Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Development and application of new ^{1}H-detected solid-state NMR methods to expand dimensionality and study protein-ligand interactions

Alexander Klein

aus

Villingen-Schwenningen, Deutschland

2021

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Rasmus Linser betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 02.11.2021

Alexander Klein

.....

Dissertation eingereicht am	10.11.2021
1. Gutachter:	Prof. Dr. Rasmus Linser
2. Gutachter:	Prof. Dr. Henry Dube
Mündliche Prüfung am	21.12.2021

Acknowledgements

The end of my Ph.D. marks the end of a journey in a literal and metaphorical sense, and I could not have completed this journey without the assistance of many persons along the way, to which I want to say 'Thank you!'.

First, I want to thank Prof. Rasmus Linser for being my supervisor and teacher. He always motivated me and is an inspiring researcher that gave me many great ideas, of which some were not as "fast" as initially thought. When projects seemed to be stuck, he often gave the necessary leap forward and with that also fresh motivation. From the first interview, he created an open-minded environment both personally and scientifically that I did not take for granted. He was always willing to spend as much time as possible discussing the latest results or sit together at the spectrometers. I also want to express my gratitude for the many workshops and conferences I was allowed to visit worldwide.

My second thank goes to Prof. Henry Dube for taking the role of being my second reviewer on short notice and unforeseen circumstances and to Prof. Konstantin Karaghiosoff, whom I wish all the best and fast recovery at this time.

I found an excellent research atmosphere and a very welcoming environment in the group from the very beginning. Special thanks go to Dr. Suresh Vasa and Dr. Petra Rovó, who taught me many things about solid-state NMR, pulse programming, spectrometer hardware and were always open to discuss problems at any time. They contributed a lot to the direction this journey took.

I want to thank Dr. Kristof Grohe and Dr. Diana Pöllmann for helping me to take the first steps with protein expression, and Sara Medina Gómez, Julia Kotschy, Dr. Laura Kukuk, and Dr. Himanshu Singh for ongoing assistance in this field.

I also want to thank all the other fellow PhDs and PostDocs in the group or scientific family in Munich and Dortmund: Ekatarina Burakova, Benedikt Söldner, Romeo Dubini, Hanna Kavaleuskaya, and Dr. Elias Akoury, as well as Maria Sergani and Christine Wehbrink for administrative support. It has been a pleasure talking to you and discussing with you, having coffee breaks, and delicious international brunches.

I want to thank my parents for encouraging me all the time beyond any doubt, for helping me move from one city to the other three times, and for guiding me my entire way. I would not have made it that far without their support.

A special thank goes to Sara, who joined the group as a colleague and is now the wonderful partner at my side. She is there in times of doubt and constantly changes how I see myself in the best possible way.

Last but not least, I want to acknowledge all the many other persons that I have encountered along the way, even though not mentioned, which all helped in one way or another to bring this journey to an end.

Contents

Publications

1. **Klein, A.**, Vasa, S. K., Linser, R. Automated projection spectroscopy in solid-state NMR *J. Biomol. NMR* **72**, 163-170, 2018, doi:10.1007/s10858-018-0215-0.

2. E. Burakova, **Klein, A.**, Vasa, S. K., Linser, R. Non-uniform sampling in quantitative assessment of heterogeneous solid-state NMR line shapes *J. Biomol. NMR* **74**, 71–82, 2020, doi: 10.1007/s10858-019-00291-z.

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

4. **Klein, A.**, Rovó, P., Sakhrani, V.V., Wang, Y., Holmes, J., Liu, V., Skowronek, P., Kukuk, L., Vasa, S.K., Güntert, P., Mueller, L.J., Linser, R. 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.* **119**, e2114690119, 2022, doi: 10.1073/pnas.2114690119.

5. **Klein, A.**, Vasa, S.K., Söldner, B., Grohe, K., Linser, R. Unambiguous Side-Chain Assignments for Solid-State NMR Structure Elucidation of Nondeuterated Proteins via a Combined 5D/4D Side-Chain-to-Backbone Experiment. *J. Phys. Chem. Lett.* **13**, 1644–1651, 2022, doi:10.1021/acs.jpclett.1c04075

Summary

^{1}H detected solid-state NMR has emerged as a powerful tool to study all aspects of proteins. Facilitated by significant advances both on the side of sample preparation through partial protonation of exchangeable sites and on the side of hardware with the advent of solid-state probes capable of spinning speeds of 50-100 kHz and above the ^{1}H nucleus becomes more readily accessible as detection nucleus enabling all types of protein studies ranging from resonance assignment to structure calculation as well as the characterization of dynamics. Supported through these improvements, techniques formerly developed for solution-state NMR become increasingly adaptable for solid-state NMR allowing for more sophisticated approaches to study challenging proteins such as those of high molecular weight, which are usually inaccessible for solution-state NMR due to their slow correlation times. The work presented here introduces ^{1}H-detected solid-state NMR experiments with the focus of studying proteins of high molecular weight or interactions between the protein and a ligand. The emphasis lies in developing high-dimensional experiments with more than three dimensions using non-uniform sampling. At the same time, also the concept of projection spectroscopy is adapted from solution-state NMR and analyzed towards its feasibility in solid-state NMR. This work presents 5D experiments for backbone and side-chain assignment and applies them to the SH3-domain of chicken α-spectrin as a model system for method development and additionally uses 5D backbone experiments for the *de novo* assignment of the tryptophan synthase of *Typhimurium salmonella*, a 2x72 kDa protein greatly exceeding the size of proteins commonly studied by (solid-state) NMR. Assignment strategies involving the newly developed experiments are presented and evaluated, both based on manual assignments and in combination with algorithms for automated assignments, which recently experienced an emerging usage. These new approaches facilitate substantial assignments of tryptophan synthase and provide unique insights into its β-subunit active site for the first time directly observed from the proteins without special labeling schemes, through chemical shifts, as direct reports on their chemical environment, as well as first relaxation data acquired on tryptophan synthase. Additionally, the developed 5D experiments present an easy and near-complete resonance assignment of the backbone and side-chain of the SH3-domain proposing an alternative approach over the established strategies for resonance assignment in small proteins. Ultimately, a new pulse sequence aimed at studying a ligand's structure in complex with a protein is adapted from solution to solid-state NMR. First results are presented showing the overall feasibility of such pulse sequences also in solid-state NMR.

Zusammenfasssung

^{1}H detektierte Festkörper Kernmagnetresonanz (NMR) hat sich zu einer leistungsstarken Technik zur Untersuchung von Proteinen entwickelt. Ermöglicht durch große Weiterentwicklungen sowohl im Sinne der Probenpreparation durch partielle Protonierung von austauschbaren Gruppen, als auch auf der Seite der verwendeten Hardware durch die Entwicklung von Festkörperprobenköpfen mit Rotationsfrequenzen von 50-100 kHz und darüber hinaus, ist der ^{1}H Kern zunehmend zur Detektion geeignet und ermöglicht somit eine Bandbreite von Experimenten, die sich von der Resonanzzuordnung bis zur Strukturrechnung und Charakterisierung von Dynamik erstrecken. Unterstüzt durch diese Entwicklungen, werden auch ausgefeiltere Techniken, die ursprünglich für Lösungs-NMR entwickelt wurden, zunehmend anwendbar für Festkröper-NMR und ermöglichen den Zugang zu komplexeren Proteinen, z.B. solchen mit hohem Molekulargewicht, welche aufgrund der langsamen Korrelationszeiten überlicherweise nicht durch Lösungs-NMR betrachtet werden können. Die hier vorgestelle Arbeit präsentiert ^{1}H detektierte Festkörper-NMR Experimente, mit dem Fokus auf die Betrachtung von Proteinen mit hohem Molekulargewicht als auch auf Interaktionen zwischen Ligand und Protein. Der Schwerpunkt der Arbeit liegt in der Entwicklung von hoch dimensionalen Experimenten mit mehr als drei Dimensionen unter der Verwendung von non-uniform sampling. Darüber hinaus wird auch das Konzept der Projektionsspektroskopie von Lösungs-NMR adaptiert und dessen Eignung für Festkörper NMR analysiert. Diese Arbeit stellt 5D Experimente zur Proteinrückgrat- und Seitenkettenzuordnung vor und zeigt deren Anwendung an der SH3-Domäne von Hühner α-Spektrin als Modellsystem zur Methodenentwicklung als auch die Anwendung von 5D Rückgratexperimenten für die *de novo* Resonanzzuordnung der Tryptophansynthase von *Typhimurium salmonella*, einem 2x72 kDa großem Protein, welches die Größe von gängingen mit NMR untersuchten Proteinen weit überschreitet. Es werden sowohl manuelle als auch aufkommende algorithmenunterstützte automatisierte Zuordnungsstrategien unter Verwendung der vorgestellten Experimente entwickelt und analyisiert. Die neu entwickelten Strategien ermöglichen substanzielle Resonanzzuordnung für die Tryptophansynthase und zeigen vorher unzugängliche Einblicke in das aktive Zentrum ihrer β-Untereinheit, die zum ersten Mal von der Seite des Proteins ohne Zuhilfenahme von speziellen Markierungstechniken durch chemische Verschiebungen, als direkte Reporter ihrer chemischen Umgegbung, als auch durch erste aufgenommene Relaxationsexperimente betrachtet werden können. Zusätzlichen ermöglichen die entwickelten 5D Experimente die einfache und fast vollständige Zuordnung von Rückgrat- und Seitenkettensignalen auf der SH3-Domäne und zeigen eine mögliche Alternative zu gängingen Zuordnungsstrategien für kleine Proteine. Zuletzt wird eine neue Pulssequenz zur Betrachtung von Protein-Ligand Interaktionen für Festkörper-NMR adaptiert und erste Ergebnisse gezeigt, die veranschaulichen, dass solche Experimente allgemein auch für Festkörper NMR geeignet sind.

Chapter 1

Introduction

1.1 Physical principles in NMR spectroscopy

Larmor frequency and chemical shift[1,2]

In the early 1920s, Stern and Gerlach demonstrated that the orientation of the spin angular momentum in an external magnetic field is not stochastically distributed but instead quantized and therefore occupies only discrete energy states[3,4]. Despite being initially described for the angular momentum of electrons, the findings can be transferred to atomic nuclei as well. Applying an external magnetic field B_0 to a nuclear spin will orient it along the outer field, usually along the z-axis. The Hamilton's operator representing the energy of this orientation is given as $\hat{H} = -\gamma B_0 \hat{I}_z$ where γ is the gyromagnetic ratio of the spin and $\hat{I}_z$ the z component of the nuclear spin angular momentum. For spin $\frac{1}{2}$ particles associated with spin quantum numbers of $m = \pm\frac{1}{2}$ the B_0 field will cause a splitting of the otherwise equivalent eigenstates of the spin in two states in which the α state represents a parallel and energetically preferred orientation, with an energy $E_\alpha = -\hbar\gamma B_0 m = +\frac{1}{2}\hbar\gamma B_0$, while the β state corresponds to an anti-parallel and therefore less favored orientation of the spin, with an energy of $E_\beta = -\hbar\gamma B_0 m = -\frac{1}{2}\hbar\gamma B_0$. Therefore, the energy difference $\Delta E = \hbar\gamma B_0$ depends only on the applied B_0 field. The energy difference is referred to as Larmor frequency, which is defined as

$$\begin{aligned} \nu_0 &= -\frac{\gamma B_0}{2\pi} \\ \omega_0 &= -\gamma B_0 \end{aligned} \tag{1.1}$$

The Larmor frequency of nuclear spins depends not only on the B_0 field but also strongly on the surrounding electron environment, which in most cases causes a magnetic field that is directed against the outer B_0 field. Therefore, a high electron density, as found, e.g., in aliphatic groups, shields the nucleus efficiently from the outer magnetic field and creates a weaker local field, resulting in a reduced Larmor frequency. In contrast, low electron densities, e.g., in polar bonds, have the opposite effect, i.e., the nuclear spin is less shielded and experiences a higher B_0 field, resulting in a higher Larmor frequency and vice versa. This phenomenon is referred to as the chemical shift, and the effective Larmor frequency is then given by

$$\omega_{eff} = -\gamma(1-\sigma)B_0 \tag{1.2}$$

with the average isotropic shielding constant

$$\sigma = \frac{\sigma_{xx} + \sigma_{yy} + \sigma_{zz}}{3} \tag{1.3}$$

and the principal values of the chemical shift tensor σ_{xx}, σ_{yy} and σ_{zz} along their respective axis. Whereas averaged due to fast tumbling in solution, the chemical shift also contains anisotropic components: chemical shift anisotropy (CSA)

$$\Delta\sigma = \sigma_{xx} - \frac{\sigma_{yy} + \sigma_{zz}}{2} \tag{1.4}$$

and asymmetry

$$\eta = \frac{3\left(\sigma_{yy} + \sigma_{zz}\right)}{2\Delta\sigma} \tag{1.5}$$

To make the chemical shift comparable amongst different B_0 fields, it is usually given in ppm as in equation (1.6).

$$\delta(\text{ppm}) = \frac{\nu - \nu_{ref}}{\nu_{ref}} \cdot 10^6 \tag{1.6}$$

For practical reasons, the reference chemical shift $\delta = 0$ is not defined relative to a pure spin without any electrons but rather a reference compounds like Tetramethylsilane (TMS) or 2,2-Dimethyl-2-silapentane-5-sulfonate (DSS).

Quantum mechanics and product operators in NMR[1,2]

In quantum mechanics, the overall wavefunction of a system can be expressed as a linear combination of the individual eigenfunctions representing the system, i.e.

$$\psi = \sum_{n=1}^{n} c_n \psi_n \tag{1.7}$$

In the case of an isolated spin only interacting with the B_0 field, equation (1.7) can be written as

$$\begin{aligned} \psi &= c_\alpha \psi_\alpha + c_\beta \psi_\beta \\ |\psi\rangle &= c_\alpha |\alpha\rangle + c_\beta |\beta\rangle \end{aligned} \tag{1.8}$$

Using this formalism, the expectation values, i.e., average, for the x, y, and z-magnetization, are given as

$$\left\langle \hat{I}_z \right\rangle = \frac{1}{2} c_\alpha^* c_\alpha - \frac{1}{2} c_\beta^* c_\beta \tag{1.9}$$

$$\left\langle \hat{I}_x \right\rangle = \frac{1}{2} c_\alpha^* c_\beta - \frac{1}{2} c_\beta^* c_\alpha \tag{1.10}$$

$$\left\langle \hat{I}_y \right\rangle = \frac{1}{2} i c_\beta^* c_\alpha - \frac{1}{2} i c_\alpha^* c_\beta \tag{1.11}$$

Free precession In contrast to the time-independent bulk magnetization

$$M_z = \gamma N \overline{\left\langle \hat{I}_z \right\rangle} = \frac{1}{2} \gamma N \overline{\left(c_\alpha^* c_\alpha - c_\beta^* c_\beta \right)} = \frac{1}{2} \gamma \left(n_\alpha - n_\beta \right) \tag{1.12}$$

described by the time-independent Schrödinger equation, other essential processes in NMR like free precession or RF pulses are, however, time-dependent processes. Therefore the time-dependent Schördinger equation

$$\frac{d\,|\psi(t)\rangle}{dt} = -i\hat{H}\,|\psi(t)\rangle \tag{1.13}$$

has to be used. Free precession is assumed to happen around the z-axis, and in the rotating frame, the Hamiltonian is given as

$$\hat{H} = \Omega \hat{I}_z \tag{1.14}$$

, where Ω is the frequency offset relative to the Larmor frequency. Using (1.13) in its time-dependent variant, the solution of (1.14) is

$$\frac{dc_\alpha(t)}{dt} = -i\frac{1}{2}\Omega c_\alpha(t) \tag{1.15}$$

$$\frac{dc_\beta(t)}{dt} = i\frac{1}{2}\Omega c_\beta(t) \tag{1.16}$$

The solutions of those differential equations are given as

$$c_\alpha(t) = c_\alpha(0)e^{-i\frac{1}{2}\Omega t} \qquad \text{and} \qquad c_\beta(t) = c_\beta(0)e^{i\frac{1}{2}\Omega t} \tag{1.17}$$

, which describes a phase oscillation that depends on the eigenvalue $\pm\frac{1}{2}\Omega$ of the Hamiltonian and for α is clockwise while for β is counterclockwise. The effect of this trajectory on the z-component can be calculated as

$$\langle I_z\rangle(t) = \frac{1}{2}c_\alpha^*(0)c_\alpha(0) - \frac{1}{2}c_\beta^*(0)c_\beta(0) \tag{1.18}$$

and is therefore time-independent, whereas the x and y components are expressed by

$$\begin{aligned}\langle I_x\rangle(t) &= \frac{1}{2}c_\alpha^*(0)c_\beta(0)e^{i\Omega t} + \frac{1}{2}c_\beta^*(0)c_\alpha(0)e^{-i\Omega t} \\ &= \cos\Omega t\,\langle I_x\rangle(0) - \sin\Omega t\,\langle I_y\rangle(0)\end{aligned} \tag{1.19}$$

$$\begin{aligned}\langle I_y\rangle(t) &= \frac{1}{2}c_\beta^*(0)c_\alpha(0)e^{-i\Omega t} + \frac{1}{2}c_\alpha^*(0)c_\beta(0)e^{i\Omega t} \\ &= \cos\Omega t\,\langle I_y\rangle(0) + \sin\Omega t\,\langle I_x\rangle(0)\end{aligned} \tag{1.20}$$

This means the free evolution periods will solely affect x and y magnetization, whereas z-magnetization remains unaltered during those episodes.

Pulses A radiofrequency pulse that is applied on resonance, i.e., with the Larmor frequency, makes the outer B_0 apparently disappear in the rotating frame, and the only field that is experienced by the spins is the B_1 field of the RF wave with the Hamiltonian

$$\hat{H}_{RF} = \omega_1\hat{I}_x \tag{1.21}$$

The solutions of the time-dependent Schrödinger equation for this operator that describe the time dependency of the individual coefficients are

$$\frac{dc_\alpha(t)}{dt} = -i\frac{1}{2}\omega_1 c_\beta(t) \qquad c_\alpha(t) = \cos\left(\frac{1}{2}\omega_1 t\right)c_\alpha(0) - i\sin\left(\frac{1}{2}\omega_1 t\right)c_\beta(0) \tag{1.22}$$

$$\frac{dc_\beta(t)}{dt} = -i\frac{1}{2}\omega_1 c_\alpha(t) \qquad c_\beta(t) = \cos\left(\frac{1}{2}\omega_1 t\right)c_\beta(0) - i\sin\left(\frac{1}{2}\omega_1 t\right)c_\alpha(0) \tag{1.23}$$

These equations can be plugged into equations (1.10)-(1.11) to describe the single spin angular momentum components and result in

$$\langle I_z\rangle = \langle I_z\rangle(0)\cos\omega_1 t + \langle I_y\rangle(0)\sin\omega_1 t \tag{1.24}$$

$$\langle I_x\rangle = \langle I_x\rangle(0) \tag{1.25}$$

$$\langle I_y \rangle = \langle I_y \rangle (0) \cos \omega_1 t - \langle I_z \rangle (0) \sin \omega_1 t \tag{1.26}$$

As can be seen, a pulse along the x-axis does not alter the x component while the z and y components are in an oscillatory dependency. The same situation would be obtained while pulsing along the y-axis so that the y component remains unchanged while the z and x components are periodically interchanged. A 90° pulse, i.e., $\omega_1 t = \frac{\pi}{2}$ would therefore cause $\langle I_y \rangle = -\langle I_z \rangle (0)$. Considering a higher population of $|\alpha\rangle$ states in the bulk sample, that means that the entire z-magnetization would be turned into magnetization along the y-axis.

The density operator and product operators The density operator $\hat{\rho}$ is defined as the ensemble average of all combinations of coefficients.

$$\hat{\rho} = \overline{|\psi\rangle\langle\psi|} = \begin{pmatrix} \overline{\langle\alpha|\hat{\rho}|\alpha\rangle} & \overline{\langle\alpha|\hat{\rho}|\beta\rangle} \\ \overline{\langle\beta|\hat{\rho}|\alpha\rangle} & \overline{\langle\beta|\hat{\rho}|\beta\rangle} \end{pmatrix} = \begin{pmatrix} \overline{c_\alpha^* c_\alpha} & \overline{c_\beta^* c_\alpha} \\ \overline{c_\alpha^* c_\beta} & \overline{c_\beta^* c_\beta} \end{pmatrix} \equiv \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{21} \end{pmatrix} \tag{1.27}$$

and provides an easier way to compute the outcome of the problems above as no expectation values have to be calculated explicitly but are instead already contained in ρ such that, e.g., the bulk magnetization along the z-axis can be expressed as

$$M_z = \frac{1}{2} \gamma N (\rho_{11} - \rho_{22}) \tag{1.28}$$

$$\tag{1.29}$$

The Liouville-von Neumann equation ((1.30)) describes the time-dependent evolution of the density operator,

$$\frac{d\hat{\rho}t}{d(t)} = -i \left(\hat{H}\hat{\rho}(t) - \hat{\rho}(t)\hat{H} \right) \tag{1.30}$$

and has the general solution

$$\hat{\rho}(t) = e^{-i\hat{H}t} \hat{\rho}(0) e^{i\hat{H}t} \tag{1.31}$$

It is therefore used to describe all time-dependent effects using a matching Hamiltonian.

If the density operator is expressed as a linear combination of the basic operators $\hat{I}_z$, $\hat{I}_x$ and $\hat{I}_y$ in their matrix representation and the identity matrix E, every component can be analyzed independently

$$\begin{aligned} \hat{\rho} &= a_E \hat{E} + a_x \hat{I}_x + a_y \hat{I}_y + a_z \hat{I}_z \\ &= a_E \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + a_x \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} + a_y \begin{pmatrix} 0 & -i\frac{1}{2} \\ i\frac{1}{2} & 0 \end{pmatrix} + a_z \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 2a_E + a_z & a_x - ia_y \\ a_x + ia_y & 2a_e - a_z \end{pmatrix} \equiv \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{21} \end{pmatrix} \end{aligned} \tag{1.32}$$

Free precession of x-magnetization around the z-axis can therefore be expressed as

$$\begin{aligned} \hat{\rho}(t) &= e^{-i\Omega\hat{I}_z t} \hat{I}_x e^{i\Omega\hat{I}_z t} \\ &= \cos \Omega t \hat{I}_x + \sin \Omega t \hat{I}_y \end{aligned}$$

using $\hat{\rho}(0) = \hat{I}_x$ as the initial density operator and the Hamiltonian for free precession $\hat{H} = \Omega\hat{I}_z$. As equation (1.19) this shows, that the x-component interchanges into y-magnetization over time. Similar, a pulse around the x-axis is obtained as

$$\hat{\rho}(t_p) = e^{-i\omega_1\hat{I}_x t_p}\hat{I}_z e^{i\omega_1\hat{I}_x t_p} \tag{1.33}$$

$$= \cos(\omega_1 t_p)\hat{I}_z - \sin(\omega_1 t_p)\hat{I}_y \tag{1.34}$$

now using $\hat{\rho}(0) = \hat{I}_z$ as the initial density operator and the Hamiltonian for a pulse $\hat{H} = \omega_1\hat{I}_x$. This gives the same results as in equation (1.24). If $\omega_1 t_p = \beta = \frac{\pi}{2}$ then the pulse creates full $-\hat{I}_y$ magnetization. For simplicity, the product operator formalism can be used to describe these transformations by

$$\cos(\theta) \times \text{original operator} + \sin(\theta) \times \text{new operator} \tag{1.35}$$

***J*-coupling and INEPT** An important isotropic interaction between two spins $\hat{I}_1$ and $\hat{I}_2$ is the J-coupling or scalar coupling. Scalar coupling only acts during free precession periods and interconverts in-phase and anti-phase magnetization.

$$\hat{H}_J = 2\pi J_{12}\hat{I}_{1z}\hat{I}_{2z}$$

$$\hat{I}_y \xrightarrow{2\pi J_{12}\hat{I}_{1z}\hat{I}_{2z}\tau} \cos(\pi J_{12}\tau)\,\hat{I}_x + \sin(\pi J_{12}\tau)\,2\hat{I}_{1x}\hat{I}_{2z} \tag{1.36}$$

$$2\hat{I}_{1x}\hat{I}_{2z} \xrightarrow{2\pi J_{12}\hat{I}_{1z}\hat{I}_{2z}\tau} \cos(\pi J_{12}\tau)\,2\hat{I}_{1x}\hat{I}_{2z} + \sin(\pi J_{12}\tau)\,\hat{I}_y \tag{1.37}$$

As chemical shift evolution would also occur during free precession periods, J-couplings are usually evolved in a spin-echo sequence consisting of two symmetric delays, τ and a π pulse in the middle to refocus the chemical shift. The outcomes of such echo sequences are summarized in table 1.1, as found in common textbooks[1]. Using these elements of pulses, chemical

Table 1.1: Interconversion of individual terms using the product operators formalism as found in common textbooks[1].

	end state	
intial state	$\times\cos(2\pi J_{12}\tau)$	$\times\sin(2\pi J_{12}\tau)$
$\hat{I}_x$	$\hat{I}_x$	$2\hat{I}_{1y}\hat{I}_{2z}$
$\hat{I}_y$	$-\hat{I}_y$	$2\hat{I}_{1x}\hat{I}_{2z}$
$2\hat{I}_{1x}\hat{I}_{2z}$	$-2\hat{I}_{1x}\hat{I}_{2z}$	$-\hat{I}_y$
$2\hat{I}_{1y}\hat{I}_{2z}$	$2\hat{I}_{1y}\hat{I}_{2z}$	$-\hat{I}_x$

shift evolution, and J-coupling, the magnetization transfers in (solution-state NMR) can be sufficiently described, however, without the influence of relaxation on the system. The most crucial building block built from those is the INEPT (insensitive nuclei enhanced polarization transfer)[5] which is usually used to transfer the high polarization of the ^{1}H nucleus to less sensitive nuclei such as ^{13}C and ^{15}N and back. Built from a combination of INEPTs, are essential NMR experiments such as the HSQC[6] and early triple-resonance experiments as reviewed by Sattler et al.[7].

1.2 Solid-state NMR

Like liquid-state NMR, solid-state NMR has numerous applications ranging over various fields like inorganic solids to protein research. At the same time, also the nature of the sample can vary from single-crystal over gel-like compounds to micro-crystalline powders. In this work, only micro-crystalline protein samples were of interest, and the following section will serve as a short introduction to the most important theoretical aspects of solid-state NMR associated with that.

Magic angle spinning (MAS)[8,9]

Dipolar interactions are the dominant non-isotropic interactions in solid-state NMR with an interaction strength of several kHz, compared to Hz of, e.g., J-couplings, which in contrast to solution-state NMR are not averaged as molecular tumbling is absent. Dipolar coupling is the driving force behind increased line widths in solid-state NMR but also the basis for essential techniques. The Hamiltonian for the dipolar coupling is defined as

$$\hat{H}_{dd} = -\left(\frac{\mu_0}{4\pi}\right)\gamma_I\gamma_S\hbar\left(\frac{\hat{I}\hat{S}}{r^3} - 3\frac{(\hat{I}\cdot\vec{r})(\hat{S}\cdot\vec{r})}{r^5}\right) \tag{1.38}$$

, which simplifies to

$$\hat{H}_{dd} = -d\cdot\frac{1}{2}\left(3\cos^2\theta - 1)\right)\left[3\hat{I}_z\hat{S}_z - \hat{I}\cdot\hat{S}\right] \tag{1.39}$$

for homonuclear interactions under the high-field approximation, i.e., $d \ll \omega_0$, and when expressed in spherical coordinates with the dipolar coupling constant

$$d = \hbar\left(\frac{\mu_0}{4\pi}\right)\frac{1}{r^3}\gamma_I\gamma_S \tag{1.40}$$

For heteronuclear interactions, the interaction is defined by

$$\hat{H}_{dd} = -d\cdot\frac{1}{2}\left(3\cos^2\theta - 1\right)\hat{I}_z\hat{S}_z \tag{1.41}$$

Solid matter can be exposed to a strong magnetic field, i.e., an NMR spectrometer, in a static fashion in which all spins, except molecular motion like vibrations, etc., maintain a single relative orientation towards the B_0 field. However, the concepts presented in this work rely entirely on magic angle spinning (MAS) techniques, in which the spins move through all possible orientations relative to the magnetic field. MAS was initially introduced in the late 1950s independently by Andrew et al.[10] and Lowe[11] as a technique for averaging non-isotropic, i.e., orientation-dependent, spin-spin interactions like chemical shift anisotropy (CSA) and dipole-dipole interactions. As shown in figure 1.1, the sample, contained in a rotor, is spun around an axis tilted by $\theta = 54.74°$ away from the z-axis, i.e., the B_0 field. All orientations θ_m of the interaction tensors are found in a powder sample as they are fixed to the molecular frame rather than the rotor frame. Upon rotation, θ_m varies, and all possible orientations are taken at least once. Therefore,

$$\langle\cos^2\theta_m - 1\rangle = \frac{1}{2}\left(3\cos^2\theta - 1\right)\left(3\cos^2\beta - 1\right) = 0,\ \theta = 54.74° \tag{1.42}$$

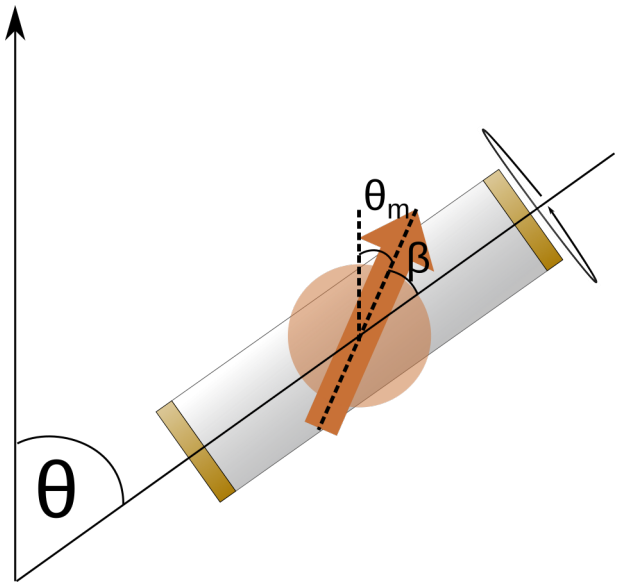


Figure 1.1: For MAS, a rotor is oriented along an axis tiled by 54.74° away from the z-axis. During one rotation, the angle θ_m, the orientation of the interaction tensor towards the B_0 field, undergoes all variations while the angle β remains constant. For each tensor and the entire powder sample, the anisotropic interactions are averaged to 0.

With β being the angle between the rotating axis and the orientation of the interaction tensor. If the spinning frequency is on the order of the interaction strength (in Hz), the anisotropic interactions are averaged to 0. As CSA interactions are comparably small, slower spinning speeds of around 5 kHz can be sufficient to achieve good results, while the averaging of homonuclear dipole-dipole interactions, especially ^{1}H-^{1}H interactions, require the fastest probeheads available by now with spinning speeds up to 110 kHz or above[12,13].

Magnetization transfer in solid-state NMR

The strength of ^{1}H-^{1}H dipole-dipole interactions renders long J-coupling-based magnetization transfer mechanisms inefficient as they are greatly dominated by the non-isotropic interactions and suffer from severe relaxation losses, especially at low spinning speeds, which nowadays, however, do not present significant hurdles anymore. As the ^{1}H nucleus was historically not suitable as a detection nucleus due to limited resolution[14] but offered significant sensitivity enhancements, the large magnetization source of the protons is transferred to a heteronucleus via cross-polarization (CP) or Hartmann-Hahn transfer[14,15], rather than INEPT[5]. Due to the different gyromagnetic ratios γ and the different Zeeman splittings, the ^{1}H nucleus and a heteronucleus do not exchange polarization. However, when applying a spin-lock field on both nuclei which matches the Harmann-Hahn condition,

$$\omega_H = \gamma_H B_{1H} = \gamma_X B_{1X} = \omega_X \tag{1.43}$$

the splittings become equal as the spins experience only the B_1 field. If before the CP a $\pi/2$ pulse is applied on the protons to create net magnetization, the spin-locked system can equilibrate, i.e., polarization is transferred to heteronucleus with no net magnetization in the rotating frame, as now the exchange is energy conserving over the dipolar coupling between the ^{1}H and heteronucleus[9,14,15]. The maximum enhancement gained on the heteronucleus

is therefore depending on the gyromagnetic ratio of the two nuclei. If magnetization is transferred to ^{13}C, a factor of 4 can be gained, whereas transfer to ^{15}N results in a 10fold enhancement in polarization. Under MAS, however, the dipolar coupling is not constant but modulated in time at $\pm\omega_r$ and $\pm 2\omega_r$ and, more importantly, is significantly attenuated. Consequently, the CP efficiency at the condition mentioned above is strongly reduced but can be restored to a large extent under conditions

$$\omega_H - \omega_X = \pm n\omega_r \tag{1.44}$$

$$\omega_H + \omega_X = \pm n\omega_r \tag{1.45}$$

referred to as zero-quantum (ZQ) or double quantum (DQ) conditions[16,17]. To account for slight differences in the Hartmann-Hahn condition during MAS, various amplitude modulated pulses have been developed that enhance CP efficiency, especially at higher spinning speeds[18–20]. Further development of CP for selective transfers from N→Cα or N→CO rather than broadband transfer is SPECIFIC-CP (spectrally induced filtering in combination with cross-polarization)[21] which introduces a chemical shift dependence in the CP condition such that an energy-conserving interaction can only occur between particular spins. A similar concept for selective transfers from H→Cα or H→CO was introduced by Laage et al.[22]. Here, the spin-locks are chosen so that the RF field on the carbon nuclei is comparably weak and therefore more selective than a high-power spin-lock. As magnetization transfer in MAS solid-state NMR usually relies on reintroducing dipole-dipole interactions, a large variety of recoupling techniques were developed for homonuclear and heteronuclear recoupling. Among the most important besides CP are REDOR[23] for accurate measurements of heteronuclear dipolar coupling constants and magnetization transfer, RFDR[24,25] for homonuclear recoupling, and recently BSH-CP[26,27] for improved CαCO transfers over DREAM[28,29] at high spinning speeds.

1.2.1 ^{1}H-detected solid-state NMR spectroscopy

Despite the severe limitations put by the ^{1}H nucleus, its superior properties in terms of sensitivity and access to additional chemical shifts have fostered the development of a broad range of techniques both on sample preparation and hardware with the goal of constantly improved ^{1}H detection in solid-state NMR. Early approaches aimed at strong dilution of the ^{1}H content, and therefore, dipole-dipole interactions through deuteration of exchangeable sites[**Chevelkov2003**, 12,30–32], as well as constantly increasing the MAS frequencies[12,13,33–36]. Both developments lead to significantly improved ^{1}H line widths and further improve the signal-to-noise ratio (SNR) gained by ^{1}H detection beyond

$$\left(\frac{\gamma_H}{\gamma_X}\right)^{\frac{3}{2}} \tag{1.46}$$

as it strongly depends on the resulting line width as shown by Zilm and co-workers for the enhancement in a H/N correlation[30]

$$\frac{SNR_H}{SNR_N} = 16.2 \cdot \sqrt{\frac{(\nu_{\frac{1}{2}})_N}{(\nu_{\frac{1}{2}})_H}} \tag{1.47}$$

Whereas high dilution of the ^{1}H-^{1}H network is required for low spinning speeds from as low as 10 % at around 14 kHz to yield narrow line widths around 20 Hz[32,37,38], the increase in MAS speeds enables higher protonation ratios at the exchangeable sites of 30 % at 24 kHz[12,39–41] and 100 % at 60 kHz with maintained narrow proton line width[12,42–44]. However, at these spinning speeds, side-chain information can merely be achieved over carbon spins through mixing schemes such as MOCCA[45,46] and TOBSY[47] or is restricted to partial protonation via random adjoint protonation[43], methyl labeling[48] and fractional deuteration[49]. The advent of solid-state MAS probes capable of spinning speeds beyond 60 kHz nowadays facilitates the use of fully protonated systems with ^{1}H line width around 80 Hz[50–52] and respective lower line widths in deuterated and back-exchanges sample preparations[52,53]. The increase of spinning speeds implies a reduction of rotor diameter, e.g., 1.3 mm for 60 kHz probes and 0.7 mm for 111 kHz probes, and therefore a reduction of sample volume. While the reduced sensitivity of smaller rotors is mainly compensated by the probe design[12,52], proteins with low expression yields can readily be accessed due to the significantly reduced sample requirements in fast MAS probes[12,52,54]. The improved averaging of non-isotropic interactions at increased MAS speeds leads to equally increased coherence lifetimes in all nuclei[42,55] affecting the line width and simultaneously enables INEPT transfers[38,41,56,57] rather than CP or any combination of those[44,58] to improve sensitivity further[58] and give access to flexible regions[38,56,57]. On the other hand, the improved averaging of dipole-dipole interactions lead to a decrease in efficiency for many recoupling techniques or alter their mode of action as rotor periods become short. This is found, e.g., for DREAM[28,29], RFDR[24,25,59] and symmetry-based sequences[60], and therefore require the implementation of new or newly combined approaches such as BSH-CP[26,27]. Nevertheless, with the overall improved performance in combination with a third nucleus providing information, solid-state NMR is now capable of adapting more techniques from solution-state NMR. Solution-like backbone experiments like hCANH, hCONH, hCAcoNH, hCOcaNH, hcaCBcaNH, and hcaCBcacoNH[38,41,44,57,58,61], as well as amide-to-amide HNCOCANH type experiments[62,63], are routinely used in the assignment of proteins of increasing size[64,65] and, as transfer efficiencies in solids are independent of molecular size, are more frequently expanded to higher dimensions[62–64,66–70] where they prospectively outperform solution-state NMR experiments[66,67]. Whereas in deuterated samples, side-chain information can only be obtained via the carbon nuclei as described above, solution-like experiments such as HCcH and hCCH and the sequences presented in this work can be employed to obtain side-chain ^{1}H as well as ^{13}C chemical shifts[51]. Along with experiments for through-bond correlations, those for through-space correlations providing NOESY-like ^{1}H-^{1}H contacts have been developed, primarily as HSQC-NOESY derivatives in three[71,72] and four dimensions[40,48,73,74] with increasing accuracy and multi-state modelling adapted from solution-state NMR[72]. Most recently, Hα nuclei have been facilitated as detection nuclei in fully protonated samples alone[51,75] and in combination with amide protons[65] to further expand the chemical shift information that can be obtained from proteins and exploiting their narrow line width.

1.3 Basic concepts of relaxation in solution and solid-state NMR[1,2]

The laws of thermodynamics predict that a system moved away from its equilibrium state will ultimately return to its equilibrium. The application of pulses in NMR spectroscopy poses such a disruption as the population states of α and β states are altered, and two types of relaxation are distinguished. The first type of relaxation is longitudinal relaxation. It describes the return of the spins to the $+z$-axis after a pulse, and the second is transversal relaxation which describes the loss of coherence in the transverse plane. Longitudinal relaxation and the non-secular part of transverse relaxation are driven by the presence of varying local B fields, which affect the x,y, and z-components while in addition, minor differences in the local field of every spin cause different frequencies or chemical shifts which lead to dephasing of the transverse magnetization. Derived from the Bloch equations, both types of relaxation are given as

$$\frac{dM_z(t)}{dt} = -\frac{M_z(t) - M_z^0}{T_1} \Rightarrow M_z(t) = \left(M_z(0) - M_z^0\right) e^{-\frac{t}{T_1}} + M_z^0 \tag{1.48}$$

$$\frac{dM_x(t)}{dt} = -\frac{M_x(t)}{T_2} \Rightarrow M_x(t) = M_x(0) e^{\frac{-t}{T_2}} \tag{1.49}$$

with T_1 as the longitudinal relaxation time and T_2 as the transverse relaxation time. The local fields leading to relaxation are mainly caused by dipole-dipole interactions and chemical shift anisotropy (CSA) and have to bear similar energy as the spins, i.e., they have to oscillate at the same frequency, which is the Larmor frequency. This oscillation is provided through rotational motion, and the amount of rotational motion at matching frequencies contributing to relaxation is given by the correlation function $G(\tau)$ or its FT, the spectral density function $J(\omega)$ or the reduced spectral density function $j(\omega)$, which ignores the amplitude of the local fields.

$$J(\omega) = \frac{2}{5} \frac{\tau_c}{1 + \omega^2 \tau_c^2} \tag{1.50}$$

, with τ_c as the correlation time. Based on this, two motion regimes can be defined. The slow-motion regime in which $\omega^2 \tau_c^2 \gg 1$ and equation (1.50) becomes

$$J(\omega) = \frac{2}{5} \frac{\tau_c}{\omega^2 \tau_c^2} \tag{1.51}$$

and the fast motion regime for which $\omega^2 \tau_c^2 \ll 1$ and therefore

$$J(\omega) = \frac{2}{5} \tau_c = J(0) \tag{1.52}$$

Knowing the different spectral density functions contributing to different relaxation rates can show to which type of motion, i.e., time regime, these relaxation rates are sensitive. In solution-state NMR, e.g., the R_1 rate, $R_1 = \frac{1}{T_1}$ is given as

$$R_1 = \left(\frac{\delta_{DD}}{4}\right)^2 \Big(3J(\omega_I) + 6J(\omega_I + \omega_S) + (\omega_I - \omega_S)\Big) + \delta_{CSA} J(\omega_S) \tag{1.53}$$

Here $\delta_{DD} = -2\frac{\mu_0}{4\pi}\frac{\gamma_I\gamma_S\hbar}{r_{IS}^3}$ and $\delta_{CSA} = (\gamma_I B_0 \Delta\sigma)^2/3$. As the R_1 rate is dominated by the Larmor frequency of the nuclei, it is mostly sensitive to ns time-scale motion. In solid-state NMR under MAS, additional modulations due to spinning are contributing. However, R_1 rates in a heteronuclear two-spin system such as HN, are given similar to the rates in solution-state NMR, and are given as[76]

$$R_1 = \left(\frac{\delta_{DD}}{4}\right)^2 (J(\omega_I - \omega_s) + 3J(\omega_s) + 6J(\omega_I + \omega_s)) + \delta_{CSA}^2 \frac{3}{4} J(\omega_S) \tag{1.54}$$

with the dipolar coupling constant $\delta_{DD} = -2\frac{\mu_0}{4\pi}\frac{\gamma_I\gamma_S\hbar}{r_{IS}^3}$, the anisotropy of the chemical shift tensor $\delta_{CSA} = -\gamma B_0(\sigma_{zz} - \sigma_{iso})$ and the spectral density function, here defined as $J(\omega) = \frac{2}{5}(1 - S^2)\frac{\tau_c}{1+(\omega\tau_c)^2}$ with the order parameter S. Here, only a dependency from the Larmor frequencies is given, and therefore, R_1 rates are mainly sensitive to motions in the ns regime. In contrast, R_2 rates are given as[76]

$$\begin{aligned} R_2 = \left(\frac{\delta_{DD}}{4}\right)^2 \Big(&\frac{4}{3}J(\omega_r) + \frac{2}{3}J(2\omega_r) + \frac{1}{2}J(\omega_I - \omega_S) \\ &+ 3J(\omega_I) + \frac{3}{2}J(\omega_S) + 3J(\omega_I + \omega_S)\Big) \\ &+ \delta_{CSA}^2\Big(\frac{1}{4}J(\omega_r) + \frac{1}{8}J(2\omega_r) + \frac{3}{8}J(\omega_S)\Big) \end{aligned} \tag{1.55}$$

With the spinning speed ω_r. Simulations show that R_2 in solids is sensitive to ns to ms motion but is most sensitive in the µs regime[76]. Depending on the spinning speed the µs time-scale can be sampled from somewhat faster µs for faster MAS frequencies and slower µs motion for slower MAS rates[76]. In this work, conformational changes are the most relevant observed dynamics, which occur primarily in the µs time regime, which can be particularly well studied using solid-state NMR. To suppress the influence of coherent contributions to transverse relaxation rates, often $R_1\rho$ rates are measured, which spin-lock the magnetization in the transverse plane. Using RF fields of varying strength and length, relaxation dispersion (RD) profiles can be obtained to give insights into the present motion regimes[76,77]. For lower RF fields, so-called Bloch-McConnell RD (BMRD) is obtained, which is sensitive to µs-ms time-scales, while additionally in solid-state NMR also near-rotary resonance RD (NERRD) can be observed for stronger RF fields approaching the MAS frequency[12,76–79]. In solid-state NMR, the $R_1\rho$ rates are defined as[76,77]

$$\begin{aligned} R_r\rho &= R_1 + \sin^2\beta_{eff}\left(R_{1\Delta} - \frac{1}{2}R_1\right) \\ R_{1\Delta} &= \left(\frac{\delta_{DD}}{4}\right)^2 \Big(3J(\omega_s) + \frac{1}{3}J(\omega_{eff} - 2\omega_r) \\ &\quad + \frac{2}{3}J(\omega_{eff} - \omega_r) + \frac{2}{3}J(\omega_{eff} + \omega_r) + \frac{1}{3}J(\omega_{eff} + 2\omega_r)\Big) \\ &\quad + \delta_{CSA}^2\Big(\frac{1}{3}J(\omega_{eff} - 2\omega_r) + \frac{2}{3}J(\omega_{eff} - \omega_r) + \frac{2}{3}J(\omega_{eff} + \omega_r) + \frac{1}{3}J(\omega_{eff} + 2\omega_r)\Big) \end{aligned} \tag{1.56}$$

where β_{eff} is the angle of the applied RF field with the resulting strength ω_{eff}. As $R_1\rho$ rates depend on ω_{eff}, as well as the ω_r, the adjustment of those allows to sample the motional regime on a fine scale from a few to hundreds of µs[76,77].

1.4 Concepts of non-uniform sampling

An in-depth theoretical description of general non-uniform sampling (NUS) concepts as well as the various reconstruction algorithms is beyond the scope of this work and can be found in various review papers[80–83] as well in NUS related work by the groups of Gerhard Wagner at Harvard Medical School, USA, Wiktor Kozminski at the University of Warsaw, Poland, Vladislav Orekhov at the University of Gothenburg, Sweden and many others. Nevertheless, a general overview of relevant aspects such as the sampling density and sampling schedules and an overview of the NUS methods used in this work should be given.

Multi-dimensional NMR experiments usually consist of equally spaced 1D experiments, i.e., FIDs, along the indirect dimensions. Incrementation of free evolution periods leads to an oscillating, time-dependent alteration of the signal amplitude detected in the direct, i.e., acquisition dimension, and step-wise Fourier-transformation (FT) in all dimensions yields an n-dimensional spectrum. For quadrature detection, two phase increments have to be acquired for each time increment in every dimension. As can be seen from (1.4), the overall number of FIDs to be recorded grows fast.

$$\mathrm{FID}_{\mathrm{total}} = 2^{n-1} \cdot j_1 \cdot j_2 \ldots j_{n-1} \tag{1.57}$$

$$j_i = \mathrm{sw}_i \cdot t_i^{max} \tag{1.58}$$

Consequently, NMR spectra with $n > 2$ dimensions usually cannot be acquired to t^{max} in the indirect dimensions with a direct impact on the resolution and the sensitivity of the resulting spectrum[82,84], while 4D (or higher) experiments are practically not feasible. Table 1.2 gives exemplary measurement times for different dimensions, with the time needed to acquire one FID being 1 s and 128 increments being acquired in every indirect dimension. Non-uniform

Table 1.2: Exemplary time estimations of uniformly sampled NMR spectra of different dimensions. A single scan experiment with a duration of 1 s and 64 time increments, i.e., 128 total increments, were assumed in every indirect dimension.

	2D	3D	4D	5D
time	2.1 m	4.5 h	24.3 d	8.5 y

sampling (NUS) can avoid these problems as only a subset of the initially required FIDs is recorded (Fig.1.2), and the full spectrum or peak positions are predicted or reconstructed through different algorithms[80–82]. NUS can be employed to increase resolution, i.e., extending the acquisition time in the indirect dimension, increasing sensitivity by weighting the number of scans in relation to the time increment, and decreasing experimental time or combining those[80–82,84].

Point spread functions and NUS schedules

Fourier transformation of a real-life NMR signal can be understood as FT of the individual components it is constituted from, which ultimately results in a more or less narrow Lorentzian peak with sinc wiggles of different intensity depending on the relaxation rate and truncation of the signal. These components are *i)* a non-decaying oscillatory function resembling only the frequency, *ii)* an exponentially decaying function due to relaxation, and in case of digital sampling *iii)* a train of delta function matched with the sampling rate. FT of *i)* and

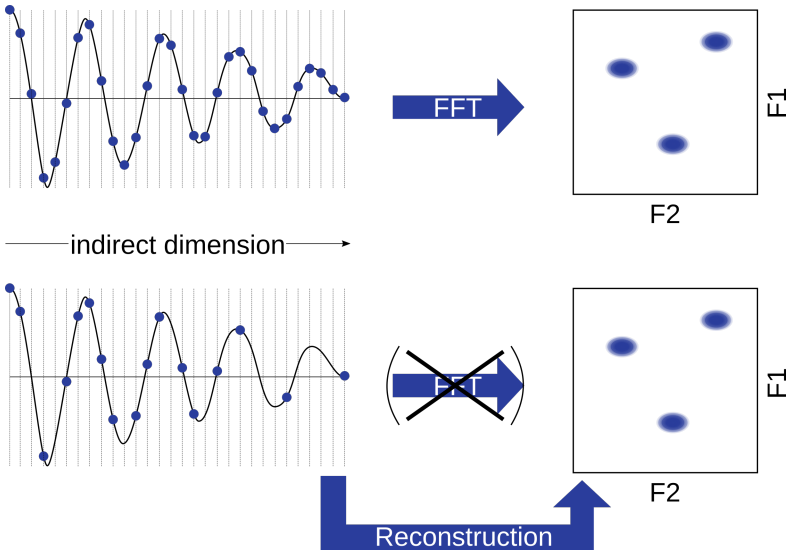


Figure 1.2: Comparison of non-uniform and uniform sampling. While in a conventional, uniformly sampled, NMR experiment, all data points are equally spaced, NUS datasets usually only contain a subset of data points that are "more or less" randomly distributed over the Nyquist-grid. Direct FT of those datasets is possible but can lead to severe artifacts in the spectrum, and therefore the spectrum usually has to be reconstructed by various algorithms.

ii) will lead to a Lorentzian peak while considering *iii)* introduces a sinc function[81,82,85]. It should be noted that adjusting the sampling rate is crucial to avoid aliasing in the spectrum. In NUS, however, several data points are omitted, and direct FT (DFT) results in deconvolution of the sparsely sampled FID and the sampling function. Omitted points are assumed to have zero intensity, which ultimately leads to artifacts in the spectrum[81,85]. The FT of the sampling function is called the point spread function (PSF), and for the mentioned reasons, the choice of a suitable sampling schedule directly impacts the NUS artifacts observed in the spectrum. Therefore, the intensity of NUS artifacts depends on the intensity of the actual signal; intense signals lead to intense artifacts but additionally will affect signals which position in the spectrum overlaps with the PSF[81]. In this context, random sampling was found to be highly suitable as the intensity of artifacts is equally distributed of the entire space leading to artifacts of lower intensities. Besides purely random schedules, several varieties exist, such as Poisson-Gap sampling[80,81,85,86]. The equal distribution of NUS artifacts in random schedules causes white noise like artifacts, and their intensity or signal-to-artifact ratio solely depends on the number N of NUS points[81,82,85] such that

$$S/A \propto \sqrt{N} \tag{1.59}$$

as well as the absolute number of peaks in the spectrum which is of great relevance, as narrow lines can be obtained through long evolution times without the necessity to increase the number of samples or, alternatively, the number of dimensions can be extended, which effectively causes a decrease of NUS density. Nevertheless, this also implies that a higher number of points has to be chosen for large molecules to improve S/A. To improve the signal-to-noise ratio, noise being white noise here, the sampling function can be probability-weighted such that more samples of shorter evolution times and higher S/N are sampled, which is essentially equivalent to the application of a window function in uniformly sampled datasets[81,85,87].

1.4.1 Automated Projection Spectroscopy (APSY)

Automated projection spectroscopy (APSY) might present a special case of NUS and was first presented by Hiller et al. in 2005[88]. APSY traces back to the cross-section theorem of Nagayama et al.[89], which states that a cross-section of an angle α, running through the origin in the two-dimensional time-space, i.e., FID, is Fourier-transformed into a cross-section in the two-dimensional frequency space, i.e., spectrum. While not directly omitting points in the full FID, all peaks in the n-dimensional space are projected onto two-dimensional planes, oriented in the n-dimensional space at different angles α, $\beta, \ldots$[88]. Therefore, a three-dimensional spectrum is the lower limit in which all 2D projections are oriented with an angle α. There is no theoretical upper limit of dimensions, and so far, APSY has been reported for up to seven dimensions[90], with sensitivity being the limiting factor[88,91]. For every additional dimension, one additional angular dependency has to be introduced[88,89,92]. As the projection

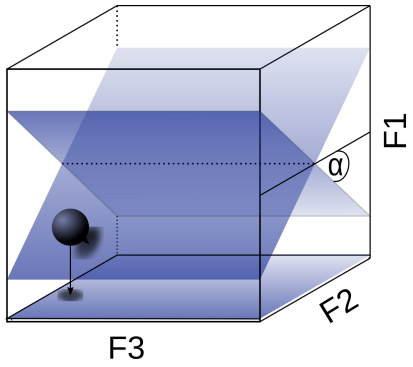


Figure 1.3: Representation of projection planes. Shown are two planes, representing ±-projections of an arbitrary angle $0° < \alpha < 90°$ as well as, framed in black, one of the special cases for $\alpha = 0°$. For a better overview, the $\alpha = 90°$ projection is not shown. A peak in the 3D space (black) would be orthogonally projected on a tilted plane with $\alpha = 0°, 90°$ representing the F3-F2 and F3-F1 projection of a conventional spectrum.

planes run through multiple indirect, i.e., $n-1$ dimensions, the chemical shift of the projected signal is mixed from all crossed indirect dimensions, and therefore, the evolution time t_{APSY} is mixed[92].

$$t_{APSY} = t_1 + t_2 = t\cos\alpha + t\sin\alpha \tag{1.60}$$

The above equation has two exceptional cases for $\alpha = 0°$ and $\alpha = 90°$, in which the time increment for one of the indirect dimensions vanishes. Those projection planes represent the orthogonal planes F3-F1 and F3-F2, also obtained from conventional experiments (Fig.1.3).

For phase-sensitive spectra, 2^{n-1} phase increments have to be acquired for each increment of t_1 and t_2, e.g., four FIDs in the 3D case, as shown below[92] .

$$S_1 = \cos(\Omega_A t\cos\alpha)\cos(\Omega_B t\sin\alpha) \tag{1.61}$$

$$S_3 = \cos(\Omega_A t\cos\alpha)\sin(\Omega_B t\sin\alpha) \tag{1.62}$$

$$S_3 = \sin(\Omega_A t\cos\alpha)\cos(\Omega_B t\sin\alpha) \tag{1.63}$$

$$S_4 = \sin(\Omega_A t\cos\alpha)\sin(\Omega_B t\sin\alpha) \tag{1.64}$$

The sums and differences of those FIDs, $S_1 - S_4$, $S_2 + S_3$, $S_2 - S_3$, and $S_1 + S_4$, can be Fourier-transformed to obtain phase-sensitive projection planes at an angle $\pm\alpha$ representing the cosine

and sine modulated signals necessary for phase-sensitive spectra[92]. A direct implication of the mixed indirect dimensions is that the spectral-width changes accordingly in every projection plane as it also consists of both indirect dimensions. The SW in the projection planes is then given as[91]

$$SW = \sqrt{\sum_{i}^{N-1} (SW_i \cdot p_i)^2} \tag{1.65}$$

with

$$p = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \tag{1.66}$$

Consequently, the resolution is lowered depending on the acquisition time when the number of acquired points is kept constant as $AQ = \frac{1}{SW} \cdot TD$, with TD representing only time increments. This can potentially have an impact on the geometric reconstruction of the peak positions by GAPRO. The sensitivity is also compromised in non-orthogonal projection planes, as not only the overall acquisition time for a non-orthogonal plane might be shorter, but also because relaxation and potential window functions act on both indirect dimensions simultaneously[91]. These limitations are putting a potential sensitivity limit to APSY, particularly in solution-state NMR, when used for higher-dimensional spectra. APSY does not reconstruct the entire spectral space but back-calculates the peak position in the n-dimensional space based on the Euler angles and the peak position on the projection plane using the GARPO algorithm[88]. Therefore, only peak positions, i.e., chemical shifts, are obtained. While determining parameters like line width and $T2$ times are possible[88,91,92], a manual assessment of peaks and artifacts remains elusive. A comprehensive overview of how APSY/GAPRO works and the parameters that have to be considered can be found in the literature[88,91,92]. Here, only a short outline for the features of APSY/GAPRO will be given, which are relevant in the later course of this work. Peak identification, and at the same time, artifact suppression in APSY are achieved by geometrical analysis of the recorded projection planes. Based on the picked peaks, potential intersections of peak positions in higher-dimensional space are created from a randomly chosen subset of projections, with each intersection marking a potential peak position in the n-dimensional space. Based on the parameter $S_{\min,1}$, called support, those points, i.e., potential peaks, which are based on the highest number of intersections S, are selected, and the respective peak leading to those are excluded from the further analysis. The process is repeated until the support for all remaining peak candidates drops below $S_{\min,1}$. The peak list obtained through this selection process is then subjected to further treatment[88], and the entire analysis will be repeated several times with a different set of projection planes from which ultimately a consensus peak list is obtained as the final result[88]. Besides the $S_{\min}$ parameter for defining the minimal support, other important parameters include the signal-to-noise threshold for the automated peak picker implemented in the GAPRO algorithm as well as the tolerances of the intersections in the direct and indirect dimensions. Due to the broader line-width present in solid-state samples, the signal-to-noise ratio is usually reduced compared to solution-state NMR experiments and with weak peaks usually being present in every spectrum, the peak picking threshold has

to be set comparably low if a complete peak list is desired. In turn, these adjustments lead to artifacts being suppressed mostly on the $S_{\min}$ as those are randomly distributed over the spectrum and will be filtered out over the many iterations performed for every GAPRO run.

In solution-state NMR, APSY has been used for various proteins with up to seven dimensions[90] in non-globular proteins and generally with different dimensions for backbone[88,90,93–98] and side-chain[90,96,99] assignments of globular proteins, high-throughput applications[94] for which time is a crucial factor, protein-ligand interaction studies[96] and IDPs[93,100,101]. Due to the previously mentioned sensitivity limit, APSY is usually not applicable for globular proteins beyond 20 kDa[66,88]. However, TROSY based sequences have been reported for proteins upto 38 kDa[98], and recently ^{1}H-detected 5D solid-state APSY has been reported[66] for small to medium size proteins at 60 kHz MAS, 100 % back-exchanged, as well as 100 kHz for fully protonated samples, published shortly after the work presented here.

1.4.2 Signal separating algorithm (SSA)

Similar to the growing measurement time requirements affected by an increase in dimensionality, the time for processing, i.e., Fourier-transforming, the spectrum, and disk space, grow[102]. Although these effects are negligible for modern computers working with 1D, 2D, and even 3D spectra, they have to be considered for experiments with more dimensions. Potential problems arising for 4D spectra usually regard only disk space, with the size of one spectrum easily exceeding $>$ 20 GB, especially in experiments with a large spectral width compared to a few MB necessary for 2D or 3D spectra. Thinking of a 5D spectrum as a stack of n 4D spectra, it requires n times the space of each 4D when no further manipulation, e.g., zero filling, is applied and therefore easily can reach TBs, rendering it impossible to handle. Additionally, the time for FT also increases, however, only logarithmically rather than quadratic[103]. SSA[102,104,105] uses the concept of sparse multidimensional Fourier transform (SMFT)[85] to avoid both problems. SMFT will only process a subset of dimensions out of the n-dimensional space for given, pre-defined coordinates for the remaining dimensions. Figure 1.4 illustrates this process exemplary for the 5D HNcoCANH experiment. For 5D experiments, a set of 2D planes would be Fourier-transformed, with each 2D plane directly linked to a 3D resonance information and therefore adding up to a peak, actually dispersed in five dimensions. To do so, a 3D experiment has to be recorded previously, processed and peak-picked. Due to the nature of SMFT, 3D peaks not provided for processing the 5D will remain virtually invisible in the end and by that putting a limitation to the technique. Therefore, it is crucial to obtain high-resolution and high-intensity 3D spectra in the first place, which in turn can readily be acquired using NUS. NUS reconstruction using SSA works in two steps. First, the removal of sampling artifacts from the sparse FID, and second, the Fourier transformation using SMFT of the cleaned FID[102,104,105]. In the case of 5D data, SSA takes only predefined regions into account for the before mentioned reasons, while lower-dimensional data can be fully processed. In the first step, the sparse data set is searched for intense points or signals in the time-domain for every point in the Fourier-transformed direct dimension. For every intense point or signal, a modeled FID is modified, and FT to best resemble the sparse FID. As NUS noise is not random but deterministic, this ultimately removes NUS artifacts from the experimental FID. Consequently, subtraction of the simulated and sparse FID will lead to a residual FID or spectrum with reduced noise. If now an inverse Fourier

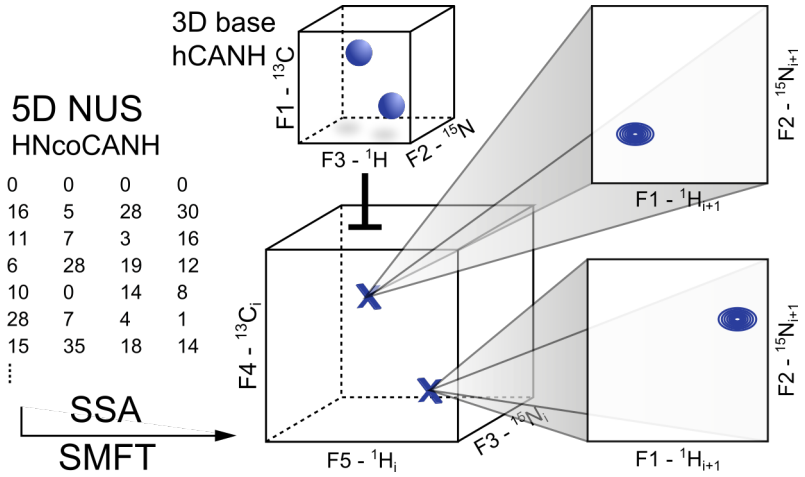


Figure 1.4: Concept of 5D processing with SSA / SMFT. First, a 3D experiment, sharing three dimensions with the 5D, has to be acquired, and peaks have to be picked. In the second step, a 5D experiment is acquired and afterwards processed. For processing (in all NUS experiments), SSA first reduces artifact noise (see main text), and afterwards, Fourier-transforms the additional two dimensions of the spectrum. However, both steps are only performed at positions pre-defined by the 3D peak list provided. As a result, a stack of 2D planes, each corresponding to a 3D peak, is created, with each 2D plane providing five-dimensional information in combination with its respective 3D peak. While NUS is advantageous in the case of the 3D to obtain high resolution, it is mandatory for the 5D.

transformed signal of a 5D model as used for fitting the sparse FID is added to the residual FID, a cleaner, i.e., artifact noise reduced signal is obtained, which can be subjected to the same process again. For the following FT, the peak information and residual signals obtained in this way a read and the spectrum can be fully transformed using FFT or SMFT in case of slicing as used for the 5D.

1.5 Studied proteins

1.5.1 Proteins as fundamental building blocks of life[106]

Proteins play a crucial role in all types of physiological processes. As a direct product of protein biosynthesis, they control important cellular processes such as proliferation or apoptosis, which is why a fundamental understanding of the structure and function is crucial for treating diseases. The structure of proteins can be classified into four categories. The primary structure describes the amino acid sequence of which the protein consists and carries no information about the protein's three-dimensional structure. The secondary structure describes individual structural elements of which the most common are the beta stand, the alpha helix, and the loop. The present ϕ and ψ angles can classify these elements in the protein backbone, but not all possible combinations are equally populated. This was found by G.N. Ramachandran in 1963 and can be visualized in the so-called Ramachandran plot[107]. Besides the most common, various additional structural elements, e.g., the 3_{10} helices, exist[108]. Beta sheets or strands are constituted from a more or less linear sequence of amino acids which often interacts over H-bonds with another beta-strand running in parallel, with side-chains pointing in the same direction, or antiparallel orientation, with side-chains pointing in the opposite direction. Despite expected dihedral angles of 180°, the commonly populated angles are $\phi = -139°$ or $\phi = -119°$ and $\psi = 135°$ or $\psi = 113°$ for antiparallel and parallel β respectively[109].

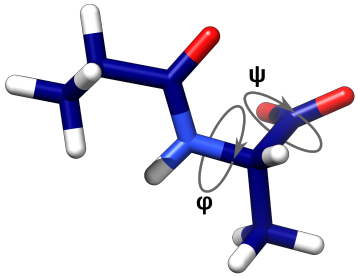


Figure 1.5: Visualization of ϕ and ψ angles shown on an Ala-Ala dipeptide. The ϕ angle describes the angle between the amide group and the Cα atom, whereas the ψ angle describes the angle between the carbonyl carbon and the Cα atom. The di-peptide was build in USCF Chimera[110].

α-helices, or 3.6_{13} helices, consist of 3.6 amino acids and 13 atoms per turn, in contrast to, e.g., 3_{10} helices which are built from 3 residues and 10 atoms per turn, and mostly occupy dihedral angles of $\phi = \pm 57°$ and $\psi = \pm 47°$[109]. The tertiary structure of proteins describes their overall three-dimensional monomeric shape build from the beforementioned elements and often grants intensive insights into the protein's function. Therefore, many structural biology techniques such as X-ray crystallography, cryo-EM, and biomolecular NMR aim to elucidate this three-dimensional structure. Beyond the tertiary structure, the quaternary structure describes the assembly of protein machinery that is constituted from multiple subunits, e.g., tryptophan synthase or the proteasome. Recombinant protein expression is one of the most used techniques in all fields of molecular biology. It allows the overexpression of a specific protein in a particular host, e.g., bacterial cells. To do so, a DNA vector carrying the gene for the target protein, usually along with genes for antibiotic resistance and a switchable promotor, is transformed into the cell. For bacteria, the heat-shock protocol is one of the most established methods to achieve this. Bacteria that took up the plasmid can easily be selected through antibiotics as only those will show resistance towards the selected antibiotic. Once in the cell, the expression of the target protein can be induced. One of the most chosen operons to control gene expression is the *lac* operon. Naturally, lactose would bind to the *lac* repressor and activate the expression of the gene, encoded after the operon. However, as lactose is gradually metabolized by bacterial cells, the expression rate decreases over time, and therefore, Isopropyl-β-D-thiogalactopyranoside (IPTG) is widely used as a replacement as it cannot be metabolized and the expression rate remains constant. If the cell density is sufficiently high at the point of induction with IPTG, the desired protein can be expressed in relatively high amounts and then isolated through a series of purification steps. These usually involve the lysis of the cells and multiple chromatography steps, e.g., affinity chromatography with His-tags and size-exclusion chromatography (SEC).

1.5.2 The SH3 domain as a model system in solid-state NMR

Src-homology 3 (SH3) domains are a, around 60 residues long, conserved motif found in several cytosolic proteins, e.g., tyrosine kinases, that were originally discovered in the c-Src tyrosine kinase of the Pous sarcoma virus[111]. Like those, they play an essential role in protein-protein interaction and are involved in many essential processes like enzymatic regulation of cell growth and many others[112–116], while their exact physiological role is relatively poorly understood[114]. Due to their regulatory properties, SH3-domains also pose an onco-

genic potential upon mutation[114]. SH3 domains usually share a high structural similarity and conserved tertiary structure, resembling a β-barrel-like structure, while the similarity in sequence remains low[114,117,118]. However, a conserved hydrophilic binding pocket recognizing of a proline-rich PxxP motif can be found throughout different protein constructs[112,114,115,119,120] and the specificity for ligand binding is regulated by the highly flexible RT and N-Src loops (Fig.1.6)[117,118,121]. While physiological relevance and functionality of SH3-domains

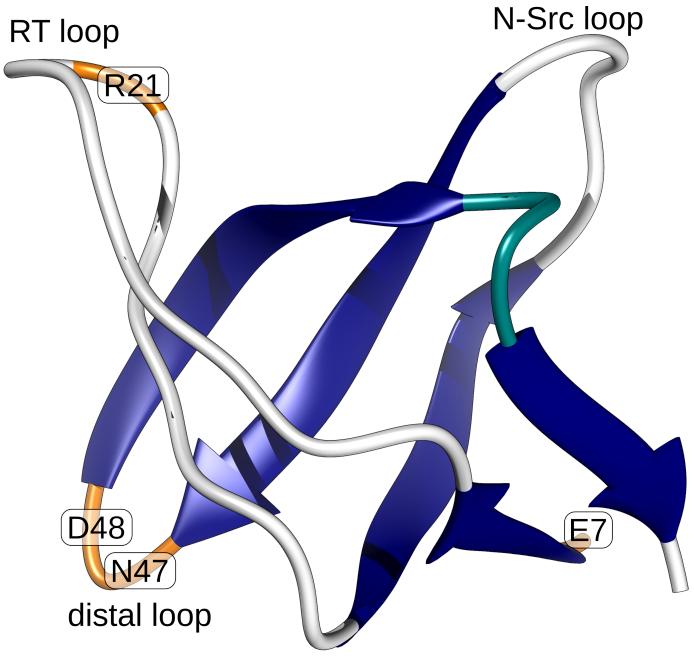


Figure 1.6: Crystal structure (PDB:2NUZ[122]) of the SH3 domain of chicken α-spectrin. Secondary structure elements are shown in dark blue (β-strand), teal (α-helix), and white (loop). Residues 55-57 show 3_{10} helical propensities while just completing a single turn, as indicated by the red color on the worm element[120,122]. Highlighted in orange are residues known to be flexible and are therefore not (M1-E7, N47, and D48) or only weakly detectable (R21) in CP-based experiments.[57,122]. Structure depiction and editing in UCSF Chimera[110].

is not completely understood, the isolated domain of chicken α-spectrin has been used extensively as a model system for NMR method development, in particular for solid-state NMR[12,26,37,38,42,43,57,68,72,121,123–126], as it can be expressed in high yield and homogeneously crystallized suing Cu-EDTA as PRE agent to shorten T_1 times[37]. In typical solid-state NMR experiments that facilitate CP for magnetization transfer, the N-terminus until residue E7 as well as residues N38 and D48 are not found due to their inherent flexibility[57,122], while also R21 shows significant line broadening and is, therefore, likely only visible in high sensitive experiments such as 2D hNH or 3D hCONH/hCANH and deuterated sample preparations, whereas both, E7 and R21, are absent in fully protonated sample preparations. The high β-strand results in well-dispersed spectra and isolated resonances (Fig.1.7), and residues in these areas can be considered rigid, while the three loops, RT-loop, n-Src loop, and distal loop, as well as the N-terminus regions, display significant flexibility[57,121,122,126–129]. In this work, the SH3-domain of chicken α domain was used as well for the development of new pulse sequences as expression and purification are well established, and assignments are readily available, enabling easy evaluation of new NMR experiments.

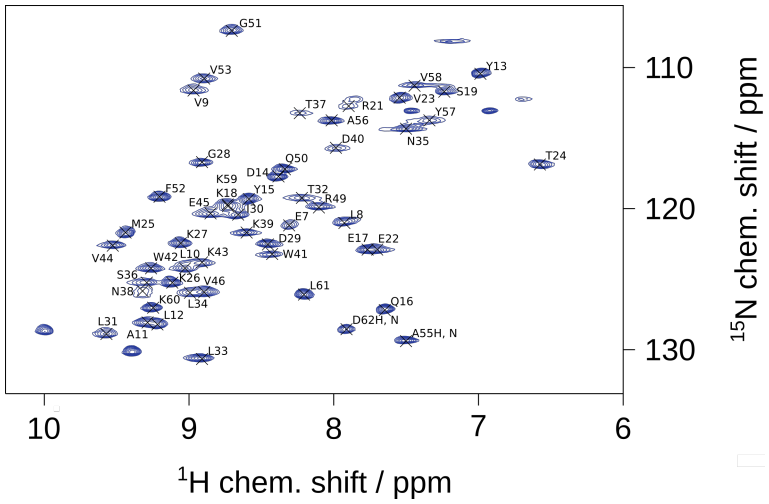


Figure 1.7: Assigned hNH spectrum of SH3.

1.5.3 Tryptophan Synthase of *S. typhimurium*

Tryptophan synthase (TS) is a 2x72 kDa enzymatic complex that catalyzes the synthesis of tryptophan in bacteria. Each 72 kDa (665 aa) asymmetric unit consists of a 29 kDa (268 aa)α-subunit and a 43 kDa (397 aa) β-subunit, which are assembled as $\alpha\beta\beta\alpha$ in the heterodimer. TS belongs to the family of pyridoxal-5'-phosphate (PLP) dependent enzymes, which catalyze a plethora of enzymatic reactions, including transaminations, racemizations, decarboxylations, eliminations, and substitution reactions[130–133]. The broad range of physiological functions taken by PLP-dependent enzymes makes them a well-studied drug target for an equal variety of diseases such as tuberculosis, epilepsy, Parkinson's disease, and more[134–139]. Furthermore, TS is a target for new antibiotics in the context of the emerging number of multi-resistant strains[138]. The β-subunit of TS has the unique ability to catalyze the reaction of tryptophan, an indole derivative, in bacteria and has therefore gained immense interest in biotechnology, foremost as a bio-catalyst for the synthesis of unnatural amino acids in the form of indole derivatives[135,140–143] made from TS mutants through directed evolution[144,145].

Figure 1.8 below shows the crystal structure of TS in complex with the ligand *N*-4'-trifluoromethoxybenzenesulfonyl)-2-aminoethyl phosphate (F9) trapping the α-subunit in its closed conformation and the β-subunit in its open conformation[146] together with the PLP cofactor, bound to K355 (βK87) in the internal aldimine state, which plays a crucial role in the catalytic cycle of TS, in particular of the β-subunit[140–142,146–150].

The synthesis of L-tryptophan in TS takes place in two consecutive steps, each step taking place in one of the subunits and without the release of the intermediate indole. Instead, the indole molecule is transported through an intramolecular tunnel between the α and β-subunit. While the α-subunit catalyzes the hydrolysis of indole-3-glycerol phosphate (IGP) into indole and glyceraldehyde-5'-phosphate (Fig.1.9 A), the β subunit catalyzes the condensation of indole and L-serine to L-tryptophan (Fig.1.9 B) in a complex catalytic cycle involving the PLP cofactor[146–149].

The entire process, especially substrate uptake, transport, and release of tryptophan, is associated with allosteric mechanisms due to large conformational changes necessary in different parts of the protein[146], in particular, the COMM domain in the β-subunit but

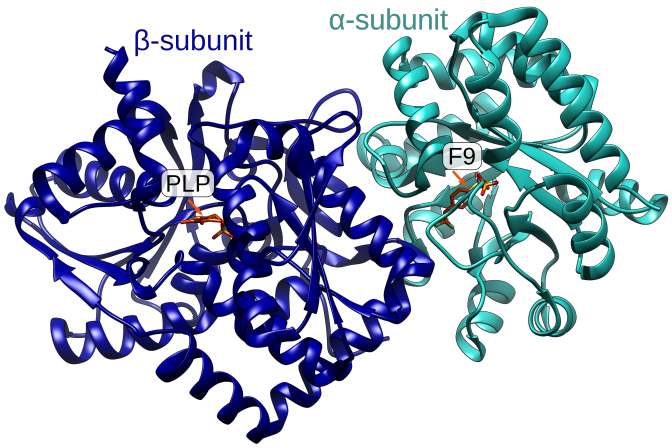


Figure 1.8: Crystal structure of tryptophan synthase (PDB:4HT3[146]) showing the α-subunit in teal with the ligand *N*-4'-trifluoromethoxybenzenesulfonyl)-2-aminoethyl phosphate (F9) in orange, and the β-subunit in dark blue with the PLP-cofactor in orange. The crystal structure shows the complete αL6 and αL2 loop with the α-subunit in the closed state and the β-subunit in the open conformation. Structure depiction and editing in UCSF Chimera[110].

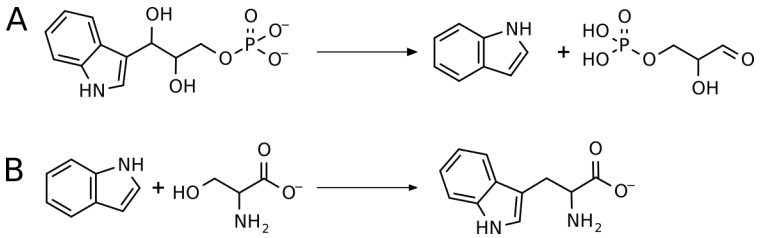


Figure 1.9: Reactions catalyzed by TS. **A)** In the α-subunit indole-3-glycerol phosphate (IGP) is hydrolyzed into indole and glyceraldehyde-3-phosphate. The indole is afterwards transported through the intra-molecular tunnel and **B)** reacts in a complex catalytic cycle with L-serine to L-tryptophan[140–142,146–150].

also loops αL2 and αL6 in the α-subunit[146,151–154]. While TS is well studied by X-ray crystallography and molecular dynamics (MD) simulations, NMR data remains mostly absent up to now, and NMR studies have been restricted to labeled substrates or the labeling of individual amino acids[146,147,149]. The high molecular weight of the asymmetric unit, as well as the allosteric and transport mechanisms, make TS an interesting and highly challenging sample for NMR method development and relaxation studies, but offers great perspectives for new insights into the chemical and dynamic machinery of TS, in particular, ^{1}H detected (solid-state NMR) could potentially shed new lights onto the ^{1}H chemistry involved in the active site of the β-subunit. Such insights can account, e.g., for the protonation states of the Schiff base in the internal aldimine state (Fig.1.10C) for which protonation of the imine has been proposed to active the C4' carbon of the cofactor before the nucleophilic attack of L-serin, but for which physicochemical details can only be estimated from first principle calculations[147,155–157], without the support of actual chemical shifts.

