 200 proteins [51].

Although secondary structural elements are present in a large number of known proteins, other protein systems with a lack of folded structures that can be fully functional have been observed in the past several decades. These proteins have been classified as intrinsically disorder proteins (IDPs); approximately 40% of the human proteome is estimated to be intrinsically disorder or contains significant length disorder regions. Due to the conformational behavior of these proteins, the study of their structure is difficult with many techniques. NMR spectroscopy is

the most powerful experimental tool for the analysis of IDPs at atomic resolution. The chemical shifts are the most easily accessible parameter and can provide ensemble-average local structural information about all the conformations present in the protein up to the millisecond time scale [52]. IDPs are characterized in NMR by low ^{1}H signal dispersion because of the similar electronic environment experience by the different protons. In a regular folded protein, having a small amount of dispersion will cause a high level of overlap, but for IDPs, this is counterbalanced by the narrower line widths due to inherent dynamic nature of the protein. By definition, the structural ensemble of an IDP is given by the primary structure, and the chemical shifts obtained are called *random coil chemical shifts* (RCCSs). As the neighboring residues will vary from one protein sequence to another, the analysis of the neighbored-corrected chemical shifts can be used to predict local changes in the structural ensemble and the propensity for some structural element. The neighbored-corrected secondary chemical shift (SCS, δ_{sec}) are the deviation from the RCCSs:

(1.75)

$$\delta_{sec} = \delta_{obs} - \delta_{ref},$$

in which δ_{obs} is the observed chemical $\delta_{ref}$is the RCCS. SCS can be used to unambiguously identify the formation of small populations of local structure and conformational propensities [53]. Similarly, for folded proteins, there are software capable of predicting the secondary structure propensity for unfolded proteins; one of the most recent developments is the Chemical shift Secondary structure Population Interference (CheSPI) software [50]. CheSPI predicts eight structural classes (SS8) defined by the DSSP classification algorithm [54], without bias to predict local structure.

In addition to the chemical shift information, experimental scalar coupling provides information about the secondary structure elements. Both, Homo- and heteronuclear three-bong J coupling constants carry information on the backbone dihedral angle ϕ and the side chain torsion angle χ_1 [4]. There are different methods for experimentally determining the values of *J*-coupling constants, and in small molecules, these values can be determined from standard one-dimensional spectra. In proteins, the process becomes complicated because of the larger number of atoms present. To overcome this problem, different pulse programs have been developed to extract the ^{3}J-coupling*;* these methods can be classified into three categories: i) direct measurement of resolved *J*-coupling from heteronuclear-edited spectra. ii) E-COSY-like methos for measurement of unresolved *J*-coupling, where a well-resolved ^{1}J-coupling in one

dimension of an n-dimensional experiments permits resolution of two components separated by the ^{3}J-coupling of interest. And iii) quantitative J correlation, in which the coupling constant is determined from the intensity ratio of two cross-peaks. [55-57].

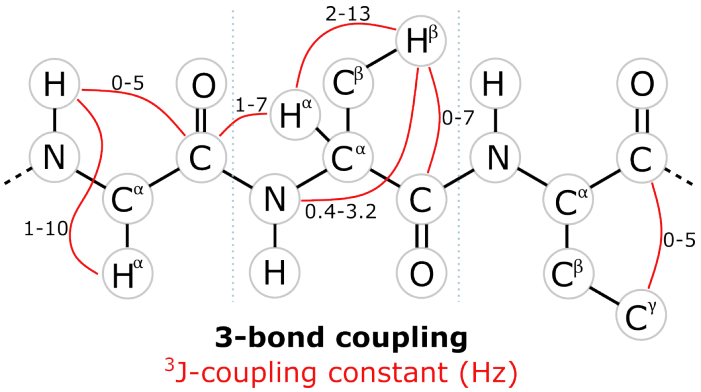


Figure 1.10. 3J-coupling constant values in Hz at three bond distance in amino acids

Provided empirical coefficients that are available, the ^{3}J-coupling values and angular constrains are interconverted by the Karplus equation [58]

(1.76)

$$3J(\theta) = Acos^2\theta + Bcos\theta + C,$$

in which A, B, and C are coefficients parameterized for each combination of nuclei.

The J couplings between different nuclei will provide information on the secondary structural elements in which the amino acid under study is located. Although these experiments provide great information, it is complicated to apply these methods to larger protein systems, as the resolution of the cross peaks becomes compromised [4, 59].

Proteins are dynamical systems, their inherent flexibility it is a crucial property to accomplish their biological function. This can be observed as conformational changes occurring at different timescales. As discussed in previous sections, NMR spectroscopy can be used to monitor the dynamic behavior of proteins at a broad range of time scales by various type of experiments.

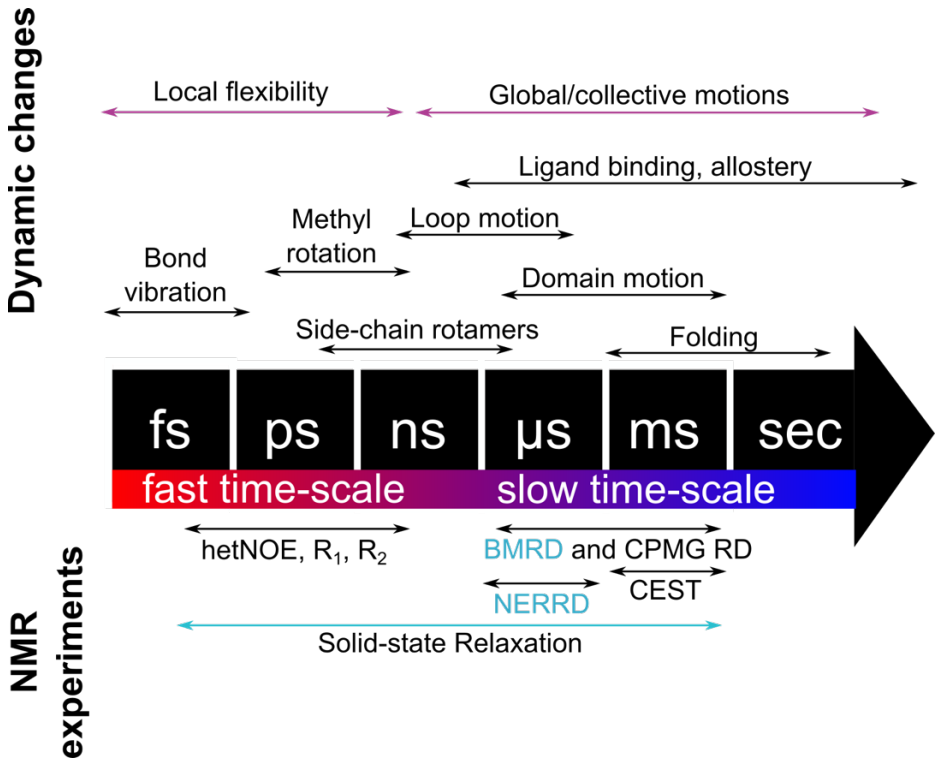


Figure 1.11. Molecular motion, time scales, and NMR experiments schematic representation.

A complete structural characterization of a protein defines the three-dimensional position of every atom in each state in the conformation landscape, the populations of each state via relative energies, and the rates and pathways of interstate conversion, also called dynamic flexibility [60]. NMR studies provide mechanistic detail at unmatched structural resolution into protein folding [61], low-populated or invisible excited state structures[62, 63], and thermodynamic quantities including entropy [60, 64].

1.6 Recombinant protein production

The first requirement for structural biology studies is sample preparation. As the visible isotopes in NMR for nitrogen and carbon are non-abundant in nature, it is necessary to use isotopically labeled samples to reach the desired level of signal in an appropriate amount of time. The most commonly used approach to obtain protein samples is the production of recombinant proteins.

Recombinant proteins are foreign proteins produced in the expression host and are encoded by a manipulated gen. These recombinant genes are a combination of one or more DNA segments that have been introduced from a different molecule or from different species [65]. The host organism whose protein synthesis machinery will produce the recombinant problem is chosen

according to the need for every case. One of the most used organisms is *Escherichia coli* because it is well characterized physiologically and metabolically, has unparalleled fast growth kinetics, high cell density can be easily achieved, can be expressed in different media, and transformation with exogenous DNA is fast and easy [66].
The general process of bacterial recombinant protein production consists of several steps. The first of them is protein expression. For this, the gene encoding the desire protein is cloned into an expression vector under the control of a promoter that will regulate the expression of the gene. The plasmid vector containing DNA is transformed into a strain of *E. coli* capable of producing the desired protein. At a specific stage of growth, the production of the recombinant protein is induced by the addition of a chemical inducer that activated the promoter in the expression vector. After induction, the recombinant polypeptide is overexpressed and can be purified[67-69].

During the production process, there are variables that can be controlled or adapt according to the needs for every particular case. At the DNA level, decisions related to the characteristics of the expression vector will include the promoter from which the recombinant gene will be expressed, the selection marker and the origin of replication. The PET system, based on the T7 RNA polymerase, is commonly used; for this system, the RNA polymerase is under the control of a *lac* operon. When a *lac* operon is present, isopropyl β-D-thiogalactopyranoside (IPTG) is commonly used to induce expression. IPTG mimics the effect of allolactose, a lactose metabolite that triggers the transcription of the *lac* operon. At the bacterial level, the selection of the *E. coli* strain can be important for the success of protein production. The choice of strain is guided by the requirements of the expression plasmid system; from a bioprocessing perspective, BL21 and derivative strains are a desirable strain due to the higher biomass yield and less acetate production compared to other E. coli strains [67, 69].

The selection of growth media for NMR sample preparation is usually restricted to the use of restrictive media that provide labeled glucose and ammonium chloride as the only source of carbon and nitrogen. One of the most widely used protocols is the use of minimal M9 media, a medium that contains salts, vitamins, and trace elements in addition to glucose and ammonium chloride [70]. Another criterion for optimization in the protein production process is the point of induction. This can be performed at different stages of growth, either at high or low biomass, and the selection of the induction point and the harvesting time will depend on the protein and the level of cell toxicity [67].

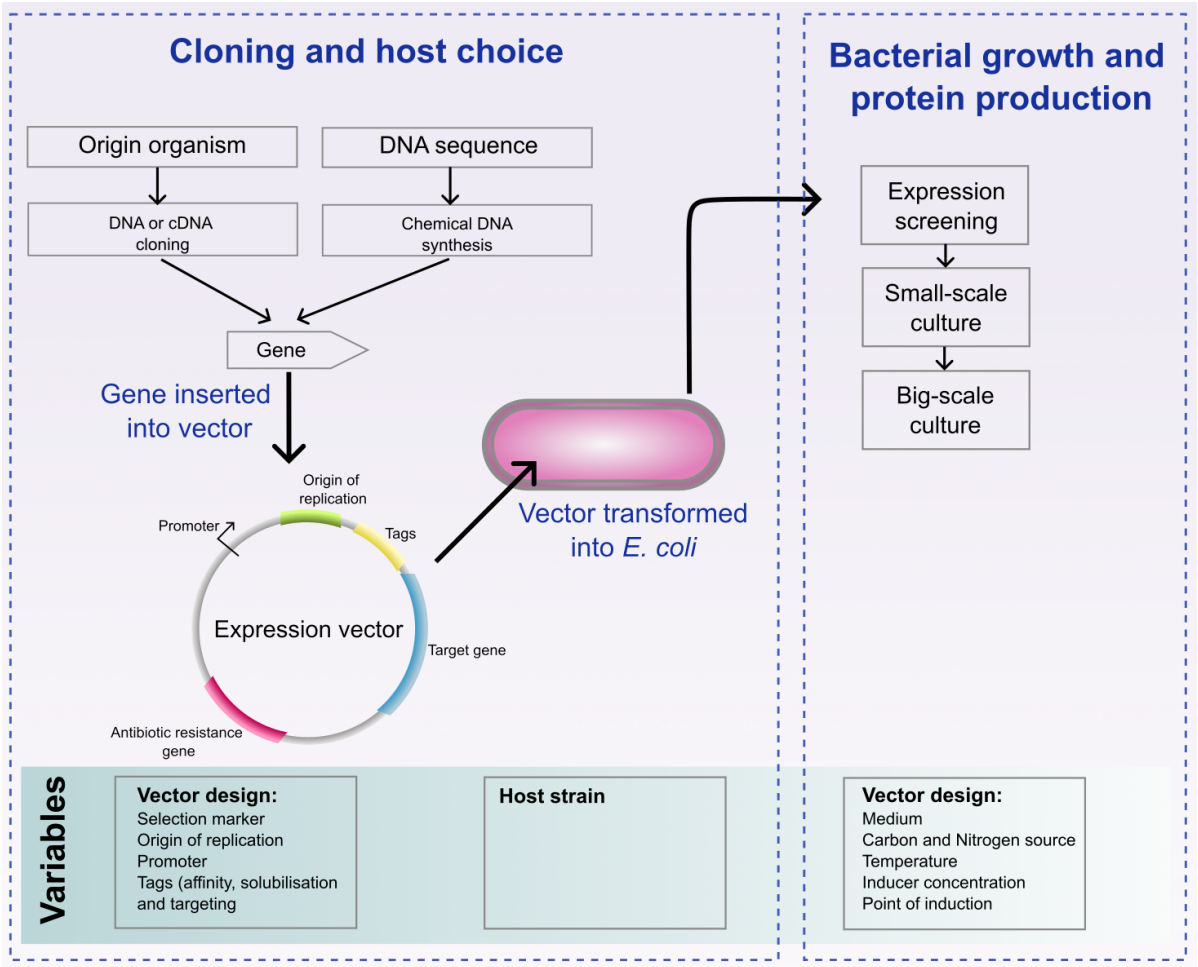


Figure 1.12. Overview of recombinant protein production process in *Escherichia coli* [67].

Once protein expression has been completed, the next step of the process is protein purification. The purification process consists of different steps that free the protein from contaminants, such as nucleic acids, viruses, residual host cell protein, and others. The separation process makes use of differences in chemical, structural or functional properties between the protein of interest and other proteins or molecules in the mixture. Properties include size, shape, charge, isoelectric point, charge distribution, hydrophobicity, solubility, density, ligand binding affinity, metal binding, reversible associations, and others [71].

For high-resolution purifications, chromatography is the method of choice. There are several chromatographic techniques that can be selected, and they can be used alone or in a series of steps to complete purification [72]. Affinity chromatography is one of the most widely used techniques, it exploits the ability of macromolecules to form specific and reversible complexes with binding partners and can be divided into two types. The first uses a natural-occurring structure or sequence of amino acids on the protein as the binding site, and the second involves

binding to a specific engineered sequence, commonly named a 'tag'. The degree of purity achieved will depend on the specific interaction [73]. Ion exchange chromatography (IE) is commonly used for the separation of molecules on the basis of their total charge. Proteins have surface charges depending on their isoelectric point (pI) and environmental pH. The ion exchanger is composed of a base matrix that provides enough surface area for adsorption, with the matrix positively or negatively charged. The protein is immobilized and by changing the buffer composition, the protein can be eluted separately from other components that were not attached to the matrix [74]. Another powerful chromatographic technique is size-exclusion chromatography (SEC), which separates molecules according to their hydrodynamic radius. The stationary phase consists of spherical porous particles with a controlled pore size, through which the molecules diffuse based on their molecular size difference (filtering) rather than their chemical properties. The main advantage of these techniques is the mild mobile phase conditions, which allow the characterization of the protein without affecting the conformational structure [75].

However, traditional isotopic labeling of proteins for NMR studies consisted in the introduction of ^{13}C and ^{15}N isotopes, as the molecular size of the protein increased the relaxation rates caused by the slower overall tumbling increase, causing a low signal-to-noise rate. The introduction of deuterium labeling improves the signal-to-noise ratio in NMR spectra by suppressing spin diffusion and by decreasing relaxation rates of ^{13}C and ^{15}N spins [76]. With the use of deuteration, some problems might arise in regions of proteins with high hydrophobicity, as the back-exchange process with ^{1}H is suppressed, leading to the invisibility of some residues. Different strategies have been developed to overcome the H-D exchange problem, as well as for selectively labeling protons in certain regions of the side chain while retaining a high level of deuteration. For nonexchangeable protons protocols such as reduce adjoining protonation (RAP) [77], stereo-array-isotope labeling (SAIL) [78], inverted fractional deuteration (iFD) [79] have been developed for solid-state samples [80].
In solution NMR, the combination of TROSY techniques, deuteration, and site-specific labeling has been developed for studying complex and big systems that do not allow for conventional methods. Metabolic pathways for amino acid biogenesis have been extensively studied with the purpose of finding adequate precursors for residue-specific methyl labeling [81, 82].

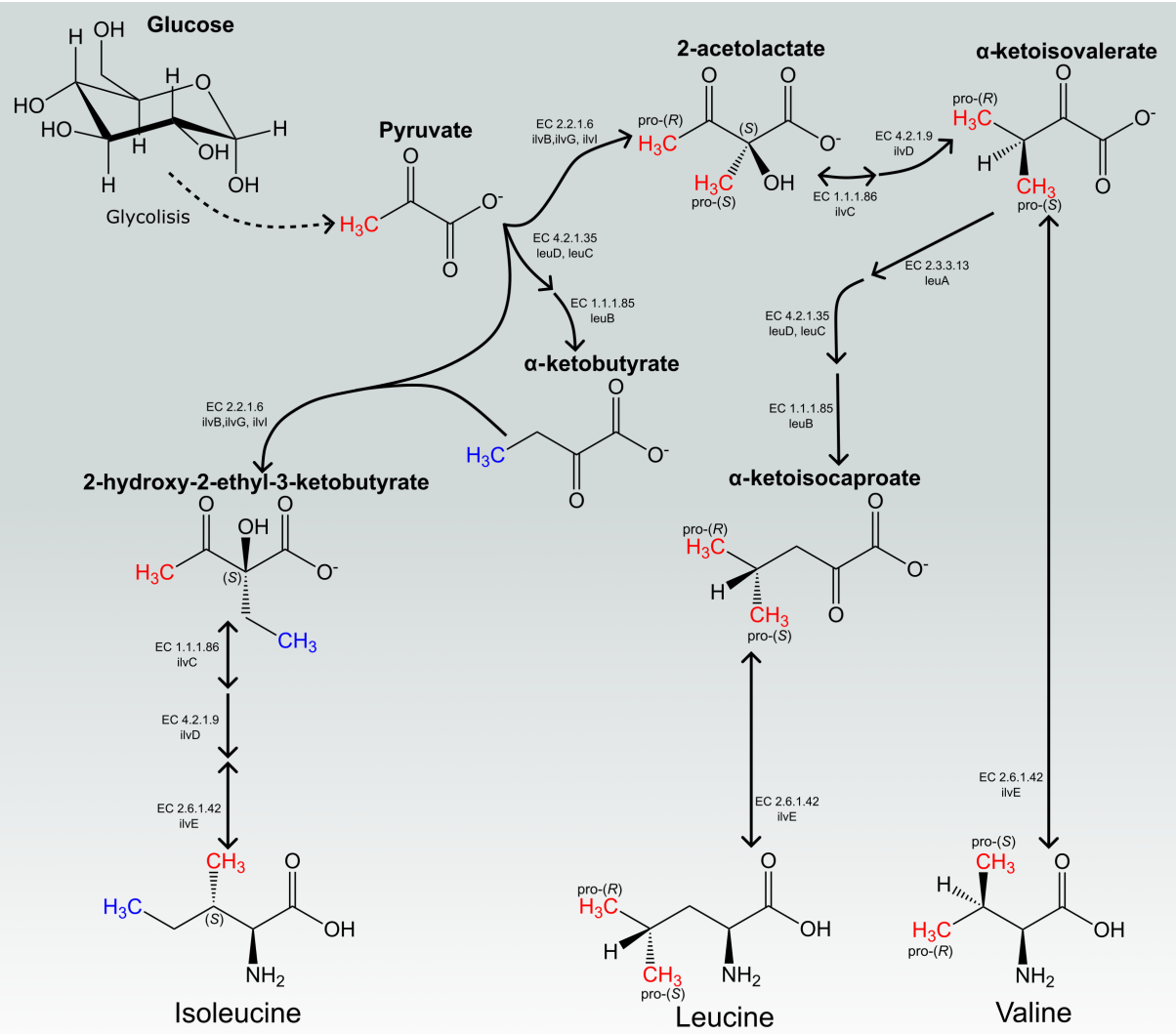


Figure 1.13. Metabolic pathways of *E. coli* involved in the biogenesis of isoleucine, leucine, and valine. The key metabolites that can be used as precursors for methyl labeling are indicated in bold. The color of the methyl groups indicates the location and stereospecificity of the molecules. The enzymes involved in biosynthesis are indicated by their EC number [81].

Chapter 2

Dynamics of p38α

2. DYNAMICS OF p38α

Protein kinases mediate many signaling processes in the cell. Mitogen-activated protein kinases (MAPKs) are a family of evolutionally conserved proteins that are activated by a wide range of extracellular stimuli as well as in response to inflammatory cytokines. The family consists of three major members: extracellular signal regulated kinases (ERK), p38 and c-Jun N-terminal kinases (JNK), representing three different signaling cascades [83, 84]. The p38 MAPK family is a highly evolutionary conserved group that mediated almost all stressful stimuli, composed of four isoforms (α, β, γ, and δ) ~60% of sequence similarity to each other but different in their expression patterns and specificities. Each of the isoforms is encoded by different genes: p38α (MAPK 14); p38β (MAPK11); p38γ (MAPK12); and p38δ (MAPK13). The p38α isoform is abundantly expressed in most cell types and is the best characterized member of the family, composed of 362 amino acids and a molecular weight of 41.3 kDa [85-87].

MAPK p38α can be activated by different extracellular signals that include cytokines, growth factors, neurotransmitters, hormones, cell stress, and cell adhesion. According to the type of tissue, the activation of the protein is carried out through different pathways. In neutrophils, tumor necrosis factor α (TNF-α) activates p38α and p38β specifically. For inflammatory cells, there are two pathways that lead to the activation of p38α; the first involves phosphorylation through MKK4, which is regulated by Rac/Cdc42 and cannot be activated by TNF-α. The second is done through MKK3 or MKK6, which can be activated by TNF-α and is not regulated by Rac/Cdc42. MAPK p38α is activated by dual phosphorylation of the tyrosine and threonine residues in a conserved Thr-Gly-Tyr motif in the activation loop. p38 MAPK signaling ends through the dephosphorylation of Thr and Tyr by phosphatases that return MAPK to the inactive state [85, 87-89].

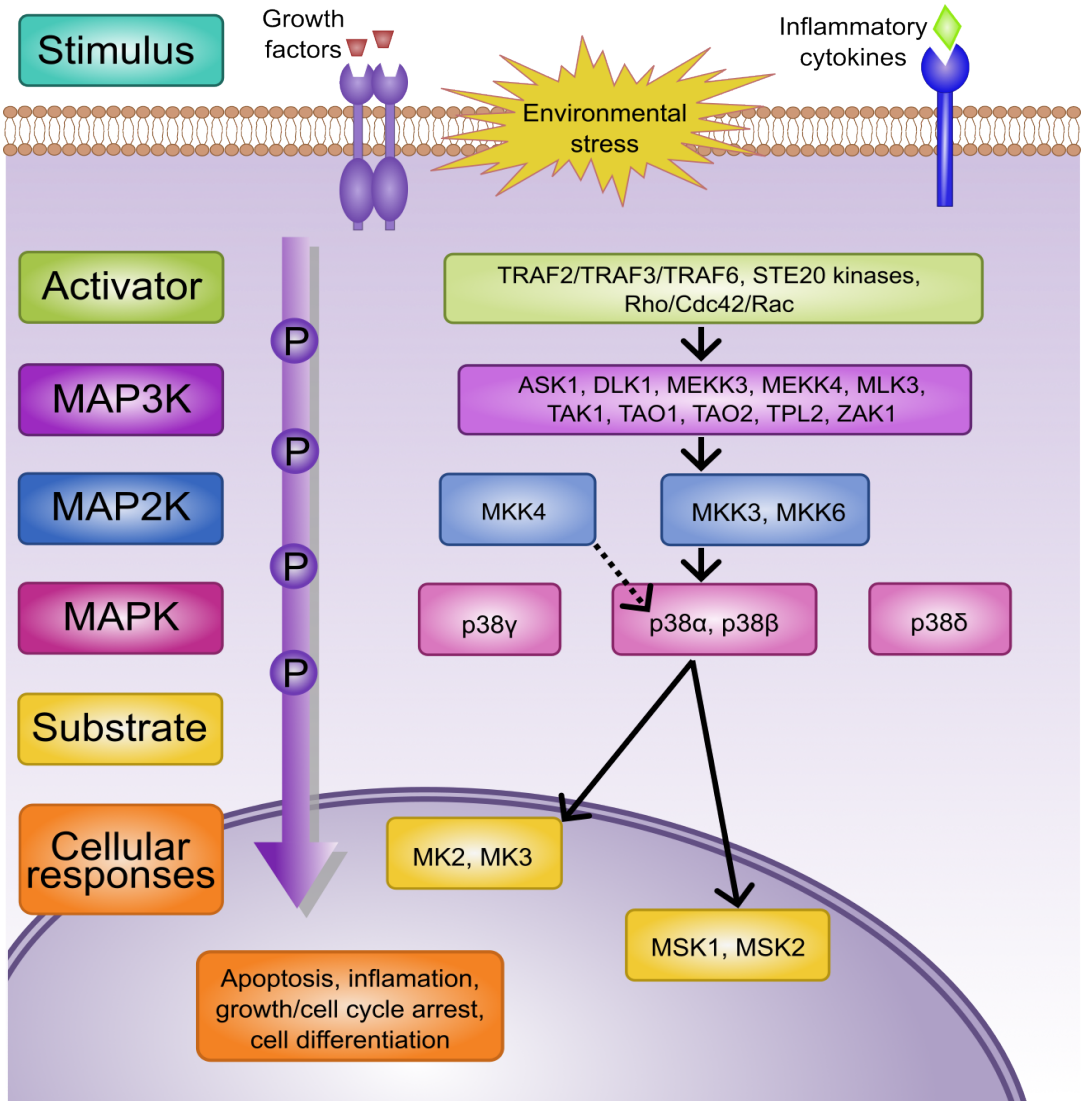


Figure 2.1 Mitogen-activated protein kinase p38 pathway. Stimuli such as growth factor, environmental stress, or inflammatory cytokines activated the phosphorylation cascade that leads to the activation of p38α, modulating cellular processes such as apoptosis, inflammation, and others.

As other protein kinases, p38α fold in two lobes: the N-terminal lobe and the C-terminal lobe connected by the hinge. The N-lobe mainly contains β-sheets and mediates ATP binding, and the C-lobe is dominated by an α-helical secondary structure and facilitates effector / substrate binding at the kinase interaction motif (KIM- or D-motif) binding site and a known allosteric ligand-binging pocket. The catalytic site of the protein is located between the two lobes, where the enzyme accepts ATP molecules, and is composed of the hinge region, the glycine-rich loop, the DFG motif, and the αC-helix (residues 63-77) [84, 87]. The region that interacts with the ATP molecule is composed of the adenine region (AR), ribose pocket (RP), and phosphate region (PR). Three hydrophobic regions can be found adjacent to the ATP binding site: hydrophobic region I, hydrophobic region II, and allosteric site (AS). Between hydrophobic

region I and AS there is a channel accessible both in 'in' and 'out' conformations of the DFG motif, also known as the linker region [87, 90]. Additionally, to the hydrophobic regions adjacent to the ATP binding site, a novel hydrophobic region is found in the C-lobe, for which there is not much information on the function yet, is known as the lipidic pocket [91].

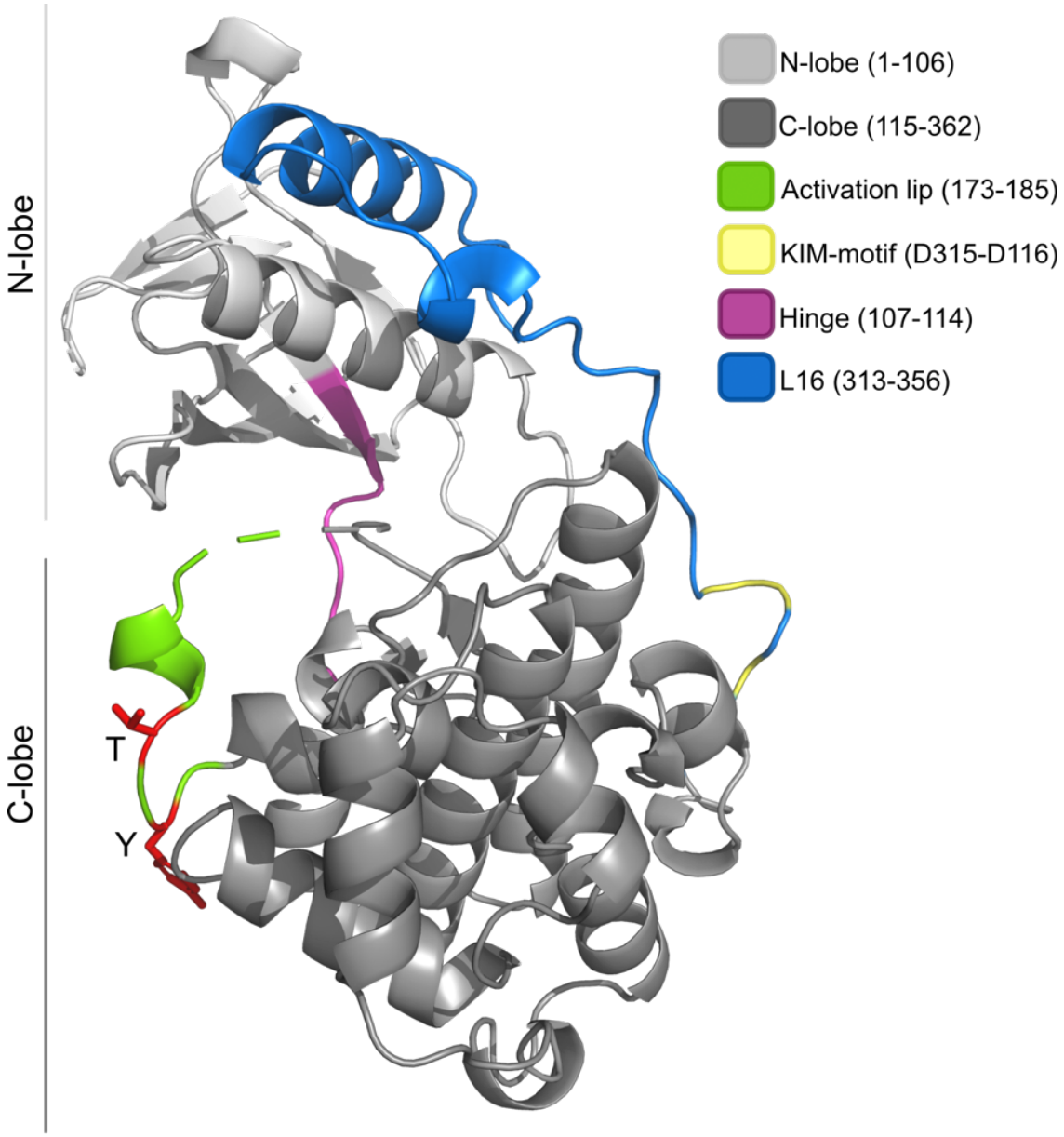


Figure 2.2. Illustration of the topology of p38α in the X-ray structure (PDB entry: 1WFC) where the different regions of the protein are illustrated in different colors.

When the protein is activated, it changes conformationally by a rigid body rotation of the N- and C-terminal lobes by 5° to 10°, which increases enzymatic activity and facilitates phosphoryl transfer. In addition to this conformational modification, crystallographic studies described the effect of the DFG-in and DFG-out conformations. In the DFG-in conformations,

the Phe side chain points towards the αC-helix occupying the hydrophobic back pocket while the Asp side chain orients towards the ATP-binding pocket, enabling the Mg^{2+} coordination of the phosphate groups of the ATP cofactor. The DFG-out conformation is represented by a flip of the Phe side chain occupying the active site, preventing this way the binding of ATP [92].

Due to the prominent expression of p38α in most cellular types and its important roles in different cellular processes, the enzyme is widely involved in the pathogenesis of various diseases. The p38α pathway is known to be connected to inflammation processes, rheumatoid arthritis, Alzheimer's disease, and inflammatory bowel disease [93-95]. In recent years, new roles for the p38 MAPK pathway have been elucidated, making members of the p38 MAPK family enter the group of canonical signaling pathways involved in the transformation process, as it has been shown that they can be involved in some of the alterations observed in the physiology of transformed cells: self-sufficiency in growth signals, unlimited replication potential, protection against apoptotic cell death, de novo angiogenesis, and tissue invasion and metastasis [85, 96]. This makes this family an important target in the drug discovery process against a wide range of diseases.

MAPK p38 inhibitors can be classified as ATP site and non-ATP site binders and according to the interaction model with the protein are classified into five types: Type I (TI-Is), which binds the region occupied by the adenine ring of ATP; type II (TII-Is), which binds to the inactive DFG-out conformation, opening a hydrophobic allosteric site, occupying both the ATP binding region and the allosteric site; type I1/2 (TI1/2Is), combine type I and type II features; type III (TIII-Is), occupy the hydrophobic region I, the linker region and the allosteric site without interaction with the ATP binding site; and type IV (TIV-Is), which bind to areas distant from the ATP binding site and normally are more selective, as they bind to regions less conserve in sequence and structure[87, 95].

Despite many efforts in medicinal chemistry and the many inhibitors approved for their use with other protein kinases, no inhibitor of the p38α MAPK has been accepted for market introduction. This is due to several reasons; from one site, the lack of efficiency or toxicity of the designed moieties plays an important role in the drug development process. However, due to the high evolutionary conservation of protein kinases, finding drugs with high specificity is complicated [83, 97]. An alternative to this problem is to direct the drug development process on the basis of structural knowledge of the protein. For this, knowing as much as possible about

the structure of the protein as well as mapping the dynamics of the protein at a broad range of time-scale regimes can provide sufficient information to direct the drug development process. Here we characterize the dynamic network of p38α at different time scale regimes in the apo form, and in complex with type II inhibitors (sorafenib and regorafenib) as well as the effect of the presence or absence of a ligand in the lipidic pocket. Through a combination of solution- and solid-state NMR methodologies.

2.1 Sample preparation

2.1.1 Protein Expression

p38α DNA was provided by our collaborators in the vector pNFG containing ampicillin and chloramphenicol resistant gen codified for rare codons, as well as a His tag with Thrombin cleavage site. The vector was transformed into BL21-DE3 *E. coli* cells in an ampicillin-containing agar plate. A single colony was used for a small overnight adaptation culture in LB. Cells were centrifuged at 6000x g for 5 min and the pellet was used to inoculate the minimal media culture, with different labeling schemes (**Table 2.1**).

Table 2.1. Labeling schemes used for solution and solid-state NMR studies

Labeling schemes for solution NMR	Labeling schemes for solid-state NMR
^{15}N, ^{13}C	^{15}N, ^{13}C
^{15}N, ^{2}H	^{15}N, ^{13}C, ^{2}H
^{15}N, ^{13}C, ^{2}H	^{15}N, ^{13}C, ^{2}H, RAP 50% D_2O
^{15}N, ^{2}H, ILV-methyl label	^{15}N, ^{13}C, ^{2}H, inverted Fractional Deuteration

The isotopically labeled recombinant protein expressed in bacteria requires the use of restrictive media enriched with labeled glucose, ammonium chloride, or precursors.

To determine the best expression conditions, an expression test was performed with two different concentrations of IPTG for induction (0.5 mM and 1 mM), two different temperatures (18°C and 30°C) for protein over expression, and two concentrations of M9 salts (1XM9 and 2XM9). The best conditions for the expression of the expression test were: 1XM9, 0.5 mM IPTG, and 18°C (**Figure 2.3**)

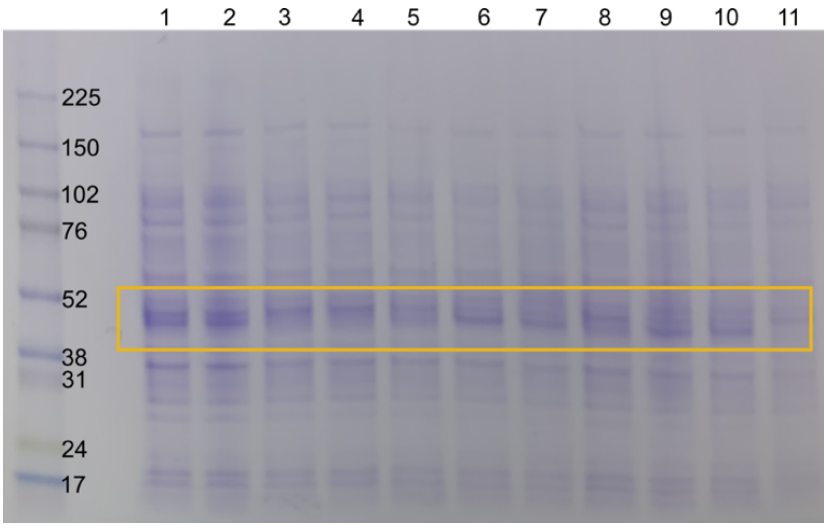


Figure 2.3. SDS Page of expression test showing the different tested conditions. 1: 1XM9, 18°C, and 0.5 mM IPTG; 2: 1XM9, 18°C and 1 mM IPTG; 3: 1xM9, 30°C, 0.5mM IPTG; 4: 1xM9, 30°C, 1mM IPTG; 5: 2xM9, 18°C, 0.5 mM IPTG; 6: 2XM9, 18°C, 1 mM IPTG; 7: 2xM9, 30°C, 0.5 mM IPTG; 8: 2XM9, 30°C, 1 mM IPTG; 9: LB, 18°C, 0.5 mM IPTG; 10: LB, 18°C, 1 mM IPTG; 11: LB before induction.

2.1.1.2 Deuterated samples

For a deuterated sample, several adaptation steps are needed to ensure bacterial growth. The adaptation is carried out by increasing the proportion of D_2O in the medium until reaching the final level of D2O, 50% D_2O or 100% D_2O (**Figure 2.4**). The adaptation and bacterial growth are carried out in restrictive medium (M9) with 1 g/L *NH_4Cl (15N) and 2 g/L D-Glucose (U-13C6; 1,2,3,4,5,6,6-D7). When the bacterial culture reached OD_{600} = 0.5, it was cooled for 15 min in an ice bath and inducible with 0.5 mM IPTG. Induced cells were expressed overnight at 18°C. After 18 hours, cells were harvested by centrifugation at 6000x g and 4°C for 20 min. Cell pellets were stored at -80°C until purification.

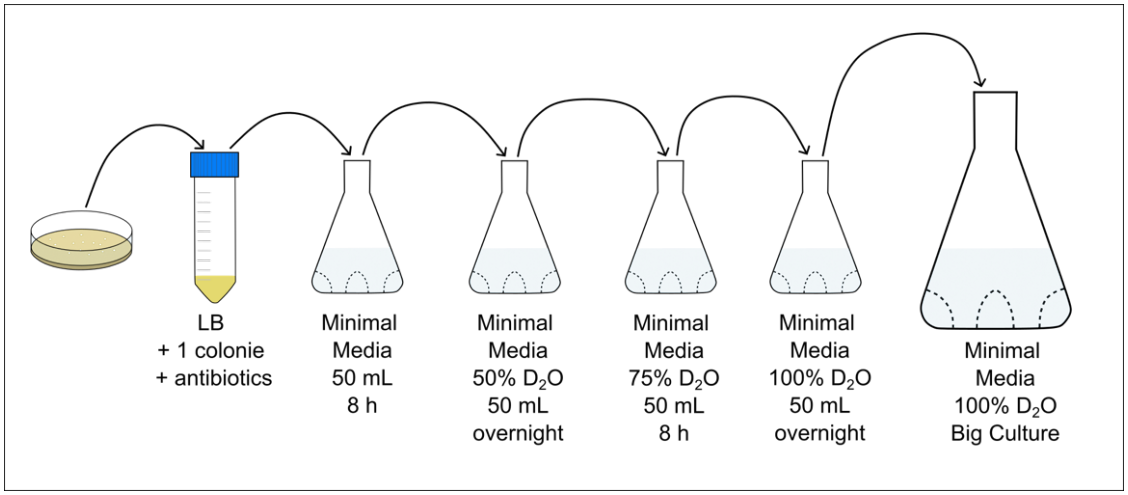


Figure 2.4. Diagram of deuterium adaptation steps for expressing ^{15}N,^{13}C,^{2}H labeled samples

2.1.1.3 Protonated samples

To incorporate protons at different sites on the side chain, there are different strategies. We can work with a fully protonated sample, where we can access the full sidechain information for structural calculations. Alternatively, for assignment purposes we can make use of the Hα protons, and we can achieve a high level of protonation at this position while keeping the rest of the sidechain deuterates with the inverted fractional deuteration (iFD). Both protocols consist of the same steps, there is no necessity of adaptation steps since the expression is carried out with H_2O. The only difference is in the substate that the bacteria is provided. We start with the freshly transform bacteria, one colony is transferred to a small bacterial culture in LB medium and antibiotics and it is grown overnight. The next day, the small culture was pelleted and transferred to the large culture containing M9 Minimal Medium supplemented with the respective label ammonium chloride and D-glucose (**Table 2.2**). When the bacteria culture reached an OD_{600}= 0.5 it was cooled for 15 min in an ice bath and inducible with 0.5 mM IPTG. Induced cells were expressed overnight at 18°C. After 18 hours, cells were harvested by centrifugation at 6000x g and 4°C for 20 min. Cell pellets were stored at -80°C until purification.

Table 2.2. Labeled supplements for labelling.

Fully Protonated	iFD
*NH_4Cl (15N)	*NH_4Cl
D-Glucose (U-13C6)	D-Glucose (U-13C6; 1,2,3,4,5,6,6-D7)

2.1.1.4 ILV-Methyl labeling

Labeling of methyl groups for specific amino acids requires an adaptation procedure to D_2O and supplementation of the medium with precursors. In the specific case of stereospecific labeling of valine and leucine methyl groups, additional treatment of the precursor is necessary through a hydrolysis reaction. For hydrolysis of ethyl 2-hydroxy-2-methyl-3-oxobutanoate, 1 g of compound was diluted with 4.08 ml of 0.6M K_2HPO_4. Approximately 350 µL of 10N NaOH were added in small aliquots until pH 11.5. After around 3 h, the pH stabilized at 10.44 and a sample was taken for a 1H 1D NMR spectrum. It was confirmed in the spectrum that the reaction was completed (**Figure 2.5**). The NMR tube content was returned to the rest of the

solution and 20 µL of 10N NaOH was added. After 1 hour, the pH was changed to 7.45 before adding the precursor to the 2L of adapted culture.

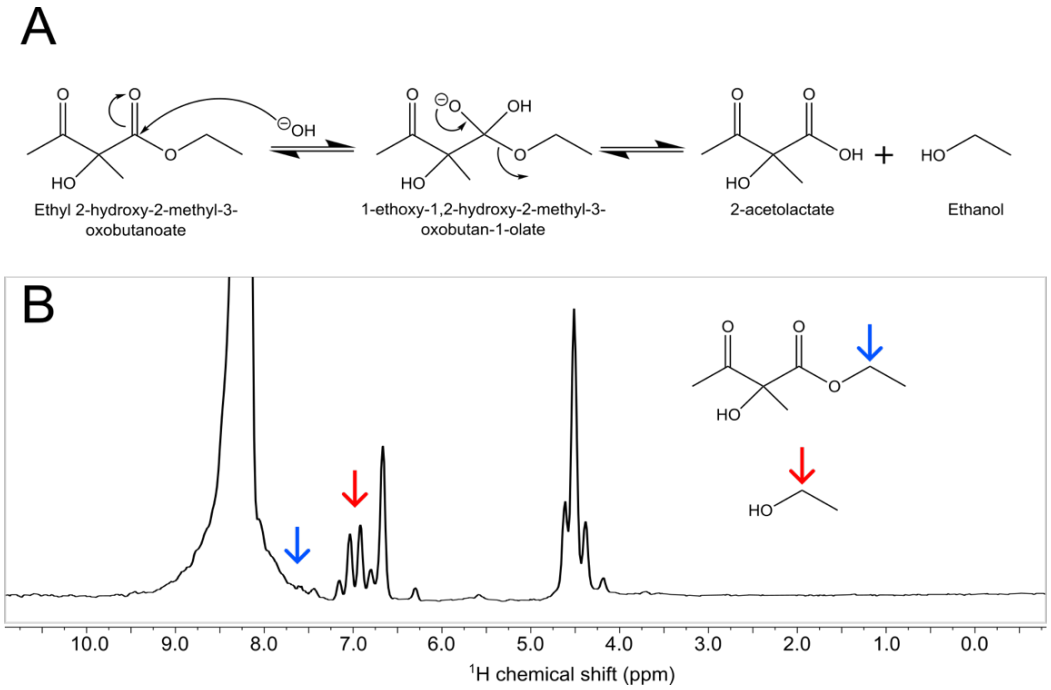


Figure 2.5. Hydrolysis process of the stereospecific leucine and valine precursor. **A)** Reaction mechanism for the hydrolysis of ethyl 2-hydroxy-2-methyl-3-oxobutanoate. **B)** ^{1}H 1D NMR spectrum of the hydrolysis reaction after 3 hours recorded at 60 MHz (^{1}H Larmor frequency) with 10% D_2O, the color arrows show the chemical shift of the respective set of protons and where we can see that the reaction has been completed.

After following the adaptation steps for the deuterated samples shown in Section 2.1.1.2, when the 2 L of the big culture reached OD_{600} = 0.3, the acetolactate was added. After 40 minutes the α-ketobutiric acid was added and after 20 minutes, cells were induced with 0.5 mM IPTG and incubated at 18°C for 20 minutes and finally harvested and stored at -80°C until purification with a final yield of 13 mg/L.

Table 2.3. Concentration of precursors added to the bacterial culture for ILV labeling.

Compound	Concentration per Liter of culture
*NH_4Cl (15N)	1 g
D-Glucose (1,2,3,4,5,6,6-D7)	2 g
2-acetolactate	500 mg
α-ketobutiric acid	70 mg

Upon purification, a ^{1}H-^{13}C HMQC spectrum was recorded to confirm the quality of the labeling. The spectrum shows a higher intensity for the Ile peaks compared to the intensity of the signal for Leu and Val. This is an expected phenomenon considering the metabolic pathway (**Figure 1.13**). The conversion from α-ketobutyrate to isoleucine has as an intermediate 2-hydroxy-2-ethyl-3-ketobutyrate, which is converted in one of the steps by ketol-acid reductoisomerase (EC 1.1.1.86). This is the same enzyme that converts 2-acetolactate, the precursor for the biosynthesis of valine and leucine. As the enzyme has a preferred affinity for 2-hydroxy-2-ethyl-3-ketobutyrate, the two precursors should be added in mistime [98]. The prepared sample will be used in further studies for obtaining eNOE contacts and residual dipolar couplings (RDCs) for structural calculations.

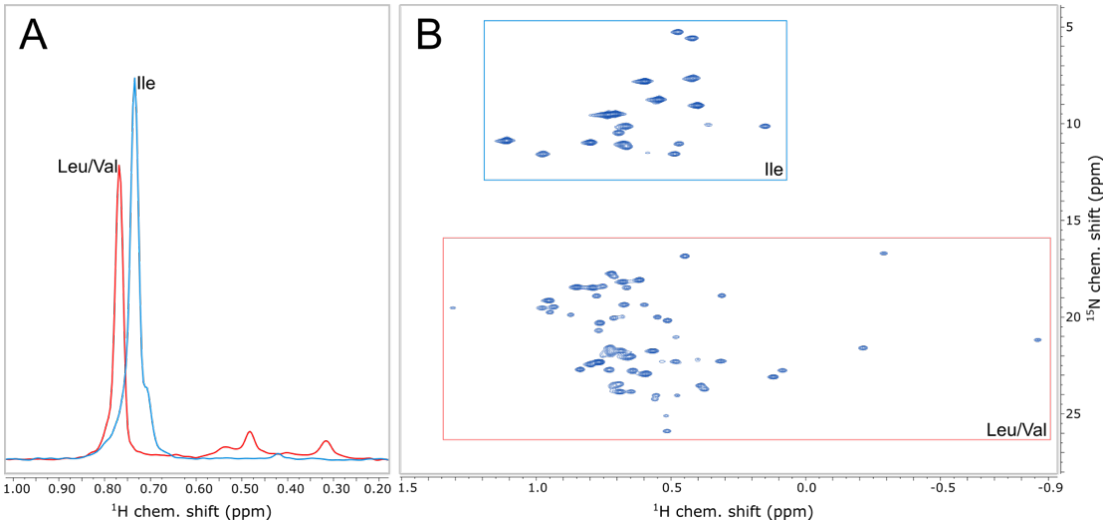


Figure 2.6. ^{1}H-^{13}C HMQC spectrum of ILV stereospecific labeled p38α. A) Slices of the 2D spectrum showing the intensity of two peaks correspond to an Ile and Leu/Val. B) 2D ^{1}H-^{13}C HMQC spectrum; Isoleucines are highlighted in blue, and Leucines/ Valines are highlighted in red.

2.1.2 Protein Purification

The 1L culture pellet was thawed and resuspended in 1: 3 volume of buffer A (**Table 2.4**). 25 mg of lysozyme per 50 ml of sample and 20 µL Benzonase were added. After 15 min of incubation on ice, the cells were disrupted through microfluidizer until the lysate became clear. The sample was centrifuged 45 min at 110,000x g and 4°C. Affinity chromatography was performed on a 2x HisTrap HP 1 ml column, the supernatant was injected at a flow rate of 1.5 ml / min through a superloop. The column was washed with 10 CV buffer A and the bounded molecules were eluted with 0-100% buffer B at a rate of 1.5 ml/min. The fractionation was followed by SDS-PAGE and the main peak was combined and dialyzed against 650 ml of dialysis buffer at 4°C, the buffer was exchanged after 1-2 h until a final time of 4 h. The dialyzed sample was transferred to a falcon tube and 300 µg/ml Thrombin was added to the

sample and incubated overnight at 4°C. The cleaved sample was centrifuged at 4000x g for 15 min and concentrated to 4-5 ml using a 10k-MWCO centricon. The sample was diluted 5: 1 in buffer QA and injected into Anion Exchange column (Anion HiTrap Q FF 5 ml) at a flow rate of 5 ml / min, non-bounded molecules were washed with 5 CV buffer A, protein was eluted with 0-100% buffer B over 20 CV. The fractions were followed by SDS-PAGE and the main peak was pooled, concentrated to 2-4 ml, and injected into an SEC column (HiLoad16/600 Superdex 75pg) with a flow rate of 1 ml/min. 2 ml fractions were collected and the major peak was identified as p38α in SDS-PAGE and solution-NMR (**Figure 2.7**).

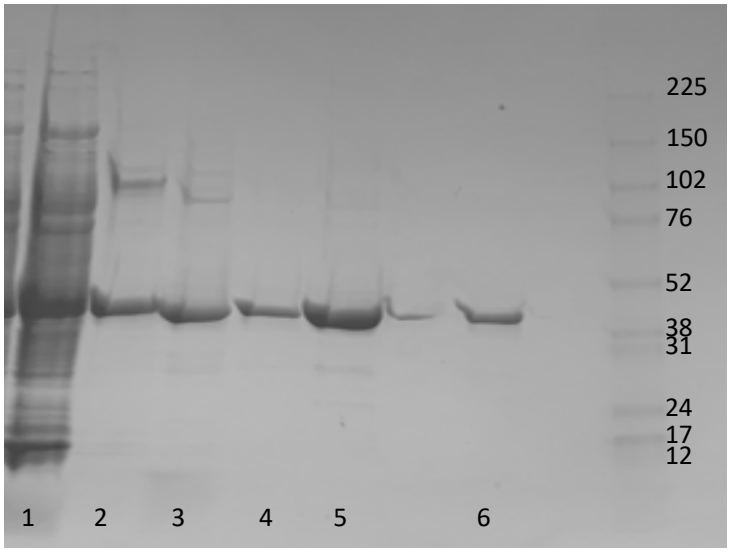


Figure 2.7. SDS Page of p38 purification process. 1: lysed pellet; 2: Ni-NTA chromatography fraction; 3: protein after cleavage with Thrombin; 4,5: Anion Exchange Chromatography fractions; 6: Size Exclusion Chromatography. The last column on the right shows the rainbow molecular marker with the respective molecular size for each band.

Table 2.4. p38α purification buffers for Affinity, Anion-Exchange and Size-Exclusion Chromatography and Dialysis

Buffer	**Composition**
A	50 mM Tris, 500 mM NaCl, 25 mM imidazole, 5% v/v glycerol, 1 mM DTT, 1 mM PMSF, pH 8
B	50 mM Tris, 500 mM NaCl, 500 mM imidazole, 5% v/v glycerol, pH 8
Cleavage/Dialysis	25 mM HEPES, 100 mM NaCl, 5% v/v glycerol, 1 mM DTT pH 7
QA	25 mM HEPES, 5% v/v glycerol, 1 mM DTT, pH 7
QB	25 mM HEPES, 1M NaCl 5% v/v glycerol, pH 7
SEC	20 mM HEPES, 50 mM NaCl, 100 mg/L methionine, 5% v/v glycerol, 10 mM DTT, pH 7

2.1.3 Crystallization

2.1.3.2 Crystallization screening

To find the crystallization conditions for the apo protein, two screening tests were carried out using the Dragonfly® crystal Screen Optimization Liquid Handler robots to mix the reservoir conditions and the mosquito® crystal to set the protein drops, both robots from SPT Labtech. For the screens, we used a plate-type TTPIQ with 96 wells, each well with a final volume of 65 µ. The buffer stock solutions used were 0.5M MES at pH 5.5, 0.5M MES at pH 6.5, 50% PEG4000 and 0.5M BOG.

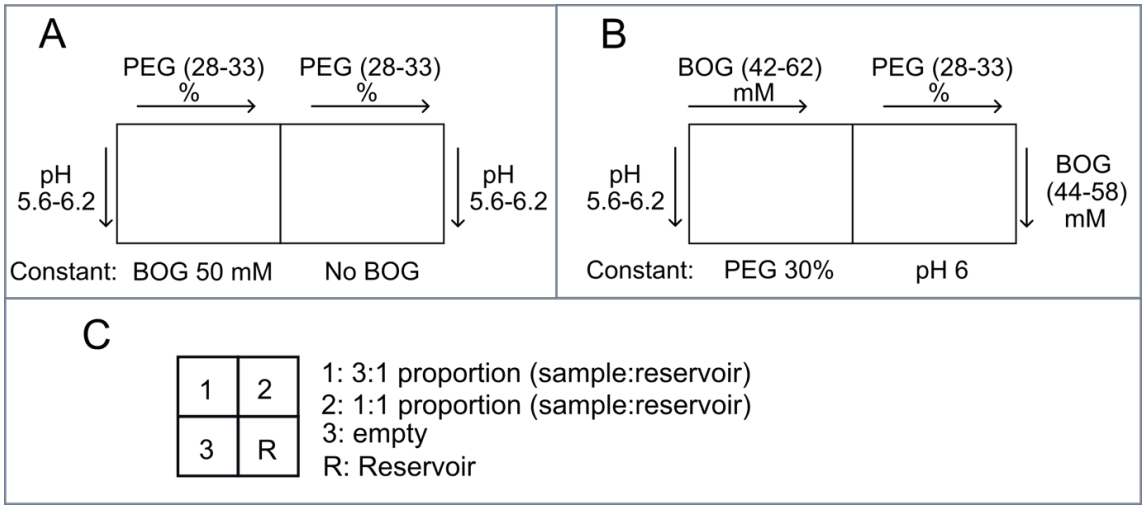


Figure 2.8. Schematic of crystallization screening for the p38 apo protein. A) Setting up for the first crystallization trial with variations of PEG400 in the presence or absence of BOG. B) Setup for the second crystallization trial with variations of PEG4000 and BOG. C) Drop proportion for each position of the well for every tested condition.

The results of these screenings were not successful as precipitation was observed for all drops. Because our collaborators had crystallized the protein in the presence of different ligands, another screening was done around the previous conditions and for testing the ligands that give the best microcrystals.

After size exclusion chromatography, the protein was concentrated at 10 mg / ml and mixed 40: 1 with the ligand compound at a concentration of 50 mM. Four ligands were tried: sorafenib (BAX), SA, BW17, and ME17. The mixture protein ligands were incubated on ice for 1 hour and centrifuged at 10,000 rpm for 1 min. The supernatant was used for siting drop crystallization. For crystallization screening, 15 conditions were tested for each ligand (**Figure 2.8**). The testing reservoir conditions were 100 mM MES at three different pH levels (5.6, 6.1 and 6.6), five concentrations of PEG4000 (20%, 22.5%, 25%, 27.5%, and 30%), 50 mM BOG and ddH_2O until a final volume of 500 µL. In each well, four 2 µL drops were set, each with

the different ligand to test, the protein: reservoir proportion in the drop was 3:1. The drops were set by first adding the protein drop followed by the reservoir drop (**Figure 2.9**). After one night, crystals were observed in most of the conditions and ligand, but for most of the drops, a high precipitation content was observed.

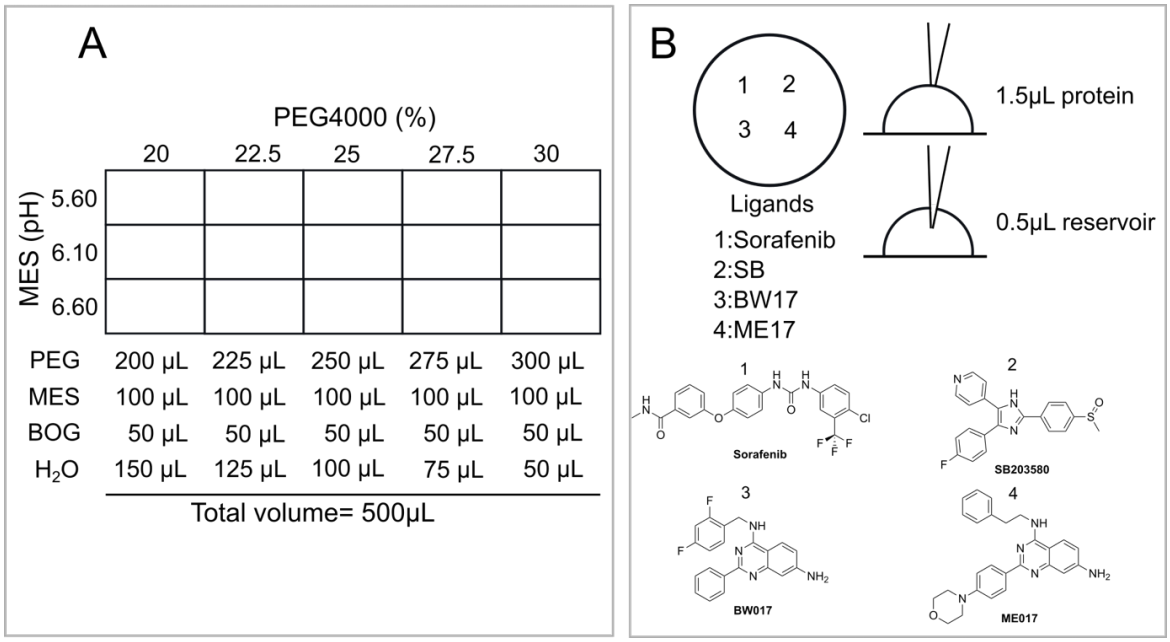


Figure 2.9. Diagram of the crystallization screening for co-crystallization of p38 with different ligands. **A)** Tested conditions for different concentrations of PEG4000 and pH for the 24 well plate. **B)** Crystallization drop setup for every well with different ligands as hanging drop.

2.1.3.3 Scaling trial

The compounds sorafenib and ME17 showed less precipitation for the conditions presented in

Table 2.5 compared to the other ligands. Conditions sorafenib + 30% PEG4000 + pH 6.1 and sorafenib + 20% PEG4000 + pH 6.6 were used for the scaling trial. In addition to these conditions, three drop sizes were used: 2 µL, 10 µL and 15 ml, as well as the absence of reservoir in the well. Crystals were instantly observed on drops of 2µL and 500µL reservoir in the well with the conditions sorafenib + 20% PEG4000 + pH 6.6. After one night, crystals were observed for the 2µL drops and 10 µL drops as well as for 10 µL drops without reservoir in the well. The selected conditions were used to test batch crystallization.

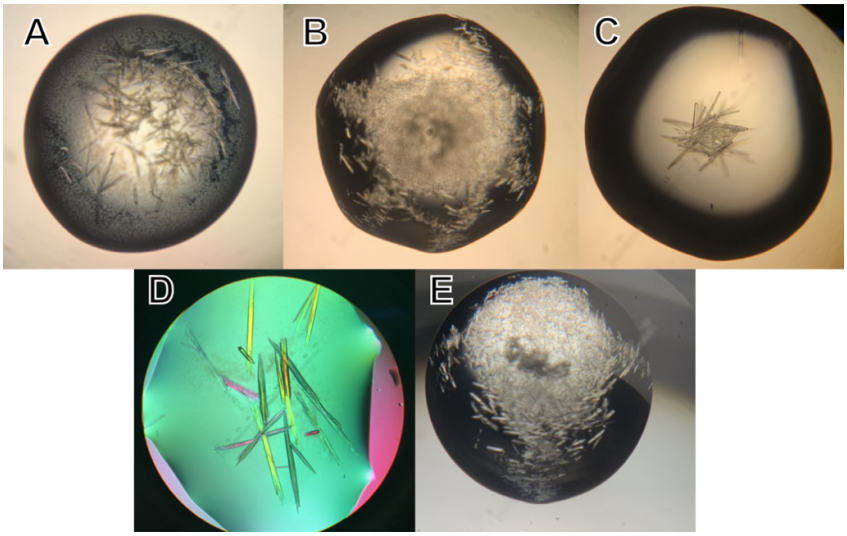


Figure 2.10. Crystals obtained under different conditions. **A)** Micro crystals obtained by cocrystallization with ME17 as ligand, 30% PEG4000 and pH 5.6. **B)** Micro crystals co-crystalized with sorafenib, 30% PEG4000 and pH 6.1. **C)** Crystals co-crystalized with sorafenib, 20% PEG4000 and pH 6.6. **D)** Crystals obtained after scalation trials with drop size of 10 µL grown without reservoir in the well after 24 h, co-crystalized with sorafenib, 20% PEG4000 and pH 6.6. **E)** Scaling trial, micro crystals obtained instantly in a drop size 2µL with reservoir present in the well, co-crystalized with sorafenib, 20% PEG4000 and pH 6.6.

Table 2.5. Crystallization screening conditions that show the less precipitant for p38α

Compound	**Conditions of the reservoir (PEG400 % and pH)**
Sorafenib	30%, 6.6
	27.5%, 6.1
	25%, 6.1
	20%, 6.6
ME17	30%, 5.6
	27.5%, 5.6

2.1.3.4 Batch Crystallization

Due to the number of micro crystals needed for filling an NMR rotor, classical techniques of crystallization represent a problem when harvesting the crystals not only because of the time demand on pipetting, but also because of the possibility of damaging the crystals. The solution proposed to overcome these difficulties is to grow the crystals directly in a larger container. This procedure is called batch crystallization and is ideal for use in protein samples that can form crystals instantly and grow as a hanging/sitting drop without a reservoir in the well [98, 99].