1.5.4 Human carbonic anhydrase II (hCAII): A model system for studying protein-ligand interactions

Human carbonic anhydrase II (hCAII) is a member of carbonic anhydrases (CA), or more specifically α-CA which catalyze the interconversion of CO_2 to HCO_3^- and are among the fastest enzymes known with a turnover rate of up 1×10^6 s^{-1}.[158] Like all CAs, hCAII contains

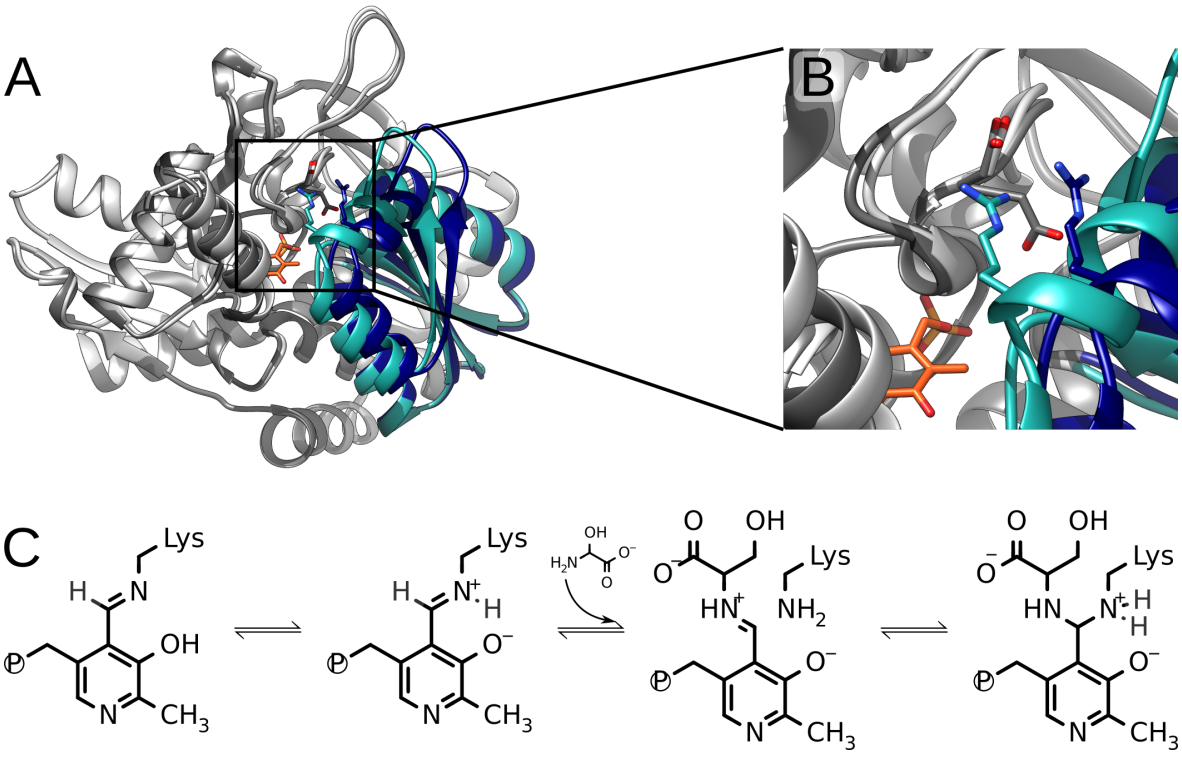


Figure 1.10: A) Conformational changes of the COMM domain (highlighted in blue tones) between the open conformation of the β-subunit (dark grey and dark blue; PDB ID: 4HT3[146]) and the closed conformation (light grey and teal; PDB ID:4HN4[146]) with the latter one showing a salt bridge between R409 and R573 (**B**)). In the open conformation, K355(βK87) is covalently bound to the PLP cofactor and in a tautomeric exchange between a neutral and zwitterionic state before L-serin diffuses in the active site and binds to the cofactor (**C**)). Structure depiction and editing in UCSF Chimera[110].

a metal center in which the reaction is catalyzed, consisting of a Zn(II) ion complexed by three histidine residues (Fig.1.11) and a solvent molecule in a tetrahedral geometry[158]. hCAII is a well studied system, inter alia, as a drug target for the treatment of glaucoma[159,160] and X-ray[161,162] as well as neutron diffraction[163–165] structures and MD simulations[166–168] are readily available. Recently, however, NMR could shed new light on the catalytic reaction, in particular involving protonation states of the His side-chains as well as dynamics and the water network involved in the reaction both in the solid and the solution state[54,64,168,169]. With a molecular weight of 29 kDa, hCAII poses as an enzyme close to the size-limit of solution-state NMR and was one of the largest proteins assigned by solid-state NMR[64] for which it now can be used as an assigned and well-understood sample for the development and application of new methods on larger proteins[54,64,169]. In the course of this work, hCAII was chosen for such purpose to evaluate the performance of a pulse program adapted from solution-state NMR to study protein-ligand interactions in the solid state (see section 4.1).

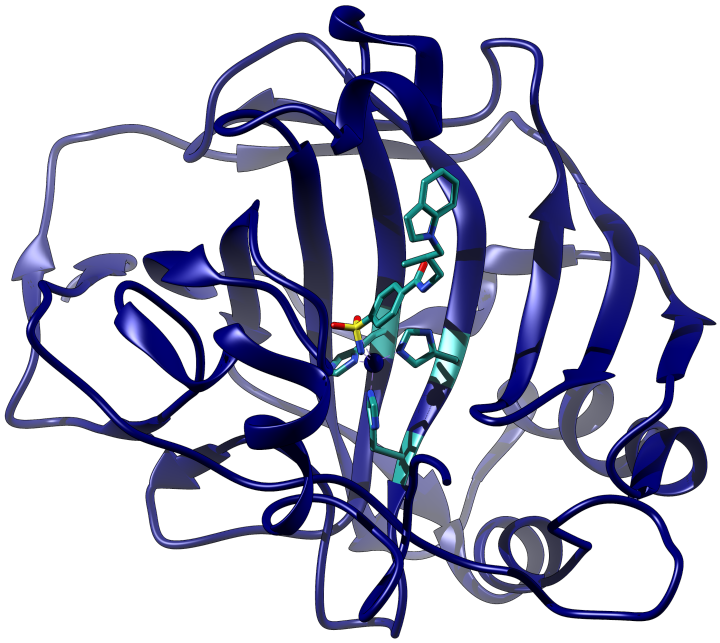


Figure 1.11: Crystal strucutre of hCAII with the ligand SBR ((R)-N-(3-Indol-1-yl-2-methyl-propyl)-4-sulfamoyl-benzamide)[170] (teal). Structure depiction and editing in UCSF Chimera[110].

Chapter 2

Projection spectroscopy in solid-state NMR

2.1 Automated projection spectroscopy (APSY) in solid-state NMR

The results of this section were published[124] as Klein, A., Vasa, S.K. & Linser, R. Automated projection spectroscopy in solid-state NMR. *J Biomol NMR* **72,** 163 170 (2018).

Introduction

As introduced in section 1.4.1, automated projection spectroscopy (APSY) only requires the acquisition of two-dimensional planes rather than the full number of FIDs in every dimension. Treating a uniformly sampled spectrum with at least three dimensions simply as a collection of 2D planes in every additional indirect dimension, APSY can lead to immense time-savings as usually a small number of planes is sufficient for the back prediction of peak positions. APSY relies, however, on sufficient resolution of the individual peaks, albeit not in every plane but at least in some of them, as the separation of resonances is not constant due to the mixed indirect dimensions. In solution-state NMR, the resolution of small to medium size proteins is often ensured due to significantly narrower line widths compared to micro-crystalline samples, whereas solid-state samples and spectra can be severely affected by broader lines due to incomplete averaging of mainly dipolar interactions as well as inhomogeneities of the sample and generally broader line widths in the projection dimensions[92]. Solid-state APSY might therefore be prone to more overlap and concomitantly reduced performance as line-width of 40-100 Hz for proton resonances and 15-20 Hz for nitrogen resonances can be observed on SH3 under the given conditions. To evaluate the performance of APSY, existing 3D ^{1}H-detected experiments commonly used for the backbone assignment of solid protein samples were recorded after having been made compatible for APSY. All experiments were recorded on a microcrystalline sample of u-[^{2}D,^{13}C,^{15}N] labeled and 100 % back-exchanged SH3-domain of chicken α-spectrin at 50 kHz MAS and a ^{1}H Larmor frequency of 700 MHz. Details on the sample preparation as well as APSY parameters can be found in section 6.2.

Pulse program and processing modifications necessary for APSY

Existing pulse programs in Topspin can be used for APSY with small modifications. Strictly not affecting the actual pulse program, a header has to be included, which is later on read by the APSY program to correctly identify the dimensions involved as well as back-predict the chemical shifts of every peak. A detailed description, especially of the *"Experiment"* parameter can found in various references[171].

Listing 2.1: Pulse program header required for APSY experiments

```
;;Dimensionality:3
;;Experiment:ONH
;;Phaseoverlap:000
;;trosy:0
;;DimProjections:2
;;Hfreq:BF1
;;Cfreq:BF3
;;Nfreq:BF2
```

```
;; Carrier1 :O3P
;; Carrier2 :CNST23
;; Carrier3 :CNST24
```

The actual pulses used for the NMR experiment remain completely unaffected by APSY, and solely the evolution of indirect dimensions as well as the routines for quadrature detection have to be adjusted as the commonly used macros in Topspin do not support APSY directly and therefore have to be manually replaced. The step size `in0` and `in10` for time evolution in both indirect dimensions is defined straightforward from equation (1.60) with `cnst51` as the projection angle that has to be individually set for every 2D plane and `in1` representing the default step size defined as $\frac{2t_1 max}{TD} = \frac{1}{SW}$ which consequently also has to be individually set for each different projection angle according to equation (1.60). `l11` only serves as a loop counter to ensure the correct size of the final FID.

Listing 2.2: Definition for simultaneous incrementation of the indirect dimensions.

```
define delay PI
"PI=3.1416"

"in0=abs(in1*cos(2*PI*cnst51/360)/2)"
"in10=abs(in1*sin(2*PI*cnst51/360)/2)"

"l11 = td1/4"
```

As mentioned, Topspin macros do not support APSY and therefore have to be replaced manually by different loops (Listing 2.1), which control the phase incrementation depending on the projection angle α. For the two special cases $\alpha = 0°$ and $\alpha = 90°$, corresponding to the pure H/N and H/C planes, only one phase has to be changed for quadrature detection, i.e. `ph17` and `ph11` respectively. For all other angles $0° < \alpha < 90°$, both phases have to change within one time increment, ultimately leading to four FIDs per time increment rather than the usual two. To ensure a constant number of four FIDs also for the orthogonal planes, simple waiting periods were introduced that leave the non-required phases untouched. Evolution of the indirect dimensions, on the other hand, is straightforwardly implemented after quadrature detection by adding `in0` and `id10` simultaneously. A direct consequence of the doubled number of FIDs per time increment is first that the overall pulse program has to be repeated 1/4 times the `TD` points instead of 1/2 and second that a direct Fourier-transformation, via the `xfb` command, of the collective FID will lead to wrong spectra, as Topspin always assumes the latter case of two FIDs per time increment. A Fourier-transformation of the unedited FID leads to a false sorting of phase and time increments, causing some signals to be moved to the edges of the spectrum, similar to artifacts moved when using States-TPPI or related methods[172,173].

Listing 2.3: Manual implementation of quadrature detection and time evolution

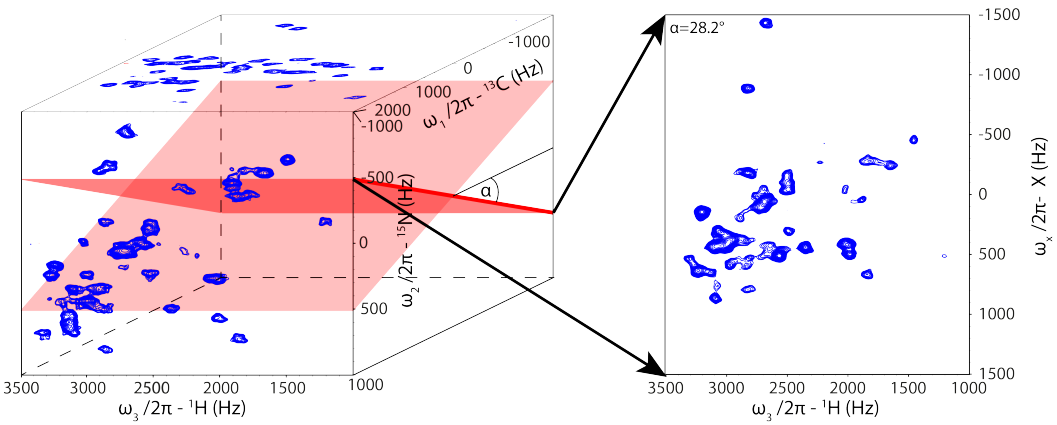
```
go=2 ph31
d11 do:f2    wr #0 if #0 zd
;---- Quad detection for Projection ----
if "cnst51==90"
```

```
{
        30m
}
else
{
        3m dp17
        lo to 3 times 2
}
if "cnst51==0"
{
        3m rp17
        3m
        3m
}
else
{
        3m rp17
        3m ip11
        lo to 4 times 2
        3m rp11
}
;----- Evolution for Projection -----
3m id0
3m id10

lo to 5 times l11

exit
```

The individual FIDs, therefore, have to be re-sorted, combined according to equation (1.61) and written to a new file, which can easily be done by a Python[174] or Topspin AU program. The APSY setup script will create two data sets for each angle intended to correspond to the $+\alpha$ and $-\alpha$ projection but is actually set up for the same angle, and therefore, the two spectra will look identical and can be summed up in the first place to increase the signal-to-noise ratio. In a second step, two new datasets have to be created in which the FIDs corresponding to the $+\alpha$ and $-\alpha$ plane are stored in a way that corresponds to the linear combinations of the FIDs belonging to one time increment (see equation (1.61)). For example, the difference of the first and fourth FID, $S_1 - S_4$, and the sum of the second and the third, $S_2 + S_3$ will create the x and y component of the first time increment in the $+\alpha$ projection and be stored in the first data set, while $S_1 + S_4$ and $S_2 - S_3$ do so for the $-\alpha$ projection and are therefore stored in the second data set. The two resulting total FIDs can then be Fourier-transformed, representing the respective 2D projection plane at the projection angle and subjected to the APSY program, involving automatic peak picking and geometric analysis by GAPRO[88]. Scripts for automatic generation of 2D APSY experiments in Topspin as well as automated peak picking and GARPO were kindly provided by Prof.



An e

Figure 2.1: xemplary representation of projection planes. Blue peaks on the cube surfaces (left) represent the $\alpha = 0°$ and $\alpha = 90°$ planes, corresponding to the pure H/N and the H/Cα projections Transparent blue planes indicate $\pm$ projections of an arbitrary angle $0° < \alpha < 90°$ with an exemplary projection depicted on the right for $\alpha = 28.2°$. The axes are shown in frequency units of Hz, with 0 Hz representing the carrier frequency, rather than in ppm to account for the mixed evolution along the indirect dimension for non-orthogonal planes.

Sebastian Hiller from the University of Basel.

Performance of APSY in backbone experiments

To assess the performance of APSY, a set of commonly used three-dimensional experiments for backbone assignments were recorded, i.e., hCANH, hCONH[35] and hCAcoNH, hCOcaNH[68], and the peak positions reconstructed based on two-dimensional H/(NC) projections. As the actual pulses remain unaltered by APSY, spectra of good quality and a sufficient signal-to-noise ratio are obtained, and the found resonances in the orthogonal planes, i.e., the H/N and H/C plane, reflect those found from conventional experiments as pure ^{15}N or ^{13}C indirect dimensions are obtained (Fig.2.1 and 2.2B). In all experiments, the $\alpha = 0°$ plane corresponds to the H/N plane, and the $\alpha = 90°$ plane equals the H/C plane. All other planes will show a mixture of those two planes, and so will the chemical shifts of the individual peaks, which might bear some resemblance but eventually cannot be analyzed manually. In this case, the chemical shift does not reflect the actual position anymore, and therefore, the analysis is done in the frequency domain with 0 Hz at the carrier position (Fig.2.1 left). Only in the final step, when the three-dimensional coordinates are known, the frequency units will be converted back to chemical shifts for the respective nuclei.

At no point does the GAPRO algorithm perform an actual reconstruction of the full spectrum, in contrast, e.g., to most NUS reconstruction algorithms, and a plain peak list with the reconstructed, i.e., geometrically analyzed, peak positions are obtained. To evaluate the accuracy as well as the precision or overall correctness of the peaks, the respective chemical shifts were compared with those known for the SH3 domain based on conventional experiments. However, the experiments will always only share the two amide proton and nitrogen chemical shifts among all of them, while the carbon resonance either deviate in their connectivity, i.e., *i* or *i-1*, or the shift region they occupy, i.e., Cα and CO are not comparable, and the best way for comparing all experiments with each other is to look only at the ^{1}H/^{15}N chemical shifts (Fig.2.2). Good accuracy is obtained as most peaks group well in regions of expected peaks, as well as high precision among the peaks in a defined region. While the accuracy generally

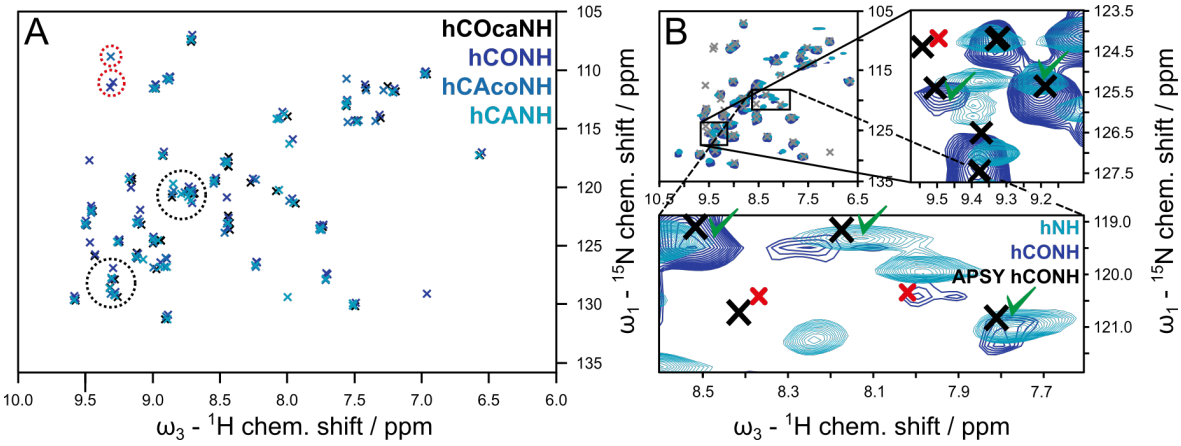


Figure 2.2: Analysis of geometric reconstruction of peak positions by GAPRO for different 3D backbone experiments. **A)** shows an overlay of H/N correlations from 3D peak lists with a different number of predicted peaks for every experiment, either due to the presence of artifacts or the involved magnetization pathways. Relatively high precision is, however, obtained for the peak positions with average deviations of 0.005 ppm in the ^{1}H dimension and 0.03 ppm along the ^{15}N axis, which is below the resolution of the recorded FID. Marked with red dotted circles are exemplary artifact peaks that do not show up in normal, uniformly sampled experiments, and dotted black circles indicate crowded and overlapping regions for which a reliable prediction of chemical shifts is challenging. **B)** shows an overlay of a conventional 2D hNH (light blue) and the H/N orthogonal plane of the 3D hCONH experiment (dark blue) together with predicted peak positions by GAPRO for the latter experiment (grey/black crosses). A magnification of the before-mentioned crowded regions highlights correctly found peak positions (green checkmarks) as well as artifacts and missing peaks (red crosses) due to unresolved peak overlap as well as inefficient magnetization transfers in the case of broadened peaks.

will reflect on whether a peak is generally found, the precision indicates that the GAPRO algorithm reliably calculates the same chemical shift of a certain peak among different runs or experiments, which is essential for assignment, especially as manual assignment based on peak lists is challenging, and automated assignment, heavily relying on good referencing, is the method of choice in combination with APSY. Isolated peaks are faithfully detected, and their chemical shifts are in good agreement among different experiments, and the RMSD for those groups is 0.005 ppm in the ^{1}H dimension and 0.03 ppm in the ^{15}N dimension. Artifacts as well as crowded regions, however, are not considered in these values as the former ones do not appear in all experiments or are generally reproducible (Fig.2.2A red circles) and the latter ones as it is not possible to define which peaks belong to one group of peaks (Fig.2.2A black circles).

Nevertheless, ≥90 % of the expected peaks are found in all APSY experiments with slight differences in the absolute number of resonances found in every experiment (Table 2.1). This is expected for certain peaks based on the magnetization pathways as, e.g., amide resonances for glutamine and arginine side-chains can be detected in hCONH or hCAcoNH experiments but remain absent in hCANH or hCOcaNH spectra. Other peaks, however, are frequently not found as they are located in crowded and overlapping regions, leading to insufficient separation in the projections, or can generally be considered weak, e.g., due to strong line broadening. While those generally can be recovered when lowering the threshold for peak detection, this usually is accompanied by an increase in artifacts found by GARPO, which cannot be suppressed by setting the $S_{min1,2}$ higher as this, in turn, causes other, correct, peaks not to be found. Especially for those experiments with a lower sensitivity, the hCAcoNH and hCOcaNH, few artifacts can be found (Fig.2.2A red circles) even for optimal GAPRO

parameters, most likely due to the reduced signal-to-noise ratio for these experiments, but likely will have little impact on the assignment procedure as they can easily be ruled out. It

Table 2.1: Number of peaks reconstructed by GAPRO/APSY for all experiments in comparison to conventional data sets

	hCANH	hCAcoNH	hCONH	hCOcaNH
APSY	50	52	61	51
Conventional	53	58	66	51
Fraction / %	94	90	92	100

can therefore be concluded that despite intrinsically broader line widths in solid-state spectra and resonance overlap in certain regions, APSY/GARPO pick up 94 % of the expected peaks when only considering the pure number of peaks, and missing peaks are mostly related to overlapped regions or weak signal intensities whereas the number of artifacts remains low.

APSY performance in smaller subsets

The biggest advantage of APSY possibly lies in the significant time savings that can be obtained even in for three-dimensional experiments but more pronounced for higher dimensional experiments. To find the minimum number of planes necessary for reliable reconstruction of peak positions, a reduced set of planes was facilitated, exemplarily for the hCANH experiment, and the obtained peak lists were analyzed similarly as done before. In the full set, 22 planes were recorded in 12.5 h in contrast to a conventional and uniformly sampled 3D hCANH with the same $t_ind$max times, i.e., the same number of increments, which would require about 78 h of measurement time. The number of necessary projections can hardly be determined after strict rules as the number of expected peaks, the number of dimensions as well as the sensitivity of the experiment play a crucial role for reliable reconstruction and the initial set-up. Nevertheless, it can be expected that less than 22 planes should lead to satisfactory results when considering that 5D APSY experiments were successfully performed with merely 18 projections[91,175].

The number of projections can be reduced by a factor of two, considering only the non-orthogonal planes, with small to moderate impact on the outcome of the geometric analysis (Fig.2.3 and table 2.2), which therefore directly translates into a time-saving of close to a factor of two taking into account the additional orthogonal planes. The RMSD, here calculated for well grouping peaks and as deviation from reference shifts obtained from a conventional hCANH, remains unchanged throughout all three different sets in all dimensions. In the ^{1}H-dimension, the RMSD is unaffected compared to the precision among the different experiments as shown in the previous section, and while the RMSD is slightly higher along the indirect dimensions, it remains on the order of the FID resolution (Table 2.2).

As can be suspected from the previous observations, the success of reconstruction for weak peaks and those located in overlapping regions of the spectrum is altered when using fewer projection planes as input for GAPRO, substantiating the initial assumption that APSY will be prospectively less suited for solid-state spectra due to the inherent lower resolution. Nevertheless, isolated peaks can be faithfully detected with a reduced number of projections with only small changes in the GAPRO parameters (Table 2.2). These adjustments do not follow trends and are rather set in a way to yield the best possible results for every set of

Table 2.2: Reconstructed peaks for reduced data sets employing a different, smaller number of planes for peak identification in the case of the 3D hCANH experiment

Number of planes	20+2	16+2	10+2	conventional 3D
Peaks found	50	57	46	53
GAPRO parameters				
$S_{min1,2}$ [a]	7	7	7	
r_{min} [b] / Hz	50.0	40.0	55.0	
$\Delta\nu$ [c]	20.0	18.0	25.0	
RMSD [d] ^{1}H/ppm	0.034	0.033	0.036	
RMSD [d] ^{15}N/ppm	0.17	0.15	0.18	
RMSD [d] ^{13}C/ppm	0.18	0.19	0.19	

[a] $S_{min1,2}$: minimal support needed for a peak candidate
[b] r_{min}: peak matching tolerance in the indirect dimension
[c] $\Delta\nu$: peak matching tolerance in the direct dimension
[d] Compared to the conventional 3D as reference

reduced projections. The values for peak matching in the direct ($\Delta\nu$) and indirect dimension (r_{min}) were initially set to the FID resolution and consecutively optimized around that while the *support* parameter ($S_{min1,2}$) has to be significantly higher than the recommended initial value of the number of dimensions[91,175]. Not only the precision but also the accuracy is maintained on a high level in all dimensions for well-defined peaks, whereas especially in crowded regions, more artifacts are picked up by the algorithm (Fig.2.3 red circles) when reducing the number of projections.

Given that the support can be relatively low in the full set, with only 7 out of 22 planes and therefore 32 % of the planes, to obtain good results in particular for isolated peaks which account for the majority of peaks in the spectrum, APSY/GAPRO performs well to identify those peaks and differentiate them from noise and artifacts. On the other hand, it becomes clear that a peak has to be identified in the majority of the projections when reducing their overall number as the $S_{min1,2}$ values remain constant with a dropping number of projections. Alteration of the parameter will ultimately yield either a higher number of artifacts when set to low or a reduced number of correctly found peaks when set to high to suppress artifacts. In solid-state APSY, it will be hence beneficial to acquire a rather high number of projection planes to support the correct peak prediction for more challenging parts of the spectrum. When spectra are already known for a certain protein, like here with the SH3-domain, or a set of experiments are acquired, it is potentially easier to tell true peaks and artifacts apart.

Potential time-savings using APSY

Using the full set of projections, APSY can provide a time-saving of about a factor of six compared to a conventional data set acquired with the same conditions, i.e., the same number of scans and acquisition times in the indirect dimensions. Uniformly sampled experiments, however, would be acquired with shorter acquisition times, i.e., fewer increments, as resolution and dispersion are often sufficient for smaller proteins in three-dimensional experiments. Even when shortened $t_{1,max}$ times of, e.g., 22 ms or 60 time increments in the ^{15}N dimension and 11 ms or 40 time increments in the $^{13}C\alpha$ dimensions of the hCANH experiment are considered, APSY can still provide time-savings of about a factor of 2.5 with simultaneously providing much higher resolution. When truncating the indirect dimensions for the projection planes

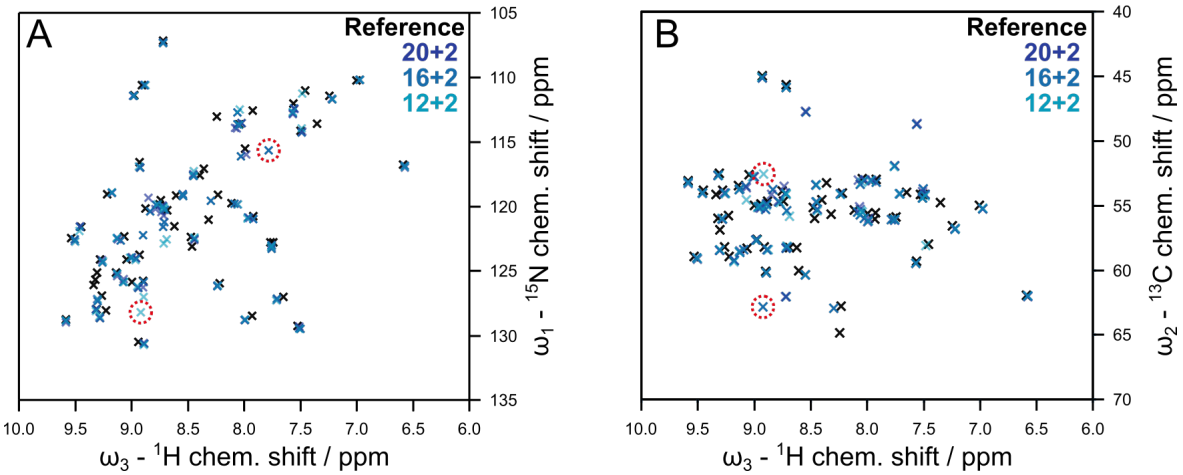


Figure 2.3: Predicted peak positions by GARPO in the hCANH experiment plotted as H/N and H/Cα planes for a different number of projections used for the calculation. Shown in black as reference are positions obtained from a conventional experiment, while peaks blue shades represent those obtained from the APSY experiment. Red dotted circles highlight exemplary artifacts found for different sets of projection planes. Whereas artifacts increase, the level of precision and accuracy for correctly identified peaks remains on a high level (see also Table 2.2).

by half, i.e., 14 ms and 8 ms for the ^{15}N and ^{13}Cα respectively, APSY performs equally well both when considering the number of found peaks as well as their individual positions and using those acquisition parameters as a new point of reference, the entire set of planes could be recorded in less than 8 h providing now significantly accelerated acquisition. Generally, the time-saving possible through the incorporation of APSY is

$$F = \frac{t_{nD}}{t_{APSY} \cdot 2(n+1)} \tag{2.1}$$

with $2(n+1)$ being the number of projection planes used each with the acquisition time t_{APSY} and the number of angles n and t_{nD} as the experimental time of the conventional experiment. The number of projection planes does increase, however not proportionally, to the dimensionality of the experiment performed, and therefore time savings obtained through APSY will be significantly larger when exceeding three dimensions.

Discussion

It has been shown that APSY is also generally applicable for micro-crystalline solid-state NMR preparations despite the inherent increased line width. Generally, a good agreement between known chemical shifts of the SH3-domain from conventional experiments and peak positions obtained from APSY/GAPRO can be found, and moreover, the precision of peak positions among different experiments can be obtained on the order of the FID resolution. APSY does not require any *a priori* knowledge about the spectrum as is sometimes required for other techniques like Hadamard spectroscopy[176], and the processing of projection planes can be performed as for conventional experiments without specialized processing scripts apart from splitting the recorded FID into the right data sets for each projection angle in contrast to many NUS algorithms and processing protocols[177–179]. Moreover, as no actual reconstruction of the full spectrum is performed, the calculation of peak positions is fast compared to most NUS reconstruction algorithms and completes within less than a minute on current computers,

which allows for fast optimization of the different GAPRO parameters to yield optimal results.

APSY offers significant time savings for solid-state NMR experiments even though a higher number of planes is beneficial to reduce artifacts and resolve peak positions in overlapped regions of the spectrum. The effective time saving will, therefore, likely differ among different samples and experiment conditions, such as the acquisition time in the indirect dimensions, which can be increased with a low measurement time penalty in APSY. The presented work only facilitated three-dimensional experiments for APSY, and even though large time savings can already be obtained, the benefits will be more pronounced in experiments with more than three dimensions. As APSY prospectively suffers from the low-sensitivity of higher-dimensional experiments[88,91], the molecular weight independent transfer mechanisms in conjunction with usually shorter transfer times of CPs over INEPTs, make solid-state NMR particularly useful for high-dimensional experiments[64,69,70], and solid-state APSY is a good potential asset for proteins of larger molecular weight under the condition of a homogeneous sample preparation, as has been recently demonstrated with five-dimensional experiments[66].

Peak overlap proves to be already a drawback for such small proteins as SH3, and even though high-dimensional experiments will lead to better peak dispersion in the n-dimensional space, all peaks are ultimately projected onto a two-dimensional plane. While non-orthogonal planes can provide better dispersion over their orthogonal counterparts, especially when more dimensions are used[66], the sheer number of residues with increased line width will put limitations to the proteins accessible with APSY, such as TS as an example for poorly dispersed spectra due to the high α-helical character. Resolution problems in APSY, however, can be avoided through special labeling schemes as demonstrated previously for solution-state studies[96,98].

The GAPRO algorithm effectively only provides a plain peak list as a result, and other information about the peak such as intensity and line widths are lost, and whereas artifacts are usually well suppressed, especially when acquiring a set of different experiments, a manual assessment of the full spectrum remains elusive. While this is certainly wanted, it can be a limitation for APSY when used with challenging samples where a careful evaluation of each spectrum is desired.

It can therefore be concluded that APSY can successfully be used in ^{1}H-detected backbone solid-state NMR experiments, despite the hurdles expected and observed for such experiments. APSY yields good accuracy and precision for most peaks and presents a valuable tool, especially in combination with automated assignment. As measurement time is not the limiting factor for APSY experiments, it will be beneficial to record higher-dimensional experiments even for smaller proteins as the prediction of peak positions in crowded regions could be improved with little cost of measurement time. Most of all, APSY does accelerate the acquisition of experiments with more than two dimensions and can therefore be a valuable tool when measurement time is the crucial factor.

Chapter 3

Higher-dimensional experiments for resonance assignments

3.1 Assessment of higher-dimensional experiments

Partial results of this section were published[67] as: Klein, A., Rovó, P., Sakhrani, V.V., Wang, Y., Holmes, J., Liu, V., Skowronek, P., Kukuk, L., Vasa, S.K., Güntert, P., Mueller, L.J., Linser, R. 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.* **119**, e2114690119, doi: 10.1073/pnas.2114690119.

Introduction

In modern NMR spectroscopy of proteins, three-dimensional experiments on triple-labeled samples are the method of choice for backbone and side-chain resonance assignment as well as to obtain restraints for structure calculation. With exceptions[180,181], the size of proteins in solution-state NMR is usually limited to around 30-40 kDa due to increasing T_2 relaxation rates because of molecular tumbling, and therefore 3D experiments are mostly sufficient for resolving resonances in the respective proteins. At the same time, long transfer times put boundaries to the application of $>$ 3D experiments for globular proteins, and hence higher dimensional experiments are mostly applicable only to systems with long coherence lifetimes like the can be found, e.g., in intrinsically disordered proteins (IDPs)[182]. Solid-state NMR is not subjected to molecular size in that respect, and in combination with comparably short dipolar coherence transfers, it is ideally suited to investigate proteins of high molecular weight. In recent years the size of studied proteins has steadily increased[44,64,65,70] alongside with the application of 4D experiments[12,64,69,70,183,184] for assignment and structure calculation[12,48,73,74]. For proteins exceeding 40 kDa, however, even four dimensions will prospectively fail to provide sufficient chemical shift dispersion and therefore making assignment more challenging. When evaluating the impact on higher-dimensional experiments, two problems have to be distinguished. First, the problem of resonance overlap, which is expected to be more pronounced in solid-state NMR due to intrinsic broader line-widths, but generally increases with the number of residues present in the protein as the chemical shift range in which resonances occur are limited[62]. Second, deriving from that, the problem of signal ambiguity which should here be understood as the non-uniqueness for linking two adjacent residues in the protein backbone. While signal overlap can be looked at for experiments of all dimensions, the problem of signal ambiguity is mostly a two-dimensional problem, as common assignment strategies will usually *i)* find a carbon resonance for an amide shift pair of residue *i* and *ii)* identify the *i+1* amide shift pair in a complementary experiment sharing the carbon dimension. Ambiguity, i.e., branches, in the assignment process will now occur if for either of those steps more than one resonance appears in the spectrum and whereas they do not necessarily have to overlap in all dimensions (pure overlap problem), the ambiguity arises from overlap in a subset of the facilitated dimensions. For example, three overlapping resonances in the H/N dimensions will give rise to three carbon peaks in step *i)* but can be well separated in the other dimensions.

The overlap problem

Besides the ambiguity and branches in the assignment arising from peak overlap, peak overlap will limit the analysis, e.g., of relaxation studies as overlapped peaks in n-dimensions will yield mixed rates obtained from those peaks or of chemical shift perturbations as often done to study protein-ligand interactions. To assess the problem of peak overlap over a wide range of molecular weight and averaged over structural features such as high α-helical or β-sheet contents, biasing the results either towards higher probability in the former and lower probability in the latter scenario, synthetic chemical shifts lists were created based on BMRB statistics. While the BMRB holds mostly globular proteins, a chemical shift distribution is biased towards structured proteins is introduced, which, however, is more accurate than, e.g., the multiplication of existing shift lists of one particular protein to reach a particular molecular weight. Moreover, most backbone nuclei, respective CO, N, and H, display a normally distributed behavior with respect to their average chemical shift and standard deviation, whereas Cα resonances of glycines deviate significantly from those of the other 19 amino acids, and also the usually downfield shifted resonances of serines and threonines have to be taken in account. Cβ resonances show similar behavior and are absent for glycine residues, and random chemical shifts created from a normal distribution seem therefore not the correct way as they would not resemble important features. To overcome these hurdles, a random primary sequence, considering the natural occurrences of amino acids in vertebrates[185], was created with a defined number of residues, including glycines and prolines, and random, normal distributed chemical shifts were created for all nuclei, i.e., H, N, CO, Cα and Cβ, of every residue based on aminoacid-specific μ and σ from the BMRB were created (Fig.3.1). Cβ shifts were excluded for glycines while no chemical shifts were created for prolines, whose amide shifts are absent in ^{1}H-detected experiments as well as for reasons of simplicity. The obtained shift list was then subjected to overlap analysis for multiple combinations of nuclei in different dimensions. The Python script for the creation of the randomized shift list and the actual analysis can be easily modified to obtain a variety of correlations other than the ones shown below, potentially also including other side-chain resonances.

The radius for overlap was set to 0.075 ppm, 0.36 ppm and 0.29 ppm for ^{1}H, ^{15}N and ^{13}C, corresponding to an approximate line width of 100 Hz, 50 Hz and 100 Hz for the respective nuclei at a ^{1}H Larmor frequency of 700 MHz. Here, two 3D correlations were analyzed, 3D hCANH and 3D hCONH, which were previously used for relaxation experiments in solid-state NMR[12]. Both experiments do not seem to be prone to spectral crowding when all three dimensions are considered, even when the length of the primary sequence accounts for 1000 residues (Fig.3.2). Nevertheless, even the approximate 5 % of resonances overlapping with one additional resonance correspond to 50 residues in the latter case as normalized values are plotted on the y-axis. For such an ideal protein, it can thus be concluded that 3D experiments should prospectively be sufficient to resolve overlap in three and, therefore, also more dimensions, rendering them adequate for relaxation series or shift perturbations. In reality, however, these considerations might drastically fail (Fig.3.3F), as non-homogeneously distributed line widths due to differential relaxation properties or a high content of α-helical elements present in a real-life system, to name only a few. The extent of overlap in two dimensions, here shown for H/N and H/Cα correlations, follows a more pronounced dependency of

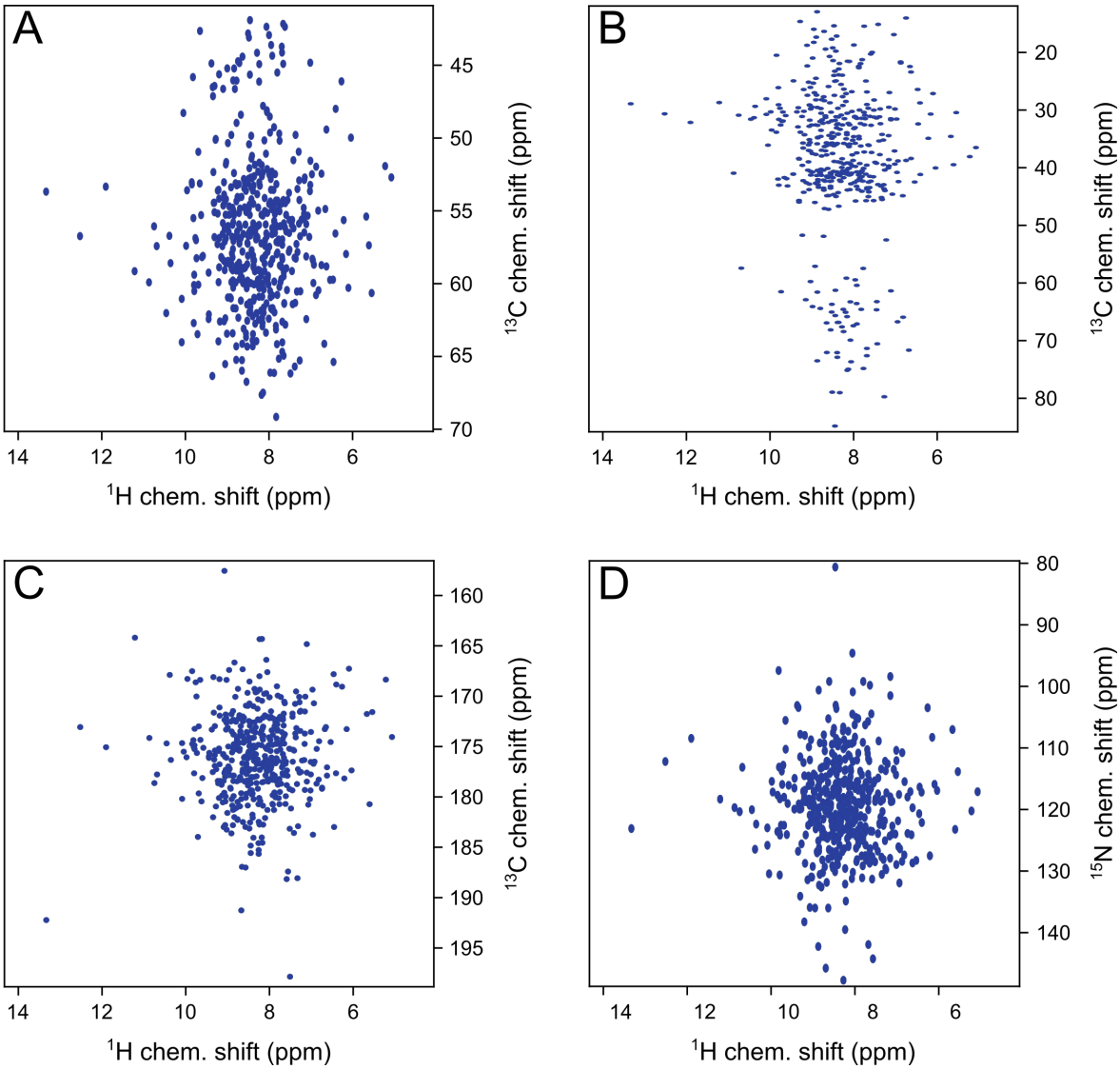


Figure 3.1: Exemplary H/X scatter plots for a synthetic shift list containing 502 residues. The H/N (**D**)) and H/CO (**C**)) planes show normally distributed shifts in contrast to the H/Cα (**A**)) and H/Cβ (**B**)) planes, which show a non-normal behavior due to the characteristic shift of glycine Cα nuclei and the Cβ shifts of serine and threonines. Prolines are not shown in the plots despite being included in the synthetic primary sequence.

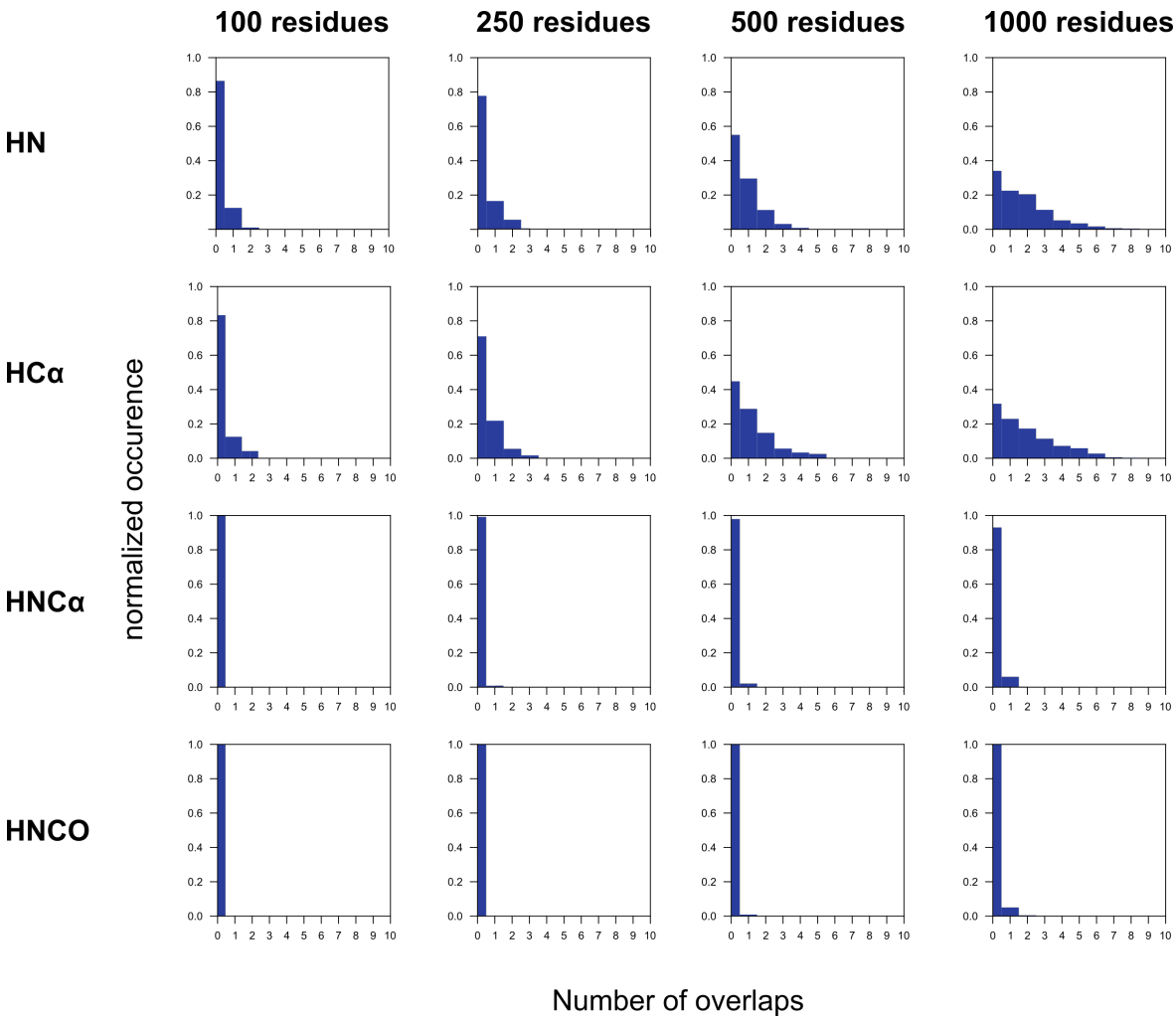


Figure 3.2: Barplots showing resonance overlap for different combinations of nuclei and in different dimensions for synthetic chemical shift lists of different sizes. The x-axis shows the number of overlapping resonances, and the y-axis is the occurrence for every event normalized to one. In 3D experiments, i.e., HNCA and HNCO, overlap presents only a small problem, and only starting from 500 residues resonance overlap can be found. The two 2D correlations show a stronger dependence on the molecular weight, and the number of isolated peaks is steadily decreasing. This is of great importance for most assignment approaches that are usually based on a 2D hNH/HSQC root experiment, as with every overlapping resonance, another branch in the backbone assignment is opened.

the molecular weight, i.e., the number of residues. Even for the small hypothetical protein made from 100 residues, merely about 85 % and 80 % of isolated resonances are found for the HN and HCα scenario, confirming the better ratio of line width and chemical shift range[68]. Despite only about 10-15 residues overlapping with one other resonance in both cases, 10-15 branches in the assignment process are implied when no complementary information will be used. Both trends get more pronounced with an increasing number of residues for which only approximately 35 %, i.e., 350 residues, are well isolated and another 40 % overlapping with one or two additional resonances (Fig.3.2). In line with the assessed 3D experiments, these trends are likely to be more pronounced for an actual protein such as tryptophan synthase.

The ambiguity problem

In contrast to the previous section, the ambiguity problem was approached based on real chemical shifts obtained from the 260 residue large hCAII and the 490 obtained shifts of TS

(section 3.3) to include two real rather synthetic and ideal scenarios in the assessment of the problem. hCAII was only recently assigned[64] and marked the size record for assigned proteins from solid-state NMR experiments at the same time, while tryptophan synthase (TS), although only partially assigned by 78 %, presents an even more recent target in the main focus of this work and with 2.5x more residues. Chemical shift lists from both proteins were analyzed with respect to the combinations for residue linking based on different 3D, 4D, and 5D experiments (Fig.3.3). To do so, all either CO, Cα, Cβ resonances, or a combination of those were counted for every given H/N shift pair of a given residue *i* within a specified range. In a second step, all H/N shifts for the previously counted combinations of carbon shifts were counted and multiplied with those to give the overall possible number of combinations for linking a H/N group of residue *i* with the adjacent *i+1* H/N group. Although hCAII shows around twice the number of unique Cα resonances for a given H/N shift pair compared to TS (Fig.3.3A/F), a linkage solely based on Cα resonances, i.e., an hCANH/hCAcoNH pair, is unsuitable for both proteins. In the case of TS, however, the distribution is strongly shifted away from 0 compared to hCAII (Fig.3.3B/G) and shows more extreme combination possibilities, indicating that a 3D-based assignment strategy is unsuitable. The introduction of additional dimensions improves the situation dramatically for hCAII. Both the hCACONH/hCOCANH and the hCACBcaNH/hCACBcaCONH pair give rise to only one or two possibilities for linking for about 50 % of the resonances. The additional Cβ dimension thereby appears to be even more impactful than the CO dimension, most likely due to better chemical shift dispersion, as the majority of H/N pairs give rise to less than five different ways for residue linking while for the CO dispersed scenario the majority lies below ten. A hypothetical 5D correlating all three carbon resonances with the amide group for residue *i* and *i+1* enables unique linking for about 70 % of the H/N resonances. The analysis for TS follows the same tendencies, however, with the absolute number of possibilities being tremendously increased overall and strongly reduced occurrences for only one or two linking possibilities. In contrast to hCAII, even the constructed 5D allows to uniquely link only around 40 % of the assigned residues of TS, and the majority of the resonances still give rise to up to ten ways for linking. It is to emphasize that these tendencies will become more pronounced for a full peak list of TS, which would add almost 200 additional resonances to the analysis. Consequently, the direct linking of residues from a single NMR experiment is essential for the success of resonances assignment. The linking of residues can most easily be achieved from straight-through transfers that directly correlate shared resonances of residue *i* with residue *i+1*, e.g., in an HNNH fashion. 4D HNNH pulse sequences were previously introduced and successfully used for the assignment of larger proteins[64], but nevertheless rely on sufficient resolution in the proton and nitrogen dimensions and are therefore of limited use for proteins like TS, in which this resolution is barely achieved (3.4). For hCAII, a 4D HNNH provides 80 % of unique links, and although an additional fifth dimension almost exclusively shows only one linking possibility, the advantage is limited here. In contrast, the fifth dimension brings significant advantages for TS in which the occurrence of a unique link is almost doubled to more than 80 % compared to the 4D HNNH experiment while the remaining 20 %, with few exceptions, will show two possibilities for linking. These findings are in good agreement with the results of the synthetic peak lists from the previous sections (Fig.3.2), as amide-to-amide experiments generally only depend on the overlap in the shift

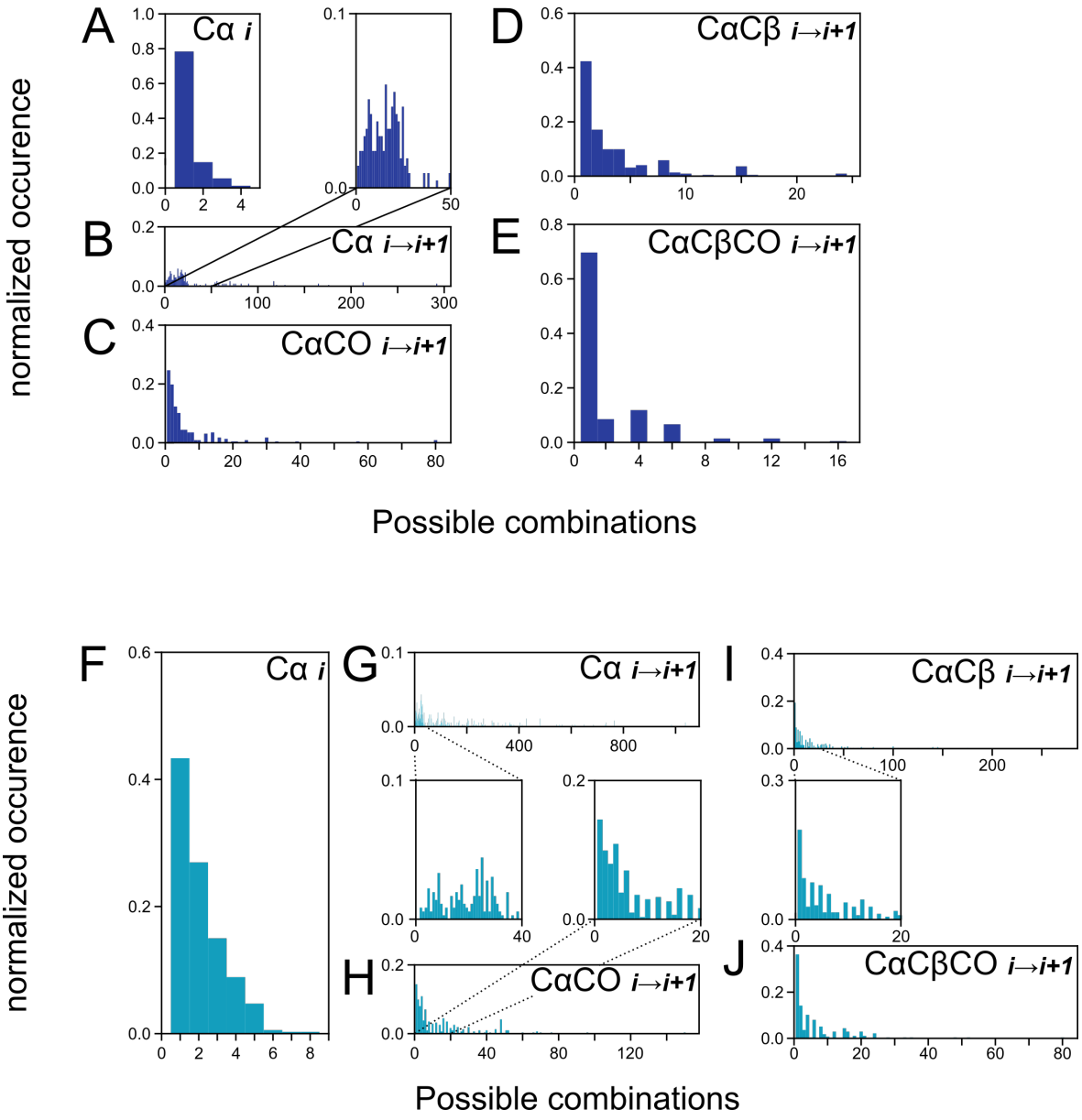


Figure 3.3: Bar plots showing the number of possible combinations for linking H/N *i* and *i+1* groups via different combinations of carbon resonances and different dimensions (**A,F** hCANH no linkage, **B,G** hCANH/hCAcoNH, **C,H** hCOCANH/hCACONH, **D,I** hCACBNH/hCACBcoNH and **E,J** hCACBCONH/hCOCACBNH, as a hypotheical 5D)for the 260 residues of hCAII (blue; **A**-**E**) and the 490 assigned residues of TS (teal, **F**-**J**). The occurrences of each plot are normalized to one.

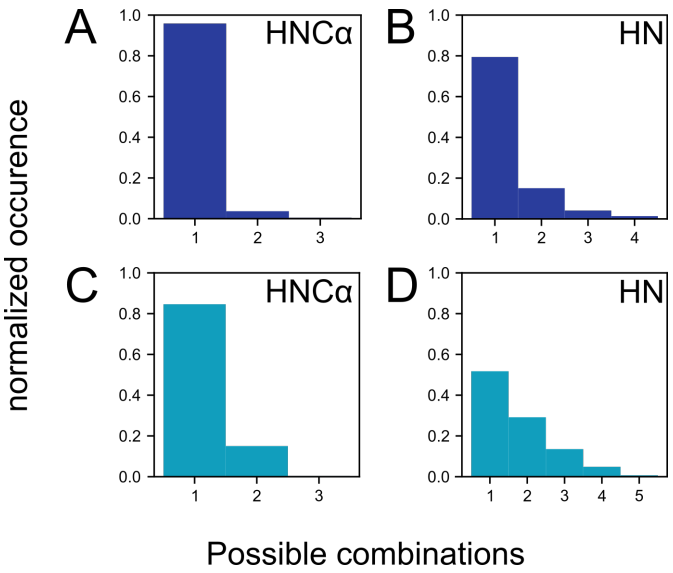


Figure 3.4: Assignment ambiguity in **A),C)** 5D HNcoCANH and **B),D)** 4D HNNH amide-to-amide experiments for hCAII (dark blue) and TS (teal). The x-axis shows the assignment possibilities based on an initial H/N shift pair or H/N/Cα shift triplet, with the y-axis being normalized to 1 as previously. No match-making between shared dimensions of two experiments is necessary, and the ambiguity solely relies on the resolution of the initial shift pair or triplet. The results are in good agreement with the predictions for resonance overlap in H/N and H/N/Cα correlations made from synthetic peak lists in the previous section.

pair or triplet of residue *i*.

The 5D HNcoCANH and 5D HNcaCONH experiments for backbone assignment introduced in this work, however, perform comparably good as a 4D HNNH when it comes to linking a residue *i+1* and *i+2*. As the 5D will only provide a H/N shift pair of residue *i+1*, only this two-dimensional information can be used for identifying the next H/N/Cα shift triplet of this residue, which then in turn would give the *i+2* residue. Nevertheless, the fifth dimension can be invaluable not only for linkage of residue but for sequential assignment. The identification of the amino acid type of a given residue is mostly achieved via Cβ chemical shifts when additional side-chain information is unavailable, as they show the most characteristic shifts of each amino acid compared to Cα or CO. To do so, the H/N resonances from an hCACBNH experiment are matched with those in other backbone experiments. Similar to the linkage of residues with the 4D HNNH experiment, this procedure strongly depends on sufficient resolution in proton and nitrogen dimension and can therefore lead to ambiguities (Fig.3.5 B, D). The 5D HNcoCANH, on the other hand, facilitates match-making in three rather than two dimensions due to the resolved Cα resonances. Similar to the linking scenario, this leads to rather moderate improvements in the case of hCAII (Fig.3.5A, B) from around 80 % to 90 % of unique matches, while it significantly reduces the number of possibly associated Cβ resonances for any given H/N/Cα triplet over a H/N shift pair (Fig.3.5C, D) for TS. Here around 70 % of the assigned resonances can now unambiguously be matched compared to 50 % before, and about additional 20 % will show only two possibilities. Considering the distribution of the above bar plots for 491 residues of TS, it can be concluded that 4D and 5D spectra are invaluable for the assignment of proteins with a molecular weight of 45 kDa or above as the described tendencies will become more emphasized for a complete assignment of TS. One might argue that high levels of assignment can also be obtained from different combinations of 4D spectra in combination with FLYA[186], as presented detailed in section

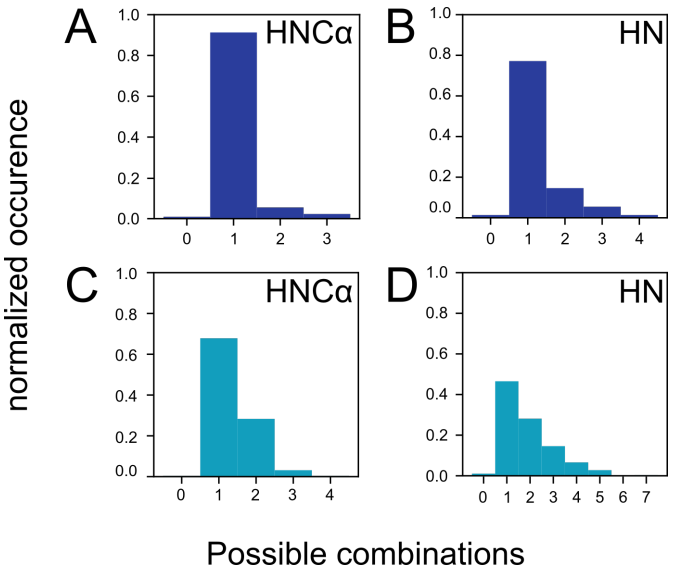


Figure 3.5: Number of H/N/Cα/Cβ resonances that can be matched onto either an H/N shift pair (right column, **B**,**D**) or an H/N/Cα shift triplet (left column, **A**/**C**) for hCAII (blue, top) and TS (teal, bottom) of a given residue *i*. As before, the y-axis was normalized to one. While the improvement is moderate for hCAII, there is a significant shift to the left in the distribution for TS.

3.3 or as recently demonstrated by a set of six 3D spectra for the 41 kDA (371 residues) large protein maltose binding protein (MBP)[65,187]. However, these approaches rely entirely on the high combinatorial power provided through FLYA and, in the latter case, might fail for proteins significantly exceeding this size. Furthermore, the 5D is not only essential for a manual assignment strategy but also for actual confirmation, i.e., proof, of assignments given by FLYA, and can reduce the number of required experiments to a minimum (see section 3.3).

Another advantage of higher-dimensional experiments, especially in the context of automated assignment, is the shift redundancy in combination with self-consistent referencing that is introduced as fewer experiments are required to obtain the same information. For example, a CO and Cα based, bi-directional backbone walk would require the set hCANH, hCAcoNH, hCONH, and hCOcaNH when three-dimensional experiments are used and which can be replaced by the pair 4D hCOCANH/hCACONH. As is demonstrated later in section 3.2.1, higher-dimensional experiments can be acquired comparably fast for smaller-sized proteins and can therefore present an alternative to the conventional assignment strategies also there.

3.2 5D experiments for backbone and side-chain assignment

3.2.1 Method developement of 5D backbone experiments

Introduction

Traditionally, the backbone assignment of proteins with ^{1}H detected solid-state NMR is achieved through a representative set of 3D experiments of which usually two complementary experiments are required to correlate two adjacent residues based on their shared carbon chemical shifts. Such experiments include the hCANH, hCONH, hCAcoNH, and hCOcaNH as well as a hcaCBcaNH to obtain the Cβ resonance often required for the identification of the

residue type. As demonstrated in the previous section (3.1), the two-step match-making strategy that is inevitable when using those experiments or their higher-dimensional counterparts leads to an error-prone assignment process as ambiguities are unavoidable when dealing with high molecular weight systems. Amide-to-amide experiments circumvent these problems, but whereas four-dimensional sequences were previously proposed and facilitated in solid-state NMR[69,70,73,74,184], 5D experiments are merely found for APSY[66] with the mentioned advantages and disadvantages. So far, 5D experiments remain almost exclusively being used for IDPs in solution-state NMR[182,188–190] where long coherence lifetimes support the extended number of transfer steps. In solid-state NMR, on the contrary, magnetization transfers are independent of molecular weight, and higher-dimensional experiments will prospectively outperform their solution-state equivalents[56,66]. In the following section, two 5D experiments for backbone assignment are introduced, and their performance with respect to sensitivity, assignment strategy, and benefits are tested on the SH3-domain as a model system.

The HNcoCANH and HNcaCONH experiment for backbone assignment

In the presented work, two different 5D experiments for backbone assignment were developed, the 5D HNcoCANH and the 5D HNcaCONH, which enable direct linking of two neighboring amide groups of different residues over an additionally evolved carbon resonance, and therefore, a straightforward backbone walk. The experiments are based on 4D HNNH sequences[62–64] but were expanded by an additional Cα or CO chemical shift evolution period. The phase cycling was adjusted and allows for a minimal phase cycle of 8 scans in the case of the HNcoCANH and 16 scans for the HNcaCONH. An additional phase incrementation was introduced to allow for clean and phase-sensitive spectra in all four indirect dimensions (Fig 3.6). As 5D spectra need a 3D base experiment for processing with SSA, the experiments were implemented as HNcoCANH and HNcaCONH rather than HNCOcaNH and HNCAcoHH to make use of the higher sensitivity of the hCANH and hCONH as base experiments compared to the hCAcoNH and hCOcaNH that would have been necessary otherwise.

As the HNcoCANH correlates two residues in a forward, N to C terminus manner, and the HNcaCONH correlates two amides in the reverse direction, each of the two experiments is sufficient for a complete backbone walk (Fig. 3.8). The HNcoCANH, however, shows slightly higher sensitivity, presumably due to increased transfer efficiency from N$\rightarrow$Cα compared to N$\rightarrow$CO[58] and turns out to be the better performing experiment (Fig 3.6D). Not only is the HNcaCONH the less sensitive experiment, but it also requires a longer phase cycle, which, however, might be shortened, but also appears to be prone to zero-frequency artifacts in the carbon dimension despite the extended phase cycle. Both experiments, however, show sensitivities comparable to other experiments like the hCACBcaNH, involve as many transfer steps as those and can therefore be considered to be well-performing and applicable to all types of proteins when considering that CP or dipolar transfers, in general, are independent of molecular weight in solid-state NMR. Compared to the 4D HNNH pulse sequence, the 5D will be as sensitive with no additional magnetization transfers in the HN bulk signal of the first FID. Due to the extra dimension, the overall intensity is reduced by $\sqrt{2}$ due to quadrature detection and additional relaxation losses during the added chemical shift evolution period.

The 5D HNcoCANH data presented here was recoreded in blocks of approximately 2 d of measurement time, after which a field reset was performed to account for the drift of the B_0

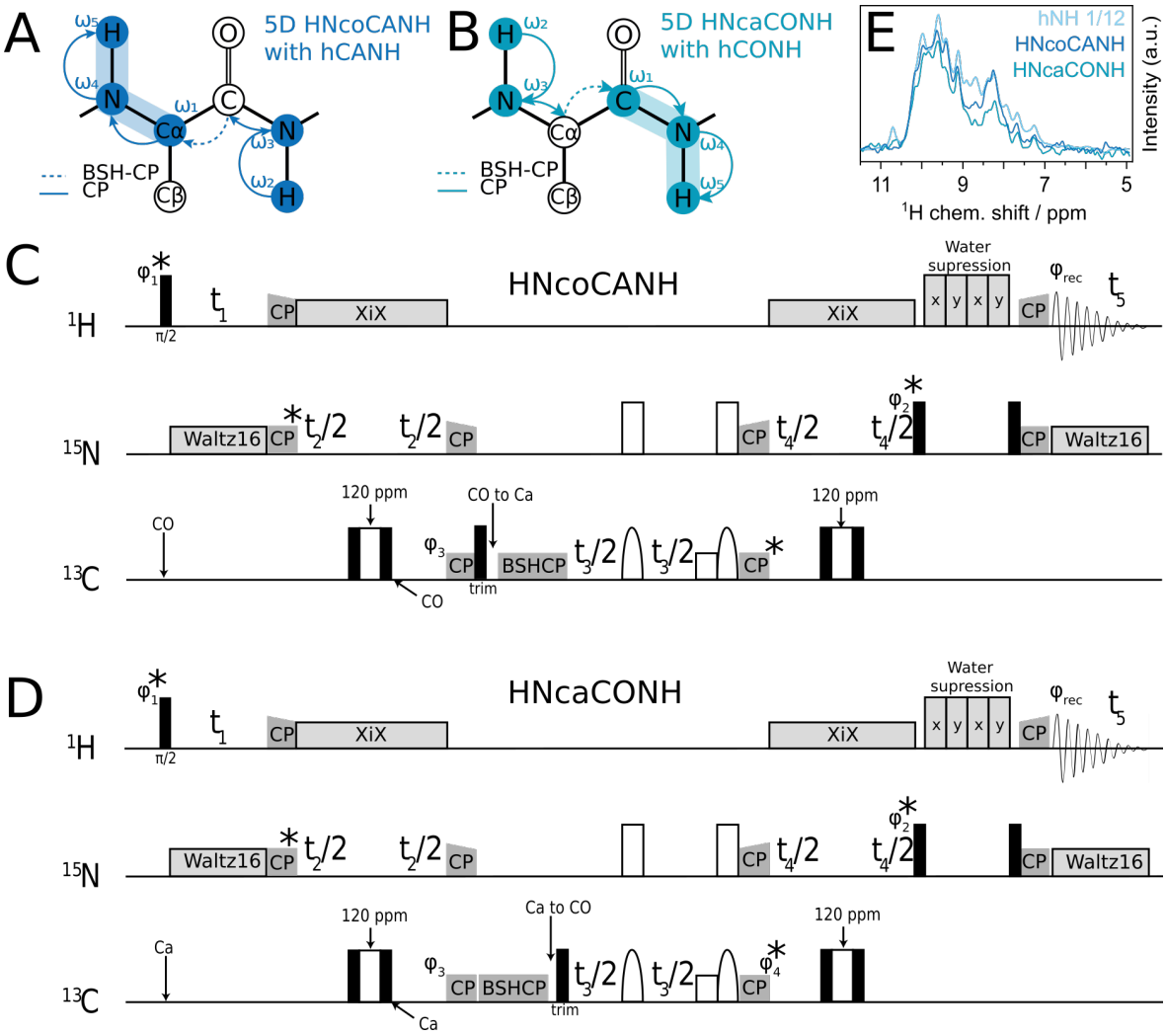


Figure 3.6: Schematic pulse sequences and transfer pathways for the developed 5D backbone experiments. All transfers are CP based, CC transfer are achieved by BSH-CP[26,27]. **A)** The 5D HNcoCANH pulse sequence correlates the amides of the *i* and the *i+1* residue with an additional chemical shift evolution period of $C\alpha_i$. As a base serves a 3D hCANH experiment as it yields better sensitivity compared to the hCAcoNH experiment[58]. Phase cycling: ϕ_1 = x,-x; ϕ_2 = x,x,-x,-x; ϕ_3 = x,x,x,x,-x,-x,-x,-x; ϕ_{rec} = x,-x,-x,x,-x,x,x,-x. B) The 5D HNcaCONH pulse sequence correlates amide to amide in a backward fashion, i.e. starting from *i* and ending on *i-1* with an additional chemical shift evolution period on $CO_i - 1$. The 3D base experiment is an hCONH due to the higher sensitivity over the hCOcaNH[58]. Phase cycling: ϕ_1 = x,-x; ϕ_2 = x,x,-x,-x; ϕ_3 = x,x,x,x,-x,-x,-x,-x; ϕ_4 = x,x,x,x,x,x,x,x,-x,-x,-x,-x,-x,-x,-x,-x; ϕ_{rec} = x,-x,-x,x,-x,x,x,-x,-x,x,x,-x,x,-x,-x,x Phases after evolution periods of indirect dimensions were changed to obtain phase-sensitive spectra by the States-TPPI method[173] (marked by asterisks). Filled and empty rectangles represent 90° and 180° pulses. The carrier changes on their carbon channel are denoted with arrows. Water suppression was achieved by MISSISSIPPI without gradients[191]. **D)**Sensitivity comparison of the 5D HNcoCANH (blue) and the 5D HNcaCONH (green) in relation to the hNH (light blue). Shown are the first FIDs of the experiments. The HNcoCANH gives ca. 7% of the hNH while the HNcaCONH only yields around 5%. Both values are in the expected range based on theoretical transfer efficiencies[58].

field and the absence of a lock signal in solid-state NMR. Each blocks consists of the same 2048 NUS points and was recorded with 8 scans. A base 3D hCANH was recorded with 896 NUS points and 16 scans in approximately 8 h, but in general any available, previously recorded, data can be facilitated. The 5D HNcaCONH was recorded in two blocks of 16 scans total with 960 NUS points of each block that were concatenated into 1920 NUS points in 3 days, while the required 3D hCONH was acquired with 1024 NUS points and 16 scans in 9 h. Taking into account the high sensitivity of the straight-through pulse sequence, the comparably high number of SH3 molecules in the rotor due to its low molecular weight and the high sparsity of 5D NUS list, these experiments can be recorded within two days of measurement time or even slightly below. When recorded with this minimal time, merely signals of residues E17 and N28 are missing, as the magnetization originates from R21 and T37, which both are severely broadend and hardly can be found in the 3D hCANH. All expected signals, including E22 and N38, could be detected after 4 d of measurement time, i.e., 16 scans. Nevertheless, the HNcoCANH was recorded for 6 d to ensure a high signal-to-noise ratio and easy identifiable peaks for the further analysis. A manual backbone walk with the HNcoCANH yields in 48 out of 48 expected and 52 assignable residues in the SH3-domain while the HNcaCONH yields in 41 out of 48 expected and 52 assignable peaks, confirming the overall better performance of the HNcoCANH experiment and showing that an almost complete backbone walk is achievable with a single experiment. However, the shorter measurement time of the HNcaCONH has to be taken into account and more correlations might be revealed after longer acquisition. Nevertheless, the HNcoCANH shows more peaks even after 2 days. Exemplary planes of the 5D spectra are shown in figure 3.7 below for residues E22 to T24 in both directions. Peaks of high S/N and reasonable line widths, in particular for ^{1}H, are obtained. While only the *i+1* or the *i-1* residue are expected for the HNcoCANH and the HNcaCONH, some planes contain an additional peak at the *i* position due to incomplete or back-transfer of magnetization (Fig. 3.7 A).

When incorporating the necessary 3D base experiments, the manual assignment can be straightforwardly done with each of the introduced 5D experiments (Fig.3.8). As mentioned before (3.1) the strength of the presented 5D experiments compared to the 4D HNNH is founded in the easy and less ambiguous matching of side-chain chemical shifts, i.e., Cβ, in the form of hcaCBcaNH experiments onto the backbone as three out of four chemical shifts can be matched (compared to only two in case of the HNNH). In the case of the HNcoCANH experiment, a 3D hcaCBcaNH with $\frac{1}{4J}$ transfer to Cβ can be utilized alternatively to a 3D hCANH, to simultaneously provide Cβ resonances. For the backbone walk, one would select a H/N/Cα shift triplet of residue *i* and maneuver to the corresponding 2D plane in the stack obtained from processing the 5D. This plane would show the *i+1* H/N shift pair, which in return, sufficient resolution provided, can be again selected in the 3D base experiment yielding in the Cα resonance of residue *i+1* and completing the next shift triplet to locate in the 5D experiment (Fig.3.8). Therefore, a reduced set of merely two experiments, 5D HNcoCANH and 3D hcaCBcaNH with $1/4J$ evolution of anti-phase, can be sufficient for near-complete resonance assignment of the protein backbone including Cβ shifts but omitting CO resonances. The backbone walk in C$\rightarrow$N-terminal direction can also be accomplished through the HNcaCONH in combination with its 3D hCONH base experiment in the same way as described for the HNcoCANH experiment (Fig.3.8B). However, the 3D hCONH experiment

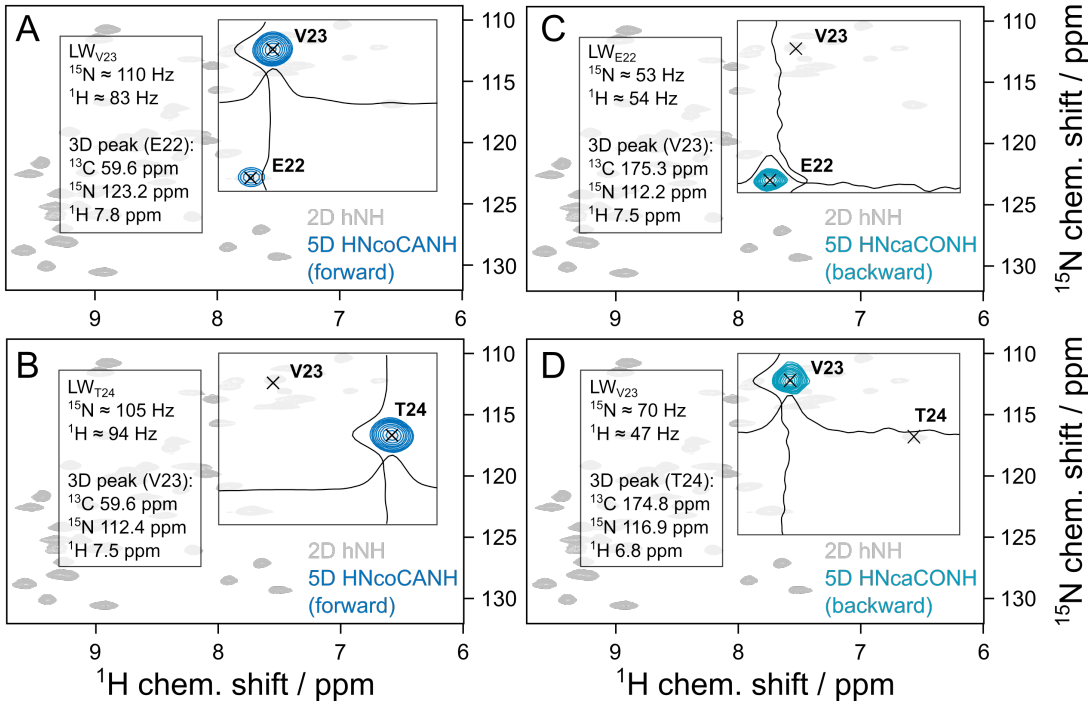


Figure 3.7: Line widths and single-residue signal-to-noise for exemplary residues E22 to T24 in the SH3 domain in the HNcoCANH experiment (**A)** and **B)**, dark blue) and using the HNcaCONH (**C)** and **F)**, teal). Line widths are estimated by using the Lorentzian peak fitting in NMRFRAM-Sparky[192] and subtracting 100 Hz and 50 Hz of line broadening applied for ^{1}H and ^{15}N during the processing. The HNcoCANH was recorded with 2048 NUS points and 24 scans, while the HNcaCONH was recorded with 1920 NUS points and 32 scans. The relatively high signal intensity indicates that experimental times can likely be shortened. Here, noise is referring to the remaining mix of NUS artifacts and thermal noise.

provides no side-chain information and the advantage of improved match-making in three dimensions for S2B experiments vanishes for this assignment strategy. Consequently, matching any side-chain correlation to the backbone will again be reduced to two dimensions with the H/N shift pair.

FLYA benefits in particular from the perfect intrinsic shift referencing through direct linkage of amides, necessary for reliable results[65], and the high redundancy introduced for H/N/Cα matching. Although 5D experiments will be most useful for more complex proteins of high molecular weight, as will be shown later on, they can be useful even for small proteins in combination with NUS and automated assignment as the necessary spectra can be recorded in essentially the same time as a set of 3D spectra required for backbone assignment in solid-state NMR (hNH, hCANH, hCONH, hCAcoNH, hCOcaNH, and hCACBcaNH) or even less when using only a minimal set of experiments (HNcoCANH, hCACBcaNH, hCONH). The SH3-domain is a small protein, and resonance overlap or ambiguity in high(er)-dimensional spectra are essentially absent, which is clearly reflected in the assignment results obtained by FLYA (Fig.3.9). Different combinations of experiments containing at least one 5D (B)-D)) in comparison with the FLYA results obtained from the full set of 3D peak lists (A)) all show close to 100 % correct resonance assignment when compared to existing chemical shifts available for the SH3-domain. All consolidated, i.e., final, chemical shifts were obtained from 100 independent FLYA runs within ca. 8 min, with tolerances for chemical shifts set to 0.05 ppm for ^{1}H and 0.5 ppm for ^{13}C and ^{15}N respectively. In all of the cases, a set containing at least one 5D works equally well or better than the 3D-based set, while in all scenarios, the rate of assignment is high and in good agreement with the manual assignment. Deviations among the different combinations of peak lists most likely occur from remaining referencing or peak picking problems, as most of the violated or *weak* peaks are broadened and therefore are even hard to detect and to assign manually[122].

Application of the HNcoCANH to fully protonated proteins

Although the benefit of fully protonated samples lies especially in the access of side-chain ^{1}H resonances (see the following section 3.2.2), they can present an alternative approach when deuteration of non-exchangeable sites cannot be achieved, and both 5D backbone experiments can equally be acquired on such samples. Due to less efficient averaging of anisotropic interaction, most of all ^{1}H–^{1}H dipolar interaction, even at MAS speeds in the 100 kHz regime, the coherence lifetimes are significantly compromised effectively, leading to less sensitive experiments that cannot be fully compensated by the coil design of commercially available probe heads[12]. Nevertheless, a similar performance of the 5D experiments can be expected at the expense of extended measurement times. To demonstrate the feasibility of the introduced pulse sequences, the 5D HNcoCANH was recorded on a u-[^{13}C,^{15}N] sample of the SH3-domain and the assignable resonances compared to those obtained from a deuterated and back-exchanged sample. As expected, the sensitivity of the HN bulk signal of the first FID (Fig.3.10A, C), as well as the signal intensity in the respective 2D planes processed from the 5D, is lowered in the fully protonated sample, preventing especially weak resonances from being found in the spectrum.

When recording the experiment for a total of 8 days, 41 out of 48 expected resonances can be directly assigned, and additional 4 resonances can be assigned indirectly from the

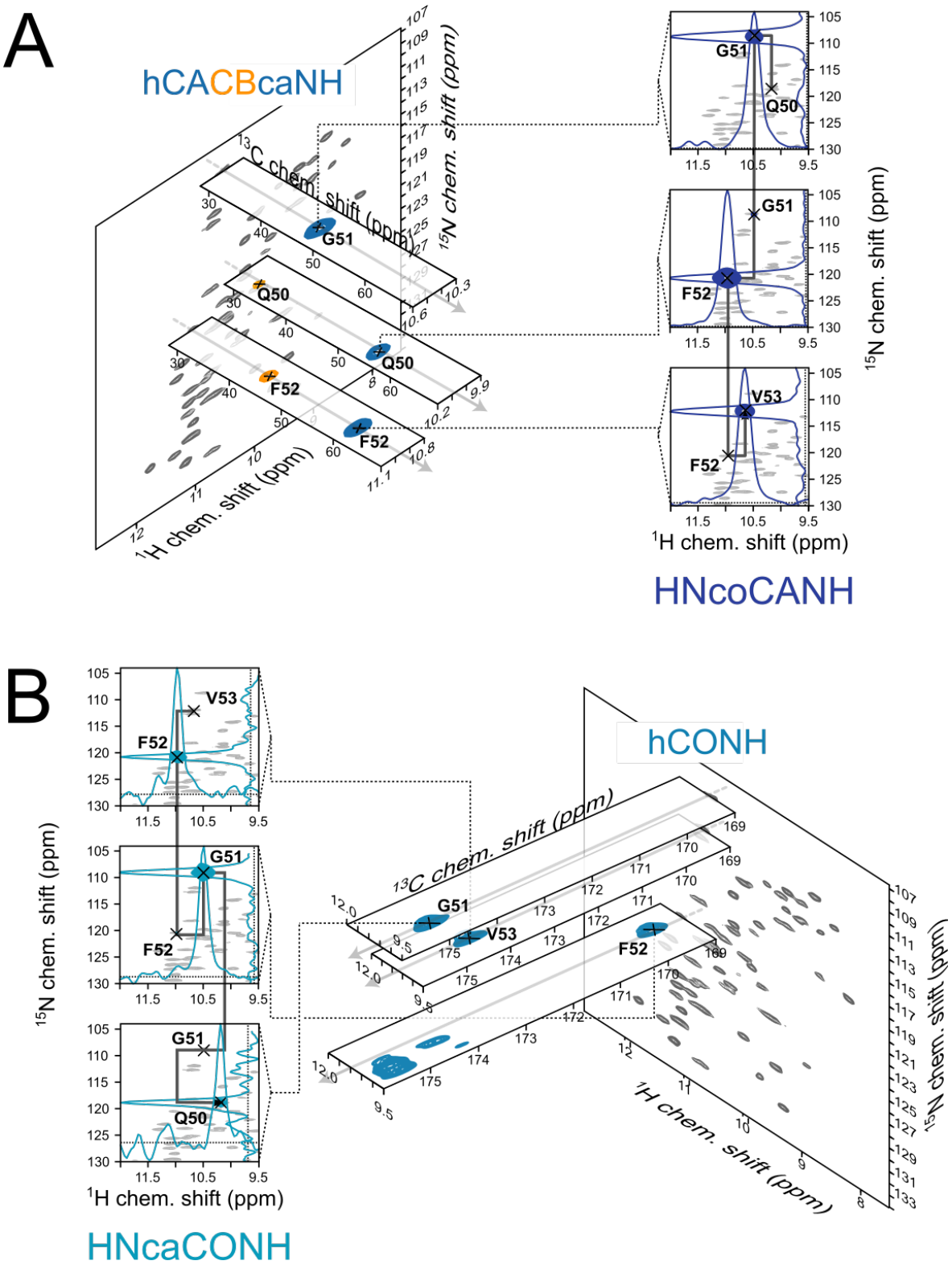


Figure 3.8: Exemplary backbone walk from Q50 to V53 using **A)** the 5D HNcoCANH to assign in the direction of the C-terminus, here using a 3D hcaCBcaNH directly as a base experiment, and **B)** using the 5D HNcaCONH with a 3D hCONH experiment to assign in the direction of the N-terminus. Shown are strips of the 3D base experiments along their carbon dimension at their respective position in the hNH spectrum (gray), and for each strip, the 2D plane processed out of the 5D showing either the *i+1* amide shift pair or the *i-1* shift pair with respect to the base experiment. 2D planes of the 5D are shown with trances running through the peak and are superimposed onto the same 2D hNH to allow for a visual comparison of line widths and cross peak intensity (details see Fig.3.7).

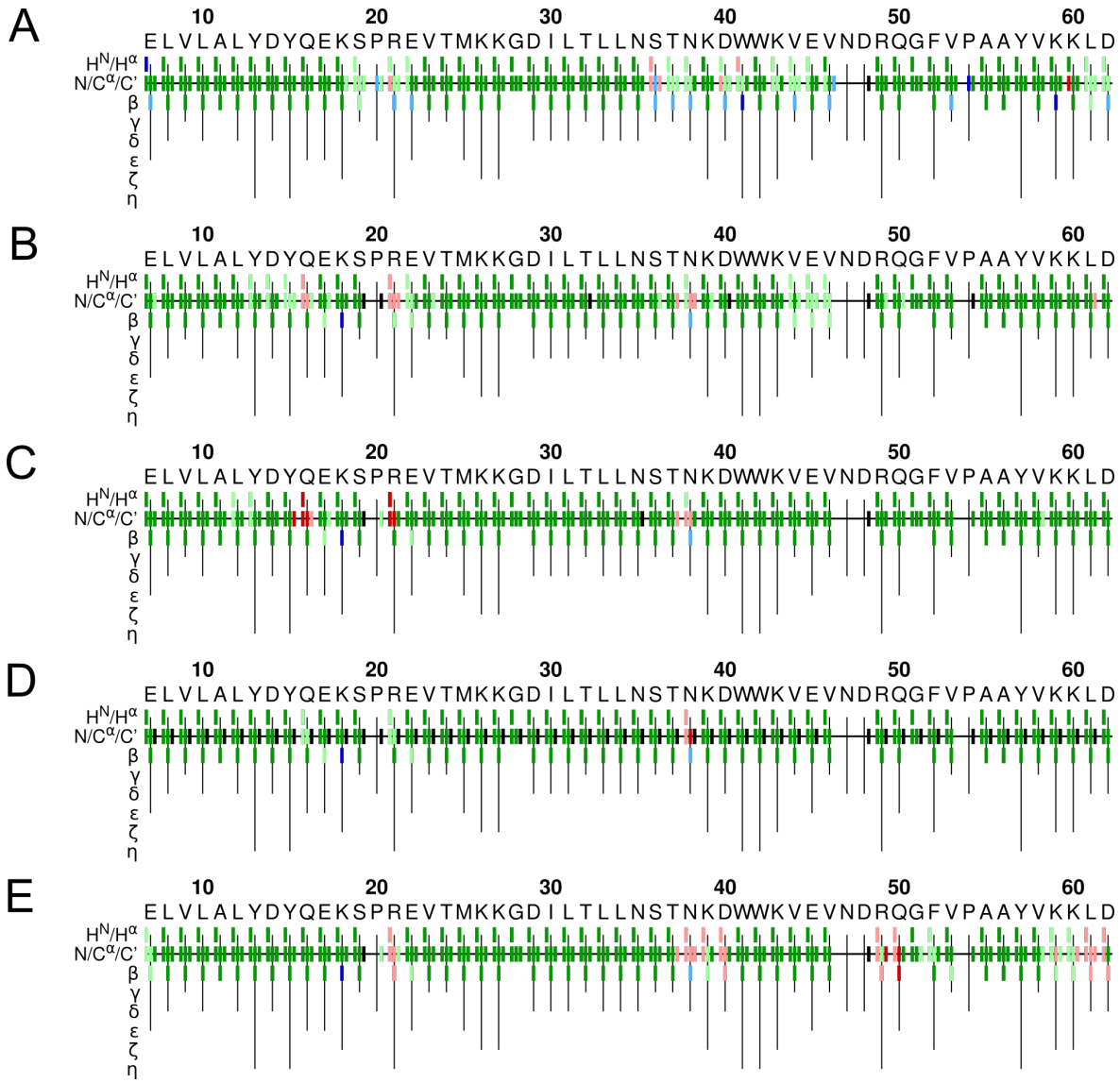


Figure 3.9: FLYA resonance assignment of the SH3 domain for a different combination of spectra. **A)** FLYA assignment based on a set of exclusively 3D based experiments, i.e. hCANH, hCAcoNH, hCONH, hCOcaNH, hCACBcaNH. **B)** A combination of the 5D HNcoCANH and the 5D HNcaCONH as well as a 3D hCACBcaNH for assigning Cβ resonances. **C)** Using a minimal set of experiments for full backbone assignment, i.e., only the better performing HNcoCANH together with a 3D hCACBcaNH and a 3D hCONH to provide CO chemical shifts were used. The minimal set yields essentially the same results as the combination of the two 5Ds. **D)** Using the minimal set without the hCONH potentially reduces the required input even further for the price of CO chemical shift assignment. This set gives the best FLYA results. **E)** Similar to **C)** but using the 5D HNcaCONH for the linkage of adjacent residues. As chemical shift redundancy is heavily reduced in this setup, the assignment quality is compromised compared to all other approaches. All combined chemical shifts were obtained from 100 independent FLYA runs using tolerances of 0.05 ppm for ^{1}H and 0.5 ppm for ^{13}C and ^{15}N. For all sets, an hNH peak list was provided additionally. Resonances determined as *strong* by FLYA are marked in dark colors (green: in agreement with reference chem. shifts; blue: no reference; red: different from reference beyond the defined tolerances; black: reference available but not assignable). N-terminal residues up to E7 as well as N47 and D48 are absent in CP-based experiments due to their inherent flexibility, and therefore inefficient magnetization transfers[57].

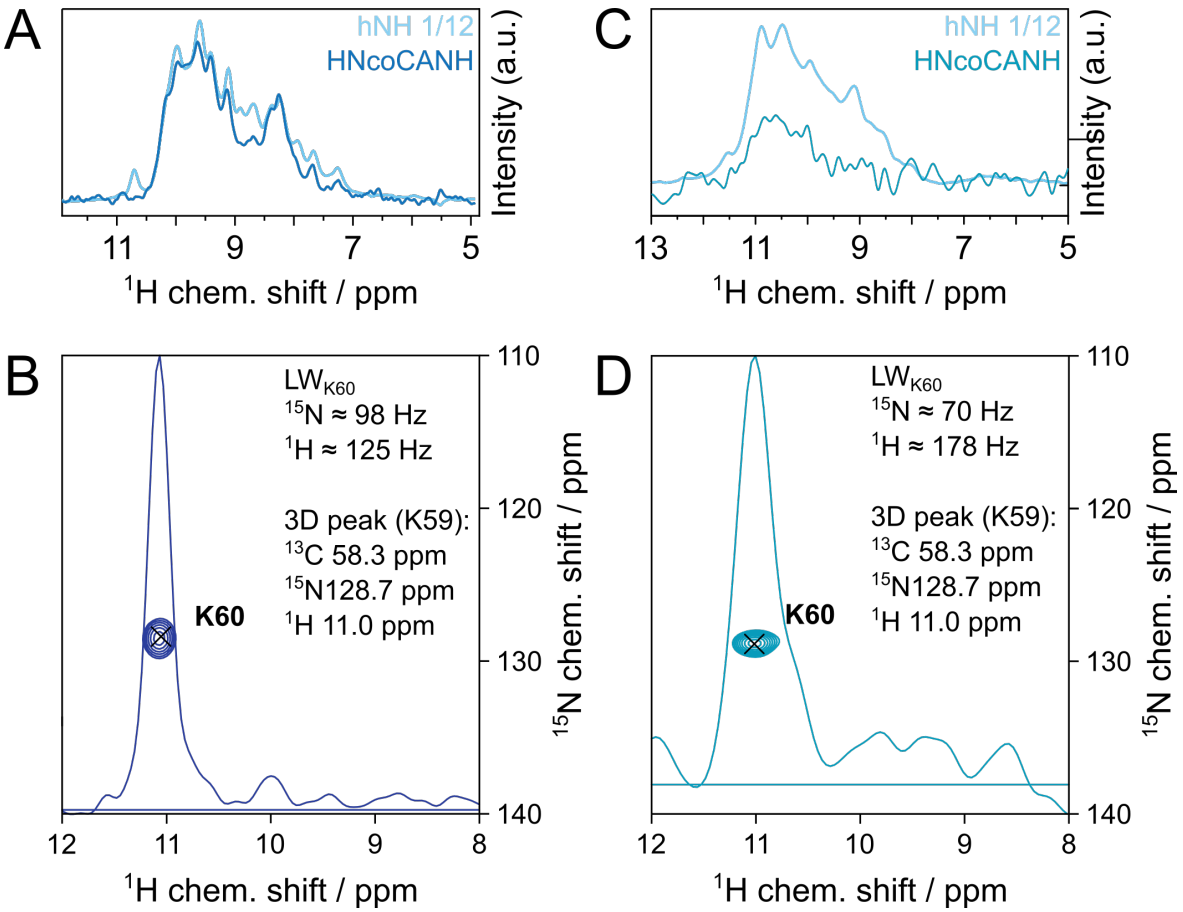


Figure 3.10: Sensitivity of the 5D HNcoCANH when **A),B)** acquired on a deuterated and 100 % back-exchanged sample for 6 d (dark blue) or **C), D)** on a fully protonated sample for 8 d (light blue). On the deuterated sample, the 5D HNcoCANH yields 8 % of the hNH sensitivity (**A)**) whereas on the fully protonated sample, only 4 % can be reached (**C)**), i.e., half of the sensitivity. The difference in sensitivity also becomes visible in the lower intense cross-peaks (**C), D)**). As expected, the ^{1}H line width increases on the fully protonated sample. The decreased ^{15}N line-width results from different ^{15}N $t_1 max$ times in the indirect dimension. Line widths were obtained from CCPN 3[193] after subtracting 50 Hz of LB in both dimensions of the deuterated sample as well as 100 Hz and 50 Hz for ^{1}H and ^{15}N respectively in case of the protonated sample.

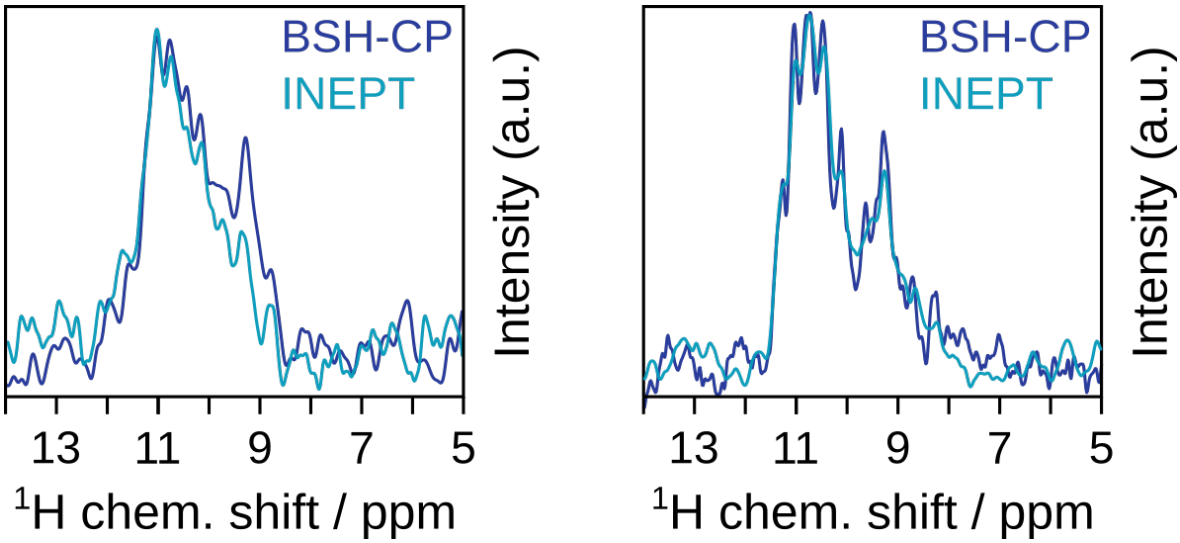


Figure 3.11: Comparison of H/N bulk intensity in the first FID of the HNcoCANH experiment using BSH-CP[26] (dark blue) or INEPT[5,38,56,194](teal) at 100 kHz acquired on a fully protonated sample (**A**)) and 55 kHz on a triple labeled, 100 % back-exchanged sample of the SH3-domain.

3D hCANH base spectrum, i.e., the *i* residue can be assigned from the 3D base when the *i+1* residue is known from other combinations. Therefore, only three residues less, D62, L33, and L8, cannot be assigned from the 5D HNcoCANH on a fully protonated sample, demonstrating that the sequence is generally applicable to such samples, probably without a reduction of assignments if the intrinsically lower sensitivity is compensated through extended measurement time, i.e., four times longer, or higher fields.

BSH-CP and INEPT versions of the HNcoCANH experiment

With increasing spinning speeds at high protonation levels, INEPT transfers become more efficient owned to the better averaging of dipolar interactions, and especially C→C transfers efficiencies are in close proximity[38,44,56,58]. At the same time, CP efficiencies decrease for the same reasons. An alternative implementation of the 5D HNcoCANH is therefore proposed in which the CO→Cα BSH-CP[26] is replaced by a CO-Cα INEPT while all other transfers are kept as CP as demonstrated previously[44,58] in lower-dimensional experiments. In the case of fully protonated SH3, INEPT and BSH-CP show no difference in sensitivity as seen from the bulk signal of the first FID (Fig.3.11A). It appears, however, that the efficiency is not the same for all residues, e.g., around 9 ppm, such that even at 100 kHz, BSH-CP provides a higher, more homogeneous performance. At 55 kHz MAS and applied to triple labeled, 100 % back-exchanged SH3 Fig.3.11B), similar results are found, but the difference is less pronounced such that either BSH-CP and INEPT appear to be equally well suited. In all cases, however, the shorter BSH-CP might be beneficial to minimize relaxation losses when long acquisition times are desired.

3.2.2 Full side-chain assignment through higher-dimensional experiments

This section is partially published[195] as: Klein, A., Vasa, S.K., Söldner, B., Grohe, K., Linser, R. Unambiguous Side-Chain Assignments for Solid-State NMR Structure Elucidation of Nondeuterated Proteins via a Combined 5D/4D Side-Chain-to-Backbone Experiment. *J.*

Phys. Chem. Lett. **13**, 1644–1651, doi:10.1021/acs.jpclett.1c04075

The list of assigned chemical shifts can be found in BMRB entry 51228.

Introduction

Correct and complete backbone assignment is the first and one of the most crucial steps for a deeper understanding of proteins studied by NMR as it enables structure calculation (in combination with NOESY or RFDR experiments) and relaxation studies[12,50,64]. The absence of side-chain assignments, proton resonances in particular, however, limit the accuracy of structural insights and can be only partially overcome by special labeling schemes such as methyl-labels[48,73] or random adjoint protonation (RAP)[43,125], each posing their own limitations regarding, e.g., resonance assignment in the former and sensitivity in the latter case. Nevertheless, full side-chain assignments are highly sought as they grant more detailed insights into the structure as well as protein-ligand interactions of the protein but require an additional proton source in the side chain for side-chain proton chemical shifts[54,196]. While Methyl-labelling relies on expensive precursors and makes assignment difficult and RAP labelling suffers from reduced sensitivity, fully protonated sample preparations pose the potentially best way protein samples. Owed to MAS frequency beyond 100 kHz such studies are nowadays feasible[12,13,52,54] and provide a plethora of information for structural studies in combination with minimal sample amounts[12,51,54,75] and no special labeling schemes, also making more delicate protein preparations approachable. Despite high sensitivities usually obtained for side-chain experiments of fully protonated proteins through the abundant proton bath, unambiguous resonance assignment is demanding even for small proteins such as the SH3 domain, due to significantly reduced redundancy compared to backbone assignment experiments as well as high spectral complexity, i.e., resonance overlap and ambiguity, resulting from the limited dispersion of ^{1}H and ^{13}C nuclei as well as homologous chemical shifts for most side-chain atoms irrespective of the amino acid type.

As high transfer efficiencies in solid-state NMR are maintained independent of the molecular weight, expanding and combining existing side-chain correlation experiments and backbone experiments to higher dimensional versions presents the possibility to disentangle ambiguous side-chain resonances and link them unequivocally to the protein backbone for a prospectively broad range of proteins. A 5D HCC(α)NH side-chain-to-backbone (S2B) experiment and its lower-dimensional counterpart, 4D HCNH, were implemented in the first place to allow easy linkage of the side-chain to the backbone resonances. In a second step, a combined 5D HCC(α)NH/4D HCCH S2B/HCCH-TOCSY experiment is introduced, which exploits orphaned magnetization on the carbon nuclei to fill in missing resonances in the 5D experiment at no cost of measurement time.

5D HCCNH and 4D HCNH S2B experiments

To assess the sensitivity and performance of a higher-dimensional S2B experiment before implementing a combined experiment to use orphaned magnetization, the 5D HCCNH was implemented as a straight-through pulse sequence starting from the side-chain protons and ending on the amide protons. Magnetization is transferred by CP transfer mostly from aliphatic protons to aliphatic carbons followed by C→C mixing and consecutive CP from Cα

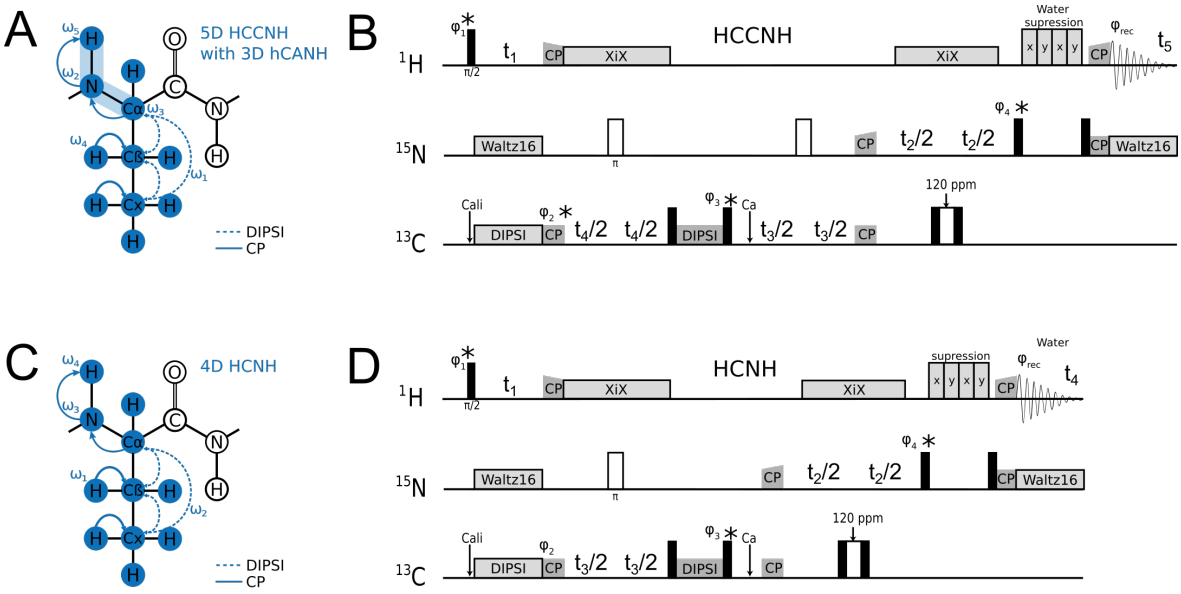


Figure 3.12: Magnetization transfers and pulse sequence **A,B)** for the 5D HCCNH and **C,D)** for the 4D HCNH S2B experiment. All transfers are CP-based, except C→C mixing, which was achieved through DIPSI-3[198,199]. In the case of the 4D HCNH, the Cα evolution period is omitted. Phase cycling: ϕ_1 = x,-x; ϕ_2 = x,x,-x,-x; ϕ_3 = -y,-y,-y,-y,y,y,y,y; ϕ_{rec} = x,-x,-x,x,-x,x,x,-x. Phases after evolution periods of indirect dimensions were changed to obtain phase-sensitive spectra by the States-TPPI method[173] (marked by asterisks). Filled and empty rectangles represent 90° and 180° pulses. The carrier changes on their carbon channel are denoted with arrows. Water suppression was achieved by two blocks of MISSISSIPPI without gradients[191] right before every acquisition block.

to nitrogen and ultimately to the amide protons for detection (Fig.3.12). A lower-dimensional 4D version was build from the same sequence, omitting the Cα evolution period to increase sensitivity for proteins with sufficient separation of H/N resonances (Fig.3.12C, D). CP transfer was employed for all heteronuclear transfers while homonuclear C→C mixing can generally be obtained from a range of either recoupling sequences[25,28,197] or sequences used for broadband isotropic mixing, e.g., DIPSI or WALTZ[198]. The results presented here were obtained through isotropic mixing using rotor-synchronized DIPSI-3[198](Fig. 3.12B, D) for 19.56 ms with a total pulse length of 25 µs at a RF field strength of 10 kHz as a compromise between mixing time and emitted heat into the rotor.

The H/N bulk signal of the first FID of the 5D HCCNH experiment accounts for 10 % of the respective FID in a 2D hNH experiment recorded at 800 MHz ^{1}H Larmor frequency (Fig.3.13A), whereas the 4D HCNH reaches up to 7 % of the hNH intensity recorded at 700 MHz ^{1}H Larmor frequency, despite the same number of transfer steps in both experiments. The difference of 3 % is therefore likely originating from slightly different experimental conditions in the two spectrometers such as RF inhomogeneities. Nevertheless, the sensitivity for both experiments can be considered in the same range as found, e.g., for hcaCBcaNH experiments on deuterated preparations (Fig.3.25, section 3.3) but providing a significantly higher impact on the assignment process. A more detailed description of the potential sensitivity losses in the HCCNH experiment is given in section 3.2.2.

To obtain mostly pure one-bond correlations, especially for the initial C to H transfer, CP contact times should be as short as possible while ensuring high magnetization transfer at the same time. For the presented results, a mixing time of around 350 µs was used as a compromise (Fig.3.14).

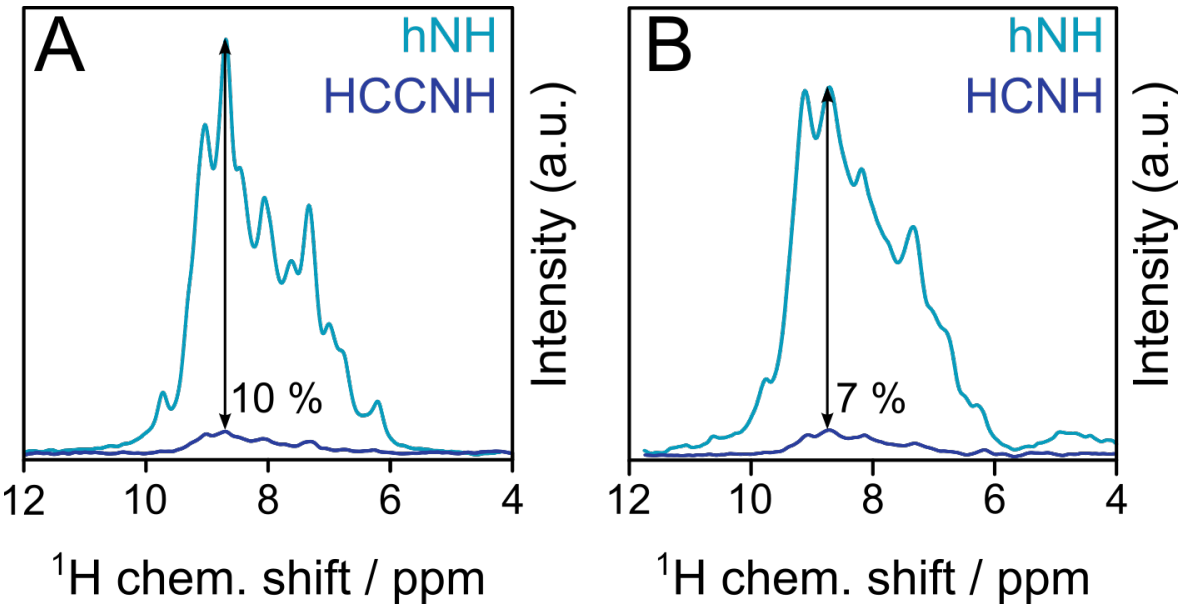


Figure 3.13: Sensitivity of the implemented 5D HCCNH (**A**)) and 4D HCNH experiment (**B**)) with respect to a 2D hNH (light blue). Shown are the H/N bulk signals of the first FID. Both experiments consist of the same magnetization pathways, as only the Cα evolution period is omitted in the 4D HCNH and should therefore show the same sensitivity. The differences probably occur from slight differences in the spectrometer setup, e.g., RF inhomogeneities. A more detailed description of potential sensitivity losses in the HCCNH experiment is given in section 3.2.2. The 5D was recorded at 800 MHz ^{1}H Larmor frequency with 1280 NUS points and 32 scans in 3 days and 6 hours, while the 4D was recorded at 700 MHz with 4096 NUS points and 16 scans in approximately 4 days and 6 hours.

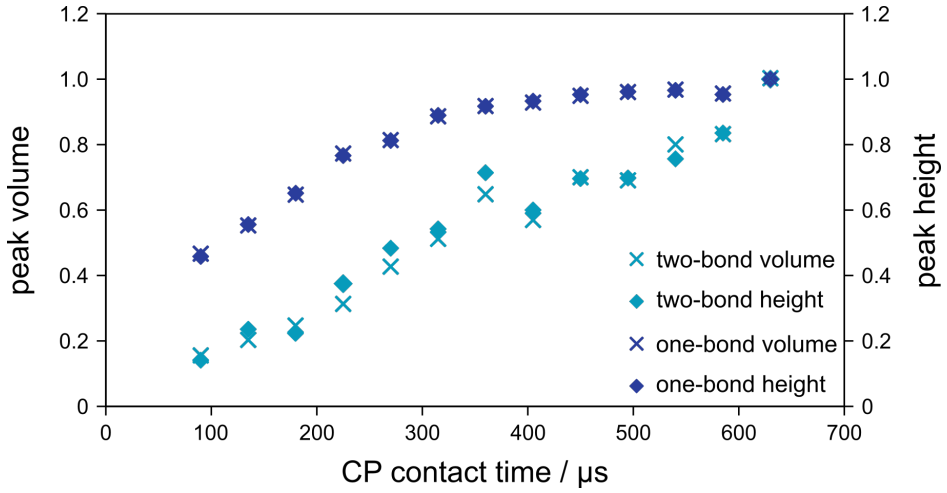


Figure 3.14: $H_{ali} \rightarrow C_{ali}$ build-up curves in relation to the CP contact time of the initial magnetization transfer given as peak intensities (squares) and peak volumes (crosses) for the one-bond transfer (dark blue) and the closes neighbor (light blue). Both curves were independently normalized to 1 for their respective highest value.

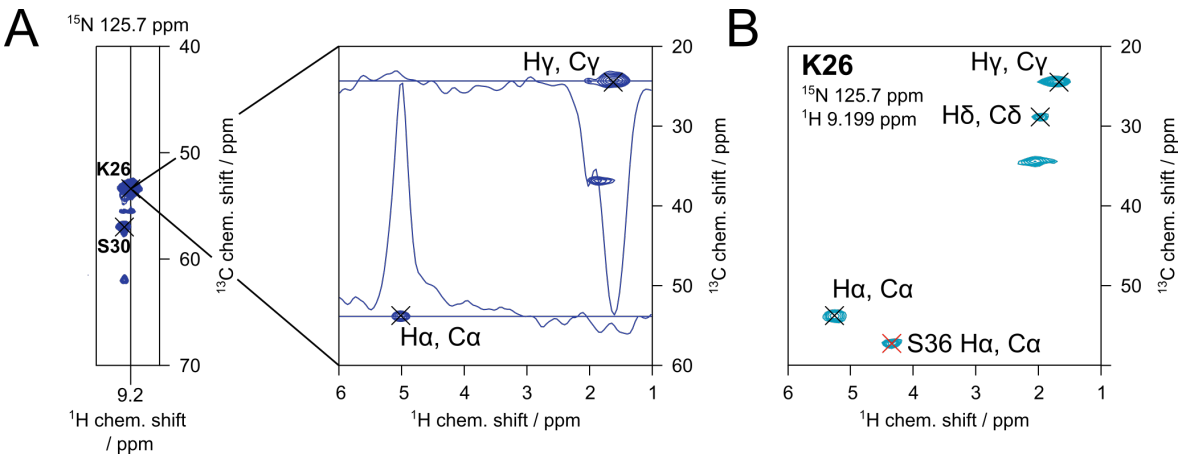


Figure 3.15: Exemplary excerpts of the spectra obtained through **A)** the 5D HCCNH and **B)** the 4D HCNH experiment. In both cases, the side-chain resonances of K26 are shown. For the 5D, the H/N/Cα is provided through the 3D hCANH base experiment and thus the peak list, while the coordinates in the 4D rely entirely on the 2D hNH with sufficient resolution and might be prone to ambiguities. This is exemplarily shown for the Cα peak of S36 adjacent to K26. However, the 4D allows for easier handling and better sensitivity, i.e., not necessarily improved SNR but eventually more side-chain correlations. Despite additional relaxation losses, the cross-peaks in the 5D HCCNH show high intensity. The 5D was recorded at 800 MHz ^{1}H Larmor frequency with 1280 NUS points and 32 scans, while the 4D was recorded at 700 MHz with 4096 NUS points and 16 scans.

The resulting spectra can be straightforwardly processed using SSA[102,104] in the case of the 5D in combination with a 3D hCANH base experiment and any NUS reconstruction algorithm feasible of reconstructing 4D data sets. Here, also SSA was used[105]. By choosing the hCANH as a base for the 5D, every Fourier-transformed 2D plane will show all $H_{ali}/_{ali}$ shift pairs of the side-chain of the residue defined by the hCANH (3.15A), which will most likely result in unique chemical shift information even for large proteins (see section 3.1) and therefore grants a straightforward resonance assignment. If sufficient resolution in the H/N shift pairs can be assumed, the fully reconstructed 4D experiment can be treated in essentially the same way. However, only a 2D H/N correlation can be used to display side-chain correlations (3.15B), which will likely lead to an ambiguous resonance assignment of the side-chain even for proteins as small as the SH3-domain, for which, e.g., residues K26 and S36 are not perfectly resolved in the 2D hNH (Fig.3.15B).

High initial magnetization through the abundant proton bath combined with merely five transfer steps, including C→C mixing, results in spectra with relatively high signal intensity through the entire side chain. The cross-peak intensity depends on the number of attached protons on every side-chain carbon as well as the broadband mixing. Considering that most aliphatic carbons carry either two or three protons whose magnetization is transferred to the bound carbon atom leads to rather uniform intensity distributions of the side-chain resonances even for distant γ or δ-groups when using broadband mixing (Fig.3.15A). The lower-dimensional HCNH experiment can expectantly be recorded within a shorter period of time, or alternatively, for the same time to improve the signal intensity of weak peaks, albeit bearing the risk of ambiguous resonance assignments. Time-equivalent acquisition, as done for the data presented here, will likely increase the chance of weak peaks to become visible and therefore facilitates higher rates of assignment, as can been seen exemplarily for the side-chain resonances of K26 (Fig.3.15), for which the 4D HCNH shows an additional Hδ,Cδ peak but simultaneously also parts of the Hα,Cα peak of S36 (Fig.3.15B). In sum-

mary, the 5D HCCNH S2B experiment shows high cross-peak intensity in general, while not all side-chain resonances are always found in the respective 2D planes. This drawback can be partially overcome by the lower-dimensional counterpart 4D HCNH, however, at risk of ambiguities. Nevertheless, the results are encouraging for the implementation of the actually intended combined 5D HCCNH/4D HCCH experiment, as no losses are expected for the magnetization pathway of the 5D experiment.

The combined 5D HCCNH/4D HCCH experiment

The concept of the combined 5D HCCNH/4D HCCH experiment is based on the recovery of orphaned magnetization from the hetero-nuclei, which was recently shown in solid-state NMR experiments at high MAS rates[65,199] and generally on concepts of separation of different FIDs acquired on the same detection-nucleus, i.e., ^{1}H. The combination of two experiments detected on the same nucleus within one is widely used in NMR through the concept of time-shared experiments[200–202] and IPAP experiments for virtual decoupling for which separation of the experiments is often implemented by sign-inversion of either of the experiments and subsequent linear combinations[203,204]. Alternatively, sequential acquisition of two separate magnetization pathways can be facilitated if the two (or more) experiments do not have to be detected at the exact same time on the same nucleus and are distinguished by certain features such as their hetero-nucleus. By that, it is possible, e.g., to detect HN and Hα protons independently from each other[65].

For the combined 5D/4D, the encoding of the two different FIDs was implemented by acquiring and storing the two FIDs in two different acquisition buffers, NBL buffers, in a similar fashion as other approaches using sequential acquisition[65] rather than sign-inversion and linear combination. Therefore, separating all FIDs belonging to one set can be easily achieved within Topspin by the `split nbl` command. Nevertheless, slight modifications in the setup of the experiment and the pulse sequence, apart from actual magnetization transfers, have to be done compared to the basic version of the previous section. Before the FIDs are actually separated, the overall size of the entire FID is two times what is expected based on all points in the indirect dimension, and a dummy loop or dimension has to be introduced to the pulse program, effectively doubling the number of TD_{ind} points, to avoid downstream errors in Topspin.

Listing 3.1: Pulse program modifications for sequential acquisition

```
...
1m st0                  ; setting the first NBL buffer
...
0.5u pl22:f2            ; 15N decoupling
0.5u cpd2:f2
goscnp ph31             ; acquire HN and do nothing else
1u do:f2                ; turn off decoupling
...
1m st                   ; switch to next NBL buffer
...
0.5u pl23:f3            ; 13C decoupling
```

```
0.5u cpd3:f3
go=2 ph31                ; acquire Hali
                         ; reset scans counter and loop

1u do:f3                 ; turn off decoupling

d11 mc #0 to 2           ; data to disk, etc.

F1PH(calph(ph11, +90), caldel(d0, +in0))
F2PH(calph(ph7, +90), caldel(d10, +in10))
F3PH(calph(ph17, -90), caldel(d20, +in20))
F4PH(calph(ph5, +90), caldel(d21, +in21))
F5QF()                   ; dummy dimension to account for
                         ; doubled FID size
```

This, however, adds an additional, although empty, dimension and the experiment has to be set up as a 6D experiment with two increments in `F5QF` dimension. Furthermore, only the `goscnp` command should be used rather than `go` to acquire the first FID, as it prevents the scan counter from resetting as well as looping back to the beginning of the sequence[205]. No other changes have to be done, and the experiment can be recorded as any other. After splitting the individual FIDs, the 5D HCCNH data set can consecutively be processed with SSA in combination with a 3D hCANH peak list without any alterations, resulting in a stack of 2D planes, with each plane showing all HC pairs for each hCANH peak (Fig.3.18) as in the basic version of the experiment introduced in the previous section. As the magnetization pathway of the 5D is unaffected through the incorporation of the 4D HCCH, the sensitivity of the H/N bulk signal remains essentially unaffected at 8 % in the first FID with respect to the hNH intensity (Fig.3.17A), which is less than half of the expected 21 % from theoretical approximations[58] when considering a similar CP efficiency for C→H and N→H CP transfers and no relaxation losses during the mixing period. For the separately stored FID acquired from aliphatic protons, 7 % of the hCH intensity are reached in the first FID (Fig.3.17B), showing lower sensitivity of what could be assumed from theoretical values[58] for what can be understood as an hCcH experiment with a C→C mixing period if no relaxation is acting during mixing. As ^{13}C T_1 times are comparably long also in fully protonated proteins at high spinning speeds[206] only little loss is expected to occur from the storage along the z-axis, and the main source of reduced sensitivity is most likely the little remaining magnetization on the carbon atoms. As the 4D HCCH is, however, effectively recorded with twice the number of scans compared to the 5D due to post-acquisition processing (see below), these losses can be partially compensated, and the gain in intensity is approximately $2\sqrt{2}$. A more detailed comparison of transfer efficiencies shows that approximately 20 % are lost during the z-storage, probably enhanced through PRE effects, as obtained from comparison of bulk signals from a normal hCH and an hCH containing such an extra z-storage for 90 ms, corresponding to ^{15}N $t_{1,max}$ in the combined experiment of 10 ms and 80 ms of water suppression. Another 40 % are lost during the isotropic mixing period as is apparent from the bulk intensities of the hCH and an HCcH experiment containing a C→C mixing period. In combination with approximately 40 % of the Cα magnetization being transferred to the

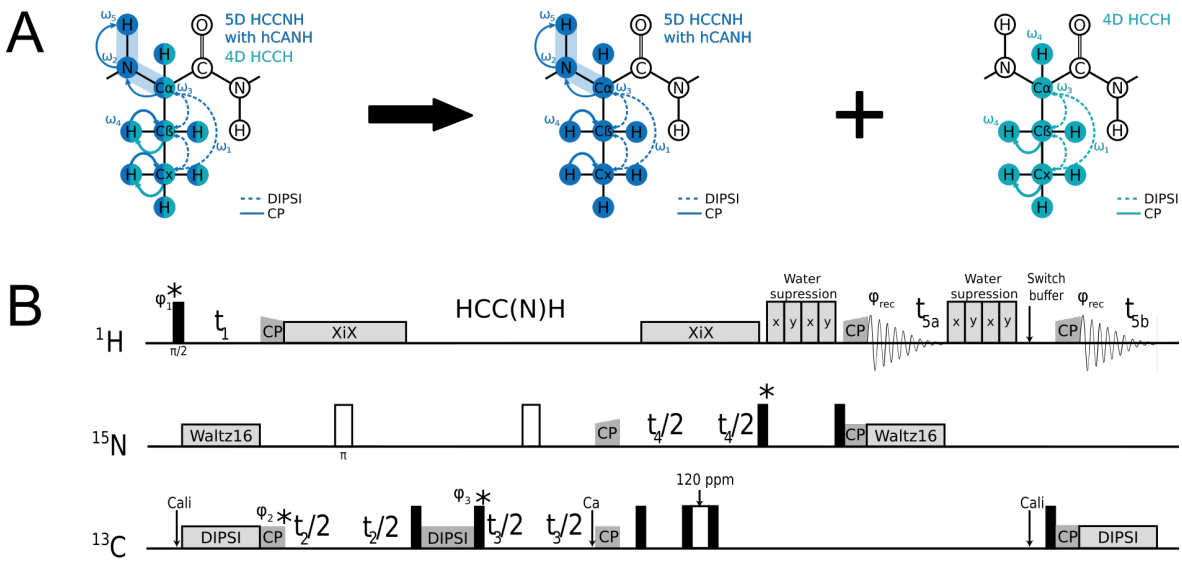


Figure 3.16: Pulse sequence for the combined 5D HCCNH / 4D HCCH experiment (left) and magnetization pathway (right). All transfers are CP-based and isotropic C→C mixing is performed with the DIPSI-3[198,199] scheme. After magnetization transfer from C to N, the carbon magnetization is stored along the z-axis during N to H transfer and acquisition and afterwards transferred back to (aliphatic) protons. Therefore, in one pathway, the nitrogen evolution is omitted, and a 4D HCCH-TOCSY is obtained, while in the other case, the magnetization is transferred to the backbone amide group yielding a 5D HCCNH S2B experiment, however, without correlating the individual HC groups in the side-chain which each other. Using a 3D hCANH as the base for the 5D, 2D planes will ultimately contain one peak for every HC group in the side-chain. Phase cycling: ϕ_1 = x,-x; ϕ_2 = x,x,-x,-x; ϕ_3 = -y,-y,-y,-y,y,y,y,y; ϕ_{rec} = x,-x,-x,x,-x,x,x,-x. Phases after evolution periods of indirect dimensions were changed to obtain phase-sensitive spectra by the States-TPPI method[173] (marked by asterisks). Filled and empty rectangles represent 90° and 180° pulses. The carrier changes on their carbon channel are denoted with arrows. Water suppression was achieved by two blocks of MISSISSIPPI without gradients[191] right before every acquisition block. The two different FIDs are stored to different NBL buffers and, after the acquisition, separated into two different data sets.

nitrogen atom[58], and attenuation down to approximately 29 % is expected for the HCCH over the hCH. Considering the losses during isotropic mixing, the expected sensitivity of the 5D HCCNH is lowered to around 13 % of the hNH bulk signal such that the experimentally obtained results are well in line with the expected results, whereas the HCCH ultimately yields only 1/4 of the expected intensity.

By storing the remaining carbon magnetization on the z-axis after Cα to N transfer and a final back-transfer to aliphatic protons, a 4D HCCH-TOCSY type of experiment is obtained alongside as the nitrogen evolution period is omitted. In contrast to the 5D, which ideally only shows one-bond correlations of each H_XC group with the backbone, the 4D HCCH-TOCSY also shows correlations among the different groups within the side chain. Depending on the view, either HH-TOCSY patterns can be seen for two given carbon chemical shifts or vice versa or correlations of any given HC shift pair with other HC pairs which resemble the same pattern obtained for the 5D HCCNH experiment. The underlying pathways were not altered with respect to the transfer mechanisms used, and still, all magnetization pathways are based on CP transfers, whereas the DIPSI-3 scheme[199] was used for isotropic C→C mixing. Despite long mixing times at an RF field strength of 10 kHz, aromatic carbons, e.g., from tyrosine or histidine side chains, remain absent, essentially independent of the carbon carrier frequency. The combined experiment was recorded in two blocks of 2 days and 20 hours with each block consisting of 8 scans and the same 2304 NUS points, which were ultimately added up to 16

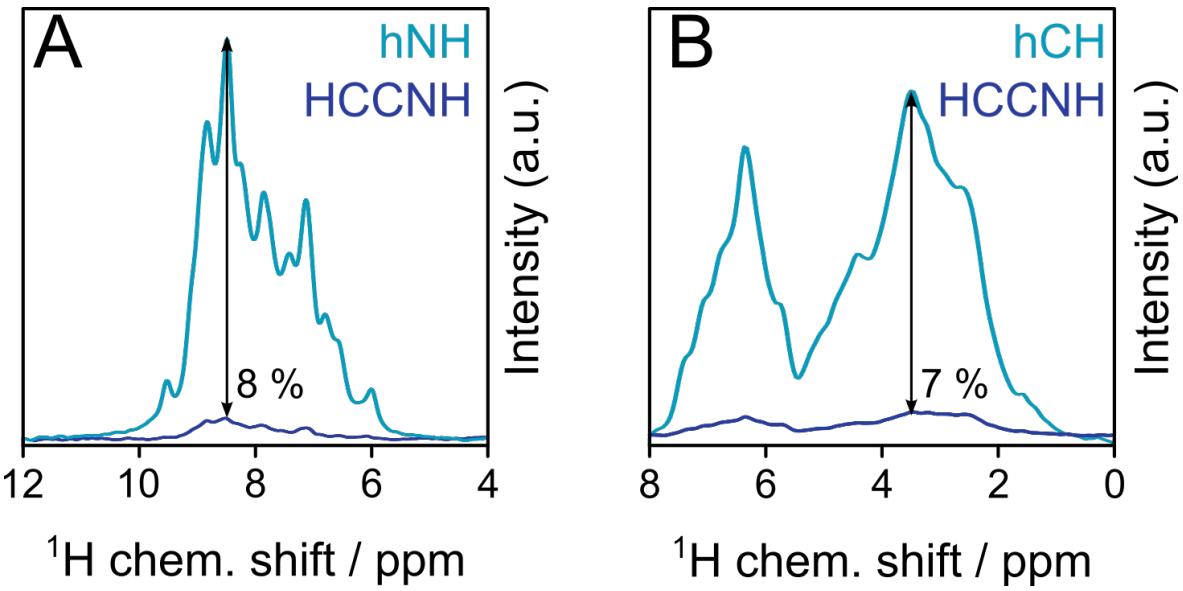


Figure 3.17: H/N and H/C bulk sensitivities of the first FID were recorded from the 5D HCCNH sequence after splitting the FIDs of different buffers with respect to either bulk signals from hNH or hCH experiments. For the amide bulk signal, 8 % of the signal are conserved in the 5D, which is in line with values obtained for the basic version of the experiment in section 3.2.2 as well as its 4D HCNH counterpart and semi-empirical expectations[58] (see text). The FID resulting from recovered carbon magnetization and H_{ali} detection, on the contrary, yields only 7 % of the hCH signal and likely results from little recovered magnetization on the carbon atoms as well as losses during the z-period and C→C mixing (see text). All spectra were recorded at 700 MHz ^{1}H Larmor frequency and scaled to the number of scans before comparison.

scans.

After the acquisition and splitting of the FIDs from the two different acquisition buffers, the data set obtained through detection of amide protons and corresponding to the 5D HCCNH can be processed without further treatment. Although the simultaneously acquired 4D HCCH does not contain any nitrogen evolution period, it is acquired as a 5D data set due to the nature of the pulse sequence and therefore requires additional data manipulation before processing. In a uniformly sampled experiment, the 4D experiment would be acquired j-times, with j being the points in the indirect dimension, and consequently would be obtained as the sum of j individual 4Ds. In the non-uniformly sampled case, however, the sparsity of the five-dimensional NUS schedule is so high that combinations of the same time increment along the four HCCH dimensions are rare, and summation of those FIDs will lead to uneven sampling within the 4D subset. Although this problem might be more of theoretical nature and with little impact on the final spectrum, the pre-processing procedure presented here will delete those FIDs (Fig. 3.18) in which all time increments except the nitrogen ones are identical to obtain a clean, not-scan weighted FID. As a consequence, the number of NUS points for the 4D HCCH will be slightly reduced compared to the 5D HCCNH, which, however, will likely result in an insignificantly lower sparsity. If the dimensions of the presented pulse sequence are lowered, e.g., a 4D HCNH/3D HCH pair, the pursued approach of deleting redundant FIDs most likely will lead to more pronounced reductions of the NUS schedule for the lower-dimensional part, which can have a significant impact on the spectral quality and it would be advisable to create a redundancy-free schedule from the beginning. While no F4 time incrementation occurs for the 4D HCCH, the data set will contain the phase increments of the nitrogen indirect dimension, which, after deleting redundant time points,

can be eliminated by the addition of the corresponding x and y components, i.e., cosine and sine modulated FIDs, ultimately resulting in an FID of half the size but twice the scans. Recombination of FIDs into a 4D data set is done by a Python script using Python 3[174] with the *nmrglue* package[207] and *NumPy*[208]. The data obtained in this way can be straightforwardly used for NUS reconstruction with SSA or any other algorithm supporting 4D NUS reconstruction, e.g., hmsIST[177].

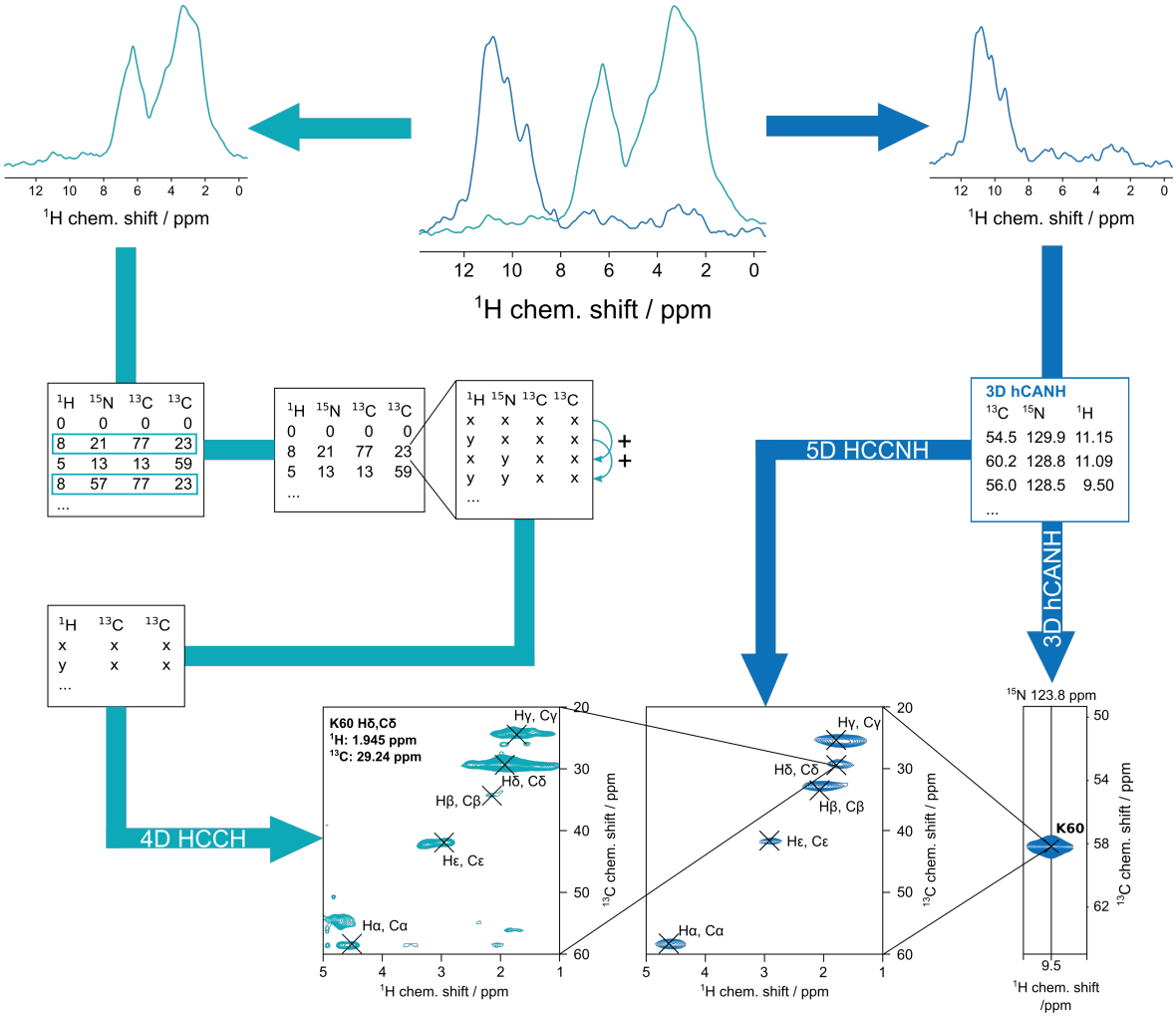


Figure 3.18: Schematic workflow of the processing for the combined 5D HCCNH (blue arrows) / 4D HCCH (teal arrows) experiment. After separating the HN detected FID from the H detected FID, the 5D can be straightforwardly processed using SSA with a 3D peak list based on a 3D hCANH. For every peak, a series of one-bond HC correlation of the respective side-chain is obtained as shown here exemplarily for residue K60 of SH3. The 4D set has to be treated by the procedure described in the text to ultimately reduce the dimensionality from 5D to 4D through the removal of redundant time increments and the addition of FIDs. The reconstructed spectrum will show correlations between all carbons and protons of the side-chain while omitting the link to the backbone like is exemplarily shown for the correlations found at the coordinates of HδCδ for K60.

Due to the properties of the 5D HCCNH experiment, the assignment of the individual side-chain peaks is straightforward, while the full assignment of the 4D HCCH-TOCSY poses a bigger and more time-consuming task, but in combination with a backbone experiment, e.g., 5D HNcoCANH and 3D hCONH, full resonance assignment can be obtained from only three experiments. The 5D provides superior dispersion of indistinct side-chain chemical shifts like they are often found for the same residue type but occasionally lacks the full resonance information of the side-chain. For example, HβCβ resonances for all threonines are not

found while generally being detectable in hcaCBcaNH experiments for no apparent reason. The 4D, on the other hand, suffers from severe overlap resulting from the before-mentioned reasons, but is a factor $\sqrt{2}$ times more sensitive, as the phase incrementation in the fifth dimension is omitted, and the scans per FID are doubled. This effectively allows to identify weak resonances that do not become apparent in the higher-dimensional counterpart, e.g., the HαCα resonance of K27 (Fig.3.19C). Whereas the 4D HCCH-TOCSY experiment alone only allows for highly ambiguous assignments or vice versa for an incomplete assignment, it can confirm and expand the assignments obtained from the 5D HCCNH in the first place. To exemplarily visualize how the two experiments work together, a plane from the 4D HCCH-TOCSY at the HδCδ resonance of K60 is compared to the planes from the 5D HCCNH of those five lysines (K18, K27, K39, K59, and K60) as well as Q16, which show either all or part of their side-chain resonances in the respective section of the 4D (Fig.3.19). Each of the 2D planes in the HCCNH show unambiguous side-chain assignments (Fig.3.19A), but when superimposed express a severe resonance overlap (Fig.3.19B) strongly resembling the 4D HCCH at the given coordinates. Therefore, those resonances cannot be assigned solely from the latter experiment despite the fact that the intensity of the cross-peaks in the 4D might hint towards multiple peaks being present (Fig.3.19C). On the other hand, the 4D additionally shows the HαCα resonance of K27, which cannot be found in the respective plane from the 5D experiment.

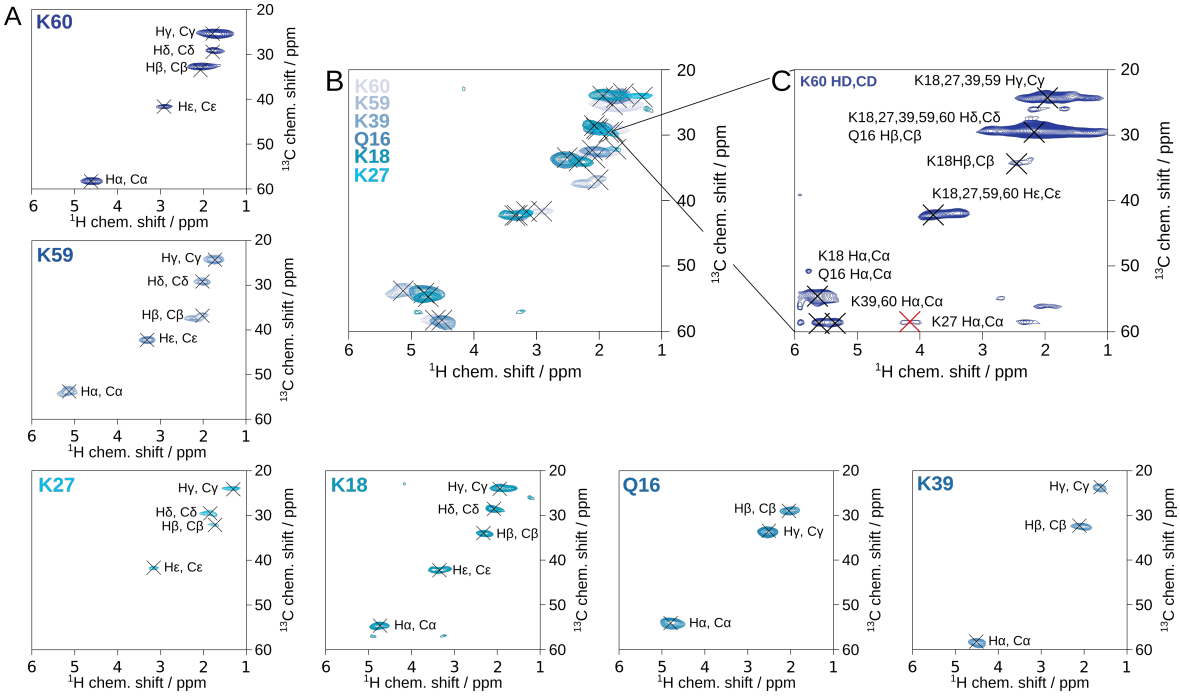


Figure 3.19: Handling and comparison of the 5D HCCNH and the 4D HCCH experiments obtained simultaneously. **A)** Individual planes from the 5D HCCNH, shown for five lysines (K18, K27, K39, K59, and K60) and Q16, enable unambiguous chemical shift assignment through the side-chain. **B)** Superimposition of these individual planes to emulate the resonance overlap that would be expected in the 4D HCCH. **C)** Resonances found at the HδCδ positions of K60 in the 4D HCCH spectrum. In addition to the 5D HCCNH, the 4D HCCH-TOCSY can provide additional, mostly weak, resonances. The combined experiment was recorded with 2304 NUS points and 16 scans in a total time of 5.6 d.

A further possible approach for full resonance assignment is to link the backbone resonances from the introduced 5D HNcoCANH (section 3.2) and its 3D hCANH base acquired from either a deuterated sample to save measurement time or a fully protonated sample as

well in case deuteration is not possible. The identification of the residue type can simply be made in the same step from the 5D HCCNH using the same base experiment, and those assignments can, in turn, be used in the 4D HCCH to identify potentially missing resonances in the 5D HCCNH experiment. Ultimately, CO resonances can be added by incorporation of a 3D hCONH, or 4D hCACONH or hCOCANH for high molecular weight proteins, if required. Following this approach, 89 % of all accessible resonances in the SH3-domain can be unambiguously assigned from merely three experiments within merely 8 d of measurement time or less, using a deuterated and 100 % back-exchanged sample for the 5D backbone experiment. If carbonyl carbons are not considered, 88 % of the expected resonances can be assigned from the two 5D experiments and their 3D hCANH base experiment. The possible resonances include the protein backbone with Hα and carbonyl carbons and the side-chain ^{1}H and ^{13}C nuclei, but exclude aromatic resonances as they cannot be reached through the facilitated mixing schemes, residues 7, 47 and 48[57] which cannot be detected even by a 2D hNH, as well as amide and carboxy side-chains. Line widths in ^{1}H–detected solid-state NMR of fully protonated proteins are on the order of 100 Hz[12,51,54], i.e., 0.14 ppm at 700 MHz, for aliphatic protons, rendering the distinction of chemically non-equivalent atoms of the same type, e.g., Hβ2 and Hβ3 often impossible. While in individual cases, two distinct resonances can be found in the spectrum, it was assumed for the general analysis that all protons of one carbon atom in the side chain are chemically equivalent and therefore have the same chemical shift. In total, 495 resonances can be expected from the backbone and side chains under the mentioned restrictions, of which 441 resonances were unambiguously assigned. Considering only potential side-chain resonances from the 5D HCCNH/4D HCCH, 287 out of 340 resonances, i.e., 84 %, can be unambiguously assigned. For most residues, a full or near-complete resonance assignment is possible (Fig. 3.20), and the missing resonances mostly account for such residues that are generally weak and hard to detect even in a normal 3D experiment like R21, as well as the N- and C-termini alongside a few isolated resonances (Fig.3.20 and table 3.1).

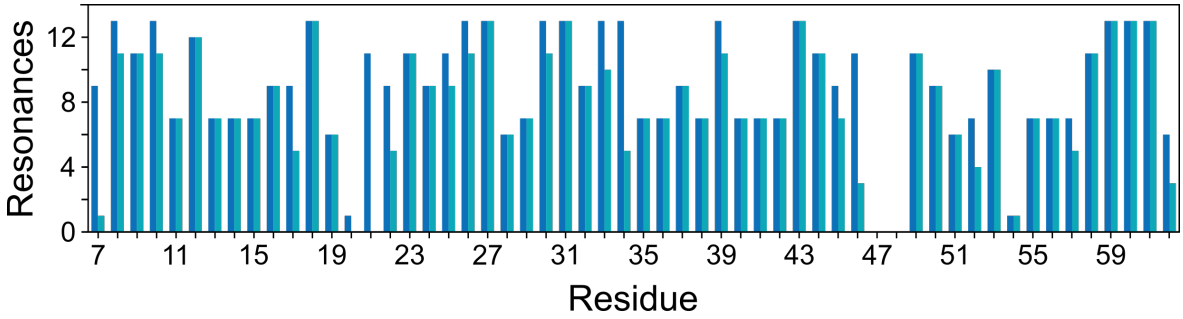


Figure 3.20: Expected (dark blue) and assigned resonances (light blue) of the SH3-domain considering the protein backbone and only chemically equivalent protons in the side-chain. Side-chain carboxyl and amides groups, as well as aromatic HC groups, are excluded as they can be identified neither through the 5D HCCNH nor the 5D HNcoCANH based on the current mixing schemes and magnetization pathways. Residues N47 and D48 are not considered due to their inherent flexibility[57].

When comparing the assignment rate of the combined 5D/4D experiment with the assignments obtained from the simpler lower-dimensional 4D HCNH, the 4D HCNH alone facilitates more assignments than the pure 5D experiment when both are recorded for the same time, and 356 and 315 out of 442 expected resonances can be assigned, corresponding to a rate of 80.5 % and 71.3 %. However, the 4D HCCH within the combined experiment effectively fills

Table 3.1: Unassigned resonances in the SH3-domain using a 5D HNcoCANH, 5D/4D HCCN-H/HCCH and 3D hCONH

Residue	missing resonances
L8	Cδ2, Hδ2
L10	Cδ1, Hδ1
E17	Cβ,γ, Hα,β,γ
P20	CO
R21	Cα,β,γ, Hα,β,γ
E22	Cβ,γ, Hα,β,γ
M25[a]	Cϵ, Hϵ
K26	Cδ, Hδ
I30	Cγ1, Hγ1
L34	C$\beta\gamma,\delta$1,δ2, H$\beta\gamma,\delta$1,δ2
K39	Cϵ, Hϵ
E45	Cβ, Hβ
V46	CO, Cβ,γ1,γ2, Hα,β,γ1,γ2
Y57	Cβ, Hβ
D62	CO, Cβ, Hα,β

[a] Cϵ might be excluded as well, as the terminal methyl group is not correlated with the rest of the side-chain

in missing resonances which in turn results in 393, i.e., 88.3 %, assigned resonances which can be considered a big improvement. Taking into account that the combined 5D/4D experiment supports improved match-making with the backbone through the added dimension and ultimately is the better performing experiment, it might be the experiment of choice even for small proteins such as SH3-domain. If sufficient resolution in the H/N plane is given, and sensitivity, time, or data handling, in the context of user friendliness, are crucial, it should be proposed here that also the 4D HCNH can be combined with a 4D HCCH which, however, would require more adjustments of the pulse sequence (Fig.3.21) as the second carbon evolution period has to be moved before the final C→H CP, i.e., right after the stored z-magnetization is flipped back. Phase incrementation can then be performed simultaneously for the ^{15}N and second ^{13}C dimension under the premise that both experiments share the same number of total incrementation steps. As the experiment will be set up as 4D with one dimension containing both a ^{15}N and ^{13}C dimension and to still allow different $t_1$max times for both nuclei, the increment time will have to be scaled by either $\frac{2}{5}$ or $\frac{5}{2}$ depending on the definition of this dimension as nitrogen or carbon dimension. The obtained spectra could, however, be used without further processing, but careful adjustment of the Cα →N CP might be necessary to increase the amount of recovered magnetization. Otherwise, the sensitivity of the 4D HCCH is likely to be compromised in comparison to the 5D/4D variant, as no recombination of FID is necessary to obtain the HCCH, and virtually only half of the scans are acquired, corresponding to a reduction of signal-to-noise by a factor of $\sqrt{2}$, which, however, is gained for the HCNH part of the experiment.

Similar to the backbone assignments of the deuterated SH3-domain, incorporating FLYA in the assignment strategy can simplify and speed up the overall process of resonance assignment. Reliable outcomes in the automated assignment of side chains are particularly interesting in the context of structure calculation and would pave the way to faster NMR protein structures as automated NOE assignment can be provided from the CYANA pack-

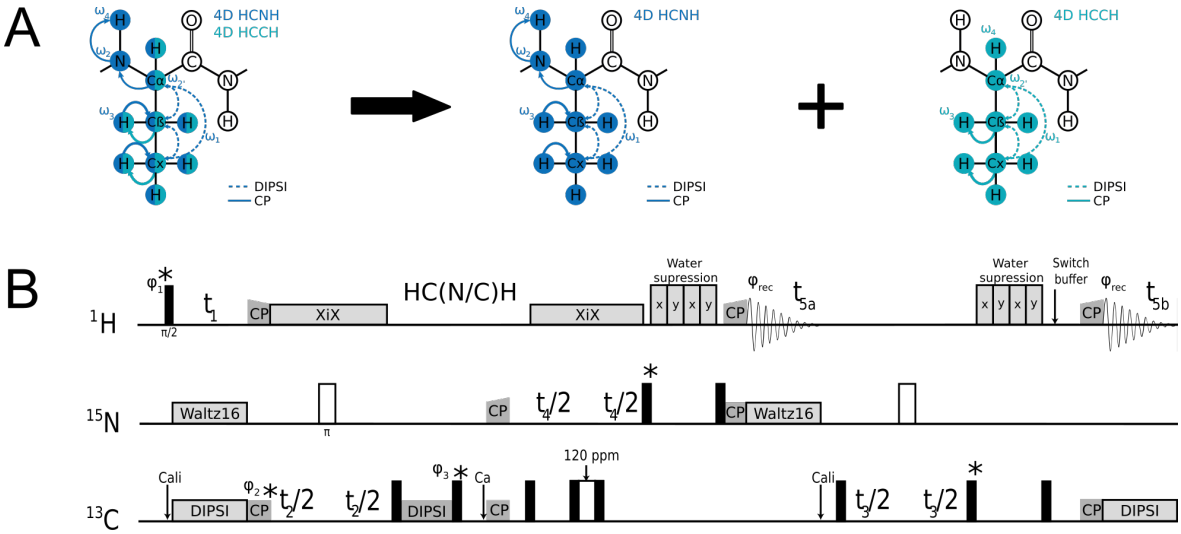


Figure 3.21: Proposed pulse sequence for a combined 4D HCNH/4D HCCH S2B/TOCSY experiment. The magnetization pathway remains unaltered compared to the 5D/4D variant, but the Cα evolution period is omitted in the case of the 4D HCNH S2B experiment, and instead, the second carbon evolution happens after the recovered carbon magnetization is flipped back to the x/y-plane. Optimal water suppression will require additional storage along the z-axis before the magnetization is ultimately transferred back to protons. As both experiments have the same dimensionality, phase incrementation can occur simultaneously, and individual acquisition times can be achieved through scaling of the increment size of the carbon dimension by $\frac{2}{5}$ of the ^{15}N step size as demonstrated previously[40,74].

age as well[209]. For full resonance assignment with FLYA, peak lists from a 5D HNcoCANH acquired on a deuterated sample were used in conjunction with peak lists from the combined 5D/4D HCCNH/HCCH, as well as a 3D hCANH list as the experiment is required as a base for the 5Ds and comes therefore with no additional costs. Additionally, an hNH experiment was incorporated, which allows an easy check of the backbone assignments and is usually available at a minimal cost of measurement time. The CYANA library was expanded by theoretical magnetization pathways for the 5D HCCNH (see section 7.2), whereas the 4D HCCH is already implemented. Similar to the manual assignment process, the default correlations expected by FLYA for the side-chain experiments were modified such that only one proton resonance is expected per carbon atom in the side-chain except for Hαs of glycine residues, and residues N47 and D48 are excluded. Final chemical shifts were consolidated from 100 independent FLYA runs as previously, where tolerances of 0.1 ppm, 0.5 ppm, and 0.8 ppm for ^{1}H, ^{15}N and ^{13}C were used.

The assignment by FLYA gives similarly good results as for the manual assignment with only 14 strong resonances, i.e., 4 % of the expected resonances by FLYA, deviating from the manual assignments and 301 strong resonances, i.e., 84 %, considered as identical. It becomes, however, clear that FLYA forcefully tries to assign certain resonances if they are not excluded, which occasionally leads to false assignments that nevertheless are considered as strong by FLYA as can be seen when comparing the assignment with and without reference assignments (Fig.3.22A right and left column). As described for the manual assignments, all Cβ resonances of threonines are absent in the 5D HCCNH and cannot be assigned. As HαCα and HβCβ resonances are sufficiently similar in their chemical shifts, FLYA assigns the HβCβ to the same peak as HαCα in a consistent manner such that both resonances are ranked as *strong*. Additionally, some though (*weak*) assignments are found for those residues

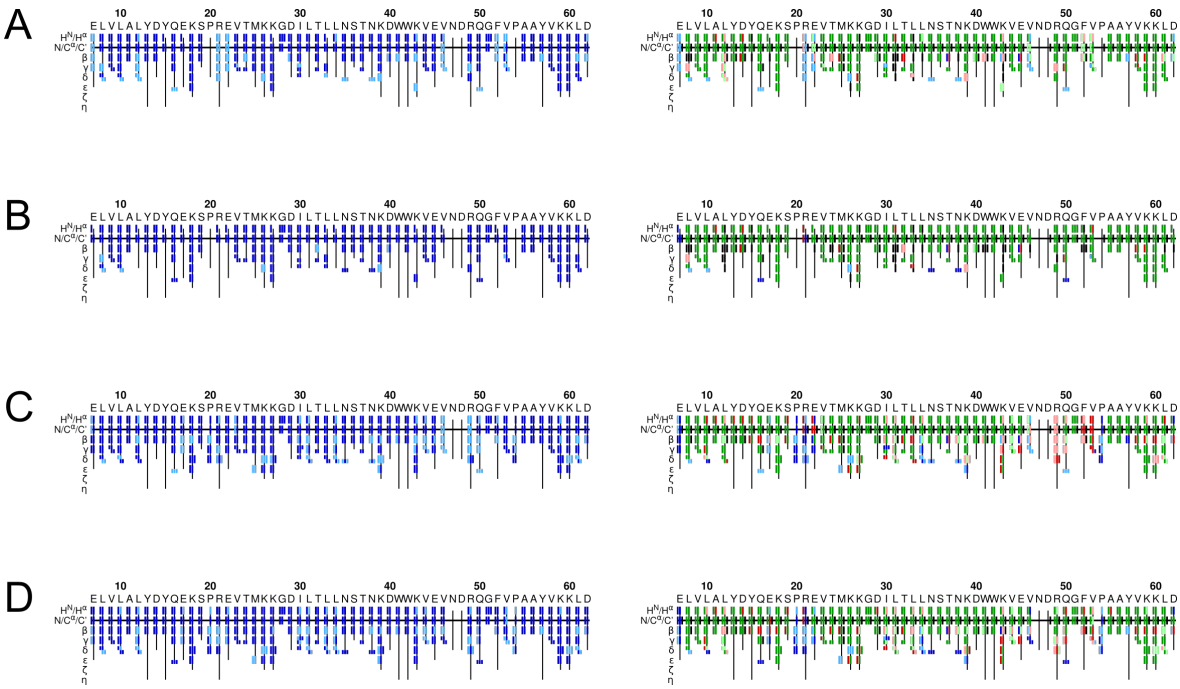


Figure 3.22: Results of different FLYA runs for full resonance assignment based on a minimal set of experiments using a 5D HNcoCANH acquired on a deuterated sample and the 5D HCCNH as well as the 3D hCANH base experiments used for both 5Ds (**A)** and **B)**) and including the 4D HCCH obtained from the combined experiments (**C)** and **D)**). An hNH peak list was additionally included, which usually comes at almost no cost of measurement time. Shown on the left side are the FLYA results without reference and on the right side using the manual assignments as reference. In **A)** and **C)** no chemical shifts were fixed, i.e. no *a priori* knowledge is required, whereas in **B)** and **D)** the backbone chemical shifts were fixed so that only side-chain resonances are added to already known backbone assignments. For all variants, 100 independent FLYA runs were brought to a consensus assignment with tolerances of 0.1 ppm, 0.5 ppm, and 0.8 ppm for ^{1}H, ^{15}N and ^{13}C respectively. Dark blue marks indicate strong assignments according to the basic FLYA definition[186] whereas light blue marks indicate non-reliable assignments.

that are too weak to be detected but are not excluded *a priori* from the run, such as E7 or R21. Most importantly, however, large amounts of side-chain resonances and the protein backbone without carbonyl carbons are correctly assigned by FLYA from essentially only two 5D experiments as ultimately 317 strong assignments are made by FLYA, corresponding to an assignment rate of 71 % of all possible resonances without carbonyl carbons. Considering a scenario in which the backbone chemical shifts are already are assigned by FLYA, e.g., from previous studies, and only the side-chain resonances are to be added, the backbone chemical shift assignments can be fixed. Running FLYA with such fixed peak assignments only brings minor improvements to the assignment quality. Ultimately, 353 strong chemical shift assignments are obtained, corresponding to an assignment of 79 %. Of those resonances, 319 agree with the manual assignments, i.e., 90%, whereas 16 are considered to be different (Fig.3.22B). This indicates that the match-making of side-chain shifts onto the backbone via the H/N/Cα triplet is reliable and mostly unambiguous such that in combination with the amide-to-amide link provided by the HNcoCANH, a unique sequential assignment is won. The major difference in those two data sets is the reduced number of weak assignments for residues such as R21, which, therefore, can be better excluded from the assignment process. Adding the 4D HCCH experiment to the assignment via FLYA (Fig.3.22C, D) should resolve conflicts in the assignments as those determined for HβCβ resonances of threonines and

provide additional shifts to close gaps in the side-chain assignment similar as demonstrated for the manual strategy. While such conflicts can be partially resolved, the sheer amount of peaks found in the HCCH spectrum with no linkage to the backbone and poor chemical shift dispersion of both aliphatic carbon and proton chemical shifts, effect, that many peaks are assigned to several, different resonances, which ultimately lowers the performance and reliability of the chemical shift assignments provided through the two 5D experiments. This becomes apparent not only visually but also in the statistics made by FLYA. Adding the HCCH peak list lowers the number of strong resonances, which correspond to the manual assignment to 276 or 74 % of all strong resonances. Although the number of total strong assignments increases to 374, i.e., 84 % complete assignment, 57 of those vary from the manual assignment, and 41 are assigned additionally. These, however, mostly occur for resonances that were found not to be assignable based on known assignments and knowledge of SH3 and, despite being *strong*, are therefore incorrect. The agreement with the manual assignments can be slightly improved when fixing the backbone assignments so that side-chain assignments are solely added but cannot alter the backbone chemical shifts (Fig.3.22D). Nevertheless, the assignment quality remains mostly unaltered, and the number of matching chemical shifts increases only slightly to 290.