For batch crystallization, the protein solution was prepared as previously described before by mixing ~[9.5 mg/mL] protein with 50 mM sorafenib in proportions 40:1, incubated 1 h on ice and centrifuged at 10,000 rpm for 2 min. The supernatant was aliquoted in fractions of 30 µL in Eppendorf microcentrifuge tubes. The reservoir solution contained 20% PEG4000, 50mM BOG, and 100mM MES at pH 6.6. Regarding the hanging drop procedure, 10 µL of reservoir were added to the protein solution and mixed by pipetting up and down twice. After mixing, the solution became cloudy, and crystals were observed after a short period of time. Eppendorf microcentrifuge tubes were incubated at room temperature for 24 h and stored at 4°C until use. In addition, the presented procedure can be performed in the absence of BOG to obtain protein crystals with the lipidic pocket empty.

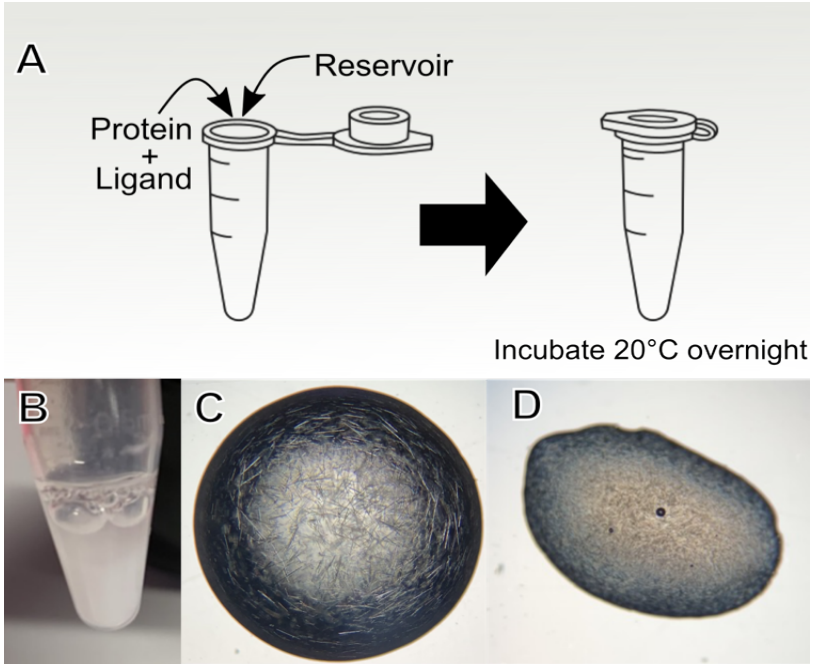


Figure 2.11. Batch crystallization. **A)** Diagram of the procedure for batch crystallization in Eppendorf micro-centrifuge tubes. **B)** Micro crystals grown in Eppendorf micro centrifuge tube after 3 days of incubation at room temperature. **C)** Micro crystals co-crystallized with regorafenib and grown by batch crystallization. **D)** Extended drop for crystals obtained by batch crystallization after one week of incubation at room temperature.

2.1.3.5 Paramagnetic Relaxation Enhancement (PRE)

As long acquisition periods are expected due to the large molecular size of the protein, it is important to consider a strategy to decrease the T_1 relaxation time and therefore the recycle delay to reduce the measurement time. This is possible with paramagnetic doping. Previous studies with solid-state NMR suggest the use of metals such as Cu (II) and Gd (III) as paramagnetic agents 99], the use of the Cu (II)-EDTA complex has been the most widely used

doping agent [98]. Following this line of knowledge, a first doping attempt with a self-made stock solution of 350 mM Cu (II)-EDTA was carried out at pH 7.47 by adding small amounts to the crystal solution until 10 µL. However, a couple of minutes after adding the doping agents, the crystals dissolved.

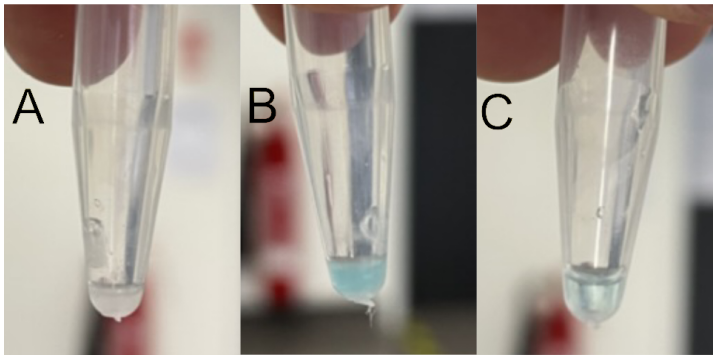


Figure 2.12. Dope crystals with Cu (II)-EDTA. **A)** Crystals without doping. **B)** Crystals with 3 µL 350 mM Cu (II)-EDTA right after adding the doping agent. **C)** Crystals with 3 µL 350 mM Cu (II)-EDTA after incubation for 1 hour, where it can be observed that all crystals are completely dissolved.

Because the stock solution used is self-made, as the first cause of the crystal dissolution, an excess of ions was assumed in the solution coming from the first step of the stock solution preparation, which is the protonation of EDTA to make it soluble. Therefore, a second test with different preparations of Cu (II)-EDTA and different ions was performed.

Among the tested conditions, we have a different preparation for the stock solutions using either water, SEC buffer, or mother liquid. Doping testing was performed using unlabeled crystals obtained by batch crystallization. The results showed negative results for all conditions where the crystals dissolved after 30 min of incubation. After all the alternatives for Cu (II)-EDTA were used without positive results, we considered the use of another type of chelate.

Table 2.6. Tested conditions for Cu (II)-EDTA stock solutions used for crystal doping.

Tested Conditions	Water	Buffer*	Mother liquid**
Cu (II)-EDTA Na^+	X	X	X
Cu (II)-EDTA NH_4^+	X	X	X
Cu (II)-EDTA Na^+ (sigma)	X	X	X
Cu (II)-EDTA NH_4^+ (sigma)	X	X	X

*SEC Buffer 20 mM HEPES, 50 mM NaCl, 100 mg/L methionine, 5% v/v glycerol, 10 mM DTT, pH 7
**Crystallization supernatant: SEC buffer plus Reservoir.
***X indicates negative result

In medical Positron Emission Tomography (PET), the use of radiopharmaceuticals that contained ^{64}Cu isotopes is widely used, and for these applications it is necessary to supplement copper with a good chelating agent. The use of tetraazamacrocyclic complexes as good chelating agents has been reported [99]. One of the selected macrocycles is CB-TE2A, after complex formation, this compound has two ions less than Cu (II)-EDTA (**Figure 2.13**), which could potentially decrease the risk of crystal damage.

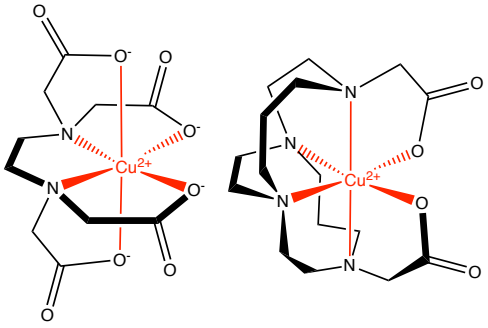


Figure 2.13. Chemical structures of Cu (II) complexes with chelating agents

Another doping test was performed using unlabeled micro crystals of p38α grown by batch crystallization with a stock solution of 310 mM Cu (II)-CB-TE2A reaching final concentrations for the final solution (**Table 2.7**) and were monitored for several days.

Table 2.7. Doping test with Cu (II)-CB-TE2A at different concentrations

Tested conditions	1h	24h	48h
Cu (II)-CB-TE2A [21.7 mM]	OK	OK	OK
Cu (II)-CB-TE2A [31 mM]	OK	OK	OK
Cu (II)-CB-TE2A [75 mM]	OK	OK	OK

The crystal stability was monitored for several days, and a good level of stability was observed after 48 hours. However, a change of appearance can be observed after a week, where the crystals appear to be more dispersed and less dense. The crystals were observed under the microscope and the shape of the crystals was maintained. After 2 months, only precipitated protein was observed in the solution, although the shape of the solution did not show a significant change with a simple eye.

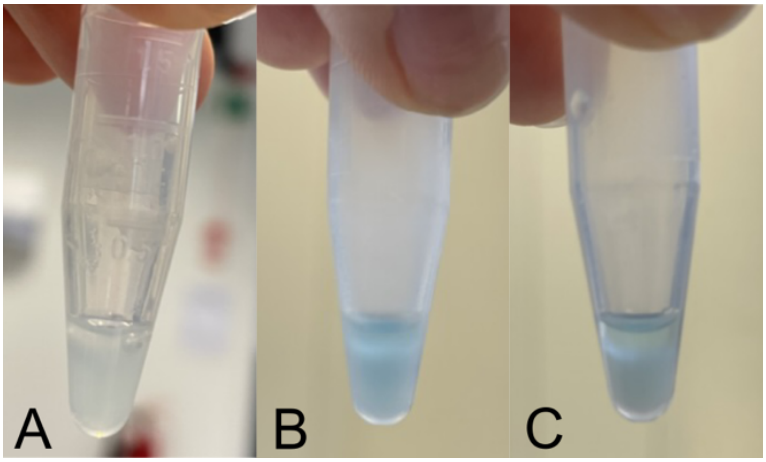


Figure 2.14. Stability test for doping with Cu (II)-CB-TE2A. **A)** Crystals without chelate. **B)** Crystals with 10 µL 310 mM Cu (II)-CB-TE2A one hour after doping. **C)** Crystals with 10 µL 310mM Cu (II)-CB-TE2A after two weeks of incubation.

From these conditions a concentration of 31mM Cu (II)-CB-TE2A was selected for doping-labeled crystals. Crystals were incubated over 24 hours before rotor filling. The predoped crystals could not be packed into the NMR rotor because they clumped together. A second attempt to fill the rotor was made with non-doped crystals and compressed at 20 kHz for 40 min, later the doped was done by diffusion by adding 10 µL of supernatant and 20 µL of supernatant doped solution with a final concentration of 60 mM Cu (II)-CB-TE2A (**Figure 2.15**). The rotor with the solutions was kept overnight. After 15 hours, a precipitation was observed at the tip of the pipette. The rotor was compressed for the second time for 25 min at 20 kHz before the final closing. The sample was looked at by the spectrometer and showed a T_1 time of 300 ms. However, the sample went bad after one week.

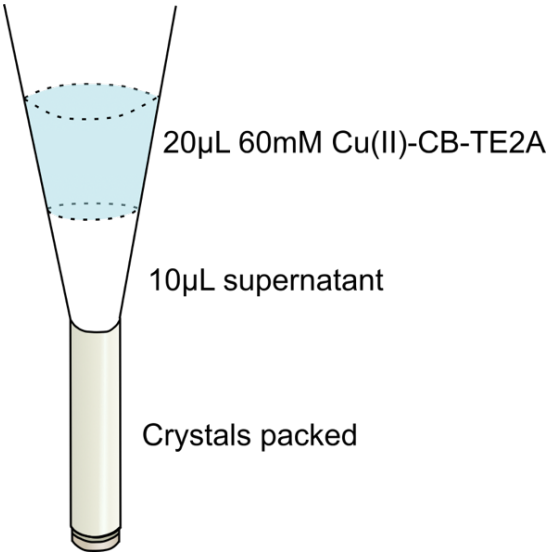


Figure 2.15. Diagram for doping while rotor filling procedure.

Polyethylene glycol (PEG) is known as a nucleation catalyst for protein crystallization but also as a cryoprotectant [100, 101]. Low molecular weight PEG, such as PEG 200, 400, and 600 are used successfully as cryoprotectants, while polymers such as PEG4000 do not penetrate crystals but coat the crystal lattice, are also used as non-penetrating cryoprotectants [101]. Following this principle, PEG4000 was thought to be used as a crystal protector for doping trials with Cu (II)-CB-TE2A. For this trial, different proportions of 50% PEG4000 were tested: protein crystals (1:10, 2:10 and 3:10) were tested in conjunction with different proportions of the chelating agent: protein crystals (0.5:10, 1:10 and 1.5:10), with positive results (**Table 2.8**)

Table 2.8. Doping test results using Cu (II)-CB-TE2A and PEG4000 as protectant.

	50% PEG4000			
Cu (II)-CB-TE2A		1:10	2:10	3:10
	1.5:10	OK (less)	OK	OK (less)
	1:10	OK	OK	OK (less)
	0.5:10	OK	OK	OK

Another set of crystals for doping was prepared using 200 µL of protein crystals, 40 µL PEG400, and 10 µL 300 mM Cu (II)-CB-TE2A. All components were mixed before rotor filling. After compressing the crystals at 20kHz for 30 min and finalizing the rotor sealing, the sample was monitored on the spectrometer. A set of bulk T_1 relaxation experiments was recorded and a final T_1 of 87 ms was obtained. However, sample degradation was observed after recording a ^{15}N-hNH. The remaining doped crystals were checked under the microscope and the presence of precipitation was observed. Later it was observed that the precipitant came from Cu (II) -CB-TE2A, which is soluble up to pH 5 at a concentration of 300 mM. Because the solution with the doping agent is blue in color, it was observed that the blue coloration persisted in the supernatant of the previous doped crystals. To recover the solution, the crystal mixture was centrifuged at 10,000x g for 2 minutes and the clear supernatant was used as the other new liquid for doping. The pH of this solution was changed by titration until a pH was measured between 6-7 with pH stripes on paper. A fresh batch of crystals was pelleted, and the supernatant was exchanged for the doping mother liquid in 10 l steps of 10 µL every time. The doping process is shown in **Figure 2.16**.

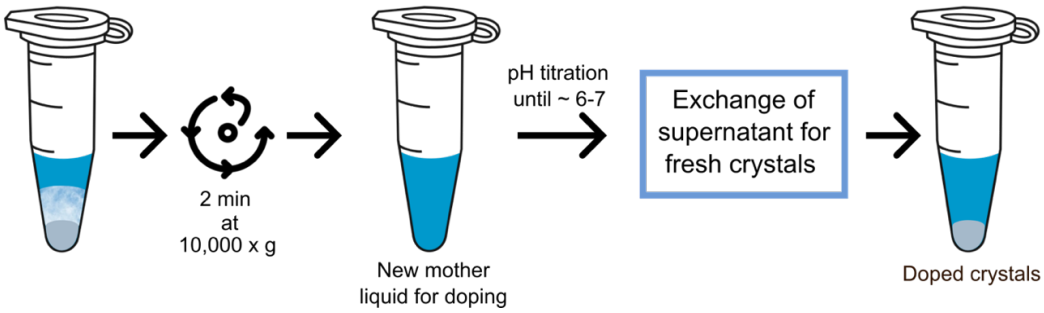


Figure 2.16. Scheme for doping procedure with Cu (II)-CB-TE2A.

The doped crystals were packed into a 1.3 mm rotor and compressed at 20 kHz for 30 min before sealing the rotor. A series of bulk T_1 experiments were performed on this sample, resulting in a $T_1 = 208$ ms in contrast to $T_1 = 873$ for the non-doped sample. Allowing to reduce de recovery delay by a factor of 5, reaching a minimum of 0.3 seconds. Furthermore, the sample was monitored by measuring a ^{15}N-hNH after four weeks, which was shown to be stable.

2.2 Solution NMR

2.2.1 Protein assignment

To gain access to site-specific information about any protein, the first step needed is assigning all the resonances. For the solution-state NMR assignment of a protein of this molecular size, the use of different alternatives for sample preparation (use of deuteration for diluting the ^{1}H network) and acquisition of experiments, like the use of transverse relaxation-optimized spectroscopy (TROSY). Partial assignment of the protein was done by Peti et al. [102], we used a set of 3D experiments to transfer those assignments for the apo protein and assign the protein in the complex with different ligands.

2.2.1.2 Experimental conditions

Before acquiring all experiments, the sample SEC buffer was exchanged for the NMR buffer consisting of 50 mM HEPES, 150 mM NaCl, and 5 mM DTT at pH 6.8. The samples were concentrated at approximately 400 µM and placed in a Shigemi NMR tube. For samples in complex with different ligands, before buffer exchange, the protein solution was concentrated at approximately 9 mg/ml and mixed with the ligand (50 mM sorafenib, 50 mM regorafenib) in proportions 40: 1 (protein: ligand). The mix was incubated on ice for one hour, then centrifugated for 2 min at 10,000x g. The clear solution was then processed as the apo protein

sample. For samples with BOG, the BOG was added in the final concentration step to reach a final concentration of 50 mM BOG.

All NMR spectra were recorded at 293K on a Bruker Avance NEO 800 MHz (^{1}H Larmor frequency) spectrometer equipped with a CP TCI proton-optimized triple resonance cryoprobe. The recorded solution NMR assignment experiment consists of 2D ^{15}N-^{1}H TROSY HSQC, 3D TROSY HNCA, 3D TROSY HNCO, and 3D TROSY HNCACB.

2.2.1.3 Assignment transference

Some nuclei are more sensitive to changes in the chemical environment, in the case of proteins, these nuclei are amide protons, as their chemical shift is influenced by the pH, temperature, and other factors. This is especially important to consider when we compare samples from different preparations and measure them in different laboratories. Therefore, for assignment transference from the BMRB list we looked with particular care at the chemical shift information of amide N, CA, and CB.

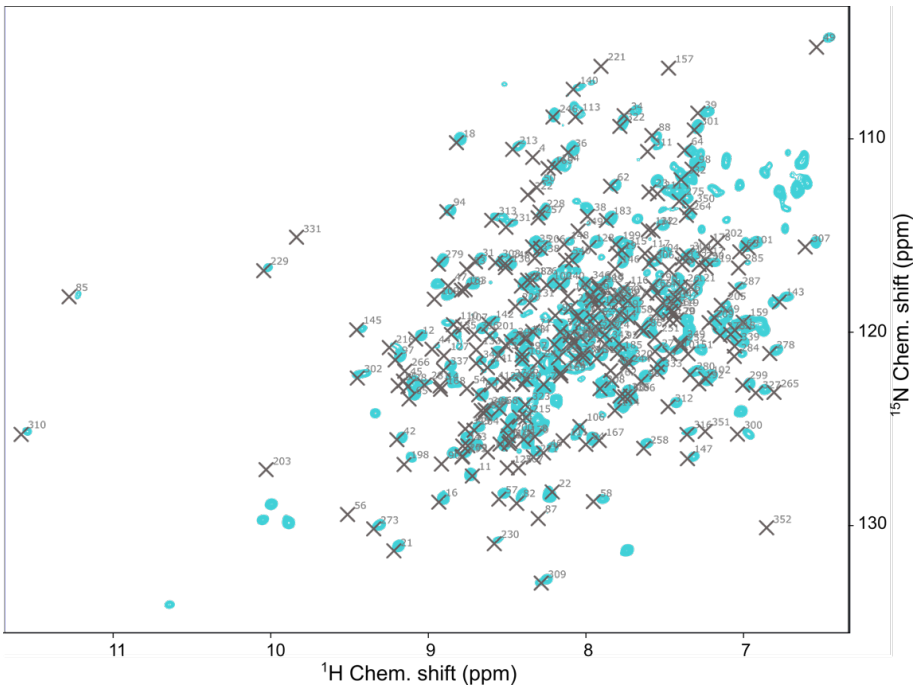


Figure 2.17. Superposition of peak list from BMRB entry (17471) with the acquire ^{15}N-^{1}H TROSY HSQC spectrum. The gray crosses represent the peak assignments from the BMRB entry, the blue peaks correspond to the spectrum acquired.

As a first step for the transfer of the assignment process, we examined the correlation of the chemical shift for the amide proton and nitrogen by overlapping the chemical shifts from the BMRB entry 17471 with our experimental ^{15}N-^{1}H TROSY HSQC (**Figure 2.17**). We observe a good agreement for the two data sets. However, given the number of expected residues, it is difficult to correctly transfer the assignments in the central region of the spectrum, where we

observe a high level of overlapping. The chemical shift information of the CA and CB was used to obtain a better resolution and cross-check the type of residue. By comparing this information, we transferred 94% of the assignments from the BMRB entry 17471.

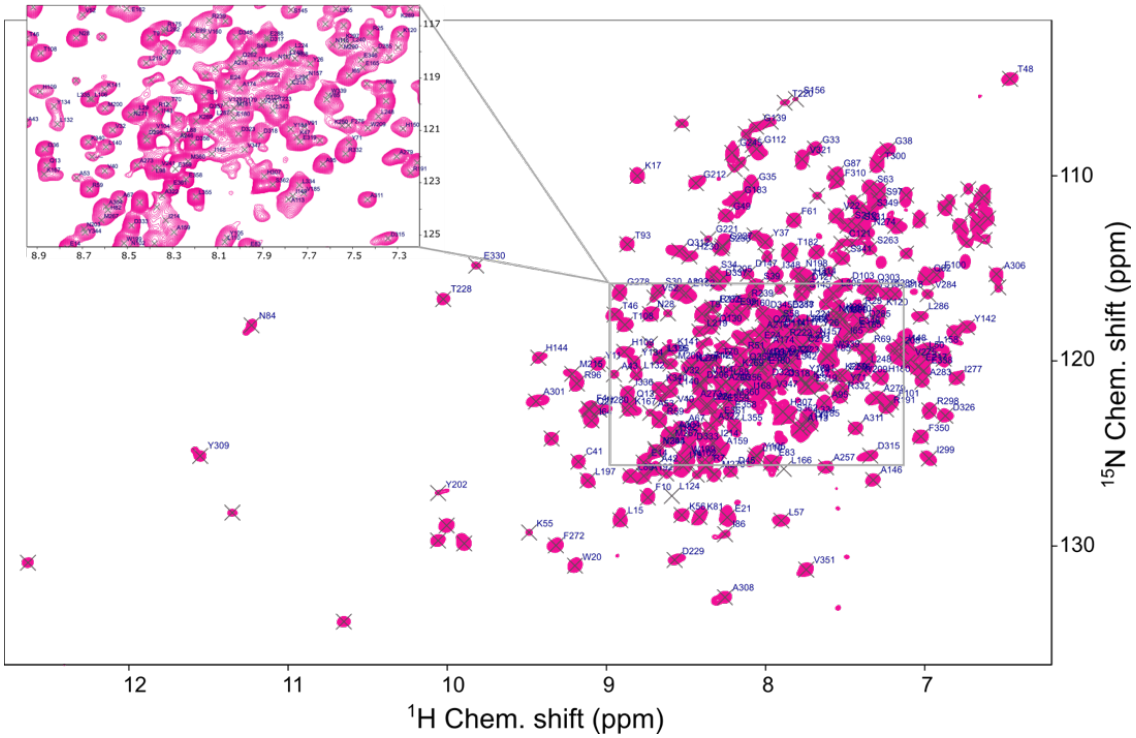


Figure 2.18. ^{15}N-^{1}H TROSY HSQC of p38α. 412 µM p38α in 50 mM HEPES, 150 mM NaCl, 5 mM DTT, pH 6.8 at 20°C in an 800 MHz (^{1}H Larmor frequency) spectrometer.

The comparison between the chemical shift from the BMRB entry and our peak list shows a good correlation for all types of chemical shift, except amide protons (**Figure 2.19**). Correlation values give a good level of confidence for the assignment process.

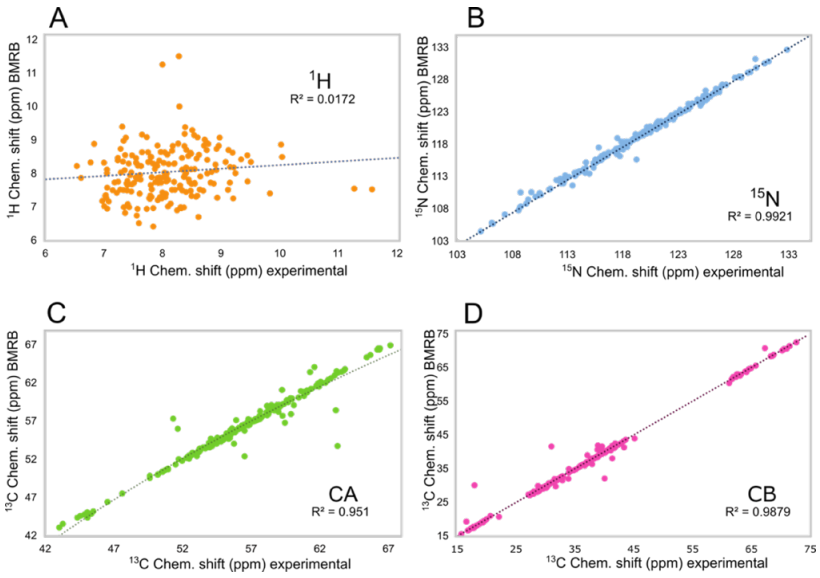


Figure 2.19. Correlation between the chemical shift from BMRB (17471) and assigned chemical shift for the p38α apo sample. **A)** Amide proton chemical shift correlation. **B)** 15N chemical shift correlation. **C)** Cα chemical shift correlation. **D)** Cβ chemical shift correlation.

2.2.2 Effect of inhibitors

To further understand the role of the allosteric binding site and the lipid pocket on the dynamic network of the protein, we analyzed the protein in complex with two allosteric binding site inhibitors (sorafenib and regorafenib) and in the presence or absence of a molecule in the lipidic pocket (Octyl β-glucoside).

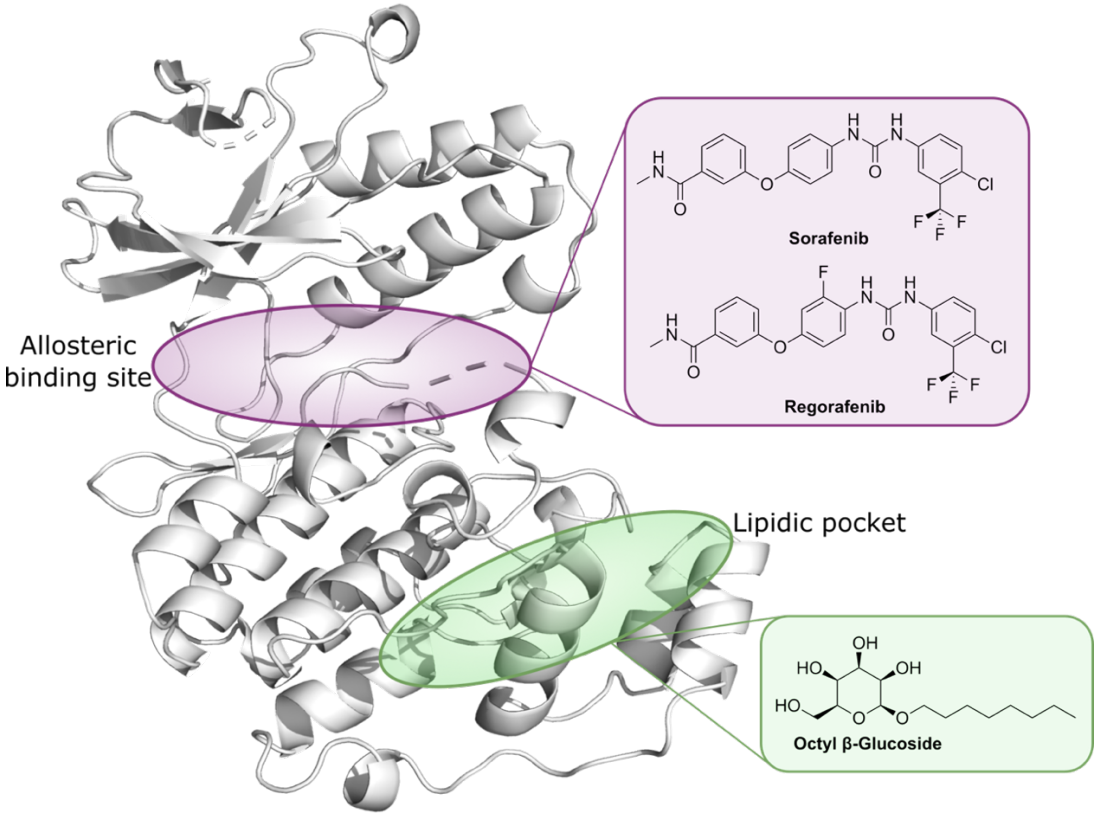


Figure 2.20. Structure of p38α indicating the allosteric binding site and the lipid pocket, with the ligands used for NMR analysis.

2.2.2.2 Chemical Shift perturbation

When we look at the effect on the chemical shift of inhibitors that occupy the allosteric binding site, we observe significant perturbations for those residues that are in close contact with the binding molecules, but we also observe some significant changes for some residues in the C-lobe, close to the lipidic pocket (**Figure 2.21**). Chemical shift perturbations can be calculated using the equation

(2.1)

$$CSP = \sqrt{(\Delta H)^2 + \left(\frac{\Delta N}{10}\right)^2}$$

where **Δ**H and **Δ**N are the difference between the chemical shift of the respective nucleus. Given the similarity between sorafenib and regorafenib, significant changes will not be expected when the chemical shift perturbation between these two samples is analyzed. However, we can observe in **Figure 2.21.C** some considerable perturbations along the whole sequence. This gives us a first hint into the allosteric communication that happens throughout the protein.

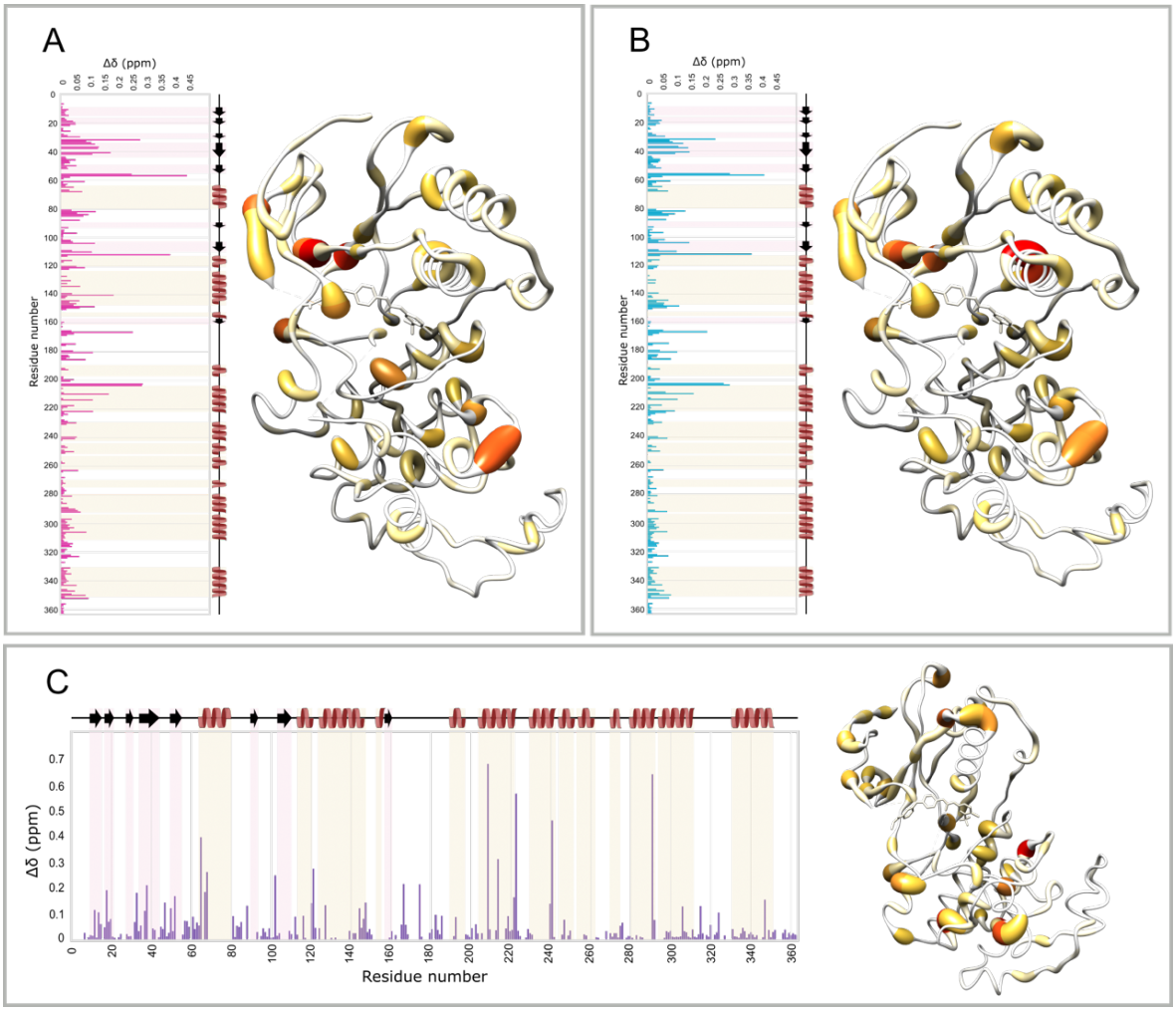


Figure 2.21. Chemical shift perturbations (CSP) for the allosteric binding site occupied by different inhibitors. **A)** CSP between p38α apo and p38α in complex with Sorafenib, with the perturbations shown in the X-ray structure. **B)** CSP for p38α apo compared to the complex p38α and Regorafenib, with the perturbations shown in the X-ray structure. **C)** CSP of the complex with Sorafenib and Regorafenib, the perturbations are shown in the X-ray structure.

In the same way that we analyze the effect of the inhibitors that bind to the allosteric binding site, we can study the effect that the lipidic pocket exerts on the rest of the molecule. For this purpose, we compare the protein apo with the protein with BOG and in complex with sorafenib and BOG, together. When we look at the chemical shift perturbation of the apo protein and p38α in the presence of BOG (**Error! Reference source not found..A**), we see significant p

erturbations for the expected residues, which are part of the lipid pocket. However, we can see some perturbation on the N-lobe.

We can assess the effect of the lipidic pocket by comparing the CSPs between the apo protein and the protein in complex with BOG, which show the expected perturbation in the region altered by the introduction of the BOG molecule, but significant perturbations can be observed as well in the N-lobe region. A second question that arises is whether the allosteric pocket and the lipidic pocket have a direct effect on each other. To analysis the effect this perturbation will cause, we compared the protein with both pockets occupy and made each pocket free at a time. In both cases, CSPs can be observed not only closed to the pockets, but also along the entire protein. Which hints on the allosteric communication between the two sites.

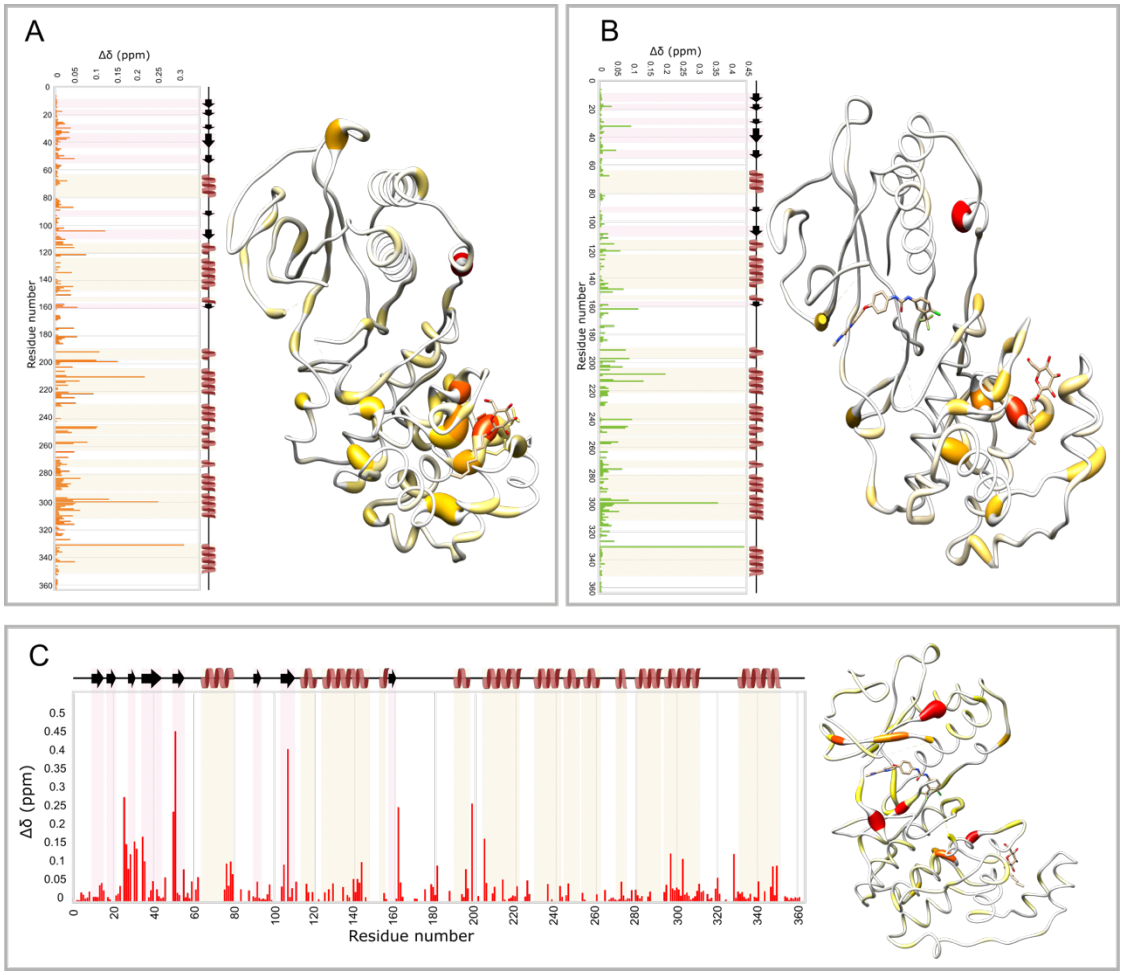


Figure 2.22. Chemical shift perturbations (CSP) for the free and occupied lipid pocket. **A)** CSP between p38α apo and p38α in the presence of BOG with the perturbation shown in the X-ray structure. **B)** CSP for p38α in complex with sorafenib and BOG compare with p38α in complex with sorafenib alone, having a full lipidic pocket and occupied while the allosteric binding site is always occupied. The perturbations are shown in the X-ray structure. **C)** CSP of p38α in complex with sorafenib and BOG compared with p38α in the presence of BOG, the perturbations are shown in the X-ray structure.

2.2.2.3 Relaxation experiments

As mentioned in Section 1.4, proteins are dynamic systems with processes that occur at different timescales. In solution NMR we can study the fast dynamics (picosecond to nanosecond timescales) with the standard experiments for the longitudinal relaxation rates (R_1), transverse relaxation rates (R_2), and heteronuclear steady state (*ssNOE* or *heteronuclear NOE*). Additionally, we can assess intermediate time scale dynamics in the µs-ms timescale, which is modulated by the transverse relaxation times of the nuclei and can be detected via changes in R_2 or the rotating frame relaxation rate ($R_{1\rho}$) as a function of the averaging process, known as relaxation dispersion. Among the most used relaxation dispersion techniques, we find Carr-Purcell-Meiboom-Gill spin-echo (CPMG) experiments and $R_{1\rho}$ spin lock experiments [103-105].

We used ^{15}N CPMG relaxation dispersion to access the dynamical processes occurring after binding to allosteric inhibitors and the changes related to lipid pocket occupancy. Constant-time ^{15}N CPMG-TROSY experiments were acquired using [^{15}N,^{2}H] p38α labeled in freshly exchanged NMR buffer (50 mM HEPES, 150 mM NaCl, 5M DTT) at a final concentration of ~400 µM. The experiments were recorded in an interleaved manner as follows for the CPMG frequencies: 0, 2000, 25, 50, 1500, 100, 1000, 250, 500, 750, 1250, 25, 100, 250 Hz. We used three repetition frequencies for error calculation, as well as a CPMG delay of total 0.022 s and a recycle delay of 1.5 seconds. The experiments were analyzed with NESSY software [106]

The ^{15}N constant-time CPMG-TROSY experiment detects the dynamics via changes in R_2 rates at different CPMG frequencies, we can extract the ^{15}N R_2 rates for the different samples. Considering that T_2 transverse relaxation times are sensitive not only to the general rotational correlation but also to the internal flexibility of individual segments in the secondary structure of the protein; we can use the individual R_2 rates to identify variations in the fast motion of every amino acid when changes are induced in the protein.

We observed changes in rates when p38α is in the apo state and when sorafenib or regorafenib are introduced as allosteric binders. In **Error! Reference source not found.** we compare the R_2 rates for the three samples and compare the differences for the individual residues. We can appreciate for many residues a slight increase in the rates when the ligands are present; this is expected behavior when we consider that the effect of the ligands is to prevent conformational exchange for the DFG motif and the hinge loop to impede the phosphorylation of the kinase.

The difference between the rate of the complex with sorafenib and regorafenib is shown in **Error! Reference source not found..C**, where it can be observed that some of the significant c hanges are located in the C-lobe, in regions considerably far from the allosteric binding site and in the region affected by the change in the structure between both ligands.

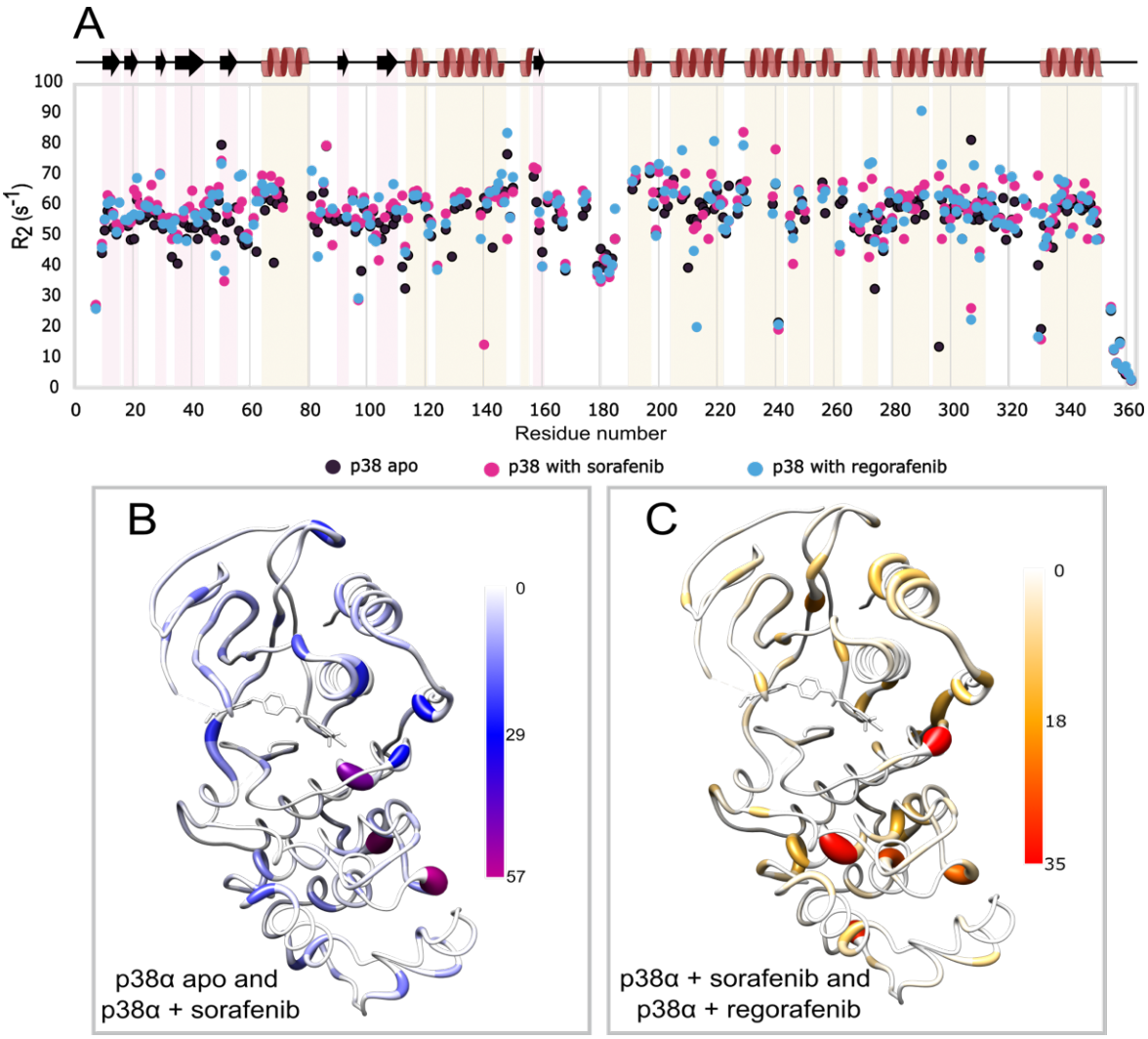


Figure 2.23. Comparison of R_2 rated for p38α apo and in complex with sorafenib and regorafenib. **A)** R_2 rates for p38α apo (black), p38α complex with sorafenib (pink), and p38α complex with regorafenib (blue) along the sequence compared to the secondary structure from the X-ray structure (PDB 3HEG). **B)** Difference between the R_2 rates of p38α apo and p38α when in complex with sorafenib shown in the X-ray structure (PDB 3HEG), the scale of the difference and the color code are indicated in the figure. **C)** Difference between the R_2 rates of p38α complex p38 with sorafenib and p38α in the complex with regorafenib plotted in the X-ray structure (PDB 3HEG), the scale of the differences and the color code is shown in the figure.

As the R_2 Relaxation rates are modulated by the field strength of the CPMG (ν_{CPMG}), the effective transverse relaxation rate R_2^{eff} can be calculated from the series of ^{15}N CPMG TROSY relaxation dispersion experiments using equation (2.2) [106]:

(2.2)

$$R_2{}^{eff} = \left(\frac{1}{\tau_{CP}}\right) In \left(\frac{I(0)}{I(\nu_{CPMG})}\right)$$

Where τ_{CP} is the constant CPMG time (0.022 s), $I(0)$ is the intensity of the peak in the reference spectrum and $I(\nu_{CPMG})$ is the intensity of the peak at the CPMG frequency, ν_{CPMG}. By fitting the difference between the measured R_2^{eff} over the range of ν_{CPMG} an exchange model can be obtained. The models describe residues that are involved or not in an exchange process and can be divided into three exchange models: model 1, no exchange, model 2, two-site fast; and model 3, two-site slow. For the first model is seen that the R_2^{eff} The rate is equal to the effective transverse relaxation rate at infinite ν_{CPMG} (equation (2.3)).

(2.3)

$$R_2{}^{eff} = R_2{}^0$$

The two-site exchange between the states A and B can be described by the equation

(2.4)

$$A \underset{k_{b\text{-}a}}{\overset{k_{a\text{-}b}}{\rightleftharpoons}} B$$

Here, k_{a-b} and k_{b-a} are the forward and backward exchange rates between states A and B. Model 2 describes fast-limit exchange, where chemical/conformational exchange is greater than the chemical shift difference between states ($K_{ex} \gg \delta\omega$)[107]. This model can be fitted to:

(2.5)

$$R_2{}^{eff} = R_2{}^0 + \frac{\Phi}{k_{ex}}\left[1 - \frac{4\nu_{CPMG}}{k_{ex}}\tanh\left(\frac{k_{ex}}{4\nu_{CPMG}}\right)\right]$$

Where $\Phi = p_a p_b \delta\omega^2$, and p_a and p_b are the populations of the two state models (p_a is the major conformation,($p_a + p_b = 1$), k_{ex} is the chemical/confirmational exchange ($k_{ex} = k_{a-b} + k_{b-a}$) constant and $\delta\omega$ is the chemical shift difference between states and $p_a p_b \delta\omega^2$.

The model 3 describes the slow-limit exchange for which $K_{ex} \ll \delta\omega$ according to the Richard-Carver equation [108]:

(2.6)

$$R_2{}^{eff} = R_2{}^0 + \frac{k_{ex}}{2} - \nu_{CPMG} cosh^{-1}[D_+ \cosh(\eta_+ - D_- \cos(\eta_-)]$$

In which

$$D_\pm = \frac{1}{2}\left[\pm 1 + \frac{\psi - 2\delta\omega^2}{(\psi^2 + \zeta)^{1/2}}\right]^{1/2}$$

$$\eta_\pm = \frac{\tau_{cp}}{2}\left[\pm\psi + (\psi^2 + \zeta^2)^{1/2}\right]^{1/2}$$

$\psi = {k_{ex}}^2 - \delta\omega^2$ and $\zeta = -2\delta\omega(p_a k_{ex} - p_b k_{ex})$

The fitting of the results is represented by three different types of plots. Model 1 (no exchange) shows a flat profile with no changes for the R_2^{eff} rates for the different CPMG frequencies, while for Model 2 (fast exchange) it can be observed a linear decay of the rate can be observed when the CPMG frequency is increased, the difference between the R_2^{eff} rate at low and high frequency can be fitted into the R_{ex} rate, and it is automatically calculated by NESSY as: $R_{ex} = p_a p_b \delta\omega^2 / k_{ex}$. On the other hand, Model 3 shows an exponential decay of the R_2^{eff} rate, and the R_{ex} rate is calculated by NESSY following the relation: $R_{ex} = p_a p_b - k_{ex} / (1 + \left(\frac{k_{ex}}{\delta\omega^2}\right))$.

This process was used for p38α apo, as well as for the complex with the different ligands. We can observe, as an example, the results obtained for R12, V129, and L120. R12 shows no dispersion in the R_2^{eff} rates, which indicate the absence of chemical / structural exchange. For V129, the R_2^{eff} rates decrease linearly along all the recorded CPMG frequencies. This can be fitted for model 2, indicating the presence of a fast chemical exchange process. In the case of L120 we observe the dispersion of the R_2^{eff} rates decreasing fast along the lower CPMG frequencies, which after fitting with Model 3 shows the presence of slow chemical exchange. Many of the residues are fitted to model 1, indicating the absence of chemical exchange. The absence of chemical exchange is only considered for dynamic events that occur on the intermediate timescale (µs-ms). For events occurring at a different timescale, either faster (nanoseconds to picoseconds) or slower (ms to second), it is necessary to complement the analysis with other methodologies.

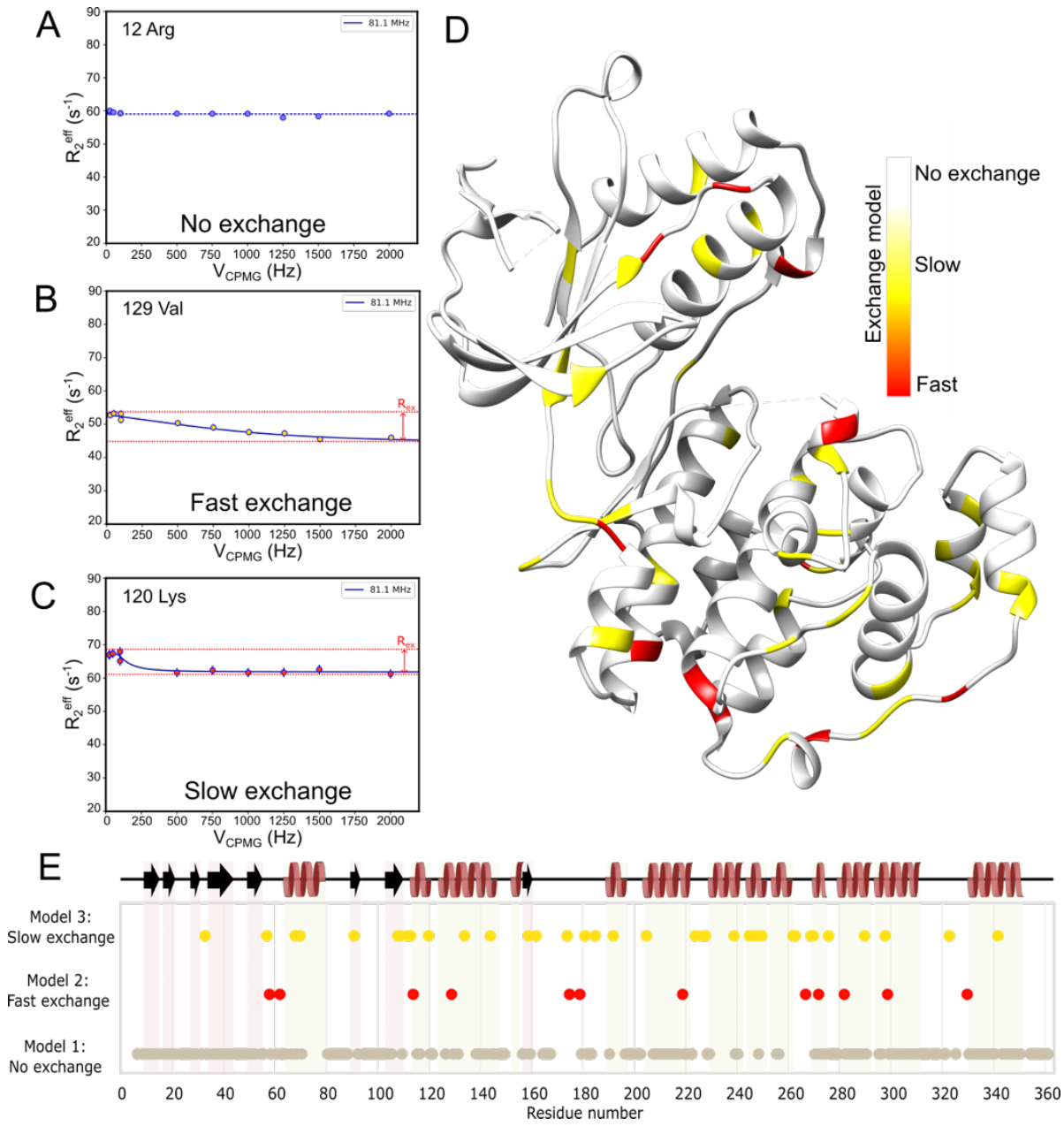


Figure 2.24. ^{15}N CPMG TROSY model fit for possible exchange processes. **A)** Fitting for the no exchange model where the R_{2eff} at the lower frequency, there are no changes compared with the last R_{2eff} at highest frequency. **B)** Example of a residue (V129) with fast exchange, where the R_{ex} is illustrated by the red lines. **C)** Fitting example for a residue (K120) that presents a fast exchange, where the R_{ex} is illustrated by the red lines. **D)** Graphical representation of the fitted models for p38α apo depicted in the X-ray structure (PDB entry: 1WFC), where residues are colored according to the fitted model, so model 1 is represented in white, model 2 is colored in red and model 3 is shown in yellow. **E)** Plot representation of the model fitted by residue along the sequence. Model one is shown in olive, model two is colored in red, and model three is presented in yellow.

As a first look at the protein in the apo form, we identify the residues that are involved in dynamic events occurring on the µs-ms timescale. From the ^{15}N CPMG TROSY analysis, we can look at the R_{ex} rates for all residues fitted with exchange in a two-site model. We can map these dynamic sites in the protein apo as a control to monitor changes upon ligand binding.

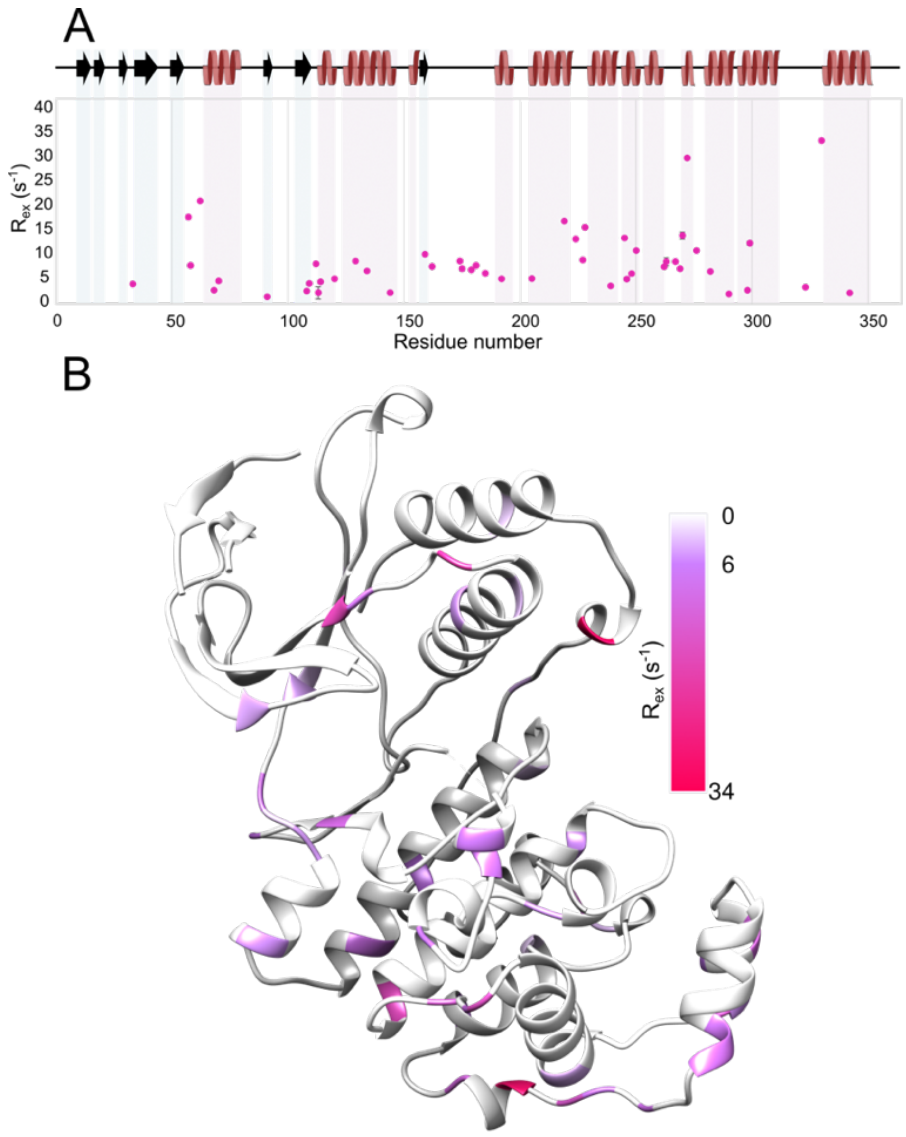


Figure 2.25. R_{ex} rates obtained from the ^{15}N CPMG TROSY analysis for the residues involved in chemical exchange of p38α apo represented along the sequence (**A**) and in the X-ray structure (**B**) where the rates correspond to the intensity of the color.

To understand the changes that occur when sorafenib and regorafenib are in complex with p38α we can compare the chemical exchange changes that occur for the two complexes between each other and with the protein in the apo form. First, we briefly look at the fitted model for every residue. In this comparison, most of the residues show no motion in this timescale, as they are fitted to model 1. For the residues that present chemical exchange, we observe changes between the type of model fitted (**Error! Reference source not found.**). We can look at in m ore detail the changes that occur in the protein when sorafenib is present at the allosteric binding site. **Error! Reference source not found..C** shows the comparison between the R_{ex} rates for p38α apo p38α and in the complex, where the most significant changes are observed for residues close to the allosteric binding site. The observed changes can be attributed to the effect

of the inhibitor over the protein, which rigidifies the system upon ligand binding, forbidding the conformational changes necessary for the protein to experience to complete the function.

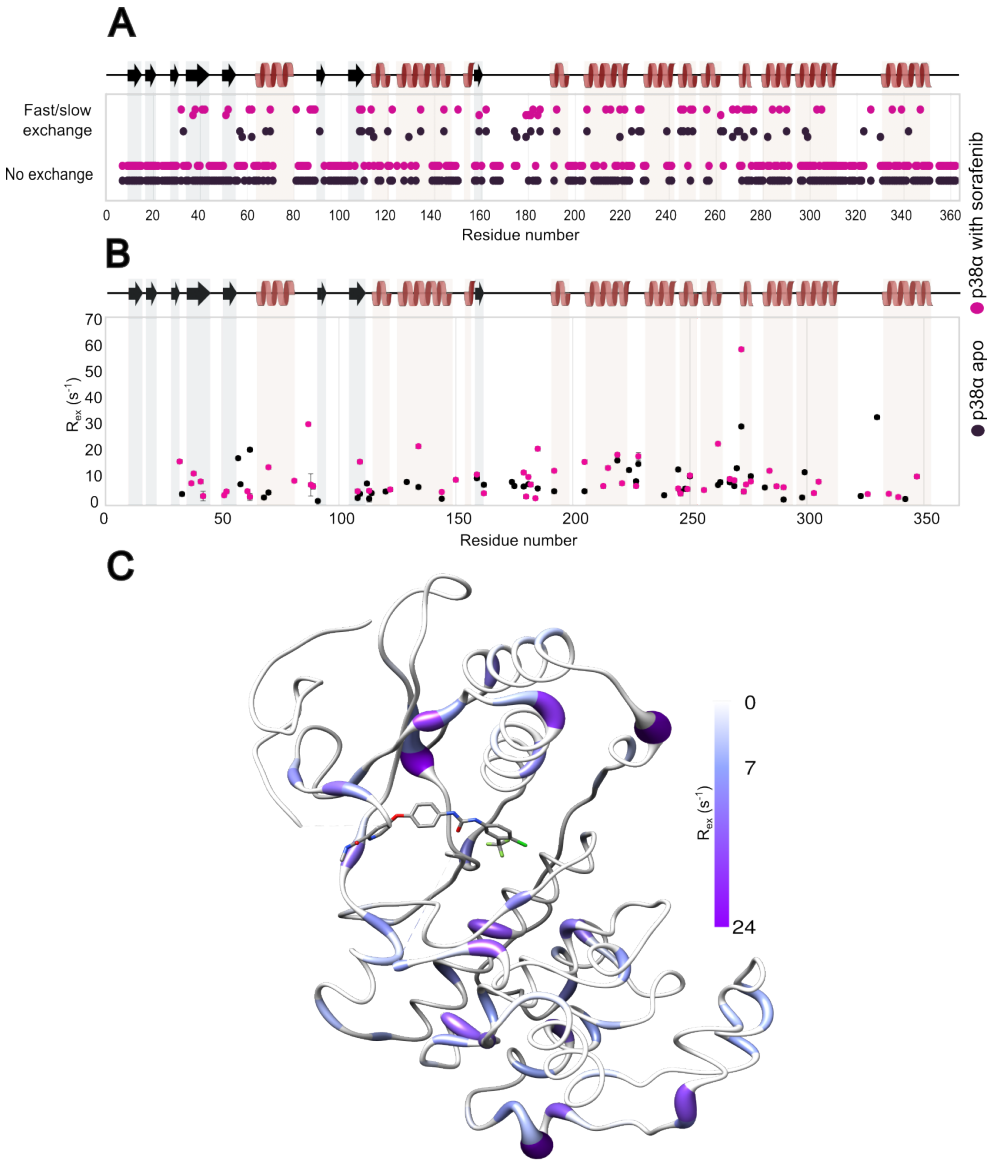


Figure 2.26. Conformational exchange between two sites for p38α apo and p38α in complex with sorafenib. **A)** Comparison of the model fitted by residues along the sequence for p38α apo (black) and p38α in complex with sorafenib (pink). **B)** Comparison of the R_{ex} values obtained for p38α apo (black) and p38α in complex with sorafenib (pink). **C)** Difference of the R_{ex} between the two samples plotted in the X-ray structure (PDB entry 3HEG), where the color code is indicated alongside the structure.

The same approach can be used to analyze the difference between sorafenib and regorafenib. As mentioned in Section 2.2.267, the structural difference between sorafenib and regorafenib is minimal and, therefore, minimal changes are expected. For some of the residues that show conformational exchange in sorafenib, this exchange is not appointed when regorafenib is present. For the majority of residues in conformational exchange in both samples the R_{ex} rate did not change significantly. However, the change in Val32 can be highlighted, for which the

R_{ex} decreases from 16.64±0.633 in the regorafenib sample to 3.26±0.312 when regorafenib is present. This residue is located in the loop region close to the fluorine group introduced in regorafenib. However, the graphical observation of the precise location is difficult to assess as this loop is not visible in the X-ray structure.