Table 3.2: FLYA results from different setups

Experiments	strong FLYA assignments matching	violations	correct/expected[a]	correct/total[b]
HNcoCANH HCCNH hCANH hNH	301	15	84%	74%
Backbone fixed HCCNH	319	16	89%	78%
HNcoCANH HCCNH HCCH hCANH hNH	276	57	78%	69%
Backbone fixed HCCNH HCCH	290	48	82%	72%

[a]expected from FLYA; manual assignment is assumed correct
[b]correct FLYA assignments vs. total protein resonances

FLYA, therefore, performs best and most correctly when only the 5D experiments are incorporated and performs almost as well as the manual assignment strategy. The automated assignment can provide a substantial amount of correct side-chain and backbone assignments from only two 5D experiments, and it is likely that the performance can be improved when additional experiments are considered, be it for the protein backbone, side-chain, or through-space correlations. Especially the latter ones might help to improve the assignment quality of the 4D HCCH as certain assignment possibilities might be strengthened based on spatial

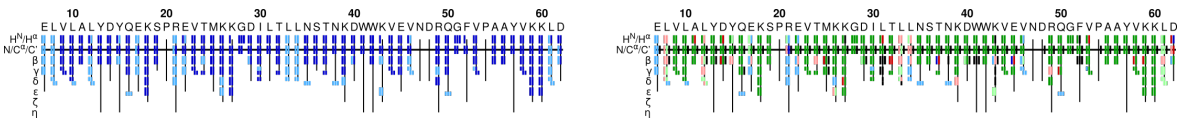


Figure 3.23: FLYA results as in figure 3.22A, but replacing the 5D HNcoCANH with a peak lists of the same experiment recorded on a fully protonated sample, resembling the assignment success with all spectra recorded on a single sample. The assignment quality is almost unaltered as changes are almost exclusively found for the residues of which no backbone assignment can be found, i.e. L8, L33, and occuring from that L34, and D62 (see also section 3.2.1).

contacts. The assignments obtained through FLYA prospectively will leverage the assignment of larger and unassigned proteins, despite being incomplete, as they provide helpful assistance for the manual assignment process and are best used in such a computationally aided assignment approach.

As mentioned in the introduction of this section, the benefits of fully protonated samples include only small sample requirements as well as access to proteins for which a deuterated preparation is not possible. As demonstrated for the manual assignment of the protein backbone from a 5D HNcoCANH recorded on a fully protonated sample, the strategy for full resonance assignment can be readily tested for such a scenario in which only a single, fully protonated sample is required. If the resonance assignment is done manually, essentially no alterations in the assignment of side chains occur since the combined 5D/4D HCCNH/HCCH is indeed already acquired on a fully protonated sample, and therefore only the three missing residues from the backbone, L8, L33, and D62, will remain elusive in addition. Incorporating FLYA once again (Fig.3.23), this is well reflected when comparing the results with those from the FLYA set-up where only the 5D HCCNH is used (Fig.3.22A). The absence but not exclusion from FLYA of L8, L33, and D62 in the 5D HNcoCANH causes a slight decrease in strong assignments as assignments are forced for the missing residues with the majority of assignments, backbone, and side chains, remaining unaltered due to the high redundancy introduced through the two 5Ds and their common 3D hCANH base. In total, 289 *strong* resonances are identical to the manual assignments, but only 16 show violations, emphasizing that the protonated 5D HNcoCANH only leads to a reduced number of strong assignments but not to an increase of violations. This demonstrates that the approach of full resonance assignment from the here introduced 5D experiments for backbone and side-chain assignments is well suited and yields large assignments even when recorded on a single, fully protonated sample.

Discussion

The resonance assignment of proteins of all sizes poses a time-consuming hurdle for downstream applications such as relaxation experiments and structure calculation. In this work, two 5D experiments for backbone assignment and a combined 5D/4D experiment for side-chain assignment were successfully implemented. The backbone experiments, the 5D HNcoCANH in particular, are essential for the assignment of large proteins such as TS (see below) but can also be an alternative to established assignment strategies as they significantly reduce the number of required experiments for full backbone assignment in smaller proteins and drastically reduce the measurement time if NUS is consequently employed. Additionally, the 5D backbone experiments are suitable for deuterated as well as fully protonated sample

preparations with no further limitations in the usual context of fully protonated samples, i.e., increased line widths and reduced sensitivity.

The combined 5D/4D experiment for side-chain assignment in fully protonated samples provides a useful and easy-to-use tool for this purpose. While the 4D HCCH is difficult to assign and suffers from heavy resonance overlap, even for a small protein, like the SH3-domain, it becomes more accessible in combination with the 5D and provides useful extra information. Despite a more compromised sensitivity than expected, the measurement times can be kept short, and in combination with a single 5D HNcoCANH, almost full resonance assignment is possible. At this point, the major bottleneck for true full resonance assignment is the absence of aromatic resonances in the spectra, which so far could not be successfully reached simultaneously with aliphatic signals. Even recent mixing schemes such as Al-FRESCO[210] do not resolve this problem despite a general successful implementation. Processing of the 5D/4D requires almost no extra interaction apart from the NUS reconstruction, as the involved scripts run essentially without user input.

Automated assignment or computationally assisted assignment, however, remains challenging and ambiguous even for small proteins and requires dedicated restrictions for FLYA and yields only good results if the 4D HCCH experiment is not incorporated. For larger proteins, this problem will only become more pronounced, but nevertheless, FLYA can probably give useful starting points for a manual assignment approach.

For future studies, continuous efforts should be invested to include aromatic resonances in the combined 5D/4D for side-chain assignment, as they often carry relevant information about, e.g., the protonation states of histidine side-chains. Furthermore, the performance of this experiment has to be evaluated on larger proteins. To push automation further forward, the set of experiments should be expanded by an experiment for distant restraints such as hNH-RFDR, and the outcome of fully automated resonance and RFDR assignment with consecutive structure calculation should be analyzed. The incorporation of such information might help to resolve the ambiguity in the assignment process as some redundant assignments can be excluded.

3.3 Assignment of and insights into 2x72 kDa tryptophan synthase

This section is published[67] as: Klein, A., Rovó, P., Sakhrani, V.V., Wang, Y., Holmes, J., Liu, V., Skowronek, P., Kukuk, L., Vasa, S.K., Güntert, P., Mueller, L.J., Linser, R. 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.* **119**, e2114690119, doi: 10.1073/pnas.2114690119. The list of assigned chemical shifts can be found in section 7.3 and in BMRB entry 51166.

Introduction

Tryptophan synthase is a well-studied system with several hundred structures solved with X-ray crystallography. Until now, the combination of X-ray crystallography, molecular dynamics (MD) simulations, mutation studies, and NMR can draw a very detailed picture of the machinery in TS. Especially NMR studies, however, are limited by the use of labeled

substrate[146,147,149,150,211] or special labeling schemes of the protein, e.g., the use of labeled His and Lys in an otherwise unlabeled protein and therefore restrict the overall assessment of the protein, especially its dynamics and the proton chemistry involved. To achieve an improved understanding of the protein's allosteric mechanisms and active site chemistry, solid-state NMR is the method-of-choice, but (full) resonance assignment is compelling.

Due to the molecular weight of 72 kDa in the asymmetric $\alpha\beta$-unit, the protein is too large to be studied by solution-state NMR as short T_2 times are anticipated that would restrict the use of necessary (see section 3.1) higher-dimensional experiments involving multiple and long INEPT steps for magnetization transfer, ultimately leading to poor intensities. Due to size-independent transfer efficiencies and generally faster magnetization transfer over dipolar mechanisms, i.e., CP, solid-state NMR proofs to be a powerful tool for assessing TS in its native $\alpha\beta\beta\alpha$ constitution without additional labeling schemes apart from deuteration.

Facilitated experiments

The manual assignment approach of TS (Fig. 3.28) is built on a set of complementing pairs of 4D spectra, hCOCANH/hCACONH and hCACBcaNH/hCACBcacoNH, each suitable for a backbone walk in the forward and backward direction, complemented by a 5D HNcoCANH experiment to achieve highly dispersed, direct residue-to-residue links and a 4D hCCNH-type experiment for side-chain carbon chemical shifts. Contrary to the implementation on the SH3 domain, the HNcaCONH was not used for the assignment of TS under consideration of its lower performance compared to the HNcoCANH (more details see below). The 4D hCOCANH, 4D hCACONH experiments were used as published previously[184], while the 4D hCCNH, 4D hCACBcaNH, 4D hCACBcacoNH [1] and 5D HNcoCANH (as introduced in section 3.2) were newly developed and build from their lower-dimensional counterparts[38,41,46,61–64]. All experiments are entirely based on CP transfers for heteronuclear transfers and employ BSH-CP[26] for homonuclear, i.e., Cα-CO transfers. INEPTs with near-complete evolved antiphase terms, $\tau = 6$ ms $< 1/(4J)$, were employed for Cα-Cβ experiments, first to transfer the magnetization fully and second to reduce offset effects as they may occur in DREAM transfers[28], without compromising sensitivity[58] and avoid unnecessary splitting of magnetization for Cα nuclei. As shown in figure 3.26A, the use of INEPTs eventually results in a higher intensity along the Cβ dimension, substantiating theoretical considerations. CC mixing in the side-chains was achieved by MOCCA mixing[45,46] in conjunction with direct excitation of the side-chain carbons (COPORADE[47]), enabling longer mixing times with comparably low power, i.e., heat emitted into the sample per time unit, no need for rotor synchronization and more magnetization available in the side chain. The delay between blocks was optimized to 60 µs and looped over 256 times, accumulating into a mixing time of 64 ms.

Fully optimized transfers in all experiments account for comparably high sensitivities. For the 5D HNcoCANH, 8 % of the hNH sensitivity could be maintained (Fig.3.25B and Fig.3.26B). The HNcaCONH, however, reaches only 5 % of the respective hNH, due to different C→C transfer efficiencies (Fig.3.26B)[58] and was therefore not used for the assignment of TS, keeping in mind that both 5Ds can be facilitated for a bidirectional backbone walk.

[1] 4D hCACBcacoNH and hCACBcaNH experiments were implemented by Dr. Petra Rovó, LMU Munich, during her time as a Postdoc in the group and all 4D spectra apart the 4D hCCNH were acquired by her as well. NUS reconstruction and data analysis was entirely done by the author and as these experiment are essential for the assignment process and context of the entire thesis, they are included here

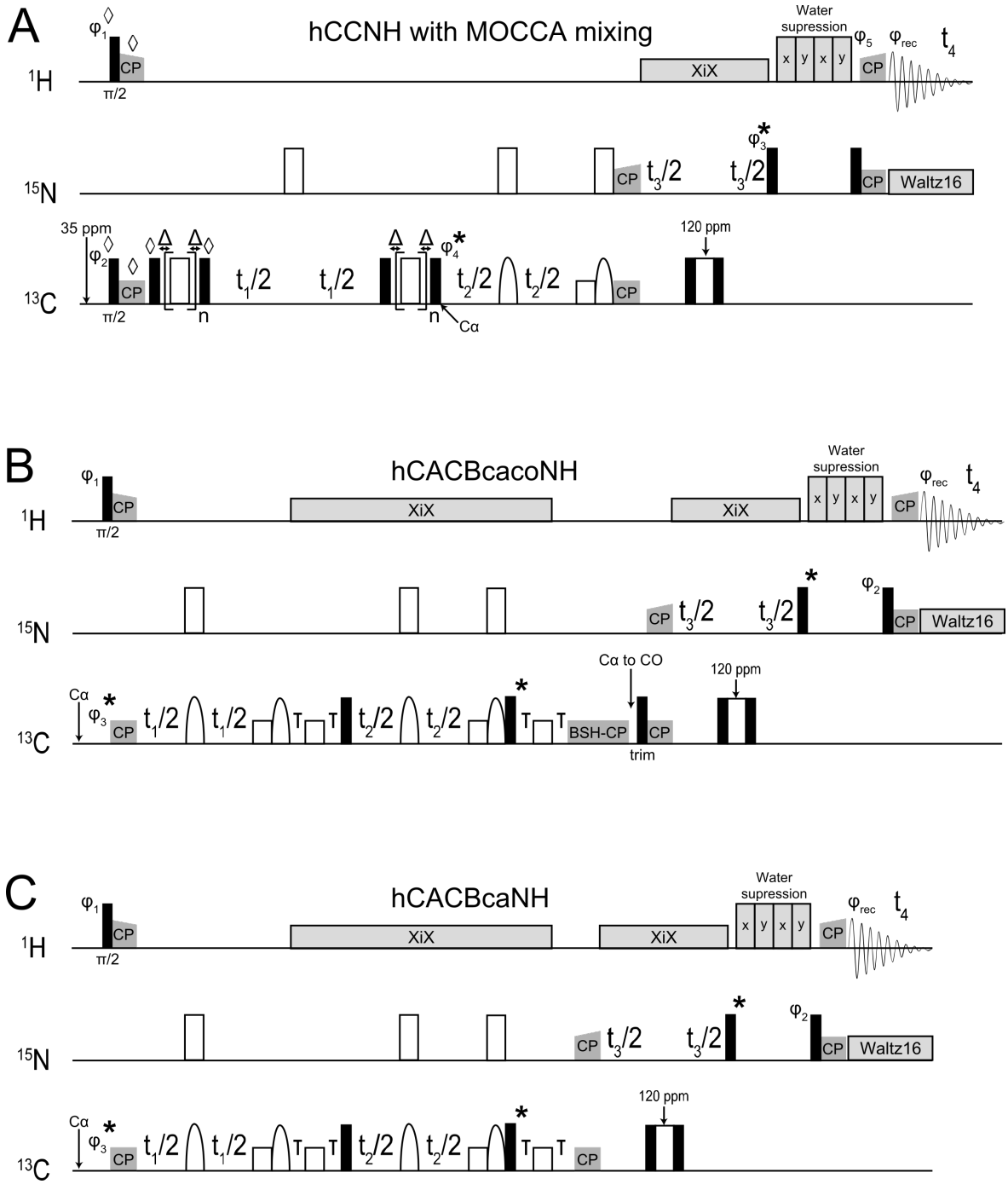


Figure 3.24: **A)** 4D hCCNH pulse sequence with MOCCA mixing (Δ = 60 µs, n = 256) and the COPORADE approach[47]. Phase cycling: ϕ_1 = -x,x; ϕ_2 = x,-x; ϕ_3 = x,x,-x-x; ϕ_4 = x,x,x,x,-x,-x,-x,-x; ϕ_5 = x,x,x,x,x,x,x,x,-x,-x,-x,-x,-x,-x,-x,-x; ϕ_{rec} = x,-x,-x,x,-x,x,x,x,-x,-x,x,x,x,-x,x,-x,-x,x. **B)** 4D hCACBcacoNH pulse sequence. Phase cycling: ϕ_1 = x,-x; ϕ_2 = x,x,-x,-x; ϕ_3 = x,x,x,x,-x,-x,-x,-x; ϕrec = x,-x,-x,x,-x,x,x,-x. The delay τ was set to 6 ms ($< 1/(4J)$) to achieve maximal transfer to Cβ while maximizing intensity. **C)** 4D hCACBcaNH pulse sequence. Phase cycling: ϕ_1 = x,-x; ϕ_2 = x,x,-x,-x; ϕ_3 = x,x,x,x,-x,-x,-x,-x; ϕ_{rec} = x,-x,-x,x,-x,x,x,x,-x. Also here the delay was set to 6 ms. Phases after evolution periods of indirect dimensions were changed to obtain phase-sensitive spectra by the States-TPPI method[173] (marked by asterisks). In case of the 4D hCCNH a block of pulses was rotated (marked by small hashes). Filled and empty rectangles represent 90°and 180°pulses. The carrier changes on the carbon channel are denoted with arrows. Water suppression was achieved by MISSISSIPPI without gradients[191].

Nevertheless, the overall intensity in the first FID is comparable to the hCACBcaNH or even higher regarding the hCACBcacoNH, emphasizing the practicability of those experiments. The hCACONH and hCOCANH reach 17 % and 20 % of the hNH, respectively, the hCACBcaNH and hCACBcacoNH 8 % and 3 % and the hCCNH up to 9 % of the hNH in the first FID.

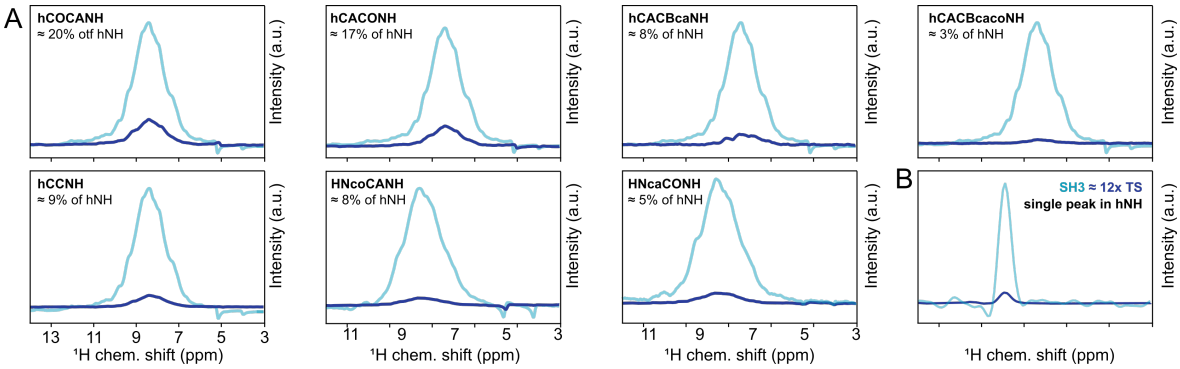


Figure 3.25: **A)** Bulk sensitivities all experiments used to assign TS (dark blue) compared to the hNH bulk signal (light blue). **B)** Isolated and well-resolved cross-sections from hNH for representative residues found in the SH3 domain and TS under identical conditions regarding preparation and measurement. As expected, this single-peak intensity is approximately reduced by a factor of 10, reflecting the 10 times higher molecular weight of TS over SH3 (see main text).

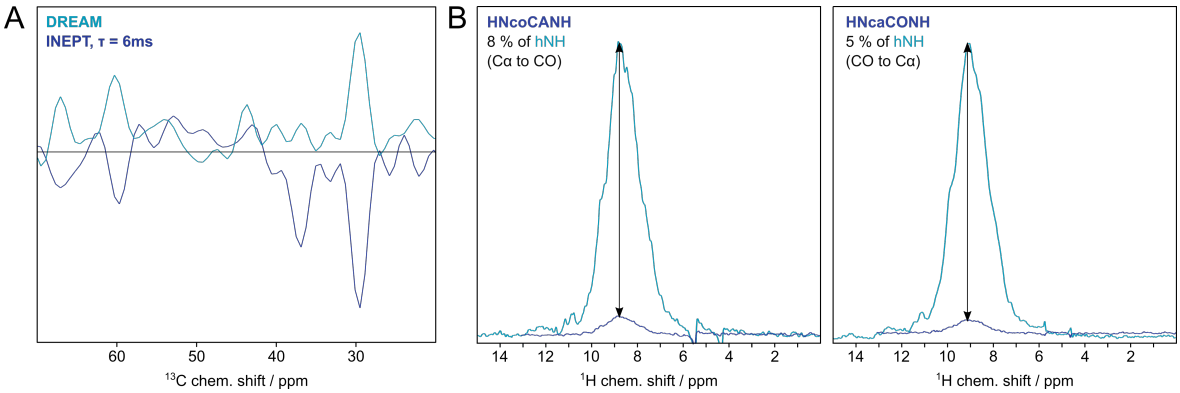


Figure 3.26: Sensitivities of the **A)** hCACBcaNH experiment regarding CC transfers, comparing the intensities in a 2D HCβ plane along the Cβ dimension obtained by either DREAM[28] transfer (light blue) or INEPT with an optimized delay $\tau = 6$ ms and **B)** HN bulk intensities comparing the HNcoCANH and HNcaCONH experiment. The magnetization pathway $H_i \rightarrow N_i \rightarrow C\alpha_i \rightarrow CO_i \rightarrow N_{i+1} \rightarrow H_{i+1}$ is favoured over $H_i \rightarrow N_i \rightarrow CO_{i-1} \rightarrow C\alpha_{i-1} \rightarrow N_{i-1} \rightarrow H_{i-1}$ as C$\alpha \leftrightarrow$ CO transfers show different efficacy depending on the direction of the transfer[58] which is ultimately reflected in the overall better performance of the HNcoCANH experiment.

All spectra were reconstructed by SSA, as implemented in the programs cleaner3d / reconstructor3d[104], cleaner4d / reconstructor4d[105] and cleaner5d / reconstructor5d[102].

Measurement time considerations

As pointed out previously (section 1.4), NMR experiments beyond three dimensions can usually not be acquired in a uniformly sampled way. Therefore, all spectra for the assignment of TS were recorded with NUS at different densities, depending on the experiment and the dimensionality, including a 3D hCANH experiment as the base for the reconstruction of the HNcoCANH. The overall acquisition time of all experiments accounts for 2 d,3.5 d, 6 d, 6 d,

33 d, and 30 d for the hCACONH, hCOCANH, hCACBcaNH, hCACBcacoNH, HNcoCANH, and hCCNH experiments, respectively. It should, however, be mentioned that the 5D HNcoCANH was recorded longer than strictly necessary to obtain a spectrum of high quality considering the aforementioned importance for the manual assignment process. Despite a similar intensity in the first FID of the hCCNH and the hCACBcaNH, the former one underlies a considerable distribution of magnetization and therefore signal intensity in the side chain over several carbon atoms in contrast to the hCACBcaNH, which will only result in a single peak and hence had to be recorded for a longer time. When discussing acquisition times and comparing intensities for (micro-crystalline) solid-state samples, a quadratically growing demand of measurement time, i.e., scans, has to be taken into account. This rises directly from the fixed rotor volume that limits the number of molecules in the rotor. Although the overall number of nuclei, e.g., protons, might be considered constant and consequently show the same bulk signal of two different proteins in the same experiment, the contribution of each individual residue per molecule will be lower for bigger molecules. For example, in the case of the SH3 domain with 62 residues and TS with 665 residues, there will be approximately ten times more SH3 molecules in the rotor, i.e., a given residue will occur ten times more often, and consequently, each individual residue will have a signal-to-noise ratio ten times lower in TS over SH3. To fully compensate for that difference, the experiments on TS would have to be acquired a hundred times longer to obtain the same SNR, or generally, for a protein being x times larger, the measurement time requirements grow by a factor of x^2. This is reflected in the longer acquisition times spent on TS as well as the reduced SNR of each individual peak (Fig.3.25B). Furthermore, it should be mentioned that the sensitivity generally can be significantly improved by increasing the B_0 field, which might be overall beneficial for high molecular weight proteins. , SNR $\propto \sqrt{B_0^3}$ and measuring on a cutting-edge high-field magnet of 1.2 GHz, as it recently became possible at the time of this work, would, in fact, increase the SNR by a factor of 2.24 or reduce the measurement time to a fifth. In the latter case, this would, however, imply a constant number of NUS points, effectively lowering the NUS density of the data set.[2]

Assignment strategy

The 4D hCOCANH and hCACONH[184] correlate the *i* and *i+1* amide shifts based on shared Cα and CO resonances (dark green and orange in figure 3.28) or alternatively the preceding and subsequent Cα and CO chemical shifts of a given H/N shift pair. The 4D hCACBcaNH and 4D hCACBcacoNH pair can be used accordingly, however, dispersing Cβ chemical shifts rather than CO chemical shifts in one of the dimensions (light green and red in figure 3.28). While the dispersion of chemical shifts in a Cβ dimension is advantageous over the dispersion of CO chemical shifts and will therefore reduce the overlap of peaks (section 3.1), both Cβ experiments show lower (HN bulk) sensitivity and might therefore be prone to missing or noisy peaks. For both pairs, the backbone-walk, linking amide to amide, involves two steps. First, two carbon resonances have to be matched in two corresponding experiments, and second, the next residue has to be identified by the resulting H/N shift pair. Ideally,

[2] If the acquisition time and spectral width in ppm in any indirect dimension are kept constant, the number of time points k has to be increased as $aq = \frac{k}{SW_{ppm}B_0} = const.$ In the case of a constant number of NUS samples, the NUS density will effectively be lowered.

only one resonance would appear in both of these steps to yield an unambiguous backbone walk. However, certain ambiguities will arise, with each extra peak marking a branch in the assignment pathway. Using a set of two pairs will reduce these ambiguities significantly, as ultimately the information provided has a five-dimensional character when each amide H/N shift is connected to three carbon resonances, Cα, Cβ and CO. Nevertheless, the two-step assignment problem remains, which can elegantly be resolved through amide-to-amide experiments as 4D HNcocaNH (HNNH) or 5D HNcoCANH. As mentioned previously, both experiments are equally powerful for the linkage of two adjacent residues. The 5D reduces the ambiguity only for one direct link due to the fifth dimension, i.e., the likelihood for only one peak showing in the processed plane (SMFT, see section 1.4.2) is three times higher (Fig.3.5, section 3.1) than for the 4D HNNH, while for the consecutive step, *i+1* to *i+2*, both experiments perform equally. While the linking of several residues only based on the 5D HNcoCANH is therefore challenging, it is utterly indispensable to resolve a single linking problem in the 4D data sets, which frequently occurs (Fig.3.27). Figure 3.27 shows impressively how the amount of peaks in a given plane of experiments with increasing dimensions is reduced. Not even the combination of two 4D spectra, hCOCANH, and hCACBcaNH, can provide a unique resonance in contrast to the 5D experiment. Additionally, the 5D increases the redundancy of chemical shifts and therefore the matching of other H/N/Cα resolved experiments is significantly improved (Fig.3.5, section 3.1), rendering side-chain assignment and with that also amino acid identification easier compared to the 4D HNNH. State-of-the-art algorithms for automated assignments like FLYA**Schmidt2012** can partly compensate the several branches occurring during the assignment process, when based solely on 4D spectra. If many 4D spectra with high chemical shift redundancy are available, the advantage of 5D experiments in the context of exclusively automated assignment, is reflected only in a reduced set of necessary experiments (for more details, see section 3.3). However, fully manual assignment strategies are significantly simplified or, moreover, only possible using additional 5D experiments. The identification of amino acids is often made by Cβ resolved

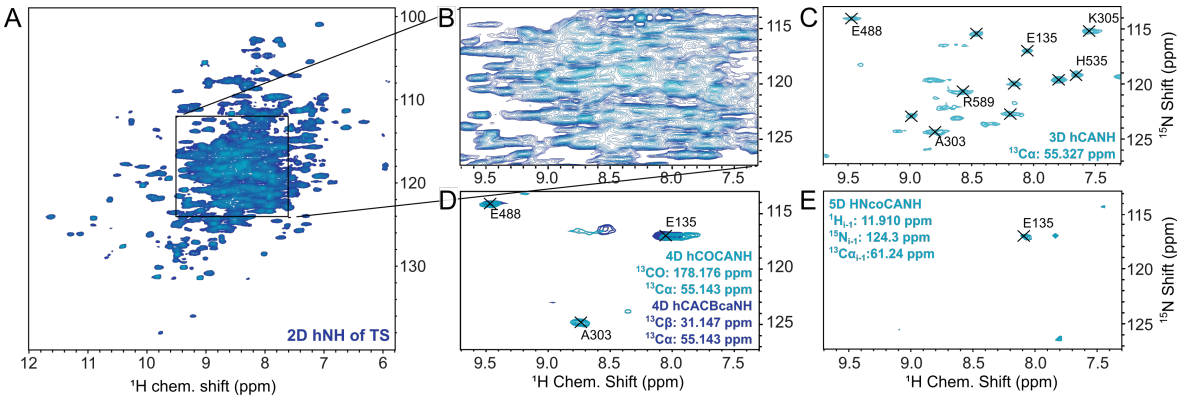


Figure 3.27: Reduction of peak ambiguity with increasing dimensionality. **A)** The hNH spectrum of $\alpha_2\beta_2$ TS with **B)** a relevant excerpt of the highly overlapped middle region due to the high α-helical content of TS. Overlap is mostly not because of increased line widths, as those observed for isolated residues are between 20-50 Hz as expected[42]. **C)-E)** show H/N planes for 3D (hCANH), 4D (hCOCANH and hCACBcaNH) and 5D (HNcoCANH) experiments. Sole the 5D plane shows a single peak and therefore unambiguous information. Crossed peaks have their respective maxima at the given ppm values while the other peaks lie outside of the planes. All spectra were recorded in a 1.3 mm rotor with 55 kHz MAS and a ^{1}H Larmor frequency of 700 MHz.

experiments due to the more characteristic chemical shifts compared to Cα or CO. Besides

the distinguished shifts of glycine, alanine, serine, and threonine, uncertainties about the correct sequential assignment will usually remain in large proteins, even with the background of already linked resonances, as the chemical shifts can overlap and vary with respect to the secondary structure. Experiments that resolve additional side-chain resonances are therefore required to narrow down the number of possible amino acids for a set of resonances. For TS, a 4D hCC(α)NH-type experiment based on MOCCA mixing was built from 3D versions of the same[45,46], which grants access to side-chain carbon resonances beyond Cβ. The correct identification of the residue type in the context of sequential assignment is often essential to correctly assign stretches of linked residues, especially in larger proteins, in which shorter sequences occur more often. In TS, for example, the short sequence GAXZ exists nine times in the protein, rendering it difficult to pin it to its correct position. Here, residue Z could be identified as lysine, with all carbon resonances down to Cϵ being visible (Fig. 3.28). The sequence GAXK is then only found a single time, ultimately allowing for the correct assignment as G352, A353, K355 and therefore also identifying X as H354.

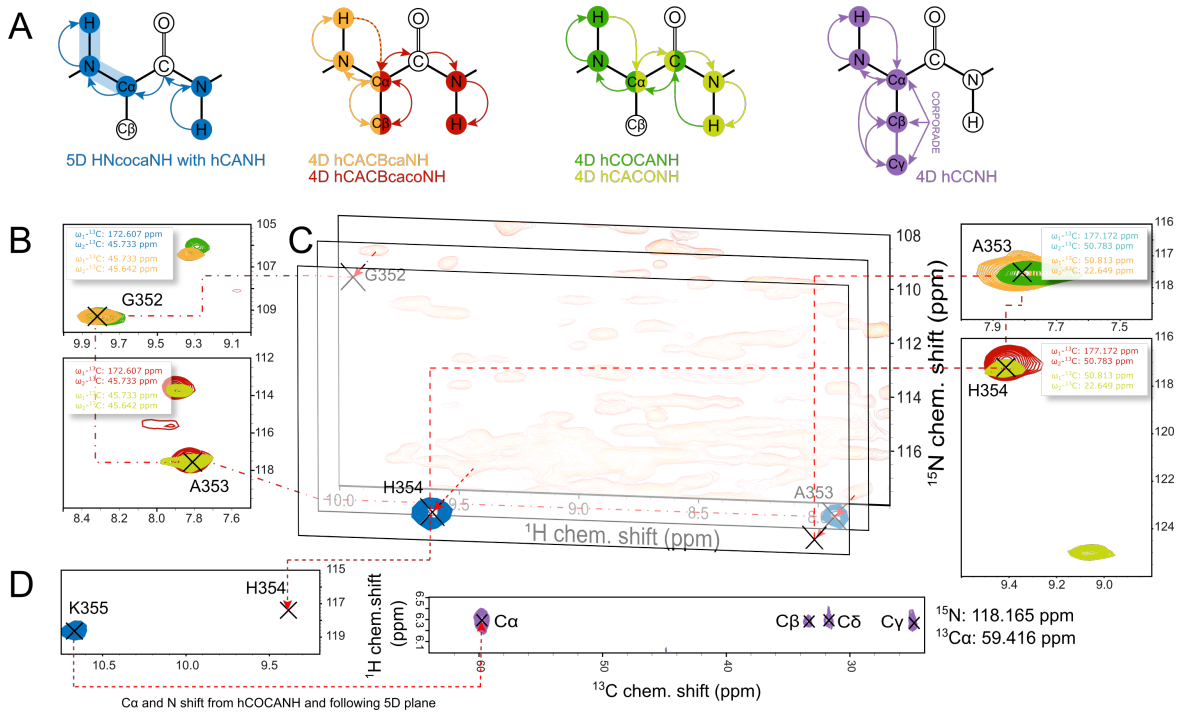


Figure 3.28: Experiments and assignment approaches for representative residues from the active site of the β-subunit. A) Magnetization pathways for the experiments employed. All experiments are CP-based, only for CαCβ transfers, INEPTs were employed. For side-chain correlations, MOCCA mixing[45,46] was used. B) Superimposed 2D planes of different 4D spectra with ambiguous peaks (left and right side of the figure). Only with the help of a 5D experiment (C) depicted as a stack of 2D planes) an unambiguous assignment is possible for many peaks. D) Link between the important active-site residues H354 (βH86) and K355 (βK87) from HNcoCANH and the corresponding hCCNH-MOCCA strip of K355. The hCCNH, triply connected to the sequential backbone link via its H/N/Cα shifts, allows to identify the residue as lysine, making a unique linkage between all the residues shown here (The sequence GAXK occurs just once while GAXX can be found nine times.) and by that also assign H354 unambiguously.

Computation-aided assignment of TS by FLYA

Through an entire manual assignment approach, about 40 % of the protein could be most likely correctly assigned. It is expected, however, that the missing 60 % of assignments are

not entirely linked to insufficient magnetization transfers or hindered H/D exchange as will be analyzed later on in section 3.3, but are rather due to loose ends in the assignment process, which arise eventually from missing signals in the 5D HNcoCANH and therefore can only be assigned based on 4Ds with the associated challenges.

To leverage the level of assignment on the one hand and verify the existing assignments on the other hand, FLYA[Schmidt2012] was used. While the analysis of the combined chemical shift assignments was done in the previous section to allow for a direct comparison between manual and combined approaches, this section serves to give more insight into the performance of FLYA using different data sets and parameters.

Although useful, peak picking any spectrum before the assignment process is not required for a manual assignment strategy as the peak can always be checked at given chemical shifts. Algorithms for automated assignment, however, work exclusively on peak lists, and therefore peak picking has to be done previous to the actual assignment. While this might appear a trivial requirement, it can be the first limitation when using FLYA (or any other algorithm) in conjunction with spectra of more than three dimensions. Although FLYA appears to be tolerant to additional peaks, i.e., surplus peak-like noise, an incomplete peak list can have a significant impact on the assignment results[Schmidt2012, 212]. Nevertheless, peak lists of high quality are beneficial for reliable assignment results. Peak picking in 2D spectra is easily performed manually by selecting fully isolated or partially overlapped resonances, as well as it is for 3D spectra in which two out of three shifts are usually fixed, and the peak picking follows the scheme of a 2D spectrum. In addition, 2D and 3D spectra are mostly not sensitivity limited, and peaks often appear with a high signal-to-noise ratio. For TS, however, the peak picking procedure is complicated by the reduced SNR of individual resonances due to its larger molecular weight on the one hand and the almost exclusive use of experiments with four and five dimensions on the other hand. In contrast to 3D spectra and because the display of spectra is limited to two dimensions, peak picking of 4D spectra is not only cumbersome but error-prone when done completely manually as two rather than one dimension have to be screened completely, which means that theoretically for every point in a given dimension another entire dimension has to be checked for peaks which most likely will lead to either missing peaks or more likely a high number of peaks being picked multiple times resulting in peak lists of lower quality for the following assignment process. Here a semi-automated approach was used. Peaks in the 4D spectra were automatically picked using the peak-picker of CCPNmr v2[213] and the parameters adjusted until the expected number of signals was approximately reached, i.e., around 630. In a second step, the peaks were manually inspected to estimate the number of noise peaks included qualitatively. When it became clear that too many noise peaks were included or well-defined signals were skipped, the parameters were further adjusted. As the 5Ds are solely processed as a stack of 2D planes, peak picking can easily be done manually as every plane can be handled like an hNH spectrum and, moreover, the absence of a three-dimensional peak shape or a highly distorted one in a pseudo-3D spectrum potentially causes problems for the peak picking algorithm. The obtained peak lists used for FLYA eventually contain a considerably different amount of peaks (Table), which nevertheless reflect the best possible results achievable within a certain time frame. Especially the 5D shows significantly fewer peaks than anticipated, which is most likely due to a high number of remaining artifacts that make it challenging to identify

Table 3.3: Peaks picked in a semi-automated way as well as peaks expected for all individual experiments.

Experiment	Peaks	
	expected	picked
hCOCANH	593	425
hCACONH	594	747
hCACBcaNH	565	808
hCACBcacoNH	568	638
HNcoCANH	592	271

the actual peak, considering that 803 peaks were provided for reconstruction based on the 3D hCANH experiment.

For the final consensus assignment, 100 independent FLYA runs instead of the 20 default runs were consolidated for all results presented here. The tolerances for chemical shift matching for ^{1}H, ^{15}N ^{13}C resonances were set to 0.1 ppm, 0.5 ppm and 0.8 ppm respectively. Furthermore, chemical shift distributions were predicted with SPARTA+[214] based on the crystal structure (PDB ID: 4HT3), rather than using average shifts from the BMRB, which significantly improves the rate and reliability of assignment. Although by default, none of the experiments used for the assignment are supported by FLYA in the first place, the algorithm is highly flexible, and new experiments, i.e., hypothetical magnetization pathways, can be added following a certain scheme. The extensions for the CYANA library can be found in the supplement (section 7.2).

FLYA cannot only be facilitated for assignment purposes, but, given the fact that 251 residues were already manually assigned, it can also be used as a tool to verify existing assignments. To do so, the manual chemical shifts were given as a reference to FLYA, which in turn will compare its own assignments with the given ones after the individual runs were brought into a consensus assignment. Unlike fixing assignments either to certain peaks or their shifts, this merely constitutes a comparison of shifts in the very last step with no impact on the actual assignment done by FLYA. Except for the resonances belonging to residues S180, G181, and V182, no shifts or peaks were fixed to obtain an unbiased assignment. Peaks belonging to those residues were checked manually against the other five occurring Ser-Gly-X combinations and found to be the only possible assignment possibility, as Ser-Gly-Val only occurs once in the primary sequence. In addition to the parameters used for the assignment, further modifications were done for the consolidation process, which yields the final chemical shift list to improve the reliability of the assignments.

The ranking of *strong* and *weak* assignments by FLYA is by default only based on how often a consensus assignment is obtained for certain resonances. If in 80 % of the runs the same consensus assignment is obtained, the resonance assignment is classified as strong and otherwise as weak. While this certainly limits assignment branches based on false peaks to be considered for the final consolidation of chemical shifts, the criterion provides little to no information about the assignment quality of the adjacent residues, which might be unreliable. The latter criterion, however, is important to validate a resonance assignment in the context of how it is linked to the adjacent residues while a classification purely based on probability can potentially rely entirely on the chemical shift, i.e., the random coil shift that is associated with certain residues, and on exclusion from other places on the primary sequence. Therefore,

the criteria for *strong* assignments were expanded, and resonances marked as strong have to fulfill the following points[3]:

(i) 80 % occurrence in the individual runs (default)

(ii) at least two *strong* resonance assignments have to present in one residue

(iii) at least four *strong* resonances have to be present in the neighboring residues

(iv) the two amide resonances, H and N, have to be assigned as *strong*

Through more strict classification criteria, it will be more likely that resonances or residues ranked as *strong* are indeed correct and "lucky shots" are more likely avoided, while on the other hand, the absolute number of resonances considered as *strong* is decreased in the end. ^{1}H and ^{15}N resonances have lower assignment tolerances compared to ^{13}C resonances, and by the inclusion criterion of *strong* ^{1}H and ^{15}N resonances, the overall chemical shift range for a *strong* resonance is reduced, reinforcing their assignment. In the case of the FLYA run, which involves all spectra used for the manual strategy, the default classification leads to 2214 *strong* resonances out of 3109 overall assigned, while the newly defined criteria only allow for 2016. On the other hand also the number of *strong* violations, i.e., disagreements between the manual and FLYA assignments are reduced from 70 to 53 resonances. Identical assignments essentially remain unaffected. Table 3.4 summarizes the results for the aforementioned FLYA setup. The performance of different set-ups for FLYA will further be evaluated based

Table 3.4: Change of *strong* resonances found by FLYA in relation to the inclusion criteria considered

	default	extended
Identical	1022	1011
Different	70	53
Additional	1122	952
Assigned	2214	2016

on the agreement with the manual assignments as well as the number of additional strong assignments obtained.

Using all the spectra involved in the manual backbone assignment process, i.e., 4D hCOCANH / hCACONH, 4D hCACBcaNH / hCACBcacoNH, and 5D HNcoCANH, the extend of assignment could essentially be doubled, and the majority of manual assignments are confirmed (dark green). Only a few resonances do agree weakly (light green) or deviate from the manual assignments (red) (Fig.3.29). In total, 2214 *strong* resonance assignments can be obtained, corresponding to 490 residues or 78 % of the non-proline residues. From those, 369 residues are fully assigned as *strong*, i.e., all of the H,N,Cα,CO and Cβ resonances are considered as *strong*, while the remaining 121 residues contain at least a subset of those resonances. From the 1199 resonances assigned manually, 1011 or 84 % are confirmed and ranked as *strong*, with additional 46 resonances agreeing weakly. Merely 5 % or 53 resonances show a *strong* disagreement with the manual assignment, and 76 maybe deviate, i.e., are ranked as *weak*. FLYA could assign 952 extra resonances ranked as *strong* (see also Table 3.4).

The confirmation of most of the manual assignments by FLYA increases the probability for those to be actually correct. However, the large extent of additional residues assigned

[3]In collaboration with Prof. Peter Güntert, University of Frankfurt

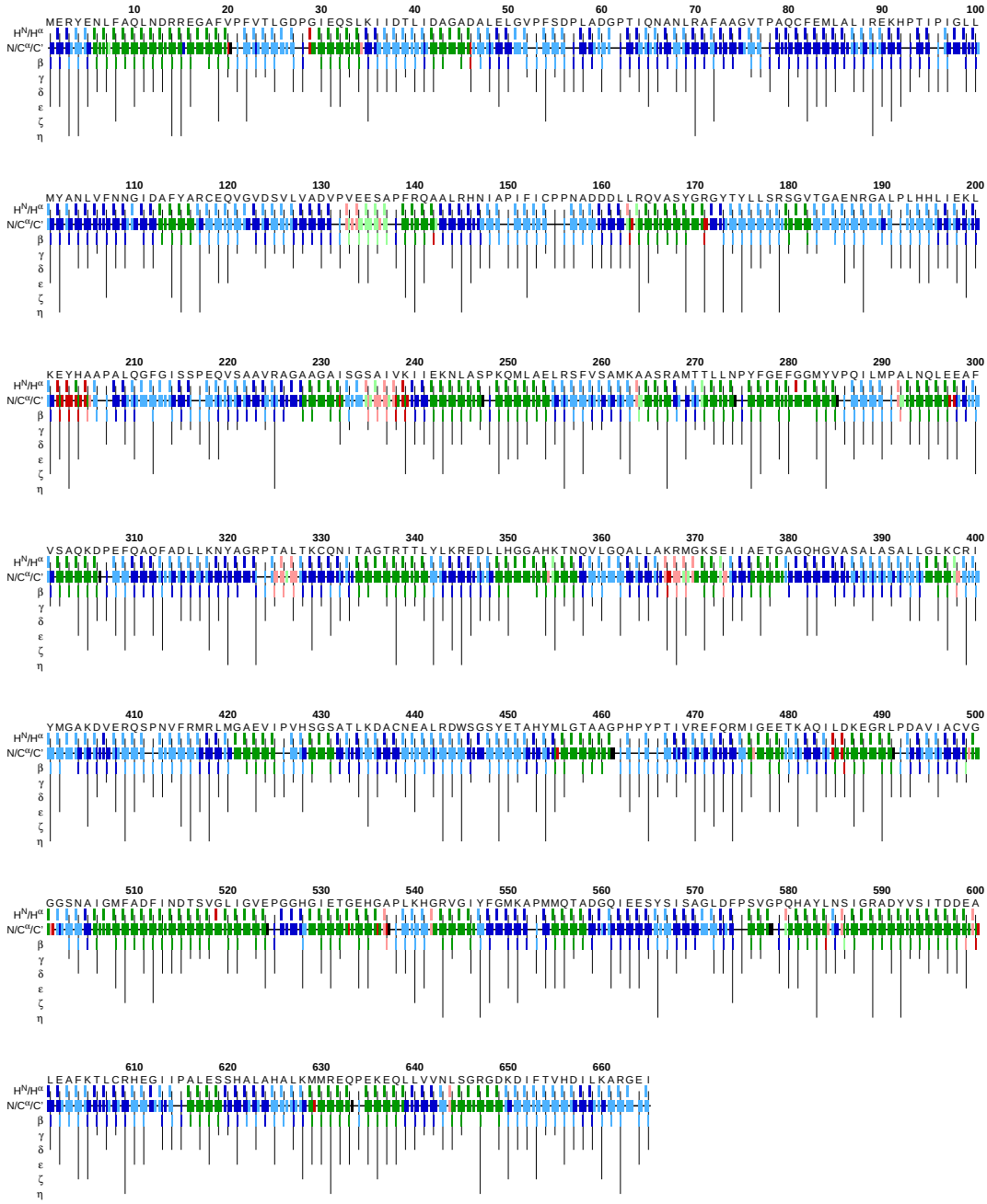


Figure 3.29: FLYA results for resonance assignment using all available experiments, i.e. 4D hCACONH, 4D hCOCANH, 4D hCACBcaNH, 4D hCACBcacoNH, and 5D HNcoCANH. Green colors indicate consistency with the manual assignment. Red indicates (large) differences beyond the defined tolerances, while additional assignments are colored in blue. For all colors, dark tones correspond to assignments classified as *strong*, whereas lighter tones are ranked as *weak* assignments. The manual assignments are almost exclusively confirmed by FLYA, with only 5 % of the assigned resonances showing big differences. In addition, 239 residues with all or some resonances ranked as *strong* could be assigned, with a subset of 118 assigned residues having all resonances, i.e., H,N,Cα,CO and Cβ, classified as *strong*.

by FLYA but not manually might come as a surprise in the first place. As pointed out before (section 3.1 and 3.3), the 5D is crucial for the manual assignment as the combined 4Ds often do not give a unique assignment possibility. When comparing the assigned peaks per experiment by FLYA with the manual assignments, this becomes evident as well as FLYA can assign 218 out of 271 (80.4 %) provided peaks for the 5D HNcoCANH compared to 251 manually assigned residues. In contrast, the 4Ds yield significantly higher rates of assigned peaks (Table 3.5). The better combinatorial power of FLYA compared to a manual strategy

Table 3.5: Assignments obtained for every individual experiment compared to the overall provided peaks

Experiment	Peaks	
	provided	assigned
HNcoCANH	271	218 (80.4 %)
hCOCANH	425	347 (81.7 %)
hCACONH	747	457 (61.2 %)
hCACBcaNH	808	437 (54.1 %)
hCACBcacoNH	638	380 (59.6 %)

likely fills the assignment gaps almost exclusively based on the peaks provided by the 4Ds and ties together those loose assignment branches that otherwise will be highly challenging to complete. Many of the assigned longer dark blue stretches start or end in sections that were previously assigned, also indicating that FLYA can continue where the manual approach more likely fails, and moreover it can be seen that often the end of green sections, respectively the beginning of dark blue sections, coincides or overlaps with the end of stretches being assigned from the HNcoCANH peaks list. Instead, only individual pairs can be matched rather than a longer chain. These results, however, also raise the question if FLYA could work equally well without the 5D HNcoCANH or when using a 4D HNNH to link adjacent residues instead of the 5D. Although not used for the manual assignment of TS, a 4D HNNH experiment was recorded by Dr. Petra Rovó using exclusively CP for magnetization transfer within approximately 8 d. The experiment was peak-picked in the same semi-automatic procedure used for all other experiments.

Visual inspection of both modifications shows similar results compared to the original set-up. Most apparent is the reduction of *strong* assignments in the α-subunit, e.g., from residues 100 to 110 or between residues 310 to 320 in the β-subunit, as well as a slight reduction of *strong* assignment matching with the manual assignments like it is found in the N-terminal region.

More in-depth analysis shows that the number of *strong* assignments is reduced by 226 resonances to 1790 and by 95 to 916 for matching assignments, respectively, when using the 4D HNNH experiment. Consequently, 117 fewer extra resonances assigned by FLYA are found, while on the other hand, the number of different resonances is decreased to 39 rather than 53 in the original setup (Table 3.6). Ultimately, the incorporation of the 4D HNNH causes only a modest decrease, about 12 %, in the assignment rate, which substantiates the assumption that the 5D is crucial for the manual assignment approach rather than for FLYA. The extent of assignment per spectrum is essentially unaffected (Table 3.7). Similar to the 5D, the number of peaks assigned in the HNNH falls far behind the other experiments for which multiple reasons can be assumed. First, the peak list might be of insufficient quality.

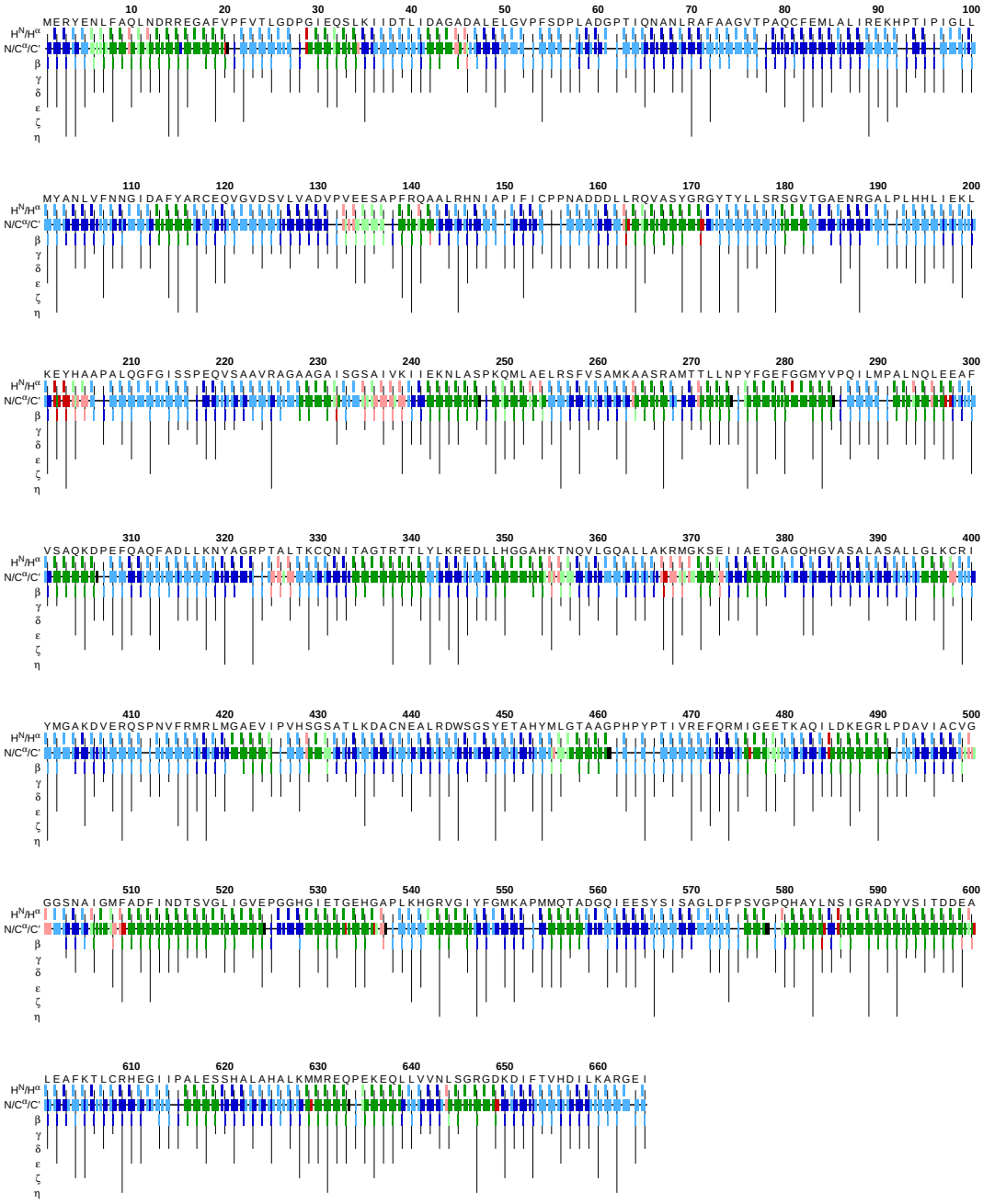


Figure 3.30: Results of the FLYA run using all 4Ds used for manual assignment and a 4D HNNH instead of the 5D HNcoCANH. Compared to the full, original set-up, only little changes can be found by visual inspection. FLYA parameters were not changed, and the color-coding is as described previously.

Table 3.6: Change of *strong* resonances found by FLYA when using the 4D HNNH experiment instead the 5D HNcoCANH for assignment purposes.

	HNcoCANH	HNNH	Δ
Identical	1011	916	-95
Different	53	39	-14
Additional	952	835	-117
Assigned	2016	1790	-226

Table 3.7: Assignments obtained for every individual experiment compared to the overall provided peaks when using the 4D HNNH and in comparison to the original set-up.

Experiment	Peaks		
	provided	assigned	Δ
HNNH	503	286 (56.9 %)	+50[a]
hCOCANH	425	350 (82.4 %)	+3 (+0.7 %)
hCACONH	747	460 (61.6 %)	+3 (+0.4 %)
hCACBcaNH	808	433 (53.6 %)	-4 (-0.5 %)
hCACBcacoNH	638	376 (59.6 %)	-4 (-0.6 %)

[a]compared to the HNcoCANH

However, this would, in principle, apply for all facilitated peak lists, and the HNNH was peak picked by the same procedure used for the other spectra. Second, and in contrast to the 5D, which had only 278 peaks provided overall, it is more likely that the H/N resonances are more difficult to match to the other peaks as for every H/N pair, more matches can be found than for an H/N/Cα triplet. This is directly confirming the findings from the histograms presented in section 3.1. Consequently, the HNNH presumably is of help for FLYA to find adjacent residues but is challenging to combine with other experiments into a consensus assignment and ultimately might only serve to increase the redundancy of assigned resonances rather than helping to identify additional ones.

Based on the original set-up and the results obtained from the 4D HNNH modification, it can be hypothesized that amide-to-amide experiments are redundant when FLYA is used. Similar to when using the 4D HNNH, the visually apparent changes are considerably small when comparing the FLYA results to the original set-up (Fig.3.31), with only little effect on the agreement with the manual assignment and with more additionally assigned stretches affected. Moreover, the direct comparison of assigned, *strong* resonances shows little influence of the absent amide-to-amide experiment on the assignment success. Moreover, it appears that in comparison to the 4D HNNH (Table ,3.6) essentially no changes can be observed. The number of identical and additional resonances is merely changing by more than 20 resonances, i.e., about 4 residues, when the full assignment of the involved nuclei is considered, indicating that the 5D HNcoCANH has slight advantages of its lower-dimensional pendant, even when using FLYA. In line with the observations made for the consolidated resonances, the effect on the peaks assigned from every spectrum essentially stays constant (Table 3.9), and a bigger change relative to the others can only be found for the hCACBcacoNH experiment, which, however, is most likely of no causality of the missing amide-to-amide experiment, as a reduction of redundancy in the *i+1* amide shifts would equally affect the hCACONH experiment. So far, it can be concluded that when using FLYA exclusively, the 5D HNcoCANH is only of little help when used in combination with all other ^{13}C-dispersed 4D experiments for

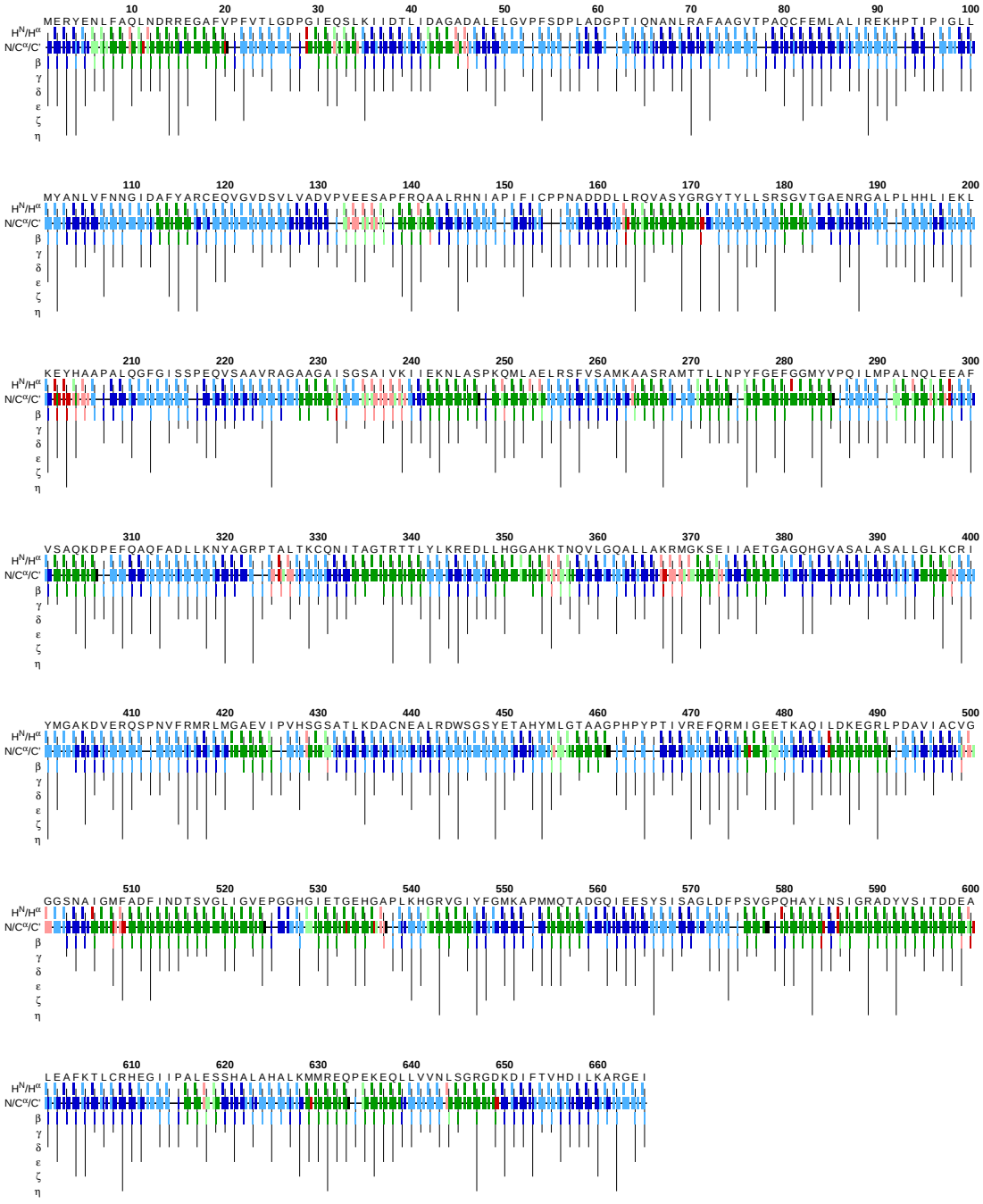


Figure 3.31: Result of the FLYA assignment without any amide-to-amide experiment, i.e., no 5D HNcoCANH or 4D HNNH experiment, being used for the assignment. Similar to the run including the 4D HNNH, almost no changes are apparent, while the comparison to the results of the 4D HNNH set-up shows effectively no change, indicating that FLYA has only little advantage from the 5D HNcoCANH and rendering the 4D HNNH redundant. The colors are as described previously.

Table 3.8: Change of *strong* resonances found by FLYA when using the 4D HNNH experiment instead the 5D HNcoCANH for assignment purposes.

	HNcoCANH	no HN→HN	Δ
Identical	1011	897	-114
Different	53	46	-7
Additional	952	859	-93
Assigned	2016	1802	-214

Table 3.9: Assignments obtained for every individual experiment compared to the overall provided peaks when using the 4D HNNH and in comparison to the original set-up.

Experiment	Peaks		
	provided	assigned	Δ
hCOCANH	425	346 (81.4 %)	-1 (-0.3 %)
hCACONH	747	452 (60.5 %)	-5 (-0.7 %)
hCACBcaNH	808	436 (54.0 %)	-1 (-0.1 %)
hCACBcacoNH	638	368 (57.7 %)	-12 (-1.9 %)

backbone assignment. Replacing the 5D with a 4D HNNH for amide-to-amide correlations brings only little effect on the success of the assignment. Nevertheless, these experiments, in particular the 5D, are indispensable for the manual assignment and, moreover, for the manual verification of the FLYA results. However, a valuable assessment will be how much the number of experiments used by FLYA can be reduced or replaced by the 5D experiment, as it conveys not only amide-to-amide correlations but also the additional carbon chemical shift. This might render some of the 4D experiments contained in the full set redundant and would contribute to a reduction of measurement time. Using a minimal set consisting of the 5D HNcoCANH in combination with the 4D hCACBcaNH for Cβ information and the 4D hCACONH for CO resonances, yields assignments on the same level when using the 4D HNNH over the 5D or when using no amide-to-amide experiment, i.e., only the four ^{13}C-dispersed 4Ds (Fig.3.32). In this context, the 5D effectively relinquishes two 4D experiments, as now only one 5D and two 4Ds are required to obtain the same amount of chemical shift assignments as was achieved before through five 4Ds (HNNH + ^{13}C-dispersed) or four 4Ds (only ^{13}C-dispersed). Although the 5D HNcoCANH contributes the biggest part to the overall measurement time for all backbone assignments (see section 3.3), replacing the two 4D experiments does decrease the overall time that has to be invested by approximately 10 days. While slightly less identical resonances are found, the minimal set exceeds the number of additional *strong* resonances assigned by FLYA as well as the number of overall resonances of the other two setups. The changes are more significant when comparing them to the full original set-up (Table 3.10), however, the smaller success rate is acceptable in the context of potential time savings as well as in comparison to the other two approaches. The redundancy of the 4D hCOCANH and 4D hCACBcacoNH also becomes clear when looking again at the individual assignments made from every peak list (Table 3.11). Compared to the full set, only two assignments less are found for the HNcoCANH, and even more peaks could be assigned from the other two experiments. Taking into account that for the full set, considerably more assignments and less violations are found, while the assignments for the hCACONH and hCACBcaNH are lower, some of the assigned peaks in the minimal set must be ruled out when adding more redundancy and the increase of assignment, therefore, should

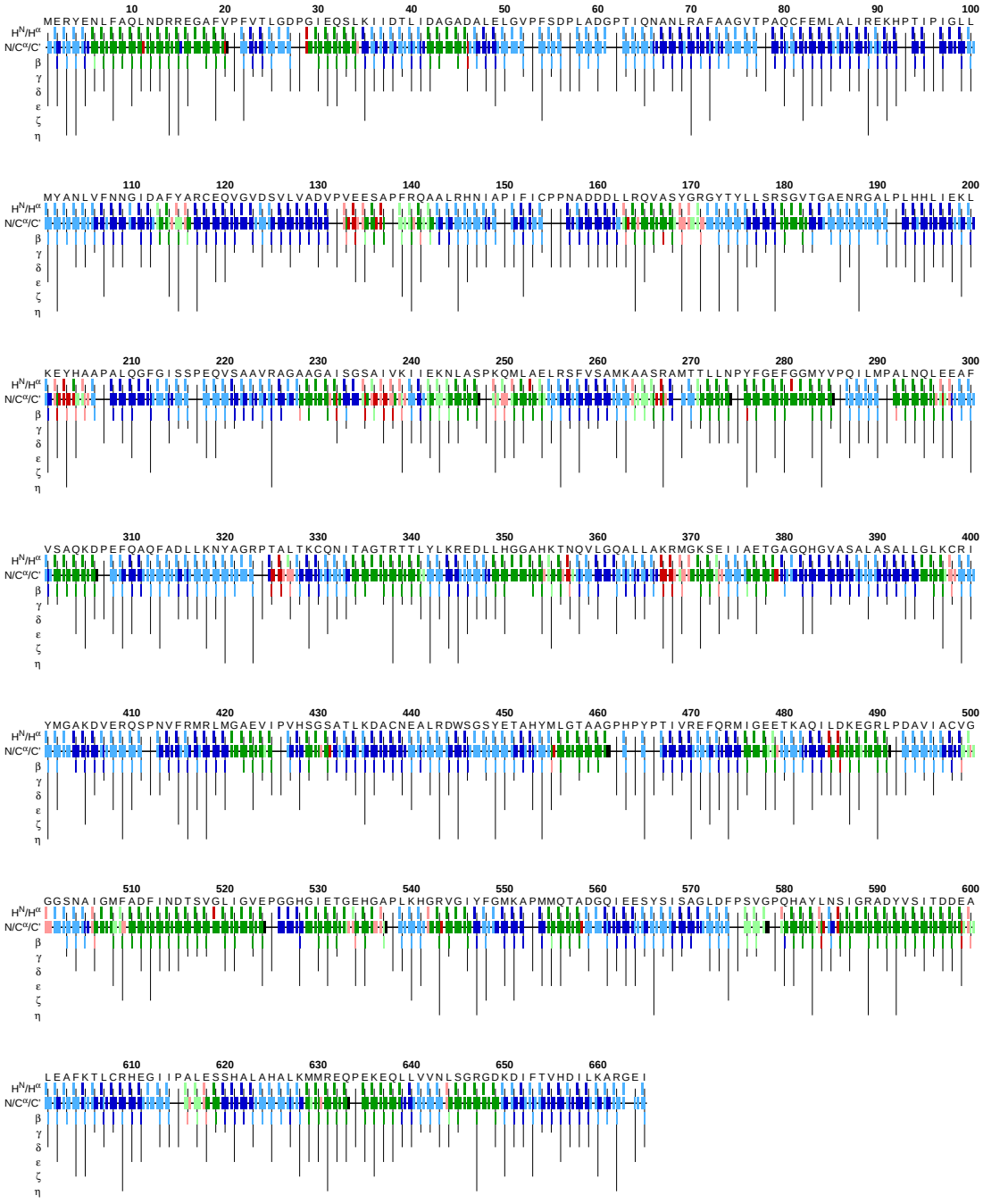


Figure 3.32: Result of 100 independent FLYA runs using only a minimal or reduced set of experiments involving the 5D HNcoCANH, 4D hCACONH, and 4D hCACBcaNH. A good agreement with the manual assignments is maintained, and the level of additional, *strong* assignments is on the level of the runs involving the 4D HNNH or no amide-to-amide correlation, pointing out that two 4Ds can almost fully be replaced by the 5D HNcoCANH. Color coding as described before.

Table 3.10: Change of *strong* resonances found by FLYA when using the 5D HNcoCANH in a minimal set of experiments, i.e. only together with a 4D hCACBcaNH and a 4D hCACONH.

	full	minimal	Δ
Identical	1011	864	-147
Different	53	77	+14
Additional	952	899	-53
Assigned	2016	1840	-176

Table 3.11: Assignments obtained for every individual experiment compared to the overall provided peaks when using a minimal set of experiments in comparison to the original set-up.

Experiment	Peaks		
	provided	assigned	Δ
HNcoCANH	271	216 (79.7 %)	-2 (-0.7 %)
hCACONH	747	480 (64.3 %)	+23 (+3.1 %)
hCACBcaNH	808	475 (58.8 %)	+38 (+4.7 %)

be treated rather carefully. To push this comparison and the reduction of measurement time even further, the HNcoCANH was replaced by the 4D HNNH experiment, which, if successful, would imply a strong reduction of measurement time along with potential simplifications in the workflow. Even a visual evaluation, however, shows already a significant change in the assignment rate and quality compared to all other set-ups. Assignments matching the manual assignments are still abundant, whereas more dark red fields are now visible along with a strong reduction of dark blue, i.e., *strong* additional assignments. Consequently, the overall amount of strong resonances drops substantially in relation not only to the full set of experiments but also compared to the minimal set involving the 5D HNcoCANH or those runs solely based on ^{13}C-dispersed 4D experiments. This result is expected and in line with the observations already made for the incorporation of the 4D HNNH and is likely due to the reduced redundancy of chemical shifts and the increase of possibilities to match resonances, in particular Cβ resonances, to the 4D HNNH peaks. Compared to the full set, 536 fewer resonances are strongly assigned and 360 less compared to the other minimal set, which therefore is a stronger reduction than from that minimal set to the full set, mostly effectuated by the reduction of *strong* identical and additional assignments as violations remain low. As

Table 3.12: Change of *strong* resonances found by FLYA when using the 4D HNNH in a minimal set of experiments, i.e., only together with a 4D hCACBcaNH and a 4D hCACONH.

	full	HNNH minimal	Δ
Identical	1011	803	-208
Different	53	59	+6
Additional	952	618	-334
Assigned	2016	1480	-536

found for the other minimal set, the assigned peaks per peak list stay constant or slightly increase (Table 3.13), but as pointed out above, this is not linked to an increase in assigned resonances or the assignment quality, which is even more obvious for this setup involving the minimal set of a 4D HNNH, 4D hCACONH, and 4D hCACBcaNH.

As the argument for the reduced assignment performance in the minimal set is the reduced redundancy of chemical shifts, a final look can be given to the results when using a reduced set

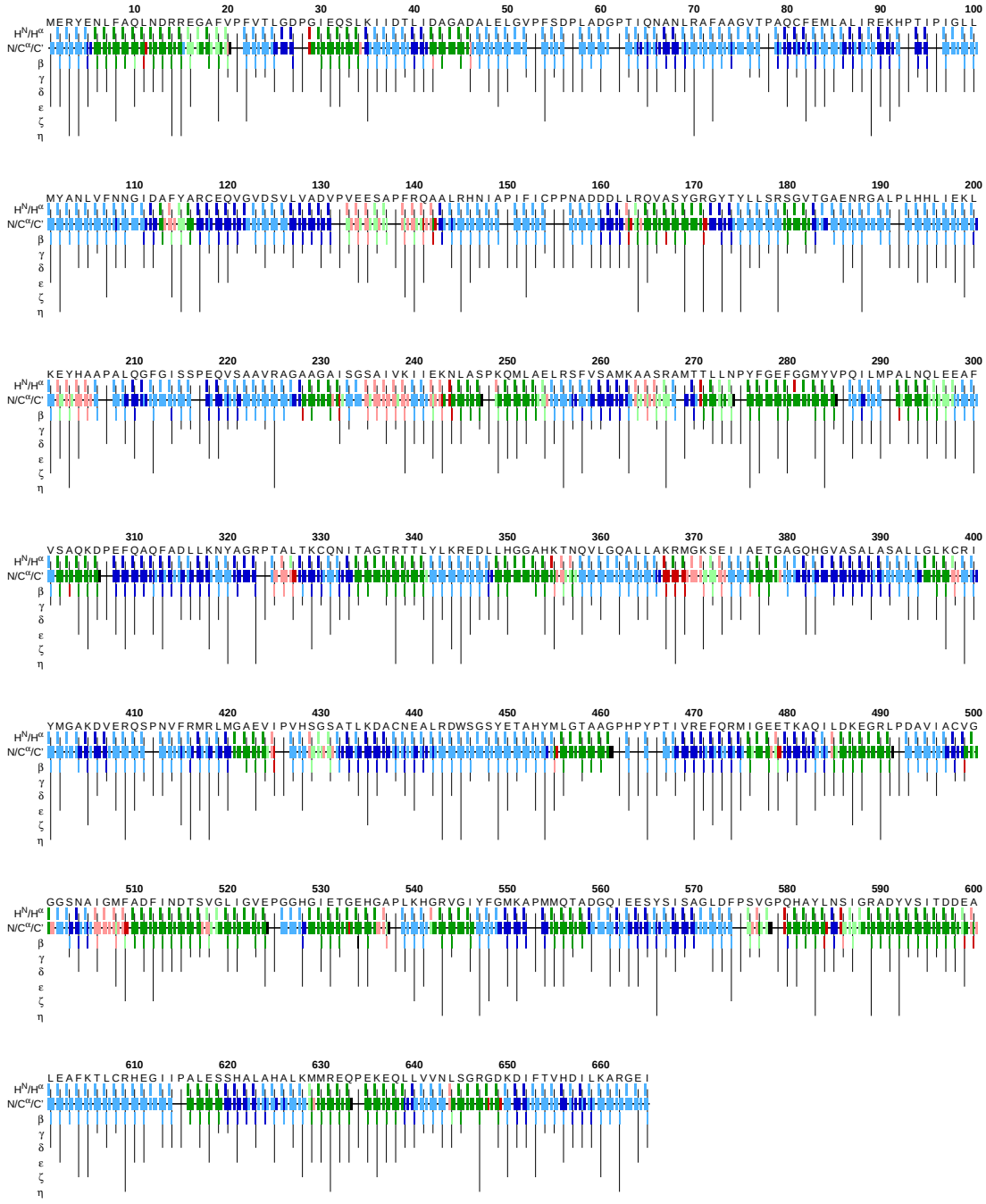


Figure 3.33: FLYA results using a minimal set similar to the previous one (Fig. 3.32), however, replacing the 5D with the 4D HNNH experiment. Here the assignment success is significantly reduced, as less strong assignment are found that match with the manual assignment as well as even more reduced additional assignment. This effect is likely due to the missing redundancy provided through the 5D as shown before for the histograms in section 3.1 (Fig.3.5

.

Table 3.13: Assignments obtained for every individual experiment compared to the overall provided peaks when using a minimal set of experiments in combination with the 4D HNNH and in comparison to the original set-up.

Experiment	Peaks		
	provided	assigned	Δ
HNNH	503	295 (58.7 %)	+9 (+1.8 %)[a] +77[b]
hCACONH	747	466 (62.4 %)	+7 (+1.2 %)
hCACBcaNH	808	447 (55.3 %)	+10 (+1.2 %)

[a]compared to the assignments involving the 4D HNNH experiments
[b]compared to the full set-up, i.e. generally how many HN links are assigned

of experiments, consisting of the hCACONH, hCOCANH, and hCACBcaNH, i.e., no amide-to-amide correlation is used and instead the redundancy of carbonyl and Cα resonances is increased. For such a combination of experiments, however, severe ambiguities even for making a single link, as could be seen exemplarily for the manual strategy as well as in general from the histograms in section 3.1 (Fig.3.3). The outcome of this set is consequently worse, also compared to the minimal set using the 4D HNNH (Table 3.14). It can therefore be followed that in a set with high redundancy in all, i.e., Cα, CO and Cβ, resonances, amide-to-amide correlation experiments are of limited use, while in a reduced set of experiments, they contribute essentially to the number of *strong* assignments that can be reached, with the 5D HNcoCANH uniting redundancy as well as amide-to-amide information proofing to be more even more useful.

Table 3.14: Change of *strong* resonances found by FLYA when using the a minimal set consisting of the 4D hCOCANH, 4D hCACONH and 4D hCACBcaNH experiments, i.e. no amide-to-amide (a2a) correlation is used.

	full	minimal w/o a2a	Δ
Identical	1011	618	-393
Different	53	66	+13
Additional	952	606	-346
Assigned	2016	1290	-726

Figure 3.34 summarizes the results of the different FLYA setups, except the last described set due to its overall insufficient performance. The full original set provides the best results, however, also the ones that would require the longest measurement time. The sets containing the 4D HNNH over the 5D HNcoCANH for amide-to-amide information and the minimal set using the 5D, the 4D hCACONH, and 4D hCACBcaNH show similar performance with only small differences. Finally, the minimal set replacing the 5D with the 4D HNNH performs considerably worse and would therefore be of limited use.

The above section demonstrates the valuable contribution FLYA can add to the assignment process, which was, however, limited to the protein backbone and Cβ resonances. As mentioned in the context of the manual assignment, the hCCNH experiment is often indispensable for the correct identification of the amino acid type, whereas FLYA can omit this information. Incorporating the hCCNH experiment into FLYA can be readily achieved with the extension of the CYANA library as given in section 7.2. Best results could be obtained when FLYA is provided a shift list containing all strong and nonviolating backbone assign-

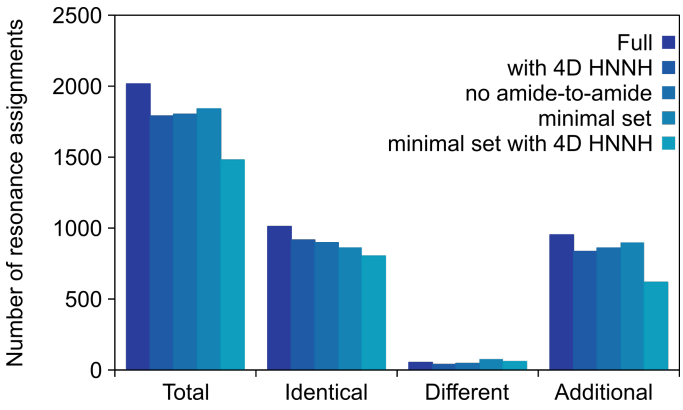


Figure 3.34: Different outcomes of the different FLYA runs depending on the peak lists provided to FLYA. Shown are the number of resonances assigned as *strong* as defined by the stricter criteria in total, in agreement with the manual assignment, different from the manual assignment and assignments additionally found by FLYA.

ments such that the side-chain carbon resonances are added to the given as well as open backbone assignment, however, without altering the backbone assignments too much. Visual evaluation of the FLYA assignments (Fig.3.35) confirms that the backbone assignments are in good agreement with the results without the side-chain experiment. Even though FLYA performs well when it is only provided with Cβ resonances, the additional resonances provided by the hCCNH experiment increase the redundancy of at most four resonances, i.e., H, N, Cα and Cβ, which ultimately leads to confirmations of assignments which were considered weak before as can exemplarily be seen for residues E135-A137 or Y4-E5. Other resonances, however, are now considered unreliable, which, however, only applies for isolated resonances within a residue that is otherwise considered reliably assigned such that incorporating the hCCNH rather closes small gaps in the assignment. Overall, the number of newly added and vanishing assignments that agree with the manual assignments is unaltered upon adding the hCCNH experiment, while the number of violations is slightly increased to 66 resonances. The number of *strong* assignments added by FLYA accounts for 1330, which is a great improvement such that now 2408 resonances are overall assigned (Table 3.15). Another indicator that the backbone assignment was only minor affected can be seen in the peaks assigned from every provided peak list (Table 3.16). Apart from the 5D HNcoCANH, the peaks assigned from every backbone experiment remain almost constant, which is, however, expected as the redundancy of the H/N/Cα shift triplet is even further increased through the side-chain experiment, as is directly reflected in the increase of *strong* assignments.