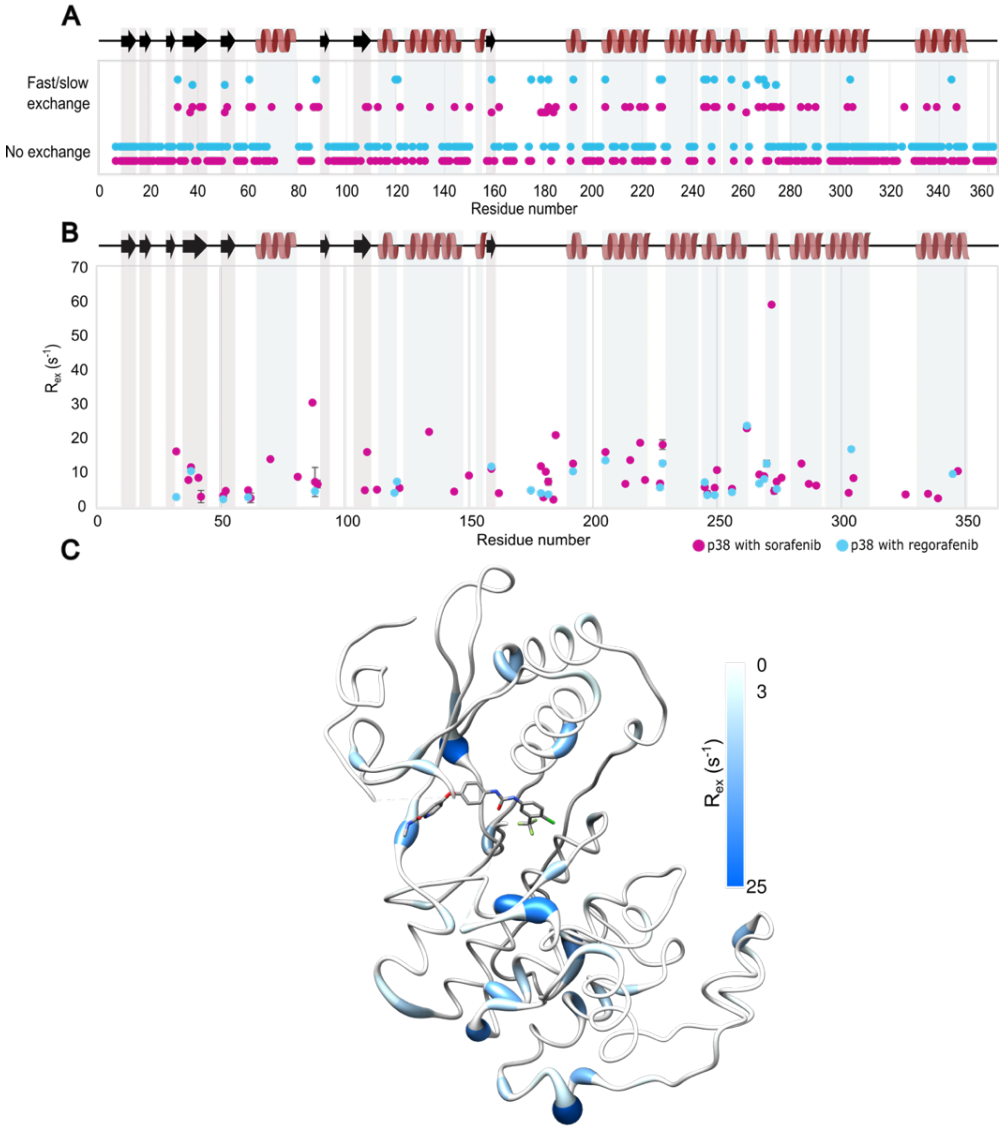


Figure 2.27. Conformational exchange between two sites for p38α in complex with sorafenib and p38α in complex with regorafenib. **A)** Comparison of the model fitted by residues along the sequence for p38α with sorafenib (pink) and p38α in complex with regorafenib (blue). **B)** Comparison of the R_{ex} values obtained for p38α with sorafenib (pink) and p38α in complex with regorafenib (blue). **C)** Difference of the R_{ex} between the two samples plotted in the X-ray structure (PDB entry 3HEG), where the color code is indicated alongside the structure.

Another aspect that can be analyzed is the effect of the lipidic pocket on the protein. It is known that the pocket can be filled with Octyl β-Glucoside (BOG), but it is not clear what effect this molecule has when it bounds the protein. When we compare the ^{15}N CPMG TROSY results from the p38α apo with p38α in complex with BOG we can see that for many residues there is no conformational exchange. In **Figure 2.28**, for many residues that show conformational

exchange in the apo form, especially in the region between residues 235 to 280 (highlighted in a red dotted rectangle), we can observe that the chemical exchange in the µs-ms timescale is not there when BOG is present, an expected and reassuring observation. However, importantly, it can be highlighted that we also observe changes in the N-lobe, close to the allosteric binding site, and the hinge, again referring to the presence of allosteric communication through the protein.

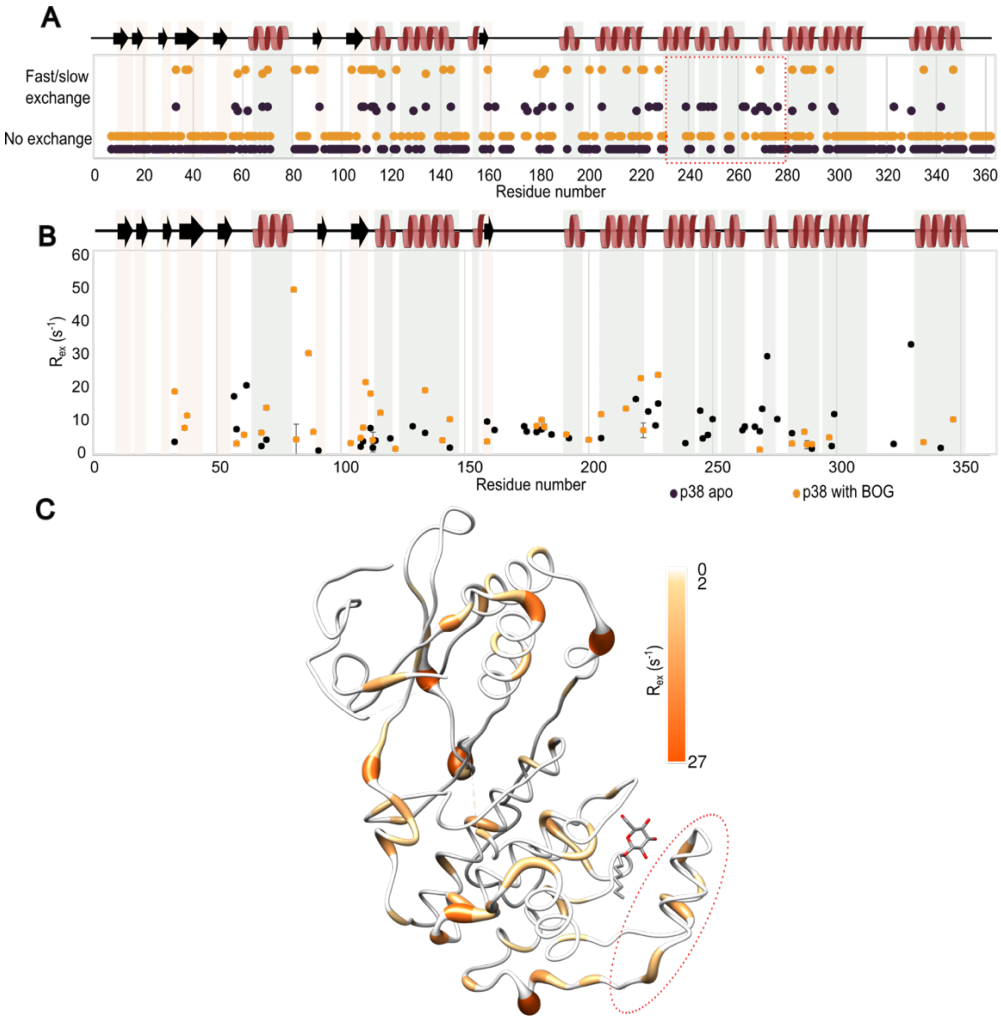


Figure 2.28. Conformational exchange between two sites for p38α apo and p38α in presence of BOG. **A)** Comparison of the fitted model by residues along the sequence for p38α apo (black) and p38α with BOG (orange). **B)** Comparison of the R_{ex} values obtained for p38α apo (black) and p38α in complex BOG (orange). **C)** Difference of the R_{ex} between the two samples plotted in the X-ray structure (PDB entry 3GCS), where the color code is indicated alongside the structure. The red dashed lines highlight the residues involved in conformational exchange in the apo protein but does not experiment conformational changes when the BOG is present

To understand the connection between the two sites, the allosteric binding site and the lipidic pocket, we can analyze the effects of a change in the lipidic pocket when the allosteric binding site is permanently occupied. For this purpose, we compare two samples: p38α in complex with sorafenib and p38α in complex with sorafenib and BOG. When BOG is absent from the complex, some residues change the time scale of motion, and dynamics events no longer occur

in the µs-ms regime. Some interesting examples are Val85 and Ile348. Val85 is at the allosteric binding site and shows conformational exchange when sorafenib is present, but conformational exchange is not observed when the lipid pocket is occupied by BOG. In the case of Ile348 when the lipidic pocket is free, there is no conformational exchange on the µs-ms timescale. However, when BOG is present in the lipidic pocket, we observe conformational exchange at around 100 ms. There are other residues, such as Ser39, Asp323, Arg332, Ile336, and Tyr344, that present motion on this time scale when sorafenib is present and the lipidic pocket is free, but the dynamic process becomes either slower or faster when BOG occupies the lipidic pocket. Other similar changes can be observed throughout the protein.

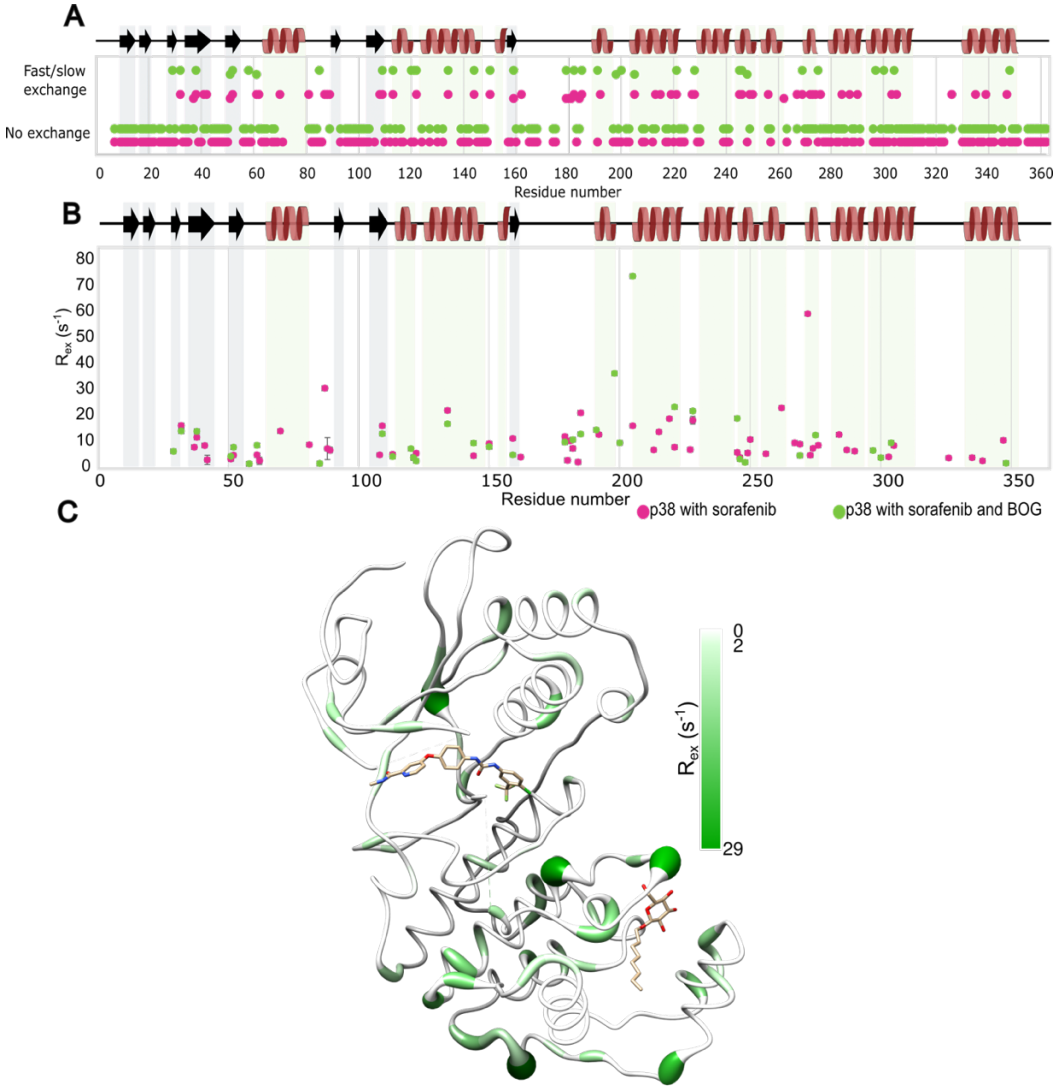


Figure 2.29. Conformational exchange between two sites for p38α in complex with sorafenib and BOG and p38α in complex with sorafenib only. **A)** Comparison of the fitted model by residues along the sequence for p38α in complex with sorafenib only (pink) and p38α with sorafenib and BOG (green). **B)** Comparison of the R_{ex} values obtained for p38α with sorafenib only (pink) and p38α in complex with sorafenib and BOG (green). **C)** Difference of the R_{ex} between the two samples plotted in the X-ray structure (PDB entry 3GCS), where the color code is indicated alongside the structure.

2.3 Solid-state NMR

2.3.1 Different Samples

Solid-state NMR spectra have a broader linewidth in comparison with that of solution. The linewidth is dependent on different factors; some of them like the spinning speed of the level of deuteration can be controlled. Fast MAS frequencies depend on the available instrumentation; recent technological progress tackling the problem of faster probes and smaller rotors had increased the spinning frequencies up to 160 kHz. We can observe a decrease in line width as a function of the spinning speed (**Error! Reference source not found.**), when w e go for faster spinning frequencies, we can observe a better resolution due to the decrease of line width.

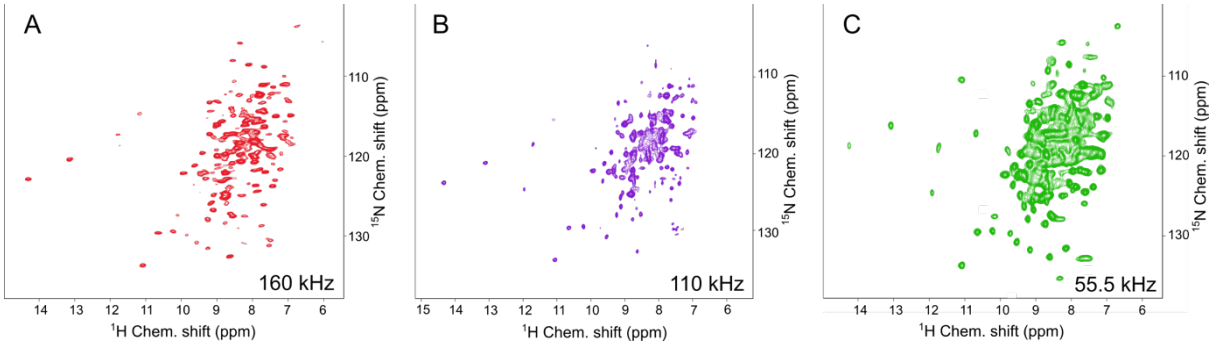


Figure 2.30. MAPK p38α ^{15}N-^{1}H hNH correlations showing different linewidth in function of spinning frequency. **A**) hNH correlation at 160 kHz spinning frequency. **B**) hNH correlation at 110 kHz spinning frequency. **C**) hNH correlation at 55.555 kHz spinning frequency.

The dilution of the proton network with deuterium brings many benefits in improving the resolution and linewidth of our spectrum. But in the process, we can lose important information that side chain protons carried. The use of strategies such as reduce adjoining protonation (RAP) [109], fractional deuteration (FD) [110], and inverted fractional deuteration (iFD) [79], allows us to have a certain level of deuteration in the side chain while accessing the information of the desired protons and reducing the problems with back-exchange of amide protons in the hydrophobic regions. By using these labeling scheme techniques, we increase the linewidth, but at 55.5 kHz the resolution seems to maintain a good compromise with respect to the information that can be obtained by the introduction of selective protons.

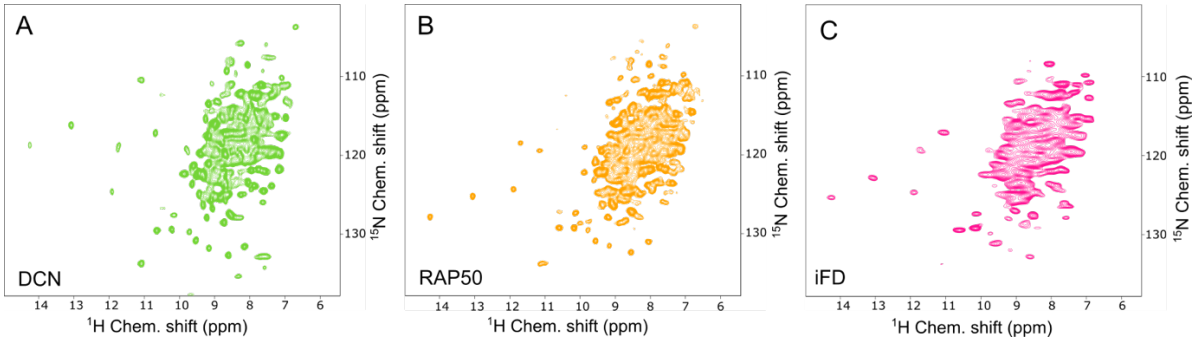


Figure 2.31. MAPK p38α ^{15}N-^{1}H hNH correlation showing different linewidth in function of deuteration level. **A)** MAPK p38α fully deuterated. **B)** MAPK p38α grown with 50% D_2O and 50% H_2O as medium solvent. **C)** MAPK p38α inverted Fractional Deuterated. All experiments were recorded in a 700 MHz (^{1}H Larmor frequency), 55.555 kHz spinning frequency at a temperature of ~25°C.

2.3.2 Assignment process

The development of ^{1}H detection at fast-MAS has enabled the acquisition of experiments that provide information similar to that in the traditional solution-state approach for protein assignment, but with the advantages that come with solid-state studies, i.e., no limit on the molecular size of the protein, access to multiple dimension experiments, use of different levels of deuteration without compromising linewidth, and others. MAPK p38α is close to the upper limit of the molecular size for solution NMR. Although part of the protein has been assigned in solution NMR, there are still regions of biological relevance that remain unassigned; this can be due to several reasons. Among the different reasons, back-exchange problems in hydrophobic regions can be present due to perdeuteration. It is observed that there are few assignments on the longer loop region; this can be attributed to exchange broadening because different conformations happening on a faster time scale can be measured by NMR.

By introducing the crystal lattice into the protein environment, we can expect a slowing down of some of the conformational exchange, and by using fast-MAS solid-state NMR in combination with different labeling methodologies and ^{1}H detection experiments, we can complete the assignment of p38α.

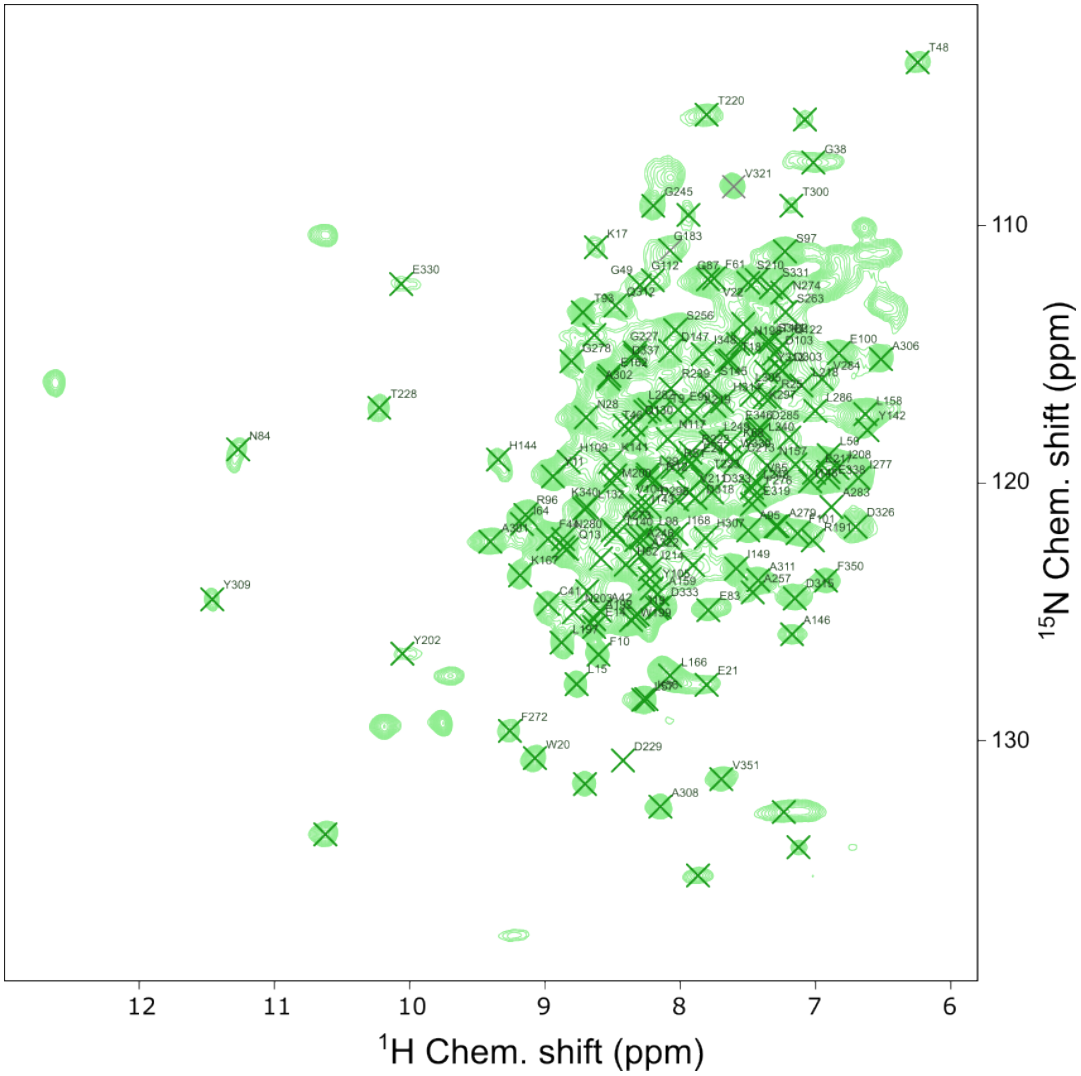


Figure 2.32. ^{1}H-^{15}N hNH spectrum of p38α fully deuterated recorded at 20°C with spinning frequency of 55.5 kHz in an 800 MHz (^{1}H Larmor frequency) spectrometer.

In a first look at the protein fingerprint in a 2D amide H-N correlation, we can appreciate that many of the amide peaks overlap around 6.5 to 9 ppm in the proton dimension and between 114 and 126 ppm in the nitrogen dimension. This level of overlap is expected for a protein of this size and number of residues, and for this reason, it is necessary to work with high-dimensional experiments. We recorded series of 3D and 4D experiments for protein assignment, these experiments contain information to connect residues through carbons (Cα and C=O), Hα protons, and amide-amide for the sequential backbone walk, also sufficient information to characterize the amino acid type based on the Cα, Cβ, and CO chemical shift information.

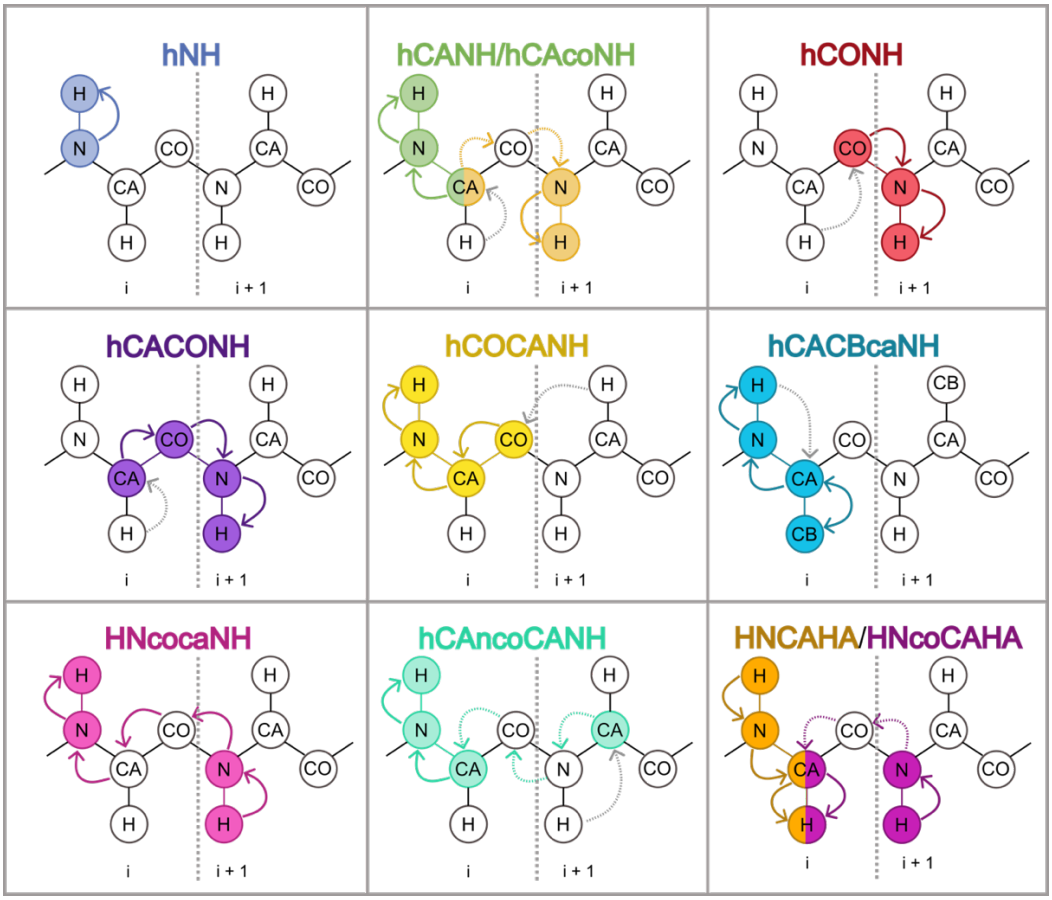


Figure 2.33. Schematic representation of the acquired experiments. The continues arrows indicate the magnetization transfer and the colored nuclei represent the atoms for which the chemical shift is evolved. The dashed arrow indicates the magnetization transfer through nuclei that did not evolve a chemical shift. For the sketches that contain two colors, each color represents a different experiment that follows a different magnetization pathway.

Traditionally, the assignment process of proteins by solution and solid-state NMR is carried out by 3D experiments, which is possible to use for proteins of low molecular size, where we have good dispersion of the peaks. However, by increasing the molecular size, proteins greater than 20 kDa show a high level of overlap, which becomes especially problematic in the solid state, where the line width of the peaks is larger than in solution. For these cases, it is necessary to increase the dimensionality of the spectra to resolve the overlapping problem. MAPK p38α is a clear example of this problem; with the 362 amino acids present, we observe a high level of overlap in the 2D spectrum. It is expected to have better resolution when working with three dimensions; however, for many of the residues, we can also find overlapping problems in the 3D experiments. By including an additional correlation, we can resolve most of the assignments. One example of this can be found in **Figure 2.34**: here we start from one position in the crowded region to the correspondent slide in the hCANH 3D where we find a single peak. But when we move to the 4D spectrum of hCACBcaNH, we can see the separation of

the three amino acids that share similar chemical shifts of H, N, and CA by including CB information for each residue. The assignment process can be performed manually or automated with the use of assignment software tools such as FLYA.

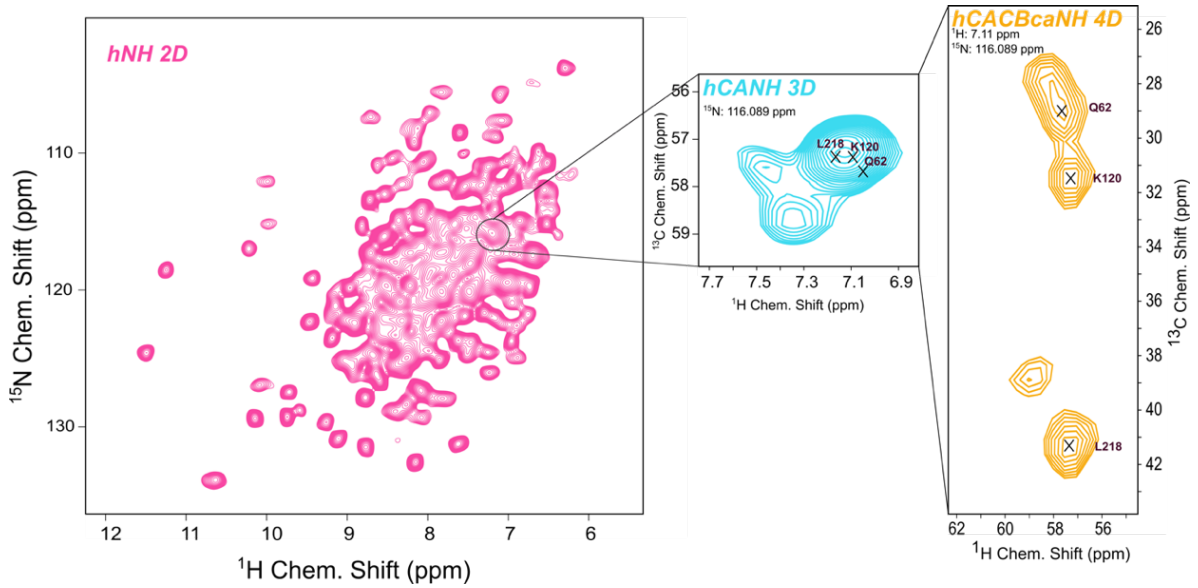


Figure 2.34. Representation of the resolution of the peak by increasing the dimensionality. In the 2D ^{1}H-^{15}H correlation it can be observe a high level of overlap, by increasing the number of dimensions overlap peaks can be resolved. The type of experiment corresponds to the color of the spectra. ^{15}N hNH is shown in pink, hCANH 3D is shown in blue, and hCACBcaNH 4D is orange.

2.3.2.2 Manual Assignment

Manual assignment was performed using the software ccpNmr AnalysisAssign [111]. hNH 2D, hCANH 3D, and hCONH 3D experiments were used as a basis for the assignment of the backbone. The connection of the residues and the sequential backbone walk, as well as the identification of the type of residue, was carried out with different 4D experiments. Direct connections from residue i to residue $i + 1$ can be made based on the chemical shift of H and N amide using the HNNH 4D experiment. Furthermore, another direct connection that includes the CA correlation between the H, N, and CA chemical shift of residue i with the CA chemical shift of the amino acid $i + 1$ can be obtained from the experiment hCACANH 4D. We can also include information on the CO chemical shift to the sequential backbone walk by using the experiments hCOCANH 4D and hCACONH 4D. The type of residue can be identified based on the combination of the Cα and Cβ chemical shift which can be obtained by hCACBcaNH 4D.

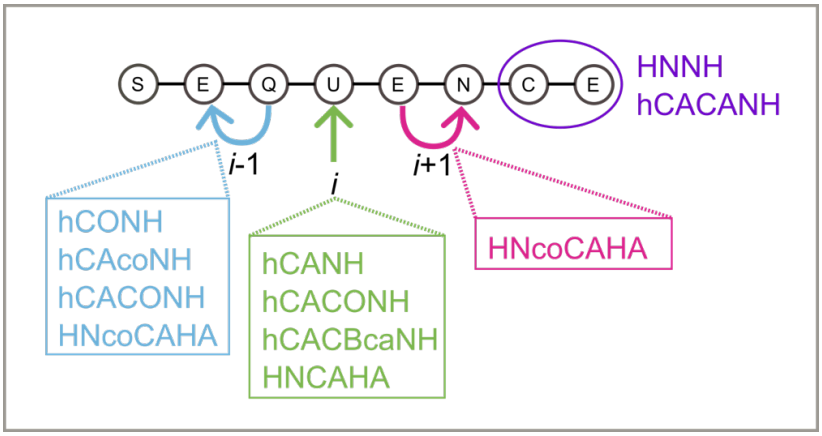


Figure 2.35. Schematic representation of the use of each experiment in the assignment process. The blue square shows the experiments that can be used for the $i-1$ backbone walk. The green square indicates experiments that provide intra-residual information. The dark pink square presents an experiment that can be used for the $i+1$ backbone walk. Purple square indicates experiments used for direct inter-residual connection. The colored circles indicate the chemical shifts that evolved for the experiments: dark blue, amide H-N; red, carbon of the carbonyl group; yellow, carbon alpha; light green, carbon beta; and light pink, proton alpha.

Protons are abundant in proteins with a high gyromagnetic ratio that enhances the spectral sensitivity, but also a powerful source of information for protein assignment. However, ^{1}H-detection techniques are usually performed in ^{1}H amide, and the amide protons are not the only nuclei that can be observed for this purpose. Hα protons carry valuable information related to the secondary structure of the protein, and they can be added to the set of assignment nuclei as another source of chemical shift to reduce ambiguity due to overlap. The use of ^{1}Hα detection-based experiments has been used as a successful strategy for IDP in solution NMR due to the insensitivity of ^{1}Hα to chemical exchange with the solvent and which are considered nonlabile protons irrespective of pH [112]. In the solid state it is also possible to use of ^{1}Hα-detection for assignment experiments, they can be implemented in fully protonated samples, [^{15}N,^{13}C] labeled, using fast spinning frequencies (110 kHz).

The ^{1}Hα detection-based HNCAHA 4D and HNcoCAHA 4D experiments were recorded for the crystallization of [^{15}N,^{13}C] p38α crystallized in the presence of sorafenib and BOG, the experiments were acquired using 1900 NUS points (3% NUS). Due to the total impossibility of suppressing the signal from the HEPES buffer, processing of the spectra was unsuccessful. As a first approach, we exchanged the buffer for deuterated HEPES (*d*-HEPES) to get a reading of the possible source of the signal; this buffer exchange was performed before crystallization using an Amicon® NMWCO filter 10 kDa cutoff. However, this attempt to decrease the buffer signal was not successful; moreover, after a couple of months, the crystals in *d*-HEPES redissolved.

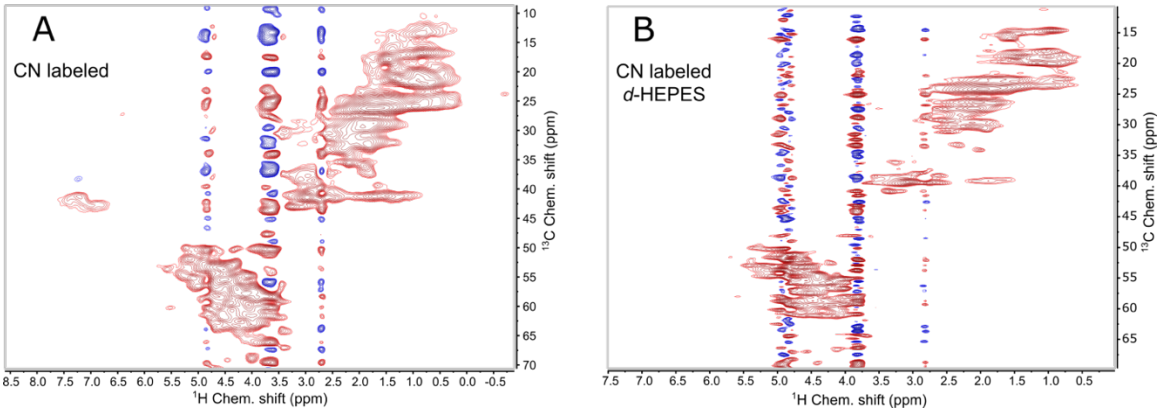


Figure 2.36. ^{1}H-^{13}C hCH spectra of [^{15}N,^{13}C] p38α with sorafenib and BOG in HEPES buffer (**A**) and *d*-HEPES (**B**) at 110 kHz in a 700 MHz (^{1}H Larmor frequency) spectrometer. In both spectra we can see strong solvent signals.

There are several methods to decrease the solvent signal by implementing a solvent suppression scheme in the pulse program, and one of the most efficient high-performance suppression methods is the following. Multiple Intense Solvent Suppression Intended for Sensitive Spectroscopic Investigation of Protonated Proteins, Instantly (MISSISSIPPI)[113]. To eliminate the HEPES signals, the carrier was changed to the HEPES signal during solvent suppression, and the length of the MISSISSIPPI scheme was increased to 200 ms (4x50 ms), using an RF saturation pulse of 25 kHz. By this slight modification, we obtained clean hNH and hCH spectrum, show **Error! Reference source not found..C.**

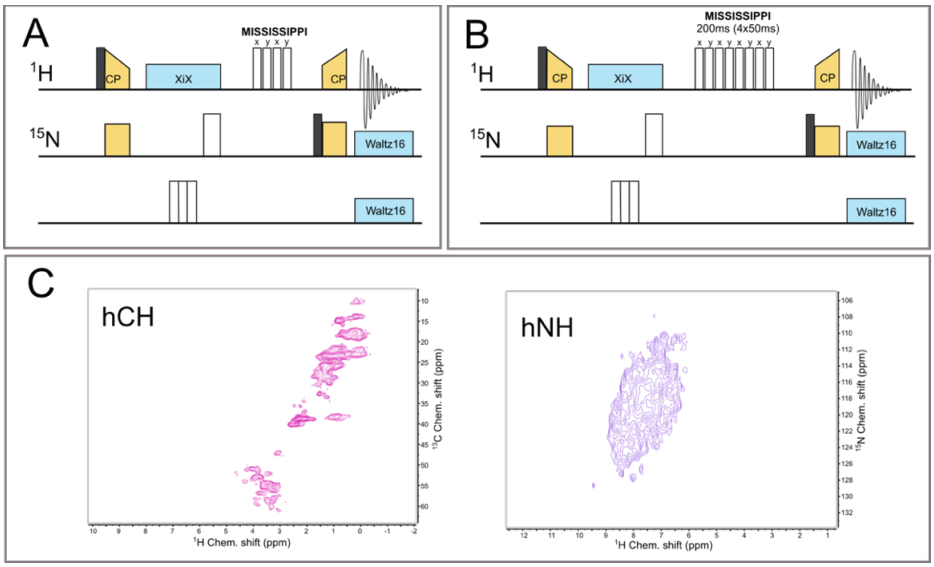


Figure 2.37. MISSISSIPI solvent suppression scheme in the context of heteronuclear correlation. **A)** hXH correlation pulse program with solvent suppression length of 80 ms and RF saturation pulse of 10 kHz. **B)** hXH correlation pulse program with modification applied to the solvent suppression scheme. **C)** ^{1}H-^{13}C (left) and ^{1}H-^{15}N (right) spectrum with modified MISSISSIPPI solvent suppression

With the HEPES solvent suppression problem resolved for p38α, it is possible to acquire assignment experiments that include ^{1}Hα. These experiments will be recorded in the near future.

2.3.2.3 Automated Assignment

The manual assignment of chemical shift can be considered possible when the number of residues is small; the larger the protein, the more complex this process can be. Even with enough information to connect residues or identify the amino acid type unambiguously, the manual process can take months or even a year of working time from an experienced spectroscopist to assign a protein [114]. The manual process becomes more complicated for assigning chemical shift for structural calculation, in those cases we are dealing with a large list of peaks coming from all the spatial correlation between the protons in the sidechain in addition to the backbone resonances and through bond correlations. There are many automated methods for resonance assignment, each with particularities depending on the application for which they have been developed [115]. The recent technological development in computer science has allowed the improvement of automated methods and helped to reduce the calculations time; all together, this has made the use of automated methos more attractive and reliable to overcome the bottleneck problem that chemical shift assignment represents in NMR studies. The Fully Automated NMR Structure Determination Algorithm (FLYA) generates a network of expected peaks from the protein sequence and magnetization transfer pathways and, following several computational calculation steps, assigns the chemical shifts from the provided experimental peak lists [115, 116]. The algorithm provides the user with a visual representation of correctness and reliability of individual assignments where every atom is represented by a color code: green, assignment by FLYA that agrees with a provided reference chemical shift within a tolerance of 0.03 ppm; red, assignment differs from reference; blue, assigned by FLYA but no reference available; black, with reference assignment but not assigned by FLYA. The respective light colors indicate assignments classified as 'unsafe' by chemical shift consolidation. [115]

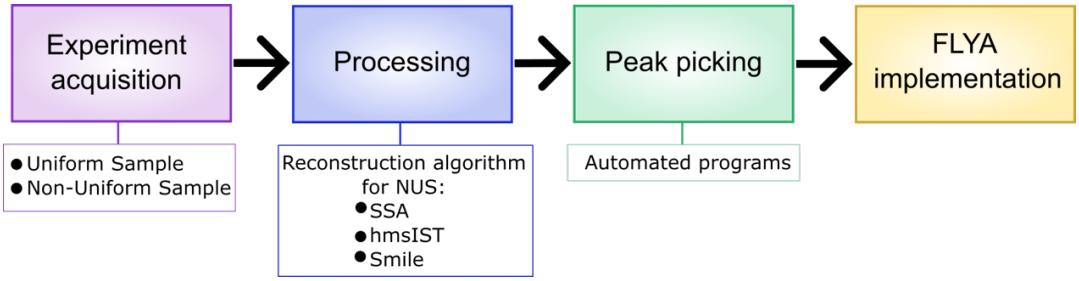


Figure 2.38. General process for FLYA implementation with the possible variable the user needs to consider.

The general process of FLYA implementation requires several steps of data preparation where a strong interaction is needed (**Error! Reference source not found.**). The two first steps, data a cquisition and data processing, are carried out in the same way as the user will do it for manual assignment. The peak picking and peak list preparation, although might seem a trivial process, is in many cases the determinant factor for a reliable assignment. This process might require several iterations and interventions, where it is important to keep in mind the importance of proper referencing among the spectra.

A first FLYA run was performed that included the information of all the acquire experiments without an external reference. The results obtained show weak assignments for all residues, with a few exceptions (**Figure 2.39**). A detailed analysis of the assignment shows a very low percentage of assignments of the 4D experiments.

Table 2.9. FLYA assignment results for p38α in the solid-state.

List	Expected*	Assigned**	Measured***	Assigned****
15N HSQC	345	319 (92.46%)	202	159 (78.81%)
hCANH	345	226 (65.52%)	391	126 (32.23%)
hCONH	344	248 (72.09%)	543	162 (29.65%)
hCACBcaNH	329	153 (46.50%)	449	94 (20.94%)
HNcocaNH	329	70 (21.28%)	394	54 (13.71%)
hCAcoNH	344	211 (61.34%)	739	141 (19.08%)
hCACONH	329	139 (42%)	629	104 (16.53%)
hCOCANH	329	142 (43.16%)	989	104 (10.52%)
hCAncoCANH	329	147 (44.68%)	807	100 (12.39%)

*Expected: total number of expected peaks
**Assigned: number of expected peaks that could be assigned
***Measured: total number of peaks in peak list
****Assigned number of measured peaks that could be assigned to expected peaks

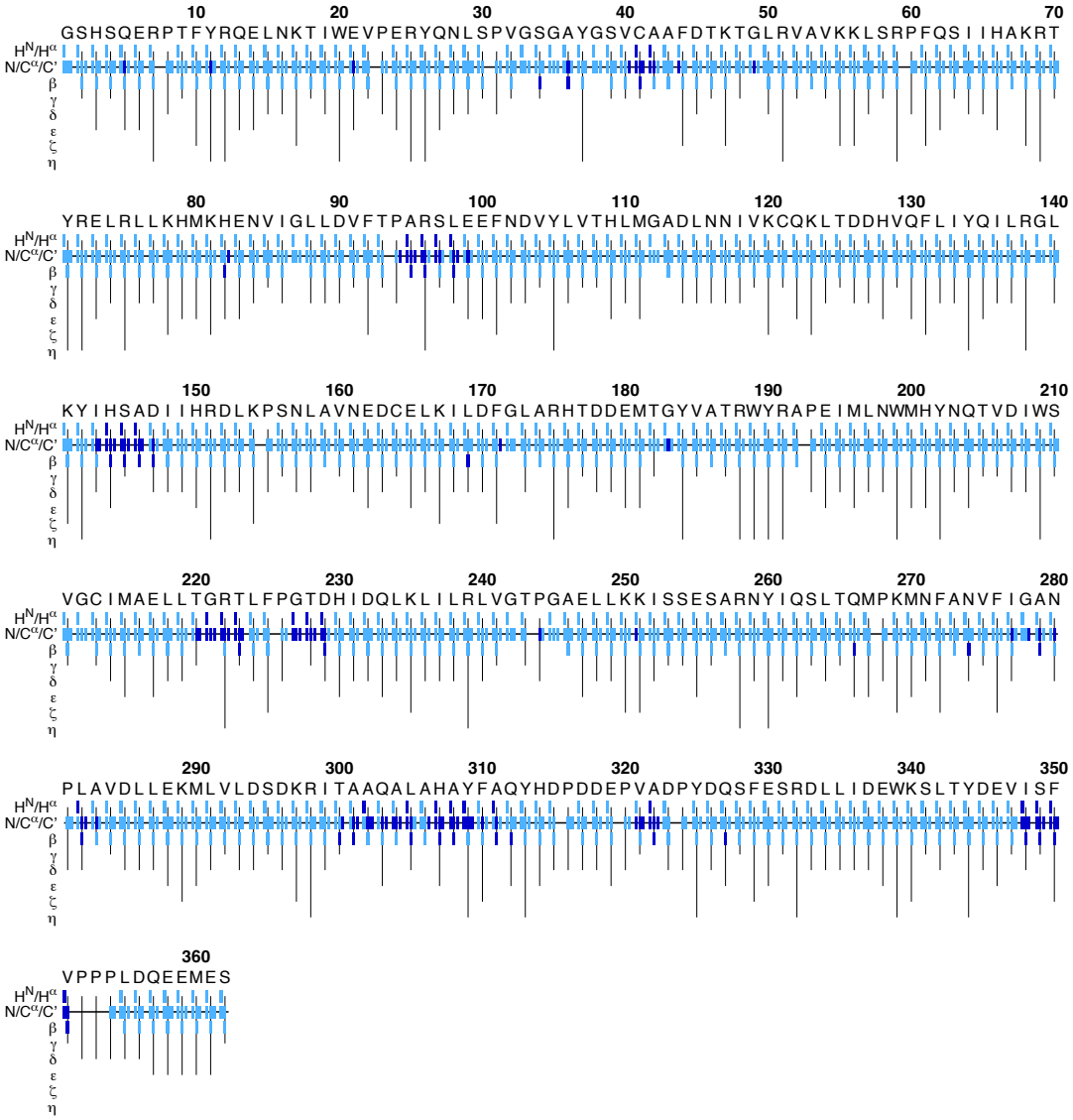


Figure 2.39. FLYA results obtained using the three and four-dimensional experiments.

FLYA allows for the introduction of predicted chemical shifts based on the X-ray structure of the protein for assisting in the assignment process. There are different types of software that can be used for this purpose. One of the available tools is SHIFTX2, which is able to rapidly and accurately calculate chemical shifts of ^{1}H, ^{13}C and ^{15N} based on their structural coordinated and experimental parameters such as deuteration status, pH, and temperature [117]. SHIFTX2 was used to predict the chemical shifts of the backbone for p38α in complex with sorafenib and BOG (PDB entry: 3GCS), which is the complex we used for solid-state NMR assignment. As partial assignments were obtained in solution NMR for the complex, we compared the correlation between the experimental and predicted chemical shift (**Error! Reference source not found.**). This comparison shows a poor correlation for the chemical shift on the amine ^{1}H and ^{15}N chemical shifts, and a good correlation for the Cα and Cβ chemical shift.

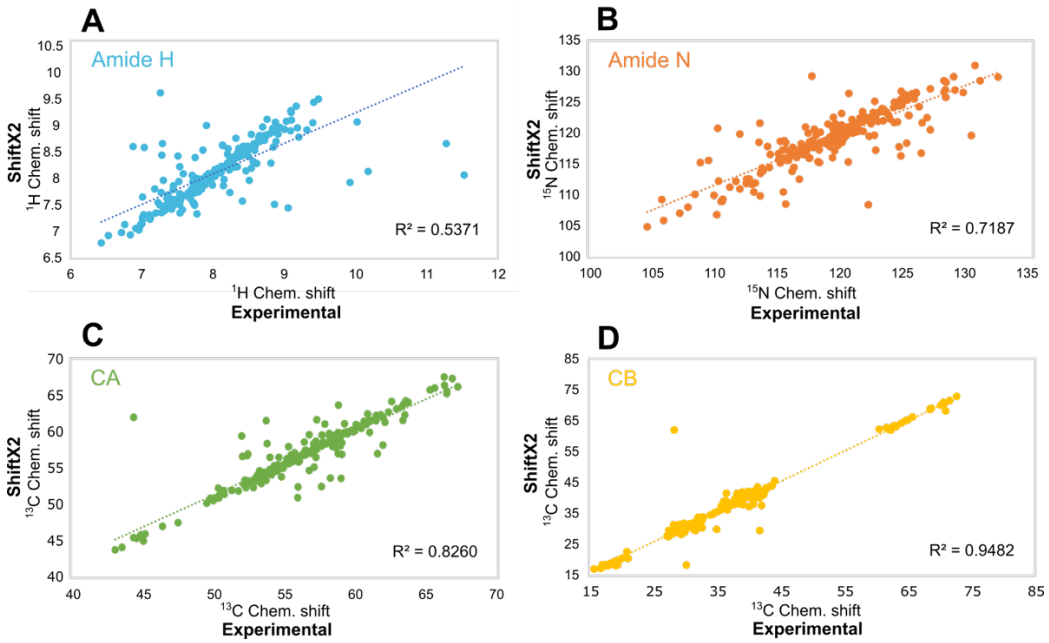


Figure 2.40. Correlation between experimental chemical-shifts of p38α in complex with sorafenib and BOG and the calculated chemical shift from ShiftX2 using the X-ray structure from PDB entry 3GCS.

When the predicted chemical-shifts were used as an input the results obtained from FLYA at a first look show a strong assignment (Error! Reference source not found.). However, a detail a nalysis of the results, shows a very low percentage of assignments for the 4D experiments, which suggest the obtained results are not correct.

Table 2.10. FLYA assignment results for p38α in the solid-state using SHIFTX2 predictions.

List	Expected*	Assigned**	Measured***	Assigned****
15N HSQC	345	312 (90.43%)	202	154 (76.24%)
hCANH	345	229 (63.48%)	391	133 (34.02%)
hCONH	344	271 (70.64%)	543	175 (32.23%)
hCACBcaNH	329	82 (24.92%)	449	65 (14.48%)
HNcocaNH	329	37 (11.25%)	394	31 (7.87%)
hCAcoNH	344	196 (56.98%)	739	150 (20.30%)
hCACONH	329	122 (37.08%)	629	108 (17.17%)
hCOCANH	329	114 (34.65%)	989	95 (9.61%)
hCAncoCANH	329	89 (27.05%)	807	73 (9.05%)

*Expected: total number of expected peaks
**Assigned: number of expected peaks that could be assigned
***Measured: total number of peaks in peak list
****Assigned number of measured peaks that could be assigned to expected peaks

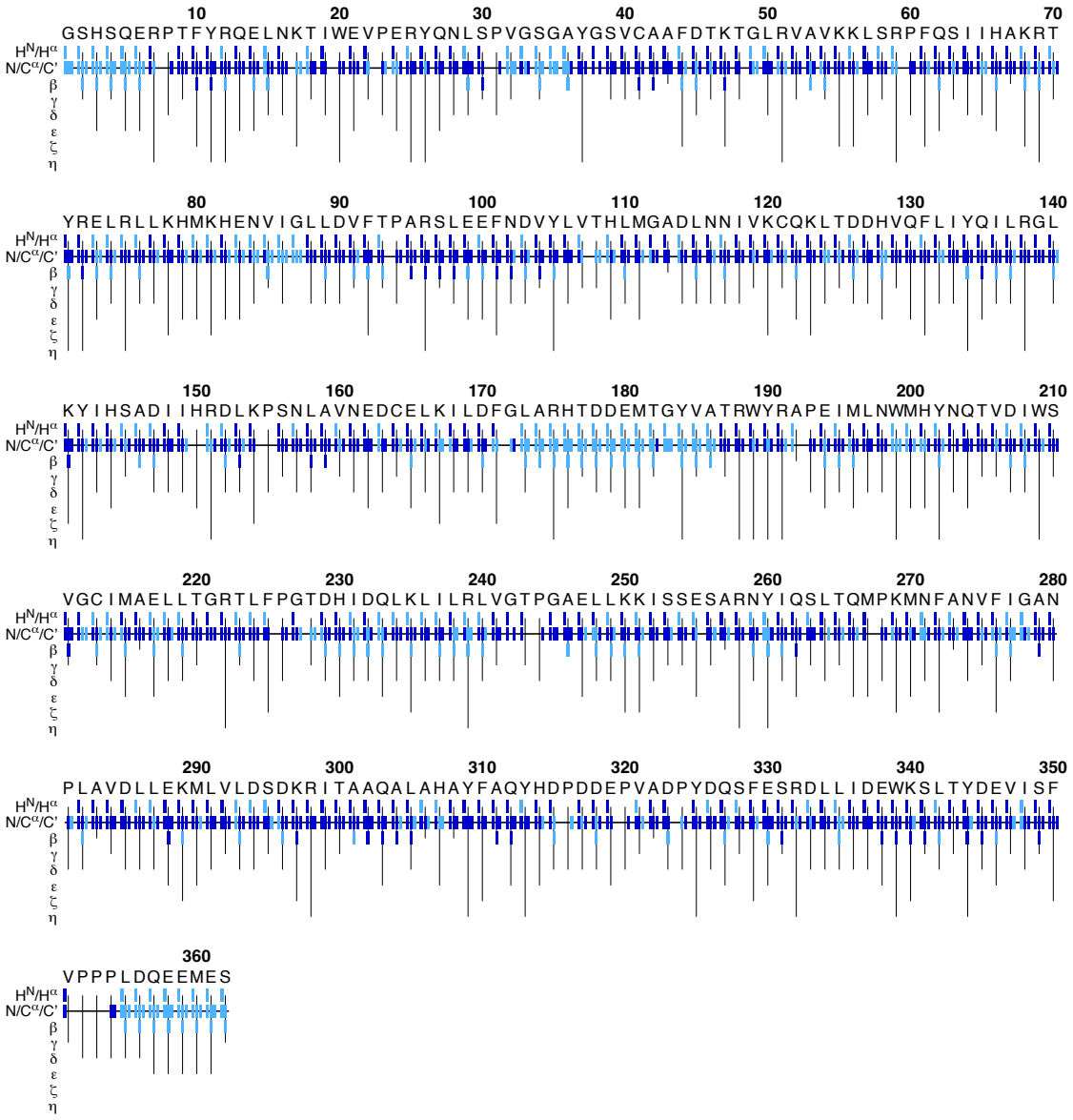


Figure 2.41. FLYA results obtained using the 3 and 4D experiments and the chemical-shifts predicted by SHIFTX2.

The previous results show the existing vias from FLYA for assigning on the basis of a provided reference chemical-shift. Further tests are needed to improve the assignment using FLYA. In addition, new experiments will be recorded in the near future and include experiments to improve the amount of data provided to FLYA, therefore improving the results.

2.3.3 ^{15}N Relaxation Dispersion

As described in Section 1.4, the lack of overall tumbling in solid samples allows us to access the information of slower motional regimes than those accessible in the solution state. Dynamic processes that occur on the µs-ms timescale can be probed by analyzing relaxation processes in the presence of a spin lock field by looking at the spin-lattice relaxation rates of the ^{15}N rotating frame spin-lattice relaxation ($R_{1\rho}$) rates.

All ^{15}N off-resonance rotating frame spin-lattice relaxation ($R_{1\rho}$) rate experiments were recorded in a Buker Avance NEO 700 MHz (^{1}H Larmor frequency) spectrometer with a PH MAS DVT 700s3 BL1.3 mm probe, spinning speed of 55.5 kHz and at ~25°C. For the relaxation dispersion analysis, six effective irradiation field strength ($\omega_e/2\pi$) were recorded: 10, 20, 30, 38, 42 and 44 kHz. Each with four relaxation delays: 10, 20, 40 and 80 ms and a recycle delay of 1.5 s for lower $\omega_e/2\pi$ and 2 s for high $\omega_e/2\pi$ (42 and 44 kHz). The off-resonance irradiation was calculated at an effective tilt angle of ±35.3° and the carrier was set at $+\Delta\Omega$ from the middle of the spectrum. To obtain the desired values, a calibration of ^{15}N RF fields was performed before every set of experiments. An example calibration is presented in **Table 2.1** and the correspondent calibration curve is shown in **Figure 2.42**.

Table 2.11. Calibration of ^{15}N RF field for $R_{1\rho}$ experiments.

^{15}N RF (dB)	**^{15}N RF (kHz)**	**T180° (µs)**
0.02	10.98	45.5
-6	22.06	22.66
-9.52	32.82	15.23
-12.02	42.99	11.63
-13.04	47.16	10.6
-13.6	50	10

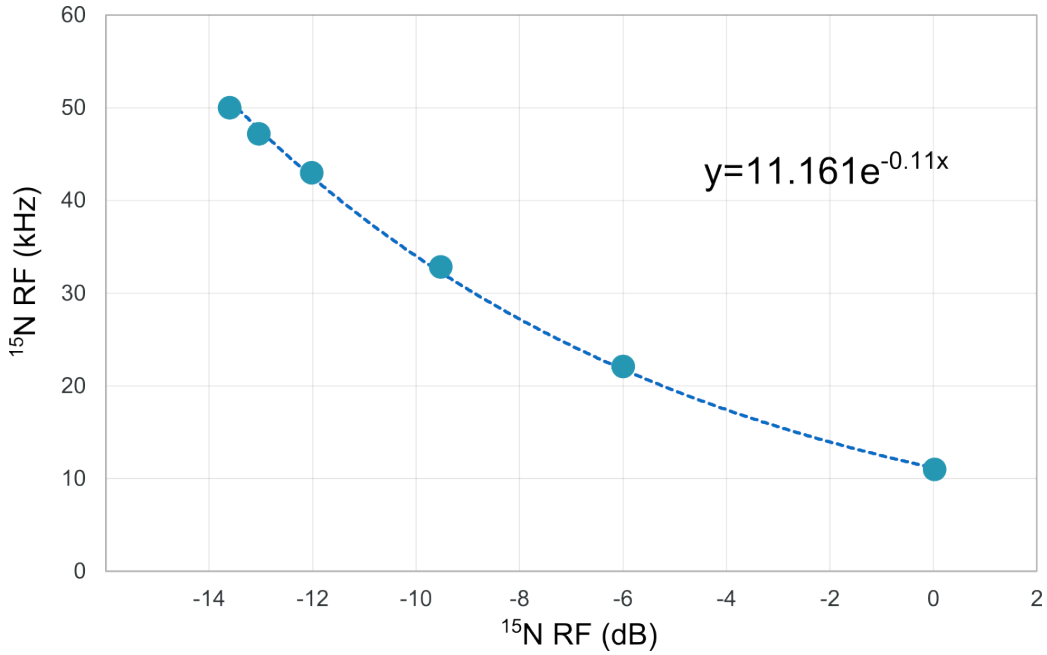


Figure 2.42. Calibration for the ^{15}N RF pulse.

From the optimized ^{15}N RF field, the exponential function obtained is used to calculate the power levels necessary to fulfill the calculated experimental parameters. For every know ω_e we can calculate the values of the off-resonance field strength ($\omega_1/2\pi$) and the and the offset frequency ($\Delta\Omega$). These values were calculated using Equations (1.71) and (1.72) with a tilt angle of 35.3°. The experimental parameters used for one data set are presented in **Table 2.12**. For all relaxation dispersion analyzes, the Relaxation Dispersion analysis the same ω_e, ω_1, and $\Delta\Omega$ were used with the respective power level.

Table 2.12. Experimental parameters for ^{15}N NERRD analysis

$\omega_e/2\pi$ **(kHz)**	$\omega_1/2\pi$ **(kHz)**	**ΔΩ (Hz)**	**Power level (dB)***
10	5.8609	8139.29	5.78
20	11.6192	16278.59	-0.36
30	17.4289	24417.89	-4
38	22.0766	30929.33	-6.12
42	24.4004	34185.05	-7.02
44	25.5623	35812.91	-7.44

*Power levels changed for every set of 15N NERRD experiments, based on the 15N RF calibration.

To access the information of a complicated and large system, as it is p38α, it is necessary to increase the number of dimensions in the experiments to resolve the overlap areas that are

found in the hNH 2D spectrum. As mentioned above, the more nuclei that are evolved during the acquisition of the experiment, the less ambiguity will be observed as the sparsity increases, especially for large systems with many amino acids. To achieve maximum resolution in a reasonable time, a third dimension was added to the $R_{1\rho}$ experiments, in the base form of the hCANH pulse program (**Figure 2.43Error! Reference source not found.**).

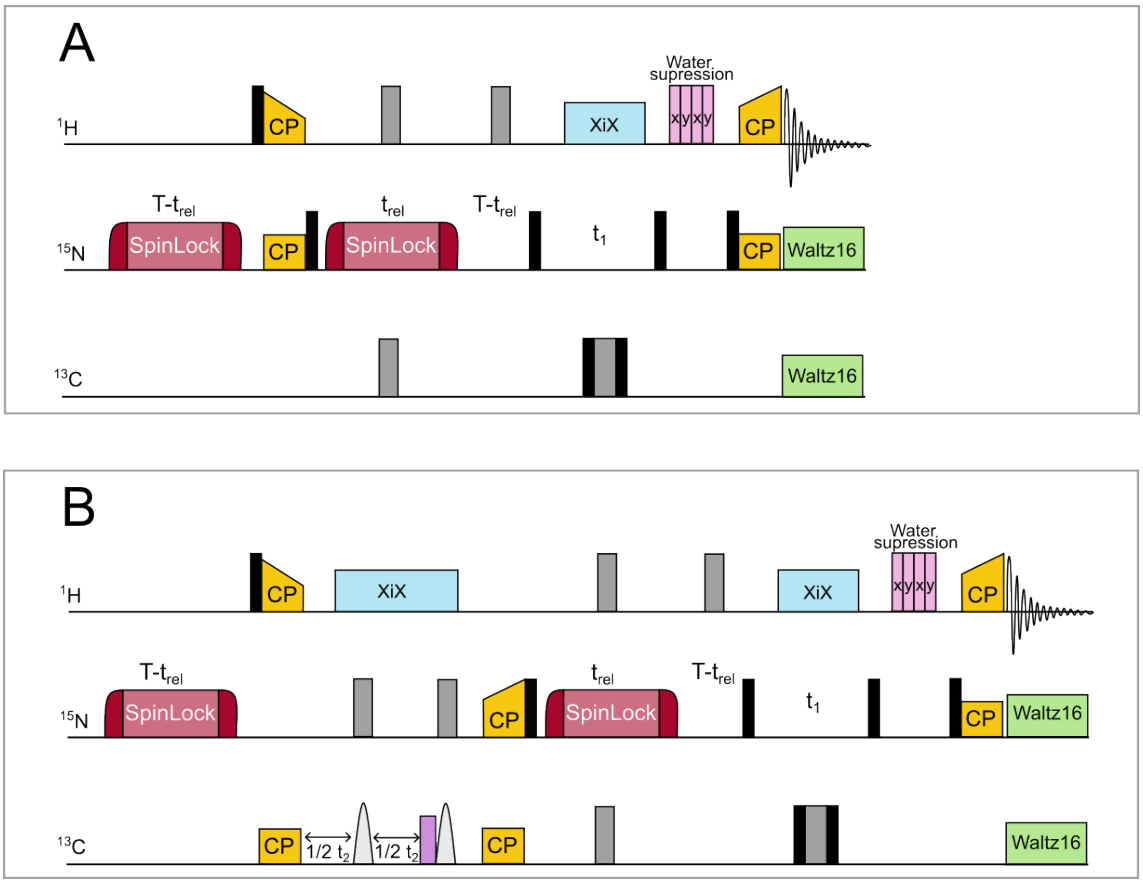


Figure 2.43. Pulse scheme for 15N off-resonance rotating frame spin-lattice relaxation rate measurement ($R_{1\rho}$) for hNH 2D (A) and hCANH 3D (B) based experiments. Filled in black and gray bars represent the 90° and 180° pulses. For the off resonance experiments the adiabatic half passage pulses before and after the spin lock has a shape of thanh/tan. The relaxation delay for the relaxation rate measurements is indicated by $t_{rel.}$

The increase in dimensionality presents the problem of a considerable increase in measurement time. For relaxation dispersion analysis, it is necessary to record the experiments under different conditions, for the NERRD analysis it is necessary to acquire 4 experiments with variations of relaxation delay per effective field. For our analysis, this is translated into 24 experiments per sample. While the total 2D experiments can be carried out in around 6 days, a single fully sampled 3D experiment will require around 2 days to be acquired with a good signal/noise ratio, resulting to be expensive and problematic due to field drift or possible sample degradation. The most common method for solving this problem is the use of non-uniform sampling (NUS), an acquisition method for two or more dimensions that decreases the

number of points measured in the indirect dimensions by skipping some of the points[118-120]. NUS is widely used for the fast acquisition of 2D, 3D, and nD experiments, especially for assignment experiments. However, the use of this methodology for relaxation experiments is not so common, as for this type of analysis, the relative signal intensity is an important parameter for relaxation fitting. An important challenge in working with NUS is the selection of the correct NUS reconstruction method because many NUS reconstruction algorithms are nonlinear and therefore can change the relative intensities for the reconstructed signals [119]. However, there are studies that prove the use of NUS for the acquisition of experiments for which the relative signal intensity is important, such as exact NOE and relaxation studies with success using sample densities of 25% [119, 121].

To test the required sample NUS density, a series of 2D ^{15}N $R_{1\rho}$ relaxation experiments for the SH3 domain of chicken α-spectrin were recorded on a Buker Avance NEO 700 MHz (^{1}H Larmor) spectrometer with a PH MAS DVT 700s3 BL1.3 mm probe, spinning frequency of 55.555 kHz and at a temperature ~25°C. As a reference, a fully sampled ^{15}N $R_{1\rho}$ relaxation series was recorded, in addition to the 30% and 40% NUS density series. Each relaxation series was recorded with four relaxation delays (10, 20, 40, 60, and 80 ms) and a 1.5 s recycle delay. **Figure 2.44** shows the comparison of the 15N R_1 rates obtained ^{15}N $R_{1\rho}$ rates for the different sampling densities; it can be seen that there are no significant changes between the rates for most of the residues.

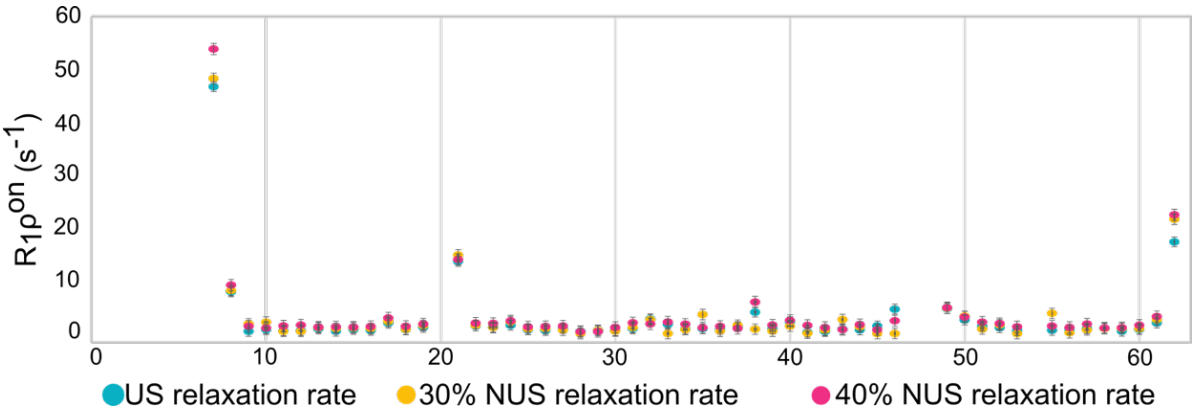


Figure 2.44. Comparison of the $R_{1\rho}$ rate obtained from experiments acquired using uniform sampling (blue), 30% NUS (yellow), and 40% NUS (pink).

For those visible resides located in flexible regions such as the N-terminal (7 and 8), RT loop (R21) and C-terminal (61 and 62) it can be observe a slight difference between the ^{15}N $R_{1\rho}$ rates from the fully sample and the NUS can be observed. The observed difference can be explained

when we consider the signal intensity for those amino acids, which are difficult to obtain in a regular hNH 2D experiment due to the fast motion they present. When long relaxation delays are included, the overall signal intensity drops down; therefore, residues that were less intense become, in some cases, difficult to distinguish from the noise. This increases the error and gives the origin to the different errors. For this reason, it is important to include as many points as possible in order to avoid reconstruction artifacts or losing signals that are inherently weak. With this consideration we have decided to use higher NUS density for the following experiments.

The Near-Rotary Resonance Relaxation Dispersion experiments for p38α were acquired in a pseudo 4D fashion (**Figure 2.45**), where the relaxation delay (dynamic dimension) acts as the third indirect dimension, while the first two indirect dimensions and the direct dimension contain information on the chemical shift correlation of the evolved nuclei per amino acid. With this methodology, it is possible to acquire a full block of ^{15}N $R_{1\rho}$ relaxation experiments without the need of setting up the experiments individually, decreasing additional sources of error.

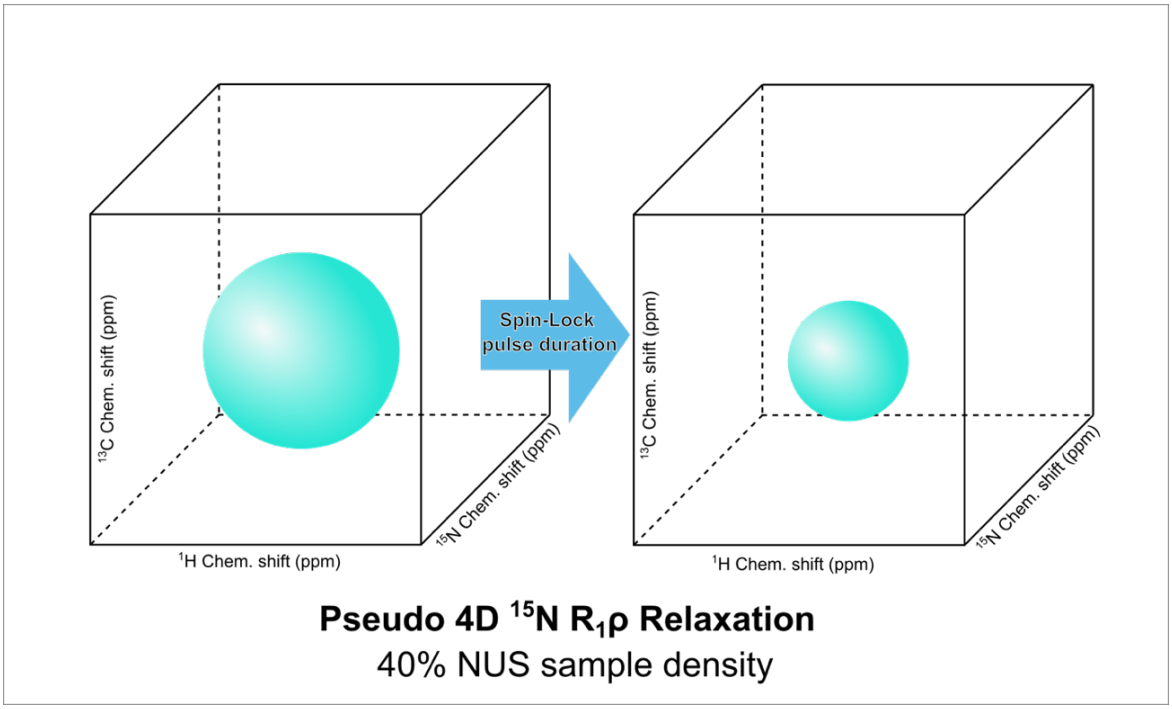


Figure 2.45. Schematic representation of the Pseudo 4D ^{15}N $R_{1\rho}$ relaxation experiments which indicates the fourth dimension as the dynamic dimension and the decrease of the signal intensity with the longer relaxation delay.

The pseudo-4D fashion experiments were recorded with randomized Gaussian weighted NUS list with 30 and 40% sampling density. With the pulse program used, it is only possible to vary the relaxation delay while keeping the same NUS list for all the implemented delays. The

pseudo-4D set of experiments was recorded in a RAP50 sample using 30% NUS density. Furthermore, a set of 3D based ^{15}N $R_{1\rho}$ experiments was recorded in a fully deuterated sample (DCN100) with 40% NUS density where all the NUS lists for the different relaxation delays were scrambled randomized Gaussian weighted lists. The experiments recorded as pseudo 4D were processed and reconstructed using the Sparce Multidimensional Interactive Line Shape Enhancement (SMILE) algorithm [122], while the 3D data were processed and reconstructed using the Signal Separation Algorithm (SSA) [123].

The relative intensity of each signal from the final processed spectral series of ^{15}N $R_{1\rho}$ was fitted to an exponential decay to extract the relaxation time:

(2.7)

$$\frac{I(\tau)}{I_0} = e^{-\frac{\tau}{R_{1\rho}}}$$

Where $I(\tau)$ is the relative intensity with a relaxation delay of time τ and I_0 is the initial intensity. Data were fitted using equations (2.6) with an in-house written Mathematica® script with the support of Dr. Petra Rovó and Dr. Suresh K. Vasa.

As more signals are expected to appear in the RAP50 sample due to the back-exchange problems present in a fully deuterated sample for the hydrophobic regions, the preferred sample to be analyzed is the RAP50 labeling scheme. However, since there is no rigorous control of the percentage of protonation present in the $^{1}H\alpha$ it is difficult to estimate the effect of the dipolar contributions that these protons will have in the analysis of the relaxation data. To investigate the suitability of the RAP50 sample for the NERRD analysis, the ^{15}N $R_{1\rho}$ rates at 10 kHz $\omega_e/2\pi$ were compared with each other (**Figure 2.46**).

This analysis shows a poor correlation between the two samples. However, it is not possible to exclude all potential sources of error. We can consider certain experimental problems, such as the NUS density, sampling scheme, and reconstruction method; nevertheless, there may also exist problems that are inherent to the labeling scheme. However, conceptually, NUS reconstruction algorithms SMILE and SSA are analogous [122], therefore, this could not be the origin of the differences. As it is difficult to explain with precision the reason for the difference between the relaxation rates for the two sample types, we can argue that the poor quality of the spectra at high fields and long relaxation delays are the main reasons for this

problem. This problem can be highlighted for the RAP50 sample, as the experiments were recorded with 30% NUS density.

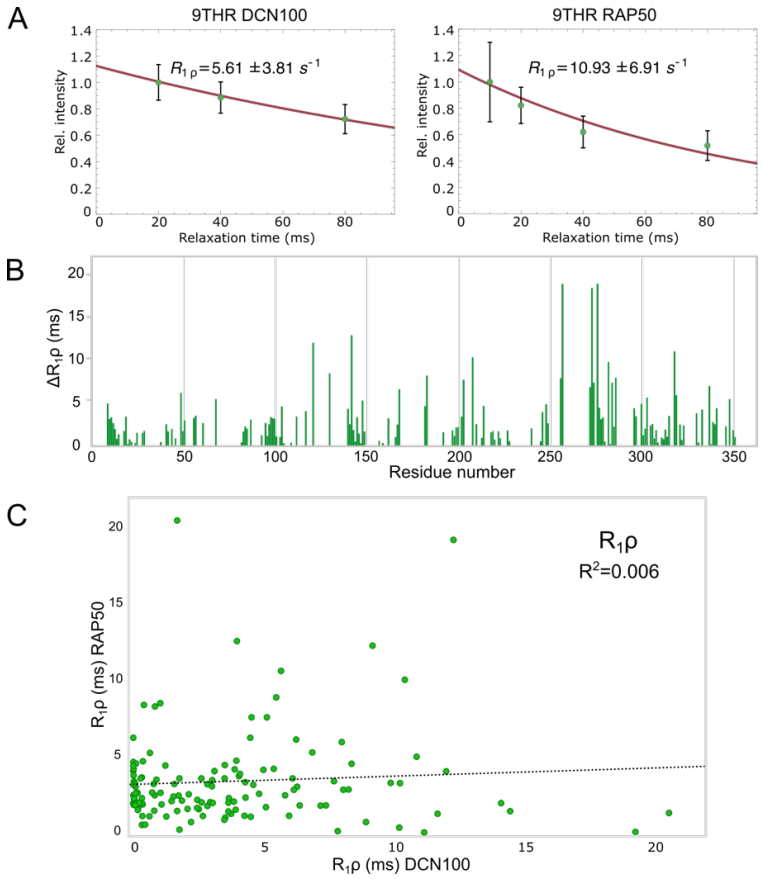


Figure 2.46. Comparison of the ^{15}N R1 rates fitted at 10 kHz $\omega_e/2\pi$ for the fully protonated sample and the RAP50 sample. **A)** Example of off-resonance ^{15}N R1 rates at 10 kHz $\omega_e/2\pi$ for Thr9 in fully protonated (left) and RAP50 (right) samples. **B)** Differences between relaxation rates ($\boldsymbol{\Delta R_{1\rho}}$) at 10 kHz $\omega_e/2\pi$ for the samples analyzed. **C)** Correlation plot between the ^{15}N $R_{1\rho}$ rates at 10 kHz $\omega_e/2\pi$ between the fully deuterated sample and the RAP50.

Considering the reconstruction methods looking for a local maximum that fulfill defined criteria to be identified and reconstructed as a peak, when the peaks of weak intensity are present, they are excluded from the further reconstruction steps. At long relaxation times the signal intensity decays close to the noise level, therefore, the possibilities of missing those peaks increase with the use of NUS data.

Being that the fitting of the relaxation follows an exponential decay, the rate between the initial intensity and the intensity with a relaxation delay can be compromised and wrongly fitted when one of the points seems to have decayed completely because of the reconstruction method. This will increase the relaxation rate value, such as the case that we observe when comparing the two samples (**Error! Reference source not found..A**). An alternative to resolve this problem i

s to acquire the experiments with a time that allows all of the residues to be present, especially for experiments with a long relaxation delay and at high effective fields.

As the fully deuterated sample was acquired using 40% NUS density, more points are recorded that are used for reconstruction. This can be translated into better spectra quality. The results of this sample were used to map the dynamic processes occurring on the µs-ms timescale. For the NERRD fit, the ^{15}N $R_{1\rho}$ rates at different ω_e are plotted. Residues that do not present motion at this particular time scale we observe a flat profile along all the ω_e. For the residues involved in µs to ms timescale, the $R_{1\rho}$ rates will increase monotonically as the ω_e frequency approached the rotary resonance condition and will decrease afterwards, giving as a result an apparent Lorentzian peak in the $R_{1\rho}$ profile. Additional information that can be obtained from the NERRD profiles is the generalized order parameter (S^2), which is related to the amplitude of the peak, and the timescale (τ_c) of motion; thus, a larger peak will indicate a faster motion. [40]

Figure 2.47 illustrates examples of the possible resulting profiles from the NERRD analysis, where we can see an example of a residue with no motion in the µs-ms timescale (A) and as residues with a positive NERRD profile; a broad peak indicating faster motion (B) and a narrower peak indicating slower motion (C) in the µs-ms regime. We can map on the structure those residues that present motion in the µs-ms regime when p38α is in complex with sorafenib and BOG.

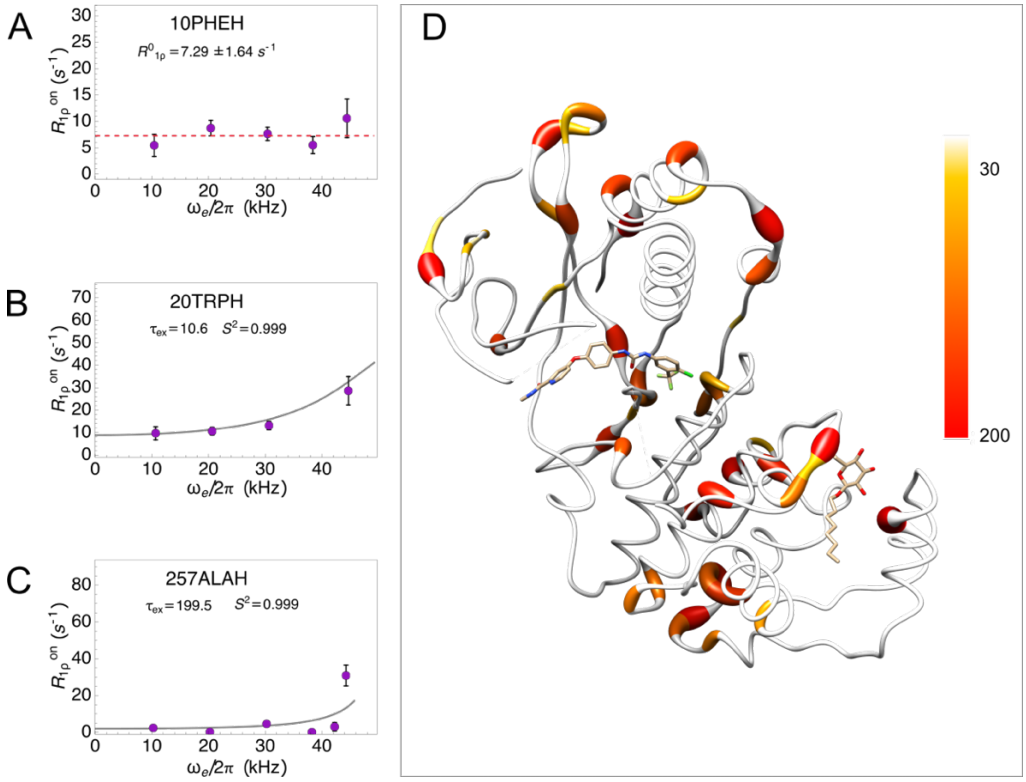


Figure 2.47. NERRD analysis of p38α in complex with sorafenib and BOG, with examples of profiles that indicates no motion in the µs-ms regime **(A)** and profiles with NERRD peaks showing a faster **(B)** and slower **(C)** residue. Those residues that show a peak in the NERRD profiles are shown in the X-ray structure (PDB: 3GCS) indicating the τ_{ex}.

NERRD analysis was performed to the p38α in complex with regorafenib and BOG and the p38α in complex with sorafenib RAP50 samples. The experiments were recorded with the same parameters as the previous samples in a pseudo 4D fashion experiment with 40% NUS density. The reconstruction of the spectra was done using the SMILE algorithm, and the data were fitted using Mathematica®.

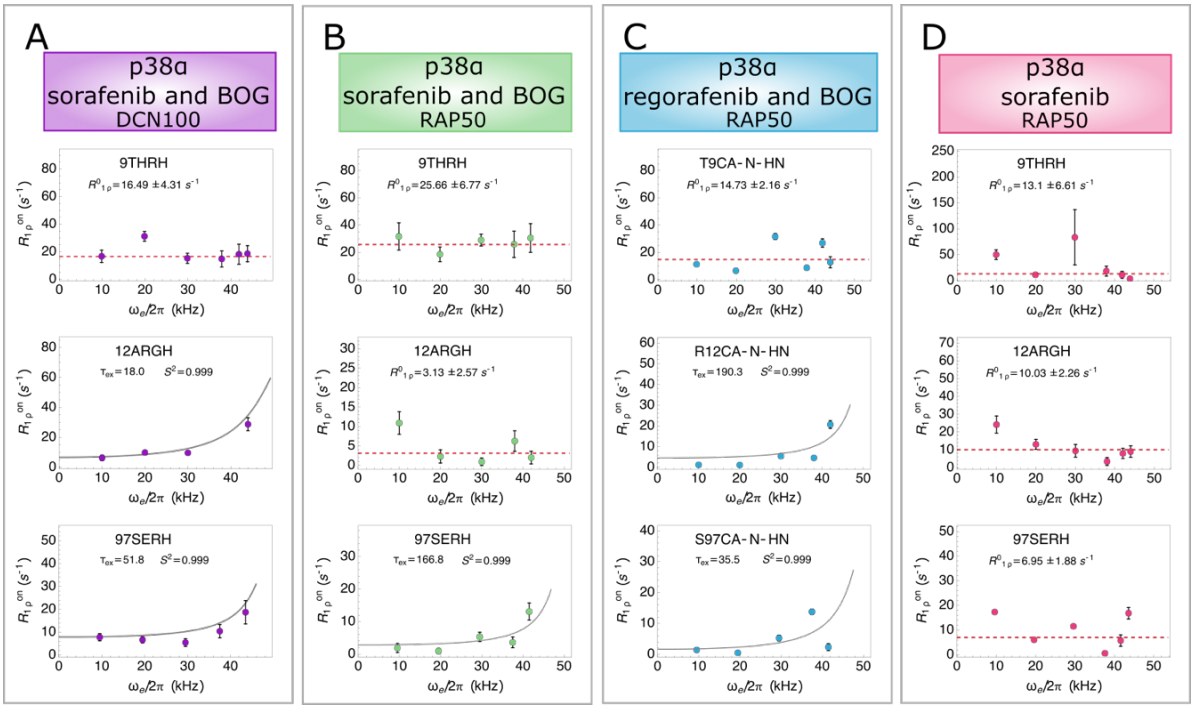


Figure 2.48. Example of ^{15}N NERRD profiles for the residues T9, R12 and S97 for the different sample conditions. **A)** ^{15}N NERRD experiments on p38α in complex with sorafenib and BOG fully deuterated recorded as 3D with 40% NUS density. **B)** ^{15}N NERRD experiments on p38α in complex with sorafenib and BOG RAP50 recorded as pseudo 4D with 30% NUS density. **C)** ^{15}N NERRD experiments on p38α in complex with regorafenib and BOG RAP50 recorded as pseudo 4D with 40% NUS density. **D)** ^{15}N NERRD experiments on p38α in complex with sorafenib RAP50 recorded as pseudo 4D with 40% NUS density.

When NERRD profiles are compared from the different sample, several inconsistencies can be observed that cannot be explained only by the effect of the ligand. **Figure 2.48** shows a few examples of the differences that can be found between the samples. Thr9 is an example of residues with no motion in the µs to ms timescale, the average off resonance $R_{1\rho}$ rate observed for the samples of p38α in complex with regorafenib and BOD and in complex with sorafenib only are consistent with the rate obtained for the fully deuterated sample of p38α in complex with sorafenib and BOG. Two more examples are Arg12 and Ser97, which in fully deuterated p38α in complex with sorafenib and BOG show a NERRD peak with different amplitudes indicating motion at different timescales. Those residues show a NERRD peak for the p38α in complex with regorafenib and BOG with a considerable change in timescale for Arg12 for slower motion. In the case of p38α in complex with sorafenib, only the same residues do not show the NERRD peak. However, it is not possible to attribute this effect to the absence of BOG in the lipidic pocket or the change of ligand at the allosteric binding site, as the data quality still shows a high error level for some of the residues.

For some of the residues in the RAP50 samples in the complex with sorafenib and BOG and the complex with sorafenib only a tendency was observed for a high $R_{1\rho}$ rate at 10 kHz $\omega_e/2\pi$ frequency that decreases with the use of stronger $\omega_e/2\pi$ frequencies, such as the type of curves are present in the BRMD analysis. One explanation for this behavior could be a nonuniform level of protonation in the side chain, because of the used expression protocol we have certain level of scrambling in the deuteration of the side chain difficult to predict. Although dipolar dephasing is less problematic for ^{15}N than for ^{1}H or ^{13}C, partial protonation at ^{1}Hα position can increase the coherent contribution to dephasing, even at a fast MAS frequency.

However, the effect of partial protonation is not present in the p38α in complex with the regorafenib and BOG sample, therefore the reason for the effect described before should be further investigated. As the results obtained from this sample are closer to the results obtained from the fully deuterated p38α in complex with sorafenib and BOG, we compare the τ_{ex} obtained from the NERRD analysis.

Figure 2.49 shows the difference between the $\boldsymbol{\Delta\tau_{ex}}$ of the residues that show a NERRD peak for both samples. For some residues, the changes are more pronounce, as for one sample there is no motion in the µs-ms timescale, and it the dynamic process is present when the ligand has a change. An example is illustrated in **Figure 2.49.A**, where the His82 residue His82 presents a flat NERRD profile when the protein is in complex with sorafenib but has a clear peak of NERRD when regorafenib is present in the complex. Another example presented in **Figure 2.49**.**B** is Glu83, for this residue, the NERRD peak is present in both samples, but with the difference in the amplitude of the peak and the calculated order parameter (S^2). Both His82 and Glu83 are close to the allosteric binding site, and even though they are not in direct contact with the Fluorine present in regorafenib, we can see that for Glu83 the motion becomes slower.

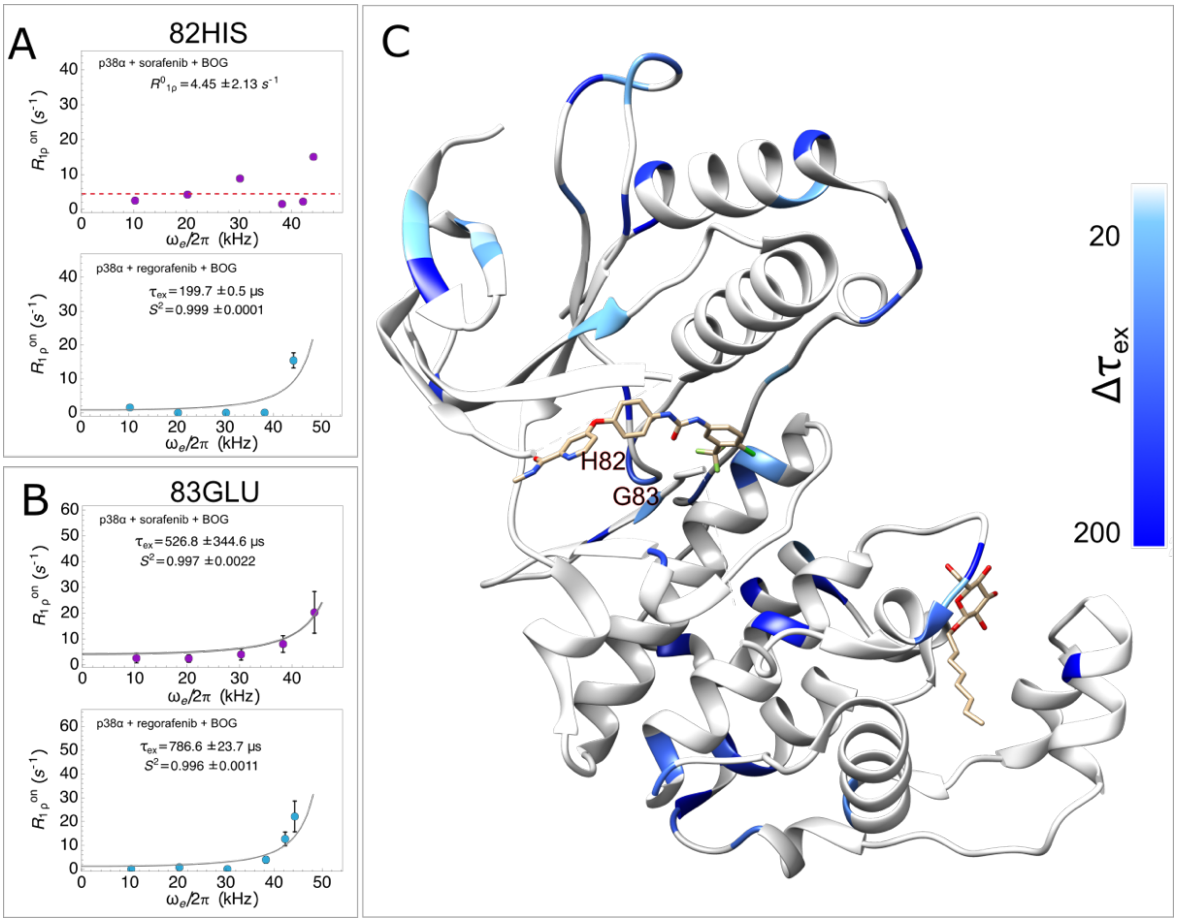


Figure 2.49. 15N NERRD profiles obtained for p38α when in complex with sorafenib and BOG (purple) compared with the profiles obtained for p38α in complex with regorafenib and BOG (light blue) for residues H82 (**A**) and G83 (**B**). **C**) Differences of τ_{ex} of both samples illustrated in the three-dimension X-ray structure (PDB entry 3CGS), where the correspondent color and values are shown in the figure

Chapter 3

SECONDARY STRUCTURE PROPENSITIES OF RHOGDI1 N-TERM DOMAIN

3 SECONDARY-STRUCTURE PROPENSITIES OF RHOGDI N-TERMINUS

The Rho family of ras-related GTPases is a family of small signaling G proteins, which includes the isoforms rho, rac, cdc42, and TC10. These proteins work as molecular switches in a broad range of important cellular functions and are well known for their action in the reorganization of the actin cytoskeleton, initiating signaling pathways for gene expression and regulation of cell growth, as well as in focal adhesion formation and stimulation of stress-activated kinases [124, 125]. Alteration of these processes can lead to neurological and immunological problems. Rho GTPases cycle between an active and an inactive conformation, which are regulated by three types of proteins: Guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs), and Guanine nucleotide dissociation inhibitors (GDIs). GEFs catalyze the exchange of GDP for GTP that activates GTPase. On the other hand, GAPs increase the intrinsic GTP hydrolysis rate of GTPase, leading to inactivation. GDIs negatively regulate this family by sequestering the GDP-bound form from [125]. The inhibitory activity comes from the ability to bind the carboxy-terminal isoprene and extract them from the membranes and from inhibition of GTPase cycling between the GTP- and GDP-bound states. (**Figure 3.1**). The N-t domain of RhoGDI plays a crucial role in these inhibition processes and has been shown to be disordered and ordered when it is in complex with Rho GTPases [123].

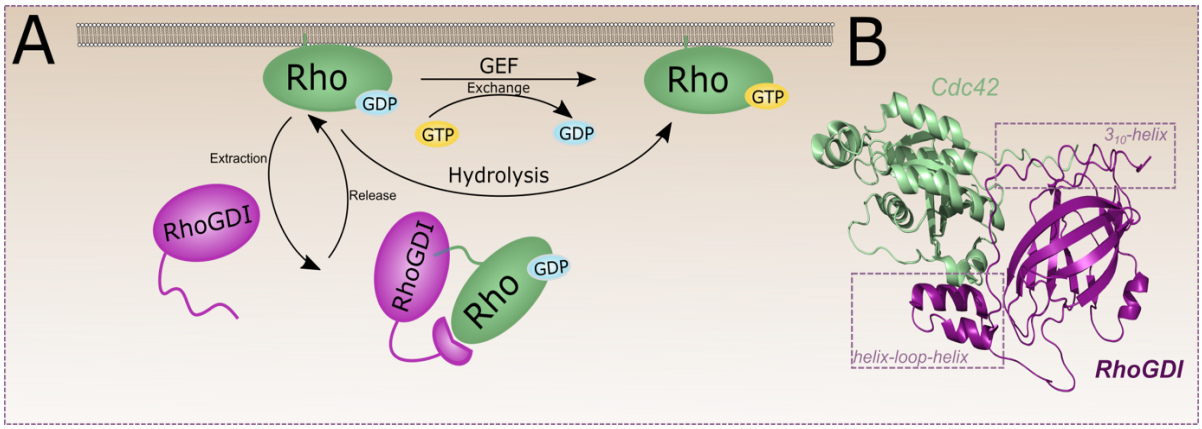


Figure 3.1. Overview of mechanism of regulation of Rho GTPase from the membrane. **A)** Process of negative regulation of the Rho GTPase family by Rho-GDP dissociation inhibitor sequester the GTPase on the cytosol and preventing it for localizing the membrane or been activated by the GEFs. **B)** X-Ray structure for the RhoGDI: Cdc42 complex where can be observed the presence of the helix-loop-helix motif as well as the 3_{10}-helix secondary structure of the N-terminal domain.

The RhoGDI family is conform by three isoforms in mammalians: RhoGDI1, RhoGDI2, and RhoGDI3. The three members of this family have different expression patterns as well as mechanisms for binding GTPases. The interaction between RhoGDI1-GTPase has been described as a two-step process, in which the amino terminal domain first binds to the switch domain of the Rho GTPase. This interaction inhibits the transition between the GDP and the GTP bound form. As a second step, the carboxyl-terminal two-third of RhoGDI1 assumes an immunoglobulin-like domain creating a lipidic pocket that provides a membrane substitute for the geranylgeranyl moiety of Cdc42, thus maintaining the protein soluble in the cytosol. [124-128]

X-ray crystallography and NMR studies suggested a change in the secondary structure of the N-terminus of RhoGDI1 as an explanation for the interaction between RhoGDI1 and GTPase. [124, 125] Here, we show the secondary structural propensities present in the N-terminal domain when bounded to Cdc42 and in the free state.

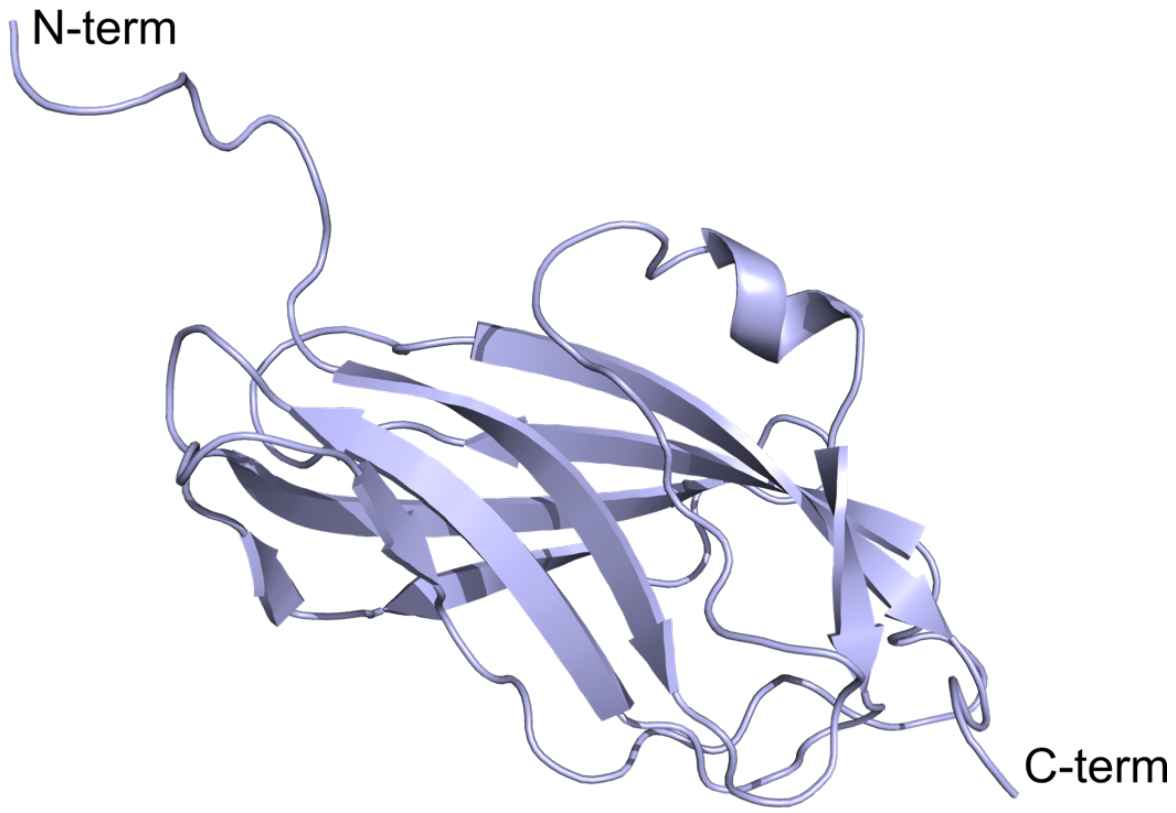


Figure 3.2. Structure of RhoGDI1 obtained by NMR where amino-terminal region is assumed unstructured in the absence of target. PBD number 1AJW.

3.1 Experimental conditions

RhoGDI1 has a molecular weight of 20 kDa. It is made up of 205 amino acids and residues 1 to 59 are part of the N-term domain of the protein. To allow site-specific characterization of secondary structural propensities, a suit of 3D solution NMR experiments was performed for assignment purposes. The experiments were carried out for the full-length RhoGDI1 as well as for the isolated N-terminal domain. The protein samples used were provided by our collaborators at the Max Planck Institute for Molecular Physiology in a form of [^{15}N,^{13}C,^{2}H] labeled for the full-length RhoGDI1 as well as for the complex between RhoGDI1 and Cdc42 and [^{15}N,^{13}C] labeled for the N-terminus. All samples were analyzed in shigemi tubes with a final volume of 0.3 mL in buffer containing 20 mM HEPES, 150 mM KCl, 0.5 mM TCEP, pH 7.5 and 10% (vol/vol) D_2O and sample concentration between 0.4-0.6 mM.

All NMR spectra were recorded at 298K on a Bruker Avance NEO 800 MHz (^{1}H Larmor) spectrometer equipped with a CP TCI proton-optimized triple resonance cryoprobe. The recorded solution NMR assignment experiment consists of 2D ^{15}N-^{1}H HSQC, 3D HNCA, 3D HNCO, 3D HN(CA)CO and HNCACB/CBCANH for both samples. ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates measurements were acquired interleaved as follows. T_1 and T_2 The experiments were acquired with a recycle delay of 1.5 s between the experiments and the follow delays for T_1: 10, 50, 100, 200, 300, 400, 500, 600, 800, 1020, 200 and 600 ms; and T_2: 10, 30, 50, 70, 90, 110, 130, 150, 170, 190, 50, 110 and 170 ms. Two and three experiments were repeated, respectively, to determine the experimental error of the measurements. For the isolated N-terminus analysis, a set of 3D 15N NOESY-HSQC with a mixing time of 200 ms and 3D HNHA with transfer time of 13.05 ms experiments were recorded.

3.2 Backbone Assignment

The *de novo* assignment was performed manually using ccpNMR v3 [111]. As the N-terminal domain is presumably a disorder, a high level of overlap is expected in the region between 7.5 and 8.5 ppm. Thus, for complete assignment of the N-terminal domain, the sequence was truncated, and the N-term was analyzed alone. The assignment process for the full-length protein was performed manually as well using 3D experiments and complemented with the assignments previously performed in the isolated N-terminal domain.

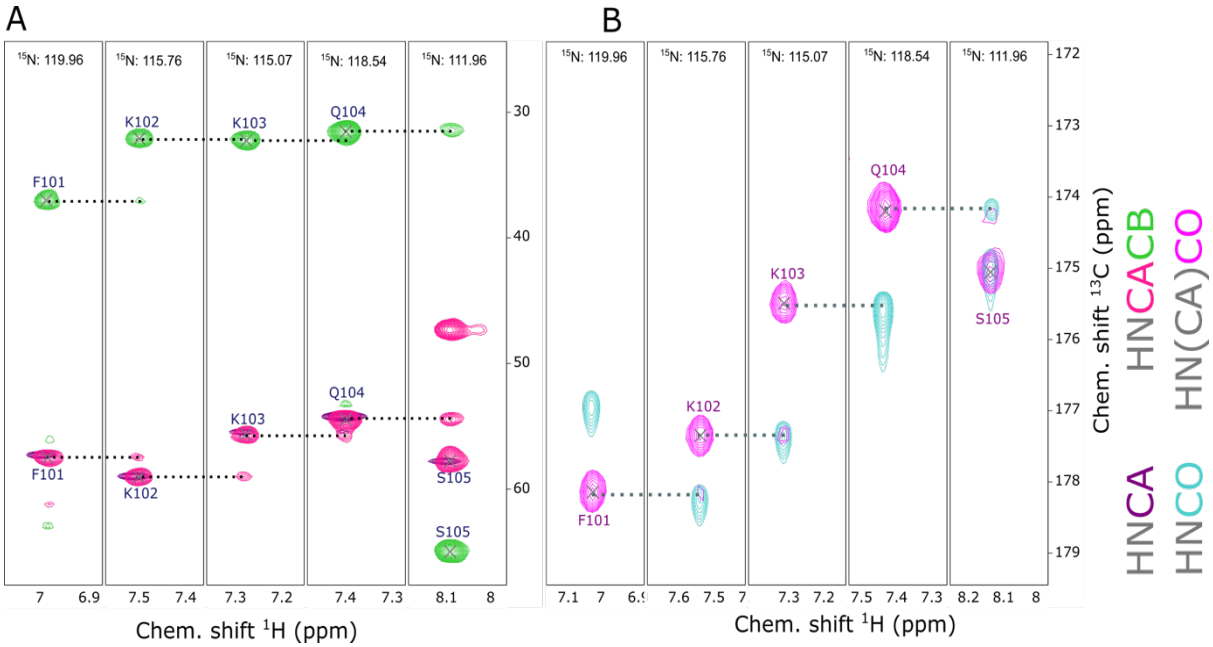


Figure 3.3. NMR assignment process. Every stripe corresponds to a slide in the ^{15}N dimension that contains the HN chemical shift information with the correspondent correlated carbon. **A)** Cα and Cβ information from HNCA and HNCACB 3D experiments, pink and purple signals correspond to the Cα chemical shift while green signals correspond to the Cβ chemical shift. The large peaks belong to the residue i and the smaller peak indicate the carbon chemical shift for the residue $i-1$. **B)** Carbonyl (CO) chemical shift information from the HNCO and HN(CA)CO 3D experiments. Blue and weak pink peaks correspond to the correlation with the residue $i-1$, while the peak strong pink peaks belong to the i residue.

By separating the N-term, it was possible to complete the assignment for the protein at full length in the center of the spectrum. 100% of the visible amide HN signals were assigned in the 15N-1H HSQC for the isolated N term, while for the full RhoGDI1 only 78% of the expected residues were assigned (**Figure 3.4**) and the chemical shift assignment was deposit in the BMRB data bank under the entry 51835.

Initial analysis of the full-length protein indicated that the N-term corresponded to an intrinsically disorder domain because most of the peaks were in the center area of the spectrum and the intensity of those peaks was greater than that of the rest of the peaks. Thus, for the rest of the analysis for the isolated N-terminal domain was analyzed as an intrinsically disorder protein (IDP).

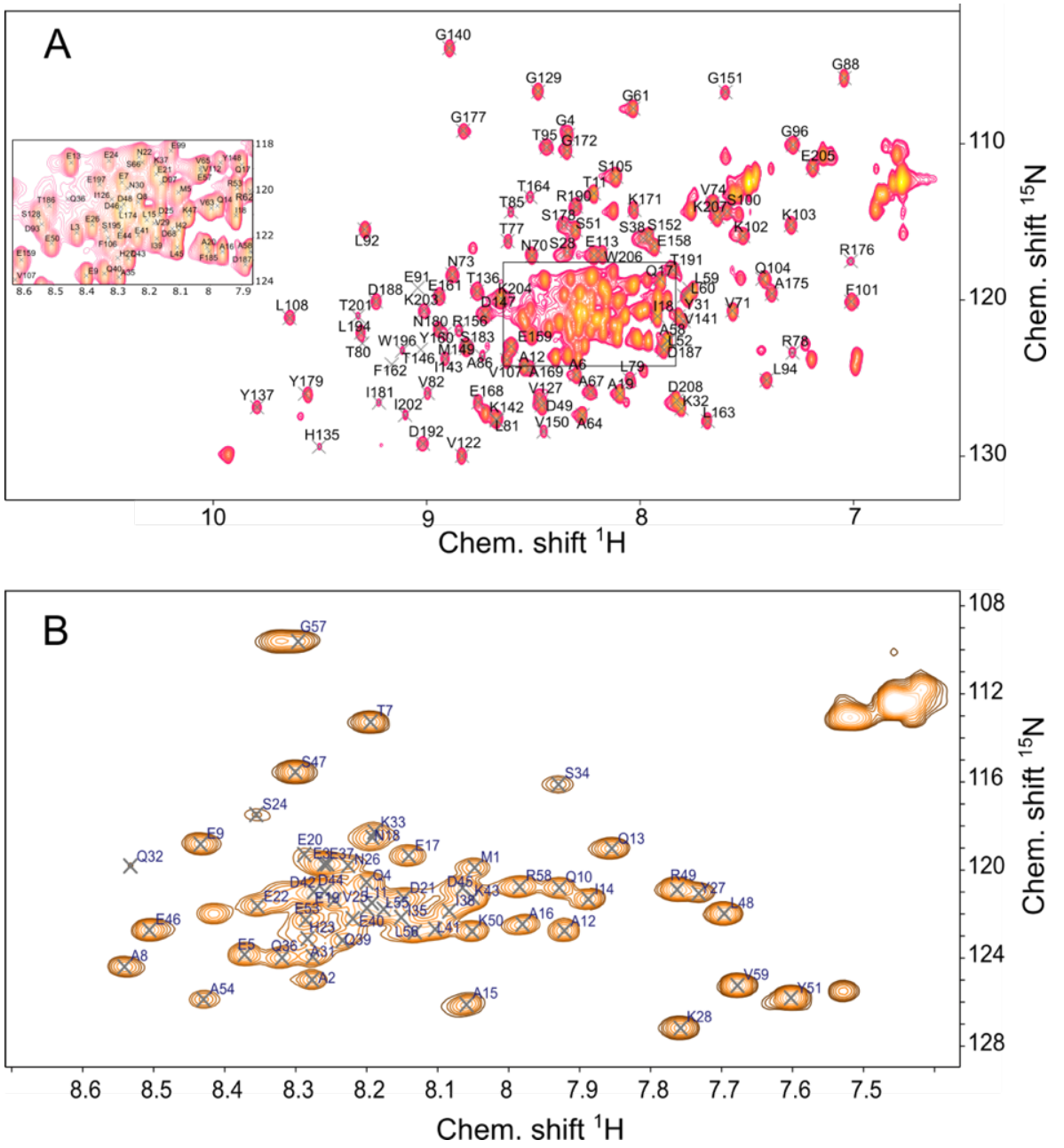


Figure 3.4. Chemical shift assignment for full-length RhoGDI1 and isolated N-terminal domain. **A)** ^{15}N-^{1}H HSQC of the full-length protein with the crowded region depicted in the excerpt shown in the small rectangle. **B)** ^{15}N-^{1}H HSQC of the isolated N-terminal domain (residues 1-59). All spectra were recorded on an 800 MHz (^{1}H-Larmor frequency) spectrometer at 298K.

3.3 Isolated N-terminal domain

As described in Section 35, NMR chemical shifts contain and report valuable information about the chemical environment of a specific nucleus. In the protein context, this information is closely related to the secondary structural elements of which an amino acid is a part. Different methods can predict these secondary structural elements; among them, CheSPI.

CheSPI provides unbiased measures of local structure and disorder, which is an ideal alternative to analyze IDPs [50]. The CheSPI software was run using the Web server (https://st-protein.chem.au.dk/chespi). The software was provided with an nmr-star v2.1 file containing the protein sequence, experimental conditions (pH of 7.5, ionic strength 1 and 25°C) and chemical shift information for HN, N, CO, Cα, and Cβ. The Z-score obtained from the CheSPI analysis indicates the level of order or disorder for different regions of the protein. A high value of the Z score indicates the presence of a stable secondary structure, while a low value indicates a disorder region. In addition to the Z-score values, CheSPI provides information for the populations of local structures corresponding to the DSSP (hydrogen bond estimation algorithm) classes.

To get a closer look of the N-terminal domain, the CheSPI analysis was performed with the experimental chemical shift information obtained from the isolated N-term. Furthermore, the sequence-based probability of disorder was calculated using ODiNPred (Prediction of Order and Disorder by Evaluation of NMR data), a method that used a deep neuronal network, trained on a comprehensive set of characteristics and an experimental Z score for 700 proteins [129, 130]. The ODiNPred software was run using the Web Server (https://st-protein.chem.au.dk/odinpred) by providing the sequence of the isolated N-terminal domain.

The analysis based on the experimental chemical shift information shows a tendency for a α-helix secondary structure from the residues 8 to 16 and indicates a tendency for the secondary structural element in the region corresponding to the helix-loop-helix domain (**Figure 3.5**). The presence of a strong tendency for the helical element in the 8 to 16 region disagrees with the probability of disorder calculated by ODiNPred, which indicates a rater mild probability for the order. However, the presence of a 3_{10}-helix in this region is known from the X-ray structure of the complex.

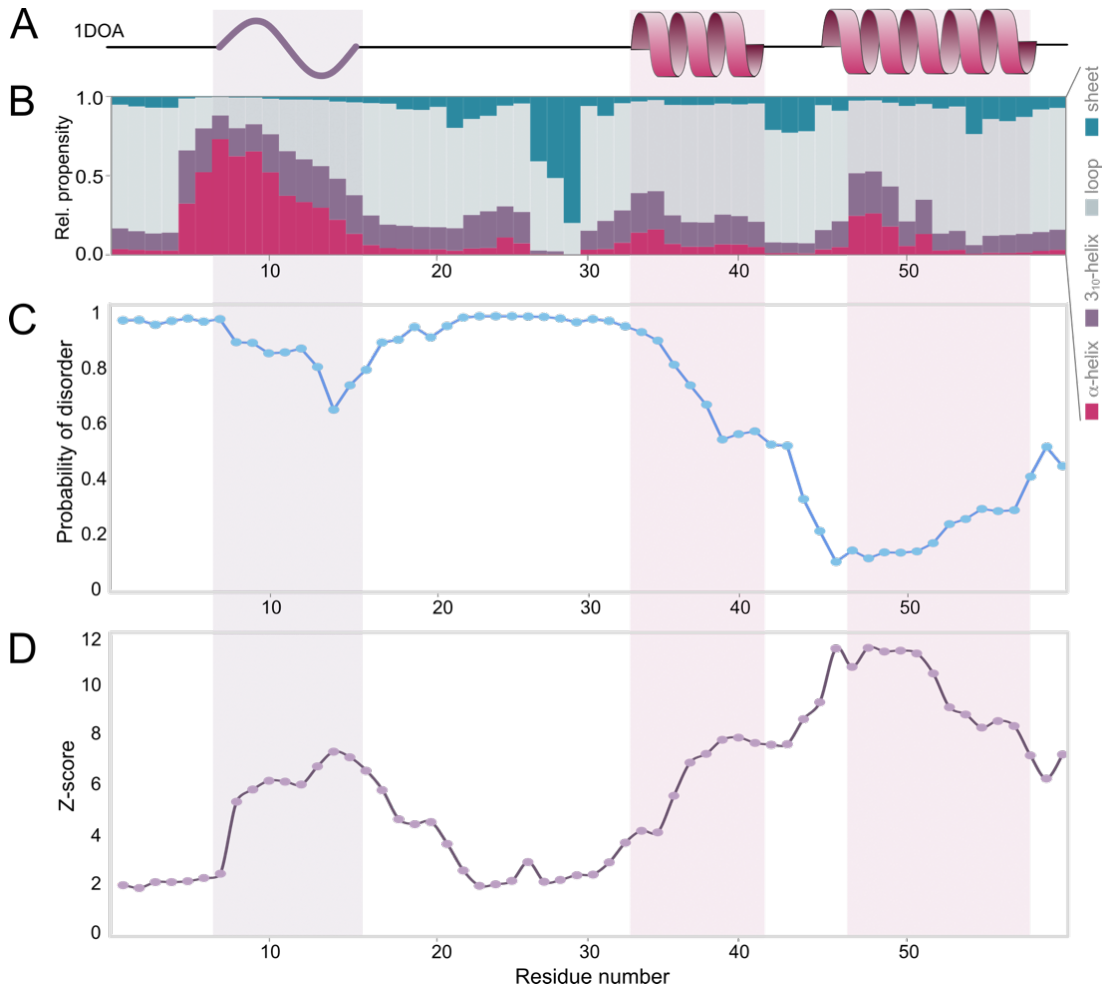


Figure 3.5. Experimental assessment of N-term with NMR chemical shifts. RhoGDI secondary structural propensity. **A)** Secondary structure found in the X-ray structure of the RhoGDI:Cdc42 complex (PDB 1DOA), showing only the part corresponding to the N-terminal domain. **B)** CheSPI relative secondary-structural propensities, **C)** ODiNPred calculation of the probability of disorder based on the isolated N-term, and **D)** CheSPI Z-score of the isolated N-terminal domain of RhoGDI. The color code for the secondary-structural analysis is shown on the right.

In addition to CheSPI, other software such as Talos+ predicts the secondary structure of a protein based on experimental chemical shift information. Talos+ predicts the ψ and ϕ backbone torsion angles using a combination of HN, Hα, Cα, Cβ, CO, and N chemical shift for a given residue sequence, by using a neural network classification scheme and a large database of 200 proteins. The neural network analyzes the chemical shift and sequence to estimate the likelihood that a residue has been in a sheet, helix, or loop conformation [51]. Talos + analysis was carried out on the Web Server (https://spin.niddk.nih.gov/bax/nmrserver/talos/) and the results obtained are compared with the CheSPI results.

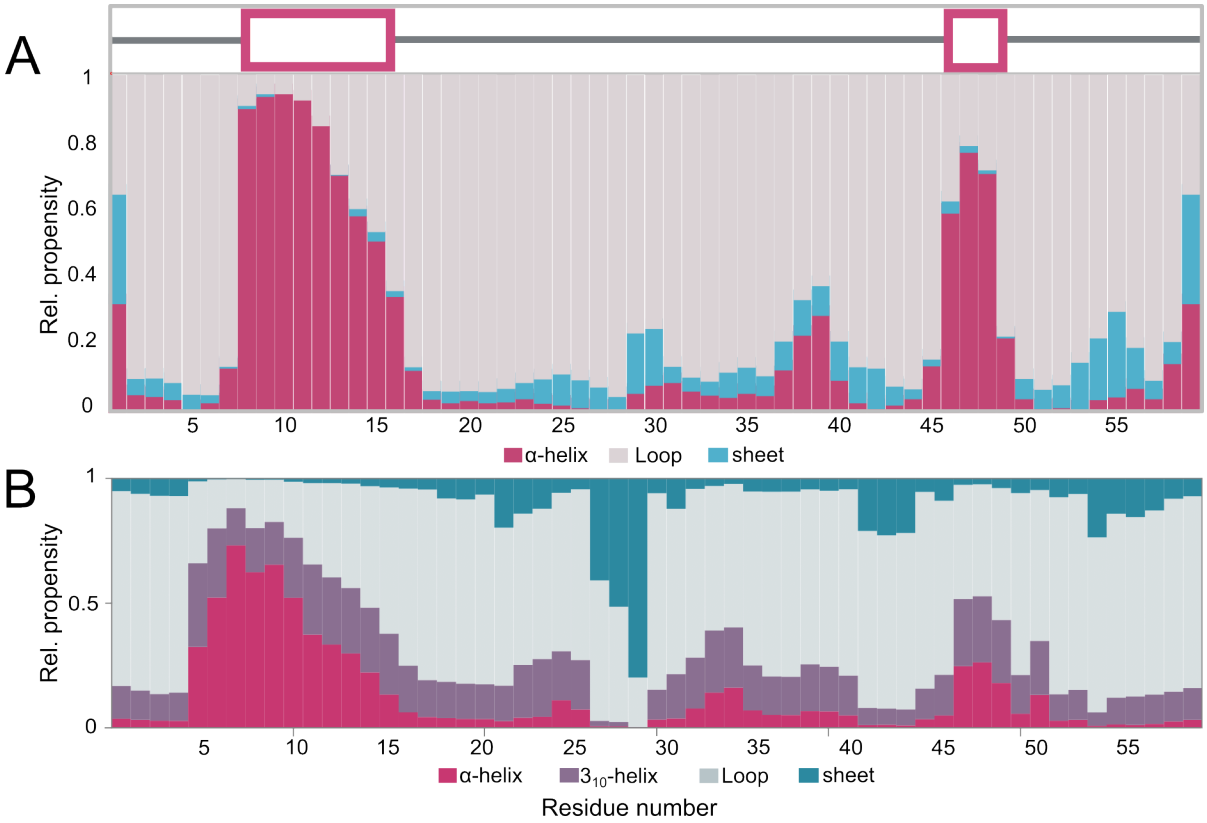


Figure 3.6. Predictions of the secondary structure on the basis of experimental chemical shifts for the isolated N-terminal domain. **A)** Talos+ prediction using HN, Hα, Cα, Cβ, CO, and N chemical shift with the representation of the proposed secondary structure. **B)** Prediction of secondary structure by CheSPI using HN, Cα, Cβ, CO, and N chemical shift.

The Talos+ prediction shows a high propensity for helical elements for residues 8 to 16 and 46 to 48. While the CheSPI prediction shows a higher propensity of helical elements for residues 5 to 14 and a low propensity for helical elements for residues 32-40 and 46-50. The rest of the sequence shows a tendency for loop conformations. Although there seems to be good agreement between the two prediction methods, a higher propensity for the secondary structural elements can be observed in the Talos+ prediction. This can be expected because of the different data sets for which the algorithms were trained; on the one hand, Talos+ uses a training data set proteins with a high level of order and the presence of clear secondary structural elements. In contrast to Talos+, CheSPI was trained using IDP s without clear secondary structure elements. Therefore, CheSPI has no bias for predicting secondary structure for the regions or proteins that have clear secondary structure.

3.3.1 Experimental determination of ^{3}J-coupling

There are experimental procedures that can be used to explore in depth the secondary structural elements in a protein. One of these methods consists in the measurement of scalar coupling constant. Homo- and heteronuclear three-bond *J*-coupling constants contain important information about the local backbone and side chain conformation, especially information about the backbone dihedral angle ϕ and side chain torsion angle χ_1. [4] One of the most widely used methods is the HNHA experiment [131] designed for relatively large protein systems. In this experiment, the HMQC signal is split into two signals appearing at δH^N and δH^α in the third proton dimension. Measurement of the $^3J_{H^N H^\alpha}$ coupling constant is based on a quantitative analysis of the diagonal to cross-peak intensity ratio ($S_{cross}/S_{diagonal}$) and can be obtained from the equation (3.1)

(3.1)

$$\frac{S_{cross}}{S_{diagonal}} = -\frac{sin^2(2\pi^3 J_{H^N H^\alpha}\zeta)}{cos^2(2\pi^3 J_{H^N H^\alpha}\zeta)} = -tan^2(2\pi^3 J_{H^N H^\alpha}\zeta)$$

Where ζ= transfer time = $\frac{1}{4}J$. For the recorded experiment $\frac{1}{4}J = 13.05$.

The magnitude of the scalar coupling constant between multiple bonds depends on the dihedral angles, therefore, can be used to establish local backbone and side chain conformations. The magnitude of ^3J is a function of the dihedral angle formed by the three covalent bonds between the couple spins. This can be described by the Karplus correlation (1.76). Calibration of the Karplus equation is important for obtaining accurate ^{3}J-constant couplings. Different experimental and theoretically calculated values for A, B, and C have been determined, for the purpose of this study the values for $^3J_{H^N H^\alpha}$ used for the calibration of the Karplus curve are A=6.51, B=-1.76, and C=1.60 [131]. The curve obtained is shown in **Figure 3.7**.

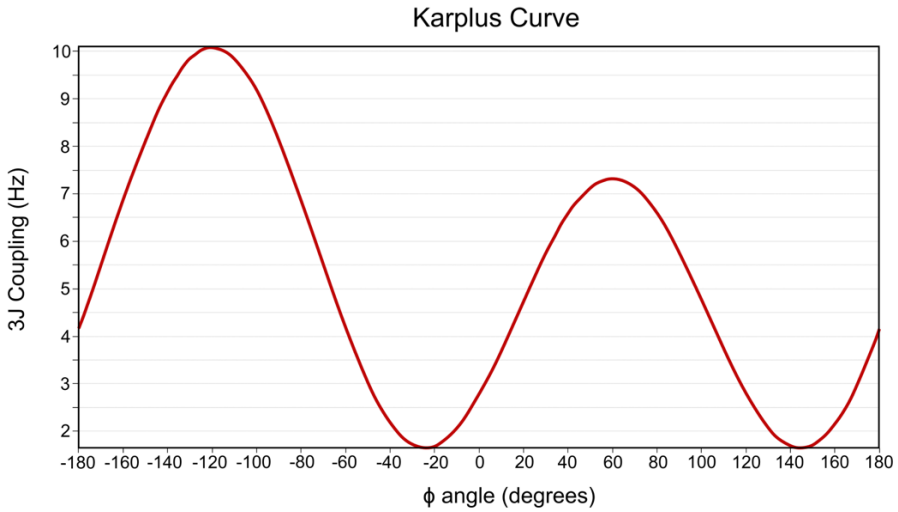


Figure 3.7. Calibration Karplus curve obtain for the calculation of $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$ with the values of A = 6.51, B=-1.76, and C = 1.60.

For proteins, the $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$ the values for coupling between the backbone amide proton and the α-proton from the same amino acid reflect the secondary structure elements. Values smaller than 5.5 Hz are indicative of α-helical secondary structure: values higher than 8.5 Hz hint at the presence of β-strand. On the other hand, those values that range between 6 to 8 Hz reflect motional averaging of multiple conformations.

The experimental $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$-coupling for the isolated N-terminal domain (**Figure 3.8**) shows that many residues have ^{3}J coupling values between 6 and 8 Hz, which can be interpreted as a random coil. It can be seen that residues between 8-15 and 43-48 have a tendency for lower values of $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$-coupling, nevertheless there is no clear tendency for a stable secondary structure. This effect can be attributed to the overall tendency of the domain for multiple conformations.