Table 3.15: *strong* resonances assigned by FLYA when incorporating the 4D hCCNH experiment to the full set of backbone experiments.

	full	with hCCNH	Δ
Identical	1011	1012	+1
Different	53	66	+13
Additional	952	1330	+378[c]
Assigned	2016	2408	+392[c]

[c]expectantly higher as side-chains beyond Cβ are assigned

Peaks from the hCCNH, however, are only assigned by 28 %, and only a few *strong* side-chain assignments beyond Cβ are found. Chemical shift redundancy is high for backbone

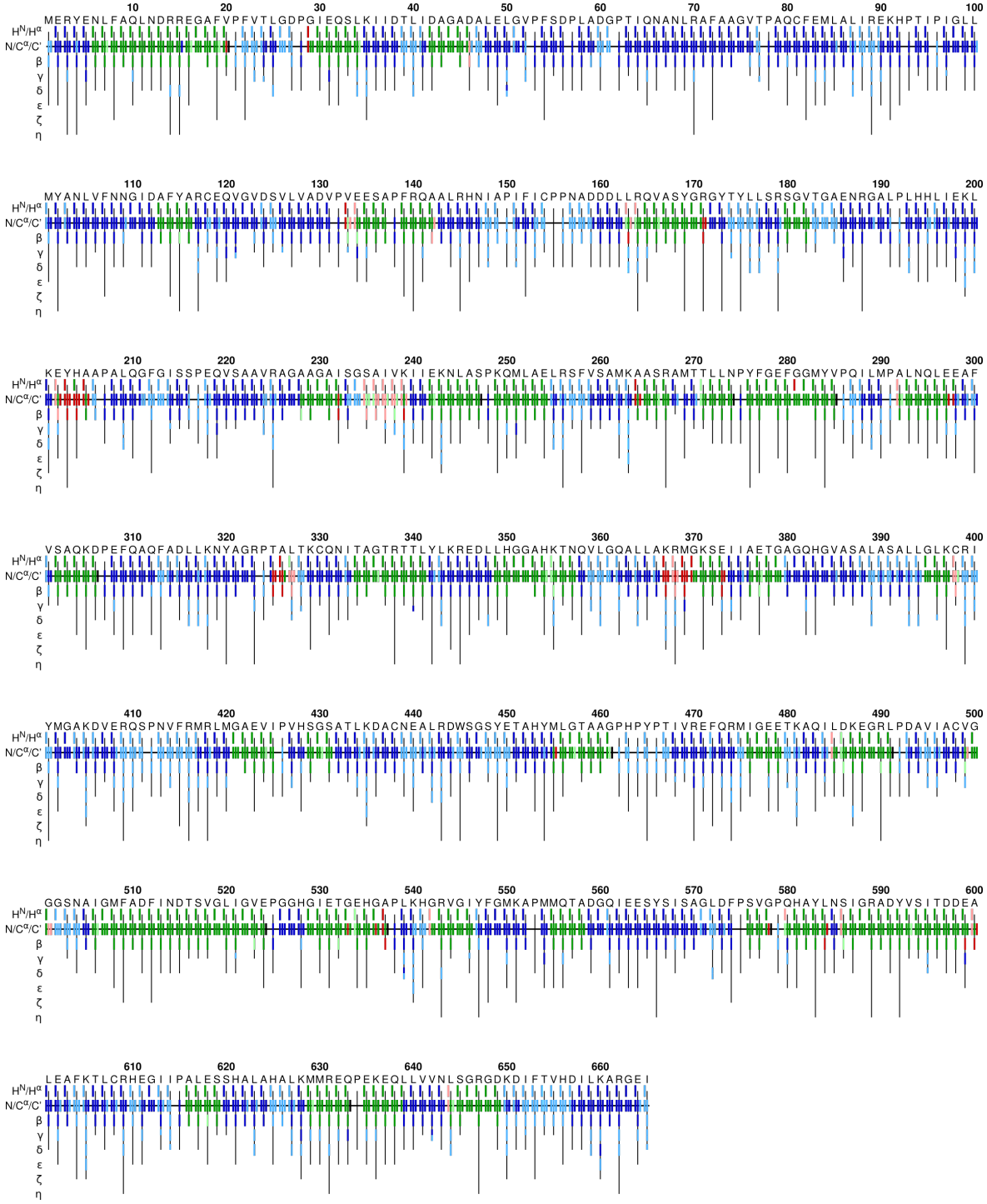


Figure 3.35: FLYA results using all available experiments for backbone assignment and including a peak list from the 4D hCCNH (MOCCA) experiment for the assignment of carbon side-chain resonances. To mimic the manual assignment process, a shift list of all *strong*, non violating resonances from the backbone assignment was provided with the intention that side-chain resonances are rather added and are less affecting the backbone assignment. The consolidation of all assignments into a consensus assignment is according to the stricter rules defined for TS.

resonances but is absent for the side-chain apart from Cβ shift as the hCCNH is the only experiment providing those. Considering the rather low sensitivity of each cross peak, as the bulk signal from every residue is split into the number of side-chain carbons and the comparably poor chemical shift dispersion of aliphatic side-chains, it is unlikely to obtain many *strong* side-chain assignments. The higher impact is likely originating from the low intensity of cross-peaks as the H/N/Cα redundancy with the backbone should generally provide good match-making with the backbone (see section 3.1). These results emphasize

Table 3.16: Assignments obtained for every individual experiment compared to the overall provided peaks when added the hCCNH experiment for side-chain carbon resonance assignment.

Experiment	Peaks		
	provided	assigned	Δ
HNcoCAH	271	205 (75.7 %)	-13 (-4.7 %)
hCOCANH	425	339 (79.8 %)	-8 (-1.9 %)
hCACONH	747	444 (59.4 %)	-13 (-1.8 %)
hCACBcaNH	808	423 (52.4 %)	-14 (-1.7 %)
hCACBcacoNH	638	366 (57.4 %)	-14 (-2.2 %)
hCCNH	1519	425 (28.0 %)	

the problems of peak picking in spectra with poor SNR and their subsequent incorporation in FLYA. While the peaks of the hCCNH were picked in the same semi-automatic approach that was sufficient for the other experiments, it is likely that several of the provided peaks were actually strong noise or artifact peaks which therefore cannot be reliably assigned. On the other hand, manual peak picking is highly challenging in any 4D spectrum, and while H/N/Cα coordinates can readily be facilitated to locate side-chain resonances for already assigned or linked residues, the obtained peak list is highly biased towards residues that are already identified. The hCCNH is, therefore, best used in a manual or computational-aided assignment strategy in which the backbone assignments made by FLYA are used to identify side-chain carbon resonances manually.

Analysis of assignment

Following the manual strategy, it was possible to assign 251 residues manually, i.e., 36 % of the 630 non-proline residues, with a higher rate of assignment of the β-subunit than in the α-subunit in absolute assignments as well as relative to the size of the subunit (Fig.3.36 and 3.38). When plotted against the crystal structure of TS, no direct correlation between structural elements and the success of resonance assignment becomes apparent. It might be concluded, however, that especially the inner regions of the protein remain unassigned as well as regions reported to be rather dynamic and acting as allosteric modulators of the enzyme activity, i.e., αL2 and αL6 in the α-subunit and the COMM-domain the β-subunit[146,215] (Fig.3.36). Residues involved as parts of the intramolecular tunnel and by that are located in a rather hydrophobic environment[148] of TS remain widely unassigned as well, perhaps due to hindered H/D back-exchange in those regions (see section 3.3).

Through the incorporation of FLYA, the rate of assignment can nearly be doubled to 491 residues or 79 % of the non-proline residues (3.40 and 3.37). Nevertheless, the tendencies for successful assignment deduced from the crystal structure remain similar to ones drawn from the manual assignment. The most significant changes might be the widely extended assign-

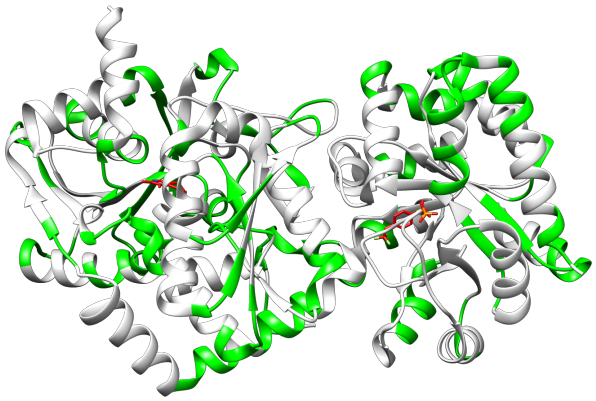


Figure 3.36: All assignments were obtained only through manual assignment (green) plotted on the crystal structure (PDB ID: 4HT3[146]) of TS. The ligand F9F, as well as the PLP cofactor, are shown in red. Structure depiction and editing in UCSF Chimera[110].

ment found for the COMM domain in the β-subunit while α-L2 and α-L6 remain unassigned, as well as most residues involved in the intramolecular tunnel. Besides the inner logical con-

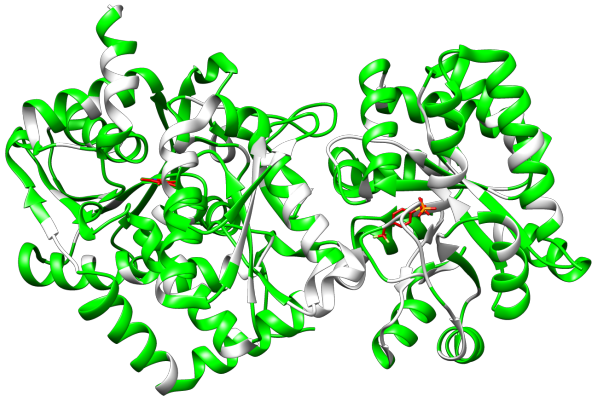


Figure 3.37: All strong assignments obtained by FLYA and manual assignment (green) plotted on the crystal structure (PDB ID: 4HT3[146]) of TS. The ligand F9F, as well as the PLP cofactor, are shown in red. Structure depiction and editing in UCSF Chimera[110].

sistency of the assignment, which ultimately can only be reached for a complete assignment of all resonances, the validation of chemical shift assignments is often done by TALOS (or its newer version TALOS-N) or secondary chemical shifts. The idea of secondary chemical shifts goes back to empirical observations made by Wishart, Sykes, and Richards[216,217] that show systematic deviations of chemical shifts, particularly of Cα and Cβ shifts, from their random coil values.

$$\Delta\delta = \delta_{measured} - \delta_{randomcoil} \tag{3.1}$$

With δ being the chemical shift and $\Delta\delta$ the chemical shift differences. For α-helical structures, the Cα chemical shifts tend to have higher chemical shift values compared to their random coil shifts, i.e., $\Delta\delta$ is positive, while they show the opposite behavior for β-sheets. For Cβ chemical shifts, one can find negative $\Delta\delta$ for α-helices and positive values for β-sheets and therefore the opposite relation of Cα nuclei. Both values can be combined as in the equation below

$$\Delta\Delta\delta = \Delta\delta_{C\alpha} - \Delta\delta_{C\beta} \tag{3.2}$$

Resulting in positive values for α-helices and negative values for β-sheets. Sections of alternating SCS or values close to zero are usually interpreted as loops or unstructured regions as they are close to the respective random coil shifts. The second common technique to validate chemical shift assignments is the before-mentioned analysis by TALOS-N[218]. TALOS-N uses a combination of data mining and an artificial neural network (ANN), the exact methodology being rather complex, but essentially compares measured chemical shifts with chemical shifts of over 9500 high-resolution structures to predict ϕ and ψ angles and other parameters. From the distribution of the conformational space, TALOS-N can provide a probability of the secondary structure, which is particularly reliable for longer stretches of assigned resonances, as not only the residue itself but also the three previous and consecutive residues are considered for the prediction. Angular restraints computed by TALOS-N are frequently used as restraints for structure calculations by NMR as they directly report on the folding and, consequently, the three-dimensional structure of the protein. High agreement of secondary structure elements predicted by TALOS-N and a crystal structure can therefore be used as an indicator for correct resonance assignment. The secondary structure prediction of TALOS, as well as the SCS, obtained from the manual assignments, agree greatly with the secondary structural elements that can be read out from the crystal structure (PDB ID: 4HT3[146]. Most of the deviations can be located at the ends of assigned stretches, e.g., residues D43 and G44 or T271 and K272, making the TALOS results less reliable as not all of the seven shifts required for prediction are available. In fact, for some cases, e.g., T272, the predictions are ranked as *Warn* by TALOS, stating that those residues should be excluded from further analysis. In total, 31 residues are annotated with that warning, form a small but significant subset of 12 % of all manual assignments. Table 3.17 gives an overview of the rankings made by TALOS in the prediction quality of every residue. A large fraction of 77 % of the residues are considered to be *Strong* (73 %) or Generous (4 %). 28 residues show strong deviations from the crystal structure, as can also be seen in figure,3.39 and therefore should be considered more careful as this could be an indication for a potentially wrong assignment in particular when several *bad* residues occur in a row and at the same time, SCS and TALOS secondary structure deviate from the crystal structure, e.g., how it can be seen for residues Y203 to A206.

Table 3.17: Ranking of TALOS predictions for manual assignments. Residues counted as *strong* or *generous* can be considered to be predicted with good confidence, while residues marked with *warn* should be excluded from further analysis. *Bad* only indicates a strong deivation from the reference crystal structure (PDB ID: 4HT3).

	Σ	Strong	Generous	Warn	Bad
abs.	253	184	10	31	28
rel / %	100	73	4	12	11

The SCS, as defined by equation (3.2) widely match with the TALOS results as well as the crystal structure, and for the mismatched cases, usually can be read in a way that substantiates and verifies the assignments rather than falsify them. The examples of T271 and K272 show slightly negative or alternating SCS, which can be interpreted as loop or random coil behavior and while TALOS predicts a loop element for residues I506 to M508, the SCS rather point towards a helical structure. Like for TALOS, more reliable trends can, however, only be drawn from longer stretches, and slight changes in SCS should not be

over-interpreted.

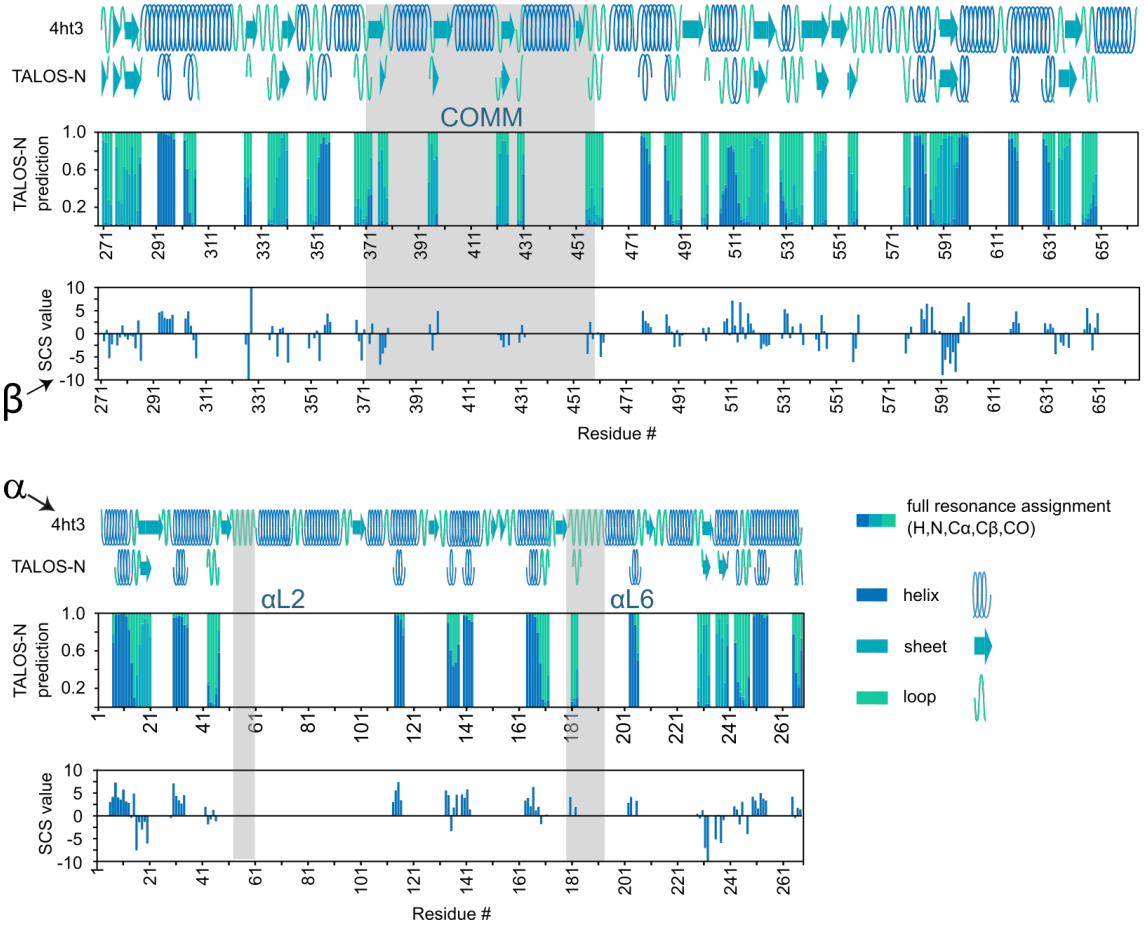


Figure 3.38: Analysis of assigned residues for the β- (top) and α-subunit (bottom) with respect to the secondary structure predicted by TALOS-N[218] (upper rows) and secondary chemical shifts (SCS; respective lower rows), shown only for manually obtained assignments. The TALOS results (blue: helix; teal: sheet; blue-green: loop) and SCS (upward stretches: helix; downward stretches: sheet; alternating: loop) are compared with the secondary structure elements found in the crystals structure 4ht3[146] (top). Highlighted in light grey are domains considered to undergo major conformational changes upon ligand binding at different states of the catalytic cycle[146]. Short stretches, e. g., 506 to 508 or 529 to 533, show minor differences between TALOS and the actual crystal structure. These deviations are thought to arise from TALOS prediction known to be improper in case of short stretches or at the edges of assigned regions.

To achieve a more general overview of the conformity of TALOS-N predictions and crystal structure, the ϕ and ψ angles extracted from the crystal structure were compared with the ones predicted by TALOS-N. While this does not enable a single-residue comparison, it is a better way to compare the outcome of TALOS and the 3D structure, as the secondary structure elements are only derived from the ϕ,ψ angles in both cases. The vast majority of the residues correlate well. However, certain outliers, as well as residues with high estimated standard deviation, are present (Fig.3.39B). Residues with large error bars can be linked mostly to unreliable predictions (*Warn*) rather than a strong deviation from the crystal structure, which becomes clear when those 31 residues are removed from the analysis (Fig.3.39A). The remaining outliers account for the 28 residues marked as *Bad*, with mostly small associated errors. These deviations might be due to incorrect assignment and should be further analyzed, i.e., verified or falsified, e.g., by FLYA (see below).

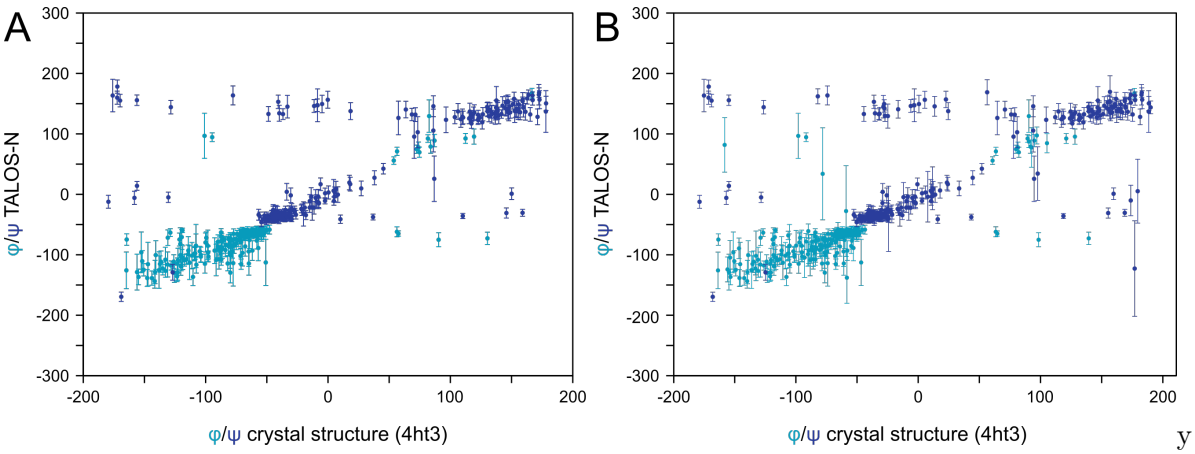


Figure 3.39: Correlation plots of ϕ (light blue) and ψ (dark blue) angles read from out from the crystal structure of TS against ϕ, ψ angles predicted by TALOS-N for manual assignment. Error bars denote the estimated standard deviation given by TALOS. **A)** shows all predictions except those 31, marked as *Warn*, while in **B)** all predictions are included. While a few outliers are removed, residues with high errors are mostly affected, indicating an unreliable prediction.

Using FLYA to assist and evaluate the manual assignment process, the rate of assignment could be increased to 369 fully assigned residues as well as 121 residues, including partially assigned residues (see section 3.3). A combined shift list of FLYA and manual assignments were used for the same analysis as previously done for the manual assignment. Also, here, the secondary structure propensities predicted by TALOS-N agree to a vast extent with the crystal structure, while certain deviations remain. These might still be linked to challenges in the prediction when the end of an assigned stretch is reached or when the predicted propensity is not dominant, e.g., for residues 259/270, which still show a significant likelihood for β-sheet. Due to the larger extent of assignment, stronger, consistent deviations can now be found also for longer sections, e.g., for 531 to 535 or 105 to 111. In addition to previous findings, the SCS rather correlate with the TALOS prediction in those cases and therefore emphasize that the secondary structure might be indeed different or the assignments wrong. As assumed for the manual assignments, the trends for SCS generally benefit from the increased assignments and help to clarify trends that were observed for the manual assignments.

When plotting the correlation between ϕ,ψ angles of the crystal structure against the ones extracted from TALOS-N, the general tendencies are conserved (Fig.3.41). As before, the larger part of residues correlates well, but several outliers, as well as residues with a large estimated standard deviation, are found (Fig.3.41 and Table 3.18). From the 444 analyzed residues, 65 are ranked as *Warn*, slightly more than twice compared to the manual assignments. However, only two additional *Bad* residues are found indicating that FLYA almost exclusively adds correct assignment, always assuming that correct assignment will lead to good correlations between the ϕ,ψ angles.

Removing the unreliable predictions from the correlation plot, it can be seen that most of the residues with a large error as well as several outliers, are removed from the plot. Despite the improvement in the correlation, it should be kept in mind that the assignments made for these residues still might be correct, and only the prediction by TALOS-N is not sufficient. Having FLYA as a tool for verification at hand, it can now be estimated if there is a correlation between the resonances marked in red, i.e., those that deviated from the manual assignments and the remaining outliers in the plot or generally the residues ranked as *Bad*. To do so,

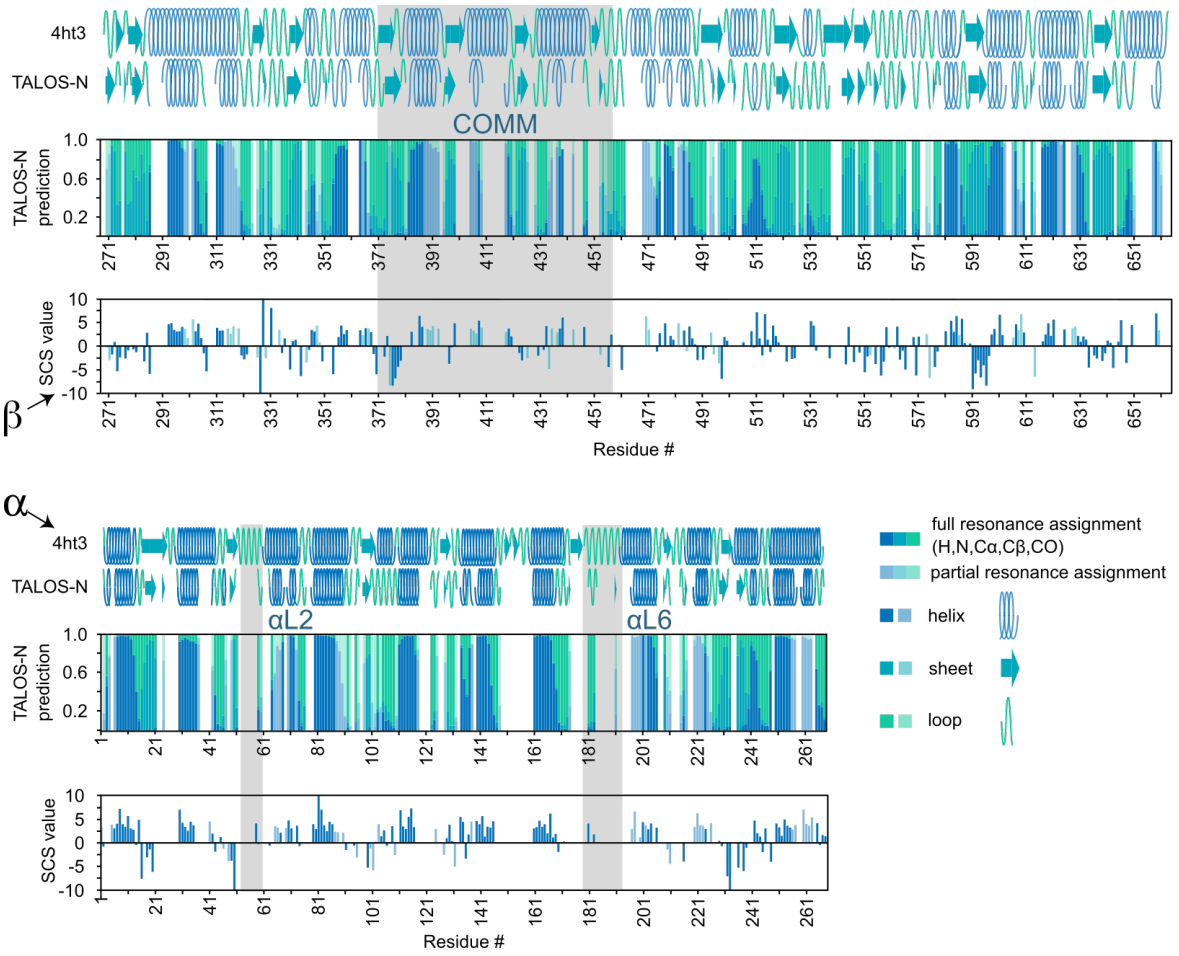


Figure 3.40: Analysis of assigned residues for the β- (top) and α-subunit (bottom) with respect to the secondary structure as in figure 3.38 for the combined assignments obtained by FLYA ("strong" assignment in FLYA nomenclature) as well as manually. Not-opaque colors denote all of HN, N, Cα, CO, and Cβ assigned as "strong", opaque colors denote that at least part of the nuclei have "strong" assignment). The coloring is accordingly to figure 3.38. Still, most of the residues show a good agreement between prediction and crystal structure, but some stretches indicate deviations, such as for residues 101-110.

all residues were excluded that contain any resonances which do not agree with the manual assignment irrespective of being marked as *strong*. In total, 45 residues were affected by that. While 11 outliers are removed, the other 34 removed residues are accounted for in the two categories that show good agreement with the crystal structure. Under the assumption that the FLYA assignments are correct, i.e., resonances or residues marked in red are indeed incorrect, and all others are correct, it can be concluded that indeed some residues show significantly different torsion angles compared to the crystal structure, e.g., residues 101-110 show strong TALOS predictions that correlate with SCS and FLYA assigns a long stretch. On the other hand, however, it can also be assumed that the (reliable) predictions made by TALOS are actually correct, and additionally, no strong deviations in the structure are to be expected, which would indicate that more or even other residues than the red-marked ones are incorrectly assigned by FYLA or manually. This scenario can be found, for example, for residues 234-239, where the FLYA assignments tend to be inconclusive and match only partially with the manual assignment while simultaneously the predictions are ranked as *Bad* or *Warn* and are ultimately removed in figure 3.41C. As expected, all outliers, except three WHY, are disappearing when the *Bad* ranked residues are removed, leaving a cleaned-up

Table 3.18: Ranking of TALOS predictions for combined FLYA and manual assignments. Residues counted as *strong* or *generous* can be considered to be predicted with good confidence, while residues marked with *warn* should be excluded from further analysis. *Bad* only indicates a strong deviation from the reference crystal structure (PDB ID: 4HT3).

	Σ	Strong	Generous	Warn	Bad
abs.	484	370	19	65	30
rel / %	100[a]	76	4	13	6

[a] Deviations due to rounded numbers

Table 3.19: Ranking of TALOS predictions for combined FLYA and manual assignments with all residues marked as *Warn* by TALOS and marked as red by FLYA removed.

	Σ	Strong	Generous	Warn	Bad
abs.	373	339	15	0	19
rel / %	100	91	4	0	5

correlation plot. While this is a trivial finding, it is more interesting to see where those residues actually are found in the sequence as well as what the FLYA assignment quality is for those (Table 3.20). Interestingly, the assignment is most likely correct, based not only on the analysis by FLYA (dark green or dark blue) but also by manual inspection of the assignment as far as possible, for most of the residues ranked as *Bad* while three to four should be rather considered as a tentative assignment. Nevertheless, manual inspection, in the sense of following or confirming the assignment route done by FLYA of the additionally assigned residues, does confirm them. It should, however, be emphasized again that the stringed requirement for the uniqueness of an assignment or link is mostly rendered through the 5D experiment in the manual process and of which assignments are often absent for the additionally assigned residues (see section 3.3), and therefore, the assignments made by FLYA can rather be followed than proven.

As becomes clear, a final and entirely conclusive verification or falsification of the assignments is hardly possible as neither the tools used for assignment nor the tools used for verification are perfect. Nevertheless, an overwhelming majority of residues most likely can be considered to be assigned correctly while at the same time, a few residues or assigned regions show possibly less reliable assignments.

Comparison with solution-state NMR assignments of the α-subunit

In contrast to the full $\alpha_2\beta_2$ construct, backbone chemical shift assignments are available for the isolated α-subunit in solution[215] which can be used to compare and verify the NMR assignment (Fig.3.42). Significant changes in the assignment are apparent, highlighted in (dark) red colors, which account for 306 out of 756 strong resonances, while 393 resonances agree between the two samples. The differences mostly occur in rather defined regions, i.e., sections of the primary sequence of the protein, as can be seen, e.g., for residues N5 to V20 or I97 to N109 , indicating a correlation between the similarity of chemical shifts and those sections.

Disagreement can be found at different places throughout the primary sequence. However, especially the N- and C-termini of the α-subunit are in good agreement. The isolated α-subunit in solution can tumble freely, presumably without any permanent interactions be-

Table 3.20: Analysis of the remaining, 19 *Bad* residues

Residue	Assignment results	Rating
2Glu	dark blue short stretch	Probably correct
28Pro	dark blue isolated; no redundancy	Tentative
173Tyr	dark blue end of long stretch no Cβ	Probably correct
180Ser	dark green fixed assignment (manually) $\rightarrow$ not affected by FLYA	assumed as correct
181Gly	see Ser180	
272Leu	dark green short stretch	Probably correct
275Pro	dark blue isolated; no redundancy	Tentative
282Gly	strong green long stretch	Probably correct
330Cys	dark blue long stretch	Probably correct
338Arg	strong green long stretch	Probably correct
372Ser	strong green short stretch	Probably correct
379Gly	strong green long stretch	Probably correct
494Ala	partially dark blue mixed / isolated stretch	Tentative
500Gly	dark green mixed assignments in stretch	Tentative / correct
548Phe	dark blue end of stretch	Probably correct
552Ala	dark blue end of stretch	Probably correct
575Pro	not assigned[a]	Ignore
589Arg	strong green long stretch	Probably correct
634Pro	not assigned[a]	Ignore

[a]Prediction filled in automatically by TALOS-N

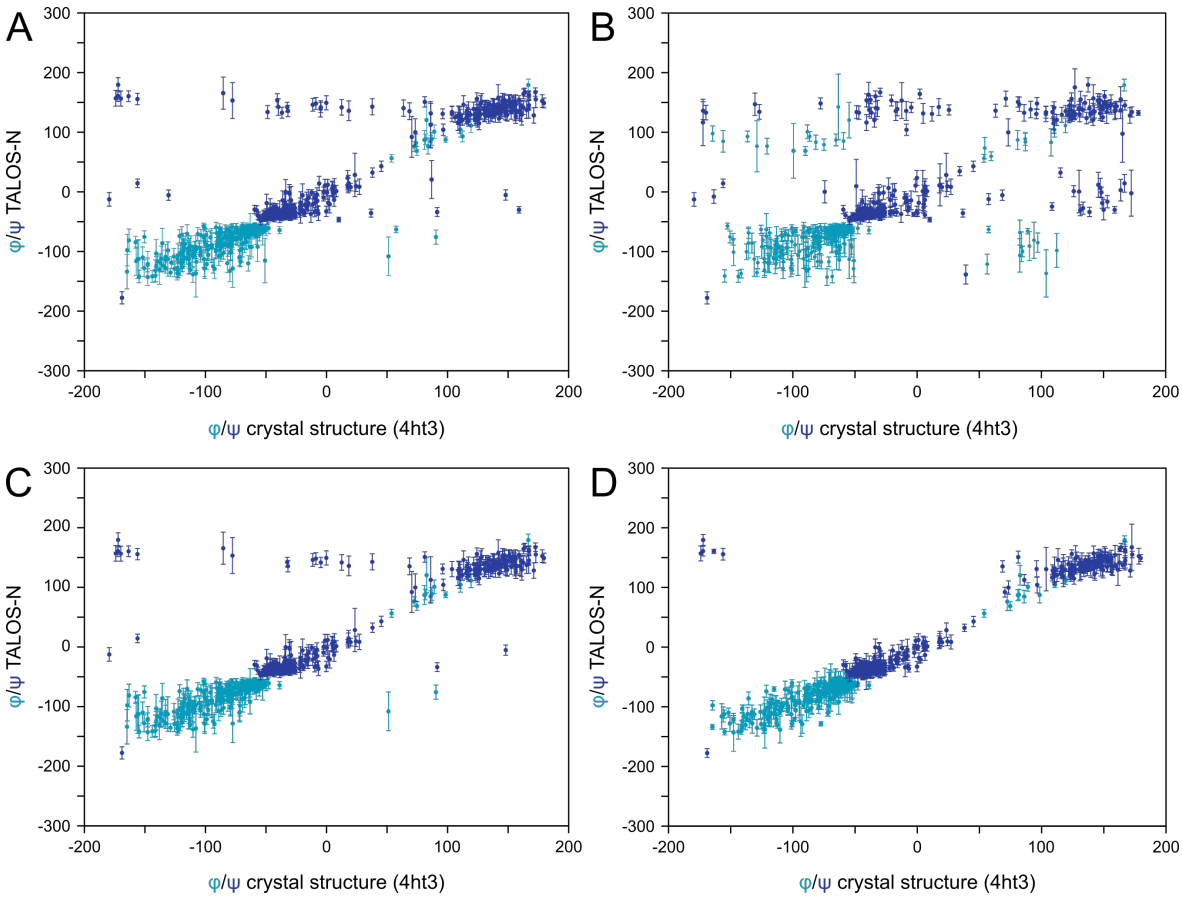


Figure 3.41: ϕ (light blue) and ψ (dark blue) correlations plots of angles obtained from the crystal structure (4ht3) plotted against values predicted by TALOS-N as in figure 3.39 for the combined FLYA and manual assignments. **A)** all predictions marked as not-reliable (*Warn*) by TALOS are excluded while in **B)** all residues are included. In **C)** additionally, all residues in which FLYA and the manual assignment disagree were removed, however, leaving several outliers. The error bars denote the standard deviations suggested by TALOS-N. **D)** ultimately shows the correlation when all *Bad* residues are removed, only leaving three outliers.

tween the individual molecules. The microcrystalline sample, although, has no free tumbling and the additional β-subunit and therefore will have one interface each to the adjacent β-subunit as well as to the next α-subunit in the crystal, i.e., the unit cell (Fig.3.43). It is therefore expected that the chemical shifts of residues located close to the interfaces, as reporters of the direct chemical environment, will be affected and eventually change compared to a solution-state sample.

Residues located at the $\alpha\alpha$ interface include, among others, R140-N147 on one molecule and E83 to E90 on the neighboring molecule with E83 as the residue, which is the most distant one (Fig.3.43A). The corresponding regions in the primary sequence show rather big differences between the two samples, and while the assignments around Q80 are still in good agreement, the deviations increase the closer the residue is located to the surface. Similar conclusions can be drawn for residues R140 to N147 that form an α-helix that runs parallel to the $\alpha\alpha$interface (Fig.3.43A), and consequently, the previously described trend for distant residues is not observed. Nevertheless, big differences in the assignment are found for residues F139-N147. More changes can be observed on the larger $\alpha\beta$ interface. Especially, the long sections from L58 to G77 as well as G98 to I111 can be located to a large extent along this interface running from top to bottom (Fig.3.43B). Ultimately, this

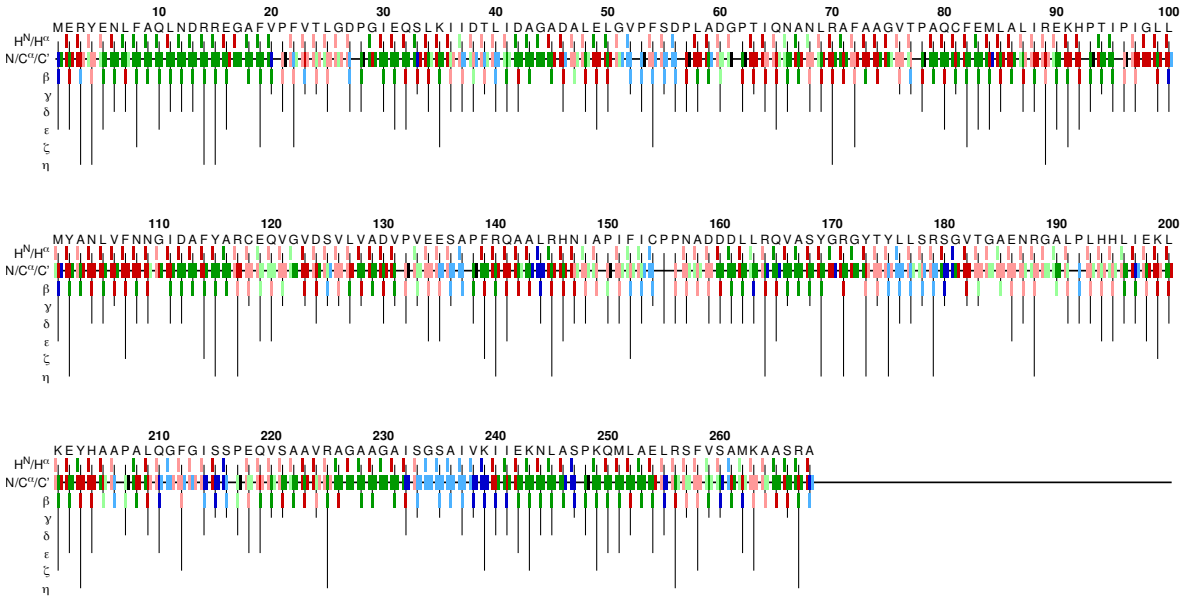


Figure 3.42: FLYA results as shown in figure 3.29 (section 3.3) in comparison to solution-state NMR assignments of the α-subunit (BMRB ID: 50139[215]). Here, the assignments of the full set of solid-state NMR experiments were simply compared with the solution-state NMR assignments of the α-subunit as the last step of a FLYA run. No other modifications were done and the color coding is as previously described. Resonances that do not match (red) are indicated as such, if the chemical shift difference is larger than the defined tolerances, i.e., 0.1 ppm, 0.5 ppm and 0.8 ppm for ^{1}H, ^{15}N and ^{13}C, respectively. When considering only *strong* resonances, about half match while the other half differs. Manual assignments tend to match better between solution and solids assignments. ^{1}H chemical shifts of the solution-state NMR assignments were re-referenced by a downfield shift of 0.6 ppm. For reasons of clarity, the β-subunit assignments are not shown.

leaves only the assignments from E199 to A205 with big differences that are not located close to some interface along with several isolated resonances or short stretches with mixed strong and weak assignments as well as disagreeing and matching resonances. Similar to the verification previously described for microcrystalline $\alpha_2\beta_2$ TS, a final conclusion about the more tentative or mixed assignment results is difficult and even more in this particular case as too many factors differ. While interactions in the crystal and along intramolecular surfaces play a significant role, temperature effects and changes in the buffer composition, i.e., crystallization buffer vs. solution-state NMR buffer, have to be considered as well. Furthermore, the assignments of the α-subunit were done through the PINE algorithm[219,220], i.e., by automated assignment, and while certainly both algorithms can provide strong support for the assignment procedure, differences in the results are likely to occur. To this point, PINE does not support the backbone experiments used for solid-state NMR assignments, and an evaluation of the two algorithms remains elusive. Nevertheless, the high correlation in regions distant from interfaces indicates the correctness of the assignment, and even more than for those sections, solid-state and solution-state samples are indeed comparable.

Assessment of H/D exchange in deuterated TS

While 4900 residues are a substantial amount of assignment, 140 non-proline residues remain elusive, assuming all other residues are assigned correctly. Several reasons come into question for those missing assignments. As can be seen from figure 3.40, a large fraction of the unassigned residues can be located in protein regions that are considered to undergo major

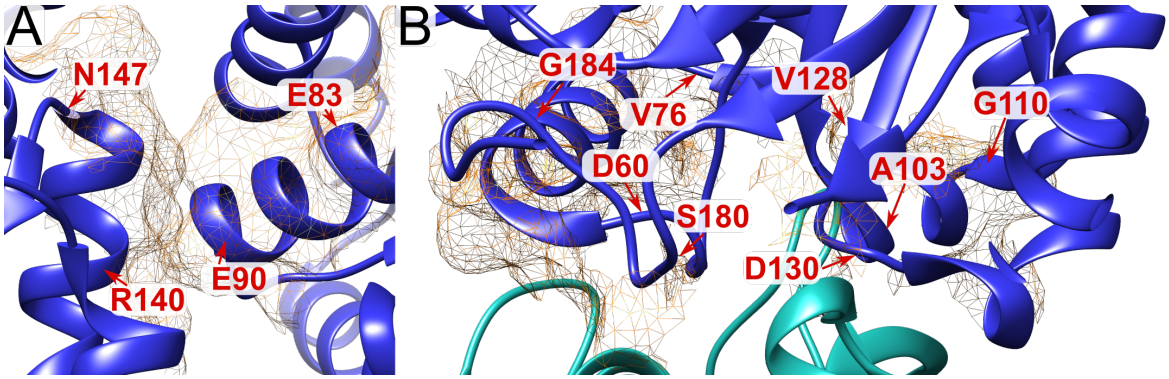


Figure 3.43: Interfaces of **A)** two α-subunits (dark blue) in the crystal cell and **B)** the α and β-subunit (light blue). Residues labeled in red mark the beginning and end of regions located near the interfaces that show deviations between the solution-state NMR assignments of the α subunit and solid-state NMR assignments of the $\alpha_2\beta_2$ construct. Mesh surfaces in orange highlight the spatial proximity. Structure depiction and editing in UCSF Chimera[110].

conformational changes in the catalytic cycle, in particular, the COMM domain and loops αL2 and αL6 (see also section 1.5.3) which will probably maintain certain flexibility, even in the present, trapped state of the internal aldimine.

Deuterated and fully back-exchanged samples, as used for the assignment, can be prone to incomplete exchange of the respective amide groups. Especially for rather isolated, hydrophobic regions in the core regions of the protein, as they are found in TS, e.g., in parts of the intra-molecular tunnel[148], a risk for remaining deuterons remains and therefore decreasing or completely quenching the signal intensity coming from those residues. To estimate the extent of incomplete back-exchange, H/D exchange experiments by mass spectrometry (MS) were performed. For reasons of practicability, the amount of exchange in water-exchangeable sites, i.e., amides, was looked at in a reversed way, i.e., a fully protonated sample of unlabelled, wild-type TS was exposed to deuterated buffer for different times from 15 sec to 14 d, and the exchange was monitored by HPLC-MS. The long incubation times were used to mimic as good as possible the conditions of the NMR sample, which ultimately is exposed to a protonated buffer during several weeks before being filled into the NMR rotor. Although the environment of the samples differs, the H/D exchange (HDX) experiments are expected to give tendencies about back-exchange problems. For experimental details and optimization of conditions for TS, refer to section 6.4.

Despite an incubation time of two weeks, the HDX experiments indicate an incomplete exchange of protons to deuterons and by that, pointing out that hindered back-exchange might generally be a problem for TS. At the same time, the relative exchange never exceeds 50 %, even for water-exposed surface regions of the protein, placing a strong contrast to the expected outcome that at least those regions, among others, will be fully exchanged after 14 d, simultaneously emphasizing that the results of the experiment rather give tendencies. Therefore, the maximum uptake of deuterium over the entire sequence is set to 100 %, and the other residues are normalized accordingly. Furthermore, it should be noticed that in theory, a single-residue resolution could be obtained while the actual outcome is mostly averaged over several residues of 5-10 amino acids (Fig. 3.45[221]. Nevertheless, some tendencies can be found when depicting the relative exchange onto the crystal structure (PDB ID: 4HT3[146]) of TS.

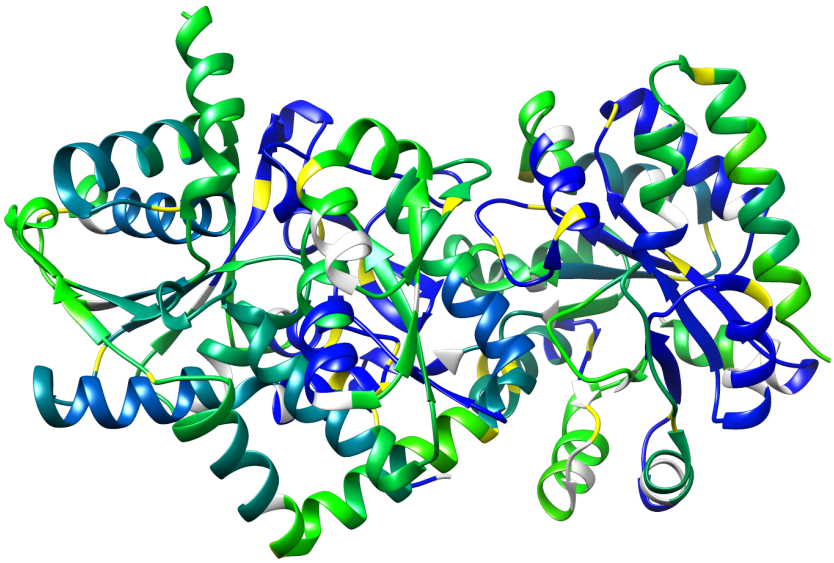


Figure 3.44: Extend of H/D exchange after 14 d (not normalized) plotted on the crystal structure of TS (PDB ID: 4HT3[146]). Blue regions indicate low exchange, and green regions indicate higher exchange rates up to 50 %. Prolines are excluded by definition and are highlighted in yellow. While the core regions expectantly show low exchange, especially the water exposed surface of α-subunit shows higher exchange levels. The β-subunit shows a similar pattern. Despite general accessibility of the protein surface, except crystal-crystal contacts, some of those solvent-exposed regions, however, do not show strong exchange as even the dimer-dimer interfaces seem not to be affected by hindered exchange. Structure depiction and editing in UCSF Chimera[110].

Analyzing the relative deuterium uptake with respect to the protein primary sequence in the context of assigned proteins can be used to deduce hindered exchange is one of the reasons for missing assignments. Taking into account cryogenic b-factors as reporters for molecular displacement as well, gives hints towards conformational flexibility, i.e., protein dynamics, in those regions with elevated b-factors, which effectively render CP transfers less efficient and therefore could lead to missing signals. For most assigned regions, a considerable deuterium uptake can be found (dark blue stretches in figure 3.45) while several regions remain unassigned despite high exchange rates, e.g., residues 183-195 or 409-415, and at the same time longer, assigned regions are found in regions that show apparently low uptakes rates placing a strong contrast to the expected outcome. A similar picture can be drawn when considering b-factors in addition. Two important loops in the α-subunit, αL2, and αL6, as well as the COMM domain in the β-subunit, remain widely unassigned and exhibit elevated b-factors, but in contrast, the region from residues 521 to 584, featuring similar b-factors as residues in the COMM domain, is assigned to a large extend while simultaneously show almost no H/D exchange.

It can therefore be concluded that hindered backexchange, as well as potential conformational flexibility, constitute pitfalls for the assignment process as they largely or completely quench the signal, but in the end, only provide tendencies rather than strong correlations and causalities.

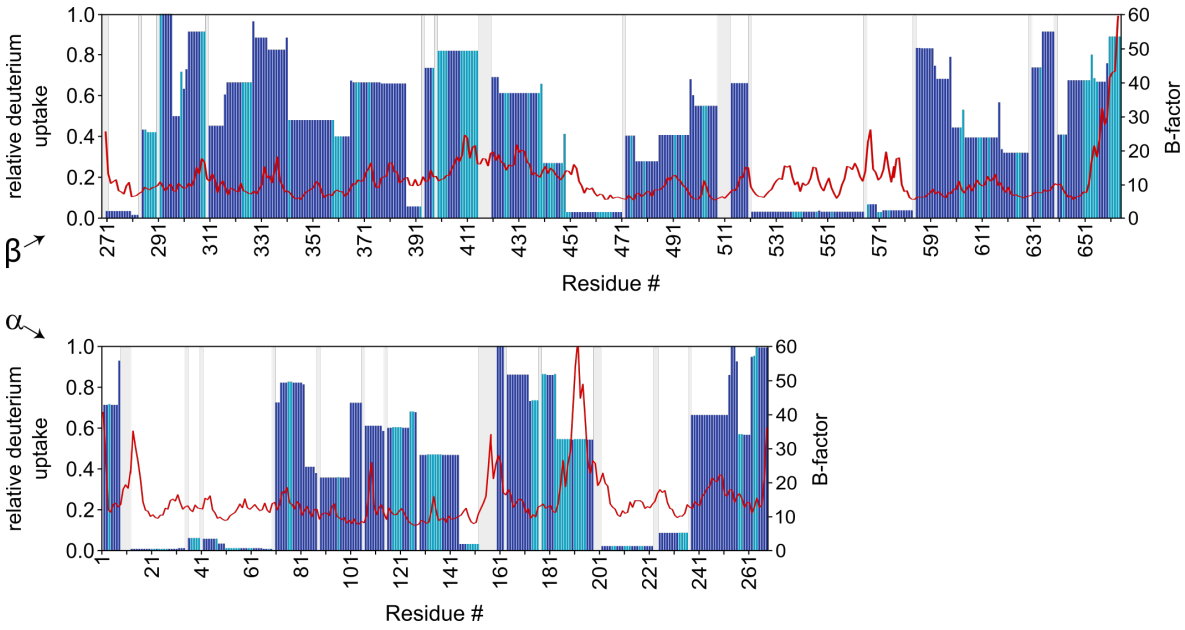


Figure 3.45: Relative deuterium uptake after 14 d(blue tones) and cryogenic b-factors (red) plotted on the sequence of TS for α-subunit (bottom) and β-subunit (top). The deuterium uptake was normalized to one instead of using the relative uptakes provided by DynamX (see text). Gaps (grey rectangles for better guidance) indicate missing HDX data rather than a missing assignment in contrast to , e.g., figure 3.40. Dark blue stretches mark assigned residues while teal ones indicate missing assignments. If missing assignment would be exclusively linked to insufficient back-exchange, assigned regions should show higher uptake compared to unassigned regions. Regions with high cryogenic b-factors might have higher flexibility and therefore can serve to indicate the presence of dynamics in those regions of TS which at the end can lead to decreased CP efficiencies.

Performance of fully protonated TS

As unfolding/refolding protocols remain elusive so far, the preferred labeling scheme to circumvent potential problems with HDX would come as a fully protonated sample, making back-exchange redundant. At the same time, Hα nuclei, as well as side-chain protons, could be facilitated as an additional information source. When increasing the MAS frequency to $>$ 100 kHz, the line widths are expected to remain comparably narrow[12,51], and in the context of the high dispersion obtained from higher dimensional experiments, peak overlap is not expected to put limitations for such a sample. However, the combination of a denser ^{1}H-^{1}H dipolar network, shortening T_2 relaxation times, and the required higher spinning speed will reduce transfer efficiencies, and even though less sample is needed, the RF coil design cannot fully compensate the smaller rotor volume[12]. Those factors will ultimately attenuate the sensitivity with a direct effect on measurement time or general feasibility. To make a qualitative assessment of the impact on measurement time as well as changes inflicted by a fully protonated sample, a small set of experiments, 2D hNH and 4D hCOCANH, were recorded on fully protonated $\alpha_2\beta_2$ TS in 0.7 mm rotor at 111 kHz MAS and a ^{1}H Larmor frequency of 800 MHz. The line widths present in the fully protonated sample range typically from 50 to 100 Hz for ^{1}H (Fig.3.46B) as expected and underline the high performance of MAS frequencies of $>$ 100 kHz. As expected, the reduced amount of sample in the rotor in combination with lower transfer efficiencies decreases the sensitivity significantly around a factor of three in the simple case of a 2D hNH experiment and about a factor of eight in the case of the 4D hCOCANH, despite the higher magnetic field used for the fully proto-

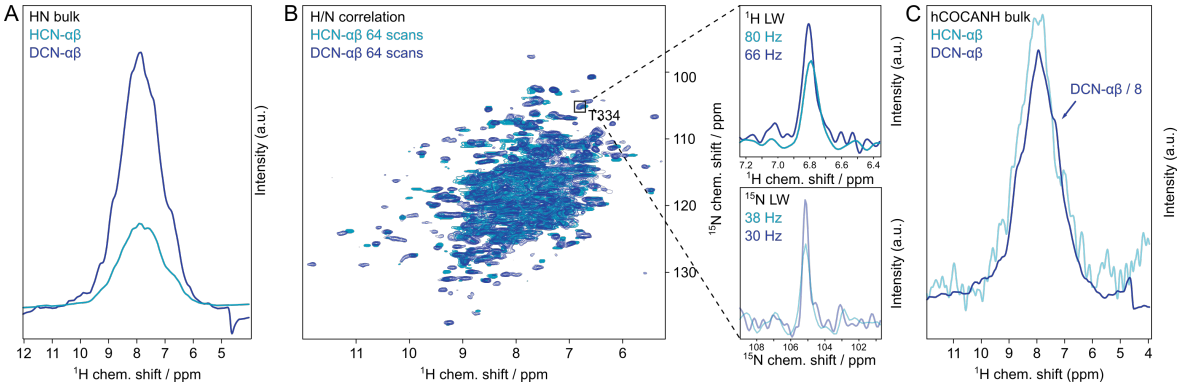


Figure 3.46: Changes of sensitivity between a 100 % back-exchanged sample at 55 kHz MAS (dark blue) and a fully protonated sample at 111 kHz MAS. **A)** H/N bulk signal of the first FID in an hNH experiment adjusted to the same noise level. Both FID were processed with 100 Hz line-broadening. **B)** Full hNH spectra of the respective samples and representative slices through the isolated signal of T334. Both spectra were processed using a squared sine bell with SSB values of 4 and 3 in the ^{1}H and ^{15}N dimension with effective acquisition times of 20 ms and 28 ms, respectively. The spectra show no apparent differences except slight differences in chemical shifts for individual signals, which most likely are due to temperature differences. The extracted slices were again scaled to the same noise level to allow for direct comparison under the given processing parameters. The line-width at half-height was, however, extracted from the same spectra without any window function. As expected[42], the fully protonated sample shows broader lines due to stronger, remaining dipolar interactions. **C)** Qualitative comparison of signal intensity in a higher dimensional experiment (4D hCOCANH). Shown are the first FIDs of the experiment processed with 100 Hz line broadening with the signal of the back-exchanged sample reduced to 1/8 for better visualization. Despite the fully protonated sample being acquired at 800 MHz ^{1}H Larmor frequency, compared to 700 MHz of the other, significant drops in intensity can be seen.

nated sample (Fig.3.46A, C). The overall shape of the two hNH spectra, considering them as a fingerprint, greatly agrees as far as it becomes apparent for isolated peaks. The slight deviations in chemical shifts might arise from small differences in temperature, a parameter always difficult to control in solid-state NMR. Taking into account that with full protonation, all residues should contribute to the bulk signal with maximum intensity, the average single-peak S/N will be slightly lower than derived from figure 3.46.

Full protonation will certainly avoid signal loss due to incomplete backexchange, but taking into account the reduced sensitivity compared to a 100 % back-exchanged sample in combination with the intrinsically longer measurement times due to the high molecular weight (as described before), the required experimental times would simply outreach a reasonable time limit if the SNR limit is to be maintained. Therefore, it might be concluded that a fully protonated sample, even though being the preferred choice, is strongly limited in terms of practicability and falls behind a deuterated and 100 % exchanged sample.

$R_1\rho$ relaxation studies on TS give first insights into conformational changes

Facilitating all *strong* assignments, including those obtained manually, downstream analysis of TS, such as relaxation experiments to receive information about the protein's dynamics, can be achieved. The complex catalytic cycle in both subunits of TS underlies allosteric modulations through conformational changes in the protein depending on the state in the catalytic cycle[146,148] (see section 1.5.3). Such conformational changes usually underlie motion on the µs timescale which in turn can be particularly well studied using $R_1\rho$ relaxation solid-state NMR experiments[12,64,76,77,79,121,129]. Due to the immense resonance overlap in

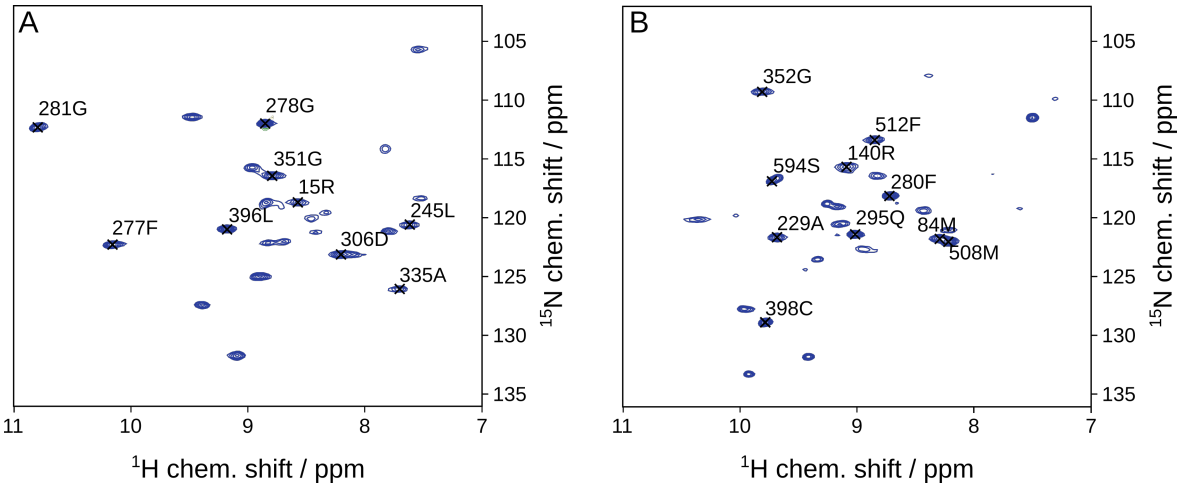


Figure 3.47: Signal dispersion and overlap as it can be found in the pseudo-4D hCONH $R_1\rho$ relaxation series at two different carbon chemical shifts. While the majority of peaks are sufficiently resolved, occasional peak overlap is found, rendering those residues only partially useful for relaxation analysis such as M84 and M508.

a 2D hNH experiment and hence following the requirement for three-dimensional relaxation experiments with a corresponding increase of measurement time, $R_1\rho$ rates were first acquired at a single RF field of 5.5 kHz in a pseudo-4D hCONH series at a spin-lock length of 5 ms, 10 ms, 20 ms, 40 ms, 60 ms and 80 ms (for details see 6.3). Residues for analysis were identified from the obtained shift lists and transferred into a template hCONH experiment from which residues were removed that do not exactly fit the present peaks after referencing. From all 490 residues, only 244 were eligible for further analysis, from which several residues are linked to the same peak (Fig.3.47). This is underlining one of the initial problems pointed out in section 3.1 about the likelihood of peak overlap in large proteins, in particular TS, which constitute a large α-helical character. For those residues, no conclusive $R_1\rho$ rates can be given as it is indistinguishable from which of the residues the observed relaxation behavior occurs.

Residues expected to show significant dynamics in the µs regime remain mostly absent due to missing assignments in regions suspected to undergo major conformational changes such as αL2, αL6 and the COMM domain[146,148,215], which is also indicated by elevated cryogenic Cα B-factors read out from the respective PDB files (Fig.3.48). Many of the 244 analyzed residues show relaxation rates, obtained through an in-house Mathematica or Python script, above an assumed baseline value for rigid residues of 2 s^{-1} as indicated by the dotted line in figure 3.48[55,64,76], pointing towards µs motion being present in TS. Depicting the site-specific $R_1\rho$ rates on the crystal structure of TS (Fig.3.49) reveals indirect indicators of the larger conformational changes assumed to be present. Despite the αL2 being essentially unassigned, the close-by interface of the β-subunit involving residue S429, which is also part of the COMM domain, shows a high relaxation rate of 10.4 s^{-1}. Similar, residues G384 and A376 show elevated rates of 10.0 s^{-1} and 7.7 s^{-1} and might therefore report on µs motion in parts of the COMM domain at room temperature whereas the cryogenic B-factors for all of those residues remain unaltered at a baseline level.

In addition to the residues of the COMM domain, richer relaxation information is obtained for residues in the active site of the β-subunit (Fig.3.50A). As expected, residues H354 and K355 show rather low rates of 2.7 s^{-1} and 3.1 s^{-1} as they hold the PLP cofactor in place via

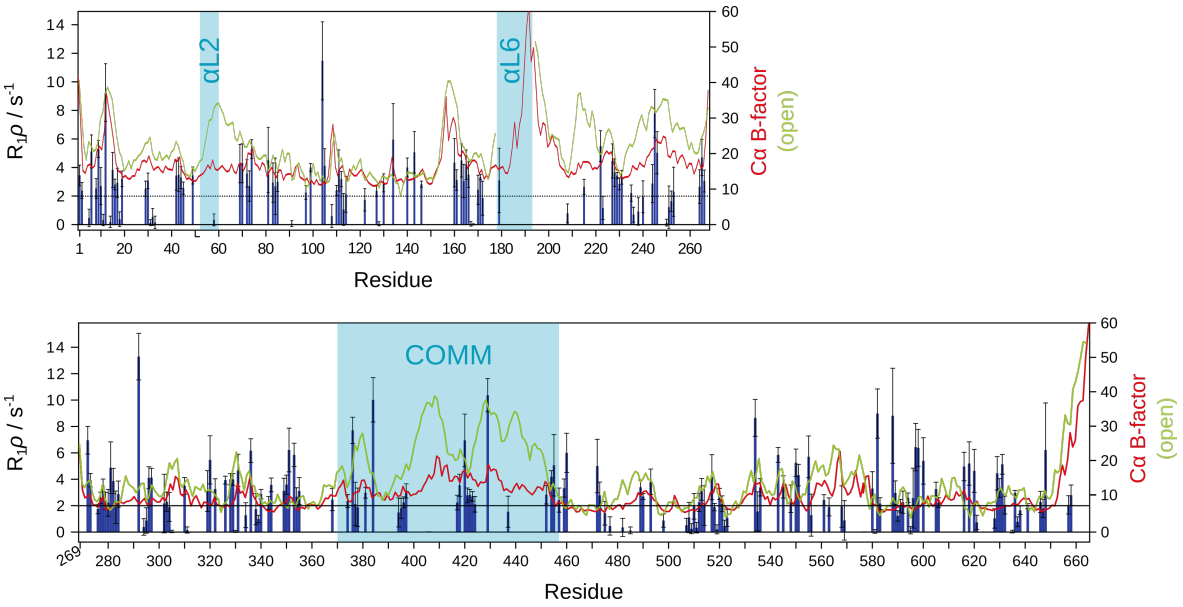


Figure 3.48: Single residue ^{15}N $R_1\rho$ relaxation rates with a spin-lock strength of 5.5 kHz plotted on the sequence for the α-subunit (top) and the β-subunit (bottom), superimposed with Cα B-factors of the crystal structure in the closed conformation (red) of the α-subunit (PDB ID: 4HT3[146]) and the open conformation (green) (PDB ID: 1KFK[222]), indicating large conformational changes occurring during the catalytic cycle and substrate tunneling, especially loops αL2 and αL6 as well as the COMM domain.

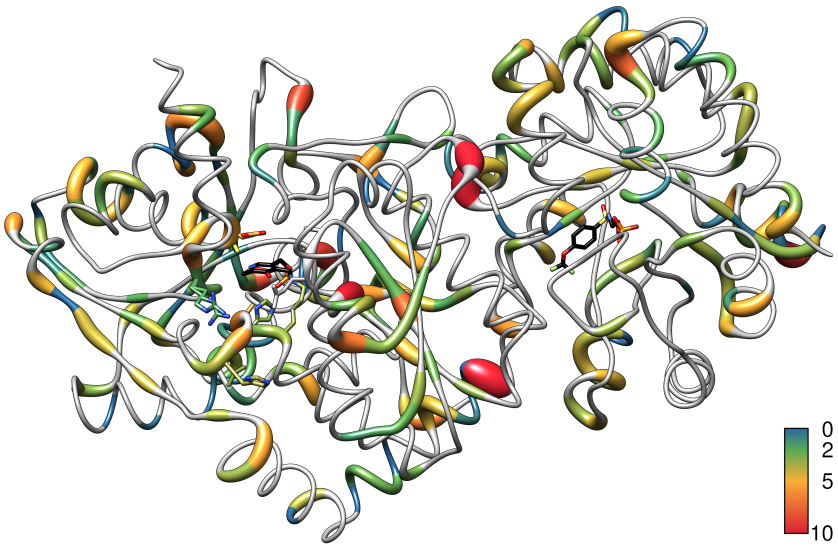


Figure 3.49: ^{15}N $R_1\rho$ relaxation rates at 5.5 kHz spin lock strength plotted on the crystal structure (PDB 4HT3[146]) of TS. The thickness of worms, as well as the color, reflect the rates where thick worms and red colors correspond to high $R_1\rho$ rates, whereas thin worms and blue colors indicate low rates. Only those residues for which a clear assignment from the obtained shifts in a template hCONH experiment was found are shown, while excluded residues are shown in grey. In total, 244 residues were used for the relaxation analysis. Structure depiction and editing in UCSF Chimera[110].

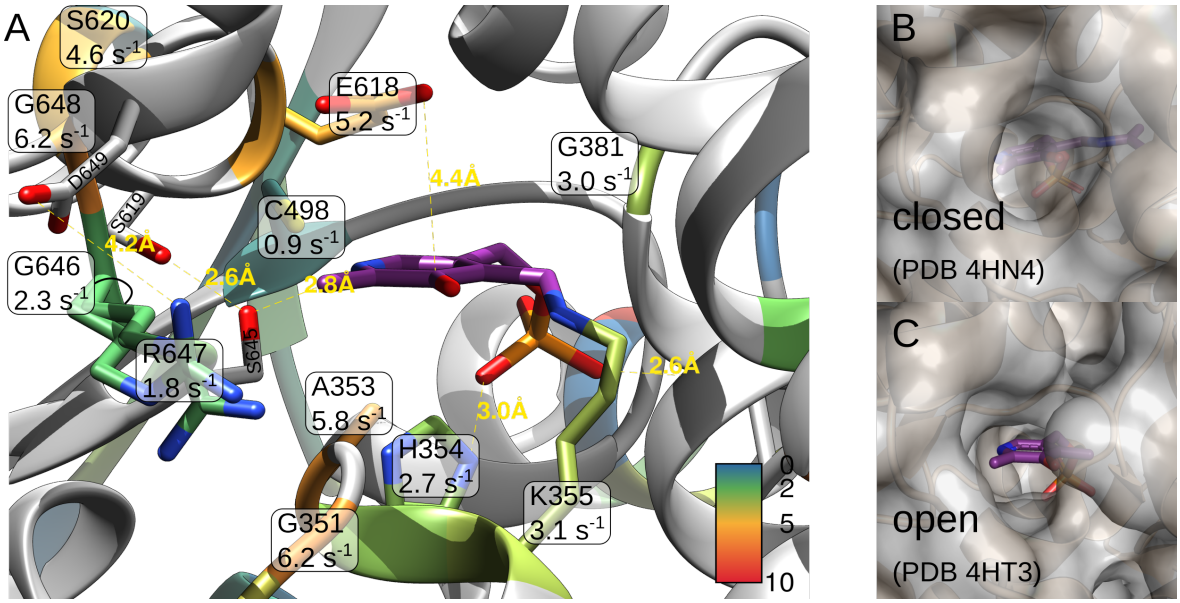


Figure 3.50: A) $R_1\rho$ rates observed for residues close to the PLP cofactor in the active site of the β-subunit. The color coding is as before and indicated in the figure. Assigned residues for which no relaxation data was obtained, but that might function as interactions partners are shown in grey. **B)** Accessibility of the PLP cofactor, shown in purple, in the open conformation of the β-subunit and the closed conformation (**C**)). These conformational rearrangements are hypothesized to be in equilibrium even in the stable internal aldimine state[147] (see section 1.5.3) and can be associated with potential high activation barriers as described in more detail in section 3.3. Especially residues R416 and R647 have to move for the portal to close, which is reflected in the elevated $R_1\rho$ rates. Structure depiction and editing in UCSF Chimera[110] (**(A)** and PyMol[223] (**B,C**).

hydrogen and covalent bonds. In contrast, E618 shows an elevated rate of 5.2 s^{-1} despite the potential anion-π interactions of the carboxyl group and the cofactor. Residues G351 and A353 show elevated rates of 6.2 s^{-1} and 5.8 s^{-1} which, together with R416 and D649 flank the portal over which the active site is accessible for solvated substrates such as L-serine. D649 is assigned but not included in the relaxation analysis. When analyzing the pseudo-4D hCONH spectra, it can, however, be seen that D649 is present although with low intensity and almost completely decays for longer spin lock times at 5.5 kHz spin lock strength, indicating that the residue is highly flexible, albeit not reliably quantifiable. Comparing crystal structures showing the β-subunit either in the open or the closed conformation reveals that residues D649 and R416 undergo significant relocation, which ultimately effects the portal to close, making the cofactor inaccessible at that stage of the catalytic cycle (Fig.3.50B,C). This mechanism is important for the controlled uptake of the substrate and release of the product at exactly the right time of the catalytic cycle and seems to be conserved even in the closed conformation at room temperature, whereas cryogenic B-factors indicate a near-complete static environment.

In a further attempt to gain a deeper understanding of the dynamics present in TS, BMRD-like relaxation dispersion experiments were acquired as before for a single field. The relaxation series was extended by pseudo-4D hCONH experiments at a spin lock RF field of 3.3 kHZ, 4.4 kHZ, and 11 kHZ to sufficiently suppress coherent contributions at lower RF fields[77]. Despite an incomplete analysis of the dispersion profiles at that point, a manual interpretation of the different $R_{1\rho}$ rates at different RF fields can give hints if relaxation dispersion is present and following from that the presence of µs time-scale motion (Fig.3.51).

Looking again at the active site of the β-subunit (Fig.3.51A), it can be seen that residues Q382, R647, D649, and R418 might present such relaxation dispersion, although less pronounced for Q382. Comparing the crystal structures of the open conformation (PDB ID: 4HT3) and the closed conformation (PDB ID: 4HN4), the large conformational changes between the two states become clear, ultimately requiring high plasticity of the residues involved in those regions of the protein.

Here, first studies of single-field and BMRD like $R_{1\rho}$ relaxation experiments indicate the presence of µs time-scale motion being present in TS, and despite unassigned loops, αL2 and αL6, individual residues in the COMM domain are assigned and present elevated relaxation rates pointing towards the conformational changes occurring in TS at room temperature, in spite of the trapped internal aldimine state. Manual analysis of BMRD series in the active site of the β-subunit does not only involve residues of the COMM domain but moreover emphasizes that residues involved in the controlled access and release of substrates and products require higher plasticity, which appears to be conserved in the open conformation.

Chemical shifts report on the direct chemical environment of the β active site

Chemical shifts are direct and sensitive reporters about the direct chemical and electrostatic environment of the observed nucleus and are, therefore, a versatile tool to gain insights into the catalytic activity of TS. Moreover, the assigned ^{1}H chemical shifts can be used as reporters of tautomerism, H-bonds, and proton-exchange processes assumed for the catalytic cycle, in contrast to ^{13}C detected experiments of labeled TS or ^{15}N labeled substrates (see also section 1.5.3)[147,149,150,211]. The identified chemical shifts give hints to an H-bond network around and involving the PLP cofactor in the β-subunit (Fig.3.52) as can be seen, e.g., for G500 interacting with the phosphate group of the cofactor and displaying a strongly downfield shifted ^{1}H amide shift of 10.61 ppm or E618, building a salt bridge to K650 as well as an H-bond to its side-chain and showing a distinct ^{1}H chemical shift of 10.98 ppm. Additional noteworthy residues in the β active site include S645, which side-chain shares H-bonds to the PLP cofactor as the donator, as well as to S619, as acceptor, resulting in an unusual upfield shift of 62.1 ppm for the Cβ nucleus while the Cα resonance is shifted to 62.5 ppm likely due to the third H-bond that is formed between the amide and the hydroxyl group of S619 (Fig.3.52). Additionally, a Cβ chemical shift of 31.1 ppm can be found for H354, which points towards higher pKa values of the imidazole side-chain[224]. More interestingly, all side-chain carbon resonances, as well as the Nζ, Hζ and backbone resonances of K355, could be assigned. The carbon side-chain resonances except Cϵ are directly obtained from the 4D hCCNH in conjunction with the backbone assignments and report already on a slight shift of the Cδ resonance (31.7 ppm vs. 29 ppm BMRB). The Cϵ, Nζ, and Hζ shifts can be obtained from 2D hNH and 2D hCH correlations under long CP contact times of 2 ms and large spectral widths of 313 ppm and 316 ppm for ^{15}N and ^{13}C respectively as well as 96 averaged FIDs (scans) to elevate weak signals. As the entire, abundant proton bath is used as initial magnetization for CP transfers to a heteronucleus, magnetization is not only transferred from the amide proton to the amide nitrogen but also from Hζ to Nζ in case of the hNH experiment as well as to spatially close carbon nuclei in case of the hCH. This effect is enhanced through the long transfer times, which can partially overcome dipolar truncation. The remote peak in

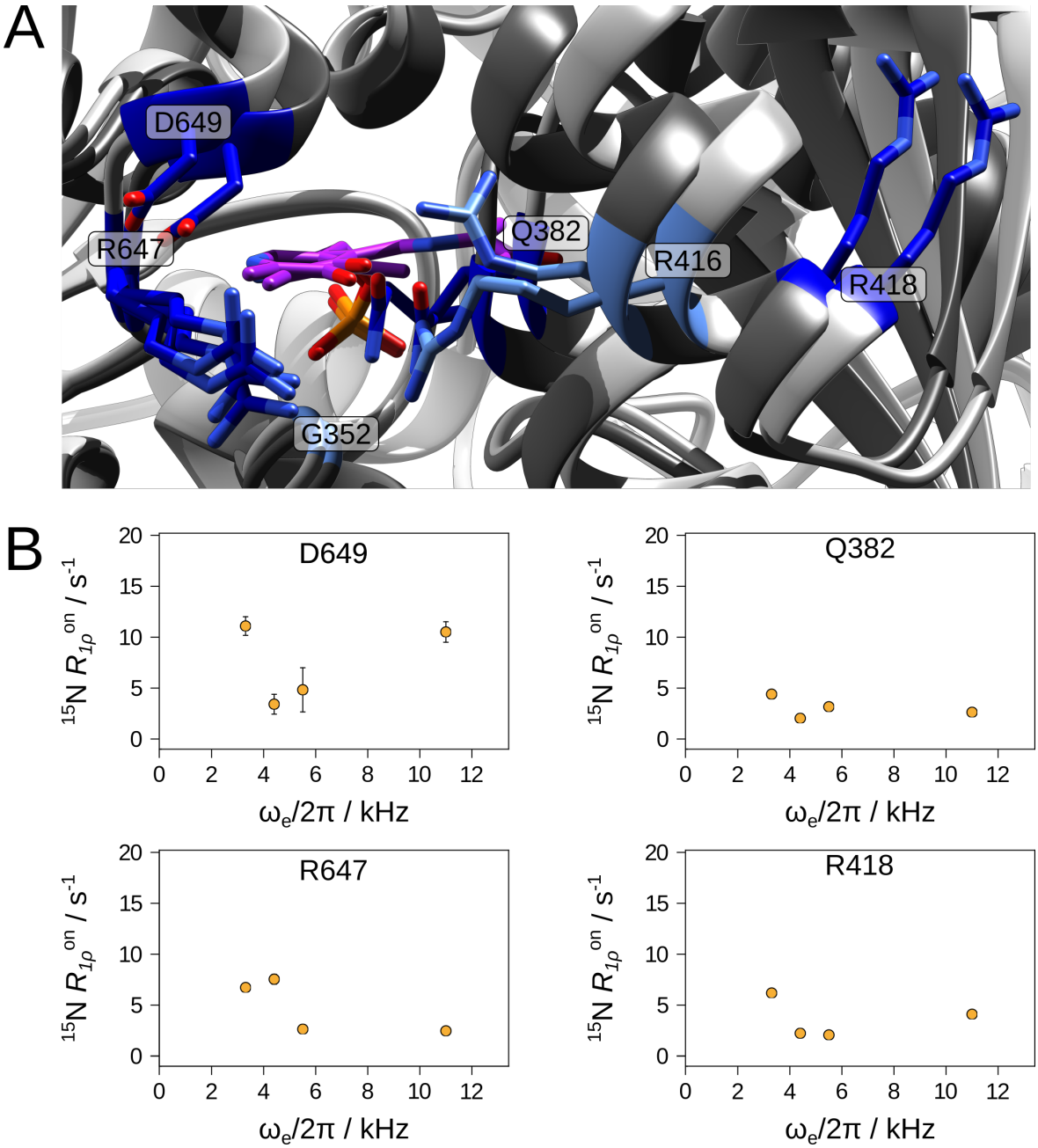


Figure 3.51: A) The active site of the β-subunit with the PLP cofactor highlighted in purple for crystal structures in the open (light grey) and closed conformation (dark grey). Shown in dark blue are residues for which BMRD like relaxation dispersion might be present (**B)** and in light blue residues that are strongly involved in the closing of the portal to the cofactor but for which relaxation analysis was not possible. In contrast to figure 3.50, D649 was included in the BMRD analysis despite being difficult to quantify at different fields in this set of experiments. The relaxation dispersion profile should therefore be considered carefully. The superimposition of the two crystal structures shows the large structural rearrangements involved between the open and closed state. Structure depiction and editing in UCSF Chimera[110].