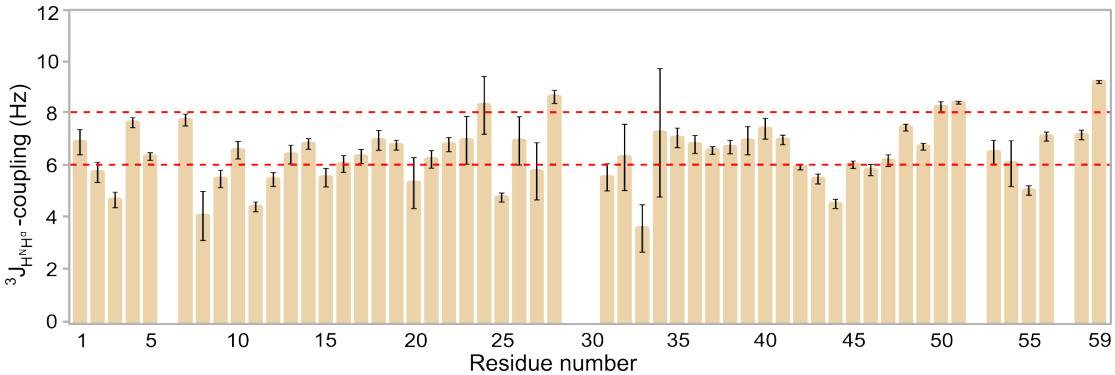


Figure 3.8. Experimental $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$-coupling values obtained from the 3D HNHA experiment.

3.3.2 Experimental confirmation of secondary structure

In addition to backbone dihedral angle ϕ and side chain torsion angle χ_1, another type of experimental restrains for structural calculation are homonuclear (^{1}H-^{1}H) NOE contacts. NOEs can be obtained from 3D HSQC-NOESY and contain a lot of information about all ^{1}H-^{1}H spatial interactions regardless of chemical connectivity [132]. These interactions are observed as cross-peaks, for which the signal intensity depends on the distance (r) between the two hydrogens. This dependance can be described as $NOE = 1/r^6$. The NOE can be described as intraresidue, medium-range, and long-range. Medium-range NOEs are used to describe the backbone conformation use to determine the secondary structure, and the long-range NOEs reflect the global folding of the protein and use for the determination of the tertiary structure. [133]

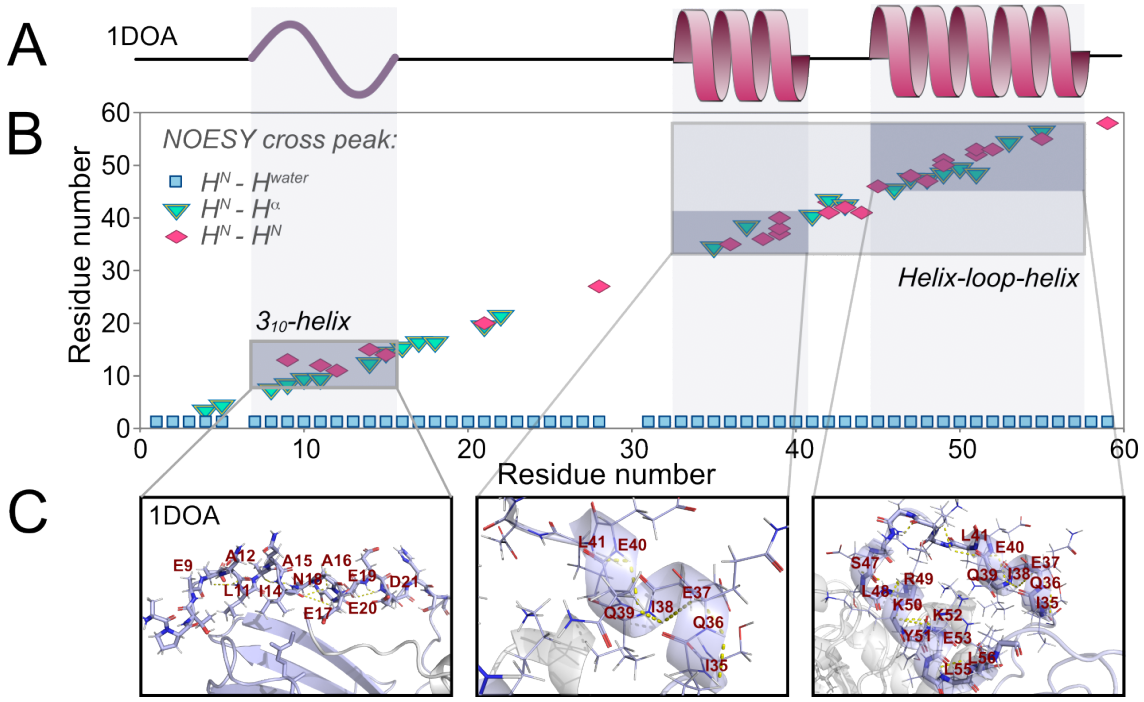


Figure 3.9. Observed NOE contacts for the isolated N-term. **A)** Close-up of the secondary structure for the N-terminal domain of RhoGDI1 when in complex with Cdc42 from PDB 1DOA. **B)** NOE cross-peaks in a ^{15}N-edited NOESY experiment showing the presence of NOE contacts between amide protons and nearby amides (red diamonds), close Hα (green triangles), and to water (blue squares). **C)** Presentation of secondary structure within the stretches of high degree of (temporary) secondary structure with H-bonds mark in yellow dashes.

The analysis of the ^{15}N-edited NOESY showed the expected inter-residue NOEs, and in addition we observe correlations between protons at 3 and 4 residue distances in the regions where helical elements were calculated by CheSPI. These experimental data suggest that the secondary structure elements are present part of the time in an exchange of conformations for the isolated amino terminal region.

3.4 Full-length apo

To better understand the secondary structure of the N-terminal domain when it is part of the full protein, the analysis of the secondary structure propensities was performed using the experimental chemical shift information for HN, Cα, Cβ, CO, and N for every residue assigned using CheSPI and Talos+.

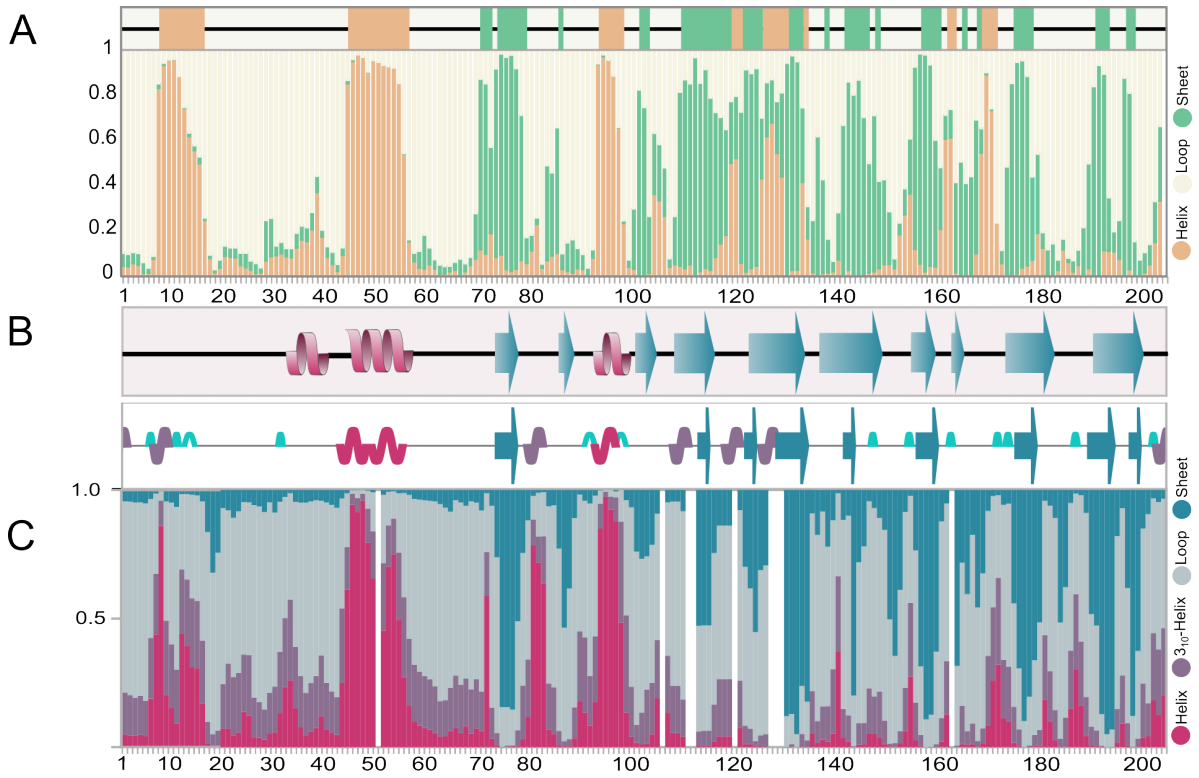


Figure 3.10. Prediction of secondary structures from experimental chemical shift information. **A)** Talos+ prediction using the chemical shift of HN, Cα, Cβ, CO, and N with the representation of the proposed secondary structure. **B)** X-ray structure of the RhoGDI1 complex with Cdc42 from PDB 1DOA. **C)** Prediction of secondary structure by CheSPI using the chemical shift of HN, Cα, Cβ, CO, and N.

Both prediction methods (**Figure 3.10**) show a propensity for a secondary structure in the N-term and agree well with the X-ray structure for the complex with Cdc42, indicating that the secondary structure for the N-term is present before the formation of the complex. It can be observed that a higher propensity for a secondary structure is observed in the prediction of Talos+, it is specially noted for the residues 8 to 17 which are predicted to be part of a strong helical element, although this part of the protein is predicted as a 3_{10}-helix. The predicted secondary structural propensities differ from previous knowledge of the N-terminal domain.

Compared to the structural elements obtained by X-ray for the complex with Cdc42, we can observe a good agreement not only for the globular part of the protein but also for the N-term. Residues 42 to 57 display a stable α-helical, element not expected to be present in the non-bound protein. The Z-score values calculated on CheSPI show a tendency for a high level of order for most of the protein in the apo form.

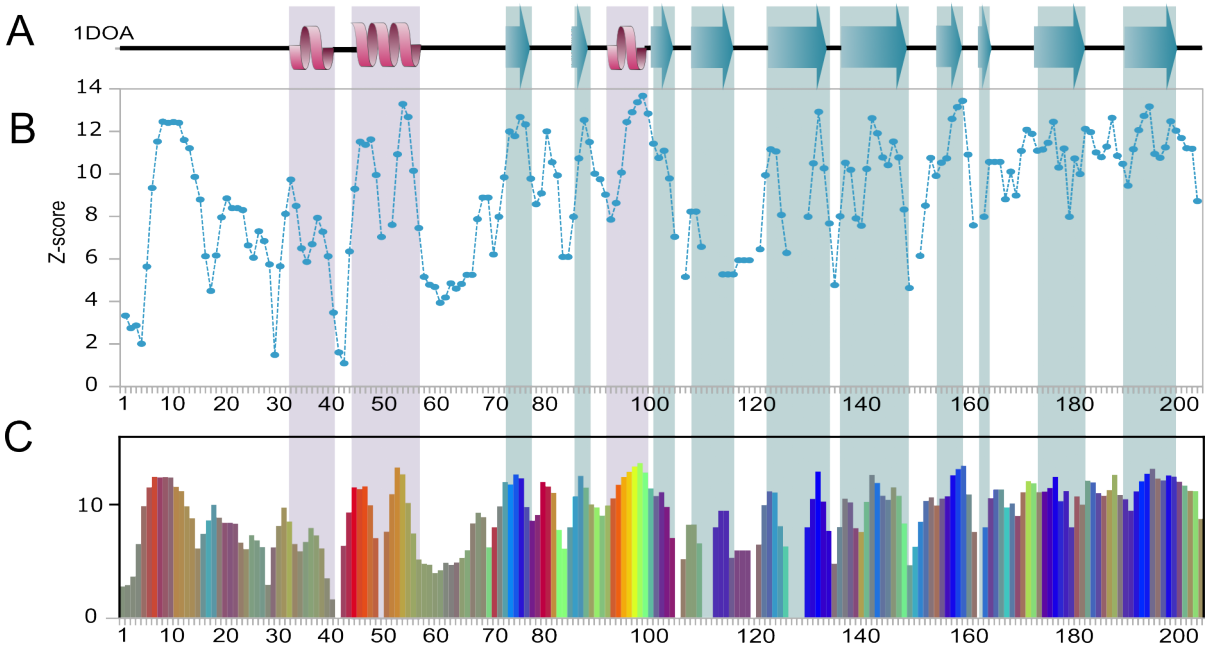


Figure 3.11. Experimental NMR-chemical-shift-based assessment of the level of disorder of fl-RhoGDI. **A)** Secondary structure found in the X-ray structure of the RhoGDI:Cdc42 complex (PDB 1DOA). **B)** CheSPI Z-score of apo RhoGDI.**C)** CheSPI Z-score of apo RhoGDI with CheSPI colors of the tendency for the DSSP classes.

In addition to the computational analysis, site-specific R_1 and R_2 relaxation rates were analyzed. As shown in **Figure 3.12**, these rates confirm the greater overall flexibility of the N-terminus. For the R_2 rates, an overall trend of around 5 s^{-1} is observed for residues 1-67 with slight increases for residues compromised in the secondary structure.

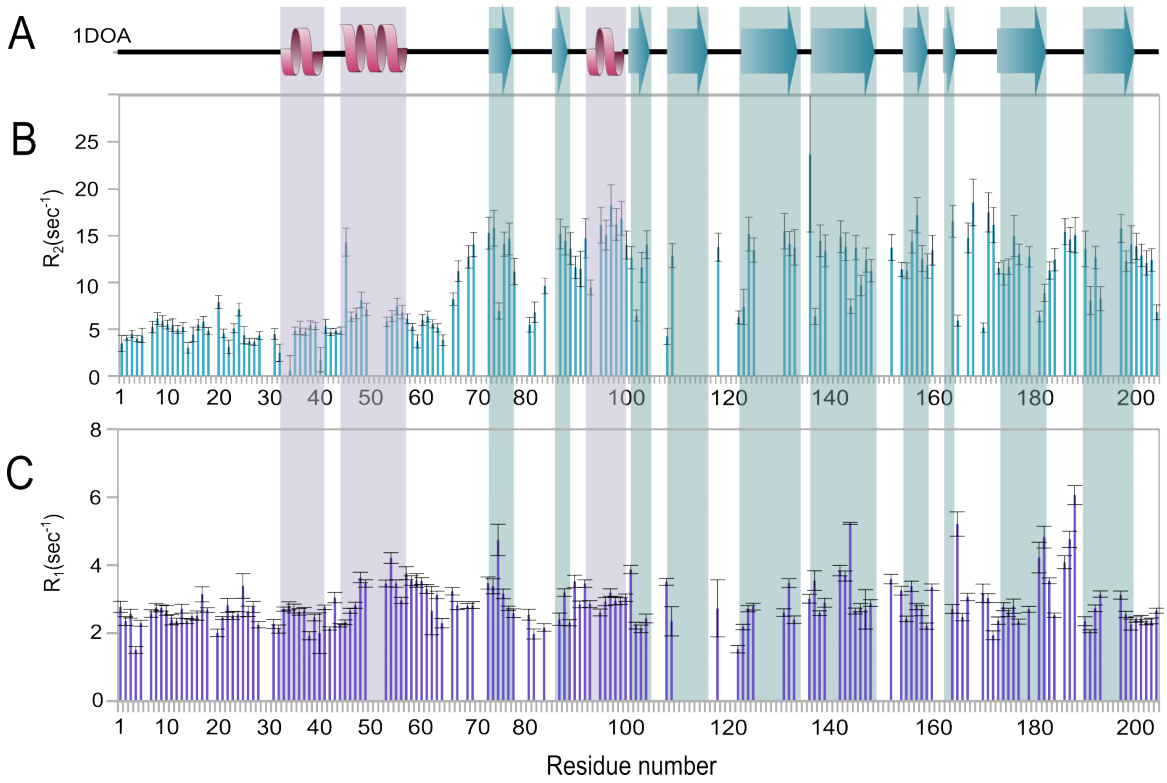


Figure 3.12. R_1 and R_2 relaxation rates for full-length RhoGDI1. **A)** X-ray structure of the RhoGDI1 complex with Cdc42 from PDB 1DOA. **B)** R_2 relaxation of the apo RhoGDI1, which confirms the existence of local (temporary) secondary structure in the sections predicted by CheSPI. **C)** R_1 relaxation rates of apo RhoGDI1.

3.5 RhoGDI1: Cdc42 complex

The amino terminal region has been described as disorder prior to formation of the complex with GTPase, and this has been assumed to be an important characteristic of the mechanism for extraction of the GTPase from the membrane. However, the experimental information obtained from the NMR studies on the full-length apo protein, and the isolated N-terminus domain disagree with this proposed mechanism.

For further understanding of the changes the N-term undergoes upon binding to the GTPase, the experimental NMR chemical shift information was analyzed by CheSPI. The predicted secondary structure for the complex does not show significant changes compared to the full-length apo protein. The comparison between both predictions can be observed in **Figure 3.13**, where we can see a good agreement with the structure obtained by X-ray for the complex in both cases. This indicates that there are no further changes in the secondary structure of the N-terminal domain after binding to Cdc42.

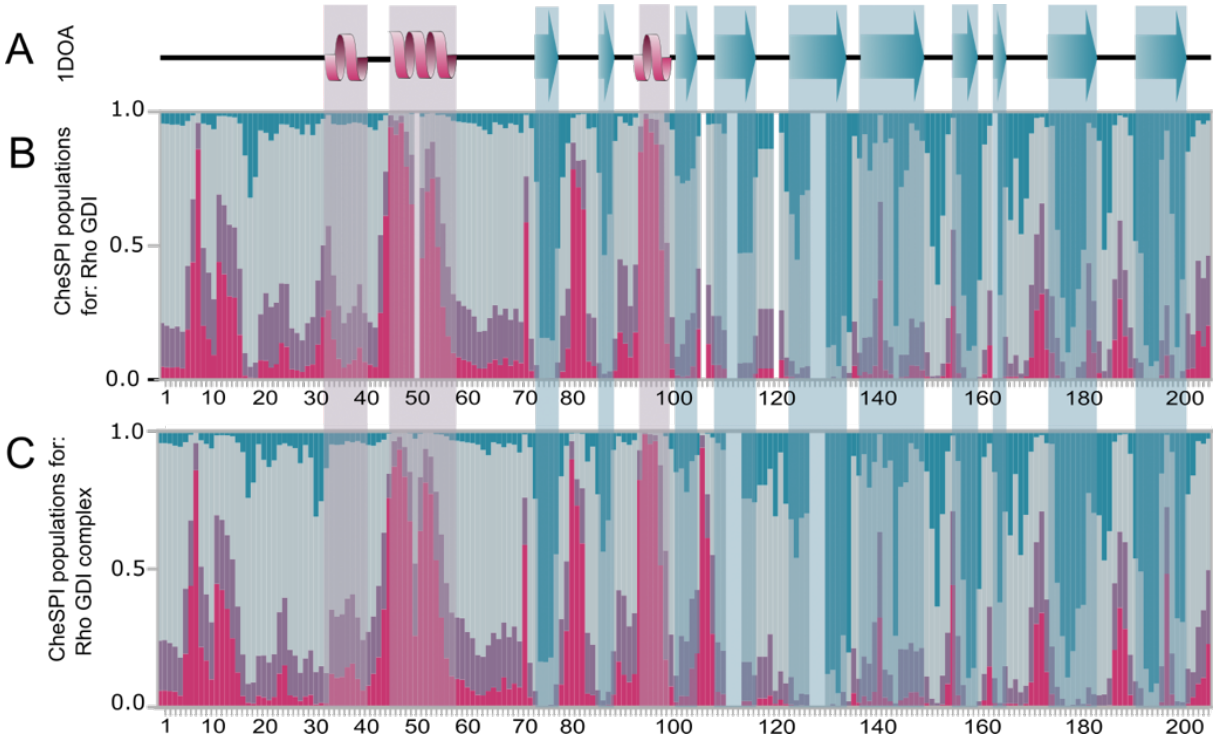


Figure 3.13. Comparison between X-Ray structure of RhoGDI1 in complex with Cdc42 (**A**), experimental chemical shift prediction by CheSPI for RhoGDI1 apo (**B**) and in complex with Cdc42 (**C**).

To corroborate the presence of the complex for the NMR measurements the R_2 rates and the chemical shift perturbation of both samples were compared. It can be observed (**Figure 3.14**) a considerable chemical shift perturbation for residues 23, 33, 45, and 46 from the amino-terminal domain as well as part of the residues in the carboxyl-terminal domain.

The R_2 rates show an overall increase of the rates, which confirms a larger molecular size of the observed protein, this due to the interaction with Cdc42 (20 kDa). An increase of the R_2 rate of residues 23, 32, 45, and 46 was observed due to the direct interaction with the switch domain of Cdc42. These residues are involved in the formation of the complex with Cdc42, on the one hand the helix-loop-helix binds to the switch domain of Cdc42. As a second step, the carboxyl creates a lipidic pocket which provides a membrane substitute for the geranylgeranyl moiety of Cdc42.

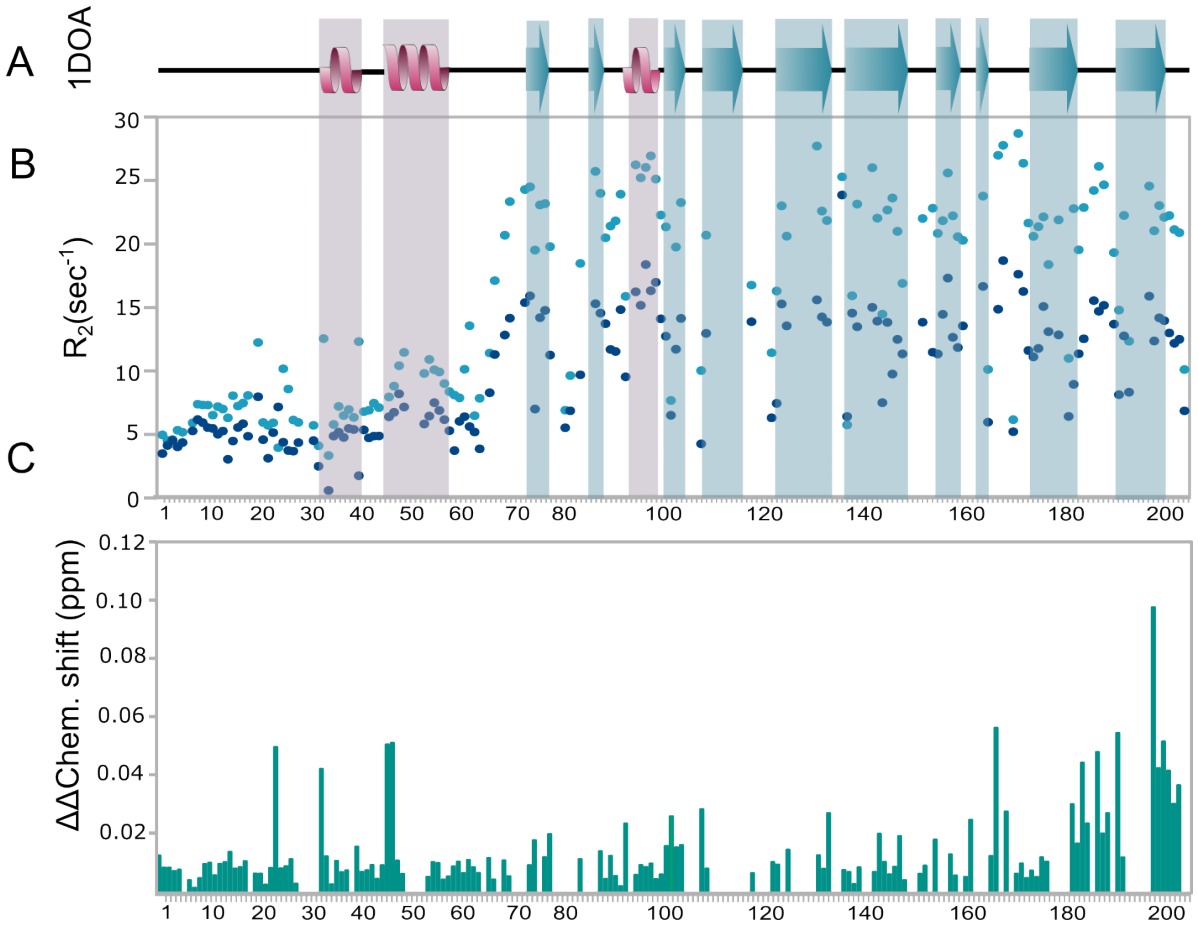


Figure 3.14. Confirmation of the presence of the complex with Cdc42. **A)** X-ray structure of the complex between RhoGDI1 and Cdc42. **B)** Comparison of the R_2 rates for RhoGD1 apo (dark blue) and RhoGDI1 in complex with Cdc42 (light blue). **C)** Chemical shift perturbation between the apo and the complex.

The analysis of the ^{15}N R_2/R_1 ratio shows an increase of the ratio for the protein in complex indicating a change in the time scale of motion. The change on the average R_2/R_1 values (minus 10% outliers) from RhoGDI1 apo to the complex with Cdc42 indicate a change in the correlation time of the molecule, which is expected due to the change in molecular size. In general, those residues with very low rations indicate fast time-scale backbone dynamics, which is for both samples the residues in the N-terminus. Residues with values substantially higher than the average hint at an intermediate timescale of motion.

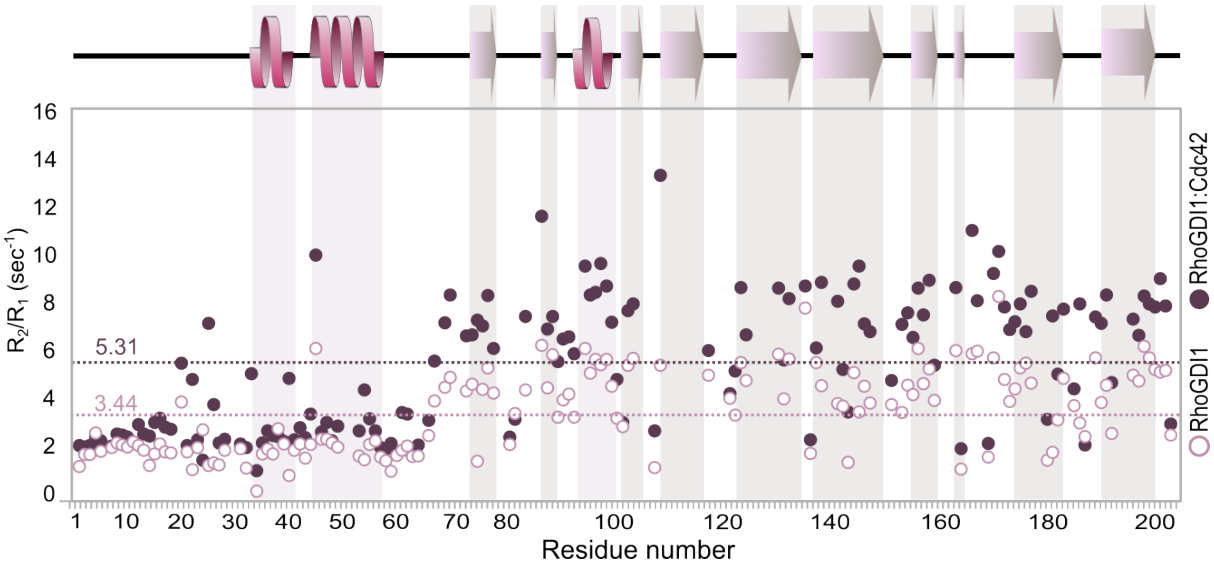


Figure 3.15. ^{15}N R_2/R_1 ratio for RhoGDI1 (open circles) and RhoGDI1 in complex with Cdc42 (dark circles) where the dashed lines indicate the average R_2/R_1 ratio value for the protein apo (pink) and the complex (dark purple) excluding 10% outliers.

In recent years, advances in bioinformatics and artificial intelligence have facilitated the development of programs that predict the structure of proteins based on the amino acid sequence. To complement the information obtained experimentally, we predicted the secondary structure using Rosetta [134]. The prediction of the secondary structure for the isolated N term (**Figure 3.16**) shows the presence of the helical element of residues 8-16 and the helix-loop-helix.

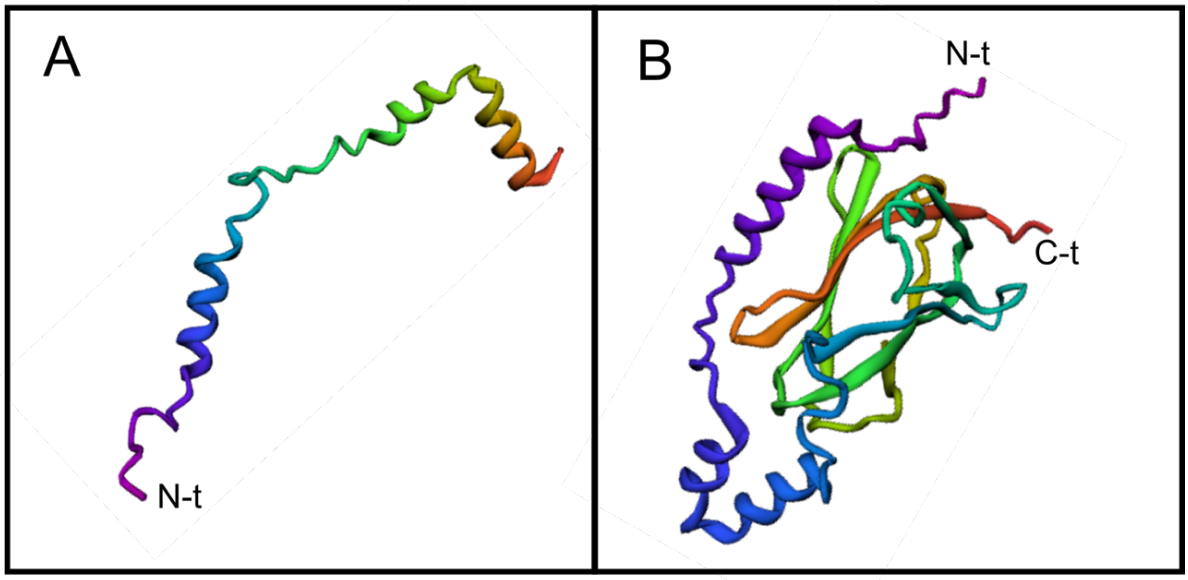


Figure 3.16. Rosetta prediction of the secondary structure for the isolated N-terminal domain (**A**) and the full-length RhoGDI1 (**B**)

Although the prediction for RhoGDI1 full length is very similar to the secondary structure observed by X-ray when the protein is in the complex, we observed the presence of the structural elements in the N-terminal domain. The predicted structures disagree with the information found in the literature, where until now the N-terminal domain of the apo protein was thought to be disorder prior to complex formation. The experimental data obtained by NMR, as well as the predictions of the secondary structure based on experimental chemical shift information, suggest the presence of the secondary structure before complex formation. Also, in combination with the Rosetta prediction, the secondary structure of the protein in the apo form contains structural elements in the N-terminal domain, leaving the question of how the mechanism of extraction of the GTPase from the membrane works open for further research.

Chapter 4

EFFECT OF MUTATION ON RT-LOOP OF SH3 DOMAIN

4 EFFECT OF MUTATION ON RT-LOOP OF SH3 DOMAIN

Spectrin is a large cytoskeletal heterodimeric protein composed of an α and a β subunit, crucial for maintaining the stability and structure of the cell membrane and the shape of the cell. Spectrin is a member of the F-actin cross-linking protein superfamily that contains proteins composed of spectrin repeats and actin. These spectrin repeats for a triple-helix supercoil connected by an α-helical linker that allows the protein to extend and contract. The protein is composed of numerous structural motifs, including actin-binding, pleckstrin homology, and *Src* homology 3 (SH3) domains that contribute to spectrin function [135]. The SH3 domain has a structure of five to eight β-strands arranged in two antiparallel β-sheets or in a β-barrel (**Figure 4.1**) [136].

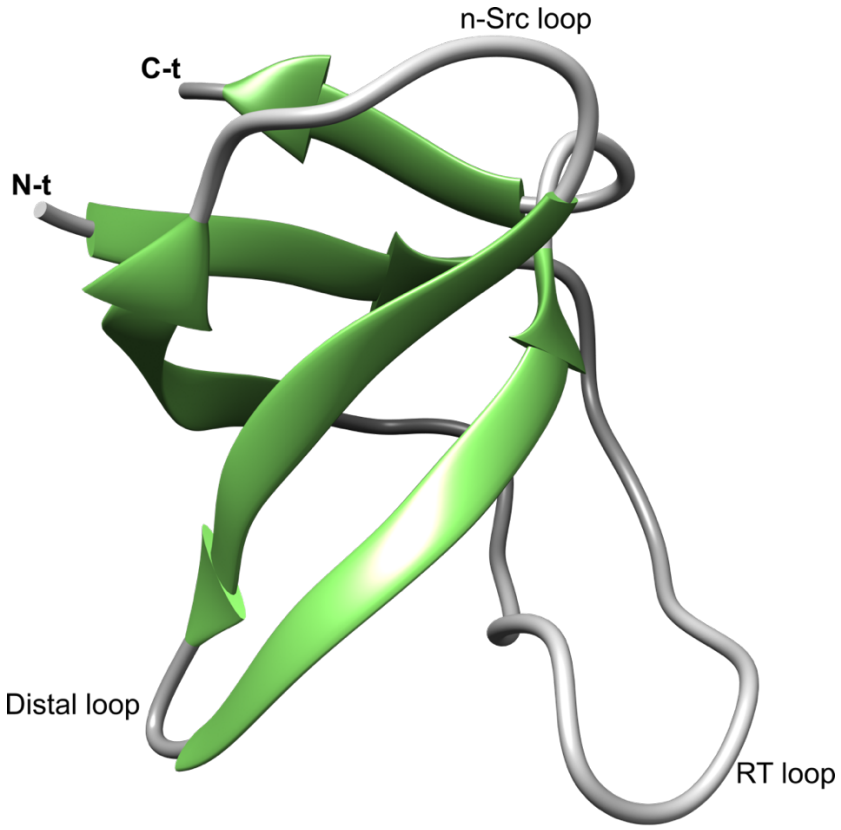


Figure 4.1. Topology of secondary structure elements of the *Src* homology 3 (SH3) domain (PDB:2NUZ).

As the SH3 domain of chicken α-spectrin has been comprehensively characterized by solution and solid-state NMR, it constitutes an ideal protein model for methods development. In this context, the SH3 domain was used by Rovó et al. [38], to study the complementary information provided by ^{1}H and ^{15}N relaxation dispersion techniques, Bloch-McConnell and near-rotary resonance, and the benefits of using this combined information in solution and solid state. A relatively low-amplitude fast microsecond motion at 0 ° C was observed for some residues that could arise from fluctuations in the intrinsic amide bond or a low-amplitude “rocking motion”,

which has been observed in previous studies for microcrystalline proteins [43, 44, 137]. To clarify the mechanistic details of the observed microsecond time-scale motion within the crystalline lattice, we pursued mutations within the RT loop and characterization of any entailed changes in the protein dynamics.

4.1 Sample preparation

4.1.1 Site-Directed Mutagenesis

Three mutations were planned for the residue arginine 21, i.e., R21A, R21E, and R21G. Mutations were introduced by site-directed mutagenesis, a method of creating specific and targeted changes in double-stranded DNA [138]. For this, three pairs of overlapped primers containing the desired mutation were designed, the mutations were introduced in the desired position by inverse PCR (**Figure 4.2**).

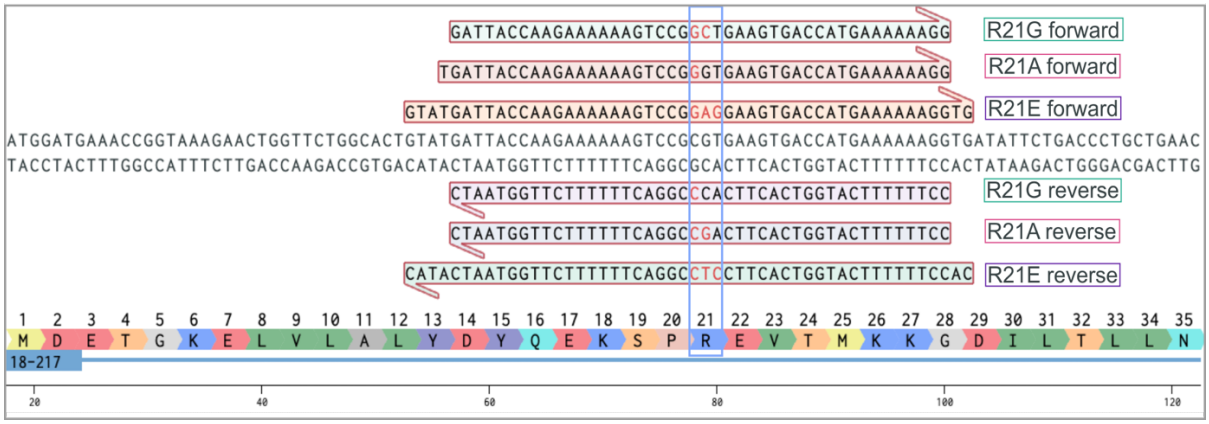


Figure 4.2. Primer design for mutation on R21 of SH3 domain. Overlapping points for single mutations are showed in red and the 3-letter codes for the amino acids implied is marked by the blue rectangle.

Inverse PCR and site-directed mutagenesis were performed using the Quick-change® Agilent kit by mixing 5µL of 10x reaction buffer, 1µL of primer forward, 1µL primer reverse, 0.5µL of original DNA [25 ng] (pET3d vector with ampicillin resistance gene) and 1µL dNTP mix. The final volume was completed to 50µL with ddH_2O. 1µl of DNA polymerase was added right before starting the PCR reaction (**Table 4.1**). Once the PCR cycling finished, 2µL of Dpn I restriction enzyme was added to the reaction PCR tube and gently mixed by pipetting several times. A transformation in XL10-Gold cells was performed in agar plates containing ampicillin and chloramphenicol. A colony for each mutant was used to start a small overnight culture in LB to extract DNA using the Miniprep Monarch® kit. The DNA extracted was sequenced, and those DNA tubes that showed a positive and correct result were retained for protein expression.

Table 4.1. Conditions used for PCR cycling upon site-directed mutagenesis.

Block	Number of cycles	Temperature	Time
1	1	95°	2 min
2	18	95°	20 sec
		60°	10 sec
		68°	2.5 min
3	1	68°	5 min

4.1.2 Protein Expression

Protein expression was carried out with freshly transformed BL21-DE3 *E. coli* cells and were grown in minimal media labeled [^{15}N, ^{13}C, ^{2}H] with 100% D_2O at 37 ° C and induced at OD_{600}~0.8 with 1 mM IPTG. Cells after induction were incubated at 20 ° C overnight and harvested after 18 h by centrifugation at 6000x g and 4°C. The pellet was stored at -80°C.

4.1.3 Protein Purification

The protein was purified by established chromatography methods. Before purification, the cell pellet was thawed at 4°C, 10 µg/mL DNAse I and 1x Complete Protease Inhibitor Cocktail tablet were added and incubated at 4°C for 1 h. Cells were disrupted in a microfluidizer at 1.5 mbar in two rounds and centrifuged at 110000x g for 45 min, the supernatant was separated by decantation. The supernatant was diluted with buffer A (**Table 4.2**) to approximately 80% of the column volume of the anion exchange column. Cell supernatant was injected into the column with a flow rate of 3 ml/min, unbound organic molecules were removed by washing with 2 CV buffer A. The bound molecules were eluted by a linear gradient of 0-9% of buffer B (8 CV). The fractions were pooled and monitored by SDS-PAGE, the SH3 containing fraction was shifted to pH 3.5 using 3M citric acid and concentrated to 10-12 mg/ml, avoiding concentrations higher than 14 mg/ml to prevent precipitation. The concentrate was further purified by Size Exclusion Chromatography (SEC) using a Superdex 75 prep grade column of a volume of 120 ml. After injection with a flow rate of 1 ml/min, 0.5 CV was eluted before fractionation of 1 CV. The SH3 fraction was poled and dialyzed against 5 L of H_2O pH 3.5 (without buffers, adjusted using HCl_{aq}).

Table 4.2. SH3 purification buffers for Anion-Exchange and Size-Exclusion Chromatography and Dialysis

Buffer	Composition
AEC Buffer A	20 mM Tris, pH 8.8 (adjusted with 30% HCl)
AEC Buffer B	20 mM Tris, 1 M NaCl, pH 8.5 (adjusted with 30% HCl)
SEC Buffer	20 mM citric acid, 150 mM NaCl, pH 3.5 (adjusted with 4M NaOH)
Dialysis buffer	H_2O, pH 3.5 (adjusted with 30% HCl)

4.1.4 Micro crystallization

The dialyzed protein was concentrated at 10 mg/ml. 100 mM $(NH_4)_2SO_4$ (from 1 mM stock solution) was added, as well as 0.03% NaN_3. Paramagnetic doping was carried out by supplementing the solution with 70 mM Cu (II)EDTA (from a 350 mM stock solution, pH 7.4). Crystallization was initialized with a pH shift to pH 8.0 by addition of ~10-15 µL NaOH (from 1M stock solution), the pH was monitored using pH indicator sticks. As soon as a pH ~8 was achieved, small crystals were visibly formed in the solution. The crystallization solution was stored in the fridge at 4°C until it was packed into the rotor. The final sample, obtained under identical procedures but from accordingly isotope-labeled M9 medium, was labeled [^{15}N, ^{13}C, ^{2}H], and 100% amide proton back-exchanged.

4.2 Assignment of mutant

4.2.1 Solution NMR

All solution NMR spectra were recorded at 298K on a Bruker Avance NEO 800 MHz (^{1}H Larmor frequency) spectrometer equipped with a CP TCI Proton Optimized Triple Resonance cryo-probe. The recorded solution NMR assignment experiment consists of 2D ^{15}N-^{1}H HSQC, 3D HNCO, and 3D HNCACB for both samples.

The SH3 domain of α-spectrin has been previously assigned in solution and solid-state [139], therefore, for the assignment of mutants in solution state NMR, the chemical shift information was transferred from previous data sets. In solution state 100% of the assigned peak were successfully transferred to both mutants, staying unassigned Met1, Pro20, and Pro54, as these residues are not observed in the ^{15}N-^{1}H HSQC.

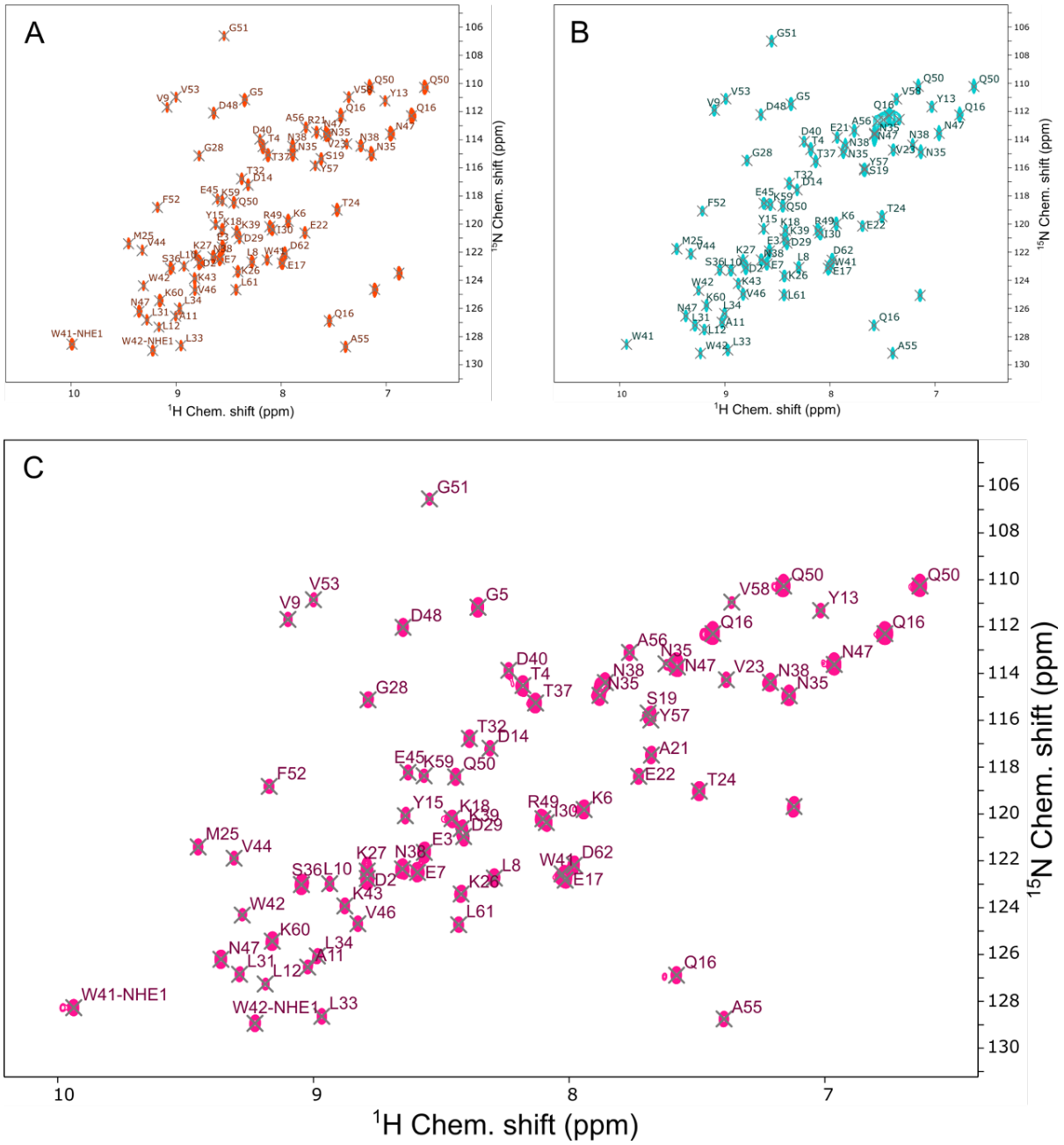


Figure 4.3. Chemical shift assignment of the 15N-1H HSQC spectra for SH3 wild-type **(A)**, SH3 R21E **(B)** and SH3 R21A **(C)**. All spectra were recorded at 25°C on an 800 MHz (^{1}H-Larmor frequency) spectrometer.

The effect of each mutation can be analyzed by looking at the chemical shift perturbations (CSP). CSPs were calculated only for the ^{1}H and ^{15}N chemical shifts information, using Equation (2.1). **Figure 4.4** illustrates the CSP for both mutants with respect to wild type (wt), for both mutants clear perturbations are observed for the mutated residue (21) and for the residues directly next to the mutation (K17, S19, and E22). While these changes in chemical

shift are expected for the mutated residue and adjacent amino acids, some slight perturbations are observed in particular for residue W41.

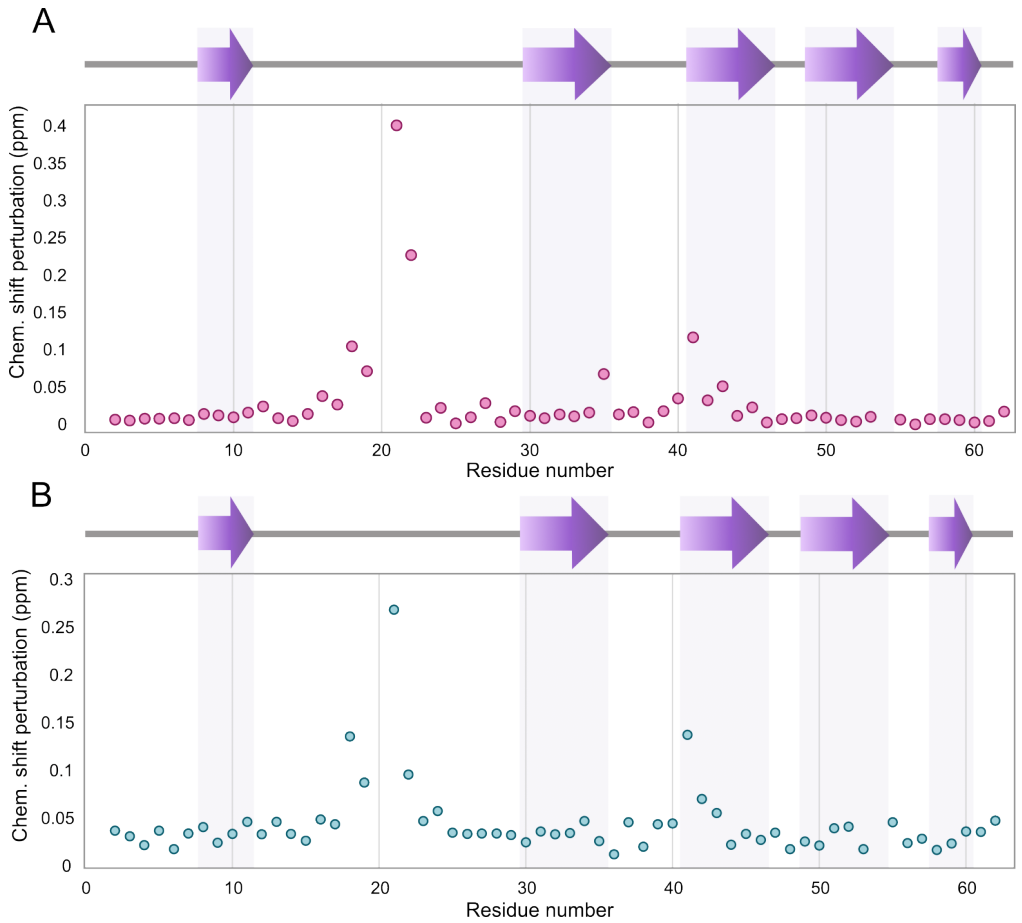


Figure 4.4. Chemical Shift Perturbations between SH3 wild-type and the mutants R21A **(A)** and R21E **(B)** plotted against the secondary structural element from the X-ray structure (PDB entry: 2NUZ).

When the CSPs between the wild-type protein and the R21A mutant are represented on the three-dimensional X-ray structure, the changes for residues N35, W41, W42 and K43 are explained by the side-chain contacts between the residues involved. **Figure 4.5** shows the superimposed X-ray structures for the wild type protein (PDB: 2NUZ) and the mutant R21A (PDB: 2F2W) with the position 21 highlighted in blue for Alanine and orange for Arginine. We can see for the mutant a shorter side chain that disrupts the interaction with the indole ring of W41, which alters the chemical environment of adjacent residues that have through-space interactions.

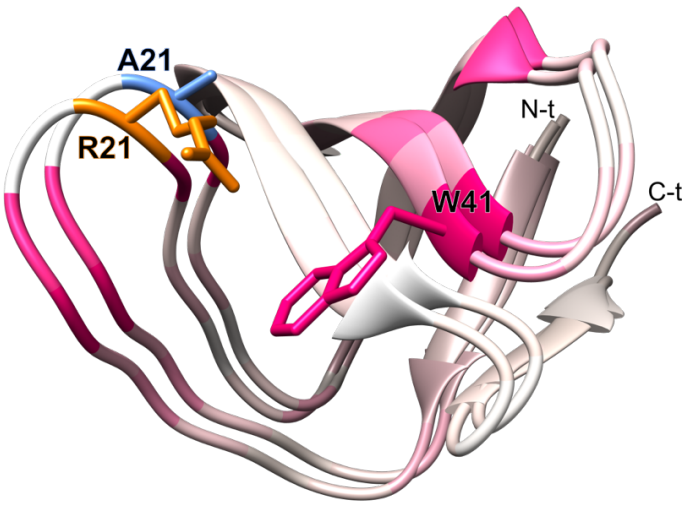


Figure 4.5. Chemical Shift Perturbations between the wild-type SH3 and the R21A mutant shown in the X-ray structures (PDB 2NUZ for SH3 WT, and 2F2W for SH3 R21A), where the sidechains for the position 21 are highlighted in orange for Arg21 and blue for Ala21.

The R21A mutation shows a larger number of perturbations, and the theoretical pI value is closer to the wild-type protein; therefore, the crystallization conditions remained similar to the wild-type. For these reasons, this sample was used for further analysis of relaxation in solution and solid-state NMR.

4.2.2 Solid-state NMR

Solid-state NMR spectra were recorded using a 1.3 mm rotor on a Bruker Avance NEO 700 MHz (^{1}H Larmor frequency) spectrometer with a PH MAS DVT 700s3 BL1.3 mm probe, a spinning frequency of 55.5 kHz and a temperature of ~25°C. Solid-state NMR assignment experiments recorded for the SH3 R21A mutant consisted of 2D 15N-1H hNH, 3D hCANH, 3D hcaCBcaNH, 3D hCAcoNH, 3D hCOcaNH and 4D hCACONH.

Although there are chemical shift assignments for wild-type SH3 [138], assignment transfer for mutant R21A required the use of a set of assignment experiments for unambiguous assignment. Assignment was carried out by linking the backbone shifts to the Cα, Cβ, and CO, enabling sequential connections and hence the completion of the backbone walk. In the solid state, only 57 residues of the 62 residues present in the protein are visible. In the nonvisible residues, we can find the two prolines (P20 and P54), in addition, the first 6 residues in the N-terminal domain, N47 and D48 are not visible due to averaging of dipolar couplings and

exchange broadening beyond detection for both, CP and INEPT-based ^{1}H-^{15}N correlations. With the set of experiments recorded, 96% of the visible residues were assigned.

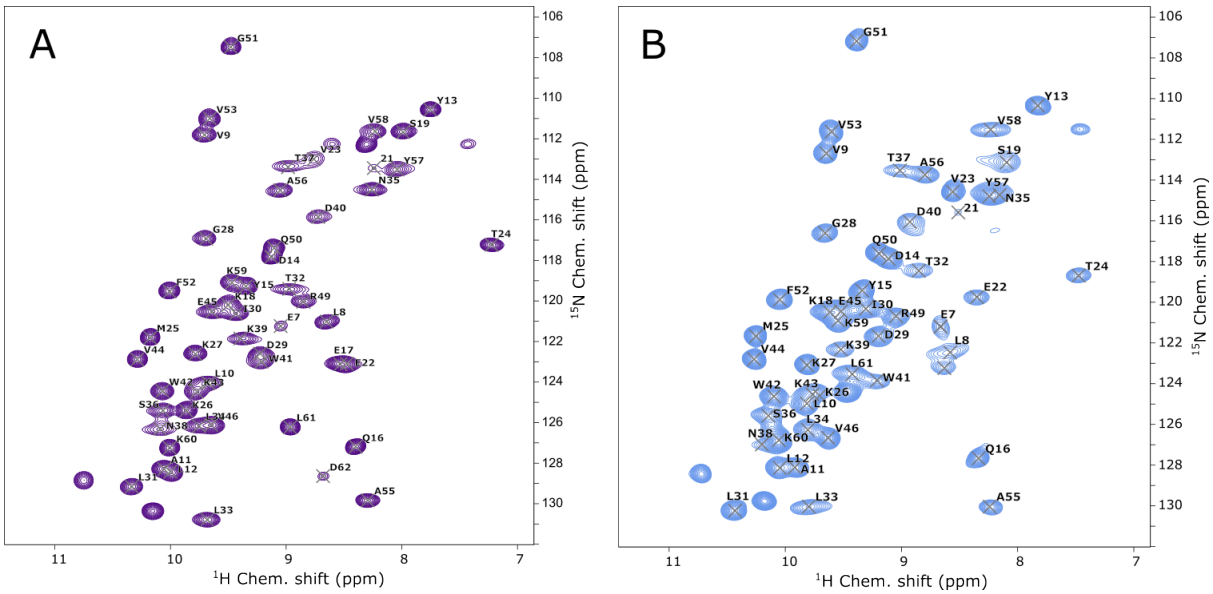


Figure 4.6. Chemical shift assignment of the ^{1}H-^{15}N hNH correlation of SH3 wild-type **(A)** and the SH3 R21A mutant **(B)**. Spectra recorded at 55.5 kHz spinning speed in a 700 MHz (1H Larmor frequency) spectrometer at ~20°C.

In contrast to the NMR chemical shift perturbations in solution, in the solid state the observed perturbations (**Figure 4.7**) are more pronounced along the entire sequence. In addition to the perturbations observed in the solution state, the largest perturbations can be observed for residues E7 and D62, residues in the N-term and C- term, respectively. The multiple chemical-shift perturbations observed in the solid state can be explained by taking into account the crystal lattice and analyzing the crystal contacts. In **Figure 4.7.C** and **D**, the most significant protein-protein interactions are illustrated. By reducing the length of the side chain in position 21, the interaction between the RT loop with the N-term and C-term residues of neighboring molecules is affected; these changes in the chemical environment are directly reflected in the chemical shift for all amino acids affected.

Protein-protein contacts in monomeric protein crystals have been studied for a long time. Although they are mainly artifacts of crystallization, they rely on the same forces involved in specific recognition in protein-protein complexes and oligomerization [140]. Statistical studies showed that the amino acids involved in this interaction are solvent accessible, and the hydrophobic residues show a higher probability of being part of the interaction. This suggests that crystal contacts do not occur in randomly selected regions of the protein. The role

hydrophobic interactions play in protein-protein recognition is well understood. On the contrary, studying the modulating effects of this steric change on the dynamics of the protein can help to better understand the intermolecular interactions within the crystal. Information on differential steric modulation of dynamics is also of interest for different applications, such as structurally guided drug designs, designer proteins, or understanding protein interactions and complex formation evolving upon evolution or altered in the disease case [141-143].

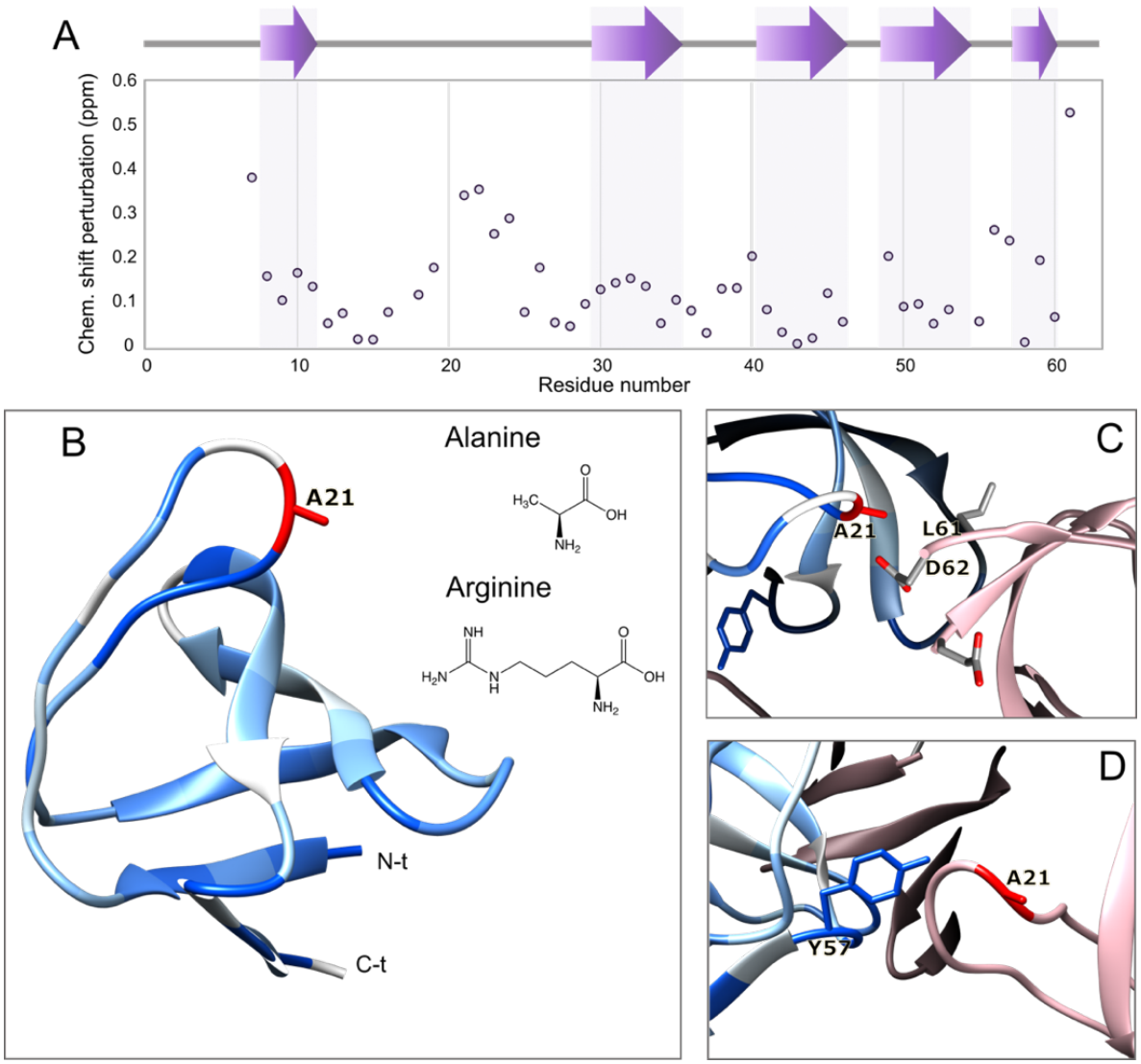


Figure 4.7. Chemical shift perturbation between SH3 wild-type and the SH3 R21A mutation **(A)** and the depiction in the X-ray structure (in blue, PDB: 2F2W) with the Ala21 highlighted in red **(B)**. **C and D)** Illustration of the crystal contacts between A21 with L61 and A21 with Y57.

4.3 Solution Relaxation

4.3.1 Experimental conditions

^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates measurements were acquired interleaved as follows. The experiments were acquired using a recycle delay of 1.5 s between the experiments and the following delays for T_1: 10, 50, 100, 200, 300, 400, 500, 600, 800, 1020, 200 and 600 ms; and for T_2: 10, 30, 50, 70, 90, 110, 130, 150, 170, 190, 50, 110 and 170 ms. Heteronuclear steady-state NOE (hetNOE) spectra were acquired with two different data sets, one recorded without an initial proton saturation and the other with an initial proton saturation period of 3 s.

4.3.2 Fast dynamics

To interrogate fast time-scale dynamics upon mutations in solution, we looked at the ^{15}N R_1, R_2, and steady-state NOE data and compared their contributions to the spectral density. When the R_1 and R_2 The relaxation rates of wild-type SH3 are compared with the rates of mutant SH3 R21A, we observed a decrease in the R_1 rates for some amino acids in the mutant. In contrast, an increase in the R_1 rate of Glu17 is observed, which indicates an increase in motion. The R_2 The mutant rate shows a slight decrease compared to that of the wild-type protein; the decrease is observed for most of the residues to a greater or lesser extent, which we ascribe to a decreasing hydration shell without the charged and extended Arg residue. In addition, the comparison of the steady-state NOE values between the wild-type and mutant shows slight differences between most of the resides. Another way to look at these values is with the R_2/R_1 rate, in which those residues with a very low ratio indicate significant fast-timescale backbone dynamics, and those higher than the average suggest intermediate timescale motion. With the calculated trim media excluding 10% of outliers, we can calculate the overall rotational correlation time, where we can appreciate slight decrees when the mutant is introduced.

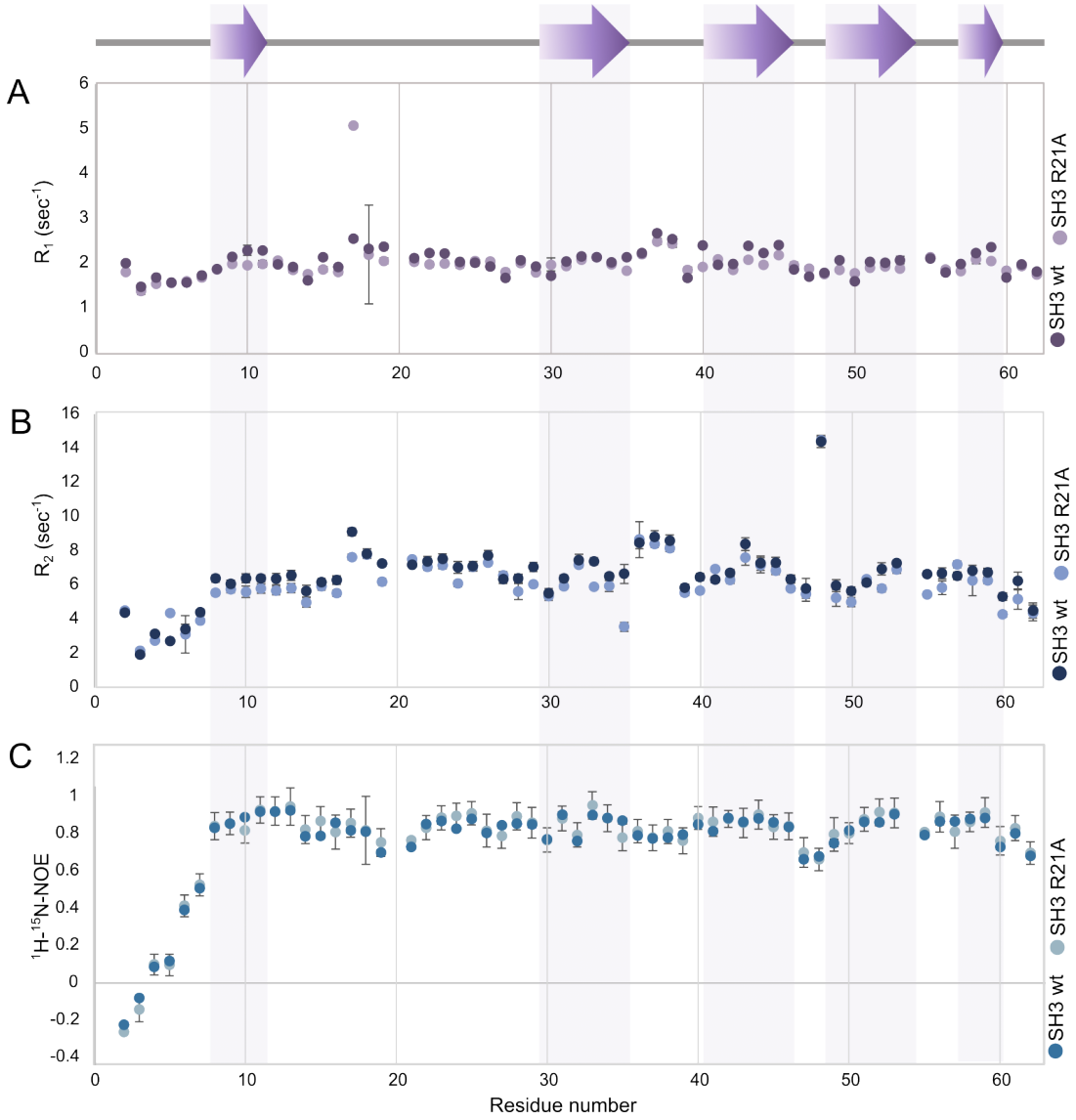


Figure 4.8. Longitudinal (R_1) and transverse (R_2) relaxation rates for SH3 wild type and R21A mutant. **A)** Comparison of the R_1 relaxation rate for SH3 wild type (dark purple) and R21A mutant (mauve). **B)** Comparison of the R_2 relaxation rate for SH3 wild type (dark blue) and R21A mutant (light blue). **C)** Comparison of steady-state NOE values for SH3 wild type (turquoise) and R21A mutant (light blue).

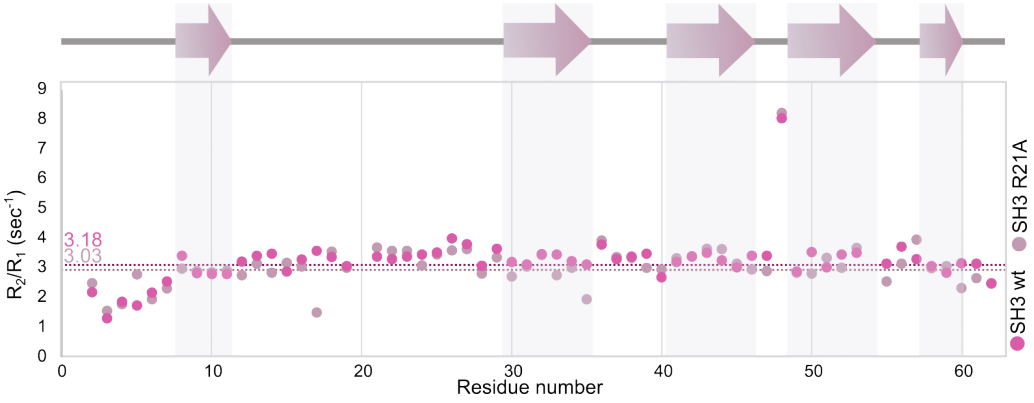


Figure 4.9. R_2/R_1 ratio for SH3 wild type (deep pink) and R21A mutant (pale pink). Dashed lines indicated the average of the ratio minus 10% outliers.

The spectral density function, $J(\omega)$, describes the dynamics of the amide bond vectors in the fast regime of motion, on the picosecond to nanosecond time scale. With R_1, R_2 and steady-state NOE measurements at one field, we can perform the spectral density mapping, a mathematical procedure that reconstructs the spectral-density contribution $J(\omega)$ at three frequencies, $J(0)$, $J\omega_N$, and $J(0.87\omega_H)$. For extracting these values, we used the model-free approach [144] with an *in-house* written script. The results obtained confirmed a decrease in $J(0)$ in the mutant sample for some residues, indicating that the global motion has become faster (maybe owing to a reduced effective radius of gyration due to fewer interactions with water molecules in the protein's hydration shell), but the local motion remains similar, as indicated by the $J(\omega_N)$ and $J(0.87\ \omega_H)$.

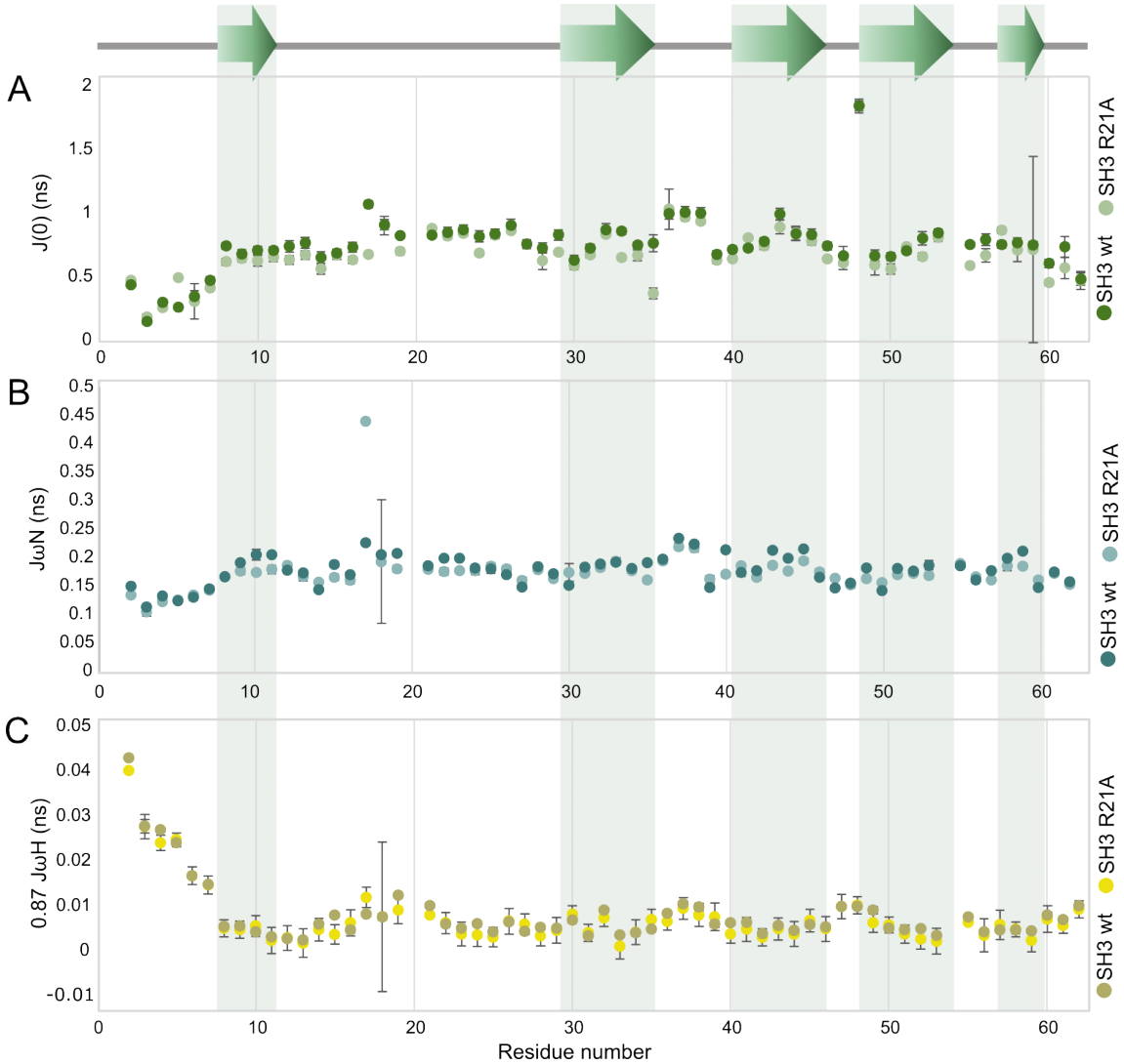


Figure 4.10. Comparison of spectral density mapping obtained by the model free approach showing the frequencies $J(0)$ (**A**), $J(\omega_N)$ (**B**), and $J(0.87\omega_H)$ (**C**) for SH3 wild-type and SH3 R21A. The corresponding color for each plot is indicated in the figure.

4.4 Solid-state Relaxation Dispersion

Using solid-state NMR spectroscopy, dynamic processes that occur on the µs-ms timescale can be probed via analysis of coherence lifetimes in the presence of a spin lock field, i.e., via ^{15}N rotating frame spin-lattice relaxation ($R_{1\rho}$) rates.

All ^{15}N off-resonance rotating frame spin-lattice relaxation ($R_{1\rho}$) rate experiments were recorded on a Bruker Avance NEO 700 MHz (^{1}H Larmor frequency) spectrometer equipped with a PH MAS DVT 700s3 BL1.3 mm probe, using spinning frequency of 55.5 kHz and an effective temperature of ~20°C. For the relaxation dispersion analysis, eleven effective irradiation field strengths (ω_e) were employed: 10, 15, 20, 25, 30, 34, 38, 42, 44, 50, and 52 kHz. Each field was used in conjunction with eight relaxation delays: 5, 10, 15, 20, 30, 40, 60, and 80 ms. A recycle delay of 1.5 s was used for lower ω_e and 2 s for high ω_e (42 and 44 kHz). Off-resonance irradiation was calculated at an effective tilt angle of ±35.3° and the carrier was set at an according offset $\pm\boldsymbol{\Delta\Omega}$ (with alternating positive or negative sign) from the middle of the spectrum. To obtain the desired values, a calibration of ^{15}N RF fields was performed before every set of experiments. The experimental parameters used for all ^{15}N off-resonance rotating frame spin-lattice relaxation ($R_{1\rho}$) rate experiments are shown in **Table 4.3**.

Table 4.3. Experimental parameters for the NERRD analysis for both samples

$\omega_e/2\pi$(kHz)	$\omega_1/2\pi$(kHz)	$\Delta\Omega$ (Hz)
10	5.8609	8139.29
15	8.7144	12208.94
20	11.6192	16278.59
25	14.5240	20348.24
30	17.4289	24417.89
34	19.7527	27673.61
38	22.0766	30929.33
42	24.4004	34185.05
44	25.5623	35812.91
50	29.0481	40696.49
52	30.2101	42324.35

The relative intensity of each signal from the final processed spectral series of ^{15}N $R_{1\rho}$ was fitted to an exponential decay for extracting the relaxation rates using equation (2.6). The difference of ^{15}N $R_{1\rho}$ rates at 10 kHz $\omega_e/2\pi$ between SH3 wild-type and SH3 R21A mutant is shown in **Error! Reference source not found.**. Rates stay the same for most of the residues e xcept for S19, A21, T24, N38, K39, W41, and R49. Interestingly, in addition to changes directly at the mutation site, some of these differences can be traced back to the different crystal contacts present in both samples, as observed in the CSP. The absence of relaxation rates for E7, E17, and D62, is due to the absence of the peaks for these residues. From the solution NMR studies for fast-timescale dynamics as mentioned above, an increase of motion is observed for E17, which can explain the absences of the signal for E17. Similarly, the absence of E7 and D62 is ascribed to the increase in motion of these two residues due to the disruption of the sterically demanding interaction with the sidechain of R21 by the introduction of the Ala mutant.

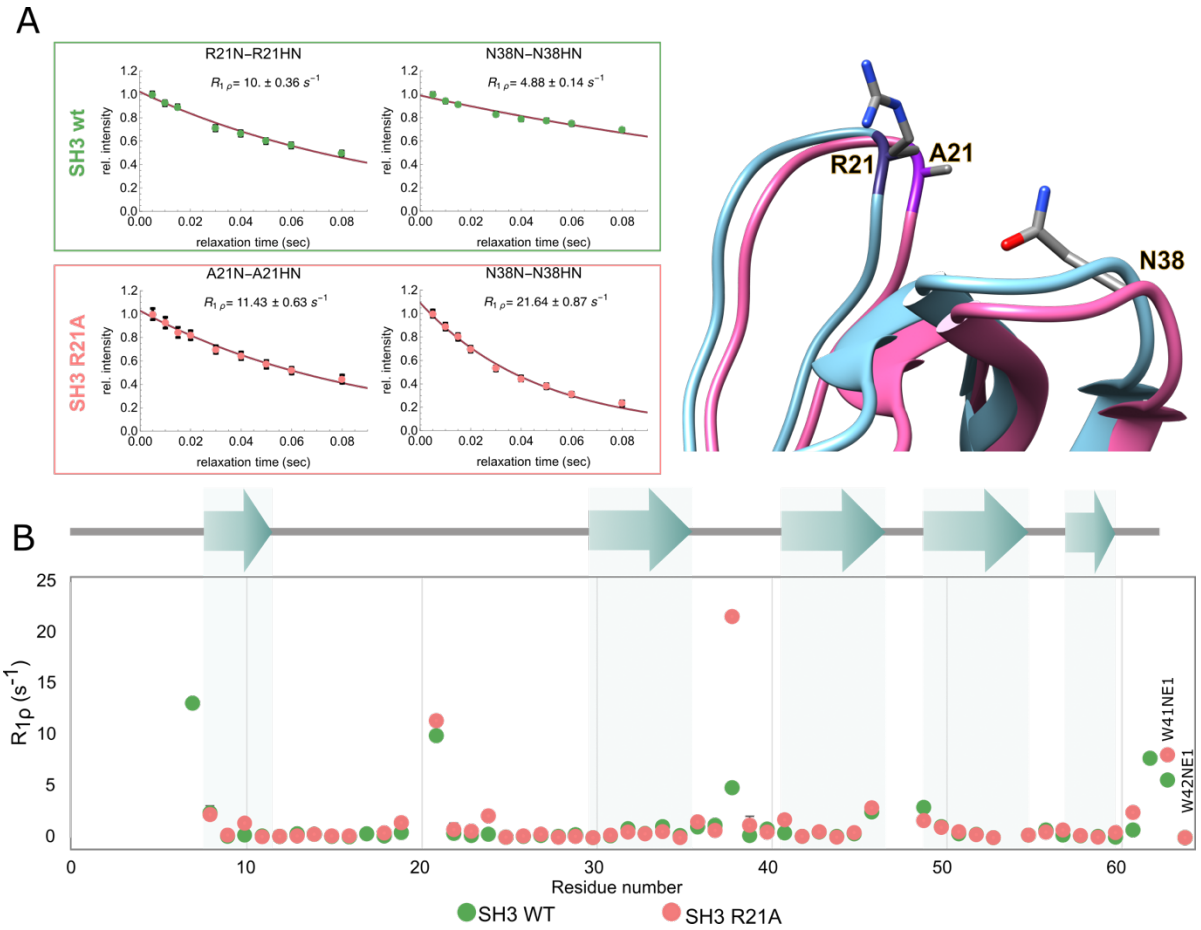


Figure 4.11. Comparison of ^{15}N off-resonance rotating frame spin-lattice relaxation ($R_{1\rho}$) rate at 10 kHz ω_e between SH3 wild-type and SH3 R21A mutant. **A)** Individual ^{15}N $R_{1\rho}$ rates for R21/A21 and N38 and the illustration of the contact in the X-ray structure (PDB:2F2W). **B)** ^{15}N $R_{1\rho}$ rates SH3 WT (green) compared with SH3 R21A mutant (pale pink) and plotted along the sequence.

For $R_{1\rho}$ experiments with an increase in the effective field strength, an increase in the relaxation rate is expected to be observed when approaching the recoupling resonance condition. Those residues involved in a slower motion will have a larger difference between the relaxation rates from the lowest field $\omega_e/2\pi$ (10 kHz) and $\omega_e/2\pi$ fields close to the resonance condition. For residues with faster motion around a few µs the difference can be measure for an $\omega_e/2\pi$ field not immediate to the recoupling condition, i.e., ~38 kHz. This approach was used to assess differences in trends after mutation. The most notorious changes are observed for residues close to the mutation site and regions close to altered crystal contacts. For N38, a negative difference for SH3 R21A is observed for both comparisons; this residue undergoes slow-timescale motion, with a $\tau_{ex} = 19.1 \pm 0.2$ µs and an $S^2 = 0.993 \pm 0.002$ (**Figure 11.4**) for the SH3 wild type protein. However, after mutation, the motion becomes faster than the detection limit of the methods, resulting in larger differences between $R_{1\rho}$ rates at different ω_e fields far away from the resonance condition. The observed increase in the motional timescale can be attributed to the disruption of the intermolecular interaction between the side chains of both amino acids.

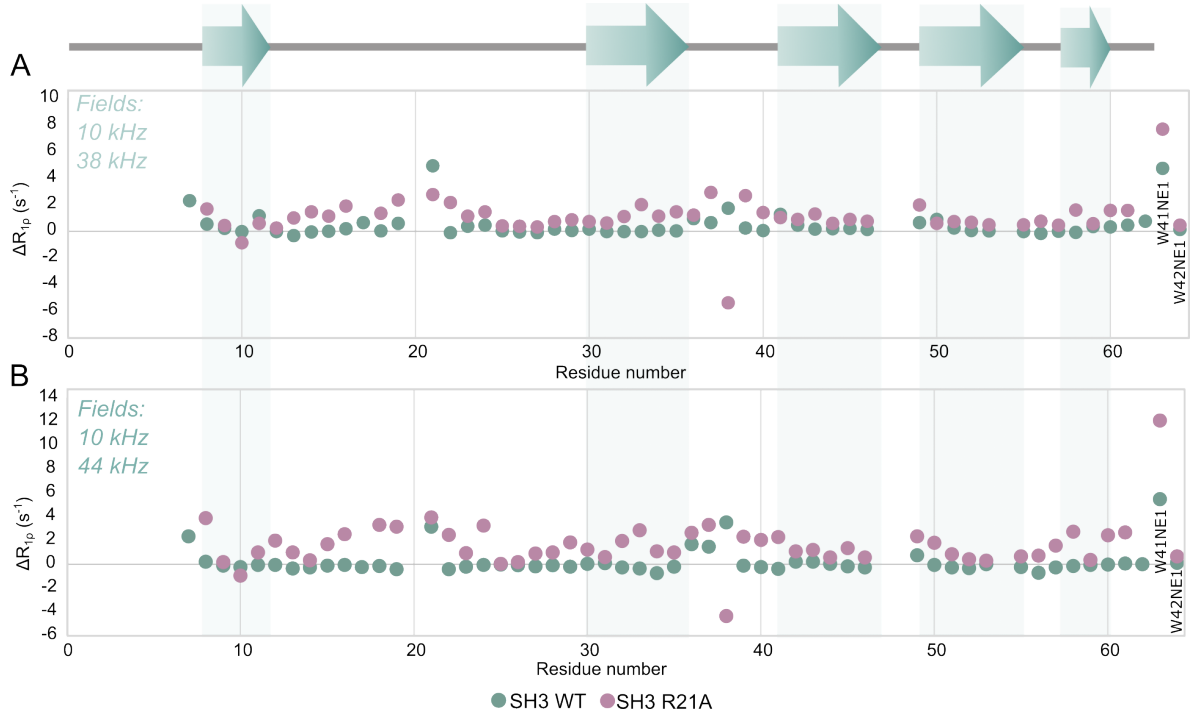


Figure 4.12. ^{15}N off-resonance rotating frame spin-lattice relaxation ($R_{1\rho}$) rates differences between 10 kHz $\omega_e/2\pi$ with 38kHz (**A**) and 44 kHz (**B**) $\omega_e/2\pi$ SH3 wild type is shown in green and SH3 R21A is represented in mauve.

The changes in dynamics introduced by the mutation can be analyzed using the generalized order parameter, which characterizes the spatial restriction of the motion, and where $S^2 = 0$

represents a system fully disordered system and $S^2 = 1$ a system completely rigid system. The order parameters obtained for both proteins showed a high level of order. However, when the mutation is introduced, a decrease in the order parameters can be observed for several residues, in which L8, N38, and W41NE1 stand out with a substantial decrease in the order parameters that indicates increased-amplitude motion. In **Figure 4.13** the order parameters are depicted as $1 - S^2$ in the three-dimensional X-ray structure, where a decrease in the level of rigidity for N38 can be observed upon mutation, as the intramolecular interaction is loosened by the change of the side chain of residue 21 toward a more hydrophobic one, replacing the former electrostatic interactions with the positive Arg.

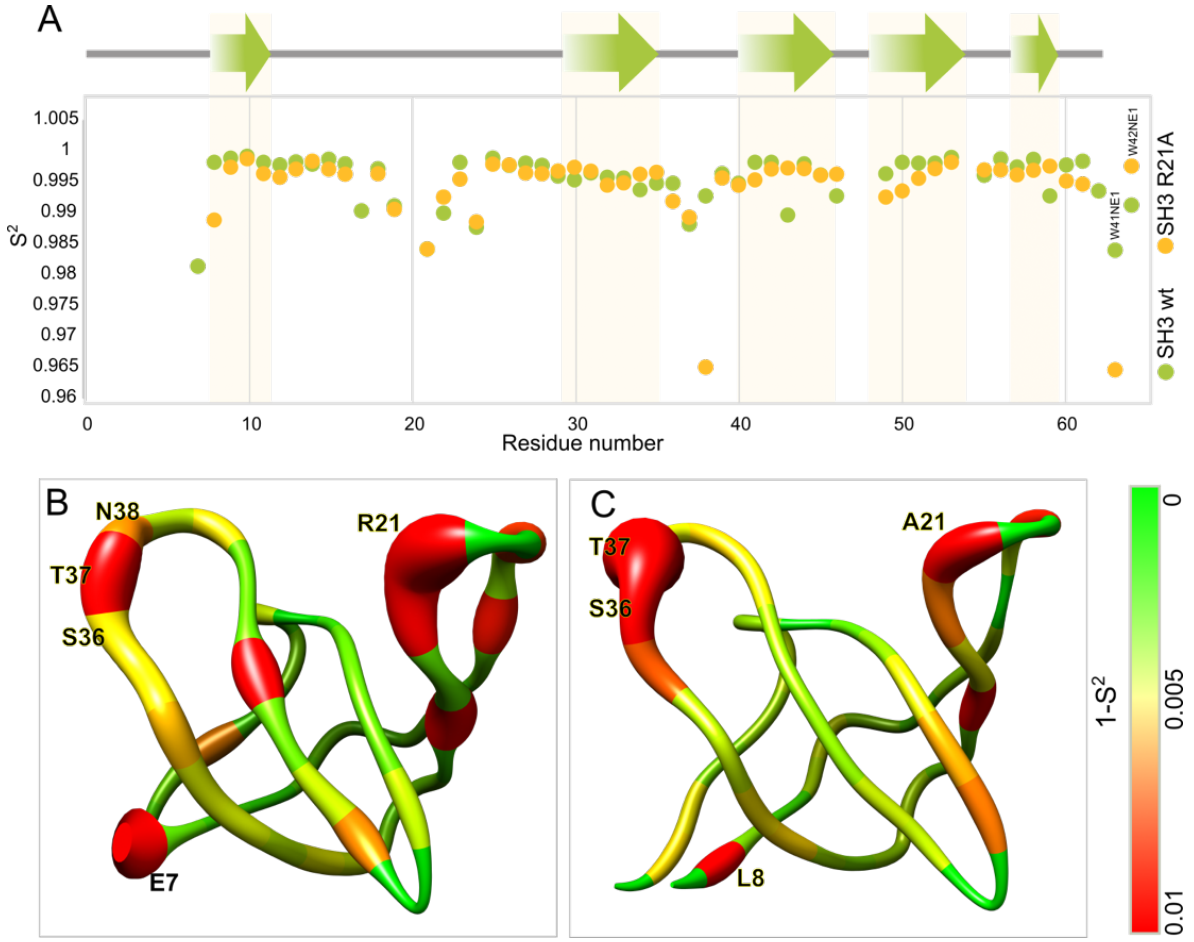


Figure 4.13. Comparison of order parameters for SH3 wild type and SH3 R21A. **A**) Obtained order parameters (S^2) from 15N NERRD analysis illustrated along the sequence, where SH3 wild type is presented in green and SH3 R21A in yellow. **B**) $1 - S^2$ of SH3 wild type depicted on the three-dimensional X-ray structure (PDB entry 2NUZ). **C**) $1 - S^2$ of SH3 R21A depicted on the three-dimensional X-ray structure (PDB entry 2WFC). The color code for the X-ray structure is shown next to the structures.

The complete ^{15}N NERRD analysis shows an increase in the intermediate time scale of motion for most of the residues that shows a NERRD bump for SH3 R21A compared to the wild-type protein. It is observed that residues L8, R21, and L61 show a NERRD profile in the SH3 wt,

and the profile remains after mutation for L8 and L61, but they show a reduced τ_{ex}. On the contrary, A21 shows an almost 'flat profile', indicating that the timescale of motion is outside (faster than) the µs to ms regime. The residues mentioned above are involved in intermolecular interaction taking part in the crystal lattice, which is modulated by the mutation.

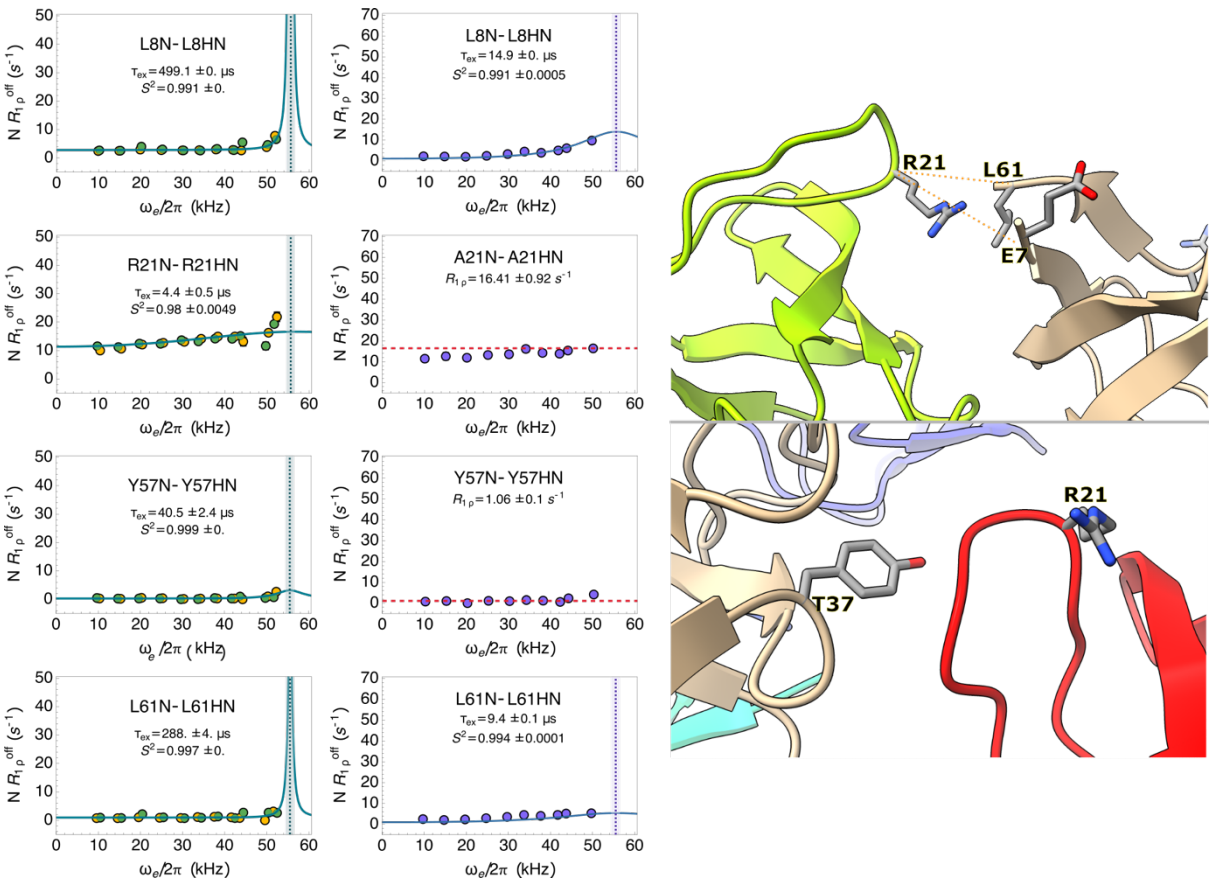


Figure 4.14. ^{15}N NERRD profiles of SH3 wild type (green) and SH3 R21A (purple) and the pictorial representation of the intermolecular interaction in the crystal lattice.

These results shed quantitative light onto at the modulation of the dynamics for a protein system by the intermolecular interactions. The implications of the observation of this effect may be of interest for the understanding of motional networks along the three-dimensional architecture within other, larger proteins, where intermolecular interactions play an important role in protein recognition or complex formation with other proteins.

Chapter 5

DISCUSSION AND OUTLOOK

5 DISCUSSION AND OUTLOOK

This work presents several tools and methodologies that use solution and solid-state NMR to characterize protein structure and dynamics at different timescales of motion. The application of these methodologies can answer questions about the dynamical changes a protein undergoes when interact with another molecule in the media, which can be another protein, the intermolecular interaction with themselves in a crystal lattice, or the interactions with small molecules that have the ultimate goal of being used to treat medical conditions.

Using solid-state NMR, we can access large protein systems, as is p38α. With the use of high-dimensional experiments, we can resolve the overlapped regions of the spectrum, which allows us to obtain information of the correspondent chemical shifts and acquire relaxation dispersion experiments that are not completely affected by the overlap level. With increasing molecular size, the correlation time of the protein increases and so does the required experimental time. To alleviate this problem, we established a paramagnetic doping protocol for p38α which causes the T_1 relaxation time to decrease and allows us to decrease the recovery delay time, and thus decreasing the experimental time. The combination of paramagnetic doping and nonuniform sampling decreases the measurement time required per experiment even further. Here we established the conditions that permit acquiring 3D-based relaxation dispersion experiments for large proteins in a reasonable time.

The solution NMR experiments provided information on the intermediate-time-scale dynamics of p38α. From the combination of chemical-shift perturbation and 15N CPMG relaxation dispersion experiments, the presence of perturbations and changes along the entire sequence was observed when disturbing one region of the protein. This hints at the presence of an allosteric dynamic network in the protein that can be further studied by using different types of inhibitors or mutations that disrupt the communication network. An additional complement to the information obtained by solution NMR, are the relaxation dispersion profiles from the NERRD analysis in the solid state. However, before more conclusions can be extracted from this analysis, it is necessary to improve the data analysis and include more residues assigned.

In the same line, to complete the solid-state NMR assignment, it is necessary to improve the data provided to FLYA. For this purpose, new experiments are planned to be recorded, including the acquisition of an HNcoCANH 5D experiment on a 1.2 GHz spectrometer and the introduction of the HNCAHA/HNcoCAHA interleaved experiments, to include the ^{1}Hα chemical shift information.

The experimental chemical shift provided information of the secondary structural elements of the N-terminal domain of RhoGDI1, and in combination with relaxation data the presence of these structural elements in the protein was established before complex formation with the GTPase client. Information was complemented by the observation of NOE contacts between residues involved in the α-helical elements of the amino terminal domain, a section of the apo protein for which the secondary structure information did not exist, as it was assumed to be disorder.

Furthermore, the vast information obtained by the combination of solution and solid-state NMR was demonstrated by the analysis of a mutation in the RT loop of the SH3 domain of chicken α-spectrin. Here, the fast-timescale motion was obtained by solution NMR and the intermediate timescale was probed by using the ^{15}N NERRD experiments. With this approach, we prove the modulation of the dynamics by the intermolecular protein-protein interactions, also known as crystal contacts.

6 PUBLICATIONS

1. K. Grohe, S. Patel, C. Hebrank, **S. Medina**, A. Klein, P. Rovó, S. K. Vasa, H. Singh, B. Vögeli, L. V. Schäfer, R. Linser, "Protein motional details revealed by complementary structural-biology techniques", *Structure (Cell Press)*, 28, 1-11, DOI: 10.1016/j.str.2020.06.001 (2020).

2. **S. Medina-Gomez**, I. Visco, F. Merino, P. Bieling, R. Linser "Transient structural properties of the Rho GDP-dissociation inhibitor. Submitted JACS

7 CONFERENCES, WORKSHOPS, AND FUNDING

Conferences

1. 62nd Experimental Nuclear Magnetic Resonance Conference (ENC), 29-31 March 2021, virtual conference.
Poster Presentation: “solid-State NMR study of the dynamics of human p38α MAP Kinase”, S. Medina-Gomez, S.K. Vasa, R. Linser

2. Scientific colloquium “Tardes de ciencia”, by the Institutional Graduate program in Chemical and Biochemical Sciences, Autonomous University of Yucatan, 16 April 2021, virtual
Invited talk: “Nuclear Magnetic Resonance of proteins and applications in Structural Biology”

3. NMR: a tool for biology conference, 2-4 May 2022, Paris, France.
Poster presentation: “Unexpected secondary structure propensity for the N-term of RhoGDI1”, S. Medina-Gomez, I. Visco, P. Bieling, R. Linser

4. Alpine conference on magnetic resonance in solids, 4-7 September 2022, Chamonix-Mont-Blanc, France.
Round table presentation: “Dynamics of human p38α MAP Kinase by solid-state NMR”, S. Medina-Gomez, S.K. Vasa, A. Klein, R. Linser

5. Academic conferences part of the XLII Chemistry week, Chemistry faculty of the Autonomous University of Yucatan, 1-2 December 2022, online.
Invited talk: “Ultrafast MAS: the magic behind solid-sate NMR of proteins”

Workshops

1. “Biomolecular NMR: Protein Dynamics”, 30th September to 4th October 2019, Swedish NMR Center, University of Gothenburg, Sweden.
2. 6th Winter School on Biomolecular solid-state NMR, 5-10 January 2020, Stowe (VT), USA.
3. Ampere Biological solid-state NMR school, 12 January to 12 May 2021, online.

Funding

1. Student grant of the Alpine conference on Magnetic Resonance in Solids, 2022.

8 REFERENCES

1. Marth, J.D., *A unified vision of the building blocks of life.* Nat Cell Biol, 2008. **10**(9): p. 1015-6.
2. Sun, P.D., C.E. Foster, and J.C. Boyington, *Overview of protein structural and functional folds.* Curr Protoc Protein Sci, 2004. **Chapter 17**: p. Unit 17 1.
3. Vasa, S.K., P. Rovó, and R. Linser, *Protons as Versatile Reporters in Solid-State NMR Spectroscopy.* Accounts of Chemical Research, 2018. **51**(6): p. 1386-1395.
4. Cavanagh, J., J. Cavanagh, and ScienceDirect, *Protein NMR spectroscopy : principles and practice*. 2nd ed. 2007, Amsterdam ; Boston: Academic Press.
5. Levitt, M.H., *Spin dynamics: basics of nuclear magnetic resonance*. 2013: John Wiley & Sons.
6. Keeler, J. and VleBooks, *Understanding NMR spectroscopy*. 2nd edition ed. 2010, Chichester, U.K: John Wiley and Sons.
7. Patching, S.G. *NMR-Active Nuclei for Biological and Biomedical Applications*. 2016.
8. Schanda, P. and M. Ernst, *Studying dynamics by magic-angle spinning solid-state NMR spectroscopy: Principles and applications to biomolecules.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2016. **96**: p. 1-46.
9. Hu, Y., et al., *NMR-Based Methods for Protein Analysis.* Analytical Chemistry, 2021. **93**(4): p. 1866-1879.
10. Pervushin, K., *Impact of Transverse Relaxation Optimized Spectroscopy (TROSY) on NMR as a technique in structural biology.* Quarterly Reviews of Biophysics, 2000. **33**(2): p. 161-197.
11. Pervushin, K., et al., *Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in 13C-Labeled Proteins.* Journal of the American Chemical Society, 1998. **120**(25): p. 6394-6400.
12. Pervushin, K. and E. Tan, *TROSY*, in *Encyclopedia of Biophysics*, G.C.K. Roberts, Editor. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 2667-2667.
13. Fernández, C. and G. Wider, *TROSY in NMR studies of the structure and function of large biological macromolecules.* Current Opinion in Structural Biology, 2003. **13**(5): p. 570-580.
14. Duer, M.J., *Introduction to Solid-State NMR Spectroscopy*. 2005: Wiley.
15. McDermott, A.E. and T. Polenova, *Solid-state NMR studies of biopolymers*. NMR handbooks. 2010, Chichester, West Sussex: John Wiley & Sons. xx, 572 pages : illustrations (some color).
16. Reif, B., et al., *Solid-state NMR spectroscopy.* Nature Reviews Methods Primers, 2021. **1**(1): p. 2.
17. Hartmann, S. and E. Hahn, *Nuclear double resonance in the rotating frame.* Physical Review, 1962. **128**(5): p. 2042.
18. Levitt, M.H., D. Suter, and R.R. Ernst, *Spin dynamics and thermodynamics in solid‐state NMR cross polarization.* The Journal of Chemical Physics, 1986. **84**(8): p. 4243-4255.
19. Gan, Z.-H. and D.M. Grant, *Rotational resonance in a spin-lock field for solid state NMR.* Chemical Physics Letters, 1990. **168**(3): p. 304-308.
20. Gan, Z., D.M. Grant, and R.R. Ernst, *NMR chemical shift anisotropy measurements by RF driven rotary resonance.* Chemical Physics Letters, 1996. **254**(5): p. 349-357.
21. Lin, J., et al., *Dipolar recoupling in solid state NMR by phase alternating pulse sequences.* Journal of Magnetic Resonance, 2009. **197**(2): p. 145-152.

22. Chevelkov, V., et al., *Ultrahigh resolution in proton solid-state NMR spectroscopy at high levels of deuteration.* Angew Chem Int Ed Engl, 2006. **45**(23): p. 3878-81.
23. Lewandowski, J.R., et al., *Enhanced Resolution and Coherence Lifetimes in the Solid-State NMR Spectroscopy of Perdeuterated Proteins under Ultrafast Magic-Angle Spinning.* The Journal of Physical Chemistry Letters, 2011. **2**(17): p. 2205-2211.
24. Knight, M.J., et al., *Structure and backbone dynamics of a microcrystalline metalloprotein by solid-state NMR.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(28): p. 11095-11100.
25. Linser, R., et al., *Sensitivity enhancement using paramagnetic relaxation in MAS solid-state NMR of perdeuterated proteins.* J Magn Reson, 2007. **189**(2): p. 209-16.
26. Klein, A., et al., *Atomic-resolution chemical characterization of (2x)72-kDa tryptophan synthase via four- and five-dimensional (1)H-detected solid-state NMR.* Proc Natl Acad Sci U S A, 2022. **119**(4).
27. Klein, A., et al., *Unambiguous Side-Chain Assignments for Solid-State NMR Structure Elucidation of Nondeuterated Proteins via a Combined 5D/4D Side-Chain-to-Backbone Experiment.* The Journal of Physical Chemistry Letters, 2022. **13**(7): p. 1644-1651.
28. Linser, R., U. Fink, and B. Reif, *Proton-detected scalar coupling based assignment strategies in MAS solid-state NMR spectroscopy applied to perdeuterated proteins.* J Magn Reson, 2008. **193**(1): p. 89-93.
29. Palmer, A.G., III, *NMR Characterization of the Dynamics of Biomacromolecules.* Chemical Reviews, 2004. **104**(8): p. 3623-3640.
30. Fischer, M.W.F., A.K. Majumdar, and E.R.P. Zuiderweg, *PROTEIN NMR RELAXATION : THEORY, APPLICATIONS AND OUTLOOK.* Progress in Nuclear Magnetic Resonance Spectroscopy, 1998. **33**: p. 207-272.
31. Jarymowycz, V.A. and M.J. Stone, *Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences.* Chemical Reviews, 2006. **106**(5): p. 1624-1671.
32. Redfield, A.G., *On the Theory of Relaxation Processes.* IBM Journal of Research and Development, 1957. **1**(1): p. 19-31.
33. Redfield, A., *Intramolecular dipolar relaxation in multi-spin systems.* Adv. Magn. Reson, 1966. **1**: p. 1-32.
34. Wangsness, R.K. and F. Bloch, *The dynamical theory of nuclear induction.* Physical Review, 1953. **89**(4): p. 728-739.
35. Abergel, D. and A.G. Palmer, *Approximate Solutions of the Bloch–McConnell Equations for Two-Site Chemical Exchange.* ChemPhysChem, 2004. **5**(6): p. 787-793.
36. Lewandowski, J.R., *Advances in Solid-State Relaxation Methodology for Probing Site-Specific Protein Dynamics.* Accounts of Chemical Research, 2013. **46**(9): p. 2018-2027.
37. Marion, D., et al., *Microsecond Protein Dynamics from Combined Bloch-McConnell and Near-Rotary-Resonance R1p Relaxation-Dispersion MAS NMR.* ChemPhysChem, 2019. **20**(2): p. 276-284.
38. Rovó, P., et al., *Mechanistic Insights into Microsecond Time-Scale Motion of Solid Proteins Using Complementary 15N and 1H Relaxation Dispersion Techniques.* Journal of the American Chemical Society, 2019. **141**(2): p. 858-869.
39. Rovó, P. and R. Linser, *Microsecond Timescale Protein Dynamics: a Combined Solid-State NMR Approach.* ChemPhysChem, 2018. **19**(1): p. 34-39.
40. Rovó, P., *Recent advances in solid-state relaxation dispersion techniques.* Solid State Nuclear Magnetic Resonance, 2020. **108**: p. 101665.

41. Quinn, C.M. and A.E. McDermott, *Monitoring conformational dynamics with solid-state R1ρ experiments.* Journal of Biomolecular NMR, 2009. **45**(1): p. 5-8.
42. Quinn, C.M. and A.E. McDermott, *Quantifying conformational dynamics using solid-state R1ρ experiments.* Journal of Magnetic Resonance, 2012. **222**: p. 1-7.
43. Ma, P., et al., *Observing the overall rocking motion of a protein in a crystal.* Nature Communications, 2015. **6**(1): p. 8361.
44. Kurauskas, V., et al., *Slow conformational exchange and overall rocking motion in ubiquitin protein crystals.* Nature Communications, 2017. **8**(1): p. 145.
45. Sun, P.D., C.E. Foster, and J.C. Boyington, *Overview of Protein Structural and Functional Folds.* Current Protocols in Protein Science, 2004. **35**(1): p. 17.1.1-17.1.189.
46. Richardson, J.S., *Describing patterns of protein tertiary structure*, in *Methods in Enzymology*. 1985, Academic Press. p. 341-358.
47. Ouellette, R.J. and J.D. Rawn, *14 - Amino Acids, Peptides, and Proteins*, in *Principles of Organic Chemistry*, R.J. Ouellette and J.D. Rawn, Editors. 2015, Elsevier: Boston. p. 371-396.
48. Skipper, L., *PROTEINS | Overview ☆*, in *Encyclopedia of Analytical Science (Third Edition)*, P. Worsfold, et al., Editors. 2005, Academic Press: Oxford. p. 412-419.
49. Wishart, D.S., et al., *1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* Journal of Biomolecular NMR, 1995. **5**(1): p. 67-81.
50. Nielsen, J.T. and F.A.A. Mulder, *CheSPI: chemical shift secondary structure population inference.* Journal of Biomolecular NMR, 2021. **75**(6): p. 273-291.
51. Shen, Y., et al., *TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts.* Journal of Biomolecular NMR, 2009. **44**(4): p. 213-223.
52. Kragelj, J., et al., *Conformational propensities of intrinsically disordered proteins from NMR chemical shifts.* Chemphyschem, 2013. **14**(13): p. 3034-45.
53. Nielsen, J.T. and F.A.A. Mulder, *Quantitative Protein Disorder Assessment Using NMR Chemical Shifts*, in *Intrinsically Disordered Proteins: Methods and Protocols*, B.B. Kragelund and K. Skriver, Editors. 2020, Springer US: New York, NY. p. 303-317.
54. Kabsch, W. and C. Sander, *Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features.* Biopolymers, 1983. **22**(12): p. 2577-2637.
55. Griesinger, C., O. Sørensen, and R.R. Ernst, *Practical aspects of the E. COSY technique. Measurement of scalar spin-spin coupling constants in peptides.* Journal of Magnetic Resonance (1969), 1987. **75**(3): p. 474-492.
56. Hu, J.-S. and A. Bax, *Determination of φ and χ1 Angles in Proteins from 13C−13C Three-Bond J Couplings Measured by Three-Dimensional Heteronuclear NMR. How Planar Is the Peptide Bond?* Journal of the American Chemical Society, 1997. **119**(27): p. 6360-6368.
57. Hu, J.-S., S. Grzesiek, and A. Bax, *Two-Dimensional NMR Methods for Determining χ1 Angles of Aromatic Residues in Proteins from Three-Bond JC'Cγ and JNCγ Couplings.* Journal of the American Chemical Society, 1997. **119**(7): p. 1803-1804.
58. Karplus, M., *Vicinal Proton Coupling in Nuclear Magnetic Resonance.* Journal of the American Chemical Society, 1963. **85**(18): p. 2870-2871.
59. Schmidt, J.M., et al., *Self-consistent 3J coupling analysis for the joint calibration of Karplus coefficients and evaluation of torsion angles.* Journal of Biomolecular NMR, 1999. **14**(1): p. 1-12.

60. Kleckner, I.R. and M.P. Foster, *An introduction to NMR-based approaches for measuring protein dynamics.* Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2011. **1814**(8): p. 942-968.
61. Neudecker, P., P. Lundström, and L.E. Kay, *Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding.* Biophys J, 2009. **96**(6): p. 2045-54.
62. Korzhnev, D.M. and L.E. Kay, *Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding.* Acc Chem Res, 2008. **41**(3): p. 442-51.
63. Baldwin, A.J. and L.E. Kay, *NMR spectroscopy brings invisible protein states into focus.* Nat Chem Biol, 2009. **5**(11): p. 808-14.
64. Yang, D. and L.E. Kay, *Contributions to Conformational Entropy Arising from Bond Vector Fluctuations Measured from NMR-Derived Order Parameters: Application to Protein Folding.* Journal of Molecular Biology, 1996. **263**(2): p. 369-382.
65. Nosaki, S. and K. Miura, *Chapter Nine - Transient expression of recombinant proteins in plants*, in *Methods in Enzymology*, W.B. O'Dell and Z. Kelman, Editors. 2021, Academic Press. p. 193-203.
66. Rosano, G.L. and E.A. Ceccarelli, *Recombinant protein expression in Escherichia coli: advances and challenges.* Frontiers in Microbiology, 2014. **5**.
67. Overton, T.W., *Recombinant protein production in bacterial hosts.* Drug Discovery Today, 2014. **19**(5): p. 590-601.
68. Andersen, D.C. and L. Krummen, *Recombinant protein expression for therapeutic applications.* Current Opinion in Biotechnology, 2002. **13**(2): p. 117-123.
69. Baneyx, F., *Recombinant protein expression in Escherichia coli.* Current Opinion in Biotechnology, 1999. **10**(5): p. 411-421.
70. Azatian, S.B., N. Kaur, and M.P. Latham, *Increasing the buffering capacity of minimal media leads to higher protein yield.* J Biomol NMR, 2019. **73**(1-2): p. 11-17.
71. Labrou, N.E., *Protein Purification: An Overview*, in *Protein Downstream Processing: Design, Development and Application of High and Low-Resolution Methods*, N.E. Labrou, Editor. 2014, Humana Press: Totowa, NJ. p. 3-10.
72. Asenjo, J.A. and B.A. Andrews, *Protein purification using chromatography: selection of type, modelling and optimization of operating conditions.* Journal of Molecular Recognition, 2009. **22**(2): p. 65-76.
73. Walters, R.R., *Affinity chromatography.* Analytical Chemistry, 1985. **57**(11): p. 1099A-1114A.
74. Jungbauer, A. and R. Hahn, *Chapter 22 Ion-Exchange Chromatography*, in *Methods in Enzymology*, R.R. Burgess and M.P. Deutscher, Editors. 2009, Academic Press. p. 349-371.
75. Fekete, S., et al., *Theory and practice of size exclusion chromatography for the analysis of protein aggregates.* Journal of Pharmaceutical and Biomedical Analysis, 2014. **101**: p. 161-173.
76. Sattler, M. and S.W. Fesik, *Use of deuterium labeling in NMR: overcoming a sizeable problem.* Structure, 1996. **4**(11): p. 1245-9.
77. Asami, S., P. Schmieder, and B. Reif, *High resolution 1H-detected solid-state NMR spectroscopy of protein aliphatic resonances: access to tertiary structure information.* J Am Chem Soc, 2010. **132**(43): p. 15133-5.
78. Wang, S., et al., *Nano-mole scale side-chain signal assignment by 1H-detected protein solid-state NMR by ultra-fast magic-angle spinning and stereo-array isotope labeling.* PLoS One, 2015. **10**(4): p. e0122714.

79. Medeiros-Silva, J., et al., *(1) H-Detected Solid-State NMR Studies of Water-Inaccessible Proteins In Vitro and In Situ.* Angew Chem Int Ed Engl, 2016. **55**(43): p. 13606-13610.
80. Xue, K., et al., *Limits of Resolution and Sensitivity of Proton Detected MAS Solid-State NMR Experiments at 111 kHz in Deuterated and Protonated Proteins.* Sci Rep, 2017. **7**(1): p. 7444.
81. Schütz, S. and R. Sprangers, *Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems.* Prog Nucl Magn Reson Spectrosc, 2020. **116**: p. 56-84.
82. Boisbouvier, J. and L.E. Kay, *Advanced isotopic labeling for the NMR investigation of challenging proteins and nucleic acids.* Journal of Biomolecular NMR, 2018. **71**(3): p. 115-117.
83. Martínez-Limón, A., et al., *The p38 Pathway: From Biology to Cancer Therapy.* International Journal of Molecular Sciences, 2020. **21**(6): p. 1913.
84. Kumar, G.S., et al., *Dynamic activation and regulation of the mitogen-activated protein kinase p38.* Proc Natl Acad Sci U S A, 2018. **115**(18): p. 4655-4660.
85. Cuenda, A. and S. Rousseau, *p38 MAP-Kinases pathway regulation, function and role in human diseases.* Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2007. **1773**(8): p. 1358-1375.
86. Francis, D.M., et al., *Structural basis of p38α regulation by hematopoietic tyrosine phosphatase.* Nat Chem Biol, 2011. **7**(12): p. 916-24.
87. Machado, T.R., T.R. Machado, and P.G. Pascutti, *The p38 MAPK Inhibitors and Their Role in Inflammatory Diseases.* Chemistryselect, 2021. **6**(23): p. 5729-5742.
88. Herlaar, E. and Z. Brown, *p38 MAPK signalling cascades in inflammatory disease.* Molecular Medicine Today, 1999. **5**(10): p. 439-447.
89. Cuadrado, A. and A.R. Nebreda, *Mechanisms and functions of p38 MAPK signalling.* Biochem J, 2010. **429**(3): p. 403-17.
90. Astolfi, A., et al., *A Comprehensive Structural Overview of p38α Mitogen-Activated Protein Kinase in Complex with ATP-Site and Non-ATP-Site Binders.* ChemMedChem, 2018. **13**(1): p. 7-14.
91. Comess, K.M., et al., *Discovery and Characterization of Non-ATP Site Inhibitors of the Mitogen Activated Protein (MAP) Kinases.* ACS Chemical Biology, 2011. **6**(3): p. 234-244.
92. Roser, P., et al., *Conformational selection vs. induced fit: insights into the binding mechanisms of p38alpha MAP Kinase inhibitors.* Chem Commun (Camb), 2020. **56**(62): p. 8818-8821.
93. Clark, A.R. and J.L. Dean, *The p38 MAPK Pathway in Rheumatoid Arthritis: A Sideways Look.* Open Rheumatol J, 2012. **6**: p. 209-19.
94. Zarubin, T. and J. Han, *Activation and signaling of the p38 MAP kinase pathway.* Cell Research, 2005. **15**(1): p. 11-18.
95. Madkour, M.M., H.S. Anbar, and M.I. El-Gamal, *Current status and future prospects of p38α/MAPK14 kinase and its inhibitors.* European Journal of Medicinal Chemistry, 2021. **213**: p. 113216.
96. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* cell, 2000. **100**(1): p. 57-70.
97. Wilson, C., et al., *Using ancient protein kinases to unravel a modern cancer drug’s mechanism.* Science, 2015. **347**(6224): p. 882-886.
98. Sun, S., et al., *A Time-Saving Strategy for MAS NMR Spectroscopy by Combining Nonuniform Sampling and Paramagnetic Relaxation Assisted Condensed Data Collection.* The Journal of Physical Chemistry B, 2012. **116**(46): p. 13585-13596.

99. Anderson, C.J., et al., *Cross-bridged macrocyclic chelators for stable complexation of copper radionuclides for PET imaging.* The quarterly journal of nuclear medicine and molecular imaging : official publication of the Italian Association of Nuclear Medicine (AIMN) [and] the International Association of Radiopharmacology (IAR), [and] Section of the Society of.. 2008. **52**(2): p. 185-192.
100. Ray, W.J. and C.E. Bracker, *Polyethylene glycol: Catalytic effect on the crystallization of phosphoglucomutase at high salt concentration.* Journal of Crystal Growth, 1986. **76**(3): p. 562-576.
101. Jang, K., et al., *A Short Review on Cryoprotectants for 3D Protein Structure Analysis.* Crystals, 2022. **12**(2): p. 138.
102. Francis, D.M., et al., *Structural basis of p38α regulation by hematopoietic tyrosine phosphatase.* Nature Chemical Biology, 2011. **7**(12): p. 916-924.
103. Mulder, F.A., et al., *Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme.* J Am Chem Soc, 2001. **123**(5): p. 967-75.
104. Chao, F.A. and R.A. Byrd, *Protein Dynamics revealed by NMR Relaxation Methods.* Emerg Top Life Sci, 2020. **2**(1): p. 93-105.
105. Millet, O., et al., *The Static Magnetic Field Dependence of Chemical Exchange Linebroadening Defines the NMR Chemical Shift Time Scale.* Journal of the American Chemical Society, 2000. **122**(12): p. 2867-2877.
106. Bieri, M. and P.R. Gooley, *Automated NMR relaxation dispersion data analysis using NESSY.* BMC Bioinformatics, 2011. **12**(1): p. 421.
107. Wang, C., M.J. Grey, and A.G. Palmer, 3rd, *CPMG sequences with enhanced sensitivity to chemical exchange.* J Biomol NMR, 2001. **21**(4): p. 361-6.
108. Carver, J.P. and R.E. Richards, *A general two-site solution for the chemical exchange produced dependence of T2 upon the carr-Purcell pulse separation.* Journal of Magnetic Resonance (1969), 1972. **6**(1): p. 89-105.
109. Asami, S. and B. Reif, *Assignment strategies for aliphatic protons in the solid-state in randomly protonated proteins.* J Biomol NMR, 2012. **52**(1): p. 31-9.
110. Cala-De Paepe, D., et al., *Is protein deuteration beneficial for proton detected solid-state NMR at and above 100 kHz magic-angle spinning?* Solid State Nuclear Magnetic Resonance, 2017. **87**: p. 126-136.
111. Skinner, S.P., et al., *CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis.* Journal of Biomolecular NMR, 2016. **66**(2): p. 111-124.
112. Karjalainen, M., et al., *HACANCOi: a new H(α)-detected experiment for backbone resonance assignment of intrinsically disordered proteins.* J Biomol NMR, 2020. **74**(12): p. 741-752.
113. Zhou, D.H. and C.M. Rienstra, *High-performance solvent suppression for proton detected solid-state NMR.* Journal of Magnetic Resonance, 2008. **192**(1): p. 167-172.
114. Güntert, P., *Automated structure determination from NMR spectra.* European Biophysics Journal, 2008. **38**(2): p. 129.
115. Schmidt, E. and P. Güntert, *A New Algorithm for Reliable and General NMR Resonance Assignment.* Journal of the American Chemical Society, 2012. **134**(30): p. 12817-12829.
116. López-Méndez, B. and P. Güntert, *Automated Protein Structure Determination from NMR Spectra.* Journal of the American Chemical Society, 2006. **128**(40): p. 13112-13122.
117. Han, B., et al., *SHIFTX2: significantly improved protein chemical shift prediction.* J Biomol NMR, 2011. **50**(1): p. 43-57.

118. Delaglio, F., et al., *Non-Uniform Sampling for All: More NMR Spectral Quality, Less Measurement Time.* Am Pharm Rev, 2017. **20**(4).
119. East, K.W., F. Delaglio, and G.P. Lisi, *A simple approach for reconstruction of non-uniformly sampled pseudo-3D NMR data for accurate measurement of spin relaxation parameters.* J Biomol NMR, 2021. **75**(6-7): p. 213-219.
120. Billeter, M., *Non-uniform sampling in biomolecular NMR.* Journal of Biomolecular NMR, 2017. **68**(2): p. 65-66.
121. Nichols, P.J., et al., *Reducing the measurement time of exact NOEs by non-uniform sampling.* J Biomol NMR, 2020. **74**(12): p. 717-739.
122. Ying, J., et al., *Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data.* J Biomol NMR, 2017. **68**(2): p. 101-118.
123. Stanek, J. and W. Koźmiński, *Iterative algorithm of discrete Fourier transform for processing randomly sampled NMR data sets.* Journal of Biomolecular NMR, 2010. **47**(1): p. 65-77.
124. Hoffman, G.R., N. Nassar, and R.A. Cerione, *Structure of the Rho family GTP-binding protein Cdc42 in complex with the multifunctional regulator RhoGDI.* Cell, 2000. **100**(3): p. 345-56.
125. Gosser, Y.Q., et al., *C-terminal binding domain of Rho GDP-dissociation inhibitor directs N-terminal inhibitory peptide to GTPases.* Nature, 1997. **387**(6635): p. 814-9.
126. Hodge, R.G. and A.J. Ridley, *Regulating Rho GTPases and their regulators.* Nature Reviews Molecular Cell Biology, 2016. **17**(8): p. 496-510.
127. Keep, N.H., et al., *A modulator of rho family G proteins, rhoGDI, binds these G proteins via an immunoglobulin-like domain and a flexible N-terminal arm.* Structure, 1997. **5**(5): p. 623-33.
128. Johnson, J.L., J.W. Erickson, and R.A. Cerione, *New insights into how the Rho guanine nucleotide dissociation inhibitor regulates the interaction of Cdc42 with membranes.* J Biol Chem, 2009. **284**(35): p. 23860-71.
129. Dass, R., F.A.A. Mulder, and J.T. Nielsen, *ODiNPred: comprehensive prediction of protein order and disorder.* Sci Rep, 2020. **10**(1): p. 14780.
130. Nielsen, J.T. and F.A.A. Mulder, *Quality and bias of protein disorder predictors.* Sci Rep, 2019. **9**(1): p. 5137.
131. Vuister, G.W. and A. Bax, *Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins.* Journal of the American Chemical Society, 1993. **115**(17): p. 7772-7777.
132. Gáspári, Z. and A. Perczel, *Chapter 2 - Protein Dynamics as Reported by NMR*, in *Annual Reports on NMR Spectroscopy*, G.A. Webb, Editor. 2010, Academic Press. p. 35-75.
133. Hu, Q., *NMR Approaches to Determine Protein Structure*, in *Molecular Life Sciences: An Encyclopedic Reference*, R.D. Wells, et al., Editors. 2018, Springer New York: New York, NY. p. 795-803.
134. Raman, S., et al., *Structure prediction for CASP8 with all-atom refinement using Rosetta.* Proteins, 2009. **77 Suppl 9**(0 9): p. 89-99.
135. Zhang, R., et al., *Spectrin: Structure, function and disease.* Science China Life Sciences, 2013. **56**(12): p. 1076-1085.
136. Kurochkina, N. and U. Guha, *SH3 domains: modules of protein-protein interactions.* Biophys Rev, 2013. **5**(1): p. 29-39.
137. Krushelnitsky, A., et al., *Microsecond motions probed by near-rotary-resonance R1ρ 15N MAS NMR experiments: the model case of protein overall-rocking in crystals.* Journal of Biomolecular NMR, 2018. **71**(1): p. 53-67.

138. Carter, P., *Site-directed mutagenesis.* Biochem J, 1986. **237**(1): p. 1-7.
139. van Rossum, B.-J., et al., *Assignment of amide proton signals by combined evaluation of HN, NN and HNCA MAS-NMR correlation spectra.* Journal of Biomolecular NMR, 2003. **25**(3): p. 217-223.
140. Janin, J. and F. Rodier, *Protein–protein interaction at crystal contacts.* Proteins: Structure, Function, and Bioinformatics, 1995. **23**(4): p. 580-587.
141. Pellicane, G., G. Smith, and L. Sarkisov, *Molecular Dynamics Characterization of Protein Crystal Contacts in Aqueous Solutions.* Physical Review Letters, 2008. **101**(24): p. 248102.
142. Fusco, D., et al., *Characterizing protein crystal contacts and their role in crystallization: rubredoxin as a case study.* Soft Matter, 2014. **10**(2): p. 290-302.
143. Carugo, O. and P. Argos, *Protein—protein crystal-packing contacts.* Protein Science, 1997. **6**(10): p. 2261-2263.
144. Lipari, G. and A. Szabo, *Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity.* Journal of the American Chemical Society, 1982. **104**(17): p. 4546-4559.

9 LIST OF TABLES

10 LIST OF FIGURES

11 APENDIX

11.1 Stock and media recipes

Table 11.1 Stock solution and media components for protein expression, purification and crystallization.