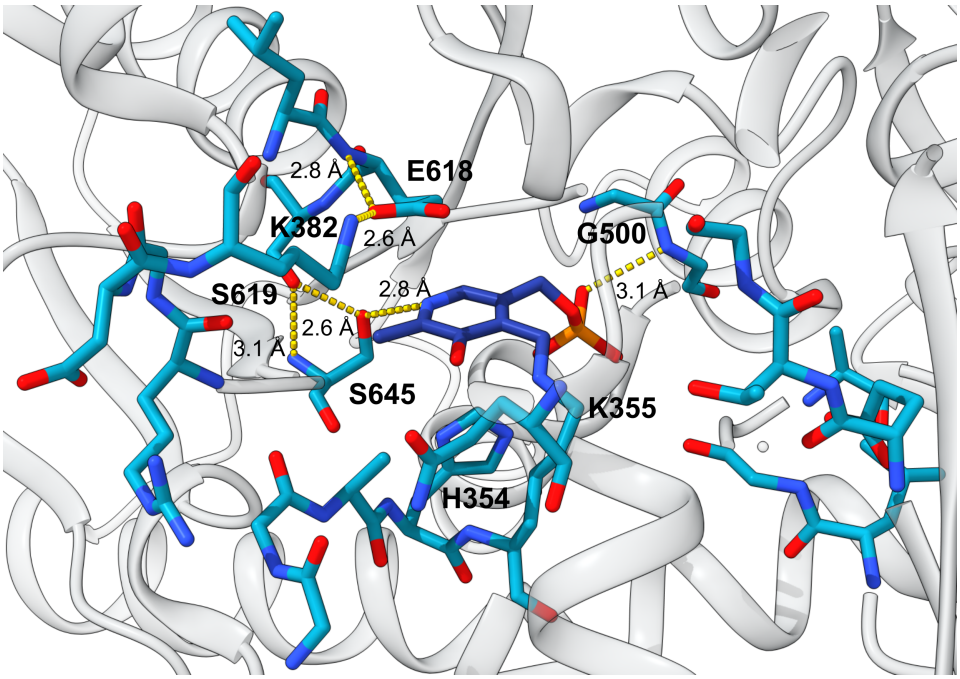


Figure 3.52: The active site of the β-subunit with the PLP cofactor in the center (dark blue) and a selection of surrounding, assigned residues (light blue). A tight H-bond network is formed between S645, S619, and the cofactor, which is reflected in the unusual shifts of S645. Similarly, the H-bond between the phosphate group of PLP and G500 causes strong downfield shifts of the amide proton as well, as the interaction between the side-chains of E618 and K650 causes a proton shift of 10.98 ppm of E618. The covalently attached K355 shows unusual Cδ, Cϵ, and Nζ shifts caused by the covalently bound cofactor. Structure depiction and editing in UCSF Chimera[110].

the H/N correlation features uncommon chemical shifts both for the ^{1}H as well as the ^{15}N dimension and is therefore unlikely part of the protein backbone or a normal side-chain based on comparison with average BMRB chemical shifts. The Nζ atom has been previously studied by selective labeling and ^{15}N-detected, solid-state NMR experiments at -10 °C[147,150] and a chemical shift of 203 ppm was found, suggesting that the here observed shift of 227 ppm likely corresponds to the same atom. A slice at the respective Hζ chemical shift through the carbon dimension in the H/C correlation shows matching chemical shifts for K355 even for the less intense carbon resonances of the side-chain when superimposed with the respective stripe of the 4D hCCNH (Fig.3.53), corroborate the findings from the H/N correlation. In contrast to the hCCNH, the long-range hCH experiment features a strong signal 52 ppm which can be linked to Cϵ, the only resonance absent in the hCCNH strip as it is the most distant residue, but the closest to the Hζ nucleus. The shift follows the tendencies already observed for the Cδ shift and is now significantly downfield shifted by 10 ppm compared to average BMRB shifts through the adjoin Schiff-base. In addition to pure chemical shifts, the line width of the respective isolated peaks can be measured to gain insights into processes associated with line broadening, e.g., chemical exchange as hypothesized for the imine formed by the PLP cofactor and K355[147,149] (see section 1.5.3). The line width at half-height measured for Nζ is on the order of 270 Hz compared to 20 Hz, as is observed for isolated resonances assumed to be not involved in exchange processes.

In cooperation with members of the Mueller group at UC Riverside [4] the newly reported chemical shifts could be used as restraints to obtain a more accurate model of the β-active site

[4]Collaboration partners and co-authors of Klein et al. PNAS DOI: 10.1073/pnas.2114690119

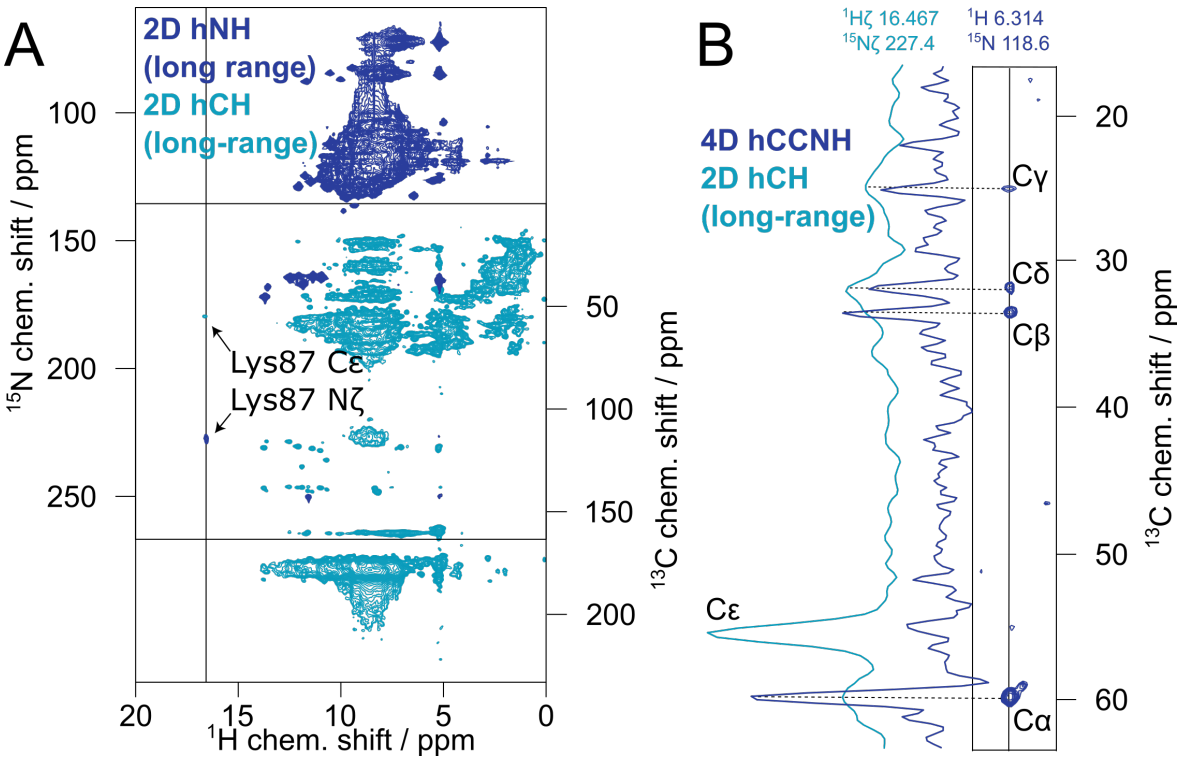


Figure 3.53: Chemical shifts of K355. The Nζ and Cϵ chemical shifts can be identified by long-range, i.e., long CP, hNH, and hCH experiments (**A**)) in combination with the assignments obtained from the hCCNH experiment (**B**)). Due to the covalent bond Hζ, Nζ and Cϵ show unusual chemical shifts compared to BMRB shifts. The 2D hNH experiment was recorded with 96 scans, a spectral width of 313 ppm and a ^{15N} acquisition time of 27 ms in 13.5 h at a ^{1}H Larmor frequency of 700 MHz. The 2D hCH experiment was recorded with 1200 scans, a spectral width of 316 ppm and a ^{15N} acquisition time of 4 ms in 4.1 d recorded in three blocks at a ^{1}H Larmor frequency of 700 MHz.

through NMR-assisted crystallography, i.e., by combing solid-state NMR chemical shifts, X-ray crystal structures and involving first-principle calculations[147,225–229]. As the computations were done by the Mueller lab, no detailed results should be given but rather a short description of the outcomes to demonstrate the utility of the achieved assignments. Facilitating the observed chemical shifts, a two-state fast exchange model can be proposed, which locates the Schiff-base proton in two tautomeric forms partially at the respective nitrogen, named PSB state, and partially at the phenolic oxygen of the PLP cofactor, named PPO state (Fig.3.54). Furthermore, a free energy difference ΔG of only 1.2 kJ/mol can be calculated based on Boltzmann statistics and populations of 62 % and 38 % for the PSB and PPO state at 30 °C. Based on the measured line widths of the Nζ resonance, an free activation energy $\Delta G‡$ of 41 kJ/mol can be assumed, which, however, seems large for a pure proton-exchange event and would rather imply larger conformational changes. The latter ones might indeed occur, as the β-subunit has to open before L-serine can diffuse into the active pocket. The conformational change from close to open is, however, associated with a shift towards a more aqueous environment in the active site, which would ultimately stabilize the zwitterionic PSB state. In contrast, a closed conformation of the β-subunit would favor the neutral PPO state over the PSB state.

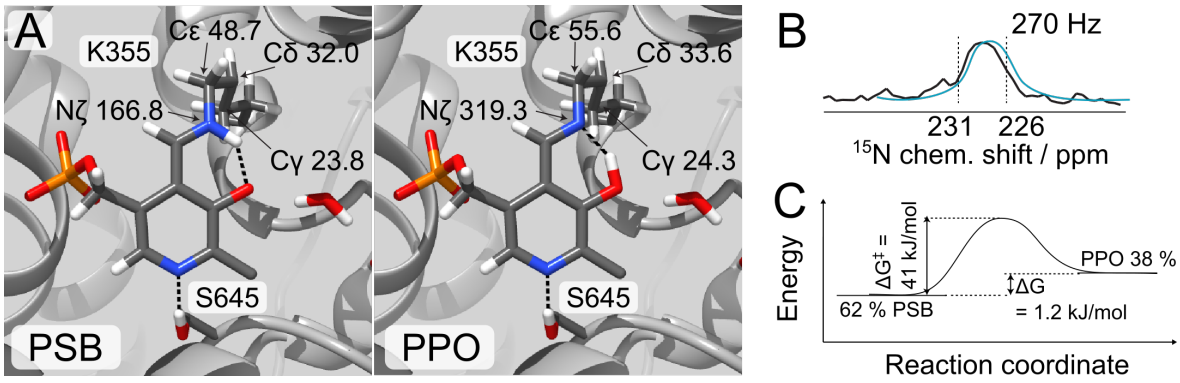


Figure 3.54: Overview of the NMR crystallography and line width analysis, done by the Mueller lab. **A)** shows the computed ab-initio chemical shifts for both the PSB and PPO state. **B)** Through fitting against the experimentally measured line width of the Nζ resonance, conceptually shown in black with a blue fit, the activation energy for the tautomeric proton exchange is obtained (**C**). **D)** Conceptual energy diagram showing the activation barrier as well as the small ΔG computed for the two-state exchange model with a higher populated PSB state at 30°C.

Discussion

Backbone and Cβ resonances of TS could be substantially assigned using a set of different four- and five-dimensional ^{1}H-detected experiments without specialized labeling schemes. Moreover, chemical shift studies on TS have so far been mainly limited to selectively labeled substrates or the PLP cofactor, and mechanistic studies, although abundant, have been restricted to those substrate studies, X-ray structures, MD simulations, and first-principle calculations[146,147,149,150,225,230,231]. Despite being incomplete, the resulting chemical shifts grant unique insights into the proton chemistry involved in the active site of the β-subunit and, in contrast to previous studies, can directly be observed. Chemical shifts have proven to be sensitive probes for the direct chemical environment but are increasingly difficult to get for large molecules in solution as every additional transfer step drastically reduces sensitivity, and even 3D routine experiments such as HNCACB pose hurdles[232,233]. Owned to size-independent magnetization transfer in solid-state NMR, proteins of high molecular weight can be explored but often suffer from severe resonance overlap. The assignments of TS, therefore, demonstrate that the molecular-weight regime beyond 70 kDa is altogether approachable through high-dimensional experiments in a computationally assisted assignment strategy. This opens access to many important proteins that play a key role in pharmaceutical research[113,233–235] and biotechnology[236].

First, relaxation data indicate the presence of relevant µs motions in regions of TS that are crucial for the allosteric modulation of the catalytic cycle in both subunits. These include especially isolated residues in the COMM domain as well as residues involved in opening and closing the portal for substrate entry or product release, respectively. However, at this point, more comprehensive relaxation data remains elusive as many residues have to be filtered out previously for analysis, and a final BMRD fitting is missing. Both will be subjects for ongoing work on TS.

Measurement time is a crucial factor in NMR and can be become a particular problem in solid-state NMR as the number of molecules per volume is fixed, in contrast to solution-state NMR. Nevertheless, high-dimensional experiments are feasible even at a nowadays moderate field strength of 700 MHz, and as pointed out, the measurement time requirements can

drastically be reduced with access to high field magnets of 1 GHz or above. It can be expected that future developments in the field of cryo-MAS probes will further improve the sensitivity and consequently decrease the time requirement, which was recently demonstrated at lower MAS frequencies which are currently available[237]. The assignment of TS was a constant process, and so were the experiments and techniques which were consulted over time. It could be shown that the initial set of 4D experiments could have delivered a similar coverage of assignments when used in a fully automated way, however, without the possibility of manual verification. The 5D HNcoCANH was demonstrated to be essential for the initially pursued manual assignment approach, and beyond that can drastically reduce the set of other experiments in a computationally aided strategy. Although hard to generalize, this will likely be important for future assignments of proteins of similar size and prospectively will shorten both experimental and assignment time.

H/D exchange problems might be surely expected in perdeuterated and back-exchanged preparations but have not been a limiting factor for resonance assignment so far. Unfortunately, for TS, important residues cannot be assigned, which maybe traces back to the incomplete exchange of deuterons. As unfolding/refolding protocols remained unsuccessful so far, the only remnant strategy involves fully protonated sample preparations, which are directly accompanied by compromised sensitivity and extended measurement requirements. However, these can likely be ameliorated with higher fields and cryo probes, as pointed out above.

Despite these limitations, the techniques presented in this work are expected to be of great help for further studies not only on TS but also for unassigned proteins of similar size, which opens great perspectives for ^{1}H-detected solid-state NMR.

Future work in TS should aim at avoiding H/D exchange problems in deuterated preparations, e.g., by incorporating deuterated growth media instead of or in addition to labeled glucose in a water-based M9 medium. While relaxation studies are ongoing, their impact will be bigger if the remaining assignment gaps can be closed through complementing sample preparations, e.g., with only one labeled subunit, which can confirm or correct the existing assignments at the same time. In that context, the usage of NUS in pseudo-5D relaxation series should be explored, which, if successful, will significantly increase the number of residues that can be used for relaxation studies for all proteins in that size regime. Histidine side chains are of particular relevance in the active site of the β-subunit, as, e.g., the protonation state of H354 is important for the catalytic activity throughout the cycle. In addition, also complementary chemical shifts of the heteronuclei in the active site will likely further improve first-principle calculations, and step-wise improve the understanding not only of TS but also other PLP-dependent enzymes. Both HN and HC signals of histidine side-chains can be seen in the same long-range hNH and hCH correlations used to identify Nζ and Cϵ of K355, but insufficient CC mixing schemes so far limit the connection of the side-chains to the protein backbone. Similar to the combined 5D/4D experiment, widening this bottleneck will likely help to investigate side-chain chemistry in a wide range of deuterated protein sample preparations.

Chapter 4

Protein-ligand interactions

4.1 The All-in-one experiment for protein-ligand interactions

Introduction

Understanding the structural properties and dynamics of proteins through a plethora of available techniques contributes essentially to a deeper knowledge about the underlying molecular mechanisms, which can ultimately help to define new drug targets and assist in the development of their inhibitors. Furthermore, the study of existing ligands and their binding proteins can lead to new insights about the mechanisms of action or alterations of the protein[168]. In both cases, structure information of the ligand, as well as the protein in their bound form, is almost always a necessary precondition. Besides *in silico* studies and X-ray crystallography, NMR, mostly in solution-state NMR, is widely used to characterize the interactions between protein and ligand from either the side of the protein, monitored by chemical shift perturbations (CSPs)[238,239] and through-space correlations, i.e. NOESY experiments[239], or from the side of the ligand, using experiments like STD[240–242], T_2-filters[241,243] and others[240,241], but also alterations of chemical shifts, often done for ^{19}F-containing compounds[243–246]. These approaches can be used to obtain binding affinities[238,241,247] as well as to identify interaction sites between the ligand and the protein and are often facilitated for fragment screenings and lead optimization[238,240,245,247,248]. Whereas most employed NOESY experiments can provide rich information about the structure of the protein[239,249] or between the ligand and the protein[239,250], structure information about the ligand itself often remains absent, and several experiments are required to get an encompassing picture. In protein NMR, edited NOESY experiments[203,204,251,252] are the experiments of choice to yield intra-protein through-space correlations, as only those ^{1}H-^{1}H contacts appear which are attached to either a ^{15}N or ^{13}C nucleus[239]. On the other hand, contacts between an unlabeled ligand and the labeled protein can be obtained through filtered NOESY experiments in which only those ^{1}H-^{1}H contacts appear for which no ^{15}N or ^{13}C nucleus is bound[239,253,254]. These experiments are usually performed in three or four dimensions and double-edited or filtered combinations exist[239,253,254] also in solid-state NMR[40,74]. Despite the insights resulting from these experiments, intra-ligand contacts can be at most indirectly observed as they are filtered out or will likely disappear in the overwhelming number of intra-protein signals. Only double-filtered experiments can provide those pure intra-ligand contacts but, in turn, will miss all contacts to or from the protein. In conclusion, several experiments will be required to obtain all types of correlations, i.e., protein-protein, protein-ligand, and ligand-ligand contacts. Here, the All-in-one experiment for solid-state NMR is introduced as adapted from concepts developed for solution-state NMR[239]. The 3D experiment contains a time-shared HSQC-NOESY[204] experiment that includes a ^{1}H-^{1}H-NOESY to show intra-ligand NOEs through an artificial chemical shift. The general pulse sequence and possible variants are introduced, and critical problems of sufficient water suppression are pointed out.

Product operators of the protein and ligand pathway

The concept of the All-in-one or All-inclusive experiment was previously introduced for solution-state NMR[239], combining a 3D filtered [^{15}N,^{13}C]-time-shared HSQC-NOESY experiment with a 2D-like ^{1}H-^{1}H NOESY of a ligand molecule (Fig.4.1). However, a detailed description of the pulse sequence itself can only indirectly be deduced from the pulse pro-

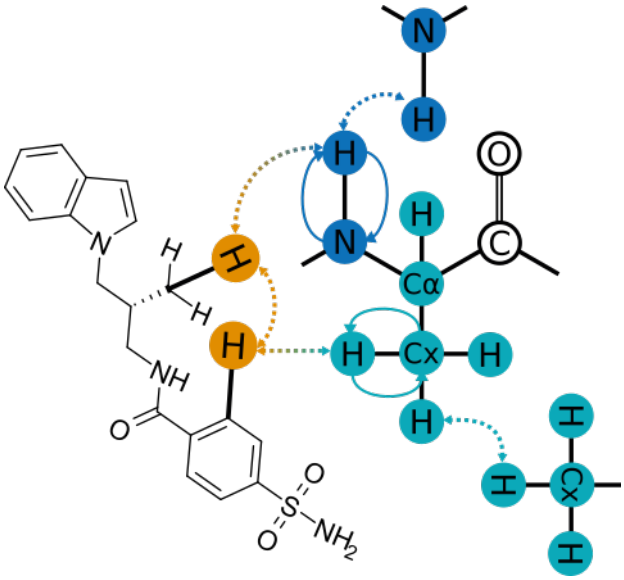


Figure 4.1: Schematic magnetization pathways as observed in the All-in-one experiment. Solid lines indicate through-bond transfers of the ^{15}N-HSQC (dark blue) and ^{13}C-HSQC (teal), dashed lines indicate ^{1}H-^{1}H through-space contacts, facilitated either by NOE in solution-state NMR or RFDR in solid-state NMR, observed starting from amide groups (dark blue), unsaturated carbons (teal) in the HSQC-NOESY pathway as well as intra-ligand contacts (orange) as well as combinations of those. The ligand compound is the compound SBR ((R)-N-(3-Indol-1-yl-2-methyl-propyl)-4-sulfamoyl-benzamide)[170], a potent non-covalent inhibitor of hCAII.

gram available for download[255]. As several modifications, especially regarding water suppression, phase cycling, and ^{1}H-^{1}H mixing had to be implemented to adapt the experiment for solid-state NMR, a detailed description of the magnetization pathways in terms of product operators should be given here. The first pathways to consider are the transfers that provide ^{15}N,^{13}C-edited ^{1}H-^{1}H contacts. As this experiment is using the concept of time-shared NMR, the timing of pulses for the ^{15}N and ^{13}C INEPTs have to be adjusted. The first π-pulse simply flips the proton magnetization in the x,y-plane, i.e., onto the -y-axis if x-phase is considered.

$$\hat{H}_z \xrightarrow{(\pi/2)\hat{H}_x} -\hat{H}_y$$

With respect to ^{13}C to overall delay until the final $\pi/2$-pulse in the INEPT to transfer the magnetization is

$$\Delta_1 - \Delta_2 + \Delta_3 = \frac{1}{4J_{HC}} - (\frac{1}{4J_{HN}} - \frac{1}{4J_{HC}}) + \frac{1}{4J_{HN}} = 2\frac{1}{4J_{HC}} = \frac{1}{2J_{HC}}$$

and similar with respect to ^{15}N

$$\Delta_1 + \Delta_2 + \Delta_3 = \frac{1}{4J_{HC}} + (\frac{1}{4J_{HN}} - \frac{1}{4J_{HC}}) + \frac{1}{4J_{HN}} = 2\frac{1}{4J_{HN}} = \frac{1}{2J_{HN}}$$

Furthermore, it can be seen that the proton chemical shift evolution is refocused with respect to

$$\begin{aligned} \Delta_1 &= -(\Delta_2 - \Delta_3) \\ \frac{1}{4J_{HC}} &= -(\frac{1}{4J_{HN}} - \frac{1}{4J_{HC}}) + \frac{1}{4J_{HN}} = \frac{1}{4J_{HC}} \end{aligned}$$

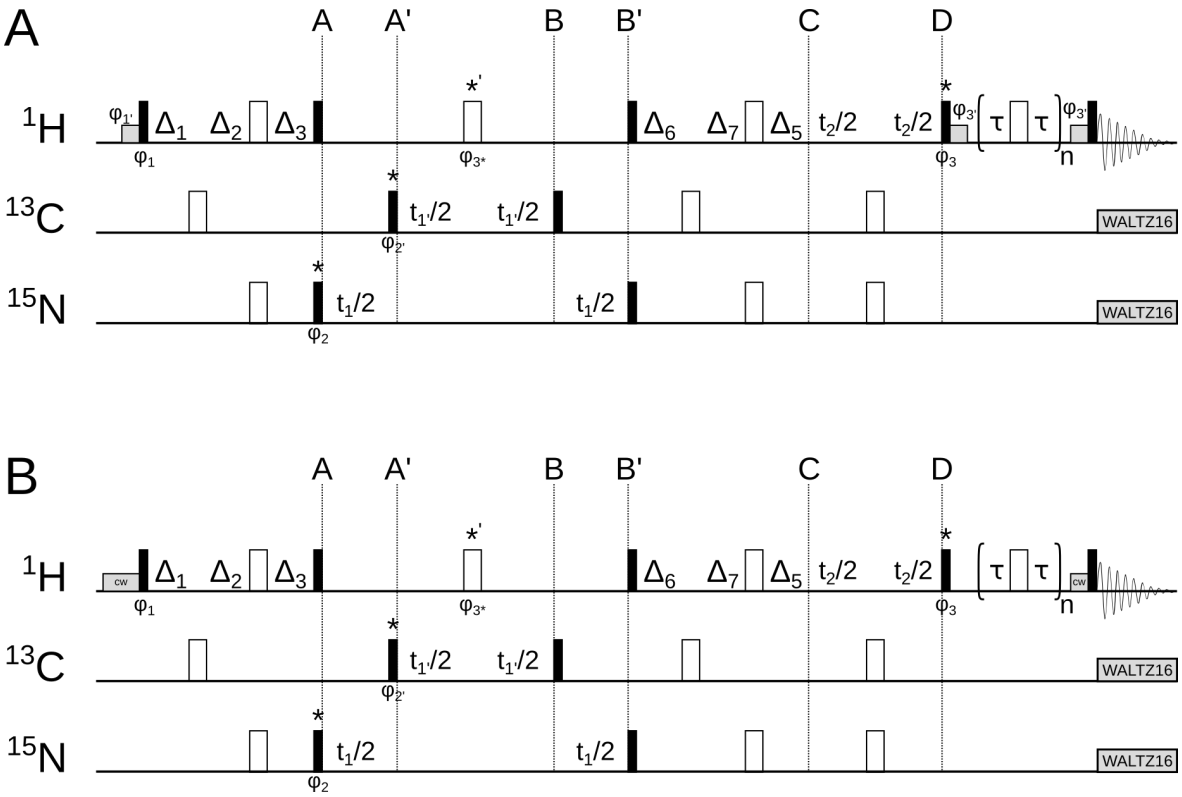


Figure 4.2: Pulse sequences for the solid-state NMR version of the All-in-one sequence[239].**A)** Hard $\pi/2$ and π pulses are denoted by filled and empty rectangles. Small, light gray squares indicate soft power rectangular pulses. The delays are set $\Delta_1 = \frac{1}{4J_{HC}}$, $\Delta_2 = k\frac{1}{4J_{HN}} - \frac{1}{4J_{HC}}$, $\Delta_3 = k\frac{1}{4J_{HN}}$, $\Delta_5 = \frac{1}{4J_{HN}}$, $\Delta_6 = \frac{1}{4J_{HC}}$ and $\Delta_7 = \frac{1}{4J_{HN}} - \frac{1}{4J_{HC}+t_2}$ with k to be optimized for optimal intensity. The pulse program including finer adjustments of the delay times can be found in the appendix (section 7.1). Phase cycle: $\phi_{1,1'}$ = x,-x; $\phi_{2,2'}$ = x,x,-x,-x; ϕ_3 = y,y,y,y,-y,-y,-y,-y; $\phi_{3'}$ = -y,-y,-y,-y,y,y,y,y; ϕ_{3*} = 45°, 45°, 135°, 135°;ϕ_{rec} = x,-x,-x,x. The RFDR π pulse was cycled with the XY-16 super cycle. Pulses marked with an asterisk are phase-shifted by 90 or 45° in case of the ^{1}H π pulse to obtain phase-sensitive States-TPPI[173] spectra in every indirect dimension including the 2D intra-ligand spectrum. **B)** As **A)**, but selective pulses on water were replaced by selective continuous wave periods during the recycle delay and a shorter 10 ms block after RFDR mixing.

and also with respect to^{15}N.

$$\Delta_1 + \Delta_2 = \Delta_3$$
$$\frac{1}{4J_{HC}} + (\frac{1}{4J_{HN}} - \frac{1}{4J_{HC}}) = \frac{1}{4J_{HN}}$$

Therefore, the $-\hat{H}_y$ term will have evolved to full anti-phase both with respect to ^{15}N and ^{13}C, which is then transferred to either of the hetero-nuclei at point **A** and **A'**.

$$-\hat{H}_y \xrightarrow{2\pi J_{HN}\tau_1 \hat{H}_z \hat{N}_z} -2\hat{H}_x\hat{N}_z \xrightarrow{\pi/2\hat{H}_y} 2\hat{H}_z\hat{N}_z \xrightarrow{\pi/2\hat{N}_x} -2\hat{H}_z\hat{N}_y$$
$$-\hat{H}_y \xrightarrow{2\pi J_{HC}\tau_1 \hat{H}_z \hat{C}_z} -2\hat{H}_x\hat{C}_z \xrightarrow{\pi/2\hat{H}_y} 2\hat{H}_z\hat{C}_z \xrightarrow{\pi/2\hat{C}_x} -2\hat{H}_z\hat{C}_y$$

After the INEPTs, chemical shift evolution of both ^{15}N and ^{13}C can occur, whereas the π pulse on ^{1}H, now assumed to be x-phase, simply refocuses the J-coupling, so that at point **B** or **B'**

$$-2\hat{H}_z\hat{N}_y \xrightarrow{\pi_x} 2\hat{H}_z\hat{N}_y$$
$$-2\hat{H}_z\hat{C}_y \xrightarrow{\pi_x} 2\hat{H}_z\hat{C}_y$$

Next, the magnetization is transferred back to ^{1}H by a reversed INEPT, and consequently, the same transformations apply, and as

$$\begin{aligned}
&\Delta_6 = \Delta_1 \\
&\Delta_7 = \Delta_2 \\
&\Delta_5 = \Delta_3 \\
&\Rightarrow \Delta_6 + \Delta_7 + \Delta_5 = \frac{1}{2J_{HN}} \\
&\Rightarrow \Delta_6 - \Delta_7 + \Delta_5 = \frac{1}{2J_{HC}} \\
&\Rightarrow \Delta_6 + \Delta_7 = \Delta_5 \\
&\rightarrow \Delta_6 += -(\Delta_7 - \Delta_5)
\end{aligned}$$

can be assumed, pure in-phase $\hat{H}_x$ magnetization is obtained at point **C**. The delays Δ are denoted differently, as in the first INEPT, a relaxation-optimized adjustment for $\frac{1}{4J_{HN}}$ can be performed to optimize the signal. However, as those adjustments do not alter the general concept of the experiment, the approximations above can be considered valid. The $\hat{H}_x$ magnetization is allowed to undergo chemical shift evolution, while the J-coupling to ^{15}N and ^{13}C are refocused by two π pulses, so that at point **D**

$$\hat{H}_x \xrightarrow{\pi\hat{N}_x} \hat{H}_x$$
$$\hat{H}_x \xrightarrow{\pi\hat{C}_x} \hat{H}_x$$

To obtain through-space ^{1}H-^{1}H contacts, a NOESY block would now be introduced, while in the solid-state version, an RFDR[24,25] mixing block is introduced, which reintroduces ^{1}H-^{1}H dipolar interactions, which ultimately lead to a through-space magnetization transfer with similar results as NOEs, however, mediated through different mechanisms. RFDR mixing, as NOESY mixing, is performed on z-magnetization, which is easily created through

$$\hat{H}_x \xrightarrow{\pi/2(-\hat{H}_y)} \hat{H}_z$$

For detection, the proton magnetization is ultimately transferred back to x,y-plane after the mixing period

$$\hat{H}_z \xrightarrow{\pi/2\hat{H}_y} \hat{H}_x$$

The benefit of the All-in-one experiment is that, simultaneously to the 3D time-shared edited HSQC-NOESY experiment, a 2D-like intra-ligand NOESY spectrum is obtained as in contrast to conventional experiments, the ^{1}H magnetization that does not undergo the HSQC path is preserved. In a conventional HSQC-NOESY experiment, the ^{1}H magnetization is kept and refocused along the y-axis, and the term $-\hat{H}_y$ is preserved. In combination with gradient, water suppression schemes like WATERGATE[256] and phase cycling, all non-edited signals essentially vanish. Here, however, $-\hat{H}_y$ is eventually turned into $-\hat{H}_x$ through an alteration in the phase of the refocusing π pulse during ^{15}N and ^{13}C chemical shift evolution from 0°, i.e., x-phase, to 45°, causing a 90° shift of magnetization (Fig.4.3C) so that at point **B'**, the pathways of protons that underwent the HSQC path and the non-edited ones are eventually recombined and undergo RFDR mixing. The pathway for non-edited, e.g., ligand and water protons, until point **B'** is therefore

$$\hat{H}_z \xrightarrow{\pi/2\hat{H}_x} -\hat{H}_y \xrightarrow{2\pi J_{HX}\tau_1\hat{H}_z\hat{S}_z,\pi\hat{H}_x} -\hat{H}_y \xrightarrow{\pi\hat{H}_{45°}} -\hat{H}_x$$

If the phase cycle of the respective π pulse is matched to the one of the HSQC pathway (Table 4.1 and Fig. 4.2), all non-edited signals are maintained throughout the entire experiment, which eventually leads to difficulties in the water suppression (see section 4.1). As the protons

Table 4.1: ^{1}H magnetization in the All-in-one and conventional filtered HSQC-NOESY experiments at points **A**,**B** and **D** created through the phase of the decoupling π pulse during ^{15}N and ^{13}C chemical shift evolution. For reasons of simplification only the phase of the terms are denoted.

Phase	A	B	D	rec
45°	±y	x,-x,-x,x	x,-x,-x,x	x,-x,-x,x
0°(x)		±y	±y	

and the resulting NOESY-like spectrum of this path omit the chemical shift evolution period on the heteronuclei, they will appear as zero-frequency artifacts at the respective carrier frequency of the dimension set up for the time-shared experiment, e.g., 120 ppm in case of a ^{15}N dimension (FIG) and therefore will most likely overlap with ^{1}H-^{1}H contacts resulting from residues of a similar chemical shift. To circumvent this pitfall, an artificial chemical shift in the pseudo-third, i.e., hetero-nuclear, dimension of the intra-ligand spectrum can be introduced via an additional phase shift of the before-mentioned π pulse similar to the TPPI

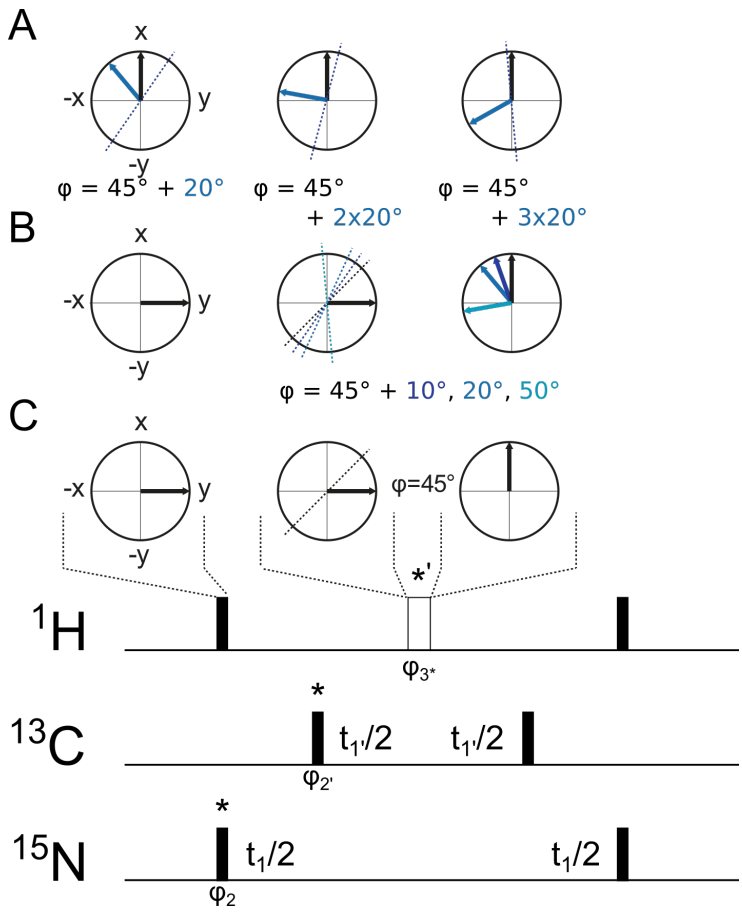


Figure 4.3: Depiction of the roles of the 45° pulse phase **(C)** as well as the added phase value ϕ causing an artificial chemical shift similar to TPPI[172,257] **(A,B)**. The 45° of the π pulse causes a phase shift of the non-edited $\hat{H}_x$ terms by $\pi/2$. Adding an additional phase value ϕ causes an additional shift, which simulates a chemical shift along the hetero-nuclei dimension when incremented simultaneously. The extent of the TPPI-like shift depends on the values of ϕ as is shown for $\phi = 45°+10°, 20°, 50°$. The black arrow follows the shift of 45°, while blue shades indicate the respective added value **(B)**. If for every incrementation of the heteronuclear dimension a fixed value (here 20° in **C**)) is constantly added, the magnetization vector is constantly moved forward or backward, simulating a Larmor frequency different from the carrier.

method[172,257]. If for every time increment of the heteronuclei, a fixed phase value is added to the 45° phase of the π pulse, the magnetization appears slightly faster or slower than it would otherwise (Fig.4.3A,B). While the protons of the edited pathway will accumulate a factor $\cos(\Omega_X t_1)$, those of the non-edited path accumulate a factor $\cos(\phi(t_1/\mu))$ with ϕ being the fixed phase-value added and t_1/μ the factor how often the phase was added with the increment-size μ (Fig.4.3A). A value of $\phi = \pm\pi/2$ will lead to a shift to the edges of the spectrum, as used in pure TPPI spectra, which is, however, not desirable for the All-in-one experiment, and ϕ should be chosen in such a way that the 2D-like non-edited spectrum is shifted in a region where no or only little overlap with the hetero-nuclei will occur (Fig.4.4). The value for ϕ can be calculated as

$$\phi = \Omega\pi\frac{1}{SW}$$

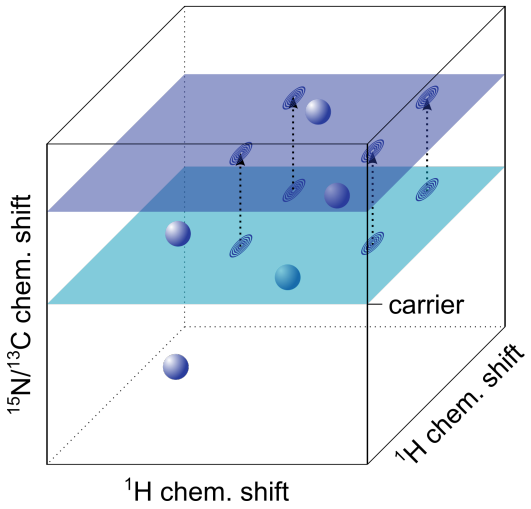


Figure 4.4: Schematic representation of a spectrum obtained by the All-in-one experiment. While signals originating from the protein (blue spheres) are dispersed in the 3D spectrum, signals originating from the ligand would usually appear as zero-frequency artifacts at the carrier frequency of the heteronuclear dimension (teal). By means of a TPPI-like artificial chemical shift, this plane can be moved up or down to such a position where overlap with signals originating from the protein is unlikely. For reasons of simplicity, the diagonal and symmetric cross-peaks of the HSQC-NOESY signals are not shown.

Water suppression

As in all HSQC-NOESY like experiments, the water signal in the NOESY or RFDR dimension cannot be suppressed through phase cycling in contrast to NOESY-HSQC-type experiments[203,204] and whereas in solution-state NMR powerful and selective water suppression methods such as WATERGATE[256] or excitation sculpting[258] can be employed for water suppression, water suppression in solid-state NMR is usually performed by a modified version of the MISSISSIPPI[191] scheme without gradients. MISSISSIPPI essentially relies on dephasing all unwanted ^{1}H magnetization while storing the selected magnetization along the z-axis on the hetero-nucleus. In the All-in-one, however, this dephased magnetization would include the ligand signal, which has to be maintained, and MISSISSIPPI cannot be used. The most promising approach might therefore be to selectively keep the water signal aligned along the z-axis as it otherwise behaves exactly like the ligand and will lead to a strong water signal and distortions in the shifted intra-ligand spectrum as well as the 3D edited HSQC-NOESY. Originally, a series of water flip-back pulses[259] were introduced in the pulse sequence (Fig.4.2A) and implemented as soft-power rectangular pulses rather than the frequently used sinc pulses . Soft power rectangular pulses are, however, slightly less selective, much shorter compared to shaped pulses like sinc and were empirically found to be the better choice for this specific experiment during optimization. Under the influence of these additional pulses, the magnetization of the water can be described as

$$\hat{H}_z \xrightarrow{(\pi/2)_{sp}\hat{H}_x} -\hat{H}_y \xrightarrow{(\pi/2)\hat{H}_x} -\hat{H}_z \xrightarrow{(\pi)\hat{H}_x} \hat{H}_z \xrightarrow{(\pi/2)\hat{H}_y} \hat{H}_x \xrightarrow{(\pi/2)\hat{H}_{45°}} \hat{H}_y \xrightarrow{(\pi/2)\hat{H}_x} \hat{H}_z$$
$$\xrightarrow{(\pi/2_sp)\hat{H}_{-y}} -\hat{H}_x \xrightarrow{(\pi/2)_{sp}\hat{H}_y} \hat{H}_z \xrightarrow{RFDR,(\pi/2)_{sp}\hat{H}_y} \hat{H}_x \xrightarrow{(\pi/2)\hat{H}_{-y}} \hat{H}_z$$

assuming that the RFDR block is an odd multiple and perfectly inverts the water magnetization. Despite expectations, the water suppression using flip-back pulses is insufficient, and a

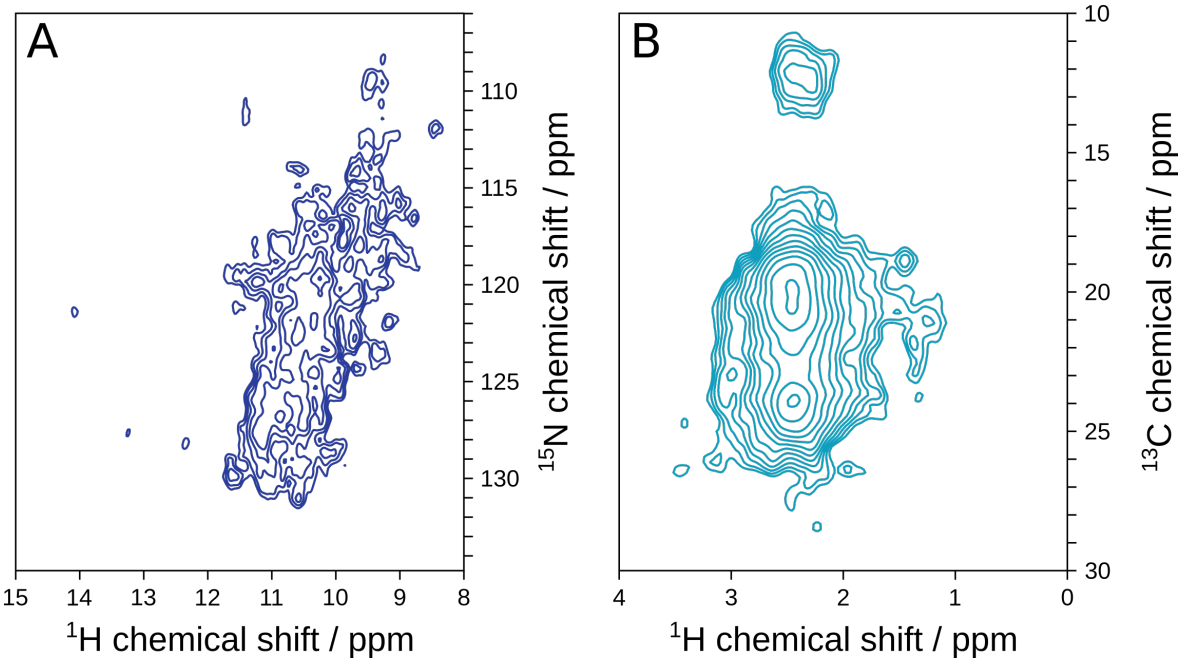


Figure 4.5: **A)** hNH and **B)** hCH spectra obtained from the hCAII sample with ILV labeling scheme. Both correlations show pronounced sample heterogeneity and indicate degradation of the crystals, rendering the sample unsuitable for the All-in-one experiment.

broad water line appears along the HSQC (and RFDR) plane of the All-in-one experiment. In a second alteration, a more brute force approach to suppress the water signal was attempted through selective saturation of the water resonance during the recycle delay (presaturation) in combination with a shorter continuous wave irradiation after RFDR mixing (Fig.4.2B). Although saturation of the water signal can lead to severe attenuation of the amide signal and other exchangeable sites, the suppression of the water signal is superior over selective pulses and is here proposed as the preferable option, as good solvent suppression is crucial for the outcome of the All-in-one experiment. Not only is the water signal almost gone in the first FID, but also the water line along the indirect dimensions is significantly decreased in intensity and width.

4.1.1 Proof of concept on the SH3-domain

The All-in-one experiment is designed to observe intra-ligand signals and signals originating from the ligand in combination with a time-shared HSQC and sample requirements, therefore, include the presence of a ligand as well as either a fully protonated sample or partial labeling, preferable of side-chains, like ILV-labeling in a deuterated and back-exchanged sample. The sample that was supposed to be used for both implementing the adapted solid-state NMR version and searching for new insights was a ^{1}D,^{13}CHD$_2$,^{15N} ILV labeled sample of hCAII as described in sections 6.1.3. As water suppression was initially the major drawback of the solid-state NMR version, the crystals were planned to grow in normal buffer, and after ligand soaking, a solvent exchange to crystallization buffer in D_2O was planned to quench the water signal as much as possible, however on the cost of the HN signals. As the crystals degraded after soaking and solvent exchange (Fig.„,4.5) a more general proof-of-concept of the adapted solid-state NMR version was done on an u-[^{2}D,^{13}C,^{15}N] 100 % back-exchanged sample of the SH3-domain. Figure 4.6 shows different planes obtained from the 3D All-in-one experiment.

Both the RFDR and HSQC planes show the expected correlations. In the case of the ^{1}H indirect dimension, no further treatment is required, whereas the heteronuclear dimension has to be changed from ^{13}C to ^{15}N through modification of relevant parameters in the *acqus* files. This step is necessary, as, despite the absence of H/C signals in the used sample, the set up of the pulse sequence remained unaltered as time-shared experiment with ^{13}C set as the hetereonucleus. The obtained plane resembles similarities to a conventional hNH spectrum of the SH3-domain (Fig.4.6B, C). Several signals, however, seem to disappear or appear shifted, while for others, more than one ^{1}H chemical shift is apparent, which accounts for RFDR contacts as the direct dimension is the RFDR or NOESY dimension in HSQC-NOESY type experiments. On top of the HSQC plane, a broad line at 104.5 ppm is found, which displays the shifted signals which would potentially be created by a ligand. As such a ligand is absent in the used sample, the signals are likely caused by water and other buffer components such as EDTA. The respective 2D RFDR plane found at 104.5 ppm, corresponding to a 70° phase shift with the ^{15}N carrier at 120 ppm, shows intense diagonal and cross-peaks in the aliphatic region, which might occur from EDTA or other buffer components. Despite missing a more suitable sample to demonstrate all aspects of the All-in-one experiment, it can be concluded that the adapted solid-state NMR version generally shows the expected correlations and features. Sufficient water suppression initially put the strongest limitation as gradients are not available in solid-state probes but could be attenuated to a level in which the water signal is not severely overlapping or distorting other signals.

Discussion

The All-in-one experiment was successfully adapted for solid-state NMR, and a robust method for satisfying water suppression was included. Although all relevant correlations appear in the spectrum, many signals appear to be missing, and the overall sensitivity remains comparably low. So far, the evaluation of the experiment hinges on the absence of a suitable sample that carries a matching labeling scheme and a ligand. The improved water suppression might make a solvent exchange of ILV labeled hCAII redundant, and a repeated preparation will most certainly be prepared for future studies. Furthermore, the use of CP transfers over INEPTs should also be considered, which will render the All-in-one experiment also suitable for slower MAS frequencies and improved sensitivity due to shorter transfers.

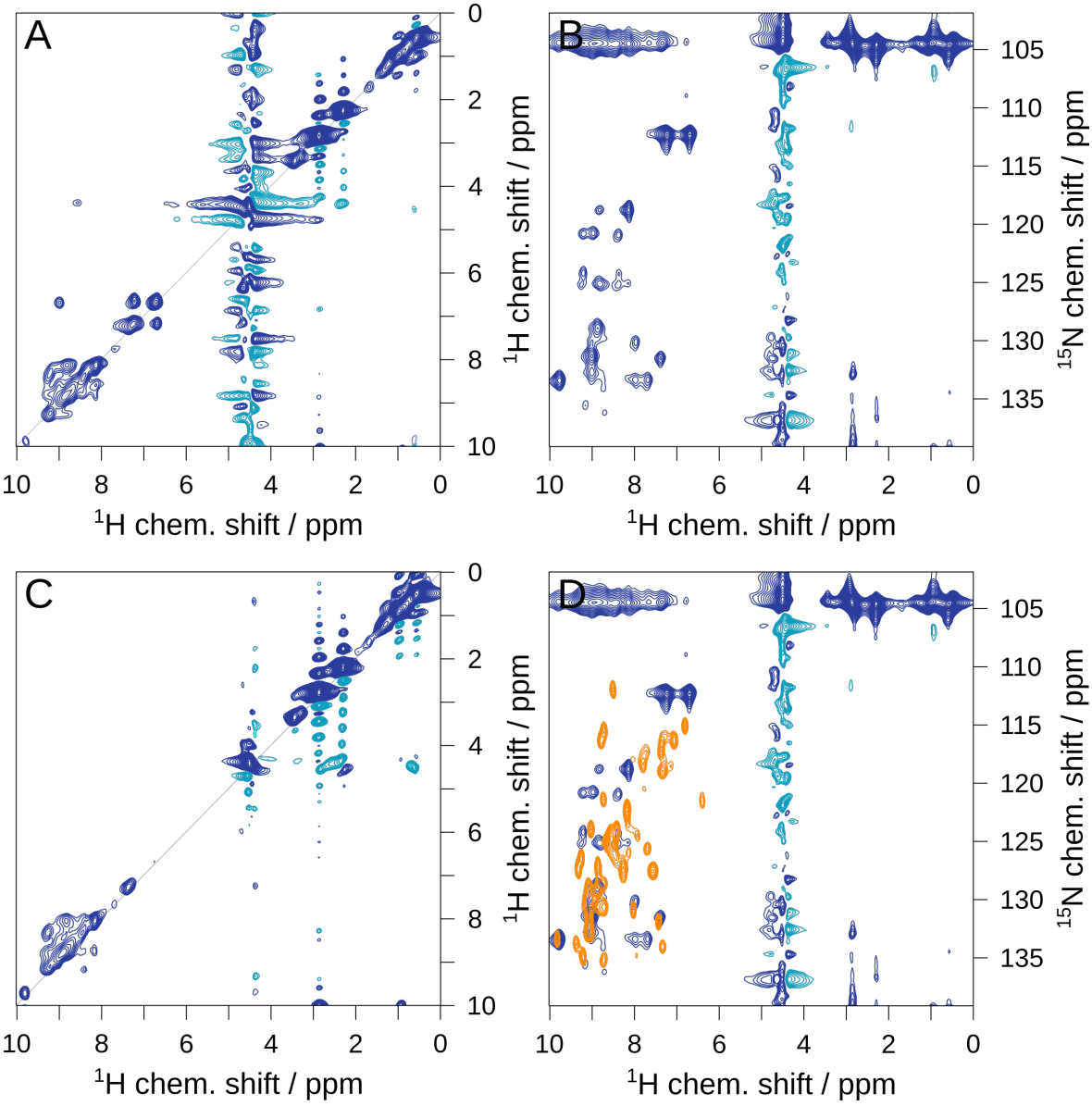


Figure 4.6: Spectra from the All-in-one sequence acquired on a u-[^{2}D,^{13}C,^{15}N] and 100 % back exchanged sample of the SH3 domain at 55 kHz MAS and a ^{1}H Larmor frequency of 700 MHz. Shown are the ^{1}H-^{1}H RFDR plane (**A**)), the ^{1}H/^{15}N HSQC plane (**B**)) processed from the full 3D spectrum. **C**) shows the shifted ^{1}H-^{1}H RFDR plane with signals originating from the ligand at ^{15}N chemical shift of 104.5 ppm. **D**) shows overlays of the HSQC plane from **B**) and an hNH (CP) spectrum recorded on the same sample and same conditions. The All-in-one was processed with 100 Hz line broadening in the direct dimension and 50 Hz in both indirect dimensions, while the hNH spectrum was processed with a squared sine bell function and a sine bell shift of 2 in both dimensions.

Chapter 5

Discussion and outlook

This work presented several new approaches in ^{1}H-detected solid-state NMR that aimed at the development and application of higher-dimensional pulse sequences, but also one new method to study protein-ligand interactions was introduced. Higher-dimensional solid-state NMR experiments have proven to be a valuable tool for proteins of all sizes. They can be either used for the fast and easy assignment of smaller proteins, including near-complete resonance assignment of fully protonated proteins, or as an indispensable tool for the manual assignment of large proteins such as TS. In addition, it was shown that incorporating the FLYA algorithm for automated resonance assignment significantly supports the assignment process of large proteins and benefits from the high chemical shift redundancy introduced by higher-dimensional experiments. However, the actual proof of the assignment remains a manual task, which proposes a hybrid approach as probably the most advisable way.

Measurement time remains a limiting factor for higher-dimensional NMR experiments despite the heavy use of NUS and shortened recycle delays through PRE. While APSY was demonstrated to be the most powerful approach of those analyzed here in terms of experimental time, it will most likely be not suitable for sample preparations with heavy resonance overlap. Although NUS enables the use of experiments with more than four dimensions through rapid data acquisition, NUS parameters, especially the number of NUS points, have to be chosen carefully, and for large proteins, a high number of points is generally advisable to obtain high-quality spectra. In combination with the limited amount of sample present in the rotor, lower limits for the total experimental time cannot be overcome without compromising the quality of the spectra. However, future improvements in probe design and the use of higher fields will or already can partially compensate those drawbacks. Furthermore, the ongoing improvement of sample preparation protocols will likely leverage the problems imposed by incomplete H/D exchange and lead to an overall improved assignment process.

Solid-state NMR experiments, which can access the chemical shifts of a bound ligand, are a valuable tool to either provide restraints for structure calculations that explicitly use those. The experiment presented here aims to simplify and speed up the process of data collection for structure calculations that involve all possible interactions between the protein and the ligand and therefore also aims in the same direction as many other pulse sequences presented here. That is, to increase the information conveyed by a single NMR experiment or to reduce the number of NMR necessary to achieve a certain goal.

Chapter 6

Materials and Methods

6.1 Preparation of protein samples

6.1.1 SH3 of chicken α-spectrin

Expression Isotopically labeled SH3-domain of chicken α-spectrin (SH3) was recombinantly expressed in *Escherichia coli* as described in[37]. The gene encoding pET-3a vector, carrying an Ampicillin resistance, was transformed into BL21-DE3 competent cells by heat shock. To do so, a 100 µl aliquot of competent cells was thawed on ice and incubated with 2 µl of plasmid solution for 10 min on ice. The heat shock was performed for 30 s at 42 °C, and the cells were incubated again for 2 min before finally grown in 500 µl preheated LB medium at 37 °C for at least 30 min. 200 µl of culture were plated on Ampicillin-containing Agar plates and incubated overnight at 37 °C. The next day, one isolated clone was given into 10 ml of LB with Ampicillin and grown overnight (ONC). Cells from 5 ml of culture were transferred to 50 ml of M9 medium and grown over day (ODC). An overnight culture of 50 ml M9, containing 50 % D_2O was prepared, and pelleted cells of 5 ml were transferred. The process was repeated with an ODC of 50 ml M9 medium with 75% D_2O as well as a consecutive ONC with 100 ml M9 containing 100% D_2O, ^{13}C,^{2}D-glucose and $^{15}NH_4Cl$ with cells from 10 ml of the ODC. The following day the ONC was transferred to 1 l of M9 of the same composition to reach a starting OD of around 0.1. For expressing [^{1}H,^{13}C,^{15}N]-labeled SH3, all D_2O adaption steps can be skipped, i.e., after growing an ONC culture in LB, the culture is grown in normal M9 medium containing labeled glucose and NH_4Cl. Expression of SH3 was induced with 1 mM IPTG, final concentration after the OD reached 0.5 and the culture was cooled to 18 °C. After 15-20 h, the cells were harvested by centrifugation for 25 min with 4250 g at 4 °C. The cell pellet can be shock frozen with liquid N_2 for further process or storage.

Purification The cell pellet was diluted 1:5 (w/v) with AEC loading buffer containing DNAse I and 1 tablet of protease inhibitor/l of culture and incubated at RT for at least 30 min. The cells were disrupted by high-pressure homogenization (Emulsiflex) and afterwards centrifuged at 75,000 g at 4 °C to separate soluble and nonsoluble cell parts. The supernatant was loaded on an AEC column (Hi Prep Q FF, GE Healthcare) using the AEC loading buffer, and the column was consecutively washed with an additional 2 CV of AEC loading buffer at a flow rate of 3 ml/min to remove the unbound material. Elution of SH3 and other bound compounds was performed by using a linear gradient from 0-9 % of AEC elution buffer over 8 CV. SH3 containing fractions were pooled, pH shifted to 3.5 with 3 M citric acid solution, and concentrated to ca. 2 ml while not exceeding a protein concentration of 10 mg/ml. Afterwards, the solution was loaded with SEC buffer on an SEC column (HiLoad Superdex 16/600 75g, GE Healthcare) at a flow rate of 1 ml/min and eluted using the same conditions. SH3 was eluted around 108 ml. Fractions were pooled, the buffer exchanged against H_2O at pH 3.5 either by dialysis or repeated washing in Amicon tubes and concentrated to a final concentration of 12 mg/ml.

Crystallization and Cu - doping Crystallization of the purified SH3-domain in H_2O at pH 3.5 was achieved by adding $NH_{42}SO_4$ together with Cu(II)-EDTA at pH 3.5 to a final concentration of 100 mM and 75 mM respectively and a step-wise pH shift to a pH of 8 using

1 M NaOH, crossing the isoelectric point of SH3 at 5.3. To conserve the protein crystals, 1 µl of protease inhibitor cocktail and 0.03 % NaN_3 (final concentration) were added before the pH shift. First microcrystals will start to form immediately, which can be recognized through clouding in the solution. The suspension can be stored for several hours or over night at 4°C until the crystal formation has reached equilibrium.

6.1.2 Tryptophan synthase of *S. typhimurium*

Uniformly [^{2}D,^{13}C,^{15}N]- as well as [^{13}C,^{15}N] labeled, full length $\alpha_2\beta_2$-tryptophan synthase was provided by the group of Leonard Mueller at UC Riverside, California. The protein was expressed and crystallized as described previously[260–262]. Triple labeled [^{2}H,^{13}C,^{15}N] and 100% back-exchanged TS was expressed in *E.coli* CB149 containing a pEBA-10 plasmid with the protein encoding gene in M9 medium. 2 g/l u-[^{2}D,^{13}C] or u-^{13}C glucose and $^{15}NH_4Cl$ each were used in combination with 20 ml/l of 10X BioExpress rich media (Cambridge Isotopes) in D_2O or H_2O respectively. Purification was performed by crystallization and re-crystallization during dialysis in the presence of Cs^+ and the α-subunit inhibitor N-(4'-trifluoro-methoxybenzenesulfonyl)-2-amino-ethyl phosphate (F9).

Cu-doping for PRE was achieved with 40 mM for and 20 mM Cu^{II}-EDTA[37] respectively. The u-[^{2}D,^{13}C,^{15}N] 100% back-exchanged sample, containing 40 mM Cu-EDTA was used for all 4D and 5D backbone assignment experiments as well as the 4D hCCNH (MOCCA) and experiments related to the assignment of histidines and βLys87 as reported in Klein et al. The samples with the labeling schemes u-[^{13}C,^{15}N], β-[^{13}C,^{15}N] and β-[^{2}D,^{13}C,^{15}N] were doped with 20 mM Cu-EDTA. To minimize dilution and the risk of dissolving the protein, all doping reagents, as well as 0.3 µl protease inhibitor (1 tablet in 1 ml of H_2O) and 0.03% NaN_3, were diluted to their final concentration with excess crystallization buffer before adding them stepwise to the suspended crystals.

6.1.3 hCAII with CHD_2-ILV labeling scheme

Expression Isotopically labeled full-length human carbonic anhydrase II (hCAII), containing a GST-tag, was recombinantly expressed in *Escherichia coli*. After thawing on ice, 100 µl of Rosetta 2 (Novagen) cells were incubated with 1 µl of the encoding pGEX plasmid for 20 min, heat-shocked at 42 °C for 30 s and transferred back on ice for an additional 10 min. 275 µl of preheated LB was added to the mixture, and the cells were grown for 2.5 h at 37 °C. 75 µl were plated on an Agar-Agar plate containing Ampicillin and Chloramphenicol and incubated overnight at 37 °C. A single colony was grown overnight in 25 ml of LB with 25 µl of Ampicillin and Chloramphenicol to an OD of around 1.7. Cells of 10 ml of culture were transferred to 50 ml of M9 media (OD 0.2) containing antibiotics and 100 µM $ZnCl_2$ and continuously incubated at 37 °C. After 8 h, the culture grew to an OD of 0.4, and cells of 25 ml were given to 50 ml of M9 containing 50 % D_2O and grown overnight under the same conditions to an OD of 1.8. All cells were pelleted and transferred to 100 ml of M9 containing 75 % D_2O (OD 0.55) and grown over the day (ODC) to an OD of 1.8. 25 ml were centrifuged and given into 1 l of M9 with 100 % D_2O, 1 g of $^{15}NH_4Cl$ and 2 g of [^{13}C,^{2}D]-D-glucose (OD 0.15) to grow overnight to an OD of about 1.4. 30 ml were directly given into 1 l of fresh M9 containing 100 % D_2O as well as 1 g of $^{15}NH_4Cl$ and 2 g of [^{13}C,^{2}D]-D-glucose. The culture

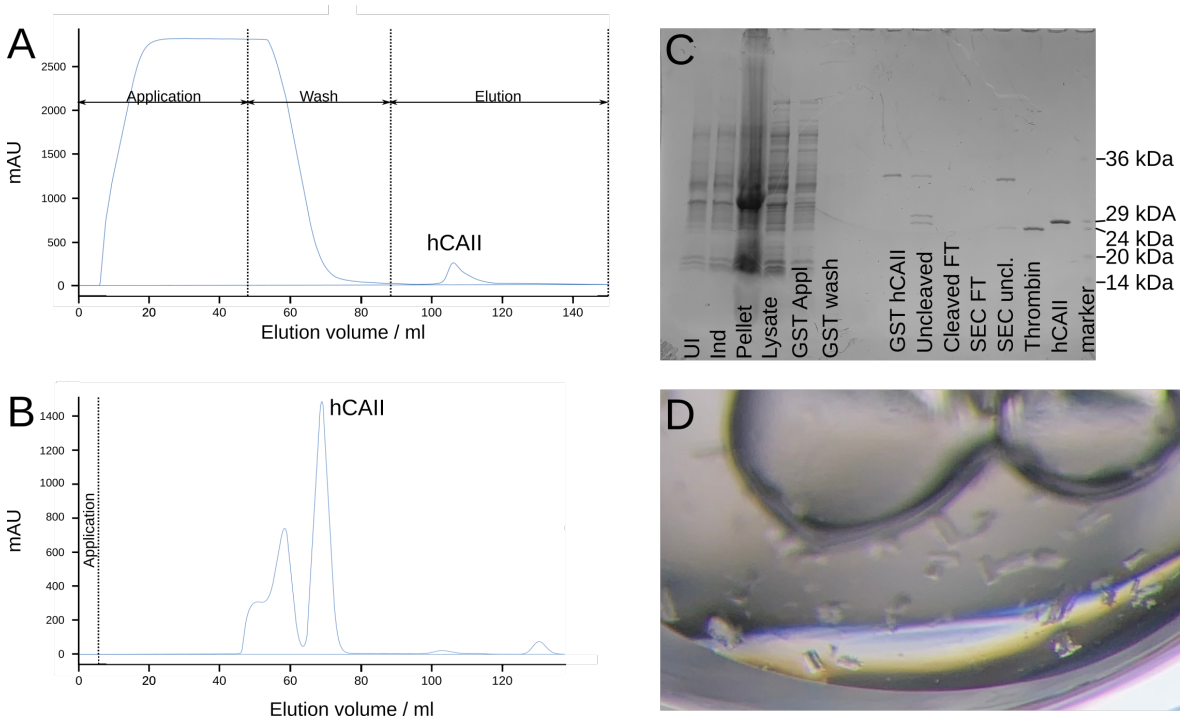


Figure 6.1: Purification and crystallization steps of ILV labeled hCAII. **A)** Chromatogram from GST affinity chromatography on a GST Prep FF 16/10 column (GE Healthcare) at 4 °C. **B)** SEC chromatogram from gel filtration on a HiLoad Superdex 16/600 75g column at 4 °C. hCAII shortly elutes after uncleaved GST-hCAII fusion protein and thrombin. **C)** SDS-PAGE showing uninduced (UI) and induced (Ind) culture as well as Lysate from cell disruption and different fractions of both chromatography steps. **D)** hCAII crystals grown by sitting drop crystallization.

was grown to an OD around 0.5, and hCAII expression was induced by adding IPTG to a final concentration of 1 mM together with 1 mM $ZnCl_2$. One hour before induction (OD 0.4) 50 mg $\alpha-$ketobutyrate as well as 80 mg $\alpha-$ketoisovalerate were added. Expression was done for 15 h at 28 °C. Cells were harvested by centrifugation at 4250 g for 20 min, and the cell pellet was shock-frozen with liquid N_2 for the further process and storage.

Purification Cells were diluted 1:5 (w/v) with lysis buffer and incubated at RT for at least 30 min. Cells were disrupted by high-pressure homogenization (Emulsiflex), and the lysate was centrifuged at 62,500 g at 4 °C to separate soluble and insoluble cell parts. The supernatant was loaded onto a GST binding column (GST Prep FF 16/10, GE Healthcare) using the binding buffer at a flow rate of 2 ml/min and further washed with 2 CV of binding buffer. Afterwards, the bound protein was eluted using 2 CV of elution buffer at a flow rate of 3 ml/min, and hCAII elution started at approximately 100 ml. Protein-containing fractions were pooled, and cleavage of the GST-tag was performed with thrombin overnight. The solution was concentrated to around 2 ml, loaded on a size exclusion column (HiLoad Superdex 16/600 75g, GE Healthcare), and eluted with size exclusion buffer at 1 ml/min. The cleaved protein elutes after ca. 70 ml. Fractions containing cleaved hCAII were pooled and stored at 4 °C.

Crystallization Pooled fractions from the SEC were concentrated to a final concentration of 15 mg/ml using Amicon tubes with an MW cutoff of 12 kDa. Crystals of hCAII were obtained through sitting drop crystallization at 4 °C in a 24 well plate after one week using

Table 6.1: 24 well plate pipetting scheme for crystallization of hCAII

	Component/µl	1	2	3	4	5	6
A	H_2O	69.14	69.14	69.14	69.14	69.14	69.14
	$(NH_4)_2SO_4$	326.9	326.9	326.9	326.9	326.9	326.9
	Tris pH 7.7	-	-	-	-	-	-
	Tris pH 8	44.00	44.00	44.00	44.00	44.00	44.00
B	H_2O	69.14	69.14	69.14	69.14	69.14	69.14
	$(NH_4)_2SO_4$	326.9	326.9	326.9	326.9	326.9	326.9
	Tris pH 7.7	5.59	5.59	5.59	5.59	5.59	5.59
	Tris pH 8	38.41	38.41	38.41	38.41	38.41	38.41
C	H_2O	69.14	69.14	69.14	69.14	69.14	69.14
	$(NH_4)_2SO_4$	326.9	326.9	326.9	326.9	326.9	326.9
	Tris pH 7.7	11.03	11.03	11.03	11.03	11.03	11.03
	Tris pH 8	32.97	32.97	32.97	32.97	32.97	32.97
D	H_2O	69.14	69.14	69.14	69.14	69.14	69.14
	$(NH_4)_2SO_4$	326.9	326.9	326.9	326.9	326.9	326.9
	Tris pH 7.7	16.28	16.28	16.28	16.28	16.28	16.28
	Tris pH 8	27.72	27.72	27.72	27.72	27.72	27.72

a stock of 3.5 M $(NH_4)_2SO_4$ pH 7.7 and different ratios of 500 mM Tris at pH 7.7 and pH 8 to obtain a pH gradient throughout the well plate. Preparation of 24 well plates with the scheme in table 6.1 was automatically done by a dragonfly©pipetting robot. A drop volume of 4 µl with a protein to buffer ratio of 1:1 was used with a 440 µl reservoir. The crystals were harvested simply by pipetting and centrifuged. The supernatant mother liquor was incubated with 2 µl of a 100 mM DMSO-d_6 stock containing the SBR ((R)-N-(3-Indol-1-yl-2-methyl-propyl)-4-sulfamoyl-benzamide) ligand leading to immediate precipitation. The DMSO-d_6 content of the mother liquor was step-wise increased to 7 %, and after centrifugation, the clear supernatant was step-wise added to the crystals for soaking. After an incubation period of 2 d, the supernatant was lyophilized and re-dissolved in the same amount of D_2O and step-wise added back to the crystals to exchange as much H_2O as possible against D_2O. The solvent exchange likely led to the degradation of the crystals. To confirm the state of the hCAII crystals, the white powder was packed into a 0.7 mM rotor, and 2D hNH, as well as hCH spectra, were recorded (see section 4.1.1).

6.2 APSY experiments on SH3

All experiments were recorded on the u-[^{2}D,^{13}C,^{15}N]-labeled and 100 % back-exchanged SH3-domain of chicken α-spectrin at 50 kHz MAS and a ^{1}H Larmor frequency of 700 MHz. The approximate sample temperature was 15 °C. Protein expression and sample preparation, including Cu-doping, were done as described in detail in sections 6.1.1. All experiments are fully based on dipolar transfer mechanisms, i.e., CP[15] for hetero-nuclear transfers and BSH-CP[26] for homonuclear CC transfers. For every experiment, a set of 12 different angles, including two orthogonal planes, were recorded, resulting in a total number of 22 spectra (Table 6.2).

The spectra were initially set up as conventional 3D experiments, out of which the set-up script created a subsequent set of 2D experiments for which the projection angle α was set accordingly and the sweep width calculated as $SW = \sqrt{(\sum_i^{N-1} p_i \cdot SW_i)}$. To compensate

Table 6.2: List of angles used for APSY

$\alpha/°$	Number of angles	Spectra per angle
0	1	1
90	1	1
38.9, 58.2, 21.9, 72.8, 11.4, 81.2, 5.8, 50.4, 28.2, 45.0	10	2

for the lower sensitivity of the projection planes with $\alpha \neq 90°, 0°$, these planes were recorded with twice the number of scans than the orthogonal planes. The number of increments in the indirect dimension, however, was kept constant, which eventually results in different resolutions of the individual projections.

Table 6.3: Acquisition parameters for all 3D APSY experiments acquired on the SH3-domain.

	3D hCONH		
	F3(^{1}H)	F3(^{15}N)	F3(^{13}C)
FID size	2048	320[a]	320[a]
SW / ppm	29.7	39.1 ppm	20.2
Increment size / µs		360	280
Aq. time / ms	49	29	23
scans	8[b]		
Experiment time	37 min/13.3 h[c]		
	3D hCANH		
	F3(^{1}H)	F3(^{15}N)	F3(^{13}C)
FID size	2048	320[a]	320[a]
SW / ppm	29.7	39.1 ppm	28.3
Increment size / µs		360	200
Aq. time / ms	49	29	16
scans	8[b]		
Experiment time	38 min/13 h[c]		
	3D hCOcaNH		
	F3(^{1}H)	F3(^{15}N)	F3(^{13}C)
FID size	2048	320[a]	320[a]
SW / ppm	29.7	39.1 ppm	31.5
Increment size / µs		360	180
Aq. time / ms	49	29	14
scans	16[b]		
Experiment time	61 min/25.5 h[c]		
	3D hCOcaNH		
	F3(^{1}H)	F3(^{15}N)	F3(^{13}C)
FID size	2048	320[a]	320[a]
SW / ppm	29.7	39.1 ppm	31.5
Increment size / µs		360	180
Aq. time / ms	49	29	14

scans	16[b]		
Experiment time	61 min/25.5 h[c]		
	3D hCAcoNH		
	F3(^{1}H)	F3(^{15}N)	F3(^{13}C)
FID size	2048	320[a]	320[a]
SW / ppm		39.1 ppm	28.3
Increment size / µs		360	200
Aq. time / ms	49	29	16
scans	16[b]		
Experiment time	73 min/21.4 h[c]		

[a] corresponding to 80 time increments rather than 160

[b] half for orthogonal planes

[c] time per projections place / total time

6.3 $R_1\rho$ relaxation experiments on TS

$R_1\rho$ experiments were recorded at on-resonance spin-lock fields of 3.3 kHz, 4.4 kHz, 5.5 kHz and 11 kHz for 5 ms, 10 ms, 20 ms, 40 ms, 60 ms and 80 ms. All spectra were processed in TopSpin using a squared sine bell window function in all dimensions with a sine bell shift of 2. hCONH shifts were created from the combined shift list, loaded in POKY[263] and filtered for matching peaks. Peak lists, including peak height, were exported from POKY, and relaxation rates were obtained through an in-house script running in Mathematica or Python developed by Dr. Petra Rovó and Dr. Suresh Vasa in the group.

Table 6.4: Experimental parameters for the $R_1\rho$ experiments performed on TS

	ps4D hCONH-based $R_1\rho$		
	F5(^{1}H)	F4 (^{15}N)	F3(^{13}CO)
FID size	2048	180	70
SW / ppm	29.7447	39.1350	15.7697
Increment size / µs		360	360
Aq. time / ms	49.1	32.4	12.6
scans	8		
Experiment time	ca. 22 h		

6.4 H/D exchange experiments on wild-type TS

An approximation of the extent of back-exchange in water-exchangeable sites, i.e., amides, was achieved by ^{1}H to ^{2}D exchange of an unlabeled sample of wild-type TS and measured by HPLC-MS at different time points. For practical reasons, the reverse experiment (^{1}H→^{2}D), compared to the actual preparation of the NMR sample (^{2}D→^{1}H) was chosen as the evaluation of the results is based on a reference protein sequence which cannot be deuterated. While this is not an exact replication of the sample preparation process, as well as the original sample buffer, it is expected that these H/D exchange (HDX) experiments can still give insights and tendencies about potential back-exchange problems. The experiments followed

Table 6.5: Buffers used for H/D exchange experiments

Protein buffer	Labelling buffer[a]	Quenching buffer[b]	Washing buffer
1 mM TCEP	1 mM TCEP	10 mM TCEP	1.5 M GuHCl
50 mM K_2HPO_4	50 mM K_2HPO_4	50 mM K_2HPO_4	4 % ACN
		6 M Urea	0.1 % formic acid
pH 7.4	pD[c] 7.0	pH 2.37	

[a] prepared in D_2O, pD adjusted with DCl and NaOD.
[b] Quenching buffers with 2 M, 4 M Urea and 0 M, 1 M, 2 M and 4 M GuHCl at different pH values were prepared and tested as well (see table 6.6).
[c] pD = pH_app + 0.4

general recommendations[264] but were specifically performed as recommended by the HPLC-MS facility of the MPI Dortmund and TU Dortmund. The buffer of the original sample was exchanged to "Protein buffer" to remove components incompatible with the further process, e.g., EDTA and DTT, to a concentration of 5 µM and afterwards diluted 1:20 with "Labelling buffer" to reach a concentration of 250 nM. Incubation with deuterated buffer was done at time points 15”, 30”, 45”, 1’, 3’, 10’, 30’, 1 h, 3 h, 6 h, 1 d, 3 d, 7 d, and 14 d. The reaction was quenched by adding the previously optimized "Quenching buffer" in a 1:1 ratio to denaturated the protein, minimize further exchange through high pH and restore conditions ideally suited for the following pepsin digest. The samples were frozen in liquid N_2 and stored at -80 °C to minimize further exchange until the HPLC-MS analysis. Each sample was quickly thawed and digested with pepsin at 20 °C, cooled down again to 0 °C, and the obtained fragments desalted and separated over a C18-HPLC column at low temperatures. Analysis of the obtained fragments was performed by high-resolution mass spectrometry (HRMAS) on a Synapt G2 Si, equipped with an Acquity nanoLC and PAL RTC HDX robot (Waters). The results were evaluated by the software package DynamX (Waters). The extent of deuteration is measured by comparing the mass of each fragment to the mass of the fragments obtained with a nonlabeled sample (see below). Each time point was measured as triplicate.

Previous to the actual HDX experiments, the conditions for the analysis as well as the digestion have to be optimized as they greatly affect the outcome of the experiment[264]. For that, the same procedure as described above is followed, except the incubation step is skipped, and a wide range of quenching buffers is used instead. Optimizing the quenching buffer will ensure optimal conditions for the pepsin digest and, therefore, fragmentation of the protein. Different concentrations of urea and guanidinium-hydrochloride (GuHCl) are screened against

Table 6.6: Optimization of quenching buffers.

Condition	pH alone	pH combined[a]	Sequence coverage	Redundancy
No GuHCl / Urea	1.72	2.63	43.8 %	2.29
1 M GuHCl	1.56	2.55	89.2 %	3.51
2 M GuHCl	1.46	2.64	99.2 %	5.12
4 M GuHCl	1.20	2.58	98.8 %	5.50
2 M Urea	1.96	2.47	92.0 %	2.80
4 M Urea	2.20	2.50	97.4 %	3.48
6 M Urea	2.37	2.64	99.4 %	6.76

[a] 1:1 mixture of "Labeling buffer" and "Quenching buffer"

the sequence coverage resulting from the analysis of the formed fragments (Table 6.6). The

redundancy indicates the readings obtained of any amide occurring in the sequence[264]. As the pH optimum for pepsin is around pH 2.5, it has to be ensured that the 1:1 mixture of "Labeling buffer" and "Quenching buffer" yields a pH around this optimum which is achieved by adjusting the pH of the "Quenching buffer". The values are reported in the table below.