Growth media	Compound	Amount	Concentration
LB Media (1L)	Tryptone	10 g	
	Yeast extract	5 g	
	NaCl	10 g	
Minimal Media (1L)	M9 (10x)	100 ml	
	Trace elements (100X)	10 ml	
	$MgSO_4$	1 ml	1 mM
	$CaCl_2$	0.3 ml	0.3 mM
	Thiamin-HCl [1 mg/ml]	1.5 mg	
	Biotin	1.5 mg	
	Glucose	2 g	
	NH_4Cl	1 g	
	Antibiotic stock	1 ml	
M9 10X	Na_2HPO_4	60 g	0.479 M
	KH_2PO_4	30 g	0.220 M
	NaCl	5 g	0.086 M
Trace elements 100X	EDTA pH 7.5	5 g	
	$FeCl_3$	0.83 g	
	$ZnCl_2$	84 mg	
	$CuCl_2$ x $2H_2O$	13 mg	
	$CoCl_2$ x $6H_2O$	10 mg	
	H_3BO_3	10 mg	
	$MnCl_2$ x $4H_2O$	1.6 mg	
Antibiotic stock			
Ampicillin 1000X	Ampicillin	2 g	
	H_2O	20 ml	
Chloramphenicol 1000X	Chloramphenicol	0.68 g	
	Ethanol	20 mL	
Other stock			
$MgSO_4$	$MgSO_4$	12.03 g	1M
	H_2O	100 ml	
$CaCl_2$	$CaCl_2$	11.09 g	1 M
	H_2O	100 ml	
IPTG	IPTG	4.76 g	1 M
	H_2O	20 ml	
Crystallization			
$(NH_4)_2SO_4$	$(NH_4)_2SO_4$	13.21g	1 M
	H_2O	100 ml	
Cu(II) EDTA	$CuCl_2$	4.70 g	350 mM
	EDTA	102.2 g	
	H_2O	100 ml	
NaOH	NaOH	4 g	1M
	H_2O	100 ml	
MES	2-(N-morpholino) ethanesulfonic acid	0.97 g	0.5 M
	H_2O	10 ml	

PEG4000	PEG4000 H_2O	5 g 10 ml	50%
BOG	Octyl β-D-glucopyranoside H_2O	0.73 g 10 ml	250 mM
Cu (II) CB TE2A	$CuCl_2$ CB TE2A H_2O	0.020 g 0.053 g 0.5 ml	310 mM
p38α purification			
Buffer A	Tris NaCl Imidazol Glycerol DTT H_2O pH 8	3.02 g 14.61 g 0.85 g 25 ml 0.077 g Fill to 500 ml	 50 mM 500 mM 25 mM 5 % 1 mM
Buffer B	Tris NaCl Imidazole Glycerol H_2O pH 8	1.5 g 7.3 g 17.01 g 12.5 ml Fill to 250 ml	50 mM 500 mM 500 mM 5 %
Buffer QA	HEPES Glycerol DTT H_2O pH 7	2.97 g 25 ml 0.077 g Fill to 500 ml	25 mM 5 % 1 mM
Buffer QB	HEPES NaCl Glycerol H_2O pH 7	1.48 g 14.61 g 12.5 ml Fill to 250 ml	25 mM 1 M 5 %
Cleavage/Dialysis	HEPES NaCl Glycerol DTT H_2O pH 7	11.91 g 11.68 g 100 ml 0.308 g Fill to 2 L	25 mM 100 mM 5 % 1 mM
SEC buffer	HEPES NaCl L-Methionine Glycerol DTT H_2O pH 7.2	4.76 g 2.92 g 100 mg 50 ml 1.54 g Fill to 1 L	20 mM 50 mM 100 mg/L 5 % 10 mM
NMR buffer	HEPES NaCl DTT H_2O pH 6.8	11.91 g 8.76 g 0.038 g/50 ml Fill to 1 L	50 mM 150 mM 5 mM
SH3 purification			
Buffer A	Tris H_2O pH 8.5	2.52 g Fill to 1 L	20 mM
Buffer B	Tris NaCl H_2O pH 8.5	2.42 g 58.44 g Fill to 1 L	20 mM 1 M

SEC buffer	Citric acid NaCl H_2O pH 3.5	3.84 g 8.76 g Fill to 1 L	20 mM 150 mM

11.2 Purification chromatograms

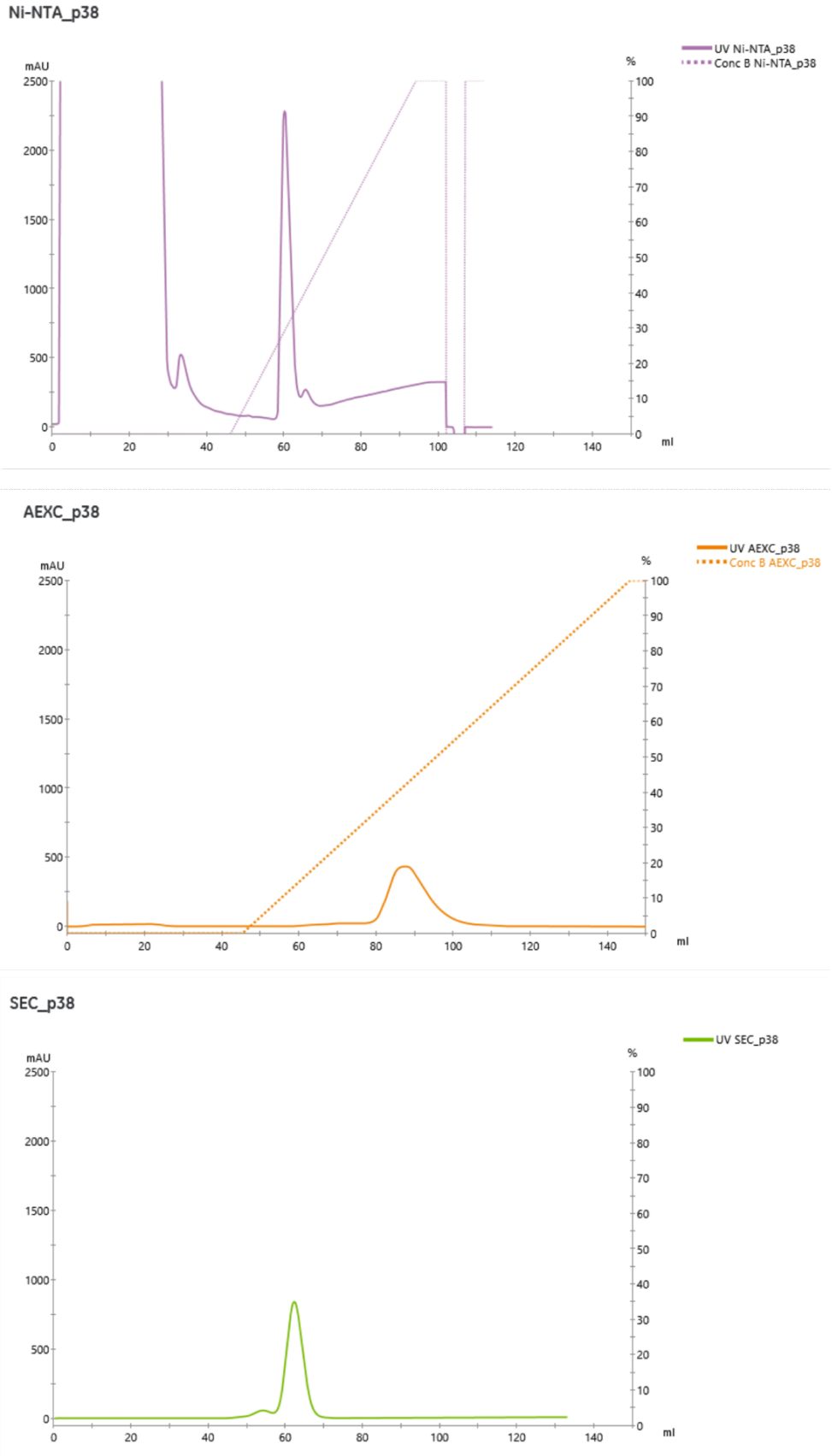


Figure 11.1. Chromatograms obtain in the purification of p38α. Ni-NTA purification (purple), Anion Exchange chromatography (orange) and Size Exclusion Chromatography (green).

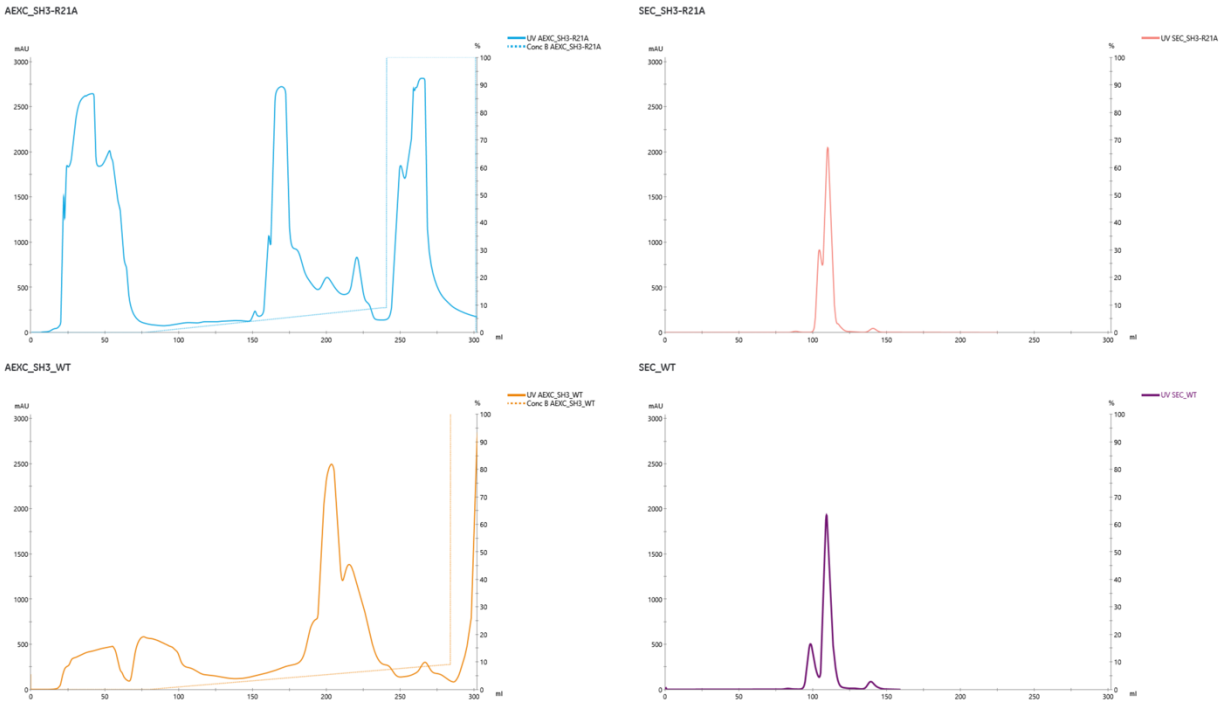


Figure 11.2. Chromatograms from purification of SH3. R21A mutant Anion exchange chromatography (blue), Size Exclusion chromatography (red). SH3 wild type Anion exchange chromatography (orange) and Size Exclusion chromatography (purple).

11.3 Chemical shift assignment

Table 11.2. Chemical shift assignment for p38α in solid state

#	Name	Atom	Chemical shift	#	Name	Atom	Chemical shift
8	PRO	C	174.800	12	ARG	CA	54.011
9	THR	C	172.510	12	ARG	CB	31.933
9	THR	CA	62.750	12	ARG	H	8.191
9	THR	CB	65.083	12	ARG	N	119.491
9	THR	H	8.174	13	GLN	C	172.254
9	THR	N	117.101	13	GLN	CA	54.728
10	PHE	C	174.715	13	GLN	CB	31.765
10	PHE	CA	56.664	13	GLN	H	8.852
10	PHE	CB	40.502	13	GLN	N	122.431
10	PHE	H	8.592	14	GLU	C	174.624
10	PHE	N	126.585	14	GLU	CA	55.372
11	TYR	C	171.697	14	GLU	H	8.615
11	TYR	CA	55.515	14	GLU	N	125.815
11	TYR	H	8.933	15	LEU	CA	53.580
11	TYR	N	119.627	15	LEU	CB	42.803
12	ARG	C	174.940	15	LEU	H	8.750

15	LEU	N	127.876
16	ASN	C	174.133
17	LYS	C	174.379
17	LYS	CA	57.648
17	LYS	CB	28.558
17	LYS	H	8.605
17	LYS	N	110.632
18	THR	C	171.558
18	THR	CA	61.092
18	THR	CB	71.418
18	THR	H	7.625
18	THR	N	114.889
19	ILE	C	174.339
19	ILE	CA	59.937
19	ILE	CB	35.528
19	ILE	H	8.319
19	ILE	N	124.849
20	TRP	C	175.698
20	TRP	CA	56.453
20	TRP	CB	27.728
20	TRP	H	9.091
20	TRP	N	130.833
21	GLU	C	172.361
21	GLU	CA	54.939
21	GLU	CB	29.797
21	GLU	H	7.866
21	GLU	N	127.795
22	VAL	C	171.526
22	VAL	CA	56.028
22	VAL	CB	44.241
22	VAL	H	7.530
22	VAL	N	112.335
23	PRO	C	176.161
24	GLU	CA	58.184
24	GLU	CB	28.361
24	GLU	H	7.945
24	GLU	N	119.212
25	ARG	CA	58.183
25	ARG	CB	28.342
25	ARG	H	7.380
25	ARG	N	116.902
27	GLN	C	173.624
27	GLN	CA	53.336
27	GLN	CB	32.071
27	GLN	H	8.888
27	GLN	N	122.345
28	ASN	C	173.251
28	ASN	CA	53.282
28	ASN	CB	35.502
28	ASN	H	8.693
28	ASN	N	117.342
29	LEU	C	178.020
29	LEU	CA	56.706
29	LEU	CB	40.166
29	LEU	H	8.226
29	LEU	N	119.661
30	SER	CA	54.613
30	SER	H	8.551
30	SER	N	116.121
36	ALA	C	177.549
37	TYR	CA	57.220
37	TYR	CB	39.989
37	TYR	H	7.726
37	TYR	N	115.717
38	GLY	CA	43.327
38	GLY	H	7.056
38	GLY	N	107.535
40	VAL	C	175.079
41	CYS	C	172.138
41	CYS	CA	57.464
41	CYS	CB	29.654
41	CYS	H	9.007
41	CYS	N	124.800
42	ALA	C	175.507
42	ALA	CA	50.152
42	ALA	CB	19.532
42	ALA	H	8.596
42	ALA	N	124.883
43	ALA	C	174.317
43	ALA	CA	50.491

43	ALA	CB	21.016
43	ALA	H	8.751
43	ALA	N	120.084
44	PHE	C	172.805
44	PHE	CA	56.685
44	PHE	CB	39.420
44	PHE	H	8.992
44	PHE	N	122.035
45	ASP	C	175.094
45	ASP	CA	51.921
45	ASP	CB	39.976
45	ASP	H	8.279
45	ASP	N	125.877
46	THR	CA	64.040
46	THR	CB	68.427
46	THR	H	8.441
46	THR	N	117.616
47	LYS	C	175.933
48	THR	C	175.608
48	THR	CA	60.735
48	THR	H	6.259
48	THR	N	103.667
49	GLY	C	172.287
49	GLY	CA	45.548
49	GLY	H	8.309
49	GLY	N	112.334
50	LEU	C	175.249
50	LEU	CA	53.690
50	LEU	CB	41.977
50	LEU	H	6.910
50	LEU	N	118.890
51	ARG	CA	55.559
51	ARG	CB	28.788
51	ARG	H	8.022
51	ARG	N	119.385
54	VAL	C	172.956
55	LYS	C	173.475
55	LYS	CA	47.244
55	LYS	H	9.536
55	LYS	N	128.747
56	LYS	C	175.797
56	LYS	CA	54.245
56	LYS	CB	32.669
56	LYS	H	8.314
56	LYS	N	128.564
57	LEU	CA	54.684
57	LEU	CB	38.267
57	LEU	H	8.280
57	LEU	N	128.455
60	PRO	C	174.473
61	PHE	C	175.751
61	PHE	CA	54.155
61	PHE	CB	36.975
61	PHE	H	7.734
61	PHE	N	111.990
62	GLN	CA	57.462
62	GLN	CB	29.001
62	GLN	H	7.068
62	GLN	N	116.072
64	ILE	C	175.443
64	ILE	CA	65.394
64	ILE	H	8.862
64	ILE	N	121.947
65	ILE	CA	64.527
65	ILE	CB	36.674
65	ILE	H	7.510
65	ILE	N	119.728
67	ALA	C	177.496
68	LYS	CA	59.474
68	LYS	CB	31.531
68	LYS	H	7.614
68	LYS	N	118.642
80	MET	C	175.781
81	LYS	C	172.352
81	LYS	CA	54.223
81	LYS	CB	31.254
81	LYS	H	8.310
81	LYS	N	128.496
82	HIS	CA	56.896
82	HIS	CB	32.319

82	HIS	H	8.403
82	HIS	N	123.210
83	GLU	C	176.977
83	GLU	CA	97.679
83	GLU	CB	28.838
83	GLU	H	7.810
83	GLU	N	124.947
84	ASN	C	172.774
84	ASN	CA	53.517
84	ASN	H	11.256
84	ASN	N	118.664
85	VAL	CA	61.304
85	VAL	CB	34.456
85	VAL	H	7.448
85	VAL	N	119.721
86	ILE	C	175.668
87	GLY	CA	43.509
87	GLY	H	7.790
87	GLY	N	112.052
92	PHE	C	171.760
93	THR	CA	57.123
93	THR	CB	70.982
93	THR	H	8.706
93	THR	N	113.321
94	PRO	C	175.463
95	ALA	C	177.456
95	ALA	CA	52.621
95	ALA	CB	19.255
95	ALA	H	7.525
95	ALA	N	121.849
96	ARG	C	174.103
96	ARG	CA	55.761
96	ARG	CB	29.061
96	ARG	H	9.117
96	ARG	N	121.027
97	SER	C	172.870
97	SER	CA	56.337
97	SER	CB	64.288
97	SER	H	7.228
97	SER	N	110.924
98	LEU	C	177.840
98	LEU	CA	56.730
98	LEU	CB	39.995
98	LEU	H	8.267
98	LEU	N	122.059
99	GLU	C	176.397
99	GLU	CA	58.878
99	GLU	CB	27.891
99	GLU	H	8.012
99	GLU	N	117.124
100	GLU	C	174.297
100	GLU	CA	54.654
100	GLU	CB	29.574
100	GLU	H	6.818
100	GLU	N	114.888
101	PHE	C	172.027
101	PHE	CA	52.748
101	PHE	H	7.133
101	PHE	N	121.859
102	ASN	C	171.427
103	ASP	C	173.525
103	ASP	CA	53.251
103	ASP	CB	42.822
103	ASP	H	7.302
103	ASP	N	114.939
104	VAL	CA	61.509
104	VAL	CB	34.064
104	VAL	H	8.252
104	VAL	N	120.642
105	TYR	C	55.151
105	TYR	CA	55.084
105	TYR	CB	38.887
105	TYR	H	8.120
105	TYR	N	124.263
108	THR	C	174.330
109	HIS	CA	59.460
109	HIS	H	8.783
109	HIS	N	119.076
112	GLY	CA	45.365
112	GLY	H	8.208

112	GLY	N	112.102
116	ASN	CA	55.933
116	ASN	CB	39.423
116	ASN	H	7.696
116	ASN	N	116.851
117	ASN	CA	57.203
117	ASN	CB	39.989
117	ASN	H	8.116
117	ASN	N	118.283
120	LYS	C	178.254
120	LYS	CA	57.300
120	LYS	CB	31.475
120	LYS	H	7.117
120	LYS	N	116.049
121	CYS	CA	58.158
121	CYS	CB	27.503
121	CYS	H	7.208
121	CYS	N	114.714
122	GLN	H	7.243
122	GLN	N	114.623
129	VAL	C	176.602
130	GLN	CA	59.192
130	GLN	CB	29.499
130	GLN	H	8.347
130	GLN	N	117.783
131	PHE	C	175.092
132	LEU	CA	58.258
132	LEU	CB	41.076
132	LEU	H	8.685
132	LEU	N	120.478
138	ARG	C	178.609
139	GLY	CA	46.881
139	GLY	H	7.910
139	GLY	N	106.865
140	LEU	C	176.504
140	LEU	CA	56.651
140	LEU	CB	39.442
140	LEU	H	8.554
140	LEU	N	121.874
141	LYS	C	178.304
141	LYS	CA	59.570
141	LYS	CB	30.619
141	LYS	H	8.532
141	LYS	N	119.164
142	TYR	C	177.212
142	TYR	CA	61.790
142	TYR	CB	36.608
142	TYR	H	6.635
142	TYR	N	117.901
143	ILE	C	177.487
143	ILE	CA	66.956
143	ILE	H	8.309
143	ILE	N	120.757
144	HIS	C	180.103
144	HIS	CA	57.164
144	HIS	H	9.375
144	HIS	N	119.168
145	SER	C	173.540
145	SER	CA	61.059
145	SER	CB	60.816
145	SER	H	7.780
145	SER	N	116.106
146	ALA	C	174.660
146	ALA	CA	51.093
146	ALA	CB	16.688
146	ALA	H	7.175
146	ALA	N	125.900
147	ASP	C	173.271
147	ASP	CA	55.424
147	ASP	CB	37.803
147	ASP	H	8.086
147	ASP	N	114.763
148	ILE	C	172.884
148	ILE	CA	60.204
148	ILE	CB	38.093
148	ILE	H	6.991
148	ILE	N	119.666
149	ILE	CA	60.048
149	ILE	CB	38.855
149	ILE	H	7.594

149	ILE	N	123.627
157	ASN	C	170.407
157	ASN	CA	49.211
157	ASN	H	7.369
157	ASN	N	119.319
158	LEU	CA	51.919
158	LEU	H	6.667
158	LEU	N	117.389
159	ALA	CA	50.255
159	ALA	CB	21.240
159	ALA	H	8.233
159	ALA	N	124.631
162	GLU	CA	58.789
162	GLU	CB	27.481
162	GLU	H	8.480
162	GLU	N	115.800
165	GLU	C	173.188
166	LEU	C	175.267
166	LEU	CA	53.053
166	LEU	CB	39.243
166	LEU	H	8.102
166	LEU	N	127.120
167	LYS	CA	54.160
167	LYS	CB	35.805
167	LYS	H	9.168
167	LYS	N	123.448
168	ILE	CA	61.993
168	ILE	CB	38.634
168	ILE	H	8.025
168	ILE	N	121.931
182	THR	C	173.994
182	THR	CA	62.225
182	THR	CB	69.459
182	THR	H	7.351
182	THR	N	114.469
183	GLY	CA	45.620
183	GLY	H	8.043
183	GLY	N	110.735
191	ARG	CA	55.466
191	ARG	CB	32.097
191	ARG	H	7.008
191	ARG	N	122.232
192	ALA	CA	49.952
192	ALA	CB	19.537
192	ALA	H	8.638
192	ALA	N	125.237
197	LEU	CA	57.736
197	LEU	CB	40.563
197	LEU	H	8.892
197	LEU	N	126.371
198	ASN	CA	53.546
198	ASN	CB	39.769
198	ASN	H	7.528
198	ASN	N	114.664
199	TRP	CA	47.420
199	TRP	H	8.397
199	TRP	N	125.476
200	MET	CA	61.857
200	MET	H	8.509
200	MET	N	119.903
201	HIS	C	175.496
202	TYR	CA	52.295
202	TYR	H	10.069
202	TYR	N	126.824
203	ASN	C	175.348
203	ASN	CA	55.551
203	ASN	CB	41.442
203	ASN	H	8.811
203	ASN	N	125.035
207	ASP	C	176.732
208	ILE	CA	60.611
208	ILE	H	6.811
208	ILE	N	119.318
209	TRP	C	177.314
210	SER	C	175.585
210	SER	CA	62.038
210	SER	CB	62.549
210	SER	H	7.472
210	SER	N	112.153
211	VAL	C	176.029

211	VAL	CA	67.478
211	VAL	CB	29.001
211	VAL	H	7.888
211	VAL	N	119.980
212	GLY	C	173.659
212	GLY	CA	47.799
212	GLY	H	8.397
212	GLY	N	110.319
213	CYS	CA	63.945
213	CYS	CB	26.130
213	CYS	H	7.618
213	CYS	N	119.078
214	ILE	C	175.209
214	ILE	CA	65.637
214	ILE	H	8.235
214	ILE	N	123.627
215	MET	C	176.029
215	MET	CA	60.280
215	MET	CB	30.590
215	MET	H	9.156
215	MET	N	120.427
216	ALA	CA	55.470
216	ALA	CB	19.060
216	ALA	H	7.967
216	ALA	N	118.239
217	GLU	CA	57.867
217	GLU	CB	28.606
217	GLU	H	6.992
217	GLU	N	119.627
218	LEU	C	178.489
218	LEU	CA	57.246
218	LEU	CB	41.310
218	LEU	H	7.146
218	LEU	N	116.137
219	LEU	C	178.366
219	LEU	CA	57.072
219	LEU	H	7.888
219	LEU	N	117.222
220	THR	CA	62.084
220	THR	CB	70.528
220	THR	H	7.824
220	THR	N	105.614
221	GLY	C	172.607
222	ARG	C	173.773
222	ARG	CA	49.470
222	ARG	CB	31.785
222	ARG	H	7.901
222	ARG	N	118.815
223	THR	CA	63.598
223	THR	CB	67.935
223	THR	H	7.836
223	THR	N	119.819
224	LEU	CA	52.289
224	LEU	CB	40.336
224	LEU	H	7.676
224	LEU	N	118.559
226	PRO	C	175.010
227	GLY	C	176.752
227	GLY	CA	45.556
227	GLY	H	8.327
227	GLY	N	114.855
228	THR	C	173.470
228	THR	CA	63.940
228	THR	H	10.189
228	THR	N	116.841
229	ASP	CA	57.809
229	ASP	H	8.387
229	ASP	N	130.815
238	LEU	C	179.305
239	ARG	CA	59.408
239	ARG	CB	30.672
239	ARG	H	8.073
239	ARG	N	116.402
240	LEU	CA	97.063
240	LEU	CB	39.963
240	LEU	H	7.430
240	LEU	N	118.148
244	PRO	C	175.232
245	GLY	C	172.805
245	GLY	CA	43.272

245	GLY	H	8.173
245	GLY	N	109.206
246	ALA	CA	55.033
246	ALA	CB	18.075
246	ALA	H	8.297
246	ALA	N	121.956
247	GLU	C	177.434
248	LEU	CA	56.732
248	LEU	CB	40.268
248	LEU	H	7.530
248	LEU	N	119.808
249	LEU	CA	57.626
249	LEU	H	7.791
249	LEU	N	118.360
255	GLU	C	176.919
256	SER	CA	60.806
256	SER	CB	60.710
256	SER	H	8.044
256	SER	N	113.805
257	ALA	CA	54.661
257	ALA	CB	17.111
257	ALA	H	7.474
257	ALA	N	124.443
263	SER	CA	59.122
263	SER	H	7.219
263	SER	N	113.359
269	LYS	CA	56.531
271	ASN	C	176.436
272	PHE	C	176.928
272	PHE	CA	59.738
272	PHE	H	9.280
272	PHE	N	129.575
273	ALA	C	177.410
273	ALA	CA	54.657
273	ALA	CB	17.905
273	ALA	H	8.516
273	ALA	N	121.851
274	ASN	CA	53.636
274	ASN	CB	37.384
274	ASN	H	7.309
274	ASN	N	112.585
275	VAL	C	175.129
276	PHE	C	174.459
276	PHE	CA	55.798
276	PHE	CB	35.474
276	PHE	H	7.440
276	PHE	N	120.256
277	ILE	CA	61.687
277	ILE	CB	36.724
277	ILE	H	6.668
277	ILE	N	119.544
278	GLY	C	172.958
278	GLY	CA	44.980
278	GLY	H	8.778
278	GLY	N	115.137
279	ALA	C	175.904
279	ALA	CA	50.911
279	ALA	CB	18.472
279	ALA	H	7.287
279	ALA	N	121.599
280	ASN	C	174.205
280	ASN	CA	51.016
280	ASN	CB	37.813
280	ASN	H	8.890
280	ASN	N	122.049
281	PRO	C	177.373
282	LEU	C	177.578
282	LEU	CA	57.042
282	LEU	CB	40.963
282	LEU	H	8.255
282	LEU	N	117.181
283	ALA	C	177.891
283	ALA	CA	53.787
283	ALA	CB	16.244
283	ALA	H	6.894
283	ALA	N	120.947
284	VAL	C	176.063
284	VAL	CA	66.512
284	VAL	CB	29.595
284	VAL	H	6.902

284	VAL	N	116.030
285	ASP	C	176.435
285	ASP	CA	57.765
285	ASP	CB	41.635
285	ASP	H	7.433
285	ASP	N	117.910
286	LEU	C	178.254
286	LEU	CA	57.490
286	LEU	CB	38.722
286	LEU	H	6.956
286	LEU	N	117.245
287	LEU	CA	57.878
287	LEU	CB	40.449
287	LEU	H	8.031
287	LEU	N	120.504
290	MET	CA	59.310
290	MET	CB	30.642
290	MET	H	7.484
290	MET	N	118.098
296	ASP	CA	55.953
296	ASP	CB	39.924
296	ASP	H	8.302
296	ASP	N	120.880
297	LYS	CA	95.188
297	LYS	CB	32.139
297	LYS	H	7.430
297	LYS	N	116.796
298	ARG	CA	57.146
298	ARG	CB	29.357
298	ARG	H	6.837
298	ARG	N	122.453
299	ILE	C	170.904
299	ILE	CA	62.336
299	ILE	H	6.889
299	ILE	N	124.808
300	THR	C	174.998
300	THR	CA	60.046
300	THR	CB	69.783
300	THR	H	7.169
300	THR	N	108.924
301	ALA	C	177.879
301	ALA	CA	56.049
301	ALA	H	9.431
301	ALA	N	122.244
302	ALA	CA	55.064
302	ALA	CB	35.840
302	ALA	H	8.511
302	ALA	N	115.943
303	GLN	CA	98.463
303	GLN	CB	27.625
303	GLN	H	7.225
303	GLN	N	115.533
304	ALA	C	177.705
304	ALA	CA	54.767
304	ALA	CB	17.948
304	ALA	H	8.536
304	ALA	N	123.292
305	LEU	C	174.721
305	LEU	CA	57.702
305	LEU	CB	40.862
305	LEU	H	7.477
305	LEU	N	116.445
306	ALA	C	175.709
306	ALA	CA	50.811
306	ALA	CB	39.697
306	ALA	H	6.430
306	ALA	N	114.858
307	HIS	C	176.300
307	HIS	CA	58.542
307	HIS	CB	31.438
307	HIS	H	7.820
307	HIS	N	122.234
308	ALA	C	178.208
308	ALA	CA	55.081
308	ALA	CB	18.077
308	ALA	H	8.126
308	ALA	N	132.449
309	TYR	C	175.209
309	TYR	CA	60.156
309	TYR	H	11.446

309	TYR	N	124.473
310	PHE	C	176.059
310	PHE	CA	56.112
310	PHE	CB	37.243
310	PHE	H	7.554
310	PHE	N	110.018
311	ALA	C	178.058
311	ALA	CA	55.752
311	ALA	CB	17.430
311	ALA	H	7.407
311	ALA	N	123.651
312	GLN	C	174.217
312	GLN	CA	57.012
312	GLN	CB	26.854
312	GLN	H	8.503
312	GLN	N	113.100
313	TYR	C	177.026
313	TYR	CA	82.238
313	TYR	CB	38.803
313	TYR	H	7.366
313	TYR	N	115.808
314	HIS	C	173.716
314	HIS	CA	55.295
314	HIS	H	7.677
314	HIS	N	116.800
315	ASP	CA	50.693
315	ASP	CB	40.190
315	ASP	H	7.358
315	ASP	N	124.345
317	ASP	C	174.884
318	ASP	C	174.935
318	ASP	CA	52.363
318	ASP	CB	40.965
318	ASP	H	7.930
318	ASP	N	120.541
319	GLU	CA	53.369
319	GLU	CB	28.570
319	GLU	H	7.447
319	GLU	N	120.830
320	PRO	C	176.316
321	VAL	C	174.099
321	VAL	CA	59.294
321	VAL	CB	32.421
321	VAL	H	7.599
321	VAL	N	108.504
322	ALA	C	177.166
322	ALA	CA	50.477
322	ALA	CB	19.038
322	ALA	H	8.294
322	ALA	N	122.781
323	ASP	CA	53.113
323	ASP	CB	37.977
323	ASP	H	7.798
323	ASP	N	120.526
326	ASP	CA	54.224
326	ASP	H	6.698
326	ASP	N	121.672
330	GLU	C	176.698
330	GLU	CA	59.193
330	GLU	H	9.974
330	GLU	N	112.385
331	SER	C	173.416
331	SER	CA	57.725
331	SER	CB	62.706
331	SER	H	7.323
331	SER	N	112.321
332	ARG	C	174.222
333	ASP	CA	52.628
333	ASP	CB	40.899
333	ASP	H	8.151
333	ASP	N	124.769
336	ILE	C	176.805
337	ASP	CA	56.999
337	ASP	CB	39.003
337	ASP	H	8.309
337	ASP	N	115.087
338	GLU	C	171.646
338	GLU	CA	57.997
338	GLU	H	6.965
338	GLU	N	119.778

339	TRP	C	179.736
339	TRP	CA	59.602
339	TRP	CB	32.115
339	TRP	H	7.677
339	TRP	N	119.391
340	LYS	C	176.284
340	LYS	CA	60.244
340	LYS	CB	31.066
340	LYS	H	8.767
340	LYS	N	121.084
341	SER	C	175.657
341	SER	CA	61.193
341	SER	CB	60.981
341	SER	H	7.362
341	SER	N	114.478
342	LEU	CA	57.423
342	LEU	H	7.951
342	LEU	N	120.071
343	THR	C	174.740
344	TYR	C	177.619
344	TYR	CA	61.987
344	TYR	CB	36.304
344	TYR	H	8.626
344	TYR	N	124.634
345	ASP	C	178.840
345	ASP	CA	57.418
345	ASP	CB	39.989
345	ASP	H	8.043
345	ASP	N	116.632
346	GLU	C	118.559
346	GLU	CA	58.459
346	GLU	CB	28.485
346	GLU	H	7.433
346	GLU	N	117.805
347	VAL	CA	66.160
347	VAL	H	7.966
347	VAL	N	121.110
348	ILE	CA	101.259
348	ILE	CB	35.980
348	ILE	H	7.837
348	ILE	N	114.842
349	SER	C	173.739
350	PHE	C	172.945
350	PHE	CA	59.435
350	PHE	CB	38.291
350	PHE	H	6.922
350	PHE	N	123.722
351	VAL	CA	58.302
351	VAL	H	7.643
351	VAL	N	131.26

Table 11.3. Chemical shift assignment for RhoGDI1

#	Name	Atom	Chemical shift	#	Name	Atom	Chemical shift
-1	LEU	C	177.9066333	1	MET	CB	32.48036539
-1	LEU	CA	55.39119353	1	MET	H	8.071567847
-1	LEU	CB	41.55948809	1	MET	N	119.8472926
-1	LEU	H	8.417785934	2	ALA	C	177.6844693
-1	LEU	N	121.7817073	2	ALA	CA	52.59754726
0	GLY	C	174.1512187	2	ALA	CB	18.76937081
0	GLY	CA	45.3241349	2	ALA	H	8.29049536
0	GLY	H	8.331031885	2	ALA	N	124.7790815
0	GLY	N	109.3728351	3	GLU	C	176.3478977
1	MET	C	176.0217521	3	GLU	CA	56.53207708
1	MET	CA	55.48725506	3	GLU	CB	29.77453481

3	GLU	H	8.273956452
3	GLU	N	119.5990783
4	GLN	C	175.5388502
4	GLN	CA	55.52772387
4	GLN	CB	29.56736209
4	GLN	H	8.217083569
4	GLN	N	120.4653927
5	GLU	C	174.4446998
5	GLU	CA	54.60983934
5	GLU	CB	29.24825976
5	GLU	H	8.384257337
5	GLU	N	123.6615226
6	PRO	C	177.1131864
6	PRO	CA	63.03930187
6	PRO	CB	31.53884701
6	PRO	H	8.204550781
7	THR	C	174.8588069
7	THR	CA	61.32514738
7	THR	CB	70.52553828
7	THR	H	8.197601657
7	THR	N	113.1251647
8	ALA	C	179.4118811
8	ALA	CA	54.17931455
8	ALA	CB	18.14964821
8	ALA	H	8.527873848
8	ALA	N	124.2068547
9	GLU	C	177.8865312
9	GLU	CA	58.39333764
9	GLU	CB	29.25962378
9	GLU	H	8.435586993
9	GLU	N	118.6871871
10	GLN	C	177.7193799
10	GLN	CA	57.40923584
10	GLN	CB	28.65261931
10	GLN	H	7.956156007
10	GLN	N	120.577033
11	LEU	C	178.2553261
11	LEU	CA	56.38652507
11	LEU	CB	41.39440431
11	LEU	H	8.195538838
11	LEU	HA	8.184776318
11	LEU	N	121.2716177
12	ALA	C	178.5578628
12	ALA	CA	53.38132427
12	ALA	CB	18.20450546
12	ALA	H	7.944644391
12	ALA	N	122.5573658
13	GLN	C	176.729348
13	GLN	CA	56.61775293
13	GLN	CB	28.62655718
13	GLN	H	7.886317808
13	GLN	N	118.760924
14	ILE	C	176.6956688
14	ILE	CA	62.12938134
14	ILE	CB	38.1169053
14	ILE	H	7.902139552
14	ILE	N	121.0778084
15	ALA	C	177.904327
15	ALA	CA	53.0657048
15	ALA	CB	18.63127634
15	ALA	H	8.074873273
15	ALA	N	125.8997581
16	ALA	C	178.1842361
16	ALA	CA	52.92504481
16	ALA	CB	18.73223501
16	ALA	H	7.992532497
16	ALA	N	122.369427
17	GLU	C	176.4979993
17	GLU	CA	57.00491967
17	GLU	CB	29.63915974
17	GLU	H	8.156107814
17	GLU	N	119.1846613
18	ASN	C	175.346095
18	ASN	CA	53.42417102
18	ASN	CB	39.08688148
18	ASN	H	8.218033664
18	ASN	N	118.5272318
19	GLU	C	173.7110445
19	GLU	CA	55.72083697
19	GLU	CB	32.58061134

20	GLU	C	176.5396887
20	GLU	CA	54.49788043
20	GLU	CB	31.65761063
20	GLU	H	8.318290838
20	GLU	N	118.6120617
21	ASP	C	176.5838798
21	ASP	CA	55.40360328
21	ASP	CB	41.26641868
21	ASP	H	8.157462987
21	ASP	N	120.9431402
22	GLU	C	176.5168534
22	GLU	CA	57.32545045
22	GLU	CB	29.28275333
22	GLU	H	8.36140319
22	GLU	N	121.4203371
23	HIS	C	176.4478455
23	HIS	CA	56.30595384
23	HIS	CB	32.74331205
23	HIS	H	8.284813029
23	HIS	N	122.7932023
24	SER	C	176.3287643
24	SER	CA	58.12774403
24	SER	CB	63.86134218
24	SER	H	8.350078385
24	SER	N	117.1137757
25	VAL	C	175.9907192
25	VAL	CA	63.0466464
25	VAL	CB	31.77403291
25	VAL	H	8.16377945
25	VAL	N	121.3145636
26	ASN	C	174.4785852
26	ASN	CA	53.11985864
26	ASN	CB	38.77440528
26	ASN	H	8.24723571
26	ASN	N	119.8741419
27	TYR	C	175.0075548
27	TYR	CA	58.53926823
27	TYR	CB	38.44515002
27	TYR	H	7.80564574
27	TYR	N	120.9531746
28	LYS	C	172.7485657
28	LYS	CA	53.15876182
28	LYS	CB	32.63426466
28	LYS	H	7.785802139
28	LYS	N	126.7556547
30	PRO	C	176.7238871
30	PRO	CA	63.05028679
30	PRO	CB	31.53339082
31	ALA	C	177.7834317
31	ALA	CA	52.42949493
31	ALA	CB	18.84256681
31	ALA	H	8.273030693
31	ALA	N	123.6411714
32	GLN	C	176.4806511
32	GLN	CA	56.79226369
32	GLN	CB	29.75840038
32	GLN	H	8.453850039
32	GLN	N	120.3232846
33	LYS	C	175.5291095
33	LYS	CA	56.84307736
33	LYS	CB	29.65054422
33	LYS	H	8.191841235
33	LYS	N	118.6282978
34	SER	C	175.1422336
34	SER	CA	58.49767856
34	SER	CB	63.77844171
34	SER	H	7.982978102
34	SER	N	116.028457
35	ILE	C	176.411293
35	ILE	CA	61.71792506
35	ILE	CB	38.13711177
35	ILE	H	8.154417945
35	ILE	N	121.9859355
36	GLN	C	175.94208
36	GLN	CA	56.09148395
36	GLN	CB	28.9887265
36	GLN	H	8.322117389
36	GLN	N	123.5683243
37	GLU	C	175.4883552
37	GLU	CA	56.85233682

37	GLU	CB	29.9510512
37	GLU	H	8.21561583
37	GLU	N	121.8373219
38	ILE	C	175.8774956
38	ILE	CA	61.66831036
38	ILE	CB	38.08552156
38	ILE	H	8.076136598
38	ILE	N	121.6835408
39	GLN	C	175.975341
39	GLN	CA	56.35840177
39	GLN	CB	28.94367997
39	GLN	H	8.244669551
39	GLN	N	122.8812898
39	GLN	ND1	122.8140519
40	GLU	C	176.5029966
40	GLU	CA	56.58499077
40	GLU	CB	29.87314901
40	GLU	H	8.268520663
40	GLU	N	121.7080498
41	LEU	C	177.1397737
41	LEU	CA	55.38773697
41	LEU	CB	41.90390514
41	LEU	H	8.083166631
41	LEU	N	122.4223054
42	ASP	C	176.3284605
42	ASP	CA	54.66079839
42	ASP	CB	41.15289968
42	ASP	H	8.275098114
42	ASP	N	120.6012824
43	LYS	C	104.4048213
43	LYS	CA	56.53446794
43	LYS	CB	32.5094668
43	LYS	H	8.054673056
43	LYS	N	120.9970029
44	ASP	C	176.9735676
44	ASP	CA	54.26120269
44	ASP	CB	40.97559343
44	ASP	H	8.256256666
44	ASP	N	120.5233821
45	ASP	C	180.1958467
45	ASP	CA	56.03424545
45	ASP	CB	40.97487547
45	ASP	H	8.44474767
45	ASP	N	126.716167
46	GLU	C	178.0143082
46	GLU	CA	58.71141024
46	GLU	CB	29.11676402
46	GLU	H	8.456770535
46	GLU	N	121.8251551
47	SER	C	177.9315665
47	SER	CA	61.21088391
47	SER	CB	62.60939842
47	SER	H	8.232048853
47	SER	N	115.3718237
48	LEU	C	178.4216595
48	LEU	CA	56.73703552
48	LEU	CB	41.14233544
48	LEU	H	7.852834292
48	LEU	N	122.5901992
49	ARG	C	177.9754126
49	ARG	CA	58.7648651
49	ARG	CB	29.61602357
49	ARG	H	7.899887995
49	ARG	N	119.7507749
53	GLU	C	178.0008593
53	GLU	CA	58.26014064
53	GLU	CB	29.47270901
53	GLU	H	8.034537979
53	GLU	N	119.0367411
54	ALA	C	178.9491124
54	ALA	CA	53.65322026
54	ALA	CB	18.07231773
54	ALA	H	7.85678404
54	ALA	N	122.6031093
55	LEU	C	178.6224056
55	LEU	CA	56.19807753
55	LEU	CB	41.47248469
55	LEU	H	7.734273181
55	LEU	N	119.4361656
56	LEU	C	178.2111798

56	LEU	CA	56.04395402
56	LEU	CB	41.46344195
56	LEU	H	7.767901331
56	LEU	N	119.8702346
57	GLY	C	174.1278541
57	GLY	CA	45.46860845
57	GLY	H	8.017920422
57	GLY	N	107.6541273
58	ARG	C	176.2910508
58	ARG	CA	56.2524023
58	ARG	CB	30.25820302
58	ARG	H	7.859070595
58	ARG	N	119.8744572
59	VAL	C	175.7769846
59	VAL	CA	62.18254199
59	VAL	CB	32.29611533
59	VAL	H	7.980786681
59	VAL	N	120.9011086
60	ALA	C	177.8167547
60	ALA	CA	52.29876095
60	ALA	CB	18.90949672
60	ALA	H	8.253285654
60	ALA	N	126.1684257
61	VAL	C	176.3892442
61	VAL	CA	62.22045043
61	VAL	CB	32.25162786
61	VAL	H	8.021490573
61	VAL	N	118.9896966
62	SER	C	173.9903725
62	SER	CA	58.18873421
62	SER	CB	63.59016825
62	SER	H	8.193083307
62	SER	N	118.4790583
63	ALA	C	176.7779938
63	ALA	CA	52.06688278
63	ALA	CB	18.99316748
63	ALA	H	8.215659401
63	ALA	N	125.8448552
64	ASP	C	175.1318696
64	ASP	CA	52.05107778
64	ASP	CB	41.20254127
64	ASP	H	8.113368379
64	ASP	N	121.5969439
65	PRO	C	176.8669252
65	PRO	CA	63.63691812
65	PRO	CB	31.74987372
66	ASN	C	175.0961742
66	ASN	CA	53.44678615
66	ASN	CB	38.76399199
66	ASN	H	8.491760415
66	ASN	N	117.0802751
67	VAL	C	173.7443105
67	VAL	CA	59.79358537
67	VAL	CB	32.13606781
67	VAL	H	7.541604219
67	VAL	N	120.6367208
68	PRO	C	175.988104
68	PRO	CA	62.84805307
68	PRO	CB	31.81339206
69	ASN	C	174.3162655
69	ASN	CA	56.79591924
69	ASN	CB	39.30212137
69	ASN	H	8.866521041
69	ASN	N	118.2999874
70	VAL	C	175.4583356
70	VAL	CA	60.15772912
70	VAL	CB	32.79967257
70	VAL	H	7.64361875
70	VAL	N	113.8096854
72	VAL	C	175.6578487
72	VAL	CB	29.58255551
73	THR	C	177.3451837
73	THR	CA	62.7616061
73	THR	CB	68.70198975
73	THR	H	8.606013929
73	THR	N	116.2135766
74	ARG	C	172.7853249
74	ARG	CA	56.36934338
74	ARG	CB	34.66386138
74	ARG	H	7.284123108

74	ARG	N	123.3767757
75	LEU	C	174.3367754
75	LEU	CA	53.61369298
75	LEU	CB	41.24309593
75	LEU	H	8.051108187
75	LEU	N	125.1517155
76	THR	C	172.7716218
76	THR	CA	61.24889606
76	THR	CB	71.42672964
76	THR	H	9.292995032
76	THR	N	122.6702904
77	LEU	C	175.0144964
77	LEU	CA	54.73404875
77	LEU	CB	41.44846918
77	LEU	H	8.672757714
77	LEU	N	127.6591183
78	VAL	C	175.9171729
78	VAL	CA	61.52359431
78	VAL	CB	32.51666914
78	VAL	H	8.980762457
78	VAL	N	125.9096142
80	SER	C	176.5165126
80	SER	CA	63.56662358
80	SER	CB	67.94601604
80	SER	H	7.844764374
80	SER	N	117.8510987
81	THR	C	174.850711
81	THR	CA	61.40473619
81	THR	CB	70.72092303
81	THR	H	8.586673558
81	THR	N	114.2305785
82	ALA	C	179.9874858
82	ALA	CA	54.66905367
82	ALA	CB	17.9052638
82	ALA	H	8.722519849
82	ALA	N	123.6377799
83	PRO	C	175.9635047
83	PRO	CA	63.71167891
83	PRO	CB	31.13921534
84	GLY	C	170.4511499
84	GLY	CA	44.307279
84	GLY	H	7.033123812
84	GLY	N	105.7940002
86	THR	C	174.0594129
87	GLU	C	172.9189027
87	GLU	CA	55.49211679
87	GLU	CB	28.74847401
87	GLU	H	9.028899373
87	GLU	N	119.2845872
88	LEU	C	173.2910002
88	LEU	CA	53.15711808
88	LEU	CB	43.08523542
88	LEU	H	9.274262222
88	LEU	N	115.5113864
89	ASP	C	176.8825768
89	ASP	CA	53.40977406
89	ASP	CB	39.42990757
89	ASP	H	8.521404485
89	ASP	N	121.4739282
90	LEU	C	177.0185035
90	LEU	CA	55.19333344
90	LEU	CB	38.86001866
90	LEU	H	7.38923306
90	LEU	N	125.0118208
91	THR	C	175.145183
91	THR	CA	62.44265956
91	THR	CB	69.62669827
91	THR	H	8.425922441
91	THR	N	110.1032813
92	GLY	C	172.6818747
92	GLY	CA	44.2152531
92	GLY	H	7.273457322
92	GLY	N	110.0742229
93	ASP	C	176.8317948
93	ASP	CA	53.66257713
93	ASP	CB	40.41936938
93	ASP	H	8.142581271
93	ASP	N	119.9591569
94	LEU	C	180.1853452
94	LEU	CA	56.63305998

94	LEU	CB	41.11607653
95	GLU	C	179.84697
95	GLU	CA	58.45597372
95	GLU	CB	28.61730392
95	GLU	H	8.127866424
95	GLU	N	118.2450079
96	SER	C	176.9881472
96	SER	CA	61.26207112
96	SER	CB	62.902228
96	SER	H	7.579010983
96	SER	N	114.295695
97	PHE	C	178.1675602
97	PHE	CA	57.42599713
97	PHE	CB	37.02726242
97	PHE	H	6.992551217
97	PHE	N	120.0399799
98	LYS	C	177.3366358
98	LYS	CA	58.97699354
98	LYS	CB	31.99788079
98	LYS	H	7.502774891
98	LYS	N	115.8449422
99	LYS	C	175.5155227
99	LYS	CA	55.58933362
99	LYS	CB	32.2208361
99	LYS	H	7.280174634
99	LYS	N	115.1630147
100	GLN	C	174.0825022
100	GLN	CA	54.38864884
100	GLN	CB	31.33477008
100	GLN	H	7.400468746
100	GLN	N	118.6277049
101	SER	C	175.0414862
101	SER	CA	57.79642286
101	SER	CB	64.94484704
101	SER	H	8.095258263
101	SER	N	112.0296597
102	PHE	C	175.5132097
102	PHE	CA	55.58817674
102	PHE	CB	40.51353223
102	PHE	H	8.310630783
102	PHE	N	122.143852
103	VAL	C	174.9294584
103	VAL	CA	61.8978404
103	VAL	CB	32.44756955
103	VAL	H	8.612929274
103	VAL	N	123.7918908
104	LEU	CA	56.00616031
104	LEU	CB	41.03496116
104	LEU	H	9.619278797
104	LEU	N	121.1295252
108	VAL	C	176.8068762
108	VAL	CA	62.19239327
108	VAL	CB	31.53037715
108	VAL	H	8.000210338
108	VAL	N	118.9326563
109	GLU	C	175.3036005
109	GLU	CA	56.23579827
109	GLU	CB	31.38368132
109	GLU	H	8.195867518
109	GLU	N	117.099159
113	LYS	C	176.9583266
113	LYS	CA	53.59563613
113	LYS	CB	32.10243852
114	ILE	C	173.5859335
114	ILE	CA	63.05211121
114	ILE	CB	41.20633138
114	ILE	H	8.711181891
114	ILE	N	127.2055515
115	SER	CA	58.58495985
115	SER	CB	65.04782825
115	SER	H	8.33125525
115	SER	N	121.7339362
117	ARG	C	174.7073177
117	ARG	CA	60.42525032
117	ARG	CB	34.41245509
118	VAL	C	175.6080175
118	VAL	CA	62.50139729
118	VAL	CB	29.80452705
118	VAL	H	8.824988563
118	VAL	N	129.9379898

121	GLU	C	174.3942614
122	ILE	C	175.5694385
122	ILE	CA	62.57315951
122	ILE	CB	39.03248737
122	ILE	H	8.313040476
122	ILE	N	120.2753541
123	VAL	C	174.4615953
123	VAL	CA	60.49810626
123	VAL	CB	33.92013593
123	VAL	H	8.457296216
123	VAL	N	126.1257939
124	SER	C	174.2808359
124	SER	CA	58.07316162
124	SER	CB	65.25599776
124	SER	H	8.536959397
124	SER	N	121.0484198
125	GLY	C	171.7514556
125	GLY	CA	43.33725401
125	GLY	H	8.465823558
125	GLY	N	106.6025079
130	GLN	C	174.9160167
131	HIS	C	174.0811759
131	HIS	CA	54.83545648
131	HIS	CB	34.27051277
131	HIS	H	9.488735309
131	HIS	N	129.3099039
132	THR	C	172.8469013
132	THR	CA	63.11938761
132	THR	CB	70.12400677
132	THR	H	8.744668844
132	THR	N	119.4006078
133	TYR	C	174.9221457
133	TYR	CA	56.49278057
133	TYR	CB	41.81587316
133	TYR	H	9.787159963
133	TYR	N	126.7937034
135	LYS	C	176.2129791
135	LYS	CB	29.55750893
136	GLY	C	173.7813014
136	GLY	CA	45.21022709
136	GLY	H	8.884725252
136	GLY	N	103.8148959
137	VAL	C	174.7412733
137	VAL	CA	61.07720401
137	VAL	CB	33.79834719
137	VAL	H	7.778977829
137	VAL	N	121.2988736
138	LYS	C	176.3131014
138	LYS	CA	57.22947476
138	LYS	CB	30.80056602
138	LYS	H	8.660945013
138	LYS	N	127.6248554
139	ILE	C	175.6027346
139	ILE	CA	61.23380815
139	ILE	CB	38.65901256
139	ILE	H	8.914777343
139	ILE	N	123.7092024
141	LYS	C	175.6679738
141	LYS	CA	57.09707815
141	LYS	CB	25.73721497
141	LYS	H	8.314776593
141	LYS	N	113.835224
142	THR	C	171.7141773
142	THR	CA	58.70972111
142	THR	CB	70.90635854
142	THR	H	9.013989761
142	THR	N	123.1317868
143	ASP	C	178.5166723
143	ASP	CA	53.61396706
143	ASP	CB	42.50712347
143	ASP	H	8.705707702
143	ASP	N	120.8522297
144	TYR	C	176.7782297
144	TYR	CA	57.21076547
144	TYR	CB	40.88056348
144	TYR	H	7.950133536
144	TYR	N	118.6584636
145	MET	C	175.1403723
145	MET	CA	55.63664493
145	MET	CB	30.89039727

145	MET	H	8.805128933
145	MET	N	122.9977916
146	VAL	C	176.7694919
146	VAL	CA	63.49103877
146	VAL	CB	30.85548274
146	VAL	H	8.443438792
146	VAL	N	128.3151397
146	VAL	ND1	128.1457714
147	GLY	C	172.3638643
147	GLY	CA	44.45911103
147	GLY	H	7.599616869
147	GLY	N	106.687945
148	SER	C	173.2586931
148	SER	CA	57.23347724
148	SER	CB	63.76669067
148	SER	H	7.948073396
148	SER	N	115.8695364
151	PRO	C	174.7669423
151	PRO	CA	61.68924806
151	PRO	CB	32.55955561
152	ARG	C	176.2297574
152	ARG	CA	53.97409853
152	ARG	CB	32.80327045
152	ARG	H	8.842722611
152	ARG	N	121.8883435
153	ALA	C	178.1062093
153	ALA	CA	54.90076672
153	ALA	CB	18.43842389
153	ALA	H	8.518027614
153	ALA	N	124.469636
154	GLU	C	177.4784007
154	GLU	CA	56.65514355
154	GLU	CB	29.06040388
154	GLU	H	7.916572194
154	GLU	N	116.4746169
155	GLU	C	176.766082
155	GLU	CA	57.27143303
155	GLU	CB	29.97856306
155	GLU	H	8.593084514
155	GLU	N	122.9570103
156	TYR	C	174.8326131
156	TYR	CA	57.01333867
156	TYR	CB	41.28271628
156	TYR	H	8.901281529
156	TYR	N	122.0948358
157	GLU	C	175.5041549
157	GLU	CA	55.39913899
157	GLU	CB	32.91324867
157	GLU	H	8.929860499
157	GLU	N	119.742696
158	PHE	C	172.2463089
158	PHE	CA	56.0300649
158	PHE	CB	42.19786469
158	PHE	H	9.150378845
158	PHE	N	124.0645021
159	LEU	C	175.968869
159	LEU	CA	53.30351919
159	LEU	CB	43.90769905
159	LEU	H	7.677010843
159	LEU	N	127.7502198
160	THR	C	172.1087691
160	THR	CA	60.68928067
160	THR	CB	67.71931356
160	THR	H	8.499308675
160	THR	N	113.2764125
163	GLU	C	175.5570367
163	GLU	CA	56.79177278
164	GLU	C	173.800896
164	GLU	CA	54.29978701
164	GLU	CB	32.90216268
164	GLU	H	8.755012862
164	GLU	N	126.3795227
165	ALA	CA	54.32169479
165	ALA	CB	18.25571987
165	ALA	H	8.505097658
165	ALA	N	124.5536797
166	PRO	C	173.6190692
166	PRO	CA	63.12885423
166	PRO	CB	32.63393008
167	LYS	C	175.3952886

167	LYS	CA	55.25494319
167	LYS	CB	36.20455012
167	LYS	H	8.016366718
167	LYS	N	114.1579309
168	GLY	C	174.5366957
168	GLY	CA	43.6903948
168	GLY	H	8.329509138
168	GLY	N	110.4196067
169	TYR	CA	58.94126351
169	TYR	CB	35.81312564
170	LEU	C	177.5248089
170	LEU	CA	56.6299311
170	LEU	CB	41.13225646
170	LEU	H	8.247373982
170	LEU	N	120.880576
171	ALA	C	178.6545754
171	ALA	CA	51.99208141
171	ALA	CB	20.1891937
171	ALA	H	7.373271467
171	ALA	N	119.524631
172	ARG	C	175.6998403
172	ARG	CA	57.65710994
172	ARG	CB	30.72650844
172	ARG	H	7.004170698
172	ARG	N	117.480941
173	GLY	C	173.3552727
173	GLY	CA	43.95279064
173	GLY	H	8.821431269
173	GLY	N	109.1406852
174	SER	C	174.3793386
174	SER	CA	59.62170894
174	SER	CB	64.3676362
174	SER	H	8.345588078
174	SER	N	115.1578651
175	TYR	C	173.7067636
175	TYR	CA	57.77972986
175	TYR	CB	40.12217887
175	TYR	H	9.547432129
175	TYR	N	125.9753345
176	ASN	C	172.1812781
176	ASN	CA	52.96565029
176	ASN	CB	42.37340774
176	ASN	H	8.935834057
176	ASN	N	121.8659942
177	ILE	C	175.2666407
177	ILE	CA	59.78076889
177	ILE	CB	36.60346158
177	ILE	H	9.200338775
177	ILE	N	127.2316
178	LYS	C	175.437767
179	SER	C	172.8377811
179	SER	CA	59.34363734
179	SER	CB	49.4151389
179	SER	H	8.813211362
179	SER	N	122.9475923
180	ARG	C	173.4252685
180	ARG	CA	25.76102376
181	PHE	C	174.5308269
181	PHE	CA	56.22509688
181	PHE	CB	34.18861122
181	PHE	H	7.995871767
181	PHE	N	122.5311985
182	THR	C	170.899711
182	THR	CA	61.7231815
182	THR	CB	71.55950453
182	THR	H	8.502815235
182	THR	N	120.5262398
183	ASP	C	174.4141288
183	ASP	CA	52.56199447
183	ASP	CB	42.48647851
183	ASP	H	7.873384257
183	ASP	N	123.1665482
184	ASP	C	175.3688237
184	ASP	CA	56.76263483
184	ASP	CB	40.5741695
184	ASP	H	9.218214066
184	ASP	N	120.0345401
185	ASP	C	175.2908804
185	ASP	CB	40.93561292
185	ASP	H	8.82283761

185	ASP	N	123.2520763
186	ARG	C	175.7445419
186	ARG	CA	57.08287525
186	ARG	CB	25.80449518
186	ARG	H	8.291270186
186	ARG	N	113.9372053
187	THR	C	174.9724431
187	THR	CA	63.68254569
187	THR	CB	67.9933258
187	THR	H	7.832227865
187	THR	N	118.008283
188	ASP	C	175.3846852
188	ASP	CA	52.18314737
188	ASP	CB	38.88105391
188	ASP	H	9.008680825
188	ASP	N	129.0854517
189	HIS	C	175.4250474
189	HIS	CA	53.40853956
189	HIS	CB	33.94496807
190	LEU	C	175.0313149
190	LEU	CA	54.62602174
190	LEU	CB	43.70642474
190	LEU	H	9.291526618
190	LEU	N	121.9750193
191	SER	C	172.9424931
191	SER	CA	58.54468972
191	SER	CB	65.02767354
191	SER	H	8.321645382
191	SER	N	121.6276224
192	TRP	C	171.9909691
192	TRP	CA	57.56228208
192	TRP	CB	32.22882172
192	TRP	H	9.097645852
192	TRP	N	123.1665275
193	GLU	C	174.0908189
193	GLU	CA	54.81199678
193	GLU	CB	33.5786737
193	GLU	H	8.343165367
193	GLU	HA	7.187713447
193	GLU	N	119.6353695
194	TRP	C	176.6089578
194	TRP	CA	59.64952202
194	TRP	CB	29.96171255
194	TRP	H	8.171592882
194	TRP	N	117.1132875
195	ASN	C	173.2783757
195	ASN	CA	55.28537311
195	ASN	CB	33.68636168
195	ASN	H	7.640730307
195	ASN	N	114.6947952
196	LEU	C	175.4995057
196	LEU	CA	53.68583453
196	LEU	CB	46.66491103
196	LEU	H	8.739022515
196	LEU	N	120.8575914
197	THR	C	172.7541104
197	THR	CA	61.71439056
197	THR	CB	70.61432635
197	THR	H	9.307838151
197	THR	N	120.9908797
198	ILE	C	176.7561513
198	ILE	CA	58.97825715
198	ILE	CB	35.84369453
198	ILE	H	9.086661955
198	ILE	N	127.3350479
199	LYS	C	174.582339
199	LYS	CA	54.72280262
199	LYS	CB	33.34745477
199	LYS	H	8.999823582
199	LYS	N	120.6886997
200	LYS	C	174.9579475
200	LYS	CA	59.50774977
200	LYS	CB	32.58636238
200	LYS	H	8.637320564
200	LYS	N	119.8319787
201	GLU	C	175.1716209
201	GLU	CA	53.33016938
201	GLU	CB	32.92510235
201	GLU	H	7.175998735
201	GLU	N	111.4665922

202	TRP	C	176.5849329	203	LYS	H	7.623261425
202	TRP	CA	59.74687846	203	LYS	N	114.6357665
202	TRP	CB	30.02060751	204	ASP	C	180.7587535
202	TRP	H	8.161022909	204	ASP	CA	56.23492654
202	TRP	N	117.0391622	204	ASP	CB	42.05495745
203	LYS	C	174.7934937	204	ASP	H	7.816562541
203	LYS	CA	55.31131051	204	ASP	N	126.3352678
203	LYS	CB	33.77173025				

Table 11.4. Chemical shift assignment of SH3 R21A mutant.

#	Name	Atom	Chemical shift	#	Name	Atom	Chemical shift
2	ASP	C	169.2269	9	VAL	CA	54.7855
2	ASP	CA	50.7108	9	VAL	H	9.1019
2	ASP	H	8.7913	9	VAL	N	111.7584
2	ASP	N	122.8353	10	LEU	C	169.8364
3	GLU	C	172.6677	10	LEU	CA	49.4079
3	GLU	CA	52.9140	10	LEU	H	8.9388
3	GLU	H	8.5679	10	LEU	N	123.0840
3	GLU	N	121.7098	11	ALA	C	174.0047
4	THR	C	173.6253	11	ALA	CA	49.2860
4	THR	CA	59.3671	11	ALA	H	9.0244
4	THR	H	8.1789	11	ALA	N	126.5966
4	THR	N	114.5489	12	LEU	C	175.5777
5	GLY	C	172.6061	12	LEU	CA	52.5113
5	GLY	CA	42.2081	12	LEU	H	9.1872
5	GLY	H	8.3540	12	LEU	N	127.2961
5	GLY	N	111.2265	13	TYR	C	172.3728
6	LYS	C	171.0751	13	TYR	CA	51.5824
6	LYS	CA	52.7150	13	TYR	H	7.0141
6	LYS	H	7.9406	13	TYR	N	111.3987
6	LYS	N	119.9289	14	ASP	C	171.1341
7	GLU	C	173.4149	14	ASP	CA	51.2822
7	GLU	CA	52.2715	14	ASP	H	8.3073
7	GLU	H	8.5899	14	ASP	N	117.2304
7	GLU	N	122.5863	15	TYR	C	173.1789
8	LEU	C	171.4487	15	TYR	CA	56.2969
8	LEU	CA	50.3271	15	TYR	H	8.6399
8	LEU	H	8.2916	15	TYR	N	120.1227
8	LEU	N	122.7801	16	GLN	C	170.1313
9	VAL	C	174.7303	16	GLN	CA	50.4761

16	GLN	H	7.5774
16	GLN	N	126.9272
17	GLU	C	171.4274
17	GLU	CA	52.9197
17	GLU	H	8.0105
17	GLU	N	122.8549
18	LYS	C	173.2379
18	LYS	CA	52.1351
18	LYS	H	8.4574
18	LYS	N	120.2840
19	SER	C	173.3653
19	SER	CA	53.2267
19	SER	H	7.6794
19	SER	N	115.6744
21	ALA	C	173.6312
21	ALA	CA	49.1672
21	ALA	H	7.6752
21	ALA	N	117.5483
22	GLU	C	175.3221
22	GLU	CA	51.8963
22	GLU	H	7.7232
22	GLU	N	118.4591
23	VAL	C	171.7098
23	VAL	CA	56.8768
23	VAL	H	7.3793
23	VAL	N	114.3135
24	THR	C	169.7820
24	THR	CA	58.4802
24	THR	H	7.4883
24	THR	N	119.1162
25	MET	C	171.6846
25	MET	CA	51.2557
25	MET	H	9.4508
25	MET	N	121.4774
26	LYS	C	171.2071
26	LYS	CA	50.5213
26	LYS	H	8.4195
26	LYS	N	123.4669
27	LYS	C	172.3728
27	LYS	CA	55.2875
27	LYS	H	8.7912
27	LYS	N	122.3909
28	GLY	C	174.4356
28	GLY	CA	41.5599
28	GLY	H	8.7859
28	GLY	N	115.1834
29	ASP	C	171.5273
29	ASP	CA	51.9986
29	ASP	H	8.4101
29	ASP	N	121.1347
30	ILE	C	171.4093
30	ILE	C	51.9304
30	ILE	CA	57.0815
30	ILE	H	8.0842
30	ILE	N	120.5330
31	LEU	C	173.5918
31	LEU	CA	50.7345
31	LEU	H	9.2904
31	LEU	N	126.8591
32	THR	C	173.0604
32	THR	CA	60.1858
32	THR	H	8.3903
32	THR	N	116.8141
33	LEU	C	171.1144
33	LEU	CA	51.6615
33	LEU	H	8.9685
33	LEU	N	128.6856
34	LEU	C	171.9599
34	LEU	CA	52.2033
34	LEU	H	8.9862
34	LEU	N	126.1697
35	ASN	C	174.9513
35	ASN	CA	51.3031
35	ASN	H	7.5717
35	ASN	N	113.6778
36	SER	C	171.7633
36	SER	CA	53.7431
36	SER	H	9.0480
36	SER	N	123.0856
37	THR	C	170.7889

37	THR	CA	62.4373
37	THR	H	8.1300
37	THR	N	115.3356
38	ASN	C	172.6481
38	ASN	CA	50.2588
38	ASN	H	8.6524
38	ASN	N	122.4169
39	LYS	C	172.0385
39	LYS	CA	55.2735
39	LYS	H	8.4163
39	LYS	N	120.6951
40	ASP	C	173.8708
40	ASP	CA	51.7598
40	ASP	H	8.2330
40	ASP	N	113.9495
41	TRP	C	173.8094
41	TRP	CA	52.9515
41	TRP	H	8.0260
41	TRP	N	122.6691
42	TRP	C	171.4487
42	TRP	CA	50.8596
42	TRP	H	9.2772
42	TRP	N	124.3272
43	LYS	C	171.7436
43	LYS	CA	52.2507
43	LYS	H	8.8758
43	LYS	N	123.9980
44	VAL	C	172.7857
44	VAL	CA	55.9463
44	VAL	H	9.3122
44	VAL	N	121.9620
45	GLU	C	170.7766
45	GLU	CA	50.9411
45	GLU	H	8.6300
45	GLU	N	118.2702
46	VAL	C	172.1368
46	VAL	CA	57.5809
46	VAL	H	8.8254
46	VAL	N	124.7261
47	ASN	C	171.9009
47	ASN	CA	52.0375
47	ASN	H	9.3652
47	ASN	N	126.2928
48	ASP	C	171.5502
48	ASP	CA	51.4869
48	ASP	H	8.6488
48	ASP	N	112.0596
49	ARG	C	171.5863
49	ARG	CA	52.0668
49	ARG	H	8.1084
49	ARG	N	120.1398
50	GLN	C	171.6729
50	GLN	CA	50.1565
50	GLN	H	8.4427
50	GLN	N	118.5065
51	GLY	C	173.1955
51	GLY	CA	42.1508
51	GLY	H	8.5464
51	GLY	N	106.6199
52	PHE	C	167.8308
52	PHE	CA	55.6146
52	PHE	H	9.1739
52	PHE	N	118.8946
53	VAL	C	173.1587
53	VAL	CA	55.0006
53	VAL	H	8.9991
53	VAL	N	110.9168
55	ALA	C	175.3024
55	ALA	CA	58.8437
55	ALA	H	7.3900
55	ALA	N	128.8380
56	ALA	C	175.8231
56	ALA	CA	49.6394
56	ALA	H	7.7615
56	ALA	N	113.1491
57	TYR	C	173.4345
57	TYR	CA	53.2267
57	TYR	H	7.6783
57	TYR	N	116.1144
58	VAL	C	171.8694

58	VAL	CA	55.1521	60	LYS	N	125.5321
58	VAL	H	7.3617	60	LYS	N	125.5330
58	VAL	N	110.9794	61	LEU	C	174.0637
59	LYS	C	170.9362	61	LEU	CA	51.8243
59	LYS	CA	50.9070	61	LEU	H	8.4313
59	LYS	H	8.5679	61	LEU	N	124.7681
59	LYS	N	118.4026	62	ASP	C	173.7098
60	LYS	C	173.4739	62	ASP	CA	50.2929
60	LYS	CA	55.4014	62	ASP	H	7.9774
60	LYS	H	9.1614	62	ASP	N	122.2445

11.4 Experimental parameters for NMR spectra

Table 11.5. Experimental parameters used for the NMR spectra in the solid state.

hCACONH 4D				
	F4 (^{1}H)	F3 (^{15}N)	F2 (^{13}C)	F1 (^{13}C)
FID size	2048	130	50	60
SW (ppm)	26.03	40.29	12.54	30.67
Increment size (µs)		306	396	162
Aq. Time (ms)	50	20	10	5
scans	8			
NUS points	1000			
NUS %	2.95			
hCOCAHN 4D				
	F4 (^{1}H)	F3 (^{15}N)	F2 (^{13}C)	F1 (^{13}C)
FID size	2048	112	28	104
SW (ppm)	29.75	40.08	29.95	20.03
Increment size (µs)		351.6	189.6	283.45
Aq. Time (ms)	50	20	2.6	14.7
scans	16			
NUS points	4000			
NUS %	9.8			
hCACBcaNH 4D				
	F4 (^{1}H)	F3 (^{15}N)	F2 (^{13}C)	F1 (^{13}C)
FID size	2048	74	96	48
SW (ppm)	29.75	35	79	40
Increment size (µs)		402.6	71.90	141.98
Aq. Time / ms	50	15	3.5	3.5
scans	8			

NUS points	2000			
NUS %	4.6			

HNcocaNH 4D				
	F4 (^{1}H)	F3 (^{15}N)	F2 (^{13}C)	F1 (^{13}C)
FID size	2048	92	92	50
SW (ppm)	26.03	38.05	38.05	6.31
Increment size (µs)		324	324	198
Aq. Time (ms)	50	15	15	5
scans	8			
NUS points	2000			
NUS %	3.78			

hCAncoCANH 4D				
	F4 (^{1}H)	F3 (^{15}N)	F2 (^{13}C)	F1 (^{13}C)
FID size	2048	110	142	78
SW (ppm)	29.75	39.14	45.07	45.07
Increment size (µs)		360	126	126
Aq. Time (ms)	50	20	9	5
scans	16			
NUS points	8192			
NUS %	5.37			

Pseudo4D R1ρ			
	F3 (^{1}H)	F2 (^{15}N)	F1 (^{13}C)
FID size	2048	100	66
SW (ppm)	29.75	35.05	35.05
Increment size (µs)		402	162
Aq. Time (ms)	50	20	5.3
scans	12		
NUS points (%)	495 and 660		
NUS %	30% and 40%		

R1ρ 3D			
	F3 (^{1}H)	F2 (^{15}N)	F1 (^{13}C)
FID size	2048	112	66
SW (ppm)	29.75	40.08	35.05
Increment size (µs)		351.6	161.9
Aq. Time (ms)	50	20	5.3
scans	8		
NUS points	739		
NUS %	40		

Relaxation Dispersion

Table 11.6. $R_{1\rho}$ relaxation rates at different effective field strength for p38α fully deuterated sample recorded as 3D with 40% NUS density.

Residue	#	10 kHz	20 kHz	30 kHz	38 kHz	42 kHz	44 kHz
9THRH	9	6.600	9.988	5.749	3.944	6.170	7.468
10PHEH	10	1.901	3.146	2.656	2.212	0.248	3.673
11TYRH	11	2.764	3.219	3.209	0.749	0.986	3.759
12ARGH	12	2.202	3.650	3.324	0.650	1.563	8.819
13GLNH	13	1.818	1.825	2.067	0.202	1.218	7.691
14GLUH	14	1.728	2.433	2.637	1.323	0.365	6.113
15LEUH	15	2.794	4.115	3.995	2.721	0.722	4.327
18THRH	18	4.031	5.523	3.810	2.394	3.778	5.031
19ILEH	19	3.033	3.198	4.426	2.602	1.029	10.894
20TRPH	20	3.027	3.309	4.921	1.605	3.297	8.843
21GLUH	21	4.164	2.003	7.451	7.996	5.568	7.353
22VALH	22	0.676	2.131	3.037	1.342	0.588	9.230
24GLUH	24	1.948	1.854	2.451	1.455	0.448	0.373
25ARGH	25	2.884	2.271	2.020	1.553	1.204	4.375
28ASNH	28	1.725	1.389	2.064	0.059	0.872	1.985
29LEUH	29	1.462	2.326	2.025	0.234	0.957	2.486
38GLYH	38	2.699	2.911	2.666	"NA"	1.693	"NA"
41CYSH	41	1.349	2.739	2.275	1.391	1.721	6.980
42ALAH	42	1.509	3.166	2.670	2.490	0.685	5.906
44PHEH	44	2.206	2.131	1.529	1.273	3.056	4.660
46THRH	46	8.668	6.315	3.493	8.125	10.156	4.296
49GLYH	49	9.822	5.068	0.285	3.385	7.669	6.007
50LEUH	50	5.408	1.158	3.443	0.002	7.495	13.800
51ARGH	51	4.128	2.600	2.710	0.127	1.543	1.169
56LYSH	56	1.320	1.274	2.558	0.158	0.020	7.009
57LEUH	57	1.042	1.160	2.444	17.088	3.729	7.566
61PHEH	61	6.001	3.673	7.970	"NA"	5.554	7.676
68LYSH	68	3.020	2.641	5.506	3.780	1.355	3.894
82HISH	82	0.560	1.613	2.901	0.409	0.738	4.396
83GLUH	83	0.825	0.806	1.331	2.679	0.651	6.764
84ASNH	84	2.191	2.367	1.358	0.379	1.032	4.202
85VALH	85	2.916	1.925	1.797	0.412	3.352	6.578
87GLYH	87	10.636	6.811	9.586	8.231	11.455	"NA"
93THRH	93	1.663	1.821	0.873	0.320	0.518	5.096
95ALAH	95	1.379	1.009	2.062	0.054	0.339	2.259
96ARGH	96	1.323	1.407	0.775	0.374	2.055	6.775

97SERH	97	2.827	2.320	1.901	3.214	1.293	6.949
98LEUH	98	2.413	3.047	2.940	0.611	2.016	0.851
99GLUH	99	1.917	3.162	1.918	1.298	1.094	10.875
100GLUH	100	1.370	2.998	6.323	0.400	1.597	12.694
101PHEH	101	1.610	1.974	0.727	0.554	0.997	2.137
103ASPH	103	5.423	4.449	4.934	0.949	1.871	16.338
104VALH	104	3.166	6.772	0.548	4.930	1.700	2.531
105TYRH	105	5.718	5.532	5.815	6.425	4.869	3.781
109HISH	109	5.151	3.955	4.043	22.205	6.355	3.886
112GLYH	112	4.025	4.181	0.472	4.039	6.764	4.027
117ASNH	117	1.716	5.020	2.616	7.341	12.811	7.981
121CYSH	121	1.318	1.932	2.682	0.527	0.646	0.820
130GLNH	130	2.048	12.578	1.227	"NA"	14.043	2.466
140LEUH	140	22.161	7.029	8.054	"NA"	17.181	1.663
141LYSH	141	0.893	2.334	3.141	1.303	1.148	0.953
142TYRH	142	1.690	2.022	2.977	3.790	1.274	4.827
143ILEH	143	0.845	2.891	5.021	3.328	5.790	9.716
144HISH	144	3.665	4.053	2.011	1.346	0.243	5.368
145SERH	145	5.777	3.882	0.955	2.378	3.111	1.650
146ALAH	146	0.766	1.358	5.074	0.769	4.047	9.297
147ASPH	147	0.675	0.570	1.120	0.003	1.166	2.164
148ILEH	148	3.157	2.165	2.584	34.451	2.669	3.848
149ILEH	149	0.261	2.664	1.802	0.373	0.599	7.453
157ASNH	157	1.416	0.987	5.456	3.015	0.843	7.175
162GLUH	162	8.347	4.510	8.891	6.216	15.086	3.158
166LEUH	166	4.674	4.258	6.884	1.776	6.980	15.098
167LYSH	167	7.035	7.738	5.999	3.816	6.501	11.613
168ILEH	168	8.230	"NA"	4.186	10.763	20.634	7.587
182THRH	182	5.078	4.264	2.782	0.750	5.651	1.185
183GLYH	183	4.557	1.216	3.904	"NA"	11.839	6.895
191ARGH	191	1.422	2.559	10.504	6.674	13.801	3.658
192ALAH	192	1.555	3.395	2.494	2.287	0.591	4.047
197LEUH	197	0.957	2.272	3.274	2.845	0.141	6.753
198ASNH	198	1.451	3.369	1.350	4.359	6.548	5.430
199TRPH	199	2.476	2.633	2.214	1.262	0.385	9.763
200METH	200	3.524	6.964	5.348	30.346	7.762	9.317
202TYRH	202	7.795	6.858	4.053	4.703	5.628	"NA"
203ASNH	203	8.160	0.139	4.110	1.099	12.388	"NA"
208ILEH	208	2.266	1.125	3.022	1.006	0.544	3.080
210SERH	210	2.134	2.388	0.791	0.367	2.666	9.370

213CYSH	213	2.785	2.413	0.002	"NA"	8.121	7.280
214ILEH	214	5.563	2.579	8.027	4.506	9.676	6.621
218LEUH	218	0.293	1.965	0.623	0.555	0.236	0.902
219LEUH	219	0.885	1.964	3.359	1.267	1.946	5.265
220THRH	220	2.686	2.410	1.465	"NA"	"NA"	2.770
222ARGH	222	3.616	0.231	3.829	3.674	4.323	8.178
223THRH	223	1.848	2.011	2.350	0.369	2.985	0.451
227GLYH	227	3.500	5.757	9.276	18.815	1.688	"NA"
228THRH	228	5.524	3.079	2.806	3.405	3.403	4.360
240LEUH	240	2.050	2.258	1.923	0.013	0.276	5.308
245GLYH	245	8.402	9.870	11.719	7.402	12.362	6.676
246ALAH	246	4.798	4.938	5.060	5.578	3.301	9.184
248LEUH	248	6.295	4.225	3.069	6.433	2.510	5.280
249LEUH	249	3.670	1.208	2.670	1.529	1.413	5.137
256SERH	256	12.112	7.953	14.902	10.049	7.995	"NA"
257ALAH	257	1.062	0.045	1.392	0.084	1.010	10.653
263SERH	263	2.763	3.720	1.072	0.514	"NA"	4.744
272PHEH	272	26.868	18.068	1.615	"NA"	"NA"	3.392
273ALAH	273	5.396	20.708	4.129	2.635	7.472	6.162
274ASNH	274	5.678	8.638	10.864	"NA"	8.514	0.738
276PHEH	276	1.288	2.065	0.364	18.938	1.014	4.164
277ILEH	277	4.155	4.782	4.431	2.300	1.717	8.048
278GLYH	278	17.961	12.804	18.993	29.005	4.011	4.818
279ALAH	279	1.476	3.734	3.171	3.076	0.405	6.577
280ASNH	280	2.344	2.897	0.682	1.514	"NA"	"NA"
282LEUH	282	1.579	0.411	0.883	0.502	2.047	6.843
283ALAH	283	1.224	1.088	2.693	0.740	1.985	8.149
284VALH	284	3.814	9.664	1.407	2.054	3.912	11.539
285ASPH	285	2.053	2.210	1.817	0.302	0.013	4.993
286LEUH	286	5.699	3.816	0.607	1.056	2.243	2.552
296ASPH	296	0.689	2.936	6.689	"NA"	5.829	5.643
297LYSH	297	0.841	2.604	2.697	0.742	2.021	6.745
300THRH	300	1.513	0.777	1.893	1.105	1.171	3.140
301ALAH	301	1.334	1.834	2.626	0.162	1.474	2.317
302ALAH	302	1.302	2.671	2.328	0.054	0.622	6.508
303GLNH	303	1.319	1.789	2.884	0.049	0.670	0.466
305LEUH	305	3.056	2.657	2.055	1.333	0.601	12.199
306ALAH	306	1.400	2.463	1.893	0.249	0.241	5.476
307HISH	307	0.758	1.377	0.294	0.163	0.075	2.335
308ALAH	308	0.639	1.833	0.403	0.114	0.380	1.893

309TYRH	309	0.498	1.077	3.676	0.508	0.792	6.180
311ALAH	311	1.031	1.492	1.385	0.524	0.708	6.814
312GLNH	312	2.624	2.196	3.230	2.192	0.818	2.152
313TYRH	313	3.585	2.132	4.323	0.604	0.214	6.543
314HISH	314	3.366	3.611	8.669	3.236	3.081	6.853
315ASPH	315	2.283	2.007	5.131	6.368	3.673	5.259
318ASPH	318	1.005	0.002	15.939	11.800	26.959	5.639
319GLUH	319	1.789	5.177	0.411	7.549	1.453	15.578
321VALH	321	2.675	3.060	3.013	2.004	1.563	4.068
322ALAH	322	3.248	1.432	2.087	0.518	0.330	7.847
323ASPH	323	0.976	1.572	2.514	1.199	0.997	6.859
326ASPH	326	6.453	3.206	9.158	10.638	"NA"	10.233
330GLUH	330	1.085	5.200	4.032	4.944	"NA"	3.145
331SERH	331	4.902	3.861	3.897	0.423	4.594	12.345
333ASPH	333	2.617	3.638	3.652	1.191	1.035	25.453
337ASPH	337	0.522	4.137	6.712	9.155	10.202	30.714
338GLUH	338	3.341	2.896	3.436	0.659	1.089	5.768
339TRPH	339	0.321	0.902	2.070	3.502	0.163	4.392
340LYSH	340	1.583	2.240	3.821	0.248	1.291	6.975
341SERH	341	3.738	4.116	2.578	1.651	1.047	1.665
346GLUH	346	2.019	2.237	1.846	0.313	0.334	5.803
348ILEH	348	0.659	1.718	1.561	0.272	0.371	7.722
350PHEH	350	0.443	1.943	0.902	0.509	2.418	16.747
351VALH	351	6.519	3.787	8.484	9.955	5.058	6.484

Table 11.7. $R_{1\rho}$ relaxation rates at different effective field strength for the RAP50 p38α in complex with regorafenib and BOG as pseudo-4D with 40% NUS density.

Residue	**#**	**10 kHz**	**20 kHz**	**30 kHz**	**38 kHz**	**42 kHz**	**44 kHz**
T9CA-N-HN	9	3.72	2.10	10.69	3.11	8.62	3.34
F10CA-N-HN	10	0.64	1.10	0.31	1.15	0.85	0.02
Y11CA-N-HN	11	0.70	0.04	3.86	3.37	6.20	11.93
R12CA-N-HN	12	0.42	0.29	1.90	1.58	7.11	0.09
Q13CA-N-HN	13	0.00	0.00	0.77	4.88	3.33	2.95
E14CA-N-HN	14	0.00	0.00	0.01	3.31	6.63	4.61
L15CA-N-HN	15	0.60	3.04	3.32	5.05	15.37	0.10
T18CA-N-HN	18	4.74	9.18	11.22	4.97	0.92	9.91
I19CA-N-HN	19	0.68	0.26	10.92	0.00	8.31	13.71
W20CA-N-HN	20	2.60	2.09	0.21	5.87	11.89	4.25
E21CA-N-HN	21	0.00	0.00	0.07	0.00	0.00	0.23
V22CA-N-HN	22	0.00	0.01	2.31	3.51	0.98	0.09

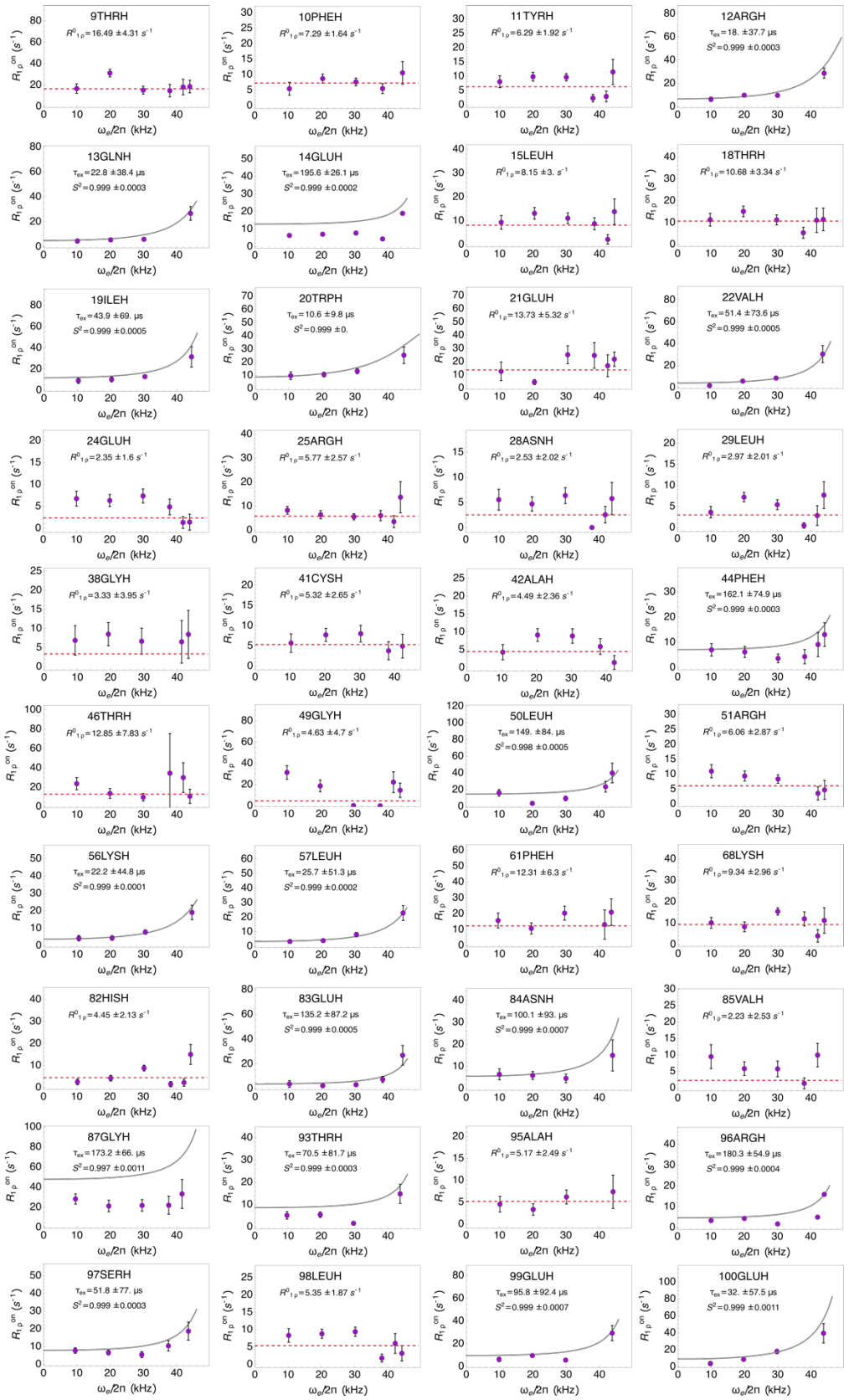
E24CA-N-HN	24	1.76	0.00	0.00	0.46	10.86	0.10
R25CA-N-HN	25	0.05	3.84	6.18	0.01	0.00	1.50
N28CA-N-HN	28	0.00	0.05	0.87	0.74	5.65	10.08
L29CA-N-HN	29	0.15	0.00	0.00	0.35	8.44	2.14
S30CA-N-HN	30	0.00	0.05	1.56	2.95	4.36	1.42
G38CA-N-HN	38	16.90	17.67	8.26	1.05	3.82	2.73
C41CA-N-HN	41	0.00	0.00	0.22	12.23	2.27	0.75
A42CA-N-HN	42	2.44	0.01	0.74	6.64	6.94	0.00
A43CA-N-HN	43	0.54	0.56	0.24	2.70	11.58	4.47
F44CA-N-HN	44	0.00	0.02	1.83	0.00	1.07	1.04
D45CA-N-HN	45	1.59	0.00	6.39	14.42	8.14	0.00
T46CA-N-HN	46	3.47	2.89	9.46	5.87	0.80	1.75
T48CA-N-HN	48	0.07	2.40	2.45	2.29	10.84	5.46
G49CA-N-HN	49	7.82	7.76	0.05	2.19	5.54	0.11
L50CA-N-HN	50	1.19	3.69	25.88	3.65	7.58	9.27
R51CA-N-HN	51	0.14	1.71	1.41	8.58	3.70	1.50
K55CA-N-HN	55	0.30	0.00	11.86	19.37	7.59	1.24
K56CA-N-HN	56	0.10	0.00	0.34	2.95	4.00	0.00
L57CA-N-HN	57	0.00	0.00	0.43	2.96	3.90	0.00
F61CA-N-HN	61	4.01	0.48	1.81	2.35	6.64	0.35
Q62CA-N-HN	62	5.42	2.68	0.22	3.61	5.15	2.83
I64CA-N-HN	64	4.27	4.29	2.82	8.09	5.07	2.18
I65CA-N-HN	65	1.49	5.30	2.47	5.59	10.80	0.01
K68CA-N-HN	68	11.13	0.50	8.01	0.56	0.38	3.56
H82CA-N-HN	82	0.62	0.01	0.00	0.00	0.10	5.14
E83CA-N-HN	83	0.00	0.38	0.00	1.48	4.09	7.30
N84CA-N-HN	84	3.74	0.00	2.18	0.00	6.00	1.35
V85CA-N-HN	85	0.00	0.00	4.83	1.36	8.73	0.48
G87CA-N-HN	87	7.77	2.87	12.12	14.60	2.96	6.97
T93CA-N-HN	93	0.00	0.00	2.40	0.00	6.12	2.09
A95CA-N-HN	95	4.08	0.00	0.52	1.38	0.00	5.61
R96CA-N-HN	96	1.80	0.02	0.04	0.08	4.30	1.24
S97CA-N-HN	97	0.48	0.16	1.83	4.72	0.61	3.03
E99CA-N-HN	99	0.13	0.00	0.00	0.06	0.14	13.20
E100CA-N-HN	100	0.00	0.00	0.03	0.12	7.76	1.53
F101CA-N-HN	101	0.00	0.00	0.00	2.21	3.80	8.58
D103CA-N-HN	103	3.24	1.74	3.16	8.31	8.41	1.54
V104CA-N-HN	104	1.12	1.99	0.04	13.37	25.52	3.94
Y105CA-N-HN	105	0.00	0.00	0.00	5.80	4.26	0.92
H109CA-N-HN	109	0.00	0.96	0.04	0.00	3.95	0.01

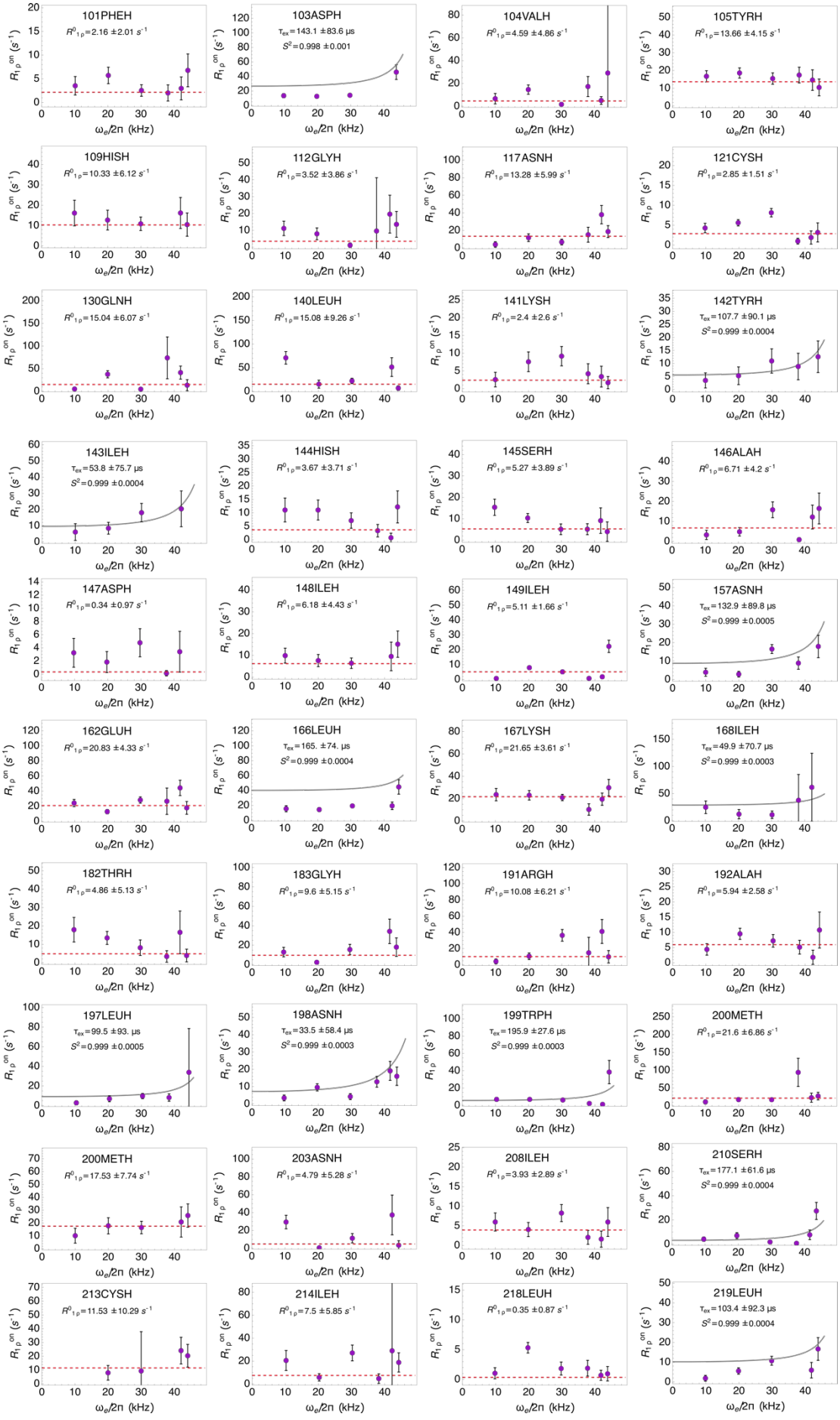
G112CA-N-HN	112	7.71	0.01	18.75	2.17	3.37	0.20
N117CA-N-HN	117	0.81	0.02	0.10	9.59	1.19	0.20
K120CA-N-HN	120	0.03	1.00	0.25	1.11	0.22	0.12
C121CA-N-HN	121	0.00	0.00	0.05	2.91	4.78	0.18
Q130CA-N-HN	130	0.00	5.27	7.12	2.38	8.48	6.25
G139CA-N-HN	139	1.67	0.00	0.15	2.81	4.55	1.64
L140CA-N-HN	140	7.53	5.35	6.52	1.49	0.77	0.19
K141CA-N-HN	141	0.30	0.05	9.55	11.94	0.00	0.21
Y142CA-N-HN	142	0.00	4.50	5.65	0.15	5.46	1.26
I143CA-N-HN	143	0.00	2.94	0.36	1.16	4.50	0.08
H144CA-N-HN	144	2.95	0.35	0.00	5.34	3.39	13.29
S145CA-N-HN	145	0.00	0.00	1.87	3.07	11.74	0.79
A146CA-N-HN	146	0.63	5.39	3.20	2.12	0.00	2.24
D147CA-N-HN	147	0.26	0.20	0.94	7.08	2.06	0.01
I148CA-N-HN	148	0.40	5.50	6.41	7.96	9.67	0.21
I149CA-N-HN	149	0.01	0.00	0.19	2.65	0.03	0.03
N157CA-N-HN	157	2.78	0.00	0.07	0.12	1.58	0.07
L158CA-N-H	158	5.18	5.67	7.00	13.62	7.78	1.25
A159CA-N-HN	159	1.96	0.00	0.79	0.02	1.48	13.35
E162CA-N-HN	162	0.00	2.37	10.20	14.95	0.23	4.93
L166CA-N-HN	166	5.51	5.64	3.84	7.14	2.15	10.47
K167CA-N-HN	167	0.69	2.70	0.57	6.43	8.25	5.48
I168CA-N-HN	168	2.08	14.93	13.09	0.31	0.02	8.26
T182CA-N-HN	182	0.00	0.18	0.01	8.87	3.87	4.42
G183CA-N-HN	183	6.56	2.48	6.74	6.19	7.42	9.27
R191CA-N-H	191	28.36	3.27	2.50	6.73	0.00	3.55
A192CA-N-HN	192	20.73	0.00	0.82	4.02	0.31	0.97
L197CA-N-HN	197	6.35	0.04	4.59	7.53	5.13	3.32
N198CA-N-HN	198	0.13	1.43	2.45	1.65	7.17	10.49
W199CA-N-HN	199	0.92	0.00	0.04	0.78	6.29	0.24
M200CA-N-HN	200	6.71	0.04	3.25	9.64	3.87	4.31
Y202CA-N-HN	202	0.41	1.44	11.98	10.37	6.93	0.03
N203CA-N-HN	203	4.65	0.00	0.05	3.20	7.16	8.42
I208CA-N-HN	208	4.28	0.85	0.21	0.25	12.54	1.07
S210CA-N-HN	210	5.08	0.00	0.88	0.00	0.74	0.20
V211CA-N-HN	211	0.00	0.36	0.03	0.39	0.13	2.77
G212CA-N-HN	212	3.90	4.40	0.00	9.64	5.69	0.05
C213CA-N-HN	213	1.41	0.08	0.00	0.66	2.21	6.07
M215CA-N-HN	215	0.18	3.89	8.37	3.77	0.28	9.81
A216CA-N-HN	216	0.13	0.25	1.51	1.89	3.62	6.55

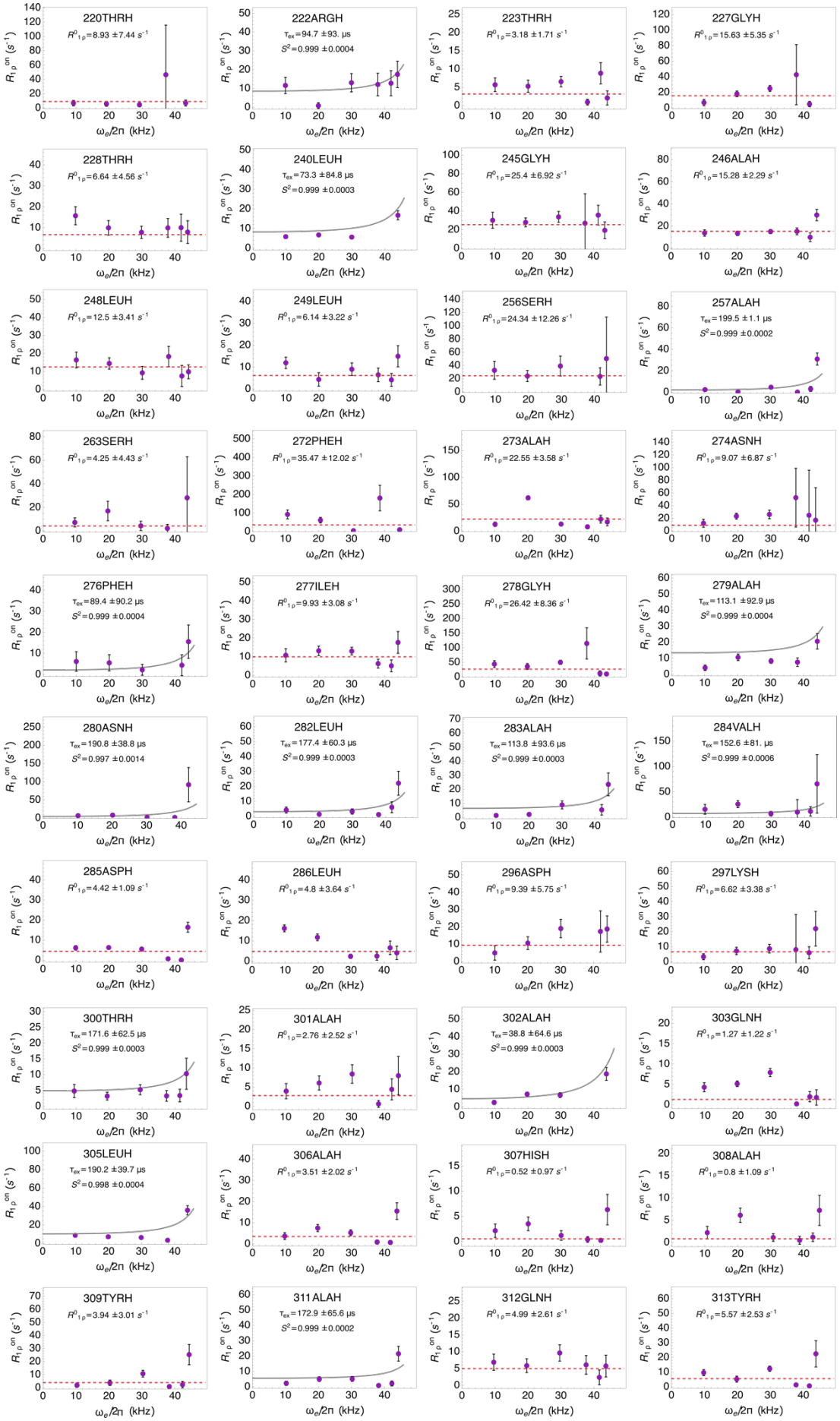
E217CA-N-HN	217	6.45	1.36	1.54	1.51	6.00	6.33
L218CA-N-HN	218	1.28	1.46	0.39	0.30	1.04	0.32
L219CA-N-HN	219	2.03	0.09	0.04	0.50	4.41	0.85
T220CA-N-HN	220	3.27	0.12	2.38	8.35	3.41	8.62
R222CA-N-HN	222	0.01	0.78	0.04	0.40	13.10	0.22
T223CA-N-HN	223	1.61	0.64	2.86	15.73	0.74	0.14
G227CA-N-HN	227	1.95	0.00	2.05	2.70	26.89	3.53
T228CA-N-HN	228	6.19	7.38	1.71	1.15	5.38	6.78
L240CA-N-HN	240	0.71	0.00	0.00	11.97	1.38	3.98
G245CA-N-HN	245	9.18	0.07	12.72	5.90	0.05	0.01
A246CA-N-HN	246	0.22	4.71	4.96	0.00	3.42	1.14
L248CA-N-HN	248	0.00	6.00	3.85	5.72	3.41	8.82
L249CA-N-HN	249	0.00	4.38	13.14	19.09	2.14	1.03
S256CA-N-HN	256	13.88	3.14	2.08	11.98	6.08	6.80
F272CA-N-HN	272	4.53	5.65	9.51	4.26	8.64	12.80
A273CA-N-HN	273	0.18	4.34	1.76	1.23	9.62	0.03
N274CA-N-HN	274	1.44	1.37	6.20	8.60	0.00	2.55
I277CA-N-HN	277	0.34	4.10	0.00	4.02	1.94	1.54
G278CA-N-HN	278	4.21	4.02	1.60	13.08	8.19	1.51
A279CA-N-HN	279	2.10	0.00	2.50	4.57	2.57	1.49
N280CA-N-HN	280	3.67	2.25	1.06	0.00	7.79	2.06
L282CA-N-HN	282	0.44	5.00	0.04	5.95	7.16	6.47
A283CA-N-HN	283	0.07	2.45	2.55	3.06	0.05	4.55
V284CA-N-HN	284	2.62	3.68	1.22	0.34	6.78	3.56
D285CA-N-HN	285	0.77	0.00	0.00	1.27	0.93	6.20
L286CA-N-HN	286	0.00	4.24	8.79	1.78	11.52	0.50
L287CA-N-HN	287	0.00	8.88	1.90	0.00	19.51	0.35
M290CA-N-HN	290	0.00	4.61	4.26	2.49	8.21	0.21
D296CA-N-HN	296	3.69	3.93	2.10	1.02	0.15	0.00
K297CA-N-HN	297	0.26	0.36	4.36	7.51	2.25	0.37
R298CA-N-HN	298	0.14	0.00	3.11	0.16	30.95	2.28
I299CA-N-HN	299	14.69	0.74	4.81	0.41	37.32	24.42
T300CA-N-HN	300	0.76	0.00	1.93	2.84	5.85	0.00
A301CA-N-HN	301	0.00	0.22	0.00	0.50	3.71	0.00
A302CA-N-HN	302	0.04	0.02	1.61	3.18	3.68	1.51
Q303CA-N-HN	303	0.00	1.55	0.09	2.88	4.94	2.55
A304CA-N-HN	304	3.98	1.57	6.48	4.03	0.20	0.02
L305CA-N-HN	305	0.10	8.89	0.00	2.08	4.37	2.85
A306CB-N-HN	306	2.51	0.00	0.16	0.82	5.18	2.05
H307CA-N-HN	307	0.02	0.01	0.00	0.43	1.74	0.37

A308CA-N-HN	308	0.32	0.00	0.20	2.55	2.20	5.25
Y309CA-N-HN	309	0.00	0.00	0.00	0.31	0.00	0.27
F310CA-N-HN	310	2.84	1.71	3.74	9.42	0.97	10.04
A311CA-N-HN	311	0.00	0.00	0.00	0.46	3.41	0.00
Q312CA-N-HN	312	2.25	0.00	0.12	2.14	2.74	4.43
Y313CA-N-HN	313	0.00	1.48	0.06	2.80	2.15	3.47
H314CA-N-HN	314	0.04	7.47	3.66	20.64	6.05	0.20
D315CA-N-HN	315	0.35	0.01	1.88	0.22	8.24	5.73
D318CA-N-HN	318	0.00	2.59	0.00	1.01	0.00	4.10
E319CA-N-HN	319	9.49	10.73	0.31	3.16	6.01	1.58
V321CA-N-HN	321	0.79	1.40	1.89	0.21	5.64	4.91
A322CA-N-HN	322	1.15	1.38	0.00	5.73	3.39	3.60
D323CA-N-HN	323	0.84	0.03	0.25	10.17	4.49	1.90
E330CA-N-HN	330	2.32	0.05	9.31	0.20	0.75	1.59
S331CA-N-HN	331	2.23	0.16	0.36	1.16	4.87	5.39
D333CA-N-HN	333	0.00	0.17	5.04	0.63	6.62	5.33
D337CA-N-HN	337	0.00	0.39	0.03	5.76	2.08	1.56
E338CA-N-HN	338	6.46	1.50	1.62	3.56	5.91	2.69
W339CA-N-HN	339	1.26	3.53	3.67	0.00	6.54	14.15
K340CA-N-HN	340	0.02	0.00	0.00	4.25	3.80	1.59
S341CA-N-HN	341	0.00	0.11	0.00	0.00	3.92	11.04
Y344CA-N-HN	344	0.00	3.23	4.36	0.71	0.15	1.12
E346CA-N-HN	346	0.04	0.00	0.00	1.39	1.44	4.60
V347CA-N-HN	347	1.85	8.03	1.53	2.30	11.33	10.47
I348CA-N-HN	348	0.31	5.54	0.08	7.04	4.02	7.10
F350CA-N-HN	350	0.00	0.01	2.99	5.07	16.91	8.49
V351CA-N-HN	351	1.96	0.00	4.48	1.55	1.65	2.78

Figure 11.3. ^{15}N NERRD profiles for p38α in complex with sorafenib and BOG, fully deuterated with 40% NUS




101PHEH
$R^0_{1\rho}$=2.16 ±2.01 s^{-1}
103ASPH
τ_{ex}=143.1 ±83.6 µs
S^2=0.998 ±0.001
104VALH
$R^0_{1\rho}$=4.59 ±4.86 s^{-1}
105TYRH
$R^0_{1\rho}$=13.66 ±4.15 s^{-1}
109HISH
$R^0_{1\rho}$=10.33 ±6.12 s^{-1}
112GLYH
$R^0_{1\rho}$=3.52 ±3.86 s^{-1}
117ASNH
$R^0_{1\rho}$=13.28 ±5.99 s^{-1}
121CYSH
$R^0_{1\rho}$=2.85 ±1.51 s^{-1}
130GLNH
$R^0_{1\rho}$=15.04 ±6.07 s^{-1}
140LEUH
$R^0_{1\rho}$=15.08 ±9.26 s^{-1}
141LYSH
$R^0_{1\rho}$=2.4 ±2.6 s^{-1}
142TYRH
τ_{ex}=107.7 ±90.1 µs
S^2=0.999 ±0.0004
143ILEH
τ_{ex}=53.8 ±75.7 µs
S^2=0.999 ±0.0004
144HISH
$R^0_{1\rho}$=3.67 ±3.71 s^{-1}
145SERH
$R^0_{1\rho}$=5.27 ±3.89 s^{-1}
146ALAH
$R^0_{1\rho}$=6.71 ±4.2 s^{-1}
147ASPH
$R^0_{1\rho}$=0.34 ±0.97 s^{-1}
148ILEH
$R^0_{1\rho}$=6.18 ±4.43 s^{-1}
149ILEH
$R^0_{1\rho}$=5.11 ±1.66 s^{-1}
157ASNH
τ_{ex}=132.9 ±89.8 µs
S^2=0.999 ±0.0005
162GLUH
$R^0_{1\rho}$=20.83 ±4.33 s^{-1}
166LEUH
τ_{ex}=165. ±74. µs
S^2=0.999 ±0.0004
167LYSH
$R^0_{1\rho}$=21.65 ±3.61 s^{-1}
168ILEH
τ_{ex}=49.9 ±70.7 µs
S^2=0.999 ±0.0003
182THRH
$R^0_{1\rho}$=4.86 ±5.13 s^{-1}
183GLYH
$R^0_{1\rho}$=9.6 ±5.15 s^{-1}
191ARGH
$R^0_{1\rho}$=10.08 ±6.21 s^{-1}
192ALAH
$R^0_{1\rho}$=5.94 ±2.58 s^{-1}
197LEUH
τ_{ex}=99.5 ±93. µs
S^2=0.999 ±0.0005
198ASNH
τ_{ex}=33.5 ±58.4 µs
S^2=0.999 ±0.0003
199TRPH
τ_{ex}=195.9 ±27.6 µs
S^2=0.999 ±0.0003
200METH
$R^0_{1\rho}$=21.6 ±6.86 s^{-1}
200METH
$R^0_{1\rho}$=17.53 ±7.74 s^{-1}
203ASNH
$R^0_{1\rho}$=4.79 ±5.28 s^{-1}
208ILEH
$R^0_{1\rho}$=3.93 ±2.89 s^{-1}
210SERH
τ_{ex}=177.1 ±61.6 µs
S^2=0.999 ±0.0004
213CYSH
$R^0_{1\rho}$=11.53 ±10.29 s^{-1}
214ILEH
$R^0_{1\rho}$=7.5 ±5.85 s^{-1}
218LEUH
$R^0_{1\rho}$=0.35 ±0.87 s^{-1}
219LEUH
τ_{ex}=103.4 ±92.3 µs
S^2=0.999 ±0.0004
$R_{1\rho}^{on}$ (s^{-1})
$\omega_e/2\pi$ (kHz)


220THRH
$R^0_{1\rho}$=8.93 ±7.44 s^{-1}
222ARGH
τ_{ex}=94.7 ±93. µs
S^2=0.999 ±0.0004
223THRH
$R^0_{1\rho}$=3.18 ±1.71 s^{-1}
227GLYH
$R^0_{1\rho}$=15.63 ±5.35 s^{-1}
228THRH
$R^0_{1\rho}$=6.64 ±4.56 s^{-1}
240LEUH
τ_{ex}=73.3 ±84.8 µs
S^2=0.999 ±0.0003
245GLYH
$R^0_{1\rho}$=25.4 ±6.92 s^{-1}
246ALAH
$R^0_{1\rho}$=15.28 ±2.29 s^{-1}
248LEUH
$R^0_{1\rho}$=12.5 ±3.41 s^{-1}
249LEUH
$R^0_{1\rho}$=6.14 ±3.22 s^{-1}
256SERH
$R^0_{1\rho}$=24.34 ±12.26 s^{-1}
257ALAH
τ_{ex}=199.5 ±1.1 µs
S^2=0.999 ±0.0002
263SERH
$R^0_{1\rho}$=4.25 ±4.43 s^{-1}
272PHEH
$R^0_{1\rho}$=35.47 ±12.02 s^{-1}
273ALAH
$R^0_{1\rho}$=22.55 ±3.58 s^{-1}
274ASNH
$R^0_{1\rho}$=9.07 ±6.87 s^{-1}
276PHEH
τ_{ex}=89.4 ±90.2 µs
S^2=0.999 ±0.0004
277ILEH
$R^0_{1\rho}$=9.93 ±3.08 s^{-1}
278GLYH
$R^0_{1\rho}$=26.42 ±8.36 s^{-1}
279ALAH
τ_{ex}=113.1 ±92.9 µs
S^2=0.999 ±0.0004
280ASNH
τ_{ex}=190.8 ±38.8 µs
S^2=0.997 ±0.0014
282LEUH
τ_{ex}=177.4 ±60.3 µs
S^2=0.999 ±0.0003
283ALAH
τ_{ex}=113.8 ±93.6 µs
S^2=0.999 ±0.0003
284VALH
τ_{ex}=152.6 ±81. µs
S^2=0.999 ±0.0006
285ASPH
$R^0_{1\rho}$=4.42 ±1.09 s^{-1}
286LEUH
$R^0_{1\rho}$=4.8 ±3.64 s^{-1}
296ASPH
$R^0_{1\rho}$=9.39 ±5.75 s^{-1}
297LYSH
$R^0_{1\rho}$=6.62 ±3.38 s^{-1}
300THRH
τ_{ex}=171.6 ±62.5 µs
S^2=0.999 ±0.0003
301ALAH
$R^0_{1\rho}$=2.76 ±2.52 s^{-1}
302ALAH
τ_{ex}=38.8 ±64.6 µs
S^2=0.999 ±0.0003
303GLNH
$R^0_{1\rho}$=1.27 ±1.22 s^{-1}
305LEUH
τ_{ex}=190.2 ±39.7 µs
S^2=0.998 ±0.0004
306ALAH
$R^0_{1\rho}$=3.51 ±2.02 s^{-1}
307HISH
$R^0_{1\rho}$=0.52 ±0.97 s^{-1}
308ALAH
$R^0_{1\rho}$=0.8 ±1.09 s^{-1}
309TYRH
$R^0_{1\rho}$=3.94 ±3.01 s^{-1}
311ALAH
τ_{ex}=172.9 ±65.6 µs
S^2=0.999 ±0.0002
312GLNH
$R^0_{1\rho}$=4.99 ±2.61 s^{-1}
313TYRH
$R^0_{1\rho}$=5.57 ±2.53 s^{-1}
$R_{1\rho}^{on}$ (s^{-1})
$\omega_e/2\pi$ (kHz)

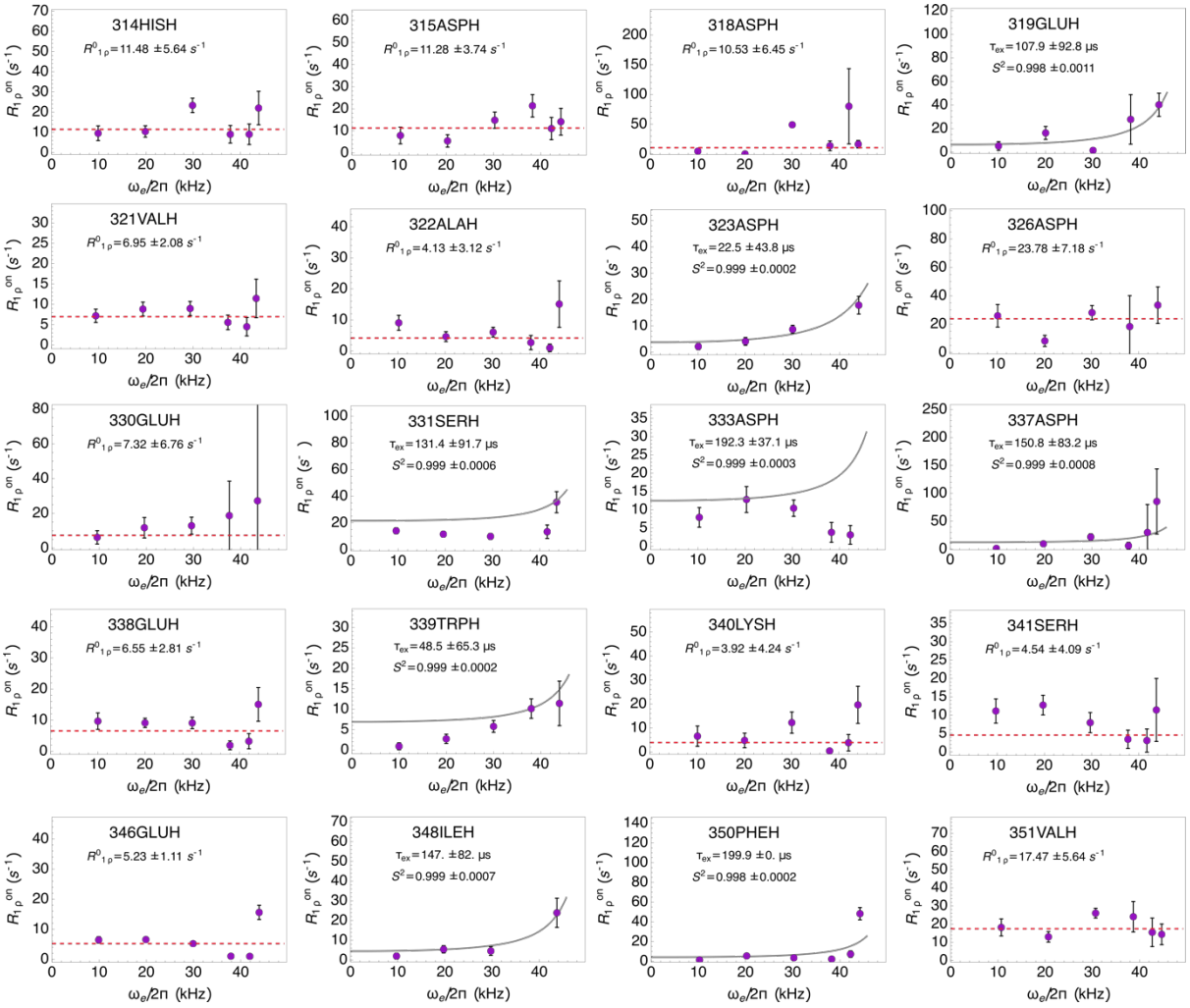

314HISH
$R^0_{1\rho}$=11.48 ±5.64 s^{-1}
315ASPH
$R^0_{1\rho}$=11.28 ±3.74 s^{-1}
318ASPH
$R^0_{1\rho}$=10.53 ±6.45 s^{-1}
319GLUH
τ_{ex}=107.9 ±92.8 µs
S^2=0.998 ±0.0011
321VALH
$R^0_{1\rho}$=6.95 ±2.08 s^{-1}
322ALAH
$R^0_{1\rho}$=4.13 ±3.12 s^{-1}
323ASPH
τ_{ex}=22.5 ±43.8 µs
S^2=0.999 ±0.0002
326ASPH
$R^0_{1\rho}$=23.78 ±7.18 s^{-1}
330GLUH
$R^0_{1\rho}$=7.32 ±6.76 s^{-1}
331SERH
τ_{ex}=131.4 ±91.7 µs
S^2=0.999 ±0.0006
333ASPH
τ_{ex}=192.3 ±37.1 µs
S^2=0.999 ±0.0003
337ASPH
τ_{ex}=150.8 ±83.2 µs
S^2=0.999 ±0.0008
338GLUH
$R^0_{1\rho}$=6.55 ±2.81 s^{-1}
339TRPH
τ_{ex}=48.5 ±65.3 µs
S^2=0.999 ±0.0002
340LYSH
$R^0_{1\rho}$=3.92 ±4.24 s^{-1}
341SERH
$R^0_{1\rho}$=4.54 ±4.09 s^{-1}
346GLUH
$R^0_{1\rho}$=5.23 ±1.11 s^{-1}
348ILEH
τ_{ex}=147. ±82. µs
S^2=0.999 ±0.0007
350PHEH
τ_{ex}=199.9 ±0. µs
S^2=0.998 ±0.0002
351VALH
$R^0_{1\rho}$=17.47 ±5.64 s^{-1}
$R_{1\rho}^{on}$ (s^{-1})
$\omega_e/2\pi$ (kHz)

Figure 11.4. ^{15}N NERRD profiles for SH3 wt.

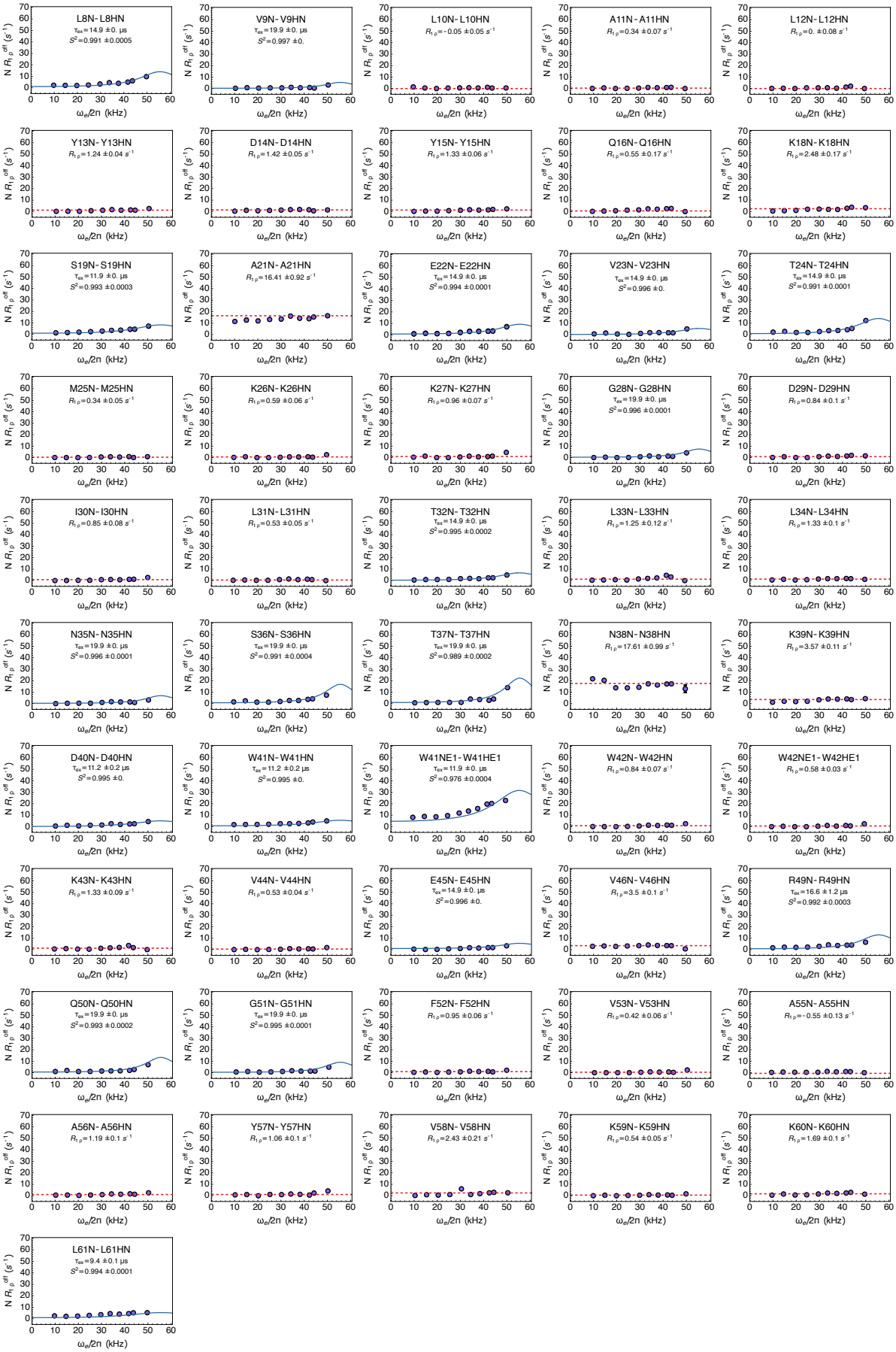


Figure 11.5. ^{15}N NERRD profiles for SH3 R21A

Eidesstattliche Versicherung (Affidavit)

Medina Gomez, Sara
Name, Vorname
(Surname, first name)

216936
Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Dortmund, 01.03.2023
Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Characterization of protein structure and dynamics by solution an solid-state NMR

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Dortmund, 01.03.2023
Ort, Datum
(Place, date)

Unterschrift
(Signature)