Chapter 7

Supplement

Bibliography

1. Keeler, J. *Understanding NMR spectroscopy* 2nd editio, 526. ISBN: 978-0-470-74608-0 (John Wiley & Sons, Ltd, 2010).
2. Cavanagh, J., Skelton, N. J., Fairbrother, W. J., Rance, M. & Palmer, A. G. *Protein NMR Spectroscopy: Principles and Practice* 2nd. ISBN: 9780080471037 (Elsevier Science, 2010).
3. Gerlach, W. & Stern, O. Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld. *Zeitschrift für Phys.* **9,** 349–352. ISSN: 14346001. https://link.springer.com/article/10.1007/BF01326983 (Dec. 1922).
4. Gerlach, W. & Stern, O. Das magnetische Moment des Silberatoms. *Zeitschrift für Phys.* **9,** 353–355. ISSN: 14346001. https://link.springer.com/article/10.1007/BF01326984 (Dec. 1922).
5. Morris, G. A. & Freeman, R. Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **101,** 760–762. https://pubs.acs.org/doi/abs/10.1021/ja00497a058 (Feb. 1979).
6. Bodenhausen, G. & Ruben, D. J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69,** 185–189. ISSN: 0009-2614 (Jan. 1980).
7. Sattler, M., in nuclear . . ., J. S. .-. P. & 1999, U. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **34,** 93–158. http://pubman.mpdl.mpg.de/pubman/item/escidoc:600636/component/escidoc:2228434/600636.pdf (1999).
8. Hennel, J. W. & Klinowski, J. in, 1–14 (Springer, Berlin, Heidelberg, 2005). https://link.springer.com/chapter/10.1007/b98646.
9. Duer, M. J. *Introduction to Solid-State NMR Spectroscopy* (Blackwell Publishing Ltd, 2004).
10. Andrew, E. R., Bradbury, A. & Eades, R. G. Nuclear magnetic resonance spectra from a crystal rotated at high speed. *Nature* **182,** 1659. ISSN: 00280836. https://doi.org/10.1038/1821659a0 (1958).
11. Lowe, I. J. Free induction decays of rotating solids. *Phys. Rev. Lett.* **2,** 285–287. ISSN: 00319007. https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.2.285 (Apr. 1959).

12. Vasa, S. K., Rovó, P. & Linser, R. Protons as Versatile Reporters in Solid-State NMR Spectroscopy. *Acc. Chem. Res.* **51,** 1386–1395. ISSN: 15204898. https://pubs.acs.org/doi/abs/10.1021/acs.accounts.8b00055 (June 2018).

13. Schledorn, M. *et al.* Protein NMR Spectroscopy at 150 kHz Magic-Angle Spinning Continues To Improve Resolution and Mass Sensitivity. *ChemBioChem* **21,** 2540–2548. ISSN: 14397633 (Sept. 2020).

14. Pine, A., Gibbyt, M. G. & Waugh, J. S. Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* **59,** 569–590 (1973).

15. Hartmann, S. R. & Hahn, E. L. Nuclear Double Resonance in the Rotating Frame. *Phys. Rev.* **128,** 2042–2053. ISSN: 0031-899X. https://link.aps.org/doi/10.1103/PhysRev.128.2042 (Dec. 1962).

16. Marks, D. & Vega, S. A Theory for Cross-Polarization NMR of Nonspinning and Spinning Samples. *J. Magn. Reson. A* **118,** 157–172. ISSN: 1064-1858 (Feb. 1996).

17. Meier, B. H. Cross polarization under fast magic angle spinning: thermodynamical considerations. *Chem. Phys. Lett.* **188,** 201–207. ISSN: 0009-2614 (Jan. 1992).

18. Hediger, S., Meier, B. H., Kurur, N. D., Bodenhausen, G. & Ernst, R. R. NMR cross polarization by adiabatic passage through the Hartmann—Hahn condition (APHH). *Chem. Phys. Lett.* **223,** 283–288. ISSN: 0009-2614 (June 1994).

19. Hediger, S., Meier, B. H. & Ernst, R. R. Adiabatic passage Hartmann-Hahn cross polarization in NMR under magic angle sample spinning. *Chem. Phys. Lett.* **240,** 449–456. ISSN: 0009-2614 (July 1995).

20. Baldus, M., Geurts, D., Hediger, S. & Meier, B. H. Efficient 15N-13C Polarization Transfer by Adiabatic-Passage Hartmann-Hahn Cross Polarization. *J. Magn. Reson. Reson. A* **118,** 140–144. https://www.researchgate.net/publication/245454016 (1996).

21. Baldus, M., Petkova, A. T., Herzfeld, J. & Griffin, R. G. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **95,** 1197–1207. https://www.tandfonline.com/doi/abs/10.1080/00268979809483251 (2009).

22. Laage, S. *et al.* Band-Selective 1H13C Cross-Polarization in Fast Magic Angle Spinning Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **130,** 17216–17217. https://pubs.acs.org/doi/abs/10.1021/ja805926d (Dec. 2008).

23. Gullion, T. & Schaefer, J. Rotational-echo double-resonance NMR. *J. Magn. Reson.* **81,** 196–200. ISSN: 0022-2364 (Jan. 1989).

24. Bennett, A. E. *et al.* Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **8624,** 8624–8627. ISSN: 00219606 (1992).

25. Bennett, A. E. *et al.* Homonuclear radio frequency-driven recoupling in rotating solids Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* **108,** 9463–9479 (1998).

26. Chevelkov, V., Giller, K., Becker, S. & Lange, A. Efficient CO-CA transfer in highly deuterated proteins by band-selective homonuclear cross-polarization. *J. Magn. Reson.* **230,** 205–211. ISSN: 10907807. http://dx.doi.org/10.1016/j.jmr.2013.02.021 (2013).

27. Shi, C. *et al.* BSH-CP based 3D solid-state NMR experiments for protein resonance assignment. *J. Biomol. NMR* **59,** 15–22. ISSN: 15735001 (2014).

28. Verel, R., Ernst, M. & Meier, B. H. Adiabatic dipolar recoupling in solid-state NMR: The DREAM scheme. *J. Magn. Reson.* **150,** 81–99. ISSN: 10907807 (2001).

29. Westfeld, T. *et al.* Properties of the DREAM scheme and its optimization for application to proteins. *J. Biomol. NMR* **53,** 103–112 (2012).

30. Paulson, E. K. *et al.* Sensitive High Resolution Inverse Detection NMR Spectroscopy of Proteins in the Solid State. *J. Am. Chem. Soc.* **125,** 15831–15836. https://pubs.acs.org/sharingguidelines (2003).

31. Reif, B. & Griffin, R. G. 1H detected 1H,15N correlation spectroscopy in rotating solids. *J. Magn. Reson.* **160,** 78–83. ISSN: 1090-7807 (Jan. 2003).

32. Chevelkov, V., Rehbein, K., Diehl, A. & Reif, B. Ultrahigh Resolution in Proton Solid-State NMR Spectroscopy at High Levels of Deuteration. *Angew. Chem., Int. Ed.* **45,** 3878–3881. ISSN: 1433-7851. http://doi.wiley.com/10.1002/anie.200600328 (June 2006).

33. Penzel, S. *et al.* Spinning faster: protein NMR at MAS frequencies up to 126 kHz. *J. Biomol. NMR* **73.** ISSN: 15735001 (2019).

34. Ishii, Y. & Tycko, R. Sensitivity Enhancement in Solid State 15N NMR by Indirect Detection with High-Speed Magic Angle Spinning. *J. Magn. Reson.* **142,** 199–204. ISSN: 1090-7807 (Jan. 2000).

35. Zhou, D. H. *et al.* Proton-detected solid-state NMR spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning. *J. Am. Chem. Soc.* **129,** 11791–11801. ISSN: 00027863 (2007).

36. Ishii, Y., Yesinowski, J. P. & Tycko, R. Sensitivity Enhancement in Solid-State 13C NMR of Synthetic Polymers and Biopolymers by 1H NMR Detection with High-Speed Magic Angle Spinning. *J. Am. Chem. Soc.* **123,** 2921–2922. https://pubs.acs.org/doi/abs/10.1021/ja015505j (2001).

37. Linser, R., Chevelkov, V., Diehl, A. & Reif, B. Sensitivity enhancement using paramagnetic relaxation in MAS solid-state NMR of perdeuterated proteins. *J. Magn. Reson.* **189,** 209–216. ISSN: 1090-7807. https://www.sciencedirect.com/science/article/pii/S1090780707002753 (Dec. 2007).

38. Linser, R., Fink, U. & Reif, B. Proton-detected scalar coupling based assignment strategies in MAS solid-state NMR spectroscopy applied to perdeuterated proteins. *J. Magn. Reson.* **193,** 89–93. ISSN: 10907807 (2008).

39. Akbey, Ü. *et al.* Optimum levels of exchangeable protons in perdeuterated proteins for proton detection in MAS solid-state NMR spectroscopy. *J. Biomol. NMR* **46,** 67–73. ISSN: 1573-5001. https://link.springer.com/article/10.1007/s10858-009-9369-0 (Aug. 2009).

40. Linser, R., Bardiaux, B., Higman, V., Fink, U. & Reif, B. Structure Calculation from Unambiguous Long-Range Amide and Methyl 1H1H Distance Restraints for a Microcrystalline Protein with MAS Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **133,** 5905–5912. ISSN: 0002-7863. https://pubs.acs.org/doi/10.1021/ja110222h (Apr. 2011).

41. Linser, R. *et al.* Proton-Detected Solid-State NMR Spectroscopy of Fibrillar and Membrane Proteins. *Angew. Chem., Int. Ed.* **50,** 4508–4512. ISSN: 14337851. http://doi.wiley.com/10.1002/anie.201008244 (May 2011).

42. Lewandowski, J. R. *et al.* Enhanced resolution and coherence lifetimes in the solid-state NMR spectroscopy of perdeuterated proteins under ultrafast magic-angle spinning. *J. Phys. Chem. Lett.* **2,** 2205–2211. ISSN: 19487185 (2011).

43. Asami, S. & Reif, B. Proton-detected solid-state NMR spectroscopy at aliphatic sites: Application to crystalline systems. *Acc. Chem. Res.* **46,** 2089–2097. ISSN: 00014842. https://pubs.acs.org/sharingguidelines (Sept. 2013).

44. Knight, M. J. *et al.* Fast resonance assignment and fold determination of human superoxide dismutase by high-resolution proton-detected solid-state MAS NMR spectroscopy. *Angew. Chem., Int. Ed.* **50,** 11697–11701. ISSN: 14337851 (2011).

45. Kulminskaya, N., Vasa, S. K., Giller, K., Becker, S. & Linser, R. Asynchronous through-bond homonuclear isotropic mixing: Application to carbon-carbon transfer in perdeuterated proteins under MAS. *J. Biomol. NMR* **63,** 245–253. ISSN: 15735001 (Nov. 2015).

46. Kulminskaya, N. *et al.* Access to side-chain carbon information in deuterated solids under fast MAS through non-rotor-synchronized mixing. *Chem. Commun.* **52,** 268–271. ISSN: 1364548X (Jan. 2016).

47. Linser, R. Side-chain to backbone correlations from solid-state NMR of perdeuterated proteins through combined excitation and long-range magnetization transfers. *J. Biomol. NMR* **51,** 221–226. ISSN: 09252738 (2011).

48. Huber, M., Böckmann, A., Hiller, S. & Meier, B. H. 4D solid-state NMR for protein structure determination. *Phys.Chem.Chem.Phys* **14,** 5239–5246. ISSN: 14639076 (Apr. 2012).

49. Mance, D. *et al.* An Efficient Labelling Approach to Harness Backbone and Side-Chain Protons in 1H-Detected Solid-State NMR Spectroscopy. *Angew. Chem., Int. Ed.* **54,** 15799–15803. ISSN: 1521-3773. https://onlinelibrary.wiley.com/doi/full/10.1002/anie.201509170%20https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201509170%20https://onlinelibrary.wiley.com/doi/10.1002/anie.201509170 (Dec. 2015).

50. Andreas, L. B. *et al.* Structure of fully protonated proteins by proton-detected magic-angle spinning NMR. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 9187–9192. ISSN: 0027-8424. http://www.pnas.org/lookup/doi/10.1073/pnas.1602248113 (Aug. 2016).

51. Stanek, J. *et al.* NMR Spectroscopic Assignment of Backbone and Side-Chain Protons in Fully Protonated Proteins: Microcrystals, Sedimented Assemblies, and Amyloid Fibrils. *Angew. Chem., Int. Ed.* **55,** 15504–15509. ISSN: 14337851. http://doi.wiley.com/10.1002/anie.201607084 (Dec. 2016).

52. Böckmann, A., Ernst, M. & Meier, B. H. Spinning proteins, the faster, the better? *J. Magn. Reson.* **253,** 71–79. ISSN: 10960856 (2015).

53. 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 Nucl. Magn. Reson.* **87,** 126–136. ISSN: 09262040. https://doi.org/10.1016/j.ssnmr.2017.07.004 (2017).

54. Vasa, S. K., Singh, H., Grohe, K. & Linser, R. Assessment of a Large Enzyme–Drug Complex by Proton-Detected Solid-State NMR Spectroscopy without Deuteration. *Angew. Chem., Int. Ed.* **58,** 5758–5762. ISSN: 1433-7851. https://onlinelibrary.wiley.com/doi/10.1002/anie.201811714 (Apr. 2019).

55. Lewandowski, J. R., Sass, H. J., Grzesiek, S., Blackledge, M. & Emsley, L. Site-Specific Measurement of Slow Motions in Proteins. *J. Am. Chem. Soc* **133,** 16762–16765. https://pubs.acs.org/sharingguidelines (2011).

56. Linser, R., Fink, U. & Reif, B. Narrow carbonyl resonances in proton-diluted proteins facilitate NMR assignments in the solid-state. *J. Biomol. NMR* **47,** 1–6. ISSN: 09252738 (May 2010).

57. Linser, R., Fink, U. & Reif, B. Assignment of Dynamic Regions in Biological Solids Enabled by Spin-State Selective NMR Experiments. *J. Am. Chem. Soc.* **132,** 8891–8893. ISSN: 0002-7863. https://pubs.acs.org/doi/10.1021/ja102612m (July 2010).

58. Penzel, S. *et al.* Protein resonance assignment at MAS frequencies approaching 100 kHz: a quantitative comparison of J-coupling and dipolar-coupling-based transfer methods. *J. Biomol. NMR* **63,** 165–186. ISSN: 0925-2738. http://link.springer.com/10.1007/s10858-015-9975-y (Oct. 2015).

59. Zhang, R., Nishiyama, Y., Sun, P. & Ramamoorthy, A. Phase cycling schemes for finite-pulse-RFDR MAS solid state NMR experiments. *J. Magn. Reson.* **252,** 55–66. ISSN: 1090-7807 (Mar. 2015).

60. Levitt, M. H., Grant, D. M., Harris, R. K. & Wiley, J. Symmetry-Based Pulse Sequences in Magic-Angle Spinning Solid-State NMR Symmetry-Based Pulse Sequences in Magic-Angle Spinning Solid-State NMR. *Encycl. Nucl. Magn. Reson.* **9,** 165–196 (2002).

61. Barbet-Massin, E. *et al.* Rapid Proton-Detected NMR Assignment for Proteins with Fast Magic Angle Spinning. *J. Am. Chem. Soc* **136,** 11. http://perso.ens-lyon.fr/torsten.herrmann/Software. (2014).

62. Xiang, S. *et al.* Sequential Backbone Assignment Based on Dipolar Amide-to-Amide Correlation Experiments. English. *J. Biomol. NMR* **62,** 303–311. http://dx.doi.org/10.1007/s10858-015-9945-4 (2015).

63. Andreas, L. B. *et al.* Protein residue linking in a single spectrum for magic-angle spinning NMR assignment. *J. Biomol. NMR* **62,** 253–261 (2015).

64. Vasa, S. K., Singh, H., Rovó, P. & Linser, R. Dynamics and Interactions of a 29 kDa Human Enzyme Studied by Solid-State NMR. *J. Phys. Chem. Lett.* **9,** 13071311 (2018).

65. Stanek, J. *et al.* Automated backbone NMR resonance assignment of large proteins using redundant linking from a single simultaneous acquisition. *J. Am. Chem. Soc.* **142,** 5793–5799. ISSN: 0002-7863. https://pubs.acs.org/doi/10.1021/jacs.0c00251 (Mar. 2020).

66. Orton, H. W. *et al.* Protein NMR resonance assignment without spectral analysis: 5D SOlid-state Automated Projection SpectroscopY (SO-APSY). *Angew. Chem., Int. Ed.* **59,** 2380. ISSN: 14337851. http://doi.wiley.com/10.1002/anie.201912211 (Oct. 2020).

67. Klein, A. *et al.* Atomic-resolution Chemical Characterization of (2x)72 kDa Tryptophan Synthase via Four- and Five-Dimensional 1H-Detected Solid-State NMR. *Proc. Natl. Acad. Sci. U.S.A.* **119,** e2114690119 (2022).

68. Xiang, S. Q., Biernat, J., Mandelkow, E., Becker, S. & Linser, R. Backbone assignment for minimal protein amounts of low structural homogeneity in the absence of deuteration. *Chem. Commun.* **52,** 4002–4005. ISSN: 1364548X. http://dx.doi.org/10.1039/C5CC09160H (2016).

69. Fraga, H. *et al.* Solid-State NMR H–N–(C)–H and H–N–C–C 3D/4D Correlation Experiments for Resonance Assignment of Large Proteins. *ChemPhysChem* **18,** 2697–2703. ISSN: 14397641. http://dx.doi.org/10.1002/cphc.201700572 (Oct. 2017).

70. Zinke, M. *et al.* Bacteriophage Tail-Tube Assembly Studied by Proton-Detected 4D Solid-State NMR. *Angew. Chem., Int. Ed.* **56,** 9497–9501. ISSN: 15213773 (2017).

71. Zhou, D. H. *et al.* Solid-State Protein-Structure Determination with Proton-Detected Triple-Resonance 3D Magic-Angle-Spinning NMR Spectroscopy. *Angew. Chem., Int. Ed.* **46,** 8380–8383. ISSN: 1521-3773. https://onlinelibrary.wiley.com/doi/full/10.1002/anie.200702905%20https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.200702905%20https://onlinelibrary.wiley.com/doi/10.1002/anie.200702905 (Nov. 2007).

72. Grohe, K. *et al.* Exact distance measurements for structure and dynamics in solid proteins by fast-magic-angle-spinning NMR. *Chem. Commun.* **55,** 7899–7902. ISSN: 1364-548X. https://pubs.rsc.org/en/content/articlehtml/2019/cc/c9cc02317h%20https://pubs.rsc.org/en/content/articlelanding/2019/cc/c9cc02317h (July 2019).

73. Huber, M. *et al.* A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination. *ChemPhysChem* **12,** 915–918. ISSN: 14394235. http://doi.wiley.com/10.1002/cphc.201100062 (Apr. 2011).

74. Linser, R. *et al.* Solid-state NMR structure determination from diagonal-compensated, sparsely nonuniform-sampled 4D proton-proton restraints. *J. Am. Chem. Soc.* **136,** 11002–11010. ISSN: 15205126 (Aug. 2014).

75. Lalli, D. *et al.* Proton-Based Structural Analysis of a Heptahelical Transmembrane Protein in Lipid Bilayers. *J. Am. Chem. Soc.* **139,** 13006–13012. https://pubs.acs.org/doi/abs/10.1021/jacs.7b05269 (Sept. 2017).

76. Schanda, P. & Ernst, M. *Studying dynamics by magic-angle spinning solid-state NMR spectroscopy: Principles and applications to biomolecules* 2016.

77. Rovó, P. Recent advances in solid-state relaxation dispersion techniques. *Solid State Nucl. Magn. Reson.* **00,** 101665. ISSN: 09262040 (2020).

78. Rovó, P. & Linser, R. Microsecond Time Scale Proton Rotating-Frame Relaxation under Magic Angle Spinning. *J. Phys. Chem. B* **121,** 6117–6130. ISSN: 15205207 (2017).

79. Kurauskas, V. *et al.* Slow conformational exchange and overall rocking motion in ubiquitin protein crystals. *Nat Commun* **8,** 1–11. ISSN: 20411723. `http://dx.doi.org/10.1038/s41467-017-00165-8` (2017).

80. Kazimierczuk, K., Stanek, J., Zawadzka-Kazimierczuk, A. & Koźmiński, W. Random sampling in multidimensional NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **57,** 420–434. ISSN: 00796565 (2010).

81. Arthanari, H. *et al. Novel Sampling Approaches in Higher Dimensional NMR* 1st ed. (eds Billeter, M. & Orekhov, V.) (Springer, Berlin, Heidelberg, 2012).

82. Nowakowski, M. *et al.* Applications of high dimensionality experiments to biomolecular NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **90-91,** 49–73. ISSN: 00796565 (2015).

83. Robson, S., Arthanari, H., Hyberts, S. G. & Wagner, G. *Nonuniform Sampling for NMR Spectroscopy* 1st ed., 263–291. ISBN: 9780128138601. `http://dx.doi.org/10.1016/bs.mie.2018.09.009` (Elsevier Inc., 2019).

84. Palmer, M. R. *et al.* Sensitivity of Nonuniform Sampling NMR. *J. Phys. Chem. B* **119,** 6502–6515. `https://pubs.acs.org/doi/full/10.1021/jp5126415` (June 2015).

85. Kazimierczuk, K., Zawadzka, A. & Koźmiński, W. Narrow peaks and high dimensionalities: Exploiting the advantages of random sampling. *J. Magn. Reson.* **197,** 219–228. ISSN: 10907807 (2009).

86. Hyberts, S. G., Takeuchi, K. & Wagner, G. Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *J. Am. Chem. Soc.* **132,** 2145–2147. ISSN: 00027863. `https://pubs.acs.org/doi/abs/10.1021/ja908004w` (Mar. 2010).

87. Kazimierczuk, K., Zawadzka, A., Koźmiński, W. & Zhukov, I. Random sampling of evolution time space and Fourier transform processing. *J. Biomol. NMR* **36,** 157–168. ISSN: 09252738 (2006).

88. Hiller, S., Fiorito, F., Wuthrich, K. & Wider, G. Automated projection spectroscopy (APSY). *Proc. Natl. Acad. Sci. U.S.A.* **102,** 10876–10881. ISSN: 0027-8424 (2005).

89. Nagayama, K., Backmann, P., Wüthrich, K. & Ernst, R. The Use of Cross-Sections and of Projections in Two-dimensional NMR Spectroscopy. *J. Magn. Reson.* **31,** 133–148 (1978).

90. Hiller, S., Wasmer, C., Wider, G. & Wüthrich, K. Sequence-specific resonance assignment of soluble nonglobular proteins by 7D APSY-NMR spectroscopy. *J. Am. Chem. Soc.* **129,** 10823–10828. ISSN: 00027863 (2007).

91. Hiller, S., Wider, G. & Wüthrich, K. APSY-NMR with proteins: Practical aspects and backbone assignment. *J. Biomol. NMR* **42,** 179–195. ISSN: 09252738 (2008).

92. Kupče, E. & Freeman, R. Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy. *J. Am. Chem. Soc.* **126,** 6429–6440. ISSN: 00027863 (2004).

93. Yao, X., Becker, S. & Zweckstetter, M. A six-dimensional alpha proton detection-based APSY experiment for backbone assignment of intrinsically disordered proteins. *J. Biomol. NMR* **60,** 231–240. ISSN: 15735001 (2014).

94. Dutta, S. K. *et al.* APSY-NMR for protein backbone assignment in high-throughput structural biology. *J. Biomol. NMR* **61,** 47–53. ISSN: 15735001 (2015).

95. Krähenbühl, B. & Wider, G. Automated Projection Spectroscopy (APSY) for the Assignment of NMR Resonances of Biological Macromolecules. *Chim. Int. J. Chem.* **66,** 770–774. ISSN: 0009-4293 (2012).

96. Gossert, A. D., Hiller, S. & Fernández, C. Automated NMR Resonance Assignment of Large Proteins for Protein-Ligand Interaction Studies. *J. Am. Chem. Soc.* **133,** 210–213. ISSN: 00027863 (2011).

97. Fiorito, F., Hiller, S., Wider, G. & Wüthrich, K. Automated resonance assignment of proteins: 6D APSY-NMR. *J. Biomol. NMR* **35,** 27–37. ISSN: 09252738. `https://doi.org/10.1007/s10858-006-0030-x` (2006).

98. Krähenbühl, B., Boudet, J. & Wider, G. 4D experiments measured with APSY for automated backbone resonance assignments of large proteins. *J. Biomol. NMR* **56,** 149–154. ISSN: 09252738 (2013).

99. Hiller, S., Joss, R. & Wider, G. Automated NMR Assignment of Protein Side Chain Resonances Using Automated Projection Spectroscopy (APSY), 12073–12079 (2008).

100. Narayanan, R. L. *et al.* Automatic assignment of the intrinsically disordered protein Tau with 441-residues. *J. Am. Chem. Soc.* **132,** 11906–11907. ISSN: 00027863. `http://pubs.acs.org.` (Sept. 2010).

101. Murrali, M. G. *et al.* 13C APSY-NMR for sequential assignment of intrinsically disordered proteins. *J. Biomol. NMR* **70,** 167–175. ISSN: 15735001. `https://doi.org/10.1007/s10858-018-0167-4` (Feb. 2018).

102. Kosiński, K., Stanek, J., Górka, M. J., Żerko, S. & Koźmiński, W. Reconstruction of non-uniformly sampled five-dimensional NMR spectra by signal separation algorithm. *J. Biomol. NMR* **68,** 129–138. ISSN: 15735001 (June 2017).

103. Brigham, E. O. *The Fast Fourier Transform and Its Applications* ISBN: 9780133075052 (Prentice Hall, 1988).

104. Stanek, J. & Koźmiński, W. Iterative algorithm of discrete Fourier transform for processing randomly sampled NMR data sets. *J. Biomol. NMR* **47,** 65–77. ISSN: 09252738 (May 2010).

105. Stanek, J., Augustyniak, R. & Koźmiński, W. Suppression of sampling artefacts in high-resolution four-dimensional NMR spectra using signal separation algorithm. *J. Magn. Reson.* **214,** 91–102. ISSN: 10907807 (2012).

106. Alberts, B. *et al. Molecular Biology of the Cell* 5th ed. (Garland Science, 2008).

107. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7,** 95–99. ISSN: 0022-2836 (July 1963).

108. Tonlolo, C. & Benedetti, E. The polypeptide 3_{10}-helix. *Trends Biochem. Sci.* **16,** 350–353. ISSN: 0968-0004 (Jan. 1991).

109. IUPAC-IUB Commission on Biochemical Nomenclature. Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains: Tentative Rules (1969). *J. Biol. Chem.* **245,** 6489–6497. ISSN: 0021-9258 (Dec. 1970).

110. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comp. Chem.* **25,** 1605–1612. ISSN: 1096-987X. https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.20084%20https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20084%20https://onlinelibrary.wiley.com/doi/10.1002/jcc.20084 (Oct. 2004).

111. Bishop, J. M. *Viral oncogenes* Aug. 1985.

112. Ball, L. J., Kühne, R., Schneider-Mergener, J. & Oschkinat, H. Recognition of proline-rich motifs by protein-protein-interaction domains. *Angew. Chem., Int. Ed.* **44,** 2852–2869. ISSN: 14337851. https://onlinelibrary.wiley.com/doi/full/10.1002/anie.200400618 (May 2005).

113. Ferguson, F. M. & Gray, N. S. Kinase inhibitors: the road ahead. *Nat. Rev. Drug Discov.* **17,** 353–377. ISSN: 1474-1784. https://doi.org/10.1038/nrd.2018.21 (2018).

114. Pawson, T. & Schlessingert, J. *SH2 and SH3 domains* July 1993.

115. Ren, R., Mayer, B., Cicchetti, P. & Baltimore, D. Identification of a ten-amino acid proline-rich SH3 binding site. *Science (80-.).* **259,** 1157–1161. ISSN: 0036-8075. https://science.sciencemag.org/content/259/5098/1157%20https://science.sciencemag.org/content/259/5098/1157.abstract (Feb. 1993).

116. Stahl, M. L., Ferenz, C. R., Kelleher, K. L., Kriz, R. W. & Knopf, J. L. Sequence similarity of phospholipase C with the non-catalytic region of src. *Nature* **332,** 269–272. ISSN: 1476-4687. https://www.nature.com/articles/332269a0 (1988).

117. Kaneko, T., Sidhu, S. S. & Li, S. S. *Evolving specificity from variability for protein interaction domains* Apr. 2011.

118. Larson, S. M., Davidson, A. R. & Davidson, A. R. The identification of conserved interactions within the SH3 domain by alignment of sequences and structures. *Protein Sci.* **9,** 2170–2180. ISSN: 09618368. https://onlinelibrary.wiley.com/doi/full/10.1110/ps.9.11.2170 (Jan. 2000).

119. Yu, H. *et al.* Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science (80-.).* **258,** 1665–1668. ISSN: 00368075. http://science.sciencemag.org/content/258/5088/1665.abstract (Dec. 1992).

120. Saksela, K. & Permi, P. *SH3 domain ligand binding: What's the consensus and where's the specificity?* Aug. 2012.

121. Rovó, P. *et al.* Mechanistic Insights into Microsecond Time-Scale Motion of Solid Proteins Using Complementary 15 N and 1 H Relaxation Dispersion Techniques. *J. Am. Chem. Soc.* **141,** 858–869. ISSN: 15205126 (Jan. 2019).

122. Chevelkov, V. *et al.* Differential line broadening in MAS solid-state NMR due to dynamic interference. *J. Am. Chem. Soc.* **129,** 10195–10200. ISSN: 00027863. https://pubs.acs.org/sharingguidelines (Aug. 2007).

123. Castellani, F. *et al.* Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420,** 99–102. ISSN: 1476-4687. https://www.nature.com/articles/nature01070 (Nov. 2002).

124. Klein, A., Vasa, S. K. & Linser, R. Automated projection spectroscopy in solid-state NMR. *J. Biomol. NMR* **72,** 163–170. ISSN: 0925-2738. http://link.springer.com/10.1007/s10858-018-0215-0 (Dec. 2018).

125. Asami, S., Schmieder, P. & Reif, B. High Resolution 1H-Detected Solid-State NMR Spectroscopy of Protein Aliphatic Resonances: Access to Tertiary Structure Information. *J. Am. Chem. Soc.* **132,** 15133–15135. https://pubs.acs.org/doi/abs/10.1021/ja106170h (Nov. 2010).

126. Grohe, K. *et al.* Protein Motional Details Revealed by Complementary Structural Biology Techniques. *Structure* **28,** 1–11 (2020).

127. Chevelkov, V., Fink, U. & Reif, B. Accurate Determination of Order Parameters from 1H,15N Dipolar Couplings in MAS Solid-State NMR Experiments. *J. Am. Chem. Soc.* **131,** 14018–14022. https://pubs.acs.org/doi/full/10.1021/ja902649u (Oct. 2009).

128. Chevelkov, V., Xue, Y., Linser, R., Skrynnikov, N. R. & Reif, B. Comparison of Solid-State Dipolar Couplings and Solution Relaxation Data Provides Insight into Protein Backbone Dynamics. *J. Am. Chem. Soc.* **132,** 5015–5017. https://pubs.acs.org/doi/full/10.1021/ja100645k (Apr. 2010).

129. Rovó, P. & Linser, R. Microsecond Timescale Protein Dynamics: a Combined Solid-State NMR Approach. *ChemPhysChem* **19,** 34–39. http://doi.wiley.com/10.1002/cphc.201701238 (Jan. 2018).

130. Eliot, A. C. & Kirsch, J. F. Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **73,** 383–415. ISSN: 0066-4154 (2004).

131. Mozzarelli, A. & Bettati, S. Exploring the pyridoxal 5-phosphate-dependent enzymes. *Chem. Rec.* **6,** 275–287. ISSN: 15278999 (2006).

132. Percudani, R. & Peracchi, A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep* **4,** 850–854. ISSN: 1469221X (2003).

133. Jansonius, J. N. Structure, evolution and action of vitamin B6-dependent enzymes. *Curr Opin Struct Biol* **8,** 759–769. ISSN: 0959440X (1998).

134. Abrahams, K. A. *et al.* Inhibiting mycobacterial tryptophan synthase by targeting the inter-subunit interface. *Sci. Rep.* **7,** 1–15. ISSN: 20452322 (Dec. 2017).

135. Rocha, J. F., Pina, A. F., Sousa, S. F. & Cerqueira, N. M. *PLP-dependent enzymes as important biocatalysts for the pharmaceutical, chemical and food industries: A structural and mechanistic perspective* Sept. 2019.

136. Daidone, F., Montioli, R., Paiardini, A., Cellini, B. & Macchiarulo, A. Identification by Virtual Screening and In Vitro Testing of Human DOPA Decarboxylase Inhibitors. *PLoS One* **7,** 31610. www.plosone.org (2012).

137. Sarup, A., Larsson, O. M. & Schousboe, A. *GABA transporters and GABA-transaminase as drug targets.* 2003.

138. Michalska, K. *et al.* Allosteric inhibitors of Mycobacterium tuberculosis tryptophan synthase. *Protein Sci.* **29,** 779–788. ISSN: 0961-8368. https://onlinelibrary.wiley.com/doi/abs/10.1002/pro.3825 (Mar. 2020).

139. Paiardini, A., Contestabile, R., Buckle, A. M. & Cellini, B. PLP-Dependent Enzymes. http://dx.doi.org/10.1155/2014/856076 (2014).

140. Buller, A. R. *et al.* Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 14599–604. ISSN: 1091-6490. http://www.ncbi.nlm.nih.gov/pubmed/26553994%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4664345 (Nov. 2015).

141. Dick, M., Sarai, N. S., Martynowycz, M. W., Gonen, T. & Arnold, F. H. Tailoring Tryptophan Synthase TrpB for Selective Quaternary Carbon Bond Formation. *J. Am. Chem. Soc.* **141,** 19817–19822. ISSN: 15205126 (Dec. 2019).

142. Schupfner, M., Straub, K., Busch, F., Merkl, R. & Sterner, R. Analysis of allosteric communication in a multienzyme complex by ancestral sequence reconstruction. *Proc. Natl. Acad. Sci. U.S.A.* **117,** 346–354. ISSN: 0027-8424. https://www.pnas.org/content/117/1/346 (Jan. 2020).

143. Romney, D. K., Murciano-Calles, J., Wehrmüller, J. E. & Arnold, F. H. Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. *J. Am. Chem. Soc.* **139,** 10769–10776. ISSN: 15205126 (Aug. 2017).

144. Arnold, F. H. & Volkov, A. A. Directed evolution of biocatalysts. *Curr Opin Chem Biol* **3,** 54–59. ISSN: 1367-5931 (Feb. 1999).

145. Arnold, F. H. Design by Directed Evolution. *Acc. Chem. Res.* **31,** 125–131. https://pubs.acs.org/doi/full/10.1021/ar960017f (1998).

146. Niks, D. *et al.* Allostery and substrate channeling in the tryptophan synthase bienzyme complex: Evidence for two subunit conformations and four quaternary states. *Biochemistry* **52,** 6396–6411. ISSN: 00062960. http://pubs.acs.org/doi/10.1021/bi400795e (Sept. 2013).

147. Caulkins, B. G. *et al.* NMR Crystallography of a Carbanionic Intermediate in Tryptophan Synthase: Chemical Structure, Tautomerization, and Reaction Specificity. *J. Am. Chem. Soc.* **138,** 15214–15226. ISSN: 15205126 (Nov. 2016).

148. Hilario, E. *et al.* Visualizing the tunnel in tryptophan synthase with crystallography: Insights into a selective filter for accommodating indole and rejecting water. *Biochim. Biophys. Acta - Proteins Proteom.* **1864,** 268–279. ISSN: 1570-9639. https://www.sciencedirect.com/science/article/pii/S1570963915003052?via%7B%5C%%7D3Dihub (Mar. 2016).

149. Caulkins, B. G. *et al.* Catalytic roles of βLys87 in tryptophan synthase: 15N solid state NMR studies. *Biochim. Biophys. Acta - Proteins Proteom.* **1854,** 1194–1199. ISSN: 18781454 (2015).

150. Caulkins, B. G. *et al.* Protonation States of the Tryptophan Synthase Internal Aldimine Active Site from Solid-State NMR Spectroscopy: Direct Observation of the Protonated Schiff Base Linkage to Pyridoxal-5-Phosphate. *J. Am. Chem. Soc.* **136,** 12824–12827. ISSN: 0002-7863. https://pubs.acs.org/doi/10.1021/ja506267d (Sept. 2014).

151. Ngo, H. *et al.* Allosteric Regulation of Substrate Channeling in Tryptophan Synthase: Modulation of the l-Serine Reaction in Stage I of the β-Reaction by α-Site Ligands†,‡. *Biochemistry* **46,** 7740–7753. https://pubs.acs.org/doi/abs/10.1021/bi7003872 (July 2007).

152. Ngo, H. *et al.* Synthesis and Characterization of Allosteric Probes of Substrate Channeling in the Tryptophan Synthase Bienzyme Complex†,‡. *Biochemistry* **46,** 7713–7727. https://pubs.acs.org/doi/abs/10.1021/bi700385f (July 2007).

153. Dunn, M. F. Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Arch. Biochem. Biophys.* **519,** 154–166. ISSN: 0003-9861 (Mar. 2012).

154. Dunn, M. F., Niks, D., Ngo, H., Barends, T. R. & Schlichting, I. Tryptophan synthase: the workings of a channeling nanomachine. *Trends Biochem. Sci.* **33,** 254–264. ISSN: 0968-0004 (June 2008).

155. Metzler, D. E. Equilibria between Pyridoxal and Amino Acids and their Imines. *J. Am. Chem. Soc.* **79,** 485–490. ISSN: 15205126. https://pubs.acs.org/doi/abs/10.1021/ja01559a068 (1957).

156. Cordes, E. H. & Jencks, W. P. Semicarbazone Formation from Pyridoxal, Pyridoxal Phosphate, and Their Schiff Bases. *Biochemistry* **1,** 773–778. ISSN: 15204995. https://pubs.acs.org/sharingguidelines (1962).

157. Heinert, D. & Martell, A. E. Pyridoxine and Pyridoxal Analogs. VI. Electronic Absorption Spectra of Schiff Bases. *J. Am. Chem. Soc.* **85,** 183–188. ISSN: 15205126. https://pubs.acs.org/sharingguidelines (Jan. 1963).

158. Lindskog, S. Structure and mechanism of carbonic anhydrase. *Pharmacol. Ther.* **74,** 1–20. ISSN: 0163-7258 (Jan. 1997).

159. Palmberg, P. A Topical Carbonic Anhydrase Inhibitor Finally Arrives. *Arch Ophthalmol.* **113,** 985–986. ISSN: 0003-9950. https://jamanetwork.com/journals/jamaophthalmology/fullarticle/641278 (Aug. 1995).

160. Maren, T. H. The Development of Topical Carbonic Anhydrase Inhibitors. *J. Glaucoma* **4,** 49–62. ISSN: 1057-0829. https://journals.lww.com/glaucomajournal/Fulltext/1995/02000/The%7B%5C_%7DDevelopment%7B%5C_%7Dof%7B%5C_%7DTopical%7B%5C_%7DCarbonic%7B%5C_%7DAnhydrase.12.aspx (1995).

161. Liljas, A. *et al.* Crystal Structure of Human Carbonic Anhydrase C. *Nat. New Biol.* **235,** 131–137. ISSN: 2058-1092. https://www.nature.com/articles/newbio235131a0 (Feb. 1972).

162. Kim, C. U. *et al.* Tracking solvent and protein movement during CO2 release in carbonic anhydrase II crystals. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 5257–5262. ISSN: 0027-8424. https://www.pnas.org/content/113/19/5257%20https://www.pnas.org/content/113/19/5257.abstract (May 2016).

163. Niimura, N., Takimoto-Kamimura, M. & Tanaka, I. Neutron Diffraction in Studies of Protein Dynamics and Functions, Application of. *Encycl. Anal. Chem.,* 1–30. https://onlinelibrary.wiley.com/doi/full/10.1002/9780470027318.a9550%20https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470027318.a9550%20https://onlinelibrary.wiley.com/doi/10.1002/9780470027318.a9550 (June 2016).

164. Fisher, S. Z., Aggarwal, M., Kovalevsky, A. Y., Silverman, D. N. & McKenna, R. Neutron Diffraction of Acetazolamide-Bound Human Carbonic Anhydrase II Reveals Atomic Details of Drug Binding. *J. Am. Chem. Soc.* **134,** 14726–14729. https://pubs.acs.org/doi/full/10.1021/ja3068098 (Sept. 2012).

165. Fisher, Z. *et al.* Neutron Structure of Human Carbonic Anhydrase II: A Hydrogen-Bonded Water Network "Switch" Is Observed between pH 7.8 and 10.0. *Biochemistry* **50,** 9421–9423. https://pubs.acs.org/doi/full/10.1021/bi201487b (Nov. 2011).

166. Toba, S., Colombo, G. & Merz, Kenneth M., J. Solvent Dynamics and Mechanism of Proton Transfer in Human Carbonic Anhydrase II. *J. Am. Chem. Soc.* **121,** 2290–2302. https://pubs.acs.org/doi/full/10.1021/ja983579y (Mar. 1999).

167. Taraphder, S., Maupin, C. M., Swanson, J. M. J. & Voth, G. A. Coupling Protein Dynamics with Proton Transport in Human Carbonic Anhydrase II. *J. Phys. Chem. B* **120,** 8389–8404. https://pubs.acs.org/doi/full/10.1021/acs.jpcb.6b02166 (Aug. 2016).

168. Singh, H. *et al.* The Active Site of a Prototypical "Rigid" Drug Target is Marked by Extensive Conformational Dynamics. *Angew. Chem., Int. Ed.* **59,** 22916–22921. ISSN: 1521-3773. https://www.onlinelibrary.wiley.com/doi/full/10.1002/anie.202009348%20https://www.onlinelibrary.wiley.com/doi/abs/10.1002/anie.202009348%20https://onlinelibrary.wiley.com/doi/10.1002/anie.202009348 (Dec. 2020).

169. Singh, H. *et al.* Fast Microsecond Dynamics of the Protein–Water Network in the Active Site of Human Carbonic Anhydrase II Studied by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **141,** 19276–19288. ISSN: 0002-7863. https://pubs.acs.org/doi/10.1021/jacs.9b05311 (Dec. 2019).

170. Grzybowski, B. A. *et al.* Combinatorial computational method gives new picomolar ligands for a known enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 1270–1273. ISSN: 0027-8424. https://www.pnas.org/content/99/3/1270%20https://www.pnas.org/content/99/3/1270.abstract (Feb. 2002).

171. *APSY User Manual* V3 (2018).

172. Marion, D. & Wüthrich, K. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1H-1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113,** 967–974. ISSN: 0006-291X (June 1983).

173. Marion, D., Ikura, M., Tschudin, R. & Bax, A. Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **85,** 393–399. ISSN: 0022-2364 (Nov. 1989).

174. Van Rossum, G. & Drake, F. L. *Python 3 Reference Manual* ISBN: 1441412697 (CreateSpace, Scotts Valley, CA, 2009).

175. Hiller, S. & Wider, G. Automated Projection Spectroscopy and Its Applications. *Top. Curr. Chem.* **316,** 21–48 (2012).

176. Kupce, E., Nishida, T. & Freeman, R. Hadmard NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **42,** 95–122 (2003).

177. Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H. & Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *J. Biomol. NMR* **52,** 315–327. ISSN: 09252738 (2012).

178. East, K. W., Delaglio, F. & Lisi, G. P. A simple approach for reconstruction of non-uniformly sampled pseudo-3D NMR data for accurate measurement of spin relaxation parameters. *J. Biomol. NMR* **75,** 213–219. https://doi.org/10.1007/s10858-021-00369-7 (2021).

179. Ying, J., Delaglio, F., Torchia, D. A. & Bax, A. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **68,** 101–118 (2017).

180. Tugarinov, V., Muhandiram, R., Ayed, A. & Kay, L. E. Four-Dimensional NMR Spectroscopy of a 723-Residue Protein: Chemical Shift Assignments and Secondary Structure of Malate Synthase G. *J. Am. Chem. Soc.* **124,** 10025–10035. https://pubs.acs.org/doi/abs/10.1021/ja0205636 (Aug. 2002).

181. Tugarinov, V. & Kay, L. E. Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *J. Am. Chem. Soc.* **125,** 13868–13878. https://pubs.acs.org/doi/abs/10.1021/ja030345s (Nov. 2003).

182. Grudziaz, K., Zawadzka-Kazimierczuk, A. & Koźmiński, W. High-dimensional NMR methods for intrinsically disordered proteins studies. *Methods* **148,** 81–87. ISSN: 10959130. https://doi.org/10.1016/j.ymeth.2018.04.031 (2018).

183. Fricke, P. *et al.* Backbone assignment of perdeuterated proteins by solid-state NMR using proton detection and ultrafast magic-Angle spinning. *Nat. Protoc.* **12.** ISSN: 17502799 (2017).

184. Xiang, S., Chevelkov, V., Becker, S. & Lange, A. Towards Automatic Protein Backbone Assignment Using Proton-Detected 4D Solid-State NMR Data. *J. Biomol. NMR* **60,** 85–90 (2014).

185. King, J. L. & Jukes, T. H. Non-Darwinian Evolution. *Science (80-.).* **164,** 788–798. ISSN: 0036-8075. https://science.sciencemag.org/content/164/3881/788%20https://science.sciencemag.org/content/164/3881/788.abstract (May 1969).

186. Schmidt, E. & Güntert, P. A new algorithm for reliable and general NMR resonance assignment. *J. Am. Chem. Soc.* **134,** 12817–12829. ISSN: 00027863. https://doi.org/10.1021/ja305091n%20http://pubs.acs.org/doi/10.1021/ja305091n (Aug. 2012).

187. Evenäs, J. *et al.* Ligand-induced structural changes to maltodextrin-binding protein as studied by solution NMR spectroscopy. *J. Mol. Biol.* **309,** 961–974. ISSN: 00222836. http://www.idealibrary.com (June 2001).

188. Novacek, J. *et al.* 5D 13C-Detected experiments for backbone assignment of unstructured proteins with a very low signal dispersion. *J. Biomol. NMR* **50,** 1–11. ISSN: 09252738 (May 2011).

189. Zawadzka-Kazimierczuk, A., Koźmiński, W., Šanderová, H. & Krásný, L. High dimensional and high resolution pulse sequences for backbone resonance assignment of intrinsically disordered proteins. *J. Biomol. NMR* **52,** 329–337. ISSN: 09252738 (Apr. 2012).

190. Zerko, S. *et al.* Five and four dimensional experiments for robust backbone resonance assignment of large intrinsically disordered proteins: application to Tau3x protein. *J. Biomol. NMR* **65,** 193–203. ISSN: 15735001 (Aug. 2016).

191. Zhou, D. H. & Rienstra, C. M. High-performance solvent suppression for proton detected solid-state NMR. *J. Magn. Reson.* **192,** 167–172. ISSN: 10907807 (May 2008).

192. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31,** 1325–1327. ISSN: 1367-4803. https://doi.org/10.1093/bioinformatics/btu830 (Dec. 2014).

193. Skinner, S. P. *et al.* CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR* **66,** 111–124. ISSN: 1573-5001. https://link.springer.com/article/10.1007/s10858-016-0060-y (Sept. 2016).

194. Barbet-Massin, E. *et al.* Out-and-back 13C-13C scalar transfers in protein resonance assignment by proton-detected solid-state NMR under ultra-fast MAS. *J. Biomol. NMR* **56,** 379–386. ISSN: 09252738 (2013).

195. Klein, A., Vasa, S. K., Söldner, B., Grohe, K. & Linser, R. Unambiguous Side-Chain Assignments for Solid-State NMR Structure Elucidation of Nondeuterated Proteins via a Combined 5D/4D Side-Chain-to-Backbone Experiment. *J. Phys. Chem. Lett.* **13,** 1644–1651. ISSN: 1948-7185. https://dx.doi.org/10.1021/acs.jpclett.1c04075 (2022).

196. Bonaccorsi, M. *et al.* Multimodal Response to Copper Binding in Superoxide Dismutase Dynamics. *J. Am. Chem. Soc.* **142,** 19660–19667. ISSN: 15205126. https://dx.doi.org/10.1021/jacs.0c09242 (Nov. 2020).

197. Bennett, A. E., Griffin, R. G., Ok, J. H. & Vega, S. Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96,** 8624. ISSN: 0021-9606. https://aip.scitation.org/doi/abs/10.1063/1.462267 (Aug. 1992).

198. Shaka, A. J., Lee, C. J. & Pines, A. Iterative schemes for bilinear operators; application to spin decoupling. *J. Magn. Reson.* **77,** 274–293. ISSN: 0022-2364 (Apr. 1988).

199. Gallo, A., Franks, W. T. & Lewandowski, J. R. A suite of solid-state NMR experiments to utilize orphaned magnetization for assignment of proteins using parallel high and low gamma detection. *J. Magn. Reson.* ISSN: 10960856 (2019).

200. Sørensen, O. W. Aspects and prospects of multidimensional time-domain spectroscopy. *J. Magn. Reson.* **89,** 210–216. ISSN: 0022-2364 (Aug. 1990).

201. Farmer, B. T. Simultaneous [13C,15N]-HMQC, a pseudo-triple-resonance experiment. *J. Magn. Reson.* **93,** 635–641. ISSN: 0022-2364 (July 1991).

202. Parella, T. & Nolis, P. Time-shared NMR experiments. *CONCEPT MAGN Reson A* **36A,** 1–23. ISSN: 1552-5023. https://onlinelibrary.wiley.com/doi/full/10.1002/cmr.a.20150%20https://onlinelibrary.wiley.com/doi/abs/10.1002/cmr.a.20150%20https://onlinelibrary.wiley.com/doi/10.1002/cmr.a.20150 (Jan. 2010).

203. Frueh, D. P., Vosburg, D. A., Walsh, C. T. & Wagner, G. Determination of all NOes in 1H–13C–Me-ILV-U2H–15N Proteins with Two Time-Shared Experiments. *J. Biomol. NMR* **34,** 31–40. ISSN: 1573-5001. https://link.springer.com/article/10.1007/s10858-005-5338-4 (Jan. 2006).

204. Frueh, D. P. *et al.* Time-shared HSQC-NOESY for accurate distance constraints measured at high-field in 15N-13C-ILV methyl labeled proteins. *J. Biomol. NMR* **45,** 311–318. ISSN: 09252738 (2009).

205. in. *TopSpin User Man.* V 013, 83–94 (Bruker Corporation, 2021).

206. Lewandowski, J. R. *et al.* Measurement of site-specific 13C spin-lattice relaxation in a crystalline protein. *J. Am. Chem. Soc.* **132,** 8252–8254. ISSN: 15205126 (June 2010).

207. Helmus, J. J. & Jaroniec, C. P. Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *J. Biomol. NMR* **55,** 355–367. ISSN: 1573-5001. https://link.springer.com/article/10.1007/s10858-013-9718-x (Mar. 2013).

208. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585,** 357–362. ISSN: 1476-4687. https://www.nature.com/articles/s41586-020-2649-2 (Sept. 2020).

209. Güntert, P. & Buchner, L. Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR* **62,** 453–471. ISSN: 1573-5001. https://link.springer.com/article/10.1007/s10858-015-9924-9 (Mar. 2015).

210. Wi, S. & Frydman, L. An Efficient, Robust New Scheme for Establishing Broadband Homonuclear Correlations in Biomolecular Solid State NMR. *ChemPhysChem* **21,** 284–294. ISSN: 14397641. https://onlinelibrary.wiley.com/doi/abs/10.1002/cphc.201901071 (Feb. 2020).

211. Holmes, J. B. *et al.* Imaging active site chemistry and protonation states: NMR crystallography of the tryptophan synthase α-aminoacrylate intermediate. *bioRxiv,* 2021.05.12.443852. http://biorxiv.org/content/early/2021/05/18/2021.05.12.443852.abstract (Jan. 2021).

212. Schmidt, E. *et al.* Automated solid-state NMR resonance assignment of protein microcrystals and amyloids. *J. Biomol. NMR* **56,** 243–254 (2013).

213. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59,** 687–696. ISSN: 08873585. http://doi.wiley.com/10.1002/prot.20449 (Apr. 2005).

214. Shen, Y. & Bax, A. SPARTA+: A modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J. Biomol. NMR* **48,** 13–22. ISSN: 09252738. http://www.ncbi.nlm.nih.gov/ (Sept. 2010).

215. Sakhrani, V. V. *et al.* Backbone assignments and conformational dynamics in the S. typhimurium tryptophan synthase α-subunit from solution-state NMR. *J. Biomol. NMR* **74,** 341–354. ISSN: 15735001. https://doi.org/10.1007/s10858-020-00320-2 (July 2020).

216. Wishart, D. S., Sykes, B. D. & Richards, F. M. The Chemical Shift Index : A Fast and Simple Method for the Assignment of Protein Secondary Structure Through NMR Spectroscopy. *Biochemistry* **31,** 1647–1651. ISSN: 15204995. https://pubs.acs.org/sharingguidelines (Feb. 1992).

217. Wishart, D. S. & Sykes, B. D. The 13C Chemical-Shift Index: A simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR* **4,** 171–180. ISSN: 09252738. https://link.springer.com/article/10.1007/BF00175245 (Mar. 1994).

218. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56,** 227–241. ISSN: 09252738. http://link.springer.com/10.1007/s10858-013-9741-y (July 2013).

219. Bahrami, A., Assadi, A. H., Markley, J. L. & Eghbalnia, H. R. Probabilistic Interaction Network of Evidence Algorithm and its Application to Complete Labeling of Peak Lists from Protein NMR Spectroscopy. *PLoS Comput. Biol.* **5** (ed Cowburn, D.) e1000307. ISSN: 1553-7358. https://dx.plos.org/10.1371/journal.pcbi.1000307 (Mar. 2009).

220. Lee, W. *et al.* I-PINE web server: an integrative probabilistic NMR assignment system for proteins. *J. Biomol. NMR* **73,** 213–222. ISSN: 15735001. https://doi.org/10.1007/s10858-019-00255-3 (May 2019).

221. Laskay, Ü. A., Lobas, A. A., Srzentić, K., Gorshkov, M. V. & Tsybin, Y. O. Proteome digestion specificity analysis for rational design of extended bottom-up and middle-down proteomics experiments. *J. Proteome Res.* **12,** 5558–5569. ISSN: 15353893. https://pubs.acs.org/sharingguidelines (Dec. 2013).

222. Kulik, V. *et al.* On the Role of αThr183 in the Allosteric Regulation and Catalytic Mechanism of Tryptophan Synthase. *J. Mol. Biol.* **324,** 677–690. ISSN: 0022-2836 (Dec. 2002).

223. *The PyMOL Molecular Graphics System*

224. Li, S. & Hong, M. Protonation, Tautomerization, and Rotameric Structure of Histidine: A Comprehensive Study by Magic-Angle-Spinning Solid-State NMR. *J. Am. Chem. Soc* **133,** 1534–1544. https://pubs.acs.org/sharingguidelines (2011).

225. Mueller, L. J. & Dunn, M. F. NMR crystallography of enzyme active sites: Probing chemically detailed, three-dimensional structure in tryptophan synthase. *Acc. Chem. Res.* **46,** 2008–2017. ISSN: 00014842 (Sept. 2013).

226. Bertarello, A. *et al.* Picometer resolution structure of the coordination sphere in the metal-binding site in a metalloprotein by NMR. *J. Am. Chem. Soc.* **142,** 16757–16765. ISSN: 15205126. https://dx.doi.org/10.1021/jacs.0c07339 (Sept. 2020).

227. Bonhomme, C. *et al. First-principles calculation of NMR parameters using the gauge including projector augmented wave method: A chemists point of view* Nov. 2012. https://pubs.acs.org/sharingguidelines.

228. Bryce, D. L. *NMR crystallography: Structure and properties of materials from solid-state nuclear magnetic resonance observables* May 2017. //scripts.iucr.org/cgi-bin/paper?lt5004.

229. Baias, M. *et al.* De novo determination of the crystal structure of a large drug molecule by crystal structure prediction-based powder NMR crystallography. *J. Am. Chem. Soc.* **135,** 17501–17507. ISSN: 00027863. https://pubs.acs.org/sharingguidelines (Nov. 2013).

230. Huang, Y.-m. M. *et al.* Protonation states and catalysis: Molecular dynamics studies of intermediates in tryptophan synthase. *Protein Sci.* **25,** 166–183. ISSN: 09618368. http://doi.wiley.com/10.1002/pro.2709 (Jan. 2016).

231. Lai, J. *et al.* X-ray and NMR crystallography in an enzyme active site: The indoline quinonoid intermediate in tryptophan synthase. *J. Am. Chem. Soc.* **133,** 4–7. ISSN: 00027863. https://pubs.acs.org/doi/abs/10.1021/ja106555c (Jan. 2011).

232. Frueh, D. P. *Practical aspects of NMR signal assignment in larger and challenging proteins* 2014. http://dx.doi.org/10.1016/j.pnmrs.2013.12.001.

233. Weikum, E. R., Liu, X. & Ortlund, E. A. The nuclear receptor superfamily: A structural perspective. *Protein Sci.* **27,** 1876–1892. ISSN: 09618368. http://doi.wiley.com/10.1002/pro.3496 (Nov. 2018).

234. Fernandez, E. J. *Allosteric pathways in nuclear receptors — Potential targets for drug design* Mar. 2018.

235. Tonks, N. K. *Protein tyrosine phosphatases: From genes, to function, to disease* Nov. 2006. www.nature.com/reviews/molcellbio.

236. Patel, R. N. *Biocatalysis in the pharmaceutical and biotechnology industries* 1–847. ISBN: 9781420019377. https://www.taylorfrancis.com/books/biocatalysis-pharmaceutical-biotechnology-industries-ramesh-patel/10.1201/9781420019377 (CRC Press, Jan. 2006).

237. Hassan, A. *et al.* Sensitivity boosts by the CPMAS CryoProbe for challenging biological assemblies. *J. Magn. Reson.* **311,** 106680. ISSN: 1090-7807 (Feb. 2020).

238. Harner, M. J., Frank, A. O. & Fesik, S. W. Fragment-based drug discovery using NMR spectroscopy. *J. Biomol. NMR* **56,** 65–75. ISSN: 1573-5001. https://link.springer.com/article/10.1007/s10858-013-9740-z (May 2013).

239. Orts, J. & Gossert, A. D. Structure determination of protein-ligand complexes by NMR in solution. *Methods* **138-139,** 3–25. ISSN: 10959130. https://doi.org/10.1016/j.ymeth.2018.01.019 (2018).

240. Dias, D. M. & Ciulli, A. NMR approaches in structure-based lead discovery: Recent developments and new frontiers for targeting multi-protein complexes. *Prog. Biophys. Mol. Biol.* **116,** 101–112. ISSN: 00796107. http://dx.doi.org/10.1016/j.pbiomolbio.2014.08.012 (2014).

241. Dalvit, C. NMR methods in fragment screening: theory and a comparison with other biophysical techniques. *Drug Discov. Today* **14,** 1051–1057. ISSN: 1359-6446 (Nov. 2009).

242. Mayer, M. & Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem., Int. Ed.* **38,** 1784–1788 (1999).

243. Dalvit, C., Gossert, A. D., Coutant, J. & Piotto, M. Rapid acquisition of 1H and 19F NMR experiments for direct and competition ligand-based screening. *Magn. Reson. Chem.* **49,** 199–202. ISSN: 07491581 (2011).

244. Dalvit, C. Ligand- and substrate-based 19F NMR screening: Principles and applications to drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **51,** 243–271. ISSN: 00796565 (2007).

245. Dalvit, C. *et al.* A General NMR Method for Rapid, Efficient, and Reliable Biochemical Screening. *J. Am. Chem. Soc.* **125,** 14620–14625. ISSN: 00027863 (2003).

246. Jordan, J. B. *et al.* Fragment based drug discovery: Practical implementation based on 19F NMR spectroscopy. *J. Med. Chem.* **55,** 678–687. ISSN: 00222623 (2012).

247. Stockman, B. J. & Dalvit, C. NMR screening techniques in drug discovery and drug design. *Prog. Nucl. Magn. Reson. Spectrosc.* **41,** 187–231. ISSN: 00796565 (2002).

248. Pellecchia, M. *et al.* Perspectives on NMR in drug discovery: a technique comes of age. *Nat. Rev. Drug Discov.* **7,** 738–745. ISSN: 1474-1784. https://www.nature.com/articles/nrd2606 (2008).

249. Clore, G. & Gronenborn, A. Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science (80-.).* **252,** 1390–1399. ISSN: 0036-8075. https://science.sciencemag.org/content/252/5011/1390%20https://science.sciencemag.org/content/252/5011/1390.abstract (June 1991).

250. Orts, J. *et al.* NMR-Based Determination of the 3D Structure of the Ligand-Protein Interaction Site without Protein Resonance Assignment. *J. Am. Chem. Soc.* **138,** 4393–4400. ISSN: 15205126 (2016).

251. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Formankay, J. D. & Kay, L. E. Simultaneous Acquisition of 15N- and 13C-Edited NOE Spectra of Proteins Dissolved in H2O. *J. Magn. Reson. Ser. B* **103,** 197–201. ISSN: 1064-1866 (Feb. 1994).

252. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4,** 845–858. ISSN: 1573-5001. https://link.springer.com/article/10.1007/BF00398413 (Nov. 1994).

253. Zwahlen, C. *et al.* Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/boxB RNA Complex. *J. Am. Chem. Soc.* **119,** 6711–6721. https://pubs.acs.org/doi/full/10.1021/ja970224q (July 1997).

254. Breeze, A. L. Isotope-filtered NMR methods for the study of biomolecularstructure and interactions. *Prog. Nucl. Magn. Reson. Spectrosc.* **36,** 323–372. www.elsevier.nl/locate/pnmrs (2000).

255. Gossert, A. D. *The Combined Ligand-TPPI NOESY: Higher Quality Structures of Protein-Ligand Complexes using just one NOESY Experiment* 2015. https://www.bruker.com/protected/en/services/bruker-user-library/liquids.html?q0=Gossert.

256. Piotto, M., Saudek, V. & Sklen, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2,** 661–665 (1992).

257. Keeler, J. & Neuhaus, D. Comparison and evaluation of methods for two-dimensional NMR spectra with absorption-mode lineshapes. *J. Magn. Reson.* **63,** 454–472. ISSN: 0022-2364 (July 1985).

258. Hwang, T. L. & Shaka, A. J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson. Ser. A* **112,** 275–279. ISSN: 1064-1858 (Feb. 1995).

259. Grzesiek, S. & Bax, A. The Importance of Not Saturating H2O in Protein NMR. Application to Sensitivity Enhancement and NOE Measurements. *J. Am. Chem. Soc.* **115,** 12593–12594. ISSN: 15205126 (1993).

260. Yang, L. H., Ahmed, S. A. & Miles, E. W. PCR mutagenesis and overexpression of tryptophan synthase from Salmonella typhimurium: On the roles of β2 subunit Lys-382. *Protein Expr. Purif.* **8,** 126–136. ISSN: 10465928 (Aug. 1996).

261. Tian, Y. *et al.* J-Based 3D sidechain correlation in solid-state proteins. *Phys.Chem.Chem.Phys* **11,** 7078–7086. ISSN: 14639076 (Aug. 2009).

262. Milest, E. W. *et al.* The β Subunit of Tryptophan Synthase - Clarification of the Roles of Histidine 86, Lysine 87, Arginine 148, Cysteine 170, and Cysteine 230*. *J. Biol. Chem.* **264,** 6280–6287 (1989).

263. Lee, W., Rahimi, M., Lee, Y. & Chiu, A. POKY: a software suite for multidimensional NMR and 3D structure calculation of biomolecules. *Bioinformatics.* `https://academic.oup.com/bioinformatics/advance-article/doi/10.1093/bioinformatics/btab180/6170655` (Mar. 2021).

264. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods* **16,** 595–602. ISSN: 15487105. `https://doi.org/10.1038/s41592-019-0459-y` (July 2019).

List of Figures

List of Tables

7.1 NMR pulse programs

HNcoCANH dual version

```
;HNcoCANH 5D for INEPT and BSH-CP (use zg options)
;H-1 (t2) to N-1 (t3) to CO to CA (t1) N (t4) to H (t5/acqu) transfer;
; F2 channel is Nitrogen
; F3 channel is carbon

;=======================
; Variables introduction
;=======================

;p1   H p90
;p2   N p90
;p3   C p90
;p6   C p180 hard pulse
;p13  C selective p180 duration
;p15  HN first CP contact time
;p16  NH CP contact time
;p17 CA N CP contact time
;p18 N CO CP contact time
;p22  water suprresion pulse length
;p28  Trim pulse on Carbonyl
;p29  CA-CO BSH CP time



;pl1  H p90 power
;pl2  N p90 power
;pl3  C hard power level
;pl5  N- NH CP power
;pl6  H- NH CP power
;pl7  N NCA-CP power
;pl8  C NCA-CP power
;pl12 H dec power
;pl13 C selective 180 power
;pl16 water suppression pulse power
;pl22 WALTZ decoupling on N (during acquisition)
;pl27 Ca H-CO CP power
;pl28 C N-CO CP power
;pl29 CA-CO BSH CP power level

;l28  (=1) ramp on H (=2) ramp on N
;l29  (=1) ramp on H (=2) ramp on CO
;l30  (=2) ramp on N (=3) ramp on CA
```

```
;l31  (=1) ramp on H (=2) ramp on N

;d0   t1 initial value
;d7 COCA J evol
;d8 COCA J evol

;cnst8 COCA J coupling
;cnst7 COCA J coupling
;cnst10 1H carrier for 1st CP
;cnst19 - offset for water suppression
;cnst30 CO offset (174 ppm)
;cnst31 C decoupling offset (at 110 ppm)
;cnst32 Ca offset (55 ppm)

;============================
; Include file for Protection
;============================

#include <Avance.incl>
#include <Delay.incl>

;==============
; Set variables
;==============

;"cnst63 = plw12"
;"cnst62 = plw22"
;"cnst61 = plw23"
;"cnst60 = plw16"
"p6 = 2.0*p3"
"p5 = 2.0*p2"
"spw2=spw8"
"d11=30m"
"d0=0.2u"
"d10=0.2u"
"d20=0.2u"
"d30=0.2u"

#ifdef NNH

"in0=inf1/2"
"in10=inf2/2"

#endif /* NNH */
```

```
#ifdef HNNH

"in20=inf1/2"
"in0=inf2/2"
"in10=inf3/2"

#endif /* HNNH */

#ifdef HNCANH

"in30=inf1/2"
"in20=inf2/2"
"in0=inf3/2"
"in10=inf4/2"

#endif

#ifdef INEPT

"d8=1/(4*cnst8)"
"d7 = 1/(4*cnst7)"
"spw14 = spw13"
"spw15 = spw13"

#endif /*INEPT*/
;============================
; Protection for parameters
;============================
1m
   if "p1 > 100u"     goto Problem
   if "p2 > 100u"     goto Problem
   if "p3 > 100u"     goto Problem
   if "p28 >10u"     goto Problem
   if "p28 >1000u"    goto Problem
   if "p15> 3000u"    goto Problem
   if "p16> 2000u"    goto Problem
   if "p17> 11001u"    goto Problem
   if "p18> 11001u"    goto Problem
   if "p29> 7000u"    goto Problem
   if "p22 >20001u"   goto Problem
   if "aq > 55m"      goto Problem
   if "d1 < 0.49s"     goto Problem
```

```
   goto PassParams
Problem, 1m
   print "Protection: Parameters not accepted, ending."
   goto HaltAcqu
PassParams, 1m

;----- define acquisition order -----

#ifdef NNH

  aqseq 321

#endif /* NNH */

;----- Relaxation & reset parameters -----

1 ze

2 d11
  40u do:f3
  40u do:f2
  40u do:f1

3 d1
  20u reset:f1     reset:f2     reset:f3
  20u fq=cnst10:f1 fq=cnst20:f2 fq=0:f3
  5u  pl1:f1       pl2:f2       pl3:f3

;----- 90 on H -----

  (p1 pl1 ph1):f1

;----- H (t1) evolution & N decoupling  -----

#ifdef HNNH

  1u pl22:f2 pl23:f3
  1u cpd2:f2 cpd3:f3
  d20*2.0
  2u do:f2 do:f3

#endif /* HNNH */

#ifdef HNCANH
```

```
  1u pl22:f2 pl23:f3
  1u cpd2:f2 cpd3:f3
  d20*2.0
  2u do:f2 do:f3

#endif /* HNNH */


;----- H/N CP -----

if (l28 == 1)
  {
  (p15 pl5 ph13):f2  (p15:sp8 ph14):f1
  }

if (l28 == 2)
  {
  (p15:sp9 ph13):f2  (p15 pl6 ph14):f1
  }

;----- N (t2) evolution & C decoupling  -----

  1u  pl12:f1  pl3:f3
  2u  fq=cnst31(bf ppm):f3
  1u  cpd1:f1
  d0
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d0
  2u fq=cnst30(bf ppm):f3
  2u do:f1

;----- N/CO CP -----

if (l29 == 2)
  {
  (p18 pl28 ph11):f3 (p18:sp6 ph10):f2
  }

if (l29 == 3)
  {
  (p18:sp7 ph11):f3 (p18 pl27 ph10):f2
```

```
  }

#ifdef BASH

;----- CO to CA BSHCP transfer  -----

  1u pl12:f1
  1u cpd1:f1
  (p28 pl3 ph28):f3
  2u fq=cnst32(bf ppm):f3
  (p29:sp29 ph29):f3

#endif /*BASH*/

#ifdef INEPT

        ;----- Selective pulse on CA  -----

  1u pl12:f1
  1u cpd1:f1
  (center (p13:sp13 ph20):f3 (p5 pl2 ph20):f2) ;shape pulse on CA

;----- Refocused INEPT transfer from CO to CA  -----

  d8
  (p13:sp15 ph20):f3            ;shape pulse on CO
  (p13:sp13 ph20):f3      ;shape pulse on CA
  d8
  2u fq=cnst31(bf ppm):f3 ;place offset at center CA CO

  (p3 pl3 ph4):f3
  d7
  (p6 pl3 ph16):f3
  d7
  2u fq=cnst32(bf ppm):f3

#endif /*INEPT*/

#ifdef HNCANH

;----- CA (t4) evolution & HC decoupling  -----

  d30
  (center (p13:sp14 ph20):f3 (p5 pl2 ph20):f2)
```

```
  d30
  2u
;----- Bloch - Siegert - shift compensation

  (p13:sp15 ph20):f3 ;shape pulse on CO
  (p13:sp14 ph20):f3 ;shape pulse on CA

#endif
        2u do:f1
;----- CA/N CP -----

if (l30 == 2)
  {
  (p17:sp5 ph2):f2 (p17 pl8 ph16):f3
  }

if (l30 == 3)
  {
  (p17 pl7 ph2):f2 (p17:sp4 ph2):f3
  }

;----- N (t3) evolution & C decoupling  -----

  1u  pl12:f1  pl3:f3
  2u  fq=cnst31(bf ppm):f3
  1u  cpd1:f1
  d10
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d10
  3u do:f1

;----- N magnetization along Z  -----

  (p2 pl2 ph17):f2

;----- Water suppression  -----

  20u fq=cnst19:f1
  2u pl16:f1

  p22:f1 ph22
  p22:f1 ph23
```

```
  p22:f1 ph22
  p22:f1 ph23

  22u fq=cnst10:f1

;----- N magnetization in XY plane  -----

  (p2 pl2 ph18):f2

;----- N/H CP -----

if (l31 == 1)
  {
  (p16 pl5 ph12):f2  (p16:sp2 ph2):f1
  }

if (l31 == 2)
  {
  (p16:sp3 ph12):f2  (p16 pl6 ph2):f1
  }

  1u pl22:f2 pl23:f3
  1u cpd2:f2  cpd3:f3
  2u fq=0:f1

  go=2 ph31
  d11 do:f2 do:f3  mc #0 to 2

#ifdef oneD

  F1QF()

#endif /*oneD*/

#ifdef NNH

  F1PH(calph(ph13, +90.0), caldel(d0, +in0))
  F2PH(calph(ph17, -90.0), caldel(d10, +in10))

#endif /* NNH */

#ifdef HNNH

  F1PH(calph(ph1, +90.0), caldel(d20, +in20))
```

```
  F2PH(calph(ph13, +90.0), caldel(d0, +in0))
  F3PH(calph(ph17, -90.0), caldel(d10, +in10))

#endif /* HNNH */

#ifdef HNCANH

  F1PH(calph(ph16, -90.0), caldel(d30, +in30))
  F2PH(calph(ph1, +90.0), caldel(d20, +in20))
  F3PH(calph(ph13, +90.0), caldel(d0, +in0))
  F4PH(calph(ph17, -90.0), caldel(d10, +in10))


#endif

HaltAcqu, 1m
exit

;----- Phase cycling -----

ph1 = 0 2
ph2 = 0
ph4 = 0
ph5 = 0
ph6 = 0
ph10 = 1
ph11 = 0 0 0 0  2 2 2 2
ph12 = 1
ph13 = 1
ph14 = 1
ph16 = 0
ph17 = 1
ph18 = 0 0 2 2
ph20 = 1
ph21 = 2
ph22 = 0
ph23 = 1
ph29 = 0
ph28 = 1
ph31 = 0 2 2 0  2 0 0 2

;=====================
; END Pulse program
;=====================
```

```
;; IMPORTANT NOTE
; On AV (avance) machines, its better to use "spf0" files instead of "sp0"
; The "sp0" files requires lot of time for power settings compared to "spf0"
```

HNcoCANH BSH-CP version

```
;HNcoCANH 5D for BSH-CP (use zg options)
;H-1 (t2) to N-1 (t3) to CO to CA (t1) N (t4) to H (t5/acqu) transfer;
; F2 channel is Nitrogen
; F3 channel is carbon

;=======================
; Variables introduction
;=======================

;p1   H p90
;p2   N p90
;p3   C p90
;p6   C p180 hard pulse
;p13  C selective p180 duration
;p14  soft 180 carbon on res.
;p15  HN first CP contact time
;p16  NH CP contact time
;p17 CA N CP contact time
;p18 N CO CP contact time
;p22  water suprresion pulse length
;p28  Trim pulse on Carbonyl
;p29  CA-CO BSH CP time

;pl1  H p90 power
;pl2  N p90 power
;pl3  C p90 power
;pl5  N- NH CP power
;pl6  H- NH CP power
;pl7  N NCA-CP power
;pl8  C NCA-CP power
;pl12 H dec power
;pl13 C selective 180 power
;pl14 soft rectangular 180 pulse 13C
;pl16 water suppression pulse power
;pl22 WALTZ decoupling on N (during acquisition)
;pl27 Ca H-CO CP power
;pl28 C N-CO CP power
;pl29 CA-CO BSH CP power level

;l28  (=1) ramp on H (=2) ramp on N
;l29  (=1) ramp on H (=2) ramp on CO
;l30  (=2) ramp on N (=3) ramp on CA
;l31  (=1) ramp on H (=2) ramp on N
```

```
;cnst10 1H carrier for 1st CP
;cnst19 - offset for water suppression
;cnst30 CO offset (174 ppm)
;cnst31 C decoupling offset (at 110 ppm)
;cnst32 Ca offset (55 ppm)

;============================
; Include file for Protection
;============================

#include <Avance.incl>
#include <Delay.incl>

;==============
; Set variables
;==============

"p6 = 2.0*p3"
"p5 = 2.0*p2"
"spw2=spw8"
"d11=30m"
"d0=0.2u"
"d10=0.2u"
"d20=0.2u"
"d30=0.2u"


#ifdef NNH

"in0=inf1/2"
"in10=inf2/2"

#endif /* NNH */

#ifdef HNNH

"in20=inf1/2"
"in0=inf2/2"
"in10=inf3/2"

#endif /* HNNH */

#ifdef HNCANH
```

```
"in30=inf1/2"
"in20=inf2/2"
"in0=inf3/2"
"in10=inf4/2"

#endif
;============================
; Protection for parameters
;============================
1m
   if "p1 > 100u"     goto Problem
   if "p2 > 100u"     goto Problem
   if "p3 > 100u"     goto Problem
   if "p28 >10u"     goto Problem
   if "p28 >1000u"    goto Problem
   if "p15> 3000u"    goto Problem
   if "p16> 2000u"    goto Problem
   if "p17> 11001u"    goto Problem
   if "p18> 11001u"    goto Problem
   if "p29> 7000u"    goto Problem
   if "p22 >20001u"   goto Problem
   if "aq > 55m"      goto Problem
   if "d1 < 0.49s"     goto Problem

   goto PassParams
Problem, 1m
   print "Protection: Parameters not accepted, ending."
   goto HaltAcqu
PassParams, 1m

;----- define acquisition order -----

#ifdef NNH

  aqseq 321

#endif /* NNH */

;----- Relaxation & reset parameters -----

1 ze

2 d11
```

```
  40u do:f3
  40u do:f2
  40u do:f1     ;decouplers off

3 d1
  20u reset:f1     reset:f2     reset:f3
  20u fq=cnst10:f1 fq=cnst20:f2 fq=0:f3
  5u  pl1:f1       pl2:f2       pl3:f3

;----- 90 on H -----

  (p1 pl1 ph1):f1

;----- H (t2) evolution & N decoupling  -----

#ifdef HNNH

 1u pl22:f2 pl23:f3
 1u cpd2:f2 cpd3:f3
 d20*2.0
 2u do:f2 do:f3

#endif /* HNNH */

#ifdef HNCANH

 1u pl22:f2 pl23:f3
 1u cpd2:f2 cpd3:f3
 d20*2.0
 2u do:f2 do:f3

#endif /* HNNH */


;----- H/N CP -----

if (128 == 1)
  {
  (p15 pl5 ph13):f2  (p15:sp8 ph14):f1
  }

if (128 == 2)
  {
  (p15:sp9 ph13):f2  (p15 pl6 ph14):f1
```

```
  }

;----- N (t3) evolution & C decoupling  -----

  1u  pl12:f1  pl3:f3
  2u  fq=cnst31(bf ppm):f3
  1u  cpd1:f1
  d0
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d0
  2u fq=cnst30(bf ppm):f3
  2u do:f1

;----- N/CO CP -----

if (l29 == 2)
  {
  (p18 pl28 ph11):f3 (p18:sp6 ph10):f2
  }

if (l29 == 3)
  {
  (p18:sp7 ph11):f3 (p18 pl27 ph10):f2
  }

;----- CO to CA BSHCP transfer  -----

  1u pl12:f1
  1u cpd1:f1
  (p28 pl3 ph28):f3
  2u fq=cnst32(bf ppm):f3
  (p29:sp29 ph29):f3

#ifdef HNCANH

;----- CA (t1) evolution & HC decoupling  -----

  d30
  (center (p13:sp12 ph20):f3 (p5 pl2 ph20):f2) ;shape pulse on CO
  d30

;----- Bloch - Siegert - shift compensation
```

```
  (p14 pl14 ph20):f3
  (center (p13:sp12 ph20):f3 (p5 ph20):f2) ;shape pulse on CO

#endif
        2u do:f1
;----- CA/N CP -----

if (l30 == 2)
  {
  (p17:sp5 ph2):f2 (p17 pl8 ph16):f3
  }

if (l30 == 3)
  {
  (p17 pl7 ph2):f2 (p17:sp4 ph2):f3
  }

;----- N (t4) evolution & C decoupling  -----

  1u  pl12:f1  pl3:f3
  2u  fq=cnst31(bf ppm):f3
  1u  cpd1:f1
  d10
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d10
  3u do:f1

;----- N magnetization along Z  -----

  (p2 pl2 ph17):f2

;----- Water suppression  -----

  p22:f1 ph22
  p22:f1 ph23
        20u fq=cnst19:f1 ; move to beginning of MISSISSIPPI when only water signal
        2u pl16:f1
  p22:f1 ph22
  p22:f1 ph23

 20u fq=0:f1
```

```
 20u fq=cnst10:f1

;----- N magnetization in XY plane  -----

 (p2 pl2 ph18):f2

;----- N/H CP -----

if (l31 == 1)
  {
  (p16 pl5 ph12):f2  (p16:sp2 ph2):f1
  }

if (l31 == 2)
  {
  (p16:sp3 ph12):f2  (p16 pl6 ph2):f1
  }

  1u pl22:f2 pl23:f3
  1u cpd2:f2  cpd3:f3
  2u fq=0:f1

  go=2 ph31
  d11 do:f2 do:f3  mc #0 to 2

#ifdef oneD
  F1QF()
#endif /*oneD*/

#ifdef NNH

  F1PH(calph(ph13, +90.0), caldel(d0, +in0))
  F2PH(calph(ph17, -90.0), caldel(d10, +in10))

#endif /* NNH */

#ifdef HNNH

  F1PH(calph(ph1, +90.0), caldel(d20, +in20))
  F2PH(calph(ph13, +90.0), caldel(d0, +in0))
  F3PH(calph(ph17, -90.0), caldel(d10, +in10))

#endif /* HNNH */
```

```
#ifdef HNCANH

  F1PH(calph(ph16, -90.0), caldel(d30, +in30))
  F2PH(calph(ph1, +90.0), caldel(d20, +in20))
  F3PH(calph(ph13, +90.0), caldel(d0, +in0))
  F4PH(calph(ph17, -90.0), caldel(d10, +in10))

#endif


HaltAcqu, 1m
exit

;----- Phase cycling -----

ph1 = 0 2
ph2 = 0
ph5 = 0
ph6 = 0
ph10 = 1
ph11 = 0 0 0 0  2 2 2 2
ph12 = 1
ph13 = 1
ph14 = 1
ph16 = 0
ph17 = 1
ph18 = 0 0 2 2
ph20 = 1
ph21 = 2
ph22 = 0
ph23 = 1
ph29 = 0
ph28 = 1
ph31 = 0 2 2 0  2 0 0 2

;=====================
; END Pulse program
;=====================
;; IMPORTANT NOTE
; On AV (avance) machines, its better to use "spf0" files instead of "sp0"
; The "sp0" files requires lot of time for power settings compared to "spf0"
```

HNcaCONH BSH-CP version

```
;HNcaCONH 5D for BSH-CP (use zg options)
;H-1 (t2) to N-1 (t3) to CA-1 to CO (t1) N (t4) to H (t5/acqu) transfer;
;F2 channel is Nitrogen
;F3 channel is carbon

;=======================
; Variables introduction
;=======================

;p1   H p90
;p2   N p90
;p3   C p90
;p6   C p180 hard pulse
;p13  C selective p180 duration
;p14  soft 180 carbon on res.
;p15  HN first CP contact time
;p16  NH CP contact time
;p17 CO N CP contact time
;p18 N CA CP contact time
;p22  water suprresion pulse length
;p28  Trim pulse on Carbonyl
;p29  CO-CA BSH CP time

;pl1  H p90 power
;pl2  N p90 power
;pl3  C hard power level
;pl5  N- NH CP power
;pl6  H- NH CP power
;pl7  N NCA-CP power
;pl8  C NCA-CP power
;pl12 H dec power
;pl13 C selective 180 power
;pl14 soft rectangular 180 pulse 13C
;pl16 water suppression pulse power
;pl22 WALTZ decoupling on N (during acquisition)
;pl27 CO H-CO CP power
;pl28 C N-CA CP power
;pl29 CO-CA BSH CP power level

;l28  (=1) ramp on H (=2) ramp on N
;l29  (=1) ramp on H (=2) ramp on CA
;l30  (=2) ramp on N (=3) ramp on CO
;l31  (=1) ramp on H (=2) ramp on N
```

```
;cnst10 1H carrier for 1st CP
;cnst19 - offset for water suppression
;cnst30 CA offset (55 ppm)
;cnst31 C decoupling offset (at 110 ppm)
;cnst32 CO offset (174 ppm)


;============================
; Include file for Protection
;============================

#include <Avance.incl>
#include <Delay.incl>

;==============
; Set variables
;==============

"p6 = 2.0*p3"
"p5 = 2.0*p2"
"spw2=spw8"
"d11=30m"
"d0=0.2u"
"d10=0.2u"
"d20=0.2u"
"d30=0.2u"


#ifdef NNH

"in0=inf1/2"
"in10=inf2/2"

#endif /* NNH */

#ifdef HNNH

"in20=inf1/2"
"in0=inf2/2"
"in10=inf3/2"

#endif /* HNNH */
```

```
#ifdef HNCONH

"in30=inf1/2"
"in20=inf2/2"
"in0=inf3/2"
"in10=inf4/2"

#endif
;============================
; Protection for parameters
;============================
1m
   if "p1 > 100u"     goto Problem
   if "p2 > 100u"     goto Problem
   if "p3 > 100u"     goto Problem
   if "p28 >10u"     goto Problem
   if "p28 >1000u"    goto Problem
   if "p15> 3000u"    goto Problem
   if "p16> 2000u"    goto Problem
   if "p17> 11001u"    goto Problem
   if "p18> 11001u"    goto Problem
   if "p29> 7000u"    goto Problem
   if "p22 >20001u"   goto Problem
   if "aq > 55m"      goto Problem
   if "d1 < 0.49s"     goto Problem

   goto PassParams
Problem, 1m
   print "Protection: Parameters not accepted, ending."
   goto HaltAcqu
PassParams, 1m

;----- define acquisition order -----

#ifdef NNH

aqseq 321

#endif /* NNH */

;----- Relaxation & reset parameters -----

1 ze
```

```
2 d11
  40u do:f3
  40u do:f2
  40u do:f1

3 d1
  20u reset:f1     reset:f2     reset:f3
  20u fq=cnst10:f1 fq=cnst20:f2 fq=0:f3
  5u  pl1:f1       pl2:f2       pl27:f3

;----- 90 on H -----

  (p1 pl1 ph1):f1

;----- H (t2) evolution & N decoupling  -----

#ifdef HNNH

 1u pl22:f2
 1u cpd2:f2
 d20*2.0
 2u do:f2

#endif /* HNNH */

#ifdef HNCONH

 1u pl22:f2
 1u cpd2:f2
 d20*2.0
 2u do:f2

#endif /* HNCONH */


;----- H/N CP -----

 0.7u pl5:f2  pl6:f1

if (128 == 1)
  {
 (p15 ph13):f2  (p15:sp8 ph14):f1
  }
```

```
if (l28 == 2)
  {
 (p15:sp9 ph13):f2  (p15 ph14):f1
  }

;----- N (t3) evolution & C decoupling  -----

  1u  pl12:f1  pl3:f3
  2u  fq=cnst31(bf ppm):f3
  1u  cpd1:f1
  d0
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d0
  2u fq=cnst30(bf ppm):f3
  2u do:f1

;----- N/CA CP -----

if (l29 == 2)
  {
  (p18 pl28 ph11):f3 (p18:sp6 ph10):f2
  }

if (l29 == 3)
  {
  (p18:sp7 ph11):f3 (p18 pl27 ph10):f2
  }

;----- CA to CO BSHCP transfer  -----

  (p29:sp29 ph29):f3
  2u fq=cnst32(bf ppm):f3
  (p28 pl3 ph28):f3

#ifdef HNCONH

;----- CO (t4) evolution & HC decoupling  -----

  ;1u  pl2:f2
  d30
  (center (p13:sp12 ph20):f3 (p5 pl2 ph20):f2) ;shape pulse on CA
  d30
```

```
;----- Bloch - Siegert - shift compensation

 (p14 pl14 ph20):f3
 (center (p13:sp12 ph20):f3 (p5 ph20):f2) ;shape pulse on CA

#endif

;----- CO/N CP -----

if (l30 == 2)
  {
  (p17:sp5 ph2):f2 (p17 pl8 ph16):f3
  }

if (l30 == 3)
  {
  (p17 pl7 ph2):f2 (p17:sp4 ph2):f3
  }

;----- N (t4) evolution & C decoupling  -----

  1u  pl12:f1  pl3:f3
  2u  fq=cnst31(bf ppm):f3
  1u  cpd1:f1
  d10
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d10
  3u do:f1

;----- N magnetization along Z  -----

  (p2 pl2 ph17):f2

;----- Water suppression  -----

  20u fq=cnst19:f1
  2u pl16:f1

  p22:f1 ph22
  p22:f1 ph23
  p22:f1 ph22
  p22:f1 ph23
```

```
 20u fq=cnst10:f1
 5u pl6:f1

;----- N magnetization in XY plane  -----

  (p2 pl2 ph18):f2

;----- N/H CP -----

  0.7u pl5:f2

if (l31 == 1)
  {
  (p16 ph12):f2  (p16:sp2 ph2):f1
  }

if (l31 == 2)
  {
  (p16:sp3 ph12):f2  (p16 ph2):f1
  }

  1u pl22:f2
  1u cpd2:f2
  2u fq=0:f1

  go=2 ph31
  d11 do:f2  mc #0 to 2

#ifdef oneD

  F1QF()

#endif /*oneD*/

#ifdef NNH

  F1PH(rp17 & rd10 & ip13, id0)
  F2PH(dp17, id10)

#endif /* NNH */

#ifdef HNNH
```

```
  F1PH(calph(ph1, +90.0), caldel(d20, +in20))
  F2PH(calph(ph13, +90.0), caldel(d0, +in0))
  F3PH(calph(ph17, -90.0), caldel(d10, +in10))

#endif /* HNNH */

#ifdef HNCONH

  F1PH(calph(ph16, -90.0), caldel(d30, +in30))
  F2PH(calph(ph1, +90.0), caldel(d20, +in20))
  F3PH(calph(ph13, +90.0), caldel(d0, +in0))
  F4PH(calph(ph17, -90.0), caldel(d10, +in10))

#endif

HaltAcqu, 1m
exit

;----- Phase cycling -----

ph1 = 0 2
ph2 = 0
ph5 = 0
ph6 = 0
ph10 = 1
ph11 = 0 0 0 0  2 2 2 2
ph12 = 1
ph13 = 1
ph14 = 1
ph16 = 0 0 0 0  0 0 0 0  2 2 2 2  2 2 2 2
ph17 = 1
ph18 = 0 0 2 2
ph20 = 1
ph21 = 2
ph22 = 0
ph23 = 1
ph29 = 0
ph28 = 3
ph31 = 0 2 2 0  2 0 0 2  2 0 0 2  0 2 2 0

;=====================
; END Pulse program
;=====================
;; IMPORTANT NOTE
```

```
; On AV (avance) machines, its better to use "spf0" files instead of "sp0"
; The "sp0" files requires lot of time for power settings compared to "spf0"
```

Combined HCCNH/HCCH

```
;HCCNH_5D.alkl
;H (t1) to Cali (t2) with CC mix (to) Ca (t3) to N (t4) to H (t5) transfer;
; F2 channel is Nitrogen
; F3 channel is carbon

;======================
; Variables introduction
;======================

;p1   H p90
;p2   N p90
;p3   C p90
;p13  C selective p180 duration
;p14 soft 180 carbon on res. (= DW/sqrt(3))
;p15  HCA CP contact time
;p16  NH CP contact time
;p17  CAN CP contact time
;p22  water suppression pulse length

;pl1  H p90 power
;pl2  N p90 power
;pl1  C p90 power
;pl5  N- NH CP power
;pl6  H- NH CP power
;pl7  N NCA-CP power
;pl8  C NCA-CP power
;pl13 C selective 180 power
;pl14 soft rectangular 180 pulse 13C
;pl16 water suppression pulse power
;pl22 WALTZ decoupling on N (during acquisition)
;pl23 (NC) decoupling on C (during acquisition)
;pl25 CC mixing power level
;pl27 Ca H-Cali CP power
;pl28 H H-Cali CP power

;l29  (=1) ramp on H (=2) ramp on CA
;l30  (=2) ramp on Nitrogen; (=3) ramp on carbon
;l31  (=1) ramp on H (=2) ramp on N
;cnst10 1H carrier for NH CP in Hz (on Hamid)
;cnst11 1H carrier for CH CP in Hz (on Hali)
;cnst19 - offset for water suppression
;cnst30 C decoupling offset (at 120 ppm)
;cnst31 13C carrier for Cali in Hz (~ 20 ppm)
```

```
;spoff13 soft 180 offset (= DW)
;d9 TOCYS mixing time (WALTZ= n*960 us@10us, DIPSI=n*2173 us@10us)
;cpd3 C dec during 1H evolution
;cpd5 CC mixing change according, BE CAREFUL WITH POWER

;==============
; Set variables
;==============

"p6 = 2.0*p3"
"p5 = 2.0*p2"
"d11=30m"

"d0=0.2u"         ; evol. time F1 (CC)
"d10=0.2u"         ; evol. time F2 (Ca/ali)
"d20=0.2u"         ; evol. time F3 (N)
"d21=0.2u"         ; evol. time F4 (1H)


"in0=inf1/2"         ; increment delay F1
"in10=inf2/2"         ; increment delay F2
"in20=inf3/2"         ; increment delay F3
"in21=inf4/2"         ; increment delay F4

"l2=1"

;============================
; Include file for Protection
;============================

#include <Avance.incl>
#include <Delay.incl>

;============================
; Protection for parameters
;============================
1m
   if "p1 > 100u"     goto Problem
   if "p2 > 100u"     goto Problem
   if "p3 > 100u"     goto Problem
   if "p13 >1001u"    goto Problem
   if "p15> 4001u"    goto Problem
   if "p16> 2000u"    goto Problem
   if "p17> 11001u"    goto Problem
```

```
   if "p22 >20001u"   goto Problem
   if "aq > 51m"      goto Problem
   if "d1 < 0.49s"     goto Problem
   if "d9 > 0.020s"     goto Problem

   goto PassParams
Problem, 1m
   print "Protection: Parameters not accepted, ending."
   goto HaltAcqu
PassParams, 1m


;----- Relaxation & reset parameters -----

1 ze
2 d11
  40u do:f3
  40u do:f2
  40u do:f1     ;decouplers off

3 d1
  20u reset:f1     reset:f2     reset:f3
  20u fq=cnst11:f1 fq=cnst20:f2 fq=0:f3 ;carrier on Ca
  5u  pl1:f1       pl2:f2       pl27:f3

  1m st0
;----- 90 on H and 1H evolution (t4) -----
  20u pl22:f3
  (p1 pl1 ph5):f1
  d21*2 cpd3:f3
        0.5u do:f3

;----- H/Cali CP ----- ; optimize for one-bond transfer

  1u fq=cnst31:f3  ;carrier on f3 should be in the middle of Cali (35 ppm)

if (l29 == 1)
  {
  (p15 pl27 ph11):f3 (p15:sp6 ph10):f1
  }

if (l29 == 3)
  {
  (p15:sp7 ph11):f3 (p15 pl28 ph10):f1
```

```
  }

;----- CX (t1) evolution & HC decoupling  -----

  0.5u pl12:f1
  d0 cpd1:f1
  (p5 pl2 ph20):f2
  d0
  0.5u
;----- 90 on C  -----

  (p3 pl3 ph6):f3

;----- CC mixing  -----

  1u do:f1
  0.5u pl25:f3
  0.5u cpd5:f3
  d9
  2u do:f3

;----- 90 on C  -----

  (p3 pl3 ph7):f3

;----- C evolution -------

  1u fq=0:f3
  0.5u pl12:f1
  0.5u cpd1:f1
  d10
  (p5 pl2 ph20):f2
  d10
  1.5u

;----- CA/N CP -----

if (l30 == 2)
  {
  (p17:sp5 ph2):f2 (p17 pl8 ph2):f3
  }

if (l30 == 3)
  {
```

```
  (p17 pl7 ph2):f2 (p17:sp4 ph2):f3
  }

;----- 90 on C (store on z for simCP)  -----

  (p3 pl3 ph10):f3

;----- N (t2) evolution & C decoupling  -----

  1u  pl3:f3
  1u  fq=cnst30:f3
  d20
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d20
  2u do:f1

;----- N magnetization along Z  -----

  (p2 pl2 ph17):f2 ;ph17

;----- Water suppression  -----

  20u fq=cnst19:f1
  2u pl16:f1

  p22:f1 ph22
  p22:f1 ph23
  p22:f1 ph22
  p22:f1 ph23

  20u fq=cnst10:f1 ;HN region
  5u pl6:f1

;----- N magnetization in XY plane  -----

  (p2 pl2 ph18):f2

;----- N/H CP -----

if (l31 == 1)
  {
  (p16 pl5 ph12):f2 (p16:sp2 ph2):f1
```

```
  }

if (l31 == 2)
  {
 (p16:sp3 ph12):f2  (p16 ph2):f1
  }

 0.5u pl22:f2
 0.5u cpd2:f2
 goscnp ph31 ;acquire HN
 1u do:f2
; ----- Water suppression ----

  20u fq=cnst19:f1
  2u pl16:f1

  p22:f1 ph22
  p22:f1 ph23
  p22:f1 ph22
  p22:f1 ph23

  20u fq=cnst11:f1 ;aliphatic region
  2u pl6:f1
;----- C/H CP -----

  1m st ; switch NBL
  (p3 pl3 ph13):f3
  (p16 pl27 ph3):f3 (p16:sp2 ph3):f1
  0.5u pl23:f3
  0.5u cpd3:f3
  go=2 ph31
  1u do:f3

  d11 mc #0 to 2

  F1PH(calph(ph11, +90), caldel(d0, +in0))
  F2PH(calph(ph7, +90), caldel(d10, +in10))
  F3PH(calph(ph17, -90), caldel(d20, +in20))
  F4PH(calph(ph5, +90), caldel(d21, +in21))
  F5QF()

HaltAcqu, 1m
exit
```

```
;----- Phase cycling -----

ph2 = 0
ph3 = 2
ph5 = 0 2
ph6 = 3
ph7 = 3 3 3 3  1 1 1 1
ph10 = 1
ph11 = 0 0 2 2
ph12 = 1
ph13 = 3
ph17 = 3
ph18 = 0 0 0 0  0 0 0 0
ph19 = 2 2 2 2  2 2 2 2
ph20 = 1
ph21 = 2
ph22 = 0
ph23 = 1
ph31= 0 2 2 0  2 0 0 2

;=====================
; END Pulse program
;=====================
;; IMPORTANT NOTE
; On AV (avance) machines, its better to use "spf0" files instead of "sp0"
; The "sp0" files requires lot of time for power settings compared to "spf0"
```

HCCNH basic version

```
;HCCNH_5D.alkl
;H (t1) to Cali (t2) with CC mix (to) Ca (t3) to N (t4) to H (t5) transfer;
; F2 channel is Nitrogen
; F3 channel is carbon

;=======================
; Variables introduction
;=======================

;=======================
; Variables introduction
;=======================

;p1   H p90
;p2   N p90
;p3   C p90
;p13  C selective p180 duration
;p14 soft 180 carbon on res. (= DW/sqrt(3))
;p15  HCA CP contact time
;p16  NH CP contact time
;p17  CAN CP contact time
;p22  water suppression pulse length

;pl1  H p90 power
;pl2  N p90 power
;pl1  C p90 power
;pl5  N- NH CP power
;pl6  H- NH CP power
;pl7  N NCA-CP power
;pl8  C NCA-CP power
;pl13 C selective 180 power
;pl14 soft rectangular 180 pulse 13C
;pl16 water suppression pulse power
;pl22 WALTZ decoupling on N (during acquisition)
;pl23 (NC) decoupling on C (during acquisition)
;pl25 CC mixing power level
;pl27 Ca H-Cali CP power
;pl28 H H-Cali CP power

;l29  (=1) ramp on H (=2) ramp on CA
;l30  (=2) ramp on Nitrogen; (=3) ramp on carbon
;l31  (=1) ramp on H (=2) ramp on N
;cnst10 1H carrier for NH CP in Hz (on Hamid)
```

```
;cnst11 1H carrier for CH CP in Hz (on Hali)
;cnst19 - offset for water suppression
;cnst30 C decoupling offset (at 120 ppm)
;cnst31 13C carrier for Cali in Hz (~ 20 ppm)
;spoff13 soft 180 offset (= DW)
;d9 TOCYS mixing time (WALTZ= n*960 us@10us, DIPSI=n*2173 us@10us)
;cpd3 C dec during 1H evolution
;cpd5 CC mixing change according, BE CAREFUL WITH POWER
;==============
; Set variables
;==============

"p6 = 2.0*p3"
"p5 = 2.0*p2"
"d11=30m"

"d0=0.2u"        ; evol. time F1 (CC)
"d10=0.2u"        ; evol. time F2 (Ca/ali)
"d20=0.2u"        ; evol. time F3 (N)
"d21=0.2u"        ; evol. time F4 (1H)


"in0=inf1/2"        ; increment delay F1
"in10=inf2/2"        ; increment delay F2
"in20=inf3/2"        ; increment delay F3
"in21=inf4/2"        ; increment delay F4



;============================
; Include file for Protection
;============================

#include <Avance.incl>
#include <Delay.incl>

;============================
; Protection for parameters
;============================
1m
   if "p1 > 100u"     goto Problem
   if "p2 > 100u"     goto Problem
   if "p3 > 100u"     goto Problem
   if "p13 >1001u"    goto Problem
```

```
   if "p15> 4001u"    goto Problem
   if "p16> 2000u"    goto Problem
   if "p17> 11001u"    goto Problem
   if "p22 >20001u"   goto Problem
   if "aq > 51m"      goto Problem
   if "d1 < 0.49s"     goto Problem
   if "d9 > 0.016s"     goto Problem

   goto PassParams
Problem, 1m
   print "Protection: Parameters not accepted, ending."
   goto HaltAcqu
PassParams, 1m

;----- Relaxation & reset parameters -----

1 ze
2 d11
  40u do:f3
  40u do:f2
  40u do:f1     ;decouplers off

3 d1
  20u reset:f1     reset:f2  reset:f3
  20u fq=cnst10:f1 fq=cnst20:f2   fq=0:f3 ;carrier on Ca
  5u  pl1:f1       pl2:f2    pl27:f3

;----- 90 on H and 1H evolution (t4) -----
        20u pl22:f3
  (p1 pl1 ph5):f1
  d21*2 cpd3:f3
        0.5u do:f3

;----- H/Cali CP ----- ; optimize for one-bond transfer

1u fq=cnst31:f3  ;carrier on f3 should be in the middle of Cali (35 ppm)

if (l29 == 1)
  {
  (p15 pl27 ph11):f3 (p15:sp6 ph10):f1
  }

if (l29 == 3)
  {
```

```
  (p15:sp7 ph11):f3 (p15 pl28 ph10):f1
  }

;----- CX (t1) evolution & HC decoupling  -----

  0.5u pl12:f1
  d0 cpd1:f1
  (p5 pl2 ph20):f2
  d0
  0.5u
;----- 90 on C  -----

  (p3 pl3 ph6):f3 ;maybe change sign here

;----- CC mixing  -----

  1u do:f1
  0.5u pl25:f3
  0.5u cpd5:f3
  d9
  2u do:f3


;----- 90 on C  -----

  (p3 pl3 ph7):f3

;----- C evolution -------

  1u fq=0:f3
  0.5u pl12:f1
  0.5u cpd1:f1
  d10
  (p5 pl2 ph20):f2
  d10
  1.5u

;----- CA/N CP -----

  0.5u pl8:f3

if (l30 == 2)
  {
  (p17:sp5 ph2):f2 (p17 ph2):f3
```

```
  }

if (l30 == 3)
  {
  (p17 pl7 ph2):f2 (p17:sp4 ph2):f3
  }

;----- N (t2) evolution & C decoupling  -----

  1u  pl3:f3
  1u  fq=cnst30:f3
  d20
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d20
  2u do:f1

;----- N magnetization along Z  -----

  (p2 pl2 ph17):f2 ;ph17

;----- Water suppression  -----

  20u fq=cnst19:f1
  2u pl16:f1

  p22:f1 ph22
  p22:f1 ph23
  p22:f1 ph22
  p22:f1 ph23

  20u fq=cnst11:f1
  5u pl6:f1

;----- N magnetization in XY plane  -----

  (p2 pl2 ph18):f2

;----- N/H CP -----

if (l31 == 1)
  {
  (p16 pl5 ph12):f2 (p16:sp2 ph2):f1
```

```
  }

if (l31 == 2)
  {
  (p16:sp3 ph12):f2  (p16 ph2):f1
  }

  1u pl23:f2 pl22:f3
  1u cpd2:f2  cpd3:f3

  go=2 ph31
  d11 do:f2  do:f3 mc #0 to 2

  F1PH(calph(ph11, +90), caldel(d0, +in0))
  F2PH(calph(ph7, +90), caldel(d10, +in10))
  F3PH(calph(ph17, -90), caldel(d20, +in20))
  F4PH(calph(ph5, +90), caldel(d21, +in21))

HaltAcqu, 1m
exit

;----- Phase cycling -----

ph2 = 0
ph5 = 0 2
ph6 = 3
ph7 = 3 3 3 3  1 1 1 1
ph10 = 1
ph11 = 0 0 2 2
ph12 = 1
ph17 = 3
ph18 = 0 0 0 0  0 0 0 0  2 2 2 2  2 2 2 2
ph20 = 1
ph21 = 2
ph22 = 0
ph23 = 1
ph31= 0 2 2 0  2 0 0 2  2 0 0 2  0 2 2 0

;=====================
; END Pulse program
;=====================
;; IMPORTANT NOTE
; On AV (avance) machines, its better to use "spf0" files instead of "sp0"
; The "sp0" files requires lot of time for power settings compared to "spf0"
```

HCNH

```
;HCNH_4D.alkl
;H (t1) to Cali (t2) with CC mix (to) Ca (t3) to N (t4) to H (t5) transfer;
; F2 channel is Nitrogen
; F3 channel is carbon

;=======================
; Variables introduction
;=======================

;p1   H p90
;p2   N p90
;p3   C p90
;p11  WURST pulse length for AL-FRESCO
;p13  C selective p180 duration
;p14 soft 180 carbon on res. (= DW/sqrt(3))
;p15  HCA CP contact time
;p16  NH CP contact time
;p17  CAN CP contact time
;p22  water suppression pulse length

;pl1  H p90 power
;pl2  N p90 power
;pl1  C p90 power
;pl5  N- NH CP power
;pl6  H- NH CP power
;pl7  N NCA-CP power
;pl8  C NCA-CP power
;pl13 C selective 180 power
;pl14 soft rectangular 180 pulse 13C
;pl16 water suppression pulse power
;pl22 WALTZ decoupling on N (during acquisition)
;pl23 (NC) decoupling on C (during acquisition)
;pl25 CC mixing power level
;pl27 Ca H-Cali CP power
;pl28 H H-Cali CP power

;l29  (=1) ramp on H (=2) ramp on CA
;l30  (=2) ramp on Nitrogen; (=3) ramp on carbon
;l31  (=1) ramp on H (=2) ramp on N
;cnst10 1H carrier for NH CP in Hz (on Hamid)
;cnst11 1H carrier for CH CP in Hz (on Hali)
;cnst19 - offset for water suppression
;cnst30 C decoupling offset (at 120 ppm)
```

```
;cnst31 13C carrier for Cali in Hz (~ 20 ppm)
;spoff13 soft 180 offset (= DW)
;d9 TOCYS mixing time (WALTZ= n*960 us@10us, DIPSI=n*2173 us@10us)
;cpd3 C dec during 1H evolution
;cpd5 CC mixing change according, BE CAREFUL WITH POWER

;==============
; Set variables
;==============

"p6 = 2.0*p3"
"p5 = 2.0*p2"
"d11=30m"

"d0=0.2u"         ; evol. time F1 (CC)
"d20=0.2u"         ; evol. time F3 (N)
"d21=0.2u"         ; evol. time F4 (1H)

"in0=inf1/2"         ; increment delay F1
"in20=inf2/2"         ; increment delay F3
"in21=inf3/2"         ; increment delay F4

#ifdef ALFRESCO

"spw22=spw21"

#endif



;============================
; Include file for Protection
;============================

#include <Avance.incl>
#include <Delay.incl>

;============================
; Protection for parameters
;============================
1m
   if "p1 > 100u"     goto Problem
   if "p2 > 100u"     goto Problem
   if "p3 > 100u"     goto Problem
```

```
   if "p13 >1001u"    goto Problem
   if "p15> 4001u"    goto Problem
   if "p16> 2000u"    goto Problem
   if "p17> 11001u"    goto Problem
   if "p22 >20001u"   goto Problem
   if "aq > 51m"      goto Problem
   if "d1 < 0.49s"     goto Problem
   if "d9 > 0.020s"     goto Problem

   goto PassParams
Problem, 1m
   print "Protection: Parameters not accepted, ending."
   goto HaltAcqu
PassParams, 1m

;----- Relaxation & reset parameters -----

1 ze

2 d11
  40u do:f3
  40u do:f2
  40u do:f1

3 d1
  20u reset:f1     reset:f2     reset:f3
  20u fq=cnst10:f1 fq=cnst20:f2 fq=0:f3 ;carrier on Ca
  5u  pl1:f1       pl2:f2       pl27:f3

;----- 90 on H and 1H evolution (t4) -----

  20u pl22:f3
  (p1 pl1 ph5):f1
  d21*2 cpd3:f3
        0.5u do:f3

;----- H/Cali CP ----- ; optimize for one-bond transfer

1u fq=cnst31:f3  ;carrier on f3 should be in the middle of Cali (35 ppm)

if (l29 == 1)
  {
  (p15 pl27 ph11):f3 (p15:sp6 ph10):f1
  }
```

```
if (l29 == 3)
  {
  (p15:sp7 ph11):f3 (p15 pl28 ph10):f1
  }

;----- CX (t1) evolution & HC decoupling  -----

  0.5u pl12:f1
  d0 cpd1:f1
  (p5 pl2 ph20):f2
  d0
  0.5u do:f1

;----- 90 on C  -----

 (p3 pl3 ph6):f3

;----- CC mixing  -----

#ifdef TOCSY

  1u do:f1
  0.5u pl25:f3
  0.5u cpd5:f3
  d9
  2u do:f3

#endif

#ifdef ALFRESCO
  (p11:sp21 ph25):f1 ;WURST pulses for AL-FRESCO mixing
  (p11:sp22 ph25):f1
  (p11:sp21 ph26):f1
  (p11:sp22 ph26):f1
  (p11:sp21 ph27):f1
  (p11:sp22 ph27):f1
  (p11:sp21 ph28):f1
  (p11:sp22 ph28):f1

#endif

;----- 90 on C  -----
```

```
 (p3 pl3 ph7):f3

;----- C evolution -------

  1u fq=0:f3
  0.5u pl12:f1
  0.5u cpd1:f1

;----- CA/N CP -----

if (l30 == 2)
  {
  (p17:sp5 ph2):f2 (p17 pl8 ph2):f3
  }

if (l30 == 3)
  {
  (p17 pl7 ph2):f2 (p17:sp4 ph2):f3
  }

;----- N (t2) evolution & C decoupling  -----

  1u  pl3:f3
  1u  fq=cnst30:f3
  d20
  (p3 ph20):f3
  (p6 ph21):f3
  (p3 ph20):f3
  d20
  2u do:f1

;----- N magnetization along Z  -----

  (p2 pl2 ph17):f2 ;ph17

;----- Water suppression  -----

  20u fq=cnst19:f1
  2u pl16:f1

  p22:f1 ph22
  p22:f1 ph23
  p22:f1 ph22
  p22:f1 ph23
```

```
  20u fq=cnst11:f1
  5u pl6:f1

;----- N magnetization in XY plane  -----

  (p2 pl2 ph18):f2

;----- N/H CP -----

if (l31 == 1)
  {
  (p16 pl5 ph12):f2 (p16:sp2 ph2):f1
  }

if (l31 == 2)
  {
  (p16:sp3 ph12):f2  (p16 ph2):f1
  }

  1u pl23:f2 pl22:f3
  1u cpd2:f2  cpd3:f3

  go=2 ph31
  d11 do:f2  do:f3 mc #0 to 2

F1PH(calph(ph11, +90), caldel(d0, +in0))
F2PH(calph(ph17, -90), caldel(d20, +in20))
F3PH(calph(ph5, +90), caldel(d21, +in21))

HaltAcqu, 1m
exit

;----- Phase cycling -----
ph2 = 0
ph5 = 0 2
ph6 = 3
ph7 = 3 3 3 3  1 1 1 1
ph10 = 1
ph11 = 0 0 2 2
ph12 = 1
ph17 = 3
ph18 = 0 0 0 0  0 0 0 0  2 2 2 2  2 2 2 2
ph25 = 0
```

```
ph26 = 1
ph27 = 2
ph28 = 3
ph20 = 1
ph21 = 2
ph22 = 0
ph23 = 1
ph31= 0 2 2 0  2 0 0 2  2 0 0 2  0 2 2 0

;====================
; END Pulse program
;====================
;; IMPORTANT NOTE
; On AV (avance) machines, its better to use "spf0" files instead of "sp0"
; The "sp0" files requires lot of time for power settings compared to "spf0"
```

7.1.1 All-in-one solid-state NMR version

```
;All in one solid state (HSQC-RFDR)
;USES 180DEG COMP PULSES FOR 13C AND 15N INSTEAD OF SHAPED PULSES ON 13C
;based on hsqcnoesygpsm193d, D. Frueh et. al J Biomol NMR 2009 Nov;45(3):311-8
;and All in one sequence by Alva Gossert (13C15N-HSQC-NOESY_TPPI)
;3D sequence with
;   homonuclear correlation via RFDR recoupling
;   H-1/X correlation via double inept transfer
;   simultaneous evolution of C-13 (aliphatics and aromatics) and N-15 chemical shift in
;          artificial chemical shift for H(!X) in t2
;
;Water suppression: Presaturation (use ZGOPT -DPRESAT) and z-filter-type after
; RFDR mixing
;
;$CLASS=HighRes
;$DIM=3D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

;=============
; Definitions
;=============

;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl9 : f1 channel - power level for presaturation
;pl11: f1 channel - RFDR (low) power level
;pl12: f2 channel - power level for CPD/BB decoupling
;pl16: f3 channel - power level for CPD/BB decoupling

;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p8 : f2 channel - 180 degree shaped pulse for inversion
;p9 : f1 channle - 90 degree low power rec. pulse (H2O on reson.)
;p11: RFDR 180 deg pulse on 1H

;p21: f3 channel -  90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse

;d0 : incremented delay (F1 in 3D)                  [3 usec]
```

```
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J)CH                                          [1.8 ms]
;d8 : cw sturation after RFDR
;d19 = (1s/(2*cnst9)) - p11/2

;d10: incremented delay (F2 in 3D)                     [3 usec]
;d11: delay for disk I/O                               [30 msec]
;d12: delay for power switching                        [20 usec]
;d21 RFDR mixing time NOT OVER 300 MS
;d26: 1/(4J(NH)                                        [2.7 ms]
;d28: incremented delay (F1 in 3D)                     [3 usec]

;cnst2: = J(CH)
;cnst4: = J(NH)
;cnst15: = SW(N) [ppm]
;cnst7: = 15N offset for decoupling (118 ppm)
;cnst6: = 15N offset in F3 (118 ppm)
;cnst5: = J(NH) optimal scaling: [0.7-1]
;cnst9: = Spinning Frequency (Hz)

;l11: RFDR loop counter (=multiple of 16)

;inf1: 1/SW(C) = 2 * DW(C)
;inf2: 1/SW(H) = 2 * DW(H)
;in0: 1/(2 * SW(C)) = DW(C)
;nd0: 2
;in10: 1/(2 * SW(H)) = DW(H)
;nd10: 2
;in28: = in30/2-in0
;in30: 1/SW(N) = 2 * DW(N)
;ns: 8 * n
;td1: number of experiments in F1
;td2: number of experiments in F2
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: States-TPPI (or TPPI) in F2
;cpd2: decoupling according to sequence defined by cpdprg2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence


;====== include files  =======

#include <Avance.incl>
```

```
#include <Delay.incl>

;===== define variables =====

"p2=p1*2"
"p4=p3*2"
"p22=p21*2"
"d4=1s/(cnst2*4)"
"d11=30m"
"d12=20u"
"d26=1s/(cnst4*4)"
"d9=d1-100ms"
"d18=1/(cnst3*4)"

"d19 = (1s/(2*cnst9)) - p11/2"
"d21 = l11*(1s/cnst9)"
"cnst13 = 24000"
"in30=1000000/(cnst15*bf3)"
"in28=in30/2-in0"


#ifdef MULTI
"in0=inf1/2"
"in10=inf2/2"
"d0=in0/2-(p3*2/PI)"
"d28=(in30/4)-(p21*2/PI)-d0-p3"
"d10=0.2u"

#else
"d0=0.2u"
"d28=0.2u"
"d10=0.2u"
#endif


;---- delays for HSQC part ----

"DELTA1=d4-0.5*(p4)"
"DELTA2=(d26*cnst5)-d4-0.5*(p4)"
"DELTA3=(d26*cnst5)"
"DELTA5=d26"
"DELTA6=d4-0.5*(p3+p3+p4)+larger(p4,p22)/2"
"DELTA7=d26-d4-0.5*(p4)+larger(p4,p22)/2+d10*2"
```

```
;========= PARAMETER PROTECTION ===========

"cnst63 = plw12"
"cnst62 = plw22"
"cnst61 = plw23"
"cnst60 = plw16"

        if "p1 > 100u"     goto Problem
        if "p2 > 100u"     goto Problem
        if "p3 > 100u"     goto Problem
        if "p21 > 200u"                goto Problem
        if "p13 >1001u"    goto Problem
        if "p15> 4001u"    goto Problem
        if "p16> 2000u"    goto Problem
        if "p17> 11000u"    goto Problem
        if "p22 >20001u"   goto Problem
        if "aq > 51m"      goto Problem
        if "d1 < 0.49s"     goto Problem
        if "cnst63>0.5"     goto Problem
        if "cnst62>1.5"     goto Problem
        if "cnst61>0.5"     goto Problem
        if "cnst60>0.7"     goto Problem

  goto PassParams
Problem, 1m
        print "Protection: Parameters not accpeted, ending."
        goto EXIT
PassParams, 1m
;=========================================


;----- acquisition order -----

aqseq 321

;====== start of pulse sequence =====

;---- reset scan counter, decoupling and recycle delay ----

1 ze
  d11 pl1:f1 pl12:f2 pl16:f3
2 d11 do:f2 do:f3
3 d11
```

```
;---- optional presat element ----

#ifdef PRESAT
  d11 pl9:f1
  d1 cw:f1 ph29^
  2u do:f1
#else
  d1
#endif /*PRESAT*/

  2u pl1:f1 pl2:f2 pl3:f3
  2u fq=cnst6(bf ppm):f3

;----- First INEPT -----

        (p1 pl1 ph10)
  DELTA1
  (p4 ph1):f2
  DELTA2
  (center (p2 pl1 ph1):f1 (p22 ph1):f3 )
  DELTA3
  (p1 ph2)

;----- combined 15N/13C evolution -----
  (p21 ph3):f3
  d28
  (center (p3 ph3 d0 d0 p3 ph4):f2 (p2 ph11):f1)
  d28
  (p21 ph5):f3

;----- reverse INEPT -----

        (p1 ph6)
  DELTA6
  (p4 ph2):f2
  DELTA7
  (center (p2 ph1) (p22 ph2):f3 )
        DELTA5


#ifdef RFDR

 ;----- H(X) evolution -----
```

```
  d10
  (center (p4 ph2):f2 (p22 ph2):f3 )
  d10

;------ RFDR mixing time -----

        (p1 pl1 ph7):f1
        (p14 pl14 ph12:r):f1
         0.2u pl11:f1
9 d19
  (p11 ph26^):f1
  d19
        lo to 9 times l11
  2u pl20:f1
        d8 cw:f1 ph29^
        2u do:f1
        (p1 pl1 ph2):f1

#endif /*RFDR*/

  2u pl12:f2
  2u pl16:f3

;----- acquisition, quad detection and indirect evolution -----

  go=2 ph31 cpd2:f2 cpd3:f3
  d11 do:f2 do:f3 mc #0 to 2

  F1PH(calph(ph3, +90) & calph(ph11, +45), caldel(d0, +in0) & caldel(d28, +in28) & calph(
  F2PH(calph(ph7, -90), caldel(d10, +in10))

EXIT, 2u
exit
; phase increment ph11:
; (+) upfield, (-) downfield. 90 deg = max chem shift, 0 deg = carrier

;----- Phase cycling -----

ph1 = 0
ph2 = 3
ph3 = 0 0 2 2
ph4 = (360) 250 ; 180 + phase increment on ph11
ph5 = (360) 250 ; 180 + phase increment on ph11
```

```
ph6 = 0
ph7 = 3 3 3 3  1 1 1 1
ph8 = 1
ph9 = 3
ph10 = 0 2
ph14 = 0
ph11 = (8) 1 1 3 3 ; 0 for testing (Ligand is not preserved)
ph12 = 1 1 1 1  3 3 3 3
ph13 = 2
ph26 = 0 1 0 1  1 2 1 2  2 3 2 3  3 0 3 0
ph29 = 0 1 0 1  1 2 1 2  2 3 2 3  3 0 3 0
#ifdef RFDR

ph31 = 0 2 2 0  2 0 0 2

#else

ph31 = 0 2 2 0

#endif

;=====================
; END Pulse program
;=====================
```

7.2 Extensions of the CYANA library

For several magnetization pathways introduced in this work, the CYANA library had to be appended by the respective experiments. Those were defined as followed:

```
SPECTRUM shCACONH HN N C CA
 0.980 H_AMI N_AMI CA:C_ALI C:C_BYL N:N_AMI HN:H_AMI

SPECTRUM shCOCANH HN N CA C
 0.980  H_AMI N_AMI C:C_BYL CA:C_ALI N:N_AMI HN:H_AMI

SPECTRUM hCACBcaNH_4D HN N CA CB
 0.980 HN:H_AMI N:N_AMI CA:C_ALI CB:C_ALI

SPECTRUM hCACBcacoNH HN N CA CB
 0.980 HN:H_AMI N:N_AMI C_BYL CA:C_ALI CB:C_ALI

SPECTRUM HNcocaNH H2 N2 H1 N1
 0.980 H2:H_AMI N2:N_AMI C_BYL C_ALI N1:N_AMI H1:H_AMI

SPECTRUM HNcoCANH_5D CA N2 H2 H1 N1
 0.980 H1:H_AMI N1:N_AMI C_BYL CA:C_ALI N2:N_AMI H2:H_AMI

SPECTRUM HNcaCONH_5D C N1 H1 H2 N2
 0.980 H2:H_AMI N2:N_AMI C_ALI C:C_BYL N1:N_AMI H1:H_AMI

SPECTRUM HCCNH H1 C1 CA N H2
 0.980 H1:H_A* C1:CA:C_ALI N:N_AMI H2:H_AMI
 0.980 H1:H_A* C1:C_A* CA:C_ALI N:N_AMI H2:H_AMI
 0.800 H1:H_A* C1:C_A* C_A* CA:C_ALI N:N_AMI H2:H_AMI
 0.300 H1:H_A* C1:C_A* C_A* C_A* CA:C_ALI N:N_AMI H2:H_AMI
 0.300 H1:H_A* C1:C_A* C_A* C_A* C_A* CA:C_ALI N:N_AMI H2:H_AMI

SPECTRUM MOCCA_4D HN N C1 C2
 0.980 HN:H_AMI N:N_AMI C1:C2:C_ALI
 0.900 HN:H_AMI N:N_AMI C1:C_ALI C2:C_ALI
 0.700 HN:H_AMI N:N_AMI C1:C_ALI C_ALI C2:C_ALI
 0.600 HN:H_AMI N:N_AMI C1:C_ALI C_ALI C_ALI C2:C_ALI
 0.500 HN:H_AMI N:N_AMI C1:C_ALI C_ALI C_ALI C_ALI C2:C_ALI
```

7.3 Assigned chemical shifts of TS

Residue	H	N	CA	CB	C
MET 1			56.4	30.0	175.3
GLU 2	9.40	127.4	56.5	30.0	175.9
ARG 3	8.73	118.2	60.1		
GLU 5			60.0	28.9	180.2
ASN 6	9.02	117.3	56.0	38.2	177.5
LEU 7	8.09	121.4	58.4	41.0	178.6
PHE 8	9.40	116.0	62.4	37.1	179.2
ALA 9	8.64	123.2	55.7	17.7	180.8
GLN 10	8.43	119.5	58.9	28.2	180.0
LEU 11	9.49	119.6	58.2	39.2	180.4
ASN 12	9.00	118.4	56.7	38.9	179.6
ASP 13	8.32	120.9	57.5	41.0	177.9
ARG 14	8.18	116.3	56.6	31.0	175.4
ARG 15	8.60	118.6	57.1	26.2	174.8
GLU 16	8.70	117.9	54.5	34.8	176.7
GLY 17	9.28	108.3	43.9		171.5
ALA 18	9.58	121.1	50.5	19.5	177.8
PHE 19	10.88	122.3	58.1	41.4	173.8
VAL 20	8.32	128.8	57.3	33.7	172.6
VAL 23			57.5		
PRO 28			65.1	32.9	
GLY 29	9.60	122.3	44.9		173.6
ILE 30	9.60	122.2	67.3	37.3	177.2
GLU 31	9.25	118.5	60.2	28.7	180.6
GLN 32	9.60	119.2	58.1	27.4	178.3
SER 33	8.83	114.8	61.0	63.5	177.0
LEU 34	8.63	117.8	58.9	41.1	178.9
LYS 35	8.17	119.8	59.7	31.9	180.5
ILE 36	8.33	122.7			
ILE 41			65.7	38.2	
ASP 42	9.11	122.8	57.7	42.0	177.8
ALA 43	8.70	116.9	52.6	20.4	177.3
GLY 44	8.07	103.1	44.5		174.2
ALA 45	8.80	121.7	53.9	18.6	177.0
ASP 46	9.18	119.3	56.8	44.3	
LEU 48			53.4	43.9	
GLU 49	8.81	122.9	56.1	32.6	176.1
LEU 50	9.80	128.9	51.3	51.6	174.3
PRO 57			62.1	37.5	
LEU 58	8.53	125.4	58.6	41.1	180.8

ALA 59	8.18	119.5	52.3	18.6	
PRO 62			65.1	33.2	
THR 63	8.90	107.7	62.2	70.2	175.6
ILE 64	7.57	125.2		38.7	
GLN 65			59.9	29.0	
ASN 66			55.8	37.8	177.9
ALA 67	8.01	126.5	55.6	19.4	179.8
LEU 69			58.3	41.8	179.0
ARG 70	8.35	118.1	59.5	28.7	179.7
ALA 71	8.39	123.6	55.5	18.4	180.2
PHE 72					179.5
ALA 73	8.75	121.4	55.3	17.6	180.0
ALA 74	7.64	118.6	53.4	20.0	177.9
GLY 75	8.38	112.6	45.3		
PRO 78			66.9	31.7	
ALA 79	9.17	118.6	55.9	17.9	182.2
GLN 80	8.61	119.3	59.0	28.8	180.1
CYS 81	9.03	119.1	64.7	25.7	177.0
PHE 82	8.82	118.8	63.5	38.4	178.9
GLU 83	8.30	121.5	59.9	29.0	180.4
MET 84	8.97	119.5	59.9	34.3	178.1
LEU 85	8.29	115.1	57.9	40.1	179.3
ALA 86	8.60	122.9	55.7	17.7	180.8
LEU 87			58.0	42.3	
ILE 88			63.5	38.3	178.7
ARG 89			59.4		
GLU 90			58.6	29.2	178.2
LYS 91	7.54	115.3	55.3	32.7	175.4
HIS 92	8.21	123.1	56.1	34.5	
PRO 93			65.1	33.2	
THR 94	8.89	107.7	62.2	70.2	175.7
ILE 95	7.56	125.1	58.7	38.8	
ILE 97			59.1		175.6
GLY 98	9.25	113.6	42.9		175.8
LEU 99	9.64	121.0	52.0	43.9	176.1
LEU 100	9.01	121.3	50.9	38.7	
MET 101			51.1	33.8	175.3
TYR 102	7.52	118.3	58.9	39.9	177.4
ALA 103			55.7	17.7	180.0
ASN 104	9.43	116.5	54.3	38.1	174.8
LEU 105	8.40	122.5	57.7	41.7	178.1
VAL 106	8.69	116.9	60.8	31.7	177.8
PHE 107	8.28	120.1	58.3	40.0	179.7
ASN 108	9.47	119.2	57.0	38.7	179.0

ASN	109	9.03	121.3	51.0	38.8	
GLY	110			44.6		173.6
ILE	111	9.62	122.2	67.3	37.4	177.2
ASP	112	9.23	118.8	59.0	41.8	177.9
ALA	113	8.37	119.2	54.6	17.6	180.8
PHE	114	8.72	121.4	62.5	38.9	178.4
TYR	115	8.47	119.2	64.1	38.0	178.1
ALA	116	8.99	123.1	54.9	17.5	180.5
ARG	117			59.4		180.9
GLY	122			46.8		175.2
VAL	123	9.21	121.3	60.6	30.9	176.8
ASP	124			57.9	41.2	177.9
LEU	127			54.1	43.3	174.7
VAL	128	9.68	128.0	60.5	29.8	176.7
ALA	129	9.23	131.1	55.7	17.8	176.7
ASP	130	9.63	111.2	51.6	37.5	172.5
VAL	131	6.88	113.4	58.3	33.6	
VAL	133	10.75	119.9	66.5	31.3	180.9
GLU	134	11.91	124.3	61.2	29.5	179.0
GLU	135	8.07	117.0	55.4	31.5	178.4
SER	136	8.35	112.5	60.1	63.4	174.2
ALA	137	8.79	124.7	55.8	17.2	
PRO	138			66.1	30.7	
PHE	139	7.99	116.1	61.0	38.3	176.3
ARG	140	9.10	115.6	59.0	29.0	177.5
GLN	141	8.69	118.7	59.2	26.2	177.6
ALA	142	7.31	119.4	55.2	19.8	179.7
ALA	143	8.16	120.3	55.2	17.7	180.6
LEU	144	8.01	119.8	57.9	41.3	179.8
ARG	145	8.29	122.7	58.7	28.1	178.3
HIS	146	8.41	112.2	54.1	31.0	172.2
ASN	147	8.65	115.0		39.6	
ASP	160	9.18	119.2	58.9	42.1	177.7
ASP	161	8.60	116.5	57.6	40.5	179.3
LEU	162	8.61	122.8	57.8	39.8	179.4
LEU	163	9.18	119.0	58.8	42.1	177.9
ARG	164	8.59	116.5	59.7	29.7	180.2
GLN	165	8.17	119.9	59.3	30.0	177.7
VAL	166	9.41	118.2	67.0	31.1	179.4
ALA	167	8.80	121.1	54.8	19.6	178.9
SER	168	7.92	110.7	60.4	63.6	177.1
TYR	169	8.99	117.7	56.0	39.1	178.9
GLY	170	8.61	104.5	45.4		175.2
ARG	171	8.83	118.8	56.3	30.0	177.0

GLY 172	9.33	105.9	46.2		172.6
TYR 173	8.21	118.9	58.0		
SER 180	8.99	110.9	62.5	63.5	174.0
GLY 181	8.21	104.9	45.3		176.1
VAL 182	8.97	118.4	60.2	28.7	180.7
ALA 190			51.7		177.0
LEU 196			57.8	41.5	178.6
ILE 197	8.24	122.9	67.2	37.6	
GLU 198			59.8		178.0
LYS 199	8.67	118.0	59.2	33.9	
LEU 200	8.32	118.1	58.6	40.9	180.3
LYS 201			59.6	31.8	180.4
GLU 202	9.05	118.7	59.1	29.0	179.8
TYR 203	7.95	122.9	61.7	38.8	177.6
HIS 204	8.65	116.1	58.6	27.5	179.6
ALA 205	8.81	121.3	55.1	17.8	180.8
ALA 208	8.90	125.4	52.7	18.6	173.6
LEU 209	9.12	121.6	54.1	42.2	
GLN 210			55.4	32.5	
ILE 214			58.7		176.2
SER 215	9.71	116.5	57.3	66.3	174.5
SER 216	8.27	123.7			
GLN 219			58.9	28.2	
VAL 220			66.8	30.8	
SER 221			61.5	62.8	177.5
ALA 222			55.5	17.8	180.6
ALA 223	8.41	119.3	54.7	17.7	180.8
VAL 224	8.73	121.2			
ARG 225			59.5	29.3	
ALA 226	8.17	120.0	52.7	18.7	176.6
GLY 227	8.42	103.5	44.6		175.8
ALA 228	8.65	123.2	52.2	17.7	176.1
ALA 229	9.69	121.7	52.1	18.7	178.1
GLY 230	7.98	103.5	46.5		170.4
ALA 231	9.56	122.5	50.5	23.5	174.0
ILE 232	9.11	121.6	53.9	42.3	176.0
SER 235	7.98	114.1	60.1	70.4	176.9
ALA 236	8.88	125.4	52.9	18.6	173.5
ILE 237	9.10	121.5	58.4	41.4	176.2
VAL 238	8.43	119.4	58.6	29.9	177.9
LYS 239	7.64	114.5	58.9	34.7	175.2
ILE 240	7.97	122.9	61.7	38.6	177.4
ILE 241	8.71	122.3	66.4	38.7	177.3
GLU 242	8.87	117.6	59.2	29.9	179.0

LYS 243	8.74	116.5	58.2	32.8	177.7
ASN 244	7.92	115.0	53.6	40.8	175.1
LEU 245	7.60	120.6	58.4	42.0	178.7
ALA 246	9.01	116.7	52.7	18.6	177.8
SER 247	8.42	112.2	54.3	63.4	171.9
PRO 248			65.9	31.8	
LYS 249	9.01	116.9	60.1	31.9	180.3
GLN 250	8.36	122.9	58.3	27.7	178.4
MET 251	9.37	118.6	59.1	34.5	178.0
LEU 252	8.26	115.1	58.2	39.9	179.4
ALA 253	8.12	122.6	55.4	17.7	181.4
GLU 254	9.30	120.0	59.6	29.1	180.5
LEU 255			58.3	42.1	178.5
ARG 256			59.6	29.8	
VAL 259			67.0	30.3	
SER 260			61.7	62.8	177.8
ALA 261			55.4	17.7	180.6
MET 262			58.5	30.0	
ALA 264	8.87	125.9	55.6	17.4	180.0
ALA 265	7.63	118.4	53.7	20.1	177.8
SER 266	8.34	112.6	60.0	63.4	174.2
ARG 267	7.92	123.3	57.3	29.9	175.3
ALA 268	8.75	118.0	58.0		
MET 269			54.7	31.8	173.9
THR 270			60.0	70.4	
THR 271	8.80	115.6	61.0	70.2	175.2
LEU 272	11.15	124.5	57.0	42.8	178.4
LEU 273	9.61	121.0	52.0	44.0	176.1
ASN 274	9.02	121.4	51.1	38.7	175.1
PRO 275			62.4	33.1	
TYR 276	8.44	119.1	57.6	41.4	175.3
PHE 277	10.23	122.3	55.7	38.5	175.3
GLY 278	8.86	112.0	47.0		175.5
GLU 279	9.34	127.5	56.5	29.9	176.1
PHE 280	8.71	118.0	57.9	41.1	175.4
GLY 281	11.42	112.3	44.9		174.2
GLY 282	8.11	109.8	44.6		174.3
MET 283	9.59	122.9	54.7	34.8	173.4
TYR 284	8.56	123.1	58.1	36.5	176.3
VAL 285	7.52	111.4	55.9	32.1	171.5
ALA 292	7.39	119.2	54.5	15.8	181.6
LEU 293	8.22	120.2	58.2	40.0	179.7
ASN 294	9.45	119.3	56.9	38.7	179.0
GLN 295	9.01	121.1	58.9	28.5	179.1

LEU 296	8.46	122.4	57.7	41.2	178.1
GLU 297	8.71	118.7	60.3	29.0	179.4
GLU 298	8.81	121.9	60.1	29.2	
ALA 299	8.82	121.9	54.4	18.6	
VAL 301			66.6	31.2	180.6
SER 302	8.14	112.2	60.9	62.8	177.8
ALA 303	8.82	124.4	55.2	16.4	180.0
GLN 304	8.62	114.4	58.2	29.2	177.8
LYS 305	7.55	115.2	55.4	32.8	175.4
ASP 306	8.21	123.2	51.0	42.6	175.9
PRO 307			64.9	32.0	
GLN 310	8.63	116.3	58.7	27.5	179.6
ALA 311	8.81	121.3	55.1	17.8	180.3
GLN 312	8.23	121.7	59.2	28.6	179.6
PHE 313			57.8		
ALA 314			55.5	17.7	180.5
ASP 315			57.8	41.4	178.5
LEU 316			58.0	40.5	180.0
LEU 317				41.2	180.2
LYS 318			59.9	32.0	180.3
ASN 319	8.37	122.7	53.7	40.9	175.4
TYR 320	7.62	120.6	58.1	42.1	178.4
ALA 321	8.70	116.8	52.6	20.3	177.5
GLY 322	8.08	103.2	44.6		174.0
ARG 323	8.80	121.8		32.8	
THR 325			63.0	72.8	174.4
ALA 326	10.11	131.7	56.3	32.5	176.9
LEU 327	9.15	127.6	61.0	28.0	175.4
THR 328			62.8	72.8	174.5
LYS 329	10.12	131.7	56.4	32.5	176.8
CYS 330	9.16	127.5	60.7	28.0	175.5
GLN 331	8.58	118.7	57.0		174.7
ASN 332	8.72	117.9			176.0
ILE 333			64.5	38.1	175.3
THR 334	7.51	105.6	60.2	69.1	175.3
ALA 335	7.75	126.1	54.3	18.6	178.9
GLY 336	9.68	111.1	45.3		174.5
THR 337	8.68	111.6	60.6	73.0	175.7
ARG 338	8.60	118.4	54.9	27.8	176.9
THR 339	8.92	120.0	63.9	70.0	174.6
THR 340	9.74	125.4	62.4	69.8	172.0
LEU 341	9.65	130.3	53.3	46.3	173.8
LEU 343			53.5	43.6	176.3
LYS 344	8.82	122.8	55.9	32.6	176.0

ARG 345	9.80	128.8	58.1	28.6	173.6
GLU 346	8.65	119.3	58.9	28.6	178.6
ASP 347			57.0	38.9	
LEU 348			54.5	40.4	178.7
LEU 349	7.96	123.9	54.4	44.3	178.7
HIS 350	9.52	125.1	58.7	32.1	175.3
GLY 351	8.79	116.5	44.3		176.2
GLY 352	9.84	109.2	46.0		172.9
ALA 353	7.75	117.6	51.1	23.0	177.3
HIS 354	9.44	117.1	58.9	31.1	179.5
LYS 355	6.25	118.6	59.7	33.7	176.7
THR 356	7.85	111.3	66.4	69.5	174.3
ASN 357	8.08	115.6	57.5	40.2	177.3
GLN 358	8.72	119.6	58.8	28.1	178.6
ALA 363	8.80	121.6	55.2	17.8	180.7
LEU 364			57.9	42.2	178.1
LEU 365	8.61	116.4	57.7	40.6	179.5
ALA 366			55.4	17.6	180.3
LYS 367	8.82	123.1	60.5	33.5	178.5
ARG 368	8.71	130.4	55.6	31.2	175.8
MET 369	8.87	118.8	53.1	35.9	174.5
GLY 370	9.33	105.9	46.2		174.4
LYS 371	8.55	116.6	55.2	31.3	175.9
SER 372	9.03	111.6	57.8	65.1	174.6
GLU 373	8.42	121.4	59.5	30.1	181.4
ILE 374			57.3	42.7	174.9
ILE 375	9.51	123.1	57.3	42.7	174.7
ALA 376	9.55	122.7	50.3	23.1	174.5
GLU 377	9.62	119.0	54.8	31.9	174.1
THR 378	8.49	110.0	59.9	70.4	171.2
GLY 379	9.03	116.2	46.5		171.1
ALA 380			52.6	18.6	176.2
GLY 381	7.38	101.4	45.9		174.4
GLN 382	9.09	118.8	59.4	29.0	179.6
HIS 383	9.86	118.4	59.7	31.7	178.2
GLY 384	8.16	106.8	47.8		174.3
VAL 385	9.02	122.4	67.5	31.4	177.2
ALA 386	8.38	121.2	55.0	16.8	179.1
SER 387				62.7	177.0
ALA 388			55.0	17.3	
LEU 389			57.9	41.2	
ALA 390			55.1	16.8	
SER 391				62.7	177.4
ALA 392			55.3	17.5	

LEU 394			54.4	41.0	
GLY 395	8.23	109.9	47.3		175.4
LEU 396	9.19	121.0	53.4	43.8	176.6
LYS 397	8.80	122.8	56.2	32.4	176.1
CYS 398	9.77	129.5	58.1	28.6	173.4
ALA 404			55.4	17.6	179.7
LYS 405			59.5	32.1	180.1
ASP 406			57.5	41.0	178.5
VAL 407	8.93	121.2	66.8	31.7	178.3
GLU 408			60.0	28.8	
MET 417			60.6	34.5	178.1
ARG 418	8.85	118.4	59.5	29.7	182.5
LEU 419	9.59	124.7	57.5	42.1	178.6
MET 420			55.5		175.8
GLY 421	8.51	107.0	45.0		174.4
ALA 422	8.37	123.6	52.0	19.4	177.0
GLU 423	8.99	122.9	55.0	30.8	175.0
VAL 424	8.97	127.5	61.5	31.8	176.2
ILE 425	9.93	133.3	59.0	38.6	
HIS 428			56.4		175.6
SER 429	7.82	114.1	57.3	64.4	173.9
GLY 430	9.24	107.2	47.2		176.1
SER 431	9.54	124.4	58.0	63.9	173.8
ALA 432	8.40	121.3	55.1	16.8	178.9
THR 433			60.7	73.0	176.0
LEU 434	8.66	118.3	58.2	41.1	
ASP 436			57.8	41.2	178.5
ALA 437	8.19	122.8	55.5	17.7	181.3
CYS 438	9.31	120.0	59.8	29.1	180.6
LEU 442			58.3	41.4	
SER 446	8.98	110.9	62.5	63.5	173.9
GLY 447	8.19	104.9	45.3		176.2
ALA 452			51.7	19.5	176.8
HIS 453	8.98	122.8	55.0	30.8	176.8
TYR 454					175.9
MET 455	9.34	123.4	52.4	33.7	177.4
LEU 456	6.61	132.6	59.1	43.3	179.3
GLY 457	9.52	109.5	44.1		174.1
THR 458	7.67	113.8	60.0	67.5	172.4
ALA 459	8.01	126.8	50.0	15.7	174.2
ALA 460	8.52	121.6	50.7	21.7	173.3
GLY 461	8.54	103.7	43.4		171.8
ILE 468	8.46	120.5			176.3
VAL 469	8.98	122.5	67.4	31.4	

ARG 470			59.4	29.9	181.3
GLU 471			58.7		178.0
PHE 472			56.9	38.7	179.3
GLN 473	8.89	118.6	56.2	30.0	177.9
ARG 474	8.37	119.6	58.3	29.5	177.6
ILE 476	8.69	122.8	66.7	38.9	180.3
GLY 477	8.16	111.9	48.0		175.5
GLU 478	8.42	121.4	59.5	30.1	181.4
GLU 479	9.14	118.0	58.7	30.1	177.9
LYS 481			59.3	30.4	
ALA 482	7.17	120.1	54.8		180.3
GLN 483	8.69	120.7	58.8	28.2	179.3
ILE 484			61.2	39.3	
LEU 485	8.52	125.4	58.5	41.0	180.9
ASP 486	8.15	119.5	57.5	42.1	177.9
LYS 487	8.59	115.9	58.7	34.2	179.0
GLU 488	9.47	114.1	55.2	30.9	178.3
GLY 489	8.44	108.8	46.1		173.2
ARG 490	7.87	113.8	54.1	30.9	172.5
LEU 491	8.64	114.9	52.4	39.5	174.7
ASP 493			58.6	42.0	177.9
ALA 494	8.66	116.6	52.6	20.3	
ILE 496			60.4	41.1	
ALA 497	9.65	127.0	50.3	23.2	176.3
CYS 498	10.36	120.1	56.5	29.9	173.9
VAL 499	9.56	123.7	62.7	31.8	176.9
GLY 500	10.58	113.0	43.7		174.2
GLY 501	9.20	108.1	46.6		177.5
GLY 502			46.0		
ALA 505	8.85	121.2	54.5	19.6	178.5
ILE 506	8.67	115.4	60.4	37.5	176.8
GLY 507	8.68	110.7	47.9		176.3
MET 508	8.31	121.6	57.2	30.9	175.3
PHE 509	9.50	111.4	56.5	38.7	180.8
ALA 510	9.01	118.6	57.7	16.5	181.0
ASP 511	10.33	118.3	56.8	41.4	176.2
PHE 512	8.87	113.4	57.6	41.5	179.0
ILE 513	8.21	122.9	67.3	37.6	178.9
ASN 514	9.40	116.6	54.3	38.2	174.9
ASP 515	8.34	122.5	53.8	41.3	175.5
THR 516	8.26	112.1	65.5	68.6	175.3
SER 517	8.96	115.7	60.2	63.1	174.2
VAL 518	8.77	124.5	62.4	31.9	176.4
GLY 519	9.18	118.7	45.4		171.8

LEU 520	8.94	125.6	53.2	40.0	174.7
ILE 521	9.00	123.0	59.3	39.6	175.7
GLY 522	9.26	113.8	43.0		172.8
VAL 523	9.26	123.3	60.6	33.6	176.4
GLU 524	9.95	127.8	54.6	29.9	176.4
PRO 525			62.9	31.8	
GLY 526	10.61	112.9	43.7		174.2
GLY 527	9.19	108.1	46.5		177.6
HIS 528	9.34	124.7			
GLY 529	9.58	111.4	44.2		176.6
ILE 530	10.05	130.5	63.0	34.7	180.1
GLU 531	10.24	119.2	58.8	27.2	177.8
THR 532	8.15	110.2	63.4	71.8	177.2
GLY 533	8.53	108.1	46.8		172.9
GLU 534	9.10	124.7	54.3	27.2	172.8
HIS 535	7.66	119.1	55.4	32.3	174.4
GLY 536	8.61	109.1	47.5		172.7
ALA 537	11.69	126.6	50.2	18.7	175.1
GLY 542	8.56	112.4	44.0		173.1
ARG 543	9.40	118.4	55.3	33.1	177.5
VAL 544	9.74	128.4	65.0	31.2	174.2
GLY 545	8.85	116.6	45.9		171.1
ILE 546	9.11	119.7	59.8	40.2	175.0
PHE 548	8.72	118.1	57.8	41.1	175.5
GLY 549	10.81	112.2	45.1		174.4
MET 550	8.67	116.7	53.7	36.1	171.8
LYS 551	10.27	122.0	54.6	32.7	177.1
ALA 552	9.80	128.9	51.6	19.5	
MET 554	9.17	120.5	54.5	35.2	175.3
MET 555	8.84	122.3	56.0	32.8	174.4
GLN 556	8.90	121.2	54.2	33.1	177.1
THR 557	10.09	114.6	60.9	71.6	177.5
ALA 558	9.81	123.9	55.3	17.2	179.0
ASP 559	8.23	121.9	54.4	41.0	176.5
GLY 560	8.22	110.0			175.0
GLN 561			53.5	26.7	176.6
ILE 562	8.81	123.1	60.6	33.5	178.4
GLU 563	8.74	130.4	55.6	31.2	175.8
GLU 564	8.84	119.1	56.2	29.9	176.8
SER 565	8.45	118.0	56.7	66.7	174.4
ILE 568	8.79	118.8	63.2	38.3	174.1
SER 569	8.54	114.5	54.3	65.6	175.2
ALA 570	9.11	131.7	54.8	18.7	180.7
ASP 573	8.83	122.6	56.2	39.9	

PHE 574	8.97	121.2	53.8	42.4	
SER 576	6.46	113.1	54.9	64.4	172.2
VAL 577	9.82	121.5	61.1	32.5	171.3
GLY 578	8.62	108.5	46.8		172.7
PRO 579			65.0	32.1	
GLN 580	8.77	119.3	59.6	29.1	178.9
HIS 581	7.32	115.8	59.4	30.5	176.7
ALA 582	7.14	120.2	54.6	15.2	179.5
TYR 583	8.53	120.7	61.3	39.5	176.9
LEU 584	9.14	119.4	57.6	37.9	177.0
ASN 585			57.4	40.3	178.3
SER 586	9.57	116.9	63.5	62.8	175.9
ILE 587	7.74	107.3	61.2	37.6	177.4
GLY 588	8.73	109.6	45.3		174.1
ARG 589	8.61	120.8	55.5	28.9	175.2
ALA 590	7.78	121.2	49.5	24.5	173.2
ASP 591	8.83	116.9	52.7	44.7	173.9
TYR 592	9.66	118.4	56.4	40.6	175.8
VAL 593	9.24	116.1	58.6	35.5	176.1
SER 594	9.71	116.7	57.3	66.4	174.5
ILE 595	8.28	123.8	56.1	41.5	176.8
THR 596	10.40	120.7	62.4	71.9	174.9
ASP 597	10.09	118.8	57.6	41.4	178.4
ASP 598	8.83	116.5	57.8	40.4	180.2
GLU 599	8.24	121.1	59.1	31.6	180.2
ALA 600	9.09	124.2	55.4	14.7	179.7
LEU 601	8.02	119.7	57.7	42.0	180.2
GLU 602	8.75	118.6	60.1		
LYS 605			59.7	33.8	176.6
THR 606	7.86	111.3	66.7	69.5	177.2
LEU 607	9.21	119.5	58.5	41.8	
CYS 608	8.63	116.3	59.2	27.8	
ARG 609	8.26	119.6	58.5	29.6	177.6
GLY 612	8.49	108.9	45.9		
ILE 613			57.7	41.2	175.1
PRO 615			61.7	31.9	
ALA 616	8.21	118.7	52.0	16.9	182.0
LEU 617	10.32	126.4	57.8	42.0	180.4
GLU 618	10.97	123.5	61.3	29.2	179.2
SER 619	7.49	113.6	59.8	62.7	177.6
SER 620	9.56	116.9	63.3	62.8	175.9
HIS 621	8.95	119.2	61.5	27.3	177.2
ALA 622	7.38	119.5	54.8		
LEU 623	8.28	120.0	57.9		

ALA	624	8.78	119.7	55.4	17.8	180.5
LEU	627			58.2	41.1	
LYS	628			59.8	31.6	177.9
MET	629	8.63	117.8	59.2	33.9	180.4
MET	630	8.28	118.0	58.4	34.3	175.5
ARG	631	8.92	116.1	58.4	30.3	180.3
GLU	632	9.36	116.2	59.2	30.6	178.1
GLN	633	6.95	111.6	52.9	30.1	174.3
GLU	635	8.48	115.5	55.3	30.0	175.5
LYS	636	8.14	123.2	56.0	34.5	176.0
GLU	637	9.45	131.9	56.4	29.0	176.2
GLN	638	8.25	122.1	54.5	30.3	172.7
LEU	639	10.78	128.0	53.9	42.1	175.0
LEU	640	10.80	128.1	56.3	41.8	178.5
VAL	641	8.77	124.7	62.3	31.9	176.3
VAL	642	9.31	118.6	59.5	34.4	176.2
LEU	644	10.10	138.3	53.9	39.6	173.3
SER	645	7.47	116.1	62.5	62.1	171.9
GLY	646	8.44	105.6	47.5		172.9
ARG	647	9.11	116.4	54.4	31.9	179.7
GLY	648	8.85	107.5	46.6		175.6
ASP	649	8.54	122.7	58.8	40.6	179.2
LYS	650			59.6		180.2
ASP	657			57.9		179.1
ILE	658	8.23	122.7	67.3	37.4	177.1
LEU	659	9.22	118.7	58.7	42.0	
LYS	660			59.8		179.8

7.4 Stock and media recipes

Growth media	Compound	Amount / l[a]	Concentration
LB medium	Yeast extract	5 g	
	Tryptone	10 g	
	NaCl	10 g-	
Agar plates	LB medium	1 l	
	Bacto Agar	15 g	
	Antibiotic stocks	1 ml	
M9 10x stock	Na_2HPO_4	60 g	0.479 M
	KH_2PO_4	30 g	0.220 M
	NaCl	5 g	0.086 M
M9 1x	M9 10x	100 ml	
Add in order to avoid precipitation	Trace elements 100x	10 ml	
	Biotin stock	1.5 ml	
	Thiamine stock	1.5 ml	
	Antibiotic stocks	1 ml	
	D-Glucose	2 g	
	NH_4Cl	1 g	
	H_2O	fill to 1 l	
	$MgSO_4$ 1M	1 ml	1 mM
	$CaCl_2$ 1M	0.3 ml	0.3 mM

Antibiotic stocks			
Ampicillin 1000x	Ampicillin-HCl	2 g	
	H_2O	20 ml	
Chloramphenicol 1000x	Chloramphenicol	0.68 g	
	EtOH	20 ml	
Kanamycin 1000x	Kanamycin	5 g	
	H_2O	20 ml	

Buffers			
PBS	NaCl	8.1816 g	140 mM
	KCl	0.2013 g	2.7 mM
	Na_2HPO_4	1.4196 g	10 mM
	KH_2PO_4	0.2449 g	1.8 mM
	adjust to pH 7.3		
Lysis buffer hCAII	PBS	1 l	
	DTT	0.1928 g	2.5 mM
	Lysozyme		
	DNAse I		

	Protease inhibitor cocktail	1 tablet	
Binding buffer hCAII	PBS	1 l	
	DTT	0.1928 g	2.5 mM
Elution buffer hCAII	Tris	6.057 g	50 mM
	DTT	0.1928 g	2.5 mM
	reduced glutathione	3.0733 g	10 mM
	pH 8		
SEC buffer hCAII	NH_4HCO_3	3.953 g	50 mM
	pH 8		
AEC loading buffer SH3	Tris	2.4227 g	20 mM
	pH 8.5		
AEC elution buffer SH3	Tris	2.4227 g	20 mM
	NaCl	58.44 g	1 M
	pH 8.5		
SEC buffer SH3	Citric acid	3.8425 g	20 mM
	NaCl	8.766 g	150 mM
	pH 3.5		

Vitamin stocks

Thiamine	Thiamine-HCl	20 mg	
	H_2O	20 ml	
Biotin	Biotin	20 mg	
	H_2O	20 ml	

Other stocks

Trace elements 100x	EDTA	5 g	17.1 mM
Consider order an pH	H_2O	fill to 1 l, pH 7.5	
	$FeCl_3 \cdot 6\,H_2O$	830 mg	3.07 mM
	$ZnCl_2$	84 mg	616 µM
	$CuCl_2 \cdot 2\,H_2O$	13 mg	76.3 µM
	$CoCl_2 \cdot 6\,H_2O$	10 mg	42.1 µM
	H_3BO_3	10 mg	162 µM
	$MnCl_2 \cdot 4\,H_2O$	1.4 mg	7.07 µM
$ZnCl_2$	$ZnCl_2$	0.136 g	0.1 M
	H_2O	10 ml	
$MgSO_4$	$MgSO_4$	6.018 g	1 M
	H_2O	50 ml	
$CaCl_2$	$CaCl_2$	5.549 g	1 M
	H_2O	50 ml	
IPTG 1000x	IPTG	4.766 g	1M
	H_2O	fill to 20 ml	

7.5 Conferences, workshops, and funding

Conferences

1. Annual network meeting of the *bio*-N^3*MR* network, 13th September 2019, Jülich, Germany.
   Talk: "Higher-dimensionality solid-state NMR on a 72 kDa protein", **A. Klein**, S.K. Vasa, P. Rovó, L. Mueller, R. Linser

2. Alpine conference on Magnetic Resonance in Solids, 15-19 September 2019, Chamonix-Mont-Blanc, France.
   Promoted talk: "Increasing dimensionality and information content for challenging systems in ^{1}H detected solid-state NMR", **A. Klein**, S.K. Vasa, P. Rovó, L. Mueller, R. Linser

3. 61st Experimental Nuclear Magnetic Resonance Conference (ENC), 8-13 March 2020, Baltimore (MD), USA.
   Poster presentation: "Assignment of 2x72 kDa Tryptophan Synthase by Proton-detected 5D Solid-State NMR Experiments", **A. Klein**, P. Rovó, V. Sakhrani, P. Skowronek, L. Kukuk, S.K. Vasa, L. Mueller, R. Linser

4. 62st Experimental Nuclear Magnetic Resonance Conference (ENC), 29-31 March 2021, virtual conference.
   Poster presentation: "Proton-detected 5D Solid-State NMR Facilitates Assessment of Large Enzyme Complexes and Easy Assignment of Side-Chains in Fully Protonated Proteins", **A. Klein**, S.K. Vasa, P. Rovó, V. Sakhrani, J. Holmes, P. Güntert, L. Mueller, R. Linser

Workshops

1. 5th US-Canada Winter School on Biomolecular Solid-State NMR, 7-12 January 2018, Stowe (VT), USA.

2. NMR Practical Course 2018: Multidimensional NMR in Structural Biology, 12-19 August 2018, Joachimsthal, Germany.
   Poster presentation: "Automated Projection Spectroscopy in Solid-State NMR Spectroscopy", **A. Klein**, S.K. Vasa, R. Linser

Funding

1. Student grant of the Alpine conference on Magnetic Resonance in Solids 2